DEUFMD 26(76) (2024) 134-140

Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi
Dokuz Eylul University Faculty of Engineering Journal of Science and Engineering
Elektronik/Online ISSN: 2547-958X

RESEARCH ARTICLE / ARAŞTIRMA MAKALESI

Performance Assessment of the Modified Kernel Ridge Predictors in the Partially Linear Mixed Measurement Error Models via Covid-19 Data Analysis

Ölçüm Hatalı Kısmi Lineer Karma Modellerde Modified Kernel Ridge Öntahmin Edicilerin Covid-19 Veri Analizi Yoluyla Performans Değerlendirmesi

Özge Kuran *, **Seçil Yalaz**

Dicle Üniversitesi Fen Fakültesi İstatistik Bölümü, Diyarbakır, TÜRKİYE

Corresponding Author / Sorumlu Yazar *: ozge.kuran@dicle.edu.tr; ozge-kuran@hotmail.com

Abstract

In this article we describe new predictors under multicollinearity situation in the partially linear mixed measurement error models. In order to achieve this aim, we refer to some preliminary information and use it in order to suggest the modified Kernel ridge predictors in the partially linear mixed measurement error models. In addition, we also attain some mean square error comparisons between our new described modified Kernel ridge predictors and predictors previously described in literature for the partially linear mixed measurement error model. In conclusion, the article showcases real data analysis and a simulation study to illusrate our theoretical findings.

Keywords: *Linear Mixed Model, Measurement Error, Multicollinearity, Partially Linear Model, Ridge Predictor*

Öz

Bu çalışmada, ölçüm hatalı kısmi lineer karma modellerde çoklu iç ilişki durumu altında yeni öntahmin ediciler tanımlanmaktadır. Bu amaca ulaşmak için, bazı ön bilgiler ele alınmıştır ve bu bilgi hesaba katılarak, ölçüm hatalı kısmi lineer karma modellerde modified Kernel ridge öntahmin edicileri önerilmiştir. Ek olarak, ölçüm hatalı kısmi lineer karma model literatüründe daha önce tanımlanan öntahmin ediciler ile yeni tanımlanan modified Kernel ridge öntahmin ediciler arasında bazı hata kareler ortalama karşılaştırmaları da yapılmıştır. Daha sonra, teorik bulgularımızı kanıtlamak için gerçek bir veri analizi ve simülasyon çalışması ile makale sonlandırılmıştır.

Anahtar Kelimeler: *Lineer Karma Model, Ölçüm Hatası, Çoklu İç İlişki, Kısmi Lineer Model, Ridge Öntahmin Ediciler*

1. Introduction

Linear mixed model (LMM) [1] is an expanded version of linear model (LM). LMM has both fixed and random effects, and are especially employed to study clustered data such as longitudinal data, repeated measures data, multilevel data and etc. Another commonly studied statistical model in literature is nonparametric model (NPM) under the measurement error problem [2]. This model introduces the functional form of LM where heterogeneity is not handled. For the purpose of taking advantage of the favorable ideas of these two favored models together, partially linear mixed measurement error model (PLMMeM), a combination of LMM and the NPM under the measurement error problem, is defined by [3].

The PLMMeM based on a sample of size n with measured error in parametric part component X_i is considered as

$$\begin{aligned} Y_i &= X_i^T\beta + g(T_i) + Z_i^T b_i + \epsilon_i, \\ W_i &= X_i + U_i, \end{aligned} \qquad (1)$$

where fixed effects design matrices are $X_i = \left(x_{i1}, \dots, x_{ip}\right)^T$ and for $t_{i1}, \dots, t_{id}$ defined on $[0,1]$ $T_i = (t_{i1}, \dots, t_{id})^T$, random effects design matrix is $Z_i = \left(z_{i1}, \dots, z_{iq}\right)^T$, a parameter vector of fixed effects design matrix is $\beta = \left(\beta_1, \dots, \beta_p\right)^T$, an unknown function defined from $\mathbb{R}^d$ to $\mathbb{R}^1$ is $g(.)$, independent and identically distributed (i.i.d.), unobservable vector of the random effects design matrix is b_i and i.i.d. random vector of errors is ϵ_i. Independent b_i and ϵ_i are chosen from a Gaussian process with mean zero and covariance matrix D_i and Σ_i, respectively.

When X_i's are observable, the conditional distribution of Y_i for a given b_i is $Y_i|b_i \sim N\left(X_i^T\beta + g(T_i) + Z_i^T b_i, \Sigma_i\right)$. However, we observe W_i instead of observing X_i in model (1), assuming that the measurement error U_i has a known i.i.d. with mean zero and covariance matrix Σ_{uu} and independent of (Y_i, X_i, T_i, Z_i).

If we introduce the conditional expectations also known as the kernel regressions of Y, X and Z with bandwidth h, respectively, as

$$\omega_y(T_i) = E(Y_i|T_i),$$

$$\omega_x(T_i) = E(X_i|T_i),$$

$$\omega_z(T_i) = E(Z_i|T_i).$$

then the matrix form of model (1) is obtained as

DOI:10.21205/deufmd.2024267615

Geliş Tarihi / Received: 13.01.2023
Kabul Tarihi / Accepted: 31.07.2023

Atıf şekli / How to cite:
Kuran, Ö., Yalaz, S. (2024). Performance Assessment of the Modified Kernel Ridge Predictors in the Partially Linear Mixed Measurement Error Models via Covid-19 Data Analysis. DEUFMD, 26(76),134-140.

$$\tilde{Y} = \tilde{X}\beta + \tilde{Z}b + \epsilon, \quad (2)$$

where $\tilde{Y} = Y - \omega_y(T) = (\tilde{Y}_1, \dots, \tilde{Y}_n)^T$, $\tilde{X} = X - \omega_x(T) = (\tilde{X}_1, \dots, \tilde{X}_n)^T, \tilde{Z} = Z - \omega_z(T) = (\tilde{Z}_1, \dots, \tilde{Z}_n)^T$.

Letting $V = \tilde{Z}D\tilde{Z}^T + \Sigma$ be the covariance matrix of $\tilde{Y}$, [3] defined the Kernel estimator and the Kernel predictor under model (2) respectively, as

$$\hat{\beta} = (\widetilde{W}^T V^{-1}\widetilde{W} - tr(V^{-1})\Sigma_{uu})^{-1}(\widetilde{W}^T V^{-1}\tilde{Y}), \quad (3)$$

$$\hat{b} = D\tilde{Z}^T V^{-1}(\tilde{Y} - \widetilde{W}\hat{\beta}). \quad (4)$$

By denoting $\omega_{ni}(t) = \frac{1}{h_n}\int_{s_{i-1}}^{s_i} K(\frac{t-s}{h_n})ds$, $1 \le i \le n$, where $s_0 = 0$, $s_n = 1$ and $s_i = \frac{1}{2}(T_i + T_{i+1}), 1 \le i \le n-1$, where $K(.)$ is a kernel function, supported to have compact support and satisfy $\mathrm{supp}(K) = [-1,1]$, $\sup|K(x)| \le C < \infty$, $\int K(s)ds = 1$ and $K(s) = K(-s)$ and h_n is a sequence of bandwidth parameters which tends to zero as $n \to \infty$, $\widetilde{W} = (\widetilde{W}_1, \dots, \widetilde{W}_n)$ with $\widetilde{W}_i = W_i - \omega_w(T_i) = W_i - \sum_{j=1}^{n} \omega_{nj}(T_i)W_i$, $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)$ with $\tilde{Y}_i = Y_i - \omega_y(T_i) = Y_i - \sum_{j=1}^{n} \omega_{nj}(T_i)Y_i$ and $\tilde{Z} = (\tilde{Z}_1, \dots, \tilde{Z}_n)$ with $\tilde{Z}_i = Z_i - \omega_z(T_i) = Z_i - \sum_{j=1}^{n} \omega_{nj}(T_i)Z_i$, we can get the nonparametric function estimation as $\hat{g}(t) = E(Y_i - X_i\beta - Z_ib|T = t) = E(Y_i - W_i\beta - Z_ib|T = t) = \sum_{j=1}^{n} \omega_{nj}(t)(Y_j - W_j\hat{\beta} - Z_j\hat{b})$.

In the real data world, it is quite natural that strong linear dependence arises between the columns of $\tilde{X}$ and this linear dependence situation is called as multicollinearity. Under multicollinearity case, we may encounter some undesirable result like a large variance of $\hat{\beta}$ that deviates from its true value. To solve this undesirable result, estimators and predictors alternative to $\hat{\beta}$ and $\hat{b}$ can be suggested.

The most commonly preferred approach to overcome multicollinearity problem is the ridge approach [4] in LMs. By following [5-6] in LMMs, the Kernel ridge predictors which are the Kernel ridge estimator and predictor in PLMMeMs for a given ridge biasing parameter $k > 0$ are derived by [7] respectively, as

$$\hat{\beta}_k = (\widetilde{W}^T V^{-1}\widetilde{W} - tr(V^{-1})\Sigma_{uu} + kI_p)^{-1}(\widetilde{W}^T V^{-1}\tilde{Y}), \quad (5)$$

$$\hat{b}_k = D\tilde{Z}^T V^{-1}(\tilde{Y} - \widetilde{W}\hat{\beta}_k). \quad (6)$$

Then, using the Kernel ridge predictors given by Eqs. (5) and (6), the estimate of the ridge nonparametric function is obtained as $\hat{g}_k(t) = \sum_{j=1}^{n} \omega_{nj}(t)(Y_j - W_j\hat{\beta}_k - Z_j\hat{b}_k)$.

Another popular attempt is Liu's approach [8] in LMs. With the help of [9] in LMs and [10] in LMMs, the Kernel Liu predictors which are the Kernel Liu estimator and predictor in PLMMeMs for a given Liu biasing parameter $0 < d < 1$ are given by [11], respectively as

$$\hat{\beta}_d = (\widetilde{W}^T V^{-1}\widetilde{W} - tr(V^{-1})\Sigma_{uu} + \mathrm{I}_p)^{-1}(\widetilde{W}^T V^{-1}\tilde{Y} + \mathrm{d}\mathrm{I}_p), \quad (7)$$

$$\hat{b}_d = D\tilde{Z}^T V^{-1}(\tilde{Y} - \widetilde{W}\hat{\beta}_d), \quad (8)$$

Then, using the Kernel Liu predictors given by Eqs. (7) and (8), the estimate of the Liu nonparametric function is obtained as $\hat{g}_d(t) = \sum_{j=1}^{n} \omega_{nj}(t)(Y_j - W_j\hat{\beta}_d - Z_j\hat{b}_d)$.

Both Kernel ridge and Kernel Liu predictors given by Eqs. (5-8) are biased prediction approaches that are suggested by using some prior information [9] in order to eliminate the negative effects of the multicollinearity problem in PLMMeMs. In addition to these two approaches, our goal in this article is to propose a new biased prediction approach in PLMMeMs by taking a convex combination of Kernel ridge and Kernel Liu predictors as prior information. This new approach called the modified Kernel ridge is such a convex combined approach that it results in unifying the advantages of the Kernel ridge prediction and Kernel Liu prediction. Since it is a combination of both Kernel ridge and Kernel Liu approaches, it is thought to be more successful than Kernel ridge and Kernel Liu approaches in minimizing the negative effects of multicollinearity. Then, the rest of this paper is structured as follows: Section 2, the new predictors in PLMMeMs are characterized. In Section 3, we make some mean square error comparisons and Covid-19 data analysis under known measurement errors and covariance matrix is done in Section 4. In Section 5, a simulation study is also done. Finally, concluding remarks are given in Section 6.

2. The Modified Kernel Ridge Predictors

Our aim in this section is to suggest the modified Kernel ridge prediction approach for PLMMeMs using the idea of the modified ridge estimation in linear models [9] and in LMMs [12]. We know that under model (2), $\begin{bmatrix} b \\ \tilde{Y} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \tilde{X}\beta \end{bmatrix}, \begin{bmatrix} D & D\tilde{Z}^T \\ \tilde{Z}D & V \end{bmatrix}\right)$ which means that b and $\tilde{Y}$ are jointly Gaussian distributed. For a given b the conditional distribution of $\tilde{Y}$ is given as $\tilde{Y}|b \sim N(\tilde{X}\beta + \tilde{Z}b, \Sigma)$. Then, the joint density of $\tilde{Y}$ and b is

$$f(\tilde{Y}, b) = f(\tilde{Y}|b)f(b) = (2\pi)^{-(n+q)/2}|\Sigma|^{-1/2}|D|^{-1/2}$$

$$\times\, exp\left\{-\frac{1}{2}\left[\left(\tilde{Y} - \tilde{X}\beta - \tilde{Z}b\right)^T\Sigma^{-1}\left(\tilde{Y} - \tilde{X}\beta - \tilde{Z}b\right) + b^T D^{-1}b\right]\right\},$$

where $|.|$ denotes the determinant of a matrix. $logf(\tilde{Y}, b)$ is derived by dropping the constant term as

$$= logf(\tilde{Y}|b) + logf(b)$$

$$= -\tfrac{1}{2}\{\left[\left(\tilde{Y} - \tilde{X}\beta - \tilde{Z}b\right)^T\Sigma^{-1}\left(\tilde{Y} - \tilde{X}\beta - \tilde{Z}b\right) + b^T D^{-1}b\right]\},$$

and so, a penalization term with regularization parameter $\delta = -\frac{1}{2} \ge 0$ is added to $logf(\tilde{Y}, b)$,

$$logf(\tilde{Y}, b) - \tfrac{1}{2}k(1+d)\beta^T\beta. \quad (9)$$

Here, we use the prior information from [9] and [13] to chose the stochastic linear restriction $0 = \sqrt{k(1+d)}\beta + \varepsilon$. The partial derivatives of Eq. (9) with respect to the elements of β and b are taken equal to zero, then, by switching $\hat{\beta}$ and $\hat{b}$ by $\hat{\beta}_{k,d}$ and $\hat{b}_{k,d}$, respectively.

$$\tilde{X}^T\Sigma^{-1}(\tilde{Y} - \tilde{X}\hat{\beta}_{k,d}) - k(1+d)\hat{\beta}_{k,d} - \tilde{X}^T\Sigma^{-1}\tilde{Z}\hat{b}_{k,d} = 0, \quad (10)$$

$$\tilde{Z}^T\Sigma^{-1}(\tilde{Y} - \tilde{X}\hat{\beta}_{k,d}) - (\tilde{Z}^T\Sigma^{-1}\tilde{Z} + D^{-1})\hat{b}_{k,d} = 0, \quad (11)$$

are obtained. Eqs. (10) and (11) are also equal the matrix form given by

$$\begin{pmatrix} \tilde{X}^T\Sigma^{-1}\tilde{X} + k(1+d)I_p & \tilde{X}^T\Sigma^{-1}\tilde{Z} \\ \tilde{Z}^T\Sigma^{-1}\tilde{X} & \tilde{Z}^T\Sigma^{-1}Z + D^{-1} \end{pmatrix} \quad (12)$$

$$\times\begin{pmatrix}\hat{\beta}_{k,d}\\ \hat{b}_{k,d}\end{pmatrix}=\begin{pmatrix}\tilde{X}^T\Sigma^{-1}\tilde{Y}\\ \tilde{Z}^T\Sigma^{-1}\tilde{Y}\end{pmatrix}.$$

With [14]'s approach, Eq. (12) can be rewritten as

$$C\hat{\varphi}=\omega^T\Sigma^{-1}\tilde{Y}+\vartheta, \quad (13)$$

where $\hat{\varphi}=(\hat{\beta}_{k,d}^T,\hat{b}_{k,d}^T)^T$, $\omega=(\tilde{X},\tilde{Z})$, $\vartheta=(k(1+d)\beta^T,0^T)^T$ and $C=\omega^T\Sigma^{-1}\omega+D^{*+}$ is full rank with $D^*=\begin{bmatrix}\frac{1}{k(1+d)}I_p & 0\\ 0 & D\end{bmatrix}$ and $D^{*+}=\begin{bmatrix}k(1+d)I_p & 0\\ 0 & D^{-1}\end{bmatrix}$ where Moore–Penrose inverse shown by the superscript '+'. With the regularization of Eq. (13) $\hat{\varphi}$ is obtained as

$$\hat{\varphi}=C^{-1}\omega^T\Sigma^{-1}\tilde{Y}+C^{-1}\vartheta, \quad (14)$$

where C^{-1} is the inverse formula of the partitioned matrix [15]. After C^{-1} is found and putting into Eq. (14), the modified Kernel ridge estimator and predictor are obtained, respectively, as

$$\hat{\beta}_{k,d}=(\tilde{X}^TV^{-1}\tilde{X}+k(1+d)I_p)^{-1}(\tilde{X}^TV^{-1}\tilde{Y}),$$

$$\hat{b}_{k,d}=D\tilde{Z}^TV^{-1}(\tilde{Y}-\tilde{W}\hat{\beta}_{k,d}).$$

Since the disturbance of measurement error U is existed, we need to carry out correction for attenuation. Thus, we redefine the modified Kernel ridge estimator and predictor, respectively, as

$$\hat{\beta}_{k,d}=(\tilde{W}^TV^{-1}\tilde{W}-tr(V^{-1})\Sigma_{uu}+k(1+d)I_p)^{-1}(\tilde{W}^TV^{-1}\tilde{Y}), \quad (15)$$

$$\hat{b}_{k,d}=D\tilde{Z}^TV^{-1}(\tilde{Y}-\tilde{W}\hat{\beta}_{k,d}), \quad (16)$$

with an estimate of the modified Kernel ridge nonparametric component $\hat{g}_{k,d}(t)=\sum_{j=1}^{n}\omega_{nj}(t)(Y_j-W_j\hat{\beta}_{k,d}-Z_j\hat{b}_{k,d})$.

3. Some Mean Square Error Comparisons

Under specific matrices $L\in\mathbb{R}^{p\times s'}$ and $M\in\mathbb{R}^{q\times s'}$, we demonstrate the prediction of PLMMeM as $\mu=L^T\beta+M^Tb$ [16,17] for $s'=1$. By using Eqs. (3)-(8) and Eqs. (15) and (16), the predictors of μ under the Kernel, the Kernel ridge, the Kernel Liu and the modified Kernel ridge predictors are definable, respectively, as

$$\hat{\mu}=L^T\hat{\beta}+M^T\hat{b}=\mathbb{Q}\hat{\beta}+M^TD\tilde{Z}^TV^{-1}\tilde{Y},$$

$$\hat{\mu}_k=L^T\hat{\beta}_k+M^T\hat{b}_k=\mathbb{Q}\hat{\beta}_k+M^TD\tilde{Z}^TV^{-1}\tilde{Y},$$

$$\hat{\mu}_d=L^T\hat{\beta}_d+M^T\hat{b}_d=\mathbb{Q}\hat{\beta}_d+M^TD\tilde{Z}^TV^{-1}\tilde{Y},$$

$$\hat{\mu}_{k,d}=L^T\hat{\beta}_{k,d}+M^T\hat{b}_{k,d}=\mathbb{Q}\hat{\beta}_{k,d}+M^TD\tilde{Z}^TV^{-1}\tilde{Y},$$

where $\mathbb{Q}=L^T-M^TD\tilde{Z}^TV^{-1}\tilde{W}$.

The matrix mean square error (MMSE) criterion is used to compare the betterness of $\hat{\mu}$, $\hat{\mu}_k$, $\hat{\mu}_d$ and $\hat{\mu}_{k,d}$. By following [18], the MMSEs for $\hat{\mu}$, $\hat{\mu}_k$, $\hat{\mu}_d$ and $\hat{\mu}_{k,d}$ are calculated, respectively, as

$$MMSE(\hat{\mu})=\mathbb{Q}MMSE(\hat{\beta})\mathbb{Q}^T+M^T(D-D\tilde{Z}^TV^{-1}\tilde{Z}D)M,$$

$$MMSE(\hat{\mu}_k)=\mathbb{Q}MMSE(\hat{\beta}_k)\mathbb{Q}^T+M^T(D-D\tilde{Z}^TV^{-1}\tilde{Z}D)M,$$

$$MMSE(\hat{\mu}_d)=\mathbb{Q}MMSE(\hat{\beta}_d)\mathbb{Q}^T+M^T(D-D\tilde{Z}^TV^{-1}\tilde{Z}D)M,$$

$$MMSE(\hat{\mu}_{k,d})=\mathbb{Q}MMSE(\hat{\beta}_{k,d})\mathbb{Q}^T+M^T(D-D\tilde{Z}^TV^{-1}\tilde{Z}D)M,$$

where

$$MMSE(\hat{\beta})=N^{-1},$$

$$MMSE(\hat{\beta}_k)=N_k^{-1}NN_k^{-1}+k^2N_k^{-1}\beta\beta^TN_k^{-1},$$

$$MMSE(\hat{\beta}_d)=N_p^{-1}N_dN^{-1}N_dN_p^{-1}+(1-d)^2N_p^{-1}\beta\beta^TN_p^{-1},$$

$$MMSE(\hat{\beta}_{k,d})=NN_{k,d}^{-2}+(N_{k,d}-I_p)\beta\beta^T(N_{k,d}-I_p),$$ where

$N=(\tilde{W}^TV^{-1}\tilde{W}-tr(V^{-1})\Sigma_{uu})$, $N_p=(N+I_p)$,

$N_k=(\tilde{W}^TV^{-1}\tilde{W}-tr(V^{-1})\Sigma_{uu}+kI_p)$,

$N_d=(\tilde{W}^TV^{-1}\tilde{W}-tr(V^{-1})\Sigma_{uu}+dI_p)$,

$N_{k,d}=(\tilde{W}^TV^{-1}\tilde{W}-tr(V^{-1})\Sigma_{uu}+k(1+d)I_p)$. Then, the following theorems can be presented:

Theorem 3.1. The estimator $\hat{\beta}_{k,d}$ dominates the estimator $\hat{\beta}$ in the MMSE sense iff

$$\beta^T(N_{k,d}-I_p)^T(N^{-1}-N_{k,d}N^{-1}N_{k,d}^T)^{-1}(N_{k,d}-I_p)\beta<1.$$

Theorem 3.2. The estimator $\hat{\beta}_{k,d}$ dominates the estimator $\hat{\beta}_k$ in the MMSE sense iff

$$\beta^T(N_{k,d}-I_p)^T((N_k^{-1}NN_k^{-1}-N_{k,d}N^{-1}N_{k,d}^T)+k^2N_k^{-1}\beta\beta^TN_k^{-1})^{-1}(N_{k,d}-I_p)\beta<1.$$

Theorem 3.3. The estimator $\hat{\beta}_{k,d}$ dominates the estimator $\hat{\beta}_d$ in the MMSE sense iff

$$\beta^T(N_{k,d}-I_p)^T((N_p^{-1}N_dN^{-1}N_dN_p^{-1}-N_{k,d}N^{-1}N_{k,d}^T)+(1-d)^2N_p^{-1}\beta\beta^TN_p^{-1})^{-1}(N_{k,d}-I_p)\beta<1.$$

For the proofs of the theorems 3.1, 3.2 and 3.3 in PLMMeM, [12, p.37] in LMM can be examined. We modified [12, p.37]'s proofs which are obtained for LMM to our model PLMMeMs.

4. Covid-19 Data Analysis

For real data analysis the data taken from the Vaccine Tracker [19] submitted to ECDC through The European Surveillance System (TESSy) twice a week by European Union/European Economic Area (EU/EEA) countries. The data includes the number of vaccine doses distributed by manufacturers to the country, the number of first, second and unspecified (number of doses not known whether it was a first or second dose) doses. These vaccines are administered by age groups which are children (<18), adolescent and adult population (18+).

In this data application, we use 187 age-specific 14-day notification rate of reported Covid-19 cases per 100000 population (rate_14_day_per_100k) selected randomly from the countries Belgium, Czechia, Denmark, Estonia, Ireland, Greece, Austria, Hungary, Italy, Spain, Slovenia, Slovakia, Portugal, Malta, Norway, Luxembourg, Netherlands (17 regions) which were regularly obtained during the week periods 2021-W01, 2021-W05, 2021-W09, 2021-W13, 2021-W17, 2021-W21, 2021-W25, 2021-W29, 2021-W33, 2021-W37, 2021-W40. To determine rate_4_week_per_100k, we employ repeated measurements from denominator (25-49 years old population), first dose, second dose and vaccine name. Our Covid-19 data can be extracted from an official website of the European centre for disease prevention and control [20].

Rate_4_week_per_100k is defined as the response (y), denominator (x_1), first dose (x_2) and second dose (x_3) are obtained as the explanatory variables (fixed effects) and vaccines (t) are expressed as nonparametric variable. When we specify nonparametric part, firstly, we look at which vaccines are used in the determined weeks in our data set and these vaccines are COM = Pfizer/BioNTech, AZ = AstraZeneca, MOD = Moderna, BECNBG = Beijing CNBG, JANSS = Janssen, SPU = Sputnik V and UNK=

UNKNOWN. Secondly, we specify the vaccine efficacy rate of each vaccine (COM=0.95, AZ=0.7469, MOD=0.941, JANSS=0.7735, SPU=0.9760, BECNBG=0.7934 and UNK=0.8583 are computed as by taking the geometric mean of the COM, AZ, MOD, BECNBG, JANSS, SPU vaccines). And lastly, we create the nonparametric part by taking the geometric mean of the efficacy rates of the vaccines. Since the regions are randomly selected from 17 countries, random effect is explained as the regions. We log-transform the variables to make the data conform more closely to the normal distribution and to improve the model fit since the distribution of Covid-19 data is right skewed. Then, the PLMMeM is written as, for $i = 1, \dots, 17, j = 1, \dots, 11$,

$$y_{ij} = \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3} + b_1 + b_2 time_{ij} + g(t_{ij}) + \varepsilon_{ij},$$

where the ith observation of the jth region of the explanatory variable (x_s, $s = 1,2,3$) is indicated as x_{ijs}, the jth observation of the ith region of the response is indicated as y_{ij} and time corresponding to y_{ij} is demonstrated as $time_{ij}$. h is chosen using optimal bandwidth selection rule given by [21]. We use the quartic kernel function $K(u) = (15/16)(1 - u_j^2)^2 I(|u_j| \leq 1)$ for kernel smoothing regression with the measurement error which has normal distribution $U \to N(0,0.25)$.

In our real data analysis, we select the restricted maximum likelihood (REML) method which has the smallest AIC/BIC values for all models from Table 1. The UN(1) variance-covariance model under AIC and BIC is seen as the best model in modeling the variance-covariance matrix structure with respect to the response.

Table 1. Variance-covariance matrix results

Cov. struc.	*Est. Met. for Cov. Par.*	*AIC*	*BIC*
UN	ML	568.68	594.35
	REML	552.50	578.35
UN(1)	ML	567.91	590.38
	REML	551.94	574.56
VC	ML	570.43	589.68
	REML	555.51	574.90
CS	ML	571.23	593.70
	REML	556.08	578.70

The abbreviations "*Cov. Struc.*" and "*Est. Met. for Cov.Par.*" refer to "*Covariance Structures*" and "*Estimation Methods for Covariance Parameters*"

$\widehat{D}_{REML}$ and $\widehat{\Sigma}_{REML}$values given in Table 2. Then, via $V = \tilde{Z} D \tilde{Z}^T + \Sigma$ formula, $\widehat{V}_{REML}$ values are found.

Table 2. Covariance structures estimates

$\widehat{D}_{REML}$	$\begin{bmatrix} 0.0742 & 0 \\ 0 & 0.4591 \end{bmatrix}$
$\widehat{\Sigma}_{REML}$	$0.8855 I_{187}$

$\lambda_1 = 0.0162e^{+03}$, $\lambda_2 = 2.1292e^{+03}$,$\lambda_3 = 0.6513e^{+03}$ and $\lambda_4 = 0.1312e^{+03}$ are obtained from the matrix $\tilde{X}^T \widehat{V}_{REML}^{-1} \tilde{X}$. The condition number calculated as $\frac{\lambda_{\max}}{\lambda_{\min}} = 132.076$ is used to measure the extent of multicollinearity and $\frac{\lambda_{\max}}{\lambda_{\min}} > 100$ shows moderate multicollinearity.

We determine the estimators of the biasing parameters k and d with a computational algorithm as follows:

1. By using for each $\lambda_{\rm i}$ value, the k value is estimated from [7] for PLMMeMs as $\hat{k} = \hat{k}_{LW} = \frac{p}{\sum_{i=1}^{p} \lambda_{\rm i} \hat{\beta}_{\rm i}^2} = 12.4055$ when the covariance parameters are estimated by REML.
2. After the $\hat{k}$ value is found from the point 1, the Liu biasing parameter d is selected as $\hat{d}_h$ which is given by Theorem 4.2 [11] where h is determined as multiplying the upper bound defined in Theorem 4.2 by 0.99 if $\sum_{i=1}^{p} \frac{1}{\lambda_{\rm i}(1+\lambda_{\rm i})} > \frac{2}{\lambda_{\rm p}(1+\lambda_{\rm p})}$.
3. If $\sum_{i=1}^{p} \frac{1}{\lambda_{\rm i}(1+\lambda_{\rm i})} < \frac{2}{\lambda_{\rm p}(1+{\rm p})}$, we determined arbitrarily $\hat{d}_h$ as 0.9705 .

The estimates of the fixed parameters and nonparametric function, the predictions of the random parameters and the scalar mean square error (SMSE) values for Kernel, Kernel ridge, Kernel Liu and modified Kernel ridge cases under PLMeM are presented in Table 3.

We see that in Table 3 the modified Kernel ridge estimator has better results in the sense of SMSE for $\hat{k}_{LW} = 12.4055$ and $\hat{d}_h = 0.9705$ than the Kernel, Kernel ridge and Kernel Liu estimators. Moreover, we calculate the conditions given by Theorems 3.1, 3.2 and 3.3, respectively, as -0.3274, -0.3273 and -0.3296, which are smaller than 1. Thus, we also say that the modified Kernel ridge estimator dominates the Kernel, Kernel ridge and Kernel Liu estimators on the MMSE criterion.

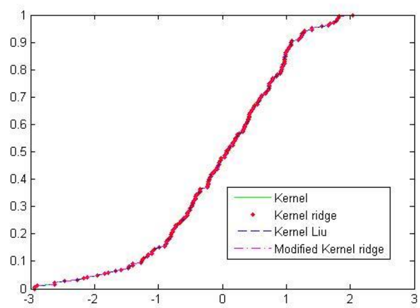

Figure 1. Comparison of the finite sample and asymptotic distributions of the estimators

Additionally, comparison between the asymptotic distributions of Kernel (green), Kernel ridge (red), Kernel Liu (blue), modified Kernel ridge (magenta) estimators and the finite sample properties are also examined. In Figure 1 where the abscissa is $Z = (Var(g(t, h_n)))^{-1/2}(g(t, h_n) - E(g(t, h_n)))$ and the ordinate is probability. The empirical cumulative distribution functions (CDFs) of the estimators agree very well with the normal CDFs.

Table 3. Data analysis results

	Kernel	*Kernel ridge*	*Kernel Liu*	*Modified Kernel ridge*
$\hat{\beta}_0$	-4.0588e^{-10}	-2.3038e^{-10}	-4.0519e^{-10}	-1.6234e^{-10}
$\hat{\beta}_1$	0.0186	0.0168	0.0186	0.0153
$\hat{\beta}_2$	-0.0172	-0.0161	-0.0172	-0.0152
$\hat{\beta}_3$	-0.0095	-0.0094	-0.0095	-0.0093
$\hat{b}_1$	-1.3563e^{-16}	-1.6294e^{-10}	-6.4321e^{-13}	-2.2625e^{-10}
$\hat{b}_2$	-0.6394	-0.6401	-0.6394	-0.6408
$\hat{g}$	6.3563	6.3727	6.3563	6.3860
SMSE	0.0713	0.0281	0.0711	0.0171

Table 4. Estimated and predicted MSE values with $g_1(t)$ function

m	γ^2	$\hat{\beta}$	$\hat{\beta}_{\hat{k}_{LW}}$	$\hat{\beta}_{\hat{d}}$	$\hat{\beta}_{\hat{k}_{LW},\hat{d}}$	$\hat{b}$	$\hat{b}_{\hat{k}_{LW}}$	$\hat{b}_{\hat{d}}$	$\hat{b}_{\hat{k}_{LW},\hat{d}}$
	0.90	1.309997	1.309915	1.309914	1.283515	0.008436	0.008435	0.008431	0.008316
15	0.95	0.779188	0.779120	0.779119	0.761362	0.009774	0.009773	0.009769	0.009614
	0.99	0.830462	0.830393	0.830392	0.811062	0.010974	0.010973	0.010969	0.009924
	0.90	0.947033	0.947019	0.947018	0.938671	0.002307	0.002306	0.002304	0.002219
30	0.95	0.696424	0.696413	0.696412	0.689989	0.001687	0.001686	0.001682	0.001537
	0.99	0.633088	0.633077	0.633076	0.626945	0.003710	0.003709	0.003705	0.003582
	0.90	0.976395	0.976391	0.976390	0.971870	0.456936$\times 10^{-3}$	0.456935$\times 10^{-3}$	0.456931$\times 10^{-3}$	0.456721$\times 10^{-3}$
60	0.95	0.912487	0.912484	0.912483	0.908417	0.420959$\times 10^{-3}$	0.420958$\times 10^{-3}$	0.420957$\times 10^{-3}$	0.420860$\times 10^{-3}$
	0.99	0.804819	0.804815	0.804814	0.800992	0.426339$\times 10^{-3}$	0.426339$\times 10^{-3}$	0.426335$\times 10^{-3}$	0.426144$\times 10^{-3}$

Table 5. Estimated and predicted MSE values with $g_2(t)$ function

m	γ^2	$\hat{\beta}$	$\hat{\beta}_{\hat{k}_{LW}}$	$\hat{\beta}_{\hat{d}}$	$\hat{\beta}_{\hat{k}_{LW},\hat{d}}$	$\hat{b}$	$\hat{b}_{\hat{k}_{LW}}$	$\hat{b}_{\hat{d}}$	$\hat{b}_{\hat{k}_{LW},\hat{d}}$
	0.90	1.063705	1.063638	1.063631	1.043362	0.008777	0.008776	0.008772	0.008604
15	0.95	0.987899	0.987841	0.987840	0.969730	0.026092	0.026091	0.026090	0.025986
	0.99	1.017534	1.017466	1.017465	0.995898	0.013327	0.013326	0.013324	0.013227
	0.90	0.840638	0.840625	0.840623	0.833181	0.001965	0.001964	0.001962	0.001945
30	0.95	0.800622	0.800609	0.800605	0.792983	0.002950	0.002949	0.002946	0.002845
	0.99	0.626170	0.626158	0.626157	0.619923	0.003990	0.003989	0.003985	0.003872
	0.90	0.985321	0.985317	0.985316	0.980775	0.419945$\times 10^{-3}$	0.419944$\times 10^{-3}$	0.419941$\times 10^{-3}$	0.419842.$\times 10^{-3}$
60	0.95	0.896444	0.896441	0.896440	0.892368	0.421394$\times 10^{-3}$	0.421393$\times 10^{-3}$	0.421391$\times 10^{-3}$	0.421290$\times 10^{-3}$
	0.99	0.843971	0.843968	0.843967	0.839810	0.475318$\times 10^{-3}$	0.475317$\times 10^{-3}$	0.475311$\times 10^{-3}$	0.475303$\times 10^{-3}$

5. A Simulation Study

In this section, we will investigate the performances of Kernel, Kernel ridge, Kernel Liu and modified Kernel ridge estimators in the sense of the estimated mean square error (EMSE) and the performances of Kernel, Kernel ridge, Kernel Liu and modified Kernel ridge predictors in the sense of the predicted mean square error (PMSE) under known covariance matrix.

By following [22], the fixed effects are calculated as

$$x_{ijk} = (1-\gamma^2)^{1/2} w_{ijk} + \gamma w_{ijp+1}, i = 1, \dots, m, \quad j = 1, \dots, n_i, k = 1, \dots, p,$$

where w_{ijk} are independent standard normal pseudo-random numbers and γ is specified so that the correlation between any two fixed effects is given by $\gamma^2 = 0.90, 0.95, 0.99$. And, the fixed effects number size is selected as $p = 3$.

We think $m = 15,30,60$ subjects and $n_i = 10$ observation per subject and then, we report the simulation results with the sample sizes of $n = \sum_{i=1}^{m} n_i = 150,300,600$. The parameter vector $\beta = (\beta_1, \dots, \beta_p)^T$ is chosen as the normalized eigenvector corresponding to the largest eigenvalue of $\tilde{X}^T V^{-1} \tilde{X}$ so that $\beta^T \beta = 1$ (see [23]). Then, the underlying model takes the following form with $q = 2$ random effects

$$y_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3} + b_1 + b_2 time_{ij} + g(t_{ij}) + \varepsilon_{ij},$$

$$b_i \overset{iid}{\sim} N(0, D), \varepsilon_{ij} \overset{iid}{\sim} N(0, I_{n_i})$$

where $D = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ is the AR(1) process with $\rho = 0.99$ and $time_{ij}$ shows time which was taken as the same set of occasions, $\{t_{ij} = j$ for $i = 1, \dots, m, \ j = 1, \dots, n_i\}$. $\hat{k}$ and $\hat{d}$ are selected as used in the Covid-19 data analysis.

We think two functions that the first is the piecewise linear continuous function $g_1(t) = S(t)$ as an example of the ordinarily smooth nonparametric function and the second is the error function $g_2(t) = erf\ (t)$ as an example of the supersmooth nonparametric function. Supposing that $T \to Uniform[0,1]$, $h_n^{-1} = 1.2(\text{In}\ n)^{0.25}$ and using the Gaussian Kernel function $K(u) = \frac{e^{-\frac{u_j^2}{2}}}{\sqrt{2\pi}}$ for the models have nonparametric part, we examine these models with the measurement error which has normal distribution $U \to N(0,0.5)$.

For each choice of m, γ and $g(t)$, the experiment is replicated 500 times by generating response variable and the EMSE for any estimator $\tilde{\beta}$ of β and the PMSE for any predictor $\tilde{b}$ of b are calculated, respectively, as

$$EMSE(\tilde{\beta}) = \frac{1}{500} \sum_{r=1}^{500} (\tilde{\beta}_r - \beta)^T (\tilde{\beta}_r - \beta),$$

$$PMSE(\tilde{b}) = \frac{1}{500} \sum_{r=1}^{500} (\tilde{b}_r - b)^T (\tilde{b}_r - b),$$

where the subscript r refers to the rth replication.

The simulation results are summarized in Tables 4 and 5. When we examine the results of Tables 4 and 5, we see that EMSE values of the modified Kernel ridge estimator and PMSE values of the modified Kernel ridge predictor are smaller than the others in all conditions. However, this superiority situation is more clearly be seen in large sample (for 600) and high correlation value (for 0.99). Additionally, we can also say that the superiority of the estimators/predictors over each other may vary depending on the selection of the biasing parameters.

6. Concluding Remarks

In this article, the modified Kernel ridge predictors have been studied with their MMSE comparisons under multicollinearity in PLMMeMs. To show the theoretical results, a Covid-19 analysis and a simulation study are given and these analyses demonstrate that although the modified Kernel ridge estimator is better than the Kernel, Kernel ridge and Kernel Liu estimators, the superiority of the modified Kernel ridge estimator depends on the chosen values of the biasing parameters.

Ethics committee approval and conflict of interest statement

This article does not require ethics committee approval.

This article has no conflicts of interest with any individual or institution.

Acknowledgment

The authors would like to thank the editor, associate editor, and anonymous referees for their constructive comments.

Author Contribution Statement

All authors contributed equally to this manuscript and they have accepted responsibility for the entire content of this manuscript and approved its submission.

References

[1] Laird, N.M., Ware, J.H. 1982. Random-Effects Models for Longitudinal Data, Biometrics, Vol. 38, No. 4 p. 963-974, DOI: 10.2307/2529876.
[2] Fuller, W.A. 1987.Measurement Error Models, John Wiley and Sons, Inc., Hoboken, NJ.
[3] Yalaz, S., Kuran, Ö. 2021. Kernel Estimator and Predictor of Partially Linear Mixed-Effect Errors-in-Variables Model, Journal of Statistical Computation and Simulation, Vol. 91, No. 5, p. 934–951, DOI:10.1080/00949655.2020.1836642.
[4] Hoerl, A.E., Kennard, R.W. 1970. Ridge Regression Biased Estimation for Nonorthogonal Problems, Technometrics, Vol. 12, No. 1, p. 55–67, DOI: 10.2307/1267351.
[5] Liu, X.Q., Hu, P. 2013. General Ridge Predictors in a Mixed Linear Model, Journal of Theoretical and Applied Statistics, Vol. 47, No. 2, p. 363-378, DOI: 10.1080/02331888.2011.592190.
[6] Özkale, M.R., Can, F. 2017. An Evaluation of Ridge Estimator in Linear Mixed Models: An Example from Kidney Failure Data, Journal of Applied Statistics, Vol. 44, No. 12, p. 2251-2269, DOI: 10.1080/02664763.2016.1252732.
[7] Kuran, Ö., Yalaz, S. 2022. Kernel Ridge Prediction Method in Partially Linear Mixed Measurement Error Model, Communications in Statistics - Simulation and Computation, Vol., No., p. 1–22, DOI:10.1080/03610918.2022.2075389.
[8] Liu, K. 1993. A New Class of Biased Estimate in Linear Regression, Communications in Statistics - Theory and Methods,Vol. 22, No.2, p. 393–402, DOI: 10.1080/03610929308831027.
[9] Swindel, B.F. 1976. Good Estimators Based on Prior Information, Communications in Statistics -Theory and Methods, Vol. 5, No.11,p. 1065-1075, DOI: 10.1080/03610927608827423.
[10] Özkale, M.R., Kuran, Ö. 2020. A Further Prediction Method in Linear Mixed Models: Liu Prediction, Communications in Statistics - Simulation and Computation, Vol. 49, No. 12, p.3171–3195, DOI: 10.1080/03610918.2018.1535071.
[11] Kuran, Ö., Yalaz, S. 2022. Kernel Liu Prediction Approach in Partially Linear Mixed Measurement Error Models, Statistics: A Journal of Theoretical and Applied Statistics, Vol. 56, No. 6, p. 1385-1408, DOI: 10.1080/02331888.2022.2152816.
[12] Kuran, Ö. 2020. The Modified Ridge Prediction Method in the Linear Mixed Models, 3nd International Congress on Statistics, Mathematics and Analytical, 12-13 March-2020, İstanbul, Turkey, 33-43.
[13] Özkale, M.R., Kaçıranlar, S. 2007. The Restricted and Unrestricted Two-Parameter Estimators, Communications in Statistics - Theory and Methods Vol. 36, No.15, p. 2707–2725, DOI: 10.1080/03610920701386877.
[14] Gilmour, A.R., Thompson, R., Cullis, B.R. 1995. Average Information REML: An Efficient Algorithm for Variance Parameter Estimation in

Linear Mixed Models, Biometrics, Vol. 51, No. 4, p. 1440-1450, DOI:10.2307/2533274.
[15] Searle, S.R. 1982. Matrix Algebra Useful for Statistics, John Wiley and Sons, New York.
[16] Pereira, L.N., Coelho, P.S. 2012. A Small Area Predictor Under Area-Level Linear Mixed Models with Restrictions, Communications in Statistics - Theory and Methods, Vol. 41, No. 13-14, p. 2524-2544, DOI:10.1080/03610926.2011.648000.
[17] Robinson, G.K. 1991. That BLUP is a Good Thing: The Estimation of Random Effects (with Discussion), Statistical Science, Vol. 6, No. 1, p. 15-51, DOI: 10.1214/ss/1177011926.
[18] Štulajter, F. 1997. Predictions in Nonlinear Regression Models, Acta Math Univ Comenian, Vol. LXVI, No. 1, p. 71–81.
[19] European Centre for Disease Prevention and Control, COVID-19 Vaccine Tracker. Available at: https://vaccinetracker.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html, 2021.
[20] European Centre for Disease Prevention and Control. Download COVID-19 Datasets.Available at: https://www.ecdc.europa.eu/en/covid-19/data, 2021.
[21] Bowman, A., Azzalini, A. 1997. Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-plus Illustrations, New York, Oxford.
[22] McDonald, G. C., and D. I. Galarneau. 1975. A Monte Carlo Evaluation of Some Ridge-Type Estimators. Journal of the American Statistical Association, Vol. 70, No. 350, p. 407–416, DOI: 10.1080/01621459.1975.10479882.
[23] Newhouse, J. P., and S. D. Oman. 1971. An Evaluation of Ridge Estimators. 1–28. Rand Corporation R-716-PR. https://www.rand.org/content/dam/rand/pubs/reports/2007/R716.pdf