Multivariate Partial Linear Modeling for Groundwater Quality Anaiysis

Huda Yahay Ahmed and Emad Hazim Aboudi
Department of Statistics, College of Administration and Economics, University of Baghdad, 10006 Baghdad, Iraq
{huda.yahia2201p, emadhazim } @coadec.uobaghdad.edu.iq

Keywords: Profile Least Squares, Profile Likelihood, Multivariate Partial Linear, Kernel Smoother.

Abstract: This research aims to estimate the multivariate partial linear model using two methods: Profile Least Squares and Profile Likelihood, with the goal of identifying the factors affecting the decline in groundwater quality in two geographical locations within Baghdad city (Abu Ghraib and Mahmudiya). The main problem lies in the presence of several response variables that are correlated with each other within the same model, which complicates the process of accurate estimation of the parameters. Additionally, there are correlations between the response variables and the explanatory variables, further challenging the estimation process. The model includes both linear and nonlinear variables, adding to the complexity of the analysis. Total dissolved solids and ion ratios were used as response variables, while the independent variables included sodium, chloride, and sulfate. The geographical location was incorporated as a nonparametric component in the analysis. The two methods were compared based on the Mean Squared Error (MSE) criterion to determine the preferred method for model estimation. Real data were used to verify the effectiveness of the methods in practical applications, and the results indicated the superiority of the Profile Least Squares method in providing more accurate and reliable estimates of the model.

1 INTRODUCTION

The study of several variables simultaneously is the focus of the field of multivariate analysis. In order to find patterns and trends that might not be apparent when examining each variable separately, it is used to investigate relationships among a group of variables. This kind of analysis is so widely used in many fields, including economics, marketing, healthcare, and the social sciences, and it is a potent tool in applied statistics. In order to make data-driven decisions and provide more accurate and trustworthy results, it offers a strong foundation for comprehending intricate interactions between independent and structured variables.

The majority of contemporary economic models are built using multivariate Partial linear regression models, which incorporate elements of both nonparametric and parametric multiple linear regressions. In order to characterize basic phenomena having direct effects, the parametric (linear) component represents a regression function that is considered to be linear in the explanatory variables. The nonparametric (nonlinear) component, on the other hand, is an unknown smoothing function that captures latent nonlinear interactions in the data (W. Härdle and M. Müller 1997) [1].

With a focus on the integration of linear and nonlinear components in the analysis of intricate interactions between explanatory and response variables, Härdle and Müller investigated semi-parametric model analysis employing kernel techniques for density estimation and multivariate regression. For multivariate data analysis,(Lawrence D. Brown et al. 2016) [2] also used kernel density estimation techniques.

Using a difference sequence approach, (Michael Levine 2019) [3] estimated partially linear regression models. In this approach, the nonlinear component is approximated using the Nadaraya–Watson estimator to capture the unknown function, while the linear component is estimated based on differences.

The idea put up by (Jun Zhang et al. 2021) [4] was to estimate the unknown function in Partial linear regression models by splitting a multidimensional grid into sub cells and computing the median inside each cell. Furthermore, because wavelet approaches are accurate and can handle complex, nonlinear interactions in multivariate data, they have been used as a strong and effective tool for estimating unknown functions.

(Meng and Huang 2024) [5] conducted more current research on variable selection in semi-functional partly linear regression models for time series data; however, they only examined temporal variables and did not take environmental multivariate applications (MDPI) into account. (Wu and Tong 2024) [6] focused on inferential mechanisms instead of environmental applications while analyzing a deep Partial linear Cox model for survival data (OUP Academic).

In light of the dearth of applied research comparing estimation techniques like profile likelihood and profile least squares, as well as the limited number of applied studies on multivariate partially linear regression models using actual environmental data for evaluating groundwater quality in Iraq, we applied this modeling framework to data from 36 wells in Baghdad. This offers an uncommon real-world case and uses the mean squared error (MSE) criterion for performance comparison, which hasn't been often discussed in environmental applications at the local level.

Thus, by providing a useful implementation of a multivariate response model, this study helps close the knowledge gap by addressing both linear and nonlinear influences on groundwater quality. In the domains of environmental science and water resources, this method enhances statistical literature and fortifies the link between research and environmental practice.

2 MULTIVARIATE PARTIAL LINEAR MODEL

The multivariate partial linear model is a regression model that has both parametric and nonparametric effects. It represents a multivariate response variable as a linear function of some explanatory variables while allowing other variables to be nonlinearly related to the response variable. Many researchers and practitioners find this model a significant multivariate semi parametric regression model, and it is also a special case of additive models.

The extension of ordinary regression techniques is a benefit of this model over nonparametric models as it solves the curse of dimensionality. This means MPLM is a better and stronger way to model complex data that involves more than one variable[7].

Even though multivariate partially linear regression is used, it gives rise to challenges like intense computation in estimation and problem in understanding results. It happens particularly when nonlinear relationships among many variables are present. Nevertheless, it has become more usable due to the development of computation. Because of this; multivariate partial linear regression has emerged as a crucial model for statistical analysis in weird and high-dimensional settings. It is applies to economics, finance, medicine, biomedical research, environment, climate studies, and many others [8].

The multivariate partial linear regression model is written according to the following equation [2]:

$$Y_k = X_i^T\beta + g(T_i) + \epsilon_k , k = 1,2, \dots , q \quad (1)$$

q represents the number of response variables, $X_i^T\beta$ Represents the parametric part of the model,X: A non-random matrix representing the observations of the parametric explanatory variables of degree $(n \times p)$ and p represents the number of the parametric explanatory variables, Β Matrix of model parameters of the degree $(p \times q)$, (T_i) Represents the nonparametric part of the model, which is an unknown smoothing function of degree$(n \times q)$, ϵ_k Matrix of random errors of degree $(n \times q)$,The random error is distributed with zero mean and variance Σ [8].

3 ESTIMATION METHODS

3.1 Profile Least Square Method

This methodology aims to improve the accuracy of predicting dependent variables, offering a computationally efficient and effective approach. It is particularly advantageous for analyzing high-dimensional datasets, as it facilitates dimensionality reduction and enhances the understanding of the interrelationships between variables. Developed by Swedish statistician Herman Wold in the 1970s, this method extends classical statistical techniques, such as multiple linear regression, with a specific emphasis on addressing issues related to Multivariate correlation among the independent variable, [9], [10]. In 2017, Yaowu Zhang & Liping Zhu estimated the multivariate partial linear model using the Profile Least Square Method. This method can be applied to the model by following the following steps [10] :

Suppose that (X_i, T_i, Y_i) $, i = 1, \dots, n$ a random sample according to (1) in respect of (X, T, Y). We introduce profile least-squares approaches to estimating β and $g(T)$ in model (1).

First, we assume that the coefficients β_k (the coefficients associated with the independent variables) are known in advance, which simplifies the estimation process, as we focus only on estimating the part $g_k(T_i)$, We use kernel estimation to estimate $g_k(T_i)$, where a kernel function (such as a Gaussian kernel) is used to estimate the values of $g_k(T_i)$ based on the values of T_i. Reformulate the model to take the following form:

$$Y_{ik} - X_i^T\beta_k = g_k(T_i) + \epsilon_{ik}, \quad i = 1,2,\dots,n, k = 1,2,\dots,q.$$

Let $Kh(\cdot) = K(\cdot/h)/h$ where $K(\cdot)$ is a kernel function and h is a bandwidth. To ease Subsequent actions, we denote by $\otimes$ the Kronecker product, 1_q a $q \times 1$ vector with all entries being one and I_q an $q \times q$ identity matrix. Let $X = (X_1,\dots,X_n)^T \otimes 1_q, Y = (Y_1^T,\dots,Y_n^T)^T$. Through pretending β_k were known, the weighted kernel estimator of $g_k(T_i)$ at an arbitrary grid point t is then defined by [10].

$$\hat{g}(t;\beta,W) = \sum_{i=1}^{n} K_h(T_i - t)1_q^T U(Y_{ik} - X_i^T\beta_k) \Big/ 1_q^T U 1_q \sum_{i=1}^{n} K_h(T_i - t).$$

Where U is a symmetric and positive-definite weight matrix. The kernel estimator of g(t) as $\hat{g}(t;\beta,U)$ this estimator depends upon β and U. To Estimate g, it thus remains to specify a weight matrix U and estimate β.

Let $K(t) = \mathrm{diag}(K_h(T_1 - t),\dots,K_h(T_n - t) \otimes U$, $K(t)$of degree $(n q \times nq)$

An equivalent matrix form of $\widehat{m}(t;\beta,U)$ is:

$$\hat{g}(t;\beta,U) = \{1_{nq}^T K(t) 1_{nq}\}^{-1}\{1_{nq}^T K(t)(Y - X\beta)\}. \quad (2)$$

We further write

$$S = \begin{pmatrix} \{1_{nq}^T K(T_1) 1_{nq}\}^{-1} 1_{nq}^T K(T_1) \\ \vdots \\ \{1_{nq}^T K(T_n) 1_{nq}\}^{-1} 1_{nq}^T K(T_n) \end{pmatrix} \otimes 1_q$$

Substituting $\hat{g}(t;\beta,U)$, for $t = T_1,\dots,T_n$, into model (1), we obtain

$$(I_{nq} - S)Y \approx (I_{nq} - S)X\beta + e.$$

Where $e = (\varepsilon_{11},\dots,\varepsilon_{1q},\dots,\varepsilon_{n1},\dots,\varepsilon_{nq})^T$ .To estimates β, we consider minimizing the following weighted least-squares loss function:

$$RSS(\beta) = \{Y - X\beta\}^T(I_{nq} - S)^T W(I_{nq} - S)\{Y - X\beta\}, \quad (3)$$

Where W is a symmetric, positive -definite weight matrix. We will show the effect of W on the resulting estimator of β in our subsequent context. By minimizing (3), the resulting estimator of β is as follows:

$$\hat{\beta}(W) = \left\{X^T(I_{nq} - S)^T W(I_{nq} - S)X\right\}^T \left\{X^T(I_{nq} - S)^T W(I_{nq} - S)Y\right\} \quad (4)$$

In other words, $\hat{\beta}(W)$ is a function of W. To estimate β, define the weight matrix W. The weight matrices W and U are defined in accordance with the regression model or any other mathematical application, where symmetric and positive-definite weight matrices are used, W is defined as the Kronecker Product between the identity matrix $I_n$and the positive-definite matrix V. This means:

$$W = I_n \otimes V$$

Whereas:

- I_n Is the identity matrix of size $n \times n$, which is a matrix that contains 1's on the main diagonal (from top left to bottom right) and 0's in the other positions.
- V Is a symmetric and positive-definite matrix of size $q \times q$. "Positive-definite" means that the eigenvalues of the matrix V are all greater than zero.
- $\otimes$ A mathematical operation that produces a new matrix by multiplying each element of the first matrix by the second matrix. In this context, I_n is the identity matrix of size $n \times n$ and V is a matrix of size $q \times q$, thus the Kronecker product $I_n \otimes V$ will result in a matrix of size $(nq \times nq)$.

We assume the observations$\{(X_i,T_i,Y_i), i = 1,\dots,n\}$ are independent. It is clear that W will be symmetric and positive-definite as long as V is so. Now, it remains to choose the working weight matrices V and U which should be selected based on the context and available data. We propose two alternatives; our first option is to choose $W = I_n \otimes I_q$ $(V = I_q)$ to produce an Initial estimator $\hat{\beta}(I_n \otimes I_q)$. we also set $U = I_q$ to obtain $\hat{g}\{t;\hat{\beta}(I_n \otimes I_q), I_q\}$. Our second Option is to choose $W = I_n \otimes \Sigma^{-1}$ $(V = \Sigma^{-1})$ and $U = \Sigma^{-1}$,where $\Sigma = E(\varepsilon\varepsilon^T)$ stands for the covariance matrix of $\varepsilon = (\varepsilon_1,\dots,\varepsilon_q)^T$.In practice Σ is typically unknown. In this case, we must estimate Σ from the observations. Towards this goal, we define $\hat{\varepsilon}_i = (\hat{\varepsilon}_{i1},\dots,\hat{\varepsilon}_{iq})^T$, where$\hat{\varepsilon}_{ik} = Y_{ik} - X_i^T\hat{\beta}(I_n \otimes I_q) - \hat{g}\{T_i;\hat{\beta}(I_n \otimes I_q), I_q\}$, for $i = 1,\dots,n$, $k = 1,\dots,q$.The moment estimator of Σ is then given by:

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_i \hat{\varepsilon}_i^{T} . \quad (5)$$

To implement the second option, we replace the weight matrix with $W = I_n \otimes \hat{\Sigma}^{-1}$ to obtain the final estimators $\hat{\beta}(I_n \otimes \hat{\Sigma}^{-1})$, and $\hat{g}\{t_i; \hat{\beta}(I_n \otimes \hat{\Sigma}^{-1}), \hat{\Sigma}^{-1}\}$.

3.2 Profile Likelihood Method

One important statistical method for estimating the parameters of multivariate partial linear models is used to simplify the estimation process by reducing the number of parameters that need to be estimated, making it easier to handle complex models. The use of likelihood is not effective in high-dimensional cases and fails particularly in semi-parametric models. Therefore, alternative methods, such as profile likelihood, are used. The concept of this method revolves around finding the parameter estimator value that maximizes the likelihood function by keeping the values of the other parameters in the model fixed. In other words, the estimator that maximizes the likelihood function is found while keeping the other parameters constant. This method aids in estimating complex multi-parameter models, the first to propose the use of profile likelihood was R.A. Fisher in 1922 in his paper titled "On the Mathematical Foundations of Theoretical Statistics." Fisher was the first to formulate this concept as a method for estimating parameters in statistical models. This concept was further developed and generalized in Fisher's subsequent works and other statistical methodologies [11]. Fisher introduced this method as a means of estimating a single parameter in a probabilistic model when other parameters are unknown. It has been widely used in both theoretical and applied statistical inference, with common applications including multiple regression modeling, survival analysis, proportional hazards models, and other advanced statistical models. It remains a powerful tool for extracting estimators and drawing statistical inferences in these contexts, reformulates the model to take the following form [7]:

$$Y_{ik} - X_i^T \beta_k = g_k(T_i) + \epsilon_{ik}$$
$$i = 1,2,\dots,n , \qquad k = 1,2,\dots,q,$$

to derive the estimators of the parameters β and Σ in model (1), we will rely on the profile likelihood approach [12], this implies that $Y \sim N_{n,q}(XB + g(T), \Sigma \otimes I_n)$. by definition [13], we have that the likelihood function is given by:-

$$L(B, g(T)) = \prod_{i=1}^{n} f(Y \mid X, g(T)) , \quad (6)$$

$$Ln(L(B, g(T)) = \sum Ln(f(Y|X, g(T)). \quad (7)$$

We prove the function g(T) when estimating B.

$L_P(B) = \max L(B, g(T))$.

Maximizing $L_P(B)$ to estimate B.

Let $Y = XB + F + U$.

- X Parametric variables matrix (n×p);
- B matrix of unknown parameters (p×q);
- F unknow function (n×q).

$$F_k = (f_k(t_1), \dots f_k(t_n)), k = 1, \dots, q.$$

U matrix of errors (n × q), To find the estimates of B, Σ in the model, we will use the profile likelihood approach, based on the assumption that the error is independent and normally distributed.

$$U \sim N_{n,q}(0, \Sigma \otimes I_n). Y \sim N_{n,q}(XB + F, \Sigma \otimes I_n).$$

By Definition, we have that the likelihood function is given by [13].

$$L = (2\pi)^{-\frac{nq}{2}} |\Sigma|^{-\frac{nq}{2}} \mathrm{etr}\left\{-\frac{1}{2}\Sigma^{-1}(Y - XB - F)'(Y - XB - F)\right\} (8)$$

For a given B, F is estimated by a linear fit and the estimator:

$\hat{F} = S(Y - XB)$, where S (Kernel Smoothing). Substituting this into (8) and denoting by $\tilde{X} = (I_n - S)X, \tilde{Y} = (I_n - S)Y$

The log-profile likelihood function is given by:

$$L = \frac{nq}{2} ln(2\pi) + \frac{n}{2} ln |\Sigma|^{-1} - \frac{1}{2} tr\{\Sigma^{-1}(\tilde{Y} - \tilde{X}B)'(\tilde{Y} - \tilde{X}B)\} (9)$$

The profile likelihood estimators of B, F and Σ are given by:

$$\hat{B} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y},$$
$$\hat{F} = S(Y - \tilde{X}\hat{B}),$$
$$\hat{\Sigma} = \frac{1}{n}(\tilde{Y} - \tilde{X}\hat{B})'(\tilde{Y} - \tilde{X}\hat{B}).$$

Differentiating (9) with respect to B and Σ^{-1}, we obtain the following system of:

$$\frac{\partial \ln(B, \hat{F}, \Sigma; Y)}{\partial B} = \Sigma^{-1}\tilde{X}'(\tilde{Y} - \tilde{X}\hat{B}) = 0$$
$$\frac{\partial \ln(B, \hat{F}, \Sigma; Y)}{\partial \Sigma^{-1}} = \frac{n}{2}(2\Sigma - 2\mathrm{diag}\,\Sigma) - \left((\tilde{Y} - \tilde{X}B)'(\tilde{Y} - \tilde{X}B) - diag(\tilde{Y} - \tilde{X}B)'(\tilde{Y} - \tilde{X}B)\right) = 0$$

By solving the above system of equations, we obtain the estimate B, Σ.

4 PRACTICAL APPLICATION

In this section, we conduct an applied study on the key factors and variables affecting well water quality through the application of semi-parametric estimation methods to determine whether the water is suitable for human consumption, agriculture, or industrial use. The study is based on a sample of 36 wells from two geographical locations within the city of Baghdad. The data were collected from the Planning Department of the General Authority for Groundwater, Ministry of Water Resources, Iraq. We will deal with the calculation of the parametric component and the nonparametric component of the model, where the parametric component includes the non-zero significant parameters of the parameter vector β, with dimensions $(p \times q)$, and the nonparametric component includes the nonparametric function.

The proposed model for analyzing well water quality is based on the most important factors or variables that will be used as follows:

$$Y_{i1} = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + g_1(T_i) + \varepsilon_{i1}, \quad (10)$$

$$Y_{i2} = \beta_4 X_{i1} + \beta_5 X_{i2} + \beta_6 X_{i3} + g_2(T_i) + \varepsilon_{i2}\,. \quad (11)$$

The variables that used in (10), (11) refers to the following:

- Y_{i1}: EC refers to the Electrical Conductivity;
- Y_{i2}: TDS refers to the Total Dissolved Solids;
- X1: Na it refers to sodium, X2: CL It refers to chloride;
- X3: SO_4 refers to sulfate. Through the real data and the adoption of the above model, the estimates for the estimation methods were obtained, as shown below:

Table 1: Estimated parameter (β) values using the profile least squares method and the profile likelihood method.

Methods	Dependent Variables	Independent Variables		
		Na	Cl	SO4
WPLS	TDS	2.87618	0.04709	0.54053
	EC	3.70270	2.55153	-3.39325
PLik	TDS	3.49767	0.26902	1.22274
	EC	4.05523	4.28007	-6.10574

Table 2: Mean Square Error (MSE).

Methods	MSE
WPLS	0.72997
PLik	1.10292

The results presented in Table 2 indicate that the Profile Least Squares Method achieves the lowest Mean Squared Error (MSE), making it the most optimal method for water quality analysis. Figures 1 - 4 display a curve comparing the real values with the estimated values of the response variables using both the Profile Least Squares Method and the Profile Likelihood Method.

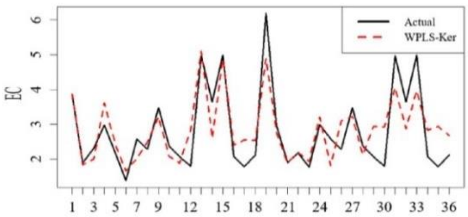

Figure 1: Real and Estimated Curves of the response variable (EC) obtained using the profile least squares method.

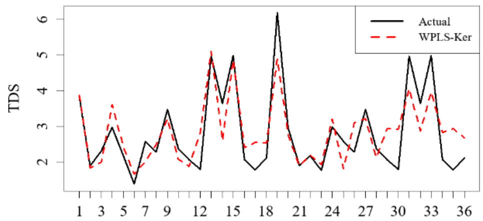

Figure 2: Real and Estimated Curves of the response variable (TDS) obtained using the profile least squares method.

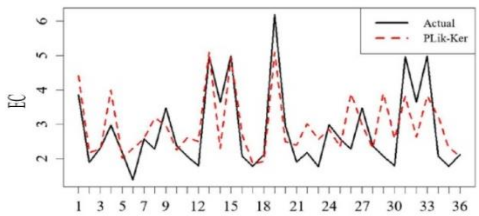

Figure 3: Real and Estimated Curves of the response variable (EC) obtained using the profile likelihood method.

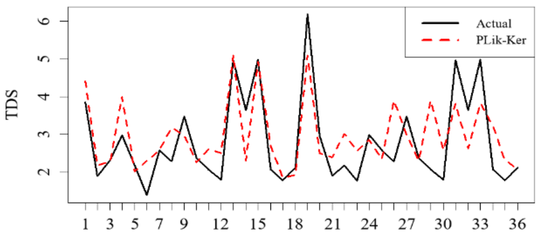

Figure 4: Real and Estimated Curves of the response variable (TDS) obtained using the profile likelihood method.

5 DISCUSSION

The findings made on the data of 36 wells in two geographical regions in the city of Baghdad indicated that the multivariate partial linear models were a good way of reflecting the relationships between the factors that influenced the quality of groundwater. The Profile Least Squares Method and the Profile Likelihood Method were adopted to estimate which one was more successful in the analysis of water quality. Table (2) shows that the Profile Least Squares Method is more superior since it produced the lowest Mean Squared Error (MSE = 0.72997) than the Profile Likelihood Method (MSE = 1.10292). This denotes the increased accuracy of the former method assessing the statistical model and water quality analysis. The actual values and the estimated values obtained with the help of Profile Least Squares Method are close to each other, which can be observed in the curves in Figures 1 - 4. This shows the precision of the estimation as well as the capacity of the model to depict the actual behavior of the response variables that are total dissolved solids (TDS) and Electrical Conductivity (EC).The proposed estimated coefficients in Table 1 show a positive relationship between the concentrations of sodium ions (Na), total dissolved solids (TDS), and Electrical Conductivity (EC). This means that as the concentration of these ions increases the quality of water becomes poor. It was also found that chloride (Cl) and sulfate (SO4) ions are also involved in the interpretation of changes in water quality to a certain degree since they are the most significant indicators of ionic pollution in groundwater. Considering the accurateness and efficiency of the Profile Least Squares Method in determining groundwater quality in Baghdad, it is also applicable to determine the water quality in other areas with similar geological and environmental features.

6 CONCLUSIONS

The results of this study showed that the performance of the Profile Least Squares (PLS) method was clearly better than the Profile Likelihood (PL) method in the estimation of the multivariate partially linear model since it showed a lower Mean Squared Error (MSE) values. This validates the power and suitability of the PLS method to analyze groundwater quality data because of its superior capacity in representing linear and non-linear relationships with higher accuracy. The results also demonstrated considerable environmental associations where increases in sodium, nitrate and calcium concentrations were associated with an increase in dissolved solids and ions, which reflects the direct effect of these variables on the declining quality of water. With regard to future research, efforts should be made to develop more advanced and efficient estimation techniques for semi-parametric models, especially for different environmental conditions. Expansion of the application of these models to other areas (within Iraq) or to other nations with similar environmental conditions would improve the generalizability of these models and validate their use in other contexts. Integrating traditional statistical models with machine learning approaches is also a promising direction to improve the accuracy of the prediction and the flexibility in the analysis. On the practical level, the results emphasize the opportunity to enhance water resource management in Baghdad using strategies to mitigate the sources of contamination. The proposed models can provide a good analytical tool for the periodic evaluation and continuous monitoring of water quality in many areas. Moreover, developing long-term monitoring initiatives and using sophisticated analytical systems would enhance the capacity to predict and make decisions. Increased collaboration between academic institutions and environmental authorities would further contribute to practical implementation as well as continued development of these models to achieve sustainable water resource management.

REFERENCES

[1] W. Häardie and M. Müller, "Multivariate and Semiparametric Kernel Regression," pp. 357-391, 2000, [Online]. Available: https://doi.org/10.1002/9781118150658.ch12.

[2] L. D. Brown, M. Levine, and L. Wang, "A semiparametric multivariate partially linear model: A difference approach," J. Stat. Plan. Inference, vol. 178, pp. 99-111, 2016, [Online]. Available: https://doi.org/10.1016/j.jspi.2016.06.005.

[3] M. Levine, "Robust functional estimation in the multivariate partial linear model," Ann. Inst. Stat. Math., vol. 71, no. 4, pp. 743-770, 2019, [Online]. Available: https://doi.org/10.1007/s10463-018-0661-1.

[4] J. Zhang, B. Lin, and Y. Zhou, "Kernel density estimation for partial linear multivariate responses models," J. Multivar. Anal., vol. 185, p. 104768, 2021, [Online]. Available: https://doi.org/10.1016/j.jmva.2021.104768.

[5] S. Meng and Z. Huang, "Variable Selection in Semi-Functional Partially Linear Regression Models with Time Series Data," Mathematics, vol. 12, no. 17, 2024, [Online]. Available: https://doi.org/10.3390/math12172778.

[6] Q. Wu, X. Tong, and X. Zhao, "Deep partially linear cox model for current status data," Biometrics, vol. 80, no. 2, pp. 1-12, 2024, [Online]. Available: https://doi.org/10.1093/biomtc/ujae024.

[7] M. Przystalski, "Estimation of the covariance matrix in multivariate partially linear models," J. Multivar. Anal., vol. 123, pp. 380-385, 2014, [Online]. Available: https://doi.org/10.1016/j.jmva.2013.09.005.

[8] B. Pateiro-López and W. González-Manteiga, "Multivariate partially linear models," Stat. Probab. Lett., vol. 76, no. 14, pp. 1543-1549, 2006, [Online]. Available: https://doi.org/10.1016/j.spl.2006.03.016.

[9] Z. Li, L. Su, and D. Zhang, "Profile least squares estimation of a partially linear time trend model with weakly dependent data," Econ. Lett., vol. 125, no. 3, pp. 404-407, 2014, [Online]. Available: https://doi.org/10.1016/j.econlet.2014.10.030.

[10] Y. Zhang and L. Zhu, "Semiparametric estimation of multivariate partially linear models," J. Stat. Comput. Simul., vol. 87, no. 11, pp. 2115-2127, 2017, [Online]. Available: https://doi.org/10.1080/00949655.2017.1318135.

[11] S. A. Murphy and A. W. van der Vaart, "On Profile Likelihood: Rejoinder," J. Am. Stat. Assoc., vol. 95, no. 450, p. 482, 2000, [Online]. Available: https://doi.org/10.2307/2669392.

[12] J. Fan and T. Huang, "Profile likelihood inferences on semiparametric varying-coefficient partially linear models," Bernoulli, vol. 11, no. 6, pp. 1031-1057, 2005, [Online]. Available: https://doi.org/10.3150/bj/1137421639.

[13] N. H. Timm, Applied Multivariate Analysis, Springer Texts in Statistics. New York, NY: Springer, 2002. doi: 10.1007/b98963.