DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

A STUDY OF SEMIPARAMETRIC ADDITIVE MODEL USING BACKFITTING ALGORITHM.

E. A. Siloko[1], and *I. U. Siloko[2]
[1]Department of Statistics, University of Benin, Nigeria.
[2]Department of Mathematics, Edo State University Uzairue, Nigeria.

Email: siloko.israel@edouniversity.edu.ng, sisraelgreat@gmail.com

Abstract: **Semiparametric density estimation is of wide applications with numerous novel approaches in its estimation. The curse of dimensionality effect which is usually associated with nonparametric density estimation is often addressed with the semiparametric density estimation. The curse of dimensionality effect in nonparametric estimation is due to the addition of more explanatory variables to a model which ultimately leads to slow convergence rate of the model. As the explanatory variables increases, the nonparametric approach in data estimation becomes difficult and hence the need for the semiparametric approach. In semiparametric estimation, the variables are considered independently in terms of their relation with the response variable through the additive techniques. This paper considers the additive model by employing the Backfitting algorithm and the kernel smoother. The Backfitting algorithm is apply on real data using the Adjusted R-Squared as the measure of performance and the results revealed dominance of the semiparametric approach over the nonparametric method. Again, the model addresses the curse of dimensionality effects that is often associated with is nonparametric counterpart.**

Keywords: *Additive Model; Backfitting Algorithm; Kernel; Nonparametric; Semiparametric*

1. INTRODUCTION

The concept of probability density estimation is often used in mathematical statistics, probability theory and other statistical related fields. Density estimation usually involves the construction of probability estimates from a data set. Probability density model are employed in description of the joint distribution of data set or explanation of the conditional distribution of a given set of observations when the values of the other observations are known [1–3]. Probability density estimates are of wide applications in terms of exploratory analysis and data visualization. Probability density estimation is very important because it helps the analyst in getting information and having a better knowledge about the distribution of

the underlying observations. Density estimation process comprises of estimating the functional form of the density from a given set of data. Density estimation involves approximating a hypothesized probability density from the observations. Other known areas were probability density estimates and derivatives of probability density estimates can be applied includes discriminant and cluster analysis; hazard rate estimation; conditional densities and score function estimation; testing for unimodality and independence [4–9].

Probability density estimation uses either the parametric model or nonparametric model in its analysis of data with the semiparametric model as the hybrid of both models. The parametric model is based on the assumption that the data sets belong to a known family of distribution with a finite number of unknown parameters to be estimated. One of the required conditions for the successful implementation of the parametric model is a small sample size. The estimation of the parameters in a parametric model demands that an accurate selection of the parametric model is made but this tends to be difficult in practice. In a parametric model, the parametric estimate converges at a $n^{-1/2}$ rate and there is no discrepancy between the true model and the fitted model. Parametric models usually have fixed structures and one of the popular parametric estimators is the maximum likelihood estimators with many applications [10].

The nonparametric model does not make assumption about the family of the distribution of the data but rather they are subjected to critical statistical examination often referred to as allowing the data to "speak for themselves" [11,12]. Nonparametric model gives a better approach to statistical analysis of data due to its ability to capture the true structure of the underlying distribution. One of the features of nonparametric models is that they produce good estimates from a very large amount of data and are useful as exploratory and visualization tools in data analysis [13,14]. On convergence rate, the nonparametric models converge at slower rate unlike their parametric counterpart that converges at the rate of $n^{-1/2}$ [15]. Nonparametric density estimators are very flexible because they do not rely on distributional assumptions and their flexibility resulted in high computational cost that has limited their wide spread uses over the years until recently. In nonparametric density estimation, it will be assumed that the distribution has a probability density function but the data will be given the opportunity to speak for themselves in the determination of the estimate of the function. Nonparametric models are of wide applications and are considered as alternative to the parametric density estimators that involves the specification of parameters which can be estimated via the likelihood principle [16, 17]. Nonparametric density estimation uses approximations which is regulated by the smoothing parameter whose choice is very important in the performance and implementation process and the techniques for its selection are usually data based [18, 19]. Due to its simplicity and wide uses, nonparametric methods have been applied in different areas of life, especially with the availability of fast computing machines. Nonparametric estimators are of great importance because they form the building blocks of different semiparametric estimators in real application.

Nonparametric models are mainly of direct application but sometimes they may be of interest only as input to another estimation problem. This other estimation problem when described by a finite number of dimensional parameters is called a semiparametric model [15]. Semiparametric models function as a bridge between parametric and nonparametric models. Semiparametric model comprises of two parts, parametric and nonparametric parts and they have the advantage over the nonparametric models for retention of finite parameterization property. A semiparametric model needs more data than the parametric model but it requires less data than its nonparametric counterpart [20]. Semiparametric model as its name implies is a hybrid of the parametric and nonparametric models use in constructing, fitting and validating statistical models. As a hybrid of both models, semiparametric estimators possess the merits and demerits of both models [21]. The estimation of a semiparametric model requires that a nonparametric estimation of some functions must be first estimated. In semiparametric models, some features of the underlying distribution of the observations are unknown while others will assume a known parametric approach [22].

In this paper, semiparametric density estimation will be considered with emphasis on the additive separable model using the Backfitting algorithm due to its simplicity, flexibility and its ability of presenting the estimates of each of the independent variables. The organisation of the rest part of this paper is as follows. Section 2 is the problem statement with a subsection that is dedicated to the semiparametric estimator. Section 3 provides a concise methodology of the study and the additive separable model with the Backfitting algorithm. Section 4 is the empirical investigation with real data of the superiority of the semiparametric model over the nonparametric model. Section 5 concludes of the paper.

2. PROBLEM FORMULATION.

The semiparametric and nonparametric methods are the two common approaches in exploratory data analysis and data fitting. Nonetheless, for some large amounts of parameters, the convergence rate of the nonparametric method becomes slow. Consequently, as the dimension of the data set increases, there tends to be difficulty in obtaining accurate results from nonparametric estimation. Hence; the need to investigate the efficacy of semiparametric estimation in overcoming the problems of nonparametric method and in effect also providing an exceptional fit. The semiparametric additive separable model is the focus of this study because the nonparametric component has only one-dimensional convergence rate which makes it more accurate than the estimation of multi-dimensional functions. Again, in real life situations, the separability ideology of the independent variables in the additive model is in consonant with the idea of devolution of decision-making process in large organizations or stages of production in industries.

2.1. Semiparametric Density Estimation.

Semiparametric models bridge the gap between parametric and nonparametric models in density estimation and they achieve a faster convergence rate for conditional mean functions estimation [22]. The problem of curse of dimensionality effect associated with nonparametric model is solved with semiparametric estimators. The problem of curse of dimensionality effect associated with the nonparametric density estimators and the conditional mean function estimators resulted in difficulties of these methods in practical applications with increase in sample sizes and dimensions. Semiparametric estimators have a faster convergence rate with conditional mean functions and other parameters of interest. The estimates of semiparametric models are more efficient than parametric and nonparametric estimates and they also achieve faster rates of convergence than both models.

DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

Semiparametric models have several applications in mathematical finance where they address the two problems of symmetry and thin tails that often occur in finance data [23]. A novel semiparametric model was developed by [24] with a nonparametric component in solving system of equations with the advantage that the regressors have independent effect on the response variable. The model was applied in solving the two-equation model structure obtained from the popular labour and returns to schooling observations. A semiparametric model that studied the relationship between semen qualities and other vital factors such as age and body mass index was investigated by [25] while [26] formulated a semiparametric model in estimating equations by generalising the linear latent variable models. As reported in their study, [27] used semiparametric discrete choice models in the analysis of panel data while [28] introduced a stochastic expansion of semiparametric models for the means of weighted residuals with the results having uniform rate of convergence. [29] proposed a semiparametric model known as the semiparametric negative binomial count data that uses the likelihood principle and the product kernels estimator approach. [30] proposed a semiparametric method in determining whether there is a statistical relationship between calorie intake and income level of different households in China with focus on panel data and cross-sectional data.

Semiparametric estimation is of great significance due to their flexibility and wide applicability [31, 32]. The flexibility of semiparametric methods has helped in displaying crucial features that parametric estimation may have hidden and do display the relationship that exists among the parameters of interest. Semiparametric estimations are mainly based on fewer parametric assumptions and restrictions to enhance its efficiency in terms of applicability and with high knowledge of nonparametric estimation in its implementation [33]. Semiparametric methods are of great importance especially when there are limited response variables in the model such as binary response model. Other fields with applications of semiparametric methods are econometrics, biomedical studies, agriculture, physics, financial mathematics, environmental sciences, biological sciences and epidemiological studies [34, 35].

3. METHODOLOGY

The efficiency and effectiveness of the additive model is demonstrated in addressing the curse of dimensionality effect that is often connected with nonparametric model. The curse of dimensionality effect is the complexity associated with nonparametric estimation as the number of variables increase. The semiparametric additive model is being used because the nonparametric component has one-dimensional convergence rate that makes them easier than estimating a multi-dimensional function, and which also help to circumvent the curse of dimensionality effects since the variables are usually estimated separately. The ideology of separating the independent variables in the additive model supports the devolution of decision-making process in organizations or stages of production in industries in real life situation.

The Backfitting algorithm is employed in the semiparametric additive model while the Gaussian kernel smoother is applied in the nonparametric component of the model. The effect of the independent variables on the response variable will be observed as the number of the explanatory variables increases. The adjusted R-square criterion will be used as the measure of performance with the introduction of extra explanatory variables to the model.

3.1. The Additive Model in Semiparametric Density Estimation

The additive model shows the relationship that exists among the individual independent variables and the response variable in smooth forms. This modelling method can be used to capture the underlying nature of the data through smooth functions base on the predictors' functions' shape that are only determined by the data themselves. The additive model is easy to interpret because of the presence of the smoothness factor that can be adjusted especially when the independent variables are nonlinear. This suggests that wiggly and noisy predictor function can be avoided with the use of the additive model due to the effect of the smoothing parameter. The additive model also known as separable model is a generalization of the multiple linear regression models and it introduces the one-dimensional nonparametric function that replaces the linear components of the model. The additive model usually restricts the function to be additively separable in the regressors. The model is of the form

$$Y = \alpha + g_1(X_1) + g_2(X_2) + g_3(X_3) + \cdots + g_k(X_k) + \varepsilon, \quad (1)$$

where α is the intercept, $g_i(X_i)$, $i = 1, 2, \ldots, k$, are unknown and are usually estimated nonparametrically and ε is the error term. The introduction of the error term is because there is no perfect fit and the error term is assumed to belong to a family of distributions with finite dimension [20]. An alternative presentation of the additive model is

$$Y = \alpha + \sum_{i=1}^{k} g_i(X_i) + \varepsilon, \qquad (2)$$

where $g_i(\cdot)$ are the smooth nonparametric functions. The function $g_i(X_i)$ shows the actual effect or contribution of the particular component of X_i on the dependent variable Y. One of the advantages of the additive separable estimators and the conditional mean function $E(Y|X)$ is that they can converge at univariate rate. Another advantage of the model is the simplicity of its graphical interpretation of results [33]. In the case of the additive model, the conditional expectation of Y given $X = (X_1, X_2, \dots, X_k)^T$ is the sum of the unknown functions of the independent variables and the intercept α such that $E(\varepsilon|X) = 0$. The conditional expectation of Y given X is of the form

$$E(Y|X) = \alpha + \sum_{i=1}^{k} g_i(X_i), \qquad (3)$$

where $g_i(X_i)$, $i = 1, 2 \dots, k$ represent smooth and nonparametric functions and with the restriction that $E\{g_i(X_i)\} = 0$. The assumption of separability in this model is not too restrictive as presented because there could be interaction amongst the regressors [22]. The intercept α is obtain from the relation

$$\alpha = E(Y) = \frac{1}{n}\sum_{i=1}^{n} Y_i. \qquad (4)$$

Semiparametric estimation requires smoothing parameter for its proper implementation because the smoothness of the estimates is dependent on the smoothing parameter and the smoother function. Generally, smoothers are referred to as the cornerstones of semiparametric models particularly the additive models due to their significant role in the estimation process. The smoothing parameter plays a vital role of balancing the bias and variance trade-off; smoother estimates tend to have more bias and less variance. Estimates with less variance due to large value of the smoothing parameter usually produce better results and validates out-of-sample tests. However, with larger values of the smoothing parameter, the estimate might be too smooth and important features of patterns present in the model may not be seen. Hence, smoothing parameter should be chosen to maintain a balance between the bias and variance and also considering the features in the observations [19].

Early work on smoothing parameter selection in semiparametric additive model could be traced to [36] rule of thumb and the plug-in algorithm introduced by [37]. In the work of [38], they proposed an algorithm for fitting the additive models known as the alternating direction method of multipliers while [32] suggested a generalized additive Markov switching process that combines the Poisson distribution and the generalized Pareto distribution which models the characteristics of random sums over time. Several algorithms for parameter selection in additive models were proposed by researchers recently and novel approaches are been suggested [39–41].

3.2 The Backfitting Algorithm.

The Backfitting algorithm is mainly used in additive models for fitting density estimates. In the application of this algorithm, a continuous smoothing function such as Spline, Loess or Kernel functions are required to estimate the nonparametric component of the model. The Backfitting method estimates the components of the additive model iteratively. The Backfitting algorithm is a reliable approach in additive model in terms of obtaining the estimates of the model [33, 42]. The Backfitting algorithm for two predictor additive models is

$$Y_i = \alpha + g_1(X_{i1}) + g_2(X_{i2}) + \varepsilon_i$$

The estimation process is as follows

1. The process will start by expressing the variables in mean deviation pattern to ensure that the partial regressions sum to zero. This is to ensure that the individual intercepts are eliminated.
2. The next step is to take the preliminary estimates of each function from the least squares regression of Y on the $X's$.
   $$y_i - \bar{y} = b_1(x_{i1} - \bar{x}_1) + b_2(x_{i2} - \bar{x}_2) + \varepsilon_i$$
   $$y_i^* = b_1 x_{i1}^* + b_2 x_{i2}^* + \varepsilon_i$$
3. The estimates in step two are then used as the first step in the iterative estimation process.
   $$\hat{g}_1^{(0)} = b_1 x_{i1}^*$$
   $$\hat{g}_2^{(0)} = b_2 x_{i2}^*$$
4. Find the partial residuals for X_1 that displaces Y from its linear relationship with X_2 but keeps the relationship between Y and X_1. The partial residuals for X_1 are

DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

$$y_i^* - b_2 x_{i2}^* = b_1 x_{i1}^* + \varepsilon_i$$

5. The process is repeated in step 4 for X_2.
6. The final step is to smooth the partial residuals against their respective $X's$ to provide a new estimate of g

$$\hat{g}_k^{(1)} = \text{Smooth}\left[Y_{(k)i}^{(1)} \text{on } x_{ik}\right]$$

$$= S_k\left\{Y_i - [\hat{g}_1^{(1)}(x_{i1}) + \hat{g}_2^{(1)}(x_{i2})]\right\},$$

where S is the smoother function.

The Backfitting algorithm is very sensitive to the choice of smoothing parameter in its estimation and hence the smoothing parameter should be appropriately select. The Backfitting algorithm for several predictors with the additive models is summarized into three steps as follows:

STEP1. Compute the initial point of the iteration denoted by $\hat{\alpha}$ and $\hat{g}_i^{(0)} = 0$ for all $i = 1, 2, \dots, k$, where $\hat{\alpha} = \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

STEP2. Cycle for $i = 1, \dots, k$. Set

$$r_i = Y - \hat{\alpha} - \sum_{j=1}^{i-1} \hat{g}_i^{(l+1)} - \sum_{j=i+1}^{k} \hat{g}_k^{(l)}.$$

$$\hat{g}_i^{(l+1)}(\cdot) = S_i(r_i).$$

STEP3. Repeat step (2) until convergence is reached, that is, there is no difference between the new estimates and the previous estimates.

The Backfitting algorithm requires a smoother and a smoothing parameter that relies on the component of X_i. Different smoothing parameters can be used for the different variables. The algorithm may require much iteration before it converges but it is not difficult in implementation. In implementing the Backfitting algorithm, we will use the kernel smoother and the choice of the smoothing parameter is the generalized normal reference rule with respect to the normal kernel. The backfitting algorithm and the additive model are of great significance in data analysis particularly in semiparametric estimation. Due to the wide applicability of semiparametric estimation in data analysis, its techniques have gained attention in statistics and other related fields of studies by several researchers recently [43–49].

3.3 The Measure of Performance.

The performance measure that will be employed in this paper is the Adjusted R-squared. The Adjusted R-squared is a model performance evaluator that tells how the explanatory variables are able to explain the variation in the response variable. The multiple R-squared is an estimate of the proportion of the variance in the data explained by the regression and is given by

$$R^2 = 1 - \frac{\sum \hat{\varepsilon}_i^2/n}{\sum(y_i - \bar{y})^2/n} \qquad (5)$$

where $\bar{y}$ is the mean of y_i, $\hat{\varepsilon}_i^2$ is the sum of squares of the error and n is the sample size. The fraction in this expression is basically an estimate of the proportion of variance not explained by the regression. The problem with R^2 is that it always increases when a new predictor is added to the model and this is because the variance estimates used in calculating R^2 are biased and that tends to inflates its value [50]. If unbiased estimators are used, then we will have the adjusted R^2 which satisfies the inequality $0 \le R^2 \le 1$ and is of the form

$$R_{\text{adj}}^2 = 1 - \frac{\sum \hat{\varepsilon}_i^2/(n-p)}{\sum(y_i - \bar{y})^2/(n-1)}, \qquad (6)$$

where p is the number of variables to be estimated and $n - p$ is the degrees of freedom. A high value of R_{adj}^2 is an indication that the model is doing well in terms of explaining the variability in the response variable.

3.4. The Kernel Density Estimator.

The implementation of the Backfitting algorithm requires a smoother function with a smoothing parameter. The kernel smoother is one of the nonparametric estimators that have being widely used in many fields of learning. The kernel estimator comprises of the kernel function and the smoothing parameter also known as the bandwidth. Amongst the classes of nonparametric estimators, the kernel estimator is the most studied estimator and frequently used estimator in nonparametric estimation [4, 13, 51, 52]. The univariate kernel density estimator with the smoothing parameter is given by

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \qquad (7)$$

where h is the bandwidth also called the smoothing parameter; K a kernel function; n is the sample size and X_i are observations obtained from real life occurrences with an unknown probability function.

The kernel function is a probability density function which must satisfy the following conditions

$$\begin{cases} \int K(\mathrm{x})d\mathrm{x} = 1, \\ \int \mathrm{x}K(\mathrm{x})d\mathrm{x} = 0 \quad \text{and} \\ \int \mathrm{x}^2 K(\mathrm{x})d\mathrm{x} = k_2(K) \neq 0. \end{cases} \tag{8}$$

In determining the smoothing parameter in kernel estimation, its selection procedure should depend on the data for a better use of kernel estimator. The kernel density estimator is primarily for data analysis either for exploratory or visualisation purposes due to the importance of both in many fields of studies [53, 54]. The main setback of the application of the kernel estimator is the complexity associated with the selection of the smoothing parameter especially with increase in dimension of the variables. The smoothing parameter determines the smoothness of the estimates and the performance of the kernel estimator. Novel smoothing parameter selectors have been proposed by researchers in providing solution to the problem of smoothing parameter selection in kernel estimation [19, 55–58].

The performance of the kernel estimator Equation (7) can be measured by several error criteria functions but the most applied criterion is the asymptotic mean integrated squared error (AMISE). The popularity of the AMISE in kernel estimation is attributed to its inclusion of dimension in its formula and this unique characteristic with potential benefits is absent in other error criteria. The AMISE is an error criterion with two components and these two components depend on the smoothing parameter for their computation and measure of performance [12–14]. The components of the AMISE are the asymptotic integrated squared bias and asymptotic integrated variance given as

$$AMISE\left(\hat{f}(\mathrm{x})\right) = E\int\left(\hat{f}(\mathrm{x}) - f(\mathrm{x})\right)^2 d\mathrm{x}$$
$$= \int Bias^2\left(\hat{f}(\mathrm{x})\right)d\mathrm{x} + \int Var\left(\hat{f}(\mathrm{x})\right)d\mathrm{x}. \tag{9}$$

The asymptotic mean integrated squared error of the univariate kernel density function when approximated using Taylors' expansion is given as

$$AMISE = \frac{R(K)}{nh} + \frac{1}{4}h^4\mu_2(K)^2 R(f''). \tag{10}$$

The smoothing parameter that minimizes the AMISE of the univariate kernel is of the form

$$h_{\text{AMISE}} = \left[\frac{dR(K)^d}{\mu_2(\mathrm{K})^2 R(f'')}\right]^{\left(\frac{1}{4+d}\right)} \times n^{-1/(4+d)}, \tag{11}$$

where d is the dimension of the kernel, $R(K)$ is the roughness of the kernel, $\mu_2(K)^2$ is the moment or variance of the kernel, n is the sample size and $R(f'')$ is the roughness of the unknown density function.

The additive model requires smoothing function and smoothing parameter. The choice of the Gaussian kernel is due to the fact that it produces smooth density estimates and simplifies the mathematical computational process. The smoothing parameter that minimizes the AMISE of the univariate kernel using the Gaussian kernel is

$$h_{\text{AMISE}} = \hat{\sigma}\left(\frac{4}{3}\right)^{1/5} \times n^{-1/5} = 1.06\hat{\sigma}n^{-1/5}, \tag{12}$$

where $\hat{\sigma}$ is the standard deviation of the observations. The bandwidth of Equation (12) is known as the normal reference rule and is mainly for data that are normally distributed and unimodality [12, 13].

4. RESULTS AND DISCUSSION.

This section focuses on the numerical and graphical results of the additive model using the Backfitting algorithm on the employee attitude survey data [59]. The data were obtained from a survey of the clerical employees of a large financial organization. The data were aggregated from the questionnaires of approximately thirty-five employees for each of 30 randomly selected Departments with one supervisor. The questions presented were about the employees' satisfaction with their supervisors. The basic reason for this survey was to measure the overall performance of a supervisor and questions that are related to specific attitude of the supervisor. The satisfaction of the employee in any organization should be paramount to the employer owing to the fact that no employee can perform optimally in uncomfortable or toxic environment. The desire of any employer of labor in an organization is the optimal performance of the employee with minimal supervision. Hence; the attitude of the supervisor towards the supervisee is a major determinant of their performance.

The analysis tends to explore and explain the relationships that exist between specific supervisor characteristic and overall satisfaction with supervisors as felt by the employees. The numbers give the percentage proportion of favorable responses to seven questions in each Department. The seven questions or

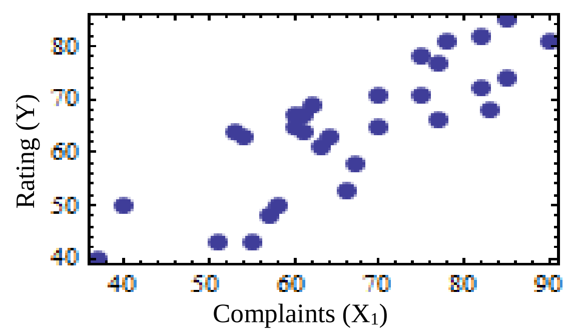
DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

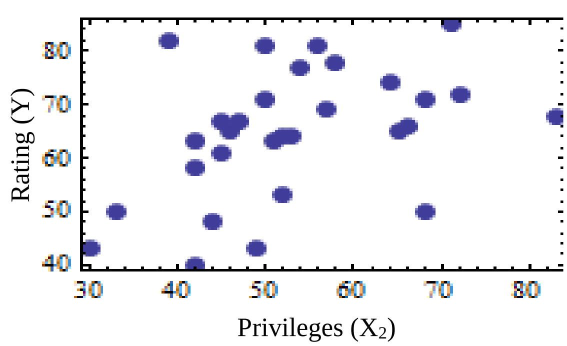
variables are Y (overall rating of the job done by Supervisor), X_1 (handle employee complaints), X_2 (does not allow special privileges), X_3 (opportunity to learn new things), X_4 (raises based on performance), X_5 (too critical to poor performance), X_6 (rate of advancing to better job). These seven questions can be briefly referred to as ratings, complaints, privileges, learning, raises, critical, and advance. The aim is primarily to understand the effect of other variables $X_1, X_2, \ldots, X_6$ on Rating (Y).

The analysis of the employee attitude survey data is primarily to investigate the impact of some variables such as complaints, privileges and learning on the percentage proportion of questions base on ratings using the semiparametric additive model. The graphical analysis of two variables oftentimes begins by considering the scatterplots of the variables in order to ascertain the existence of relationship amongst the variables. If there is existence of relationship between the two variables, such relationship is usually depicted by the data point in the scatterplot which provides a direction for further statistical analysis [60]. Although scatterplots oftentimes do not reveal hidden structures of the data due to cloudiness of the density of the data but they promote the vital function of displaying the nature of the relationship that exist amongst the variables.

The scatterplots in Fig. 1 show a positive relationship between the response variable Y and the various independent variables which are complaints, privileges and learning. The scatterplot in Fig. 1 is that of each independent variable against the response variable Y and from the plots we can observe that the effects of the studied independent variables on the response variable Y are linear. The linearity displayed by the response variable and the respective independent variables is an indication that the supervisor is properly rated by the employee. The scatterplot of rating and complaints will produce a linear graph with the line of best fit and this simply means that the supervisor handles and attends to the complaints of the employee regardless of the nature and times of the complaints. Handling of employees' complaints could improve the overall performance of employee by increasing the productivity of an organization. Again, important complaints reported by employee when not properly handle by the supervisor could reduce productivity especially for specific machines that improvising for their functions is practically impossible. Addressing the complaints of employee by the supervisor promptly will enhance the performance of the employee and improve the output of the organization; hence the positive correlation as depicted by the scatterplot of rating and complaints.

On rating and special privileges, the cloudiness of the data points shows that there exists a relationship between rating and privileges but not highly positively correlated. The scatterplot clearly indicated that the supervisor was reluctant in granting certain special privileges of the employees which is the usual behaviour of most employers in some organizations. Embargo are usually placed on certain privileges that can motivate employees in discharging their duties effectively. On the relationship of rating with learning, the scatterplot vividly indicates that the variables are highly positively correlated. The linearity of the variables means that the supervisor granted the employees the opportunity to learn new techniques that are capable of improving the productivity of the organization. Training of employees is an important role of any organization due to the fact that modern technologies and innovations are evolving in the industry. Financial organizations usually engage their employees on routine training because of the modern trends of financial transactions. Training of employees will definitely improve their productivity because it exposes the employees to new techniques which automatically affect the performances of the employees. On learning which is synonymous with training, the employees were given the opportunity to be trained on modern technologies of financial transaction.





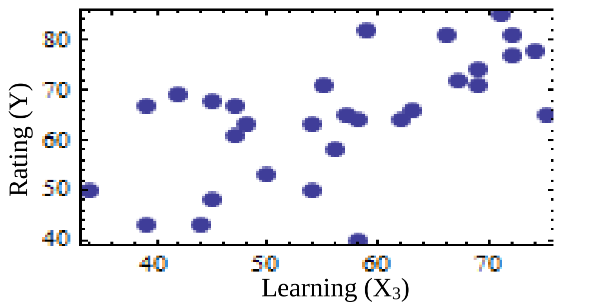


Fig. 1: *Plots of the explanatory variables against the response variable*

The results from the analysis of the observations are hereby presented using the estimates of the semiparametric and nonparametric regression and the tabulated values of the Adjusted R-squared. The analysis was carried out using the R software. The Adjusted R-squared is the performance measure for the semiparametric and nonparametric regressions. The Backfitting algorithm is applied in the semiparametric additive model.

The first variable to be considered with respect to rating is COMPLAINTS. The scatterplot of rating and complaints is linear implying that all complaints presented by the employees regarding the organization were considered. The graph of the impact of complaints on rating with the semiparametric model is presented in Fig. 2 with the graph showing a straight line which also supports the scatterplot of rating and complaints. Fig. 3 is the impact of complaints on ratings with the nonparametric component and from the graph, it could be noticed that though it is a straight line but not a perfect straight indicating that with increase in the number of variables to be estimated, the semiparametric estimation approach will demonstrate superiority over the nonparametric method in terms of performance due to cause of dimensionality effect.

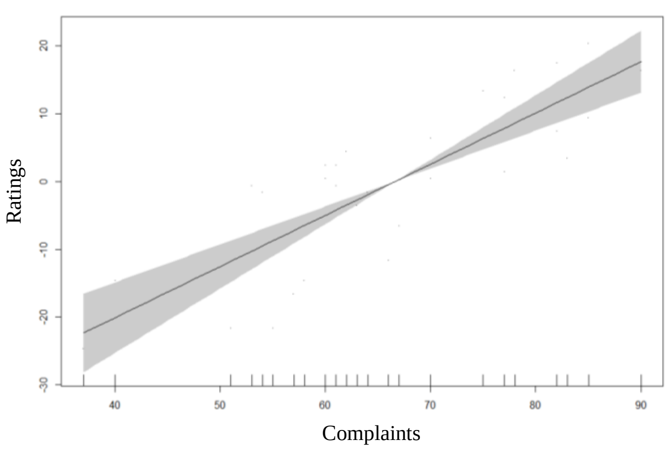


Fig. 2: *Impact of complaints on ratings with semiparametric fit.*

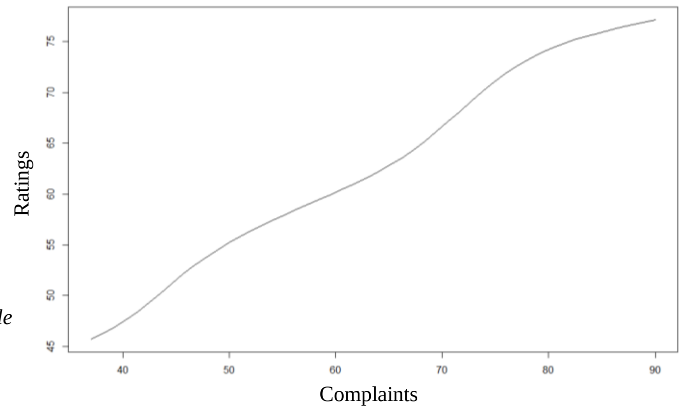


Fig. 3: *Impact of complaints on ratings with nonparametric fit.*

In Table 1, it will be observed that the adjusted R^2 of the semiparametric estimate and the nonparametric component have the same value. The adjusted R^2 is quite high with a value of 0.67 and this implies that complaints explained up to 67% of the variation in ratings. At this level, the curse of dimensionality effect has not set in because it is only complaints on rating that is being considered, hence the same adjusted R^2 value for the semiparametric estimate and nonparametric component. Fig. 3 presents the impact of complaints on rating with the nonparametric method and as can be seen from the graph, the graph is almost a straight line and it is an indication that the model offers a good approximation. The adjusted R-square value for the semiparametric model and nonparametric component are presented in Table 1.

TABLE 1: MODEL ADJUSTED R-SQUARE VALUES

Model	Semiparametric	Nonparametric
Adjusted R-Square	0.67	0.67

In order to observe the curse of dimensionality effects as induced by the addition of extra variables, another variable is introduced into the model and this variable is PRIVILEGES. The graph of the impact of complaints and privileges on rating with semiparametric model is in Fig. 4 while Fig. 5 is the impact of complaints and privileges on ratings with the nonparametric component of the model. In Fig. 4, the first graph shows that complaints have good impact on rating which also support the linearity displayed by its scatterplot while the second graph is a curve indicating that certain privileges were not granted to the employees by the supervisor and hence it is not a good fit as presented by the graph of privileges on rating. Semiparametric method considers the individual effect of the explanatory variables (complaints and privileges) on ratings but the nonparametric component considered the joint

DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

effect of both variables on ratings. Again, in Fig. 5, we can see the linearity depicted by complaints and privileges on rating with the nonparametric estimation. However, there is no interaction between complaints and privileges, that is, the slope of complaints is the same at every value of privileges. The adjusted R-square value for the semiparametric model and the nonparametric estimate are presented in Table 2 and it is observed that the adjusted R^2 value of the semiparametric model is better than that of the nonparametric estimate. The adjusted R^2 of the semiparametric model is 0.68 which implies that it could explain 68% of the model, that is, complaints and privileges explained 68% of the variations on ratings. In the case of the nonparametric estimation, the explanation of complaints and privileges on ratings is only 66%. Higher values of adjusted R^2 is an indication of better approximation or estimation. Hence; the estimate of the semiparametric model outperformed the nonparametric estimate and this could be attributed to increase in the number of the explanatory variables.

TABLE 2: MODEL ADJUSTED R-SQUARE VALUES.

Model	Semiparametric	Nonparametric
Adjusted R-Square	0.68	0.66

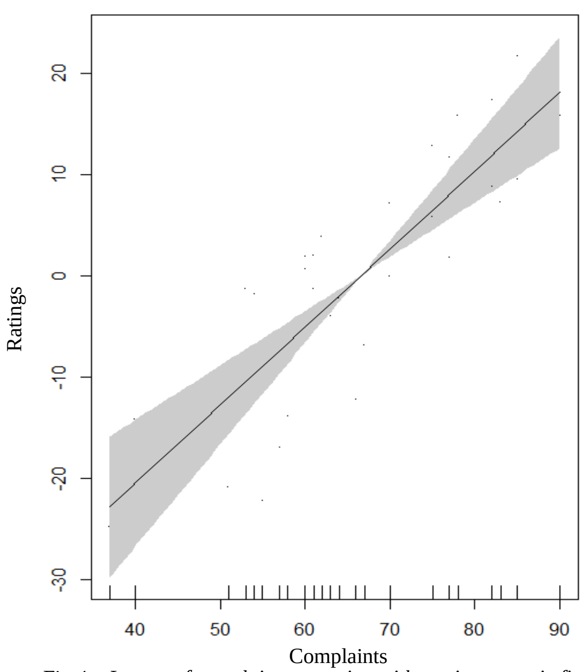


Fig.4a: *Impact of complaints on rating with semiparametric fit.*

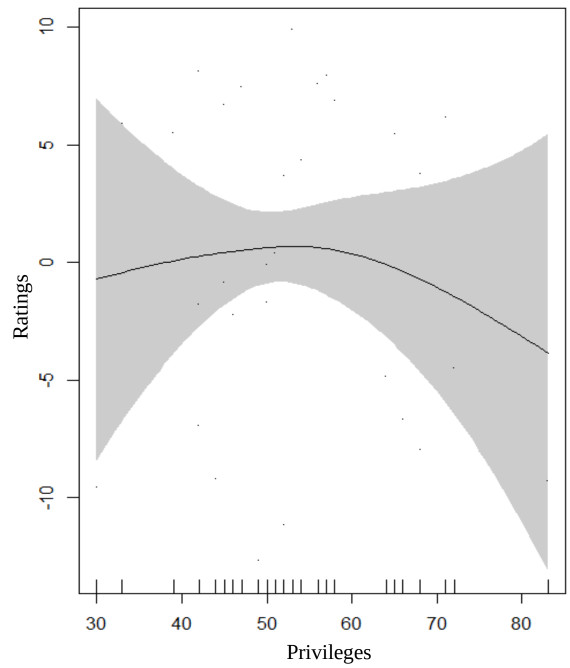


Fig. 4b: *Impact of privileges on rating with semiparametric fit.*

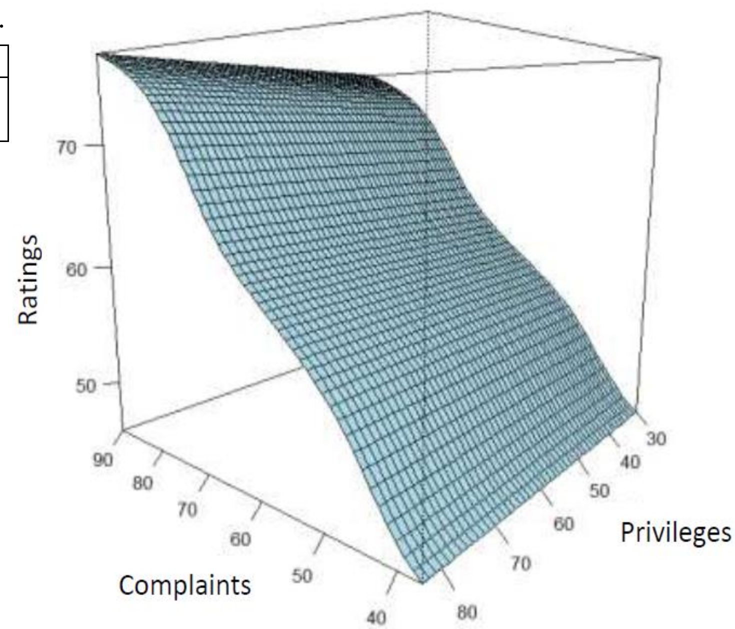


Fig. 5: *Impact of complaints and privileges on ratings with nonparametric fit.*

A third explanatory variable which is LEARNING is introduced into the model. The graph of the semiparametric fit is given in Fig. 6. The impact of complaints, privileges and learning were considered individually on ratings as presented in Fig. 6 and they indicate good fits because as the explanatory variables increases, the fit or approximation becomes better. The three variables produced estimates with straight lines with the fit of rating and privileges displaying a negative relationship and this also support the scatterplot of rating and privileges. The fit of the impact of complaints, privileges and learning on rating with the nonparametric estimate becomes difficulty to obtain due to the problem of curse of dimensionality. As the number of explanatory variable increases, it becomes practically impossible to obtain the estimates of the variable graphically with the nonparametric estimation. In Table 3, the adjusted R^2 value of the semiparametric model is better than the nonparametric estimate. The semiparametric method was able to explain up to 70% of the variations on rating while the nonparametric estimation could explain about 67% only. The fit of the nonparametric estimates with the three variables cannot be obtained due to curse of dimensionality effect usually connected with nonparametric estimation. This curse of dimensionality problem with nonparametric estimation can be addressed by the semiparametric method that considers the relationship of the response variable with each of the explanatory variables individually.

TABLE 3: MODEL ADJUSTED R-SQUARE VALUES

Model	Semiparametric	Nonparametric
Adjusted R-Square	0.70	0.67

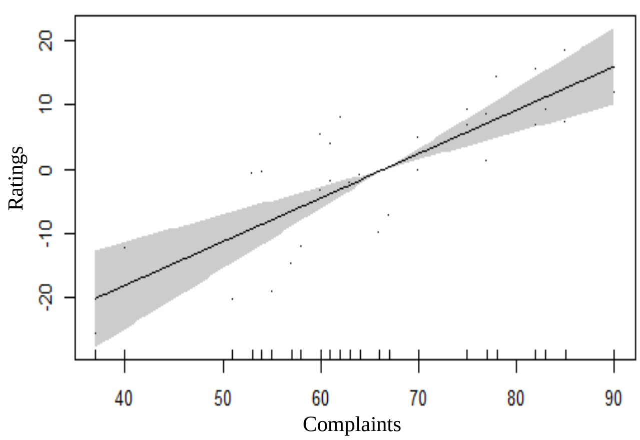


Fig. 6a: *Impact of complaints on rating with semiparametric fit.*

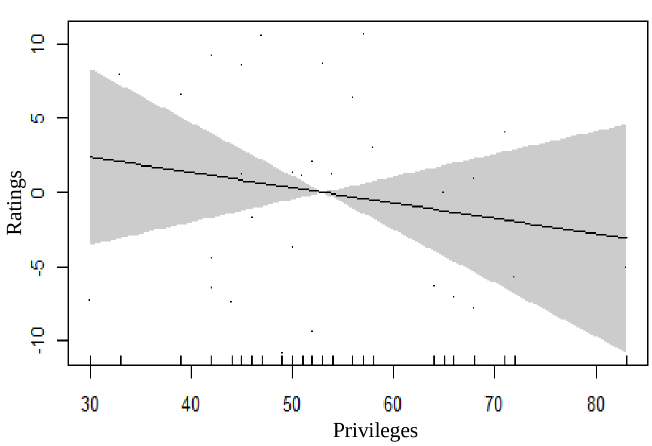


Fig. 6b: *Impact of privileges on rating with semiparametric fit.*

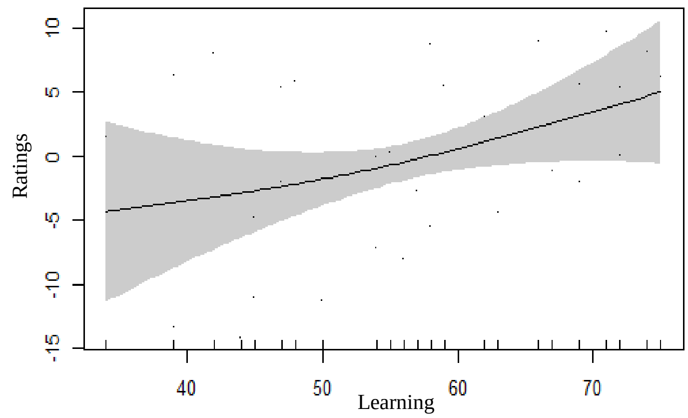


Fig. 6c: *Impact of learning on rating with semiparametric fit.*

The kernel density plots of the data are in Fig. 7 showing that these variables do have impact on the response variable. However; their level of impact will be best determined by their probability values. The kernel estimation method is one of the most applied nonparametric methods in exploratory data analysis and data visualization. The Gaussian kernel is employed in the construction of the kernel estimates for the three variables and the bandwidths for the construction of the kernel estimates in Fig. 7 are in Table 4 with the parameter estimates of the data. The bandwidth determines the smoothness of the kernel estimates. The mode of the complaint's variable is 61 with probability of 0.025 while the mode of privileges is 49 with probability of 0.030. The modal value of learning is 52 with probability of 0.025. Higher probability value shows that the supervisor is unwilling in addressing the request of the employees with respect to that variable.

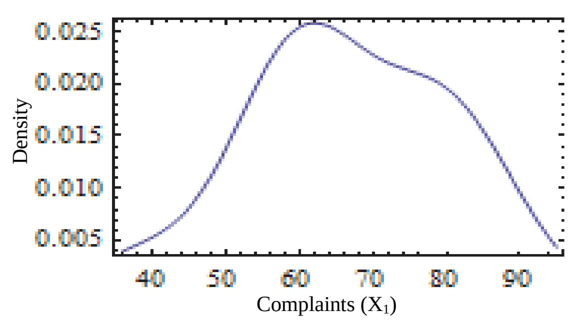


Fig. 7a: *Kernel density estimate of complaints.*

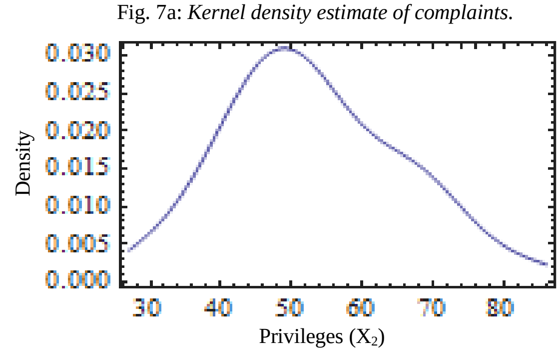


Fig. 7b: *Kernel density estimate of privileges.*

DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

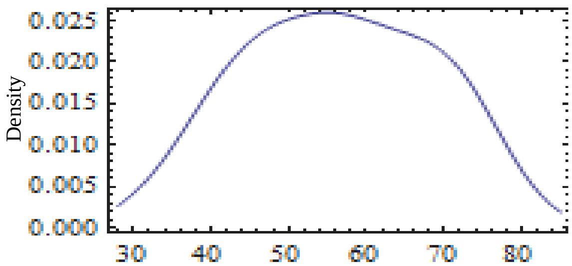


Fig. 7c: *Kernel density estimate of learning*.

As discovered from the scatterplot of rating and privileges, the supervisor seems to be reluctant in granting certain privileges of the employees which could have enhanced the performance of the employees and increase the total output of the organization. The adjusted R-square value of the semiparametric additive model and the nonparametric estimate are in Table 3 with the semiparametric model vividly displaying superiority over the nonparametric estimation in performance.

TABLE 4: PARAMETER ESTIMATES FOR THE EMPLOYEE ATTITUDE SURVEY DATA

Variables	Coefficients	Standard Error	t-values	p-values	Bandwidths
Constant	11.258305	7.3183404	1.538369	0.1360411	-------
Complaints (X_1)	0.2379762	0.1394103	1.707021	0.0997355	7.1486
Privileges (X_2)	–0.1032843	0.1293454	–0.798515	0.4318056	6.5688
Learning (X_3)	0.6824165	0.1288445	5.296434	0.0000154	6.3014

The additive model in Equation (2) with the analysed data can be written as

$$Y = \alpha + g_1(X_1) + g_2(X_2) + g_3(X_3),$$

where X_1 is Complaints, X_2 is Privileges and X_3 is Learning. Hence; the linear component of the model from the parameter estimates in Table 4 is of the form

$$Y = 11.258305 + 0.2379762X_1 - 0.1032843X_2 + 0.6824165X_3$$

The p-value of Privileges which is 0.4318056 is exceedingly high and that implies the effects on the overall performance of the organisation is insignificant since the supervisor is unwilling to grant such privileges. However; the p-values of Complaints and Learning seems small especially that of Learning and this confirms that Learning has a significant effect on the performance of the organisation since new technologies of managing an organisation are evolving. Hence; there is need for regular training of staff members to meet up with the current demand.

Again, the standard error for the three variables is small and that simply implies that the variables have means that are close. The t-value of Privilege is low when compared with the other variables and this suggests that its impact on the overall performance is minimal.

5. CONCLUSION.

The numerical and graphical results of the analysis show that as the number of the explanatory variable increases, the fit of the semiparametric additive model with the Backfitting algorithm improves. The curse of dimensionality effect is made evident as a problem of the nonparametric estimation due to the complexity associated with the graphical presentation of the explanatory variables beyond two variables. The semiparametric approach has however solved the problem of providing a better fit with fast convergence rate. The graphical plots for the two methods clearly show that the semiparametric model improves as the number of explanatory variables increase and with the predictive power of the model also increasing.

ACKNOWLEDGEMENT

The authors sincerely appreciate the editorial members and the anonymous reviewers for painstakingly going through the manuscript and making positive suggestions that improved the quality of the manuscript.

COMPETING INTERESTS.

The authors declared that they have no conflict of interest.

REFERENCES

[1]. S. U. Ejakpovi, I. U. Siloko, and E. Enoyoze, "On performance of multivariate product

kernel density estimation consistency," The Nigerian Journal of Science and Environment, 2020, 18(2), pp. 66–70.
[2]. I. U. Siloko, K. E. Ukhurebor, C. C. Ishiekwene, E. A. Siloko, O. O. Uddin, and E. Enoyoze, "Statistical estimation of some meteorological variables using the beta kernel function," Ethiopian Journal of Environmental Studies and Management, 2021a, 14(4), pp. 474–486.
[3]. J. S. Simonoff, "Smoothing methods in statistics, Springer Science & Business Media, New York, 2012.
[4]. I. U. Siloko, S. U., Ejakpovi, K. E. Ukhurebor, E. A. Siloko, and C. C. Ishiekwene, "Comparison of kernel density function in detecting effects of daily emission of sulphur (iv) oxide from an industrial plant", Ethiopian Journal of Science and Technology, 2022, 15(1), pp. 53–66.
[5]. D. J. Henderson and C. F. Parmeter, "Applied nonparametric econometrics," Cambridge University Press, 2015.
[6]. I. U. Siloko, O. Ikpotokin, and E. A. Siloko, "A key note on performance of smoothing parameterizations in kernel density estimation", Tanzania Journal of Science, 2019a 45(1), pp. 1–8.
[7]. X. Yafeng, and Z. Lirong, "Variable selection for partially linear additive model based on modal regression under high dimensional data", International Journal of Statistical Distributions and Applications, 2020, 6(1), pp. 1–9.
[8]. R. Stocker, and A. Adekpedjou, "A general class of additive semiparametric models for recurrent event data", Journal of Statistical Planning and Inference, 2020, 205, pp. 231–244.
[9]. I. U. Siloko, O. Ikpotokin, F. O. Oyegue, C. C. Ishiekwene, and B. A. Afere, "A note on application of kernel derivatives in density estimation with the univariate case", Journal of Statistics and Management Systems, 2019b, 22(3), pp. 415–423.
[10]. B. E. Hansen, "Exact mean integrated squared error of higher order kernel estimators" Econometric Theory, 2005, 21, pp. 1031–51057.
[11]. I. U. Siloko, O. Ikpotokin, and E. A. Siloko, E. A., "On evaluation of smoothing matrix performance in multivariate kernel density estimation", Edited Proceedings of 2nd International Conference of Professional Statisticians Society of Nigeria, 2018a, 2, pp. 268–272.
[12]. D. W. Scott, "Multivariate density estimation. Theory, practice and visualization", Second Edition, Wiley, New Jersey, 2015.
[13]. B. W. Silverman, "Density estimation for statistics and data analysis", Routledge, New York, 2018.
[14]. I. U. Siloko, E. A. Siloko, and O. Ikpotokin, "A mini review of dimensional effects on asymptotic mean integrated squared error and efficiencies of selected beta kernels", Jordan Journal of Mathematics and Statistics, 2020, 13(3), pp. 327–340.
[15]. B. E. Hansen, Econometrics, University of Wisconsin, Spring, 2013.
[16]. Z. I. Botev, J. F. Grotowski, and D. P. Kroese, "Kernel density estimation via diffusion", Annals of Statistics, 2010, 28, pp. 2916–2957.
[17]. M. P. Wand, and M. C. Jones, "Kernel Smoothing", Chapman and Hall, London, 1995.
[18]. J. E. Chacón, "Data–driven choice of the smoothing parameterization for kernel density estimators", Canadian Journal of Statistics, 2009, 37, pp. 249–265.
[19]. I. U. Siloko, C. C. Ishiekwene, and F. O. Oyegue, "New gradient methods for bandwidth selection in bivariate kernel density estimation", Journal of Mathematics and Statistics, 2018b, 6(1), pp. 1–8.
[20]. M. Delgado, and P. M. Robinson, "Nonparametric and Semiparametric Methods for Economic Research", Journal of Economic Surveys, 1992, 6, pp. 201–250.
[21]. J. L. Powell, "Semiparametric estimation. The New Palgrave Dictionary of Economics, Second Edition, 2008.
[22]. H. Ichimura, and P. E. Todd, "Implementing nonparametric and semiparametric estimators", Handbook of Econometrics, 2006.
[23]. N. H. Bingham, and R. Kiesel, "Semiparametric modelling in finance: Theoretical Foundations", Quantitative Finance, 2001, 1, pp. 1–10.
[24]. G. Koop, D. J. Poirier, and J. Tobias, "Semiparametric bayesian inference in multiple equations models", Department of Economics, University of Leicester, England, 2004.
[25]. H. Liang, S. Thurston, D. Ruppert, T. Apanasovich, and R. Hauser, "Additive partial linear models with measurement errors", Biometrika, 2008, 95, pp. 667–678.
[26]. Y. Ma, and M. G. Genton, "Explicit estimating equations for semiparametric generalized linear latent variable models", Journal of the

DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 18, ISSUE 2, JULY 2023
ISSN 1818-5878 (Print) 2408-8498 (Online)

Royal Statistical Society, Series B (Methodological), 2010, 72(4), pp. 475–495.

[27]. V. Chernozhukov, I. Fernández-Val, J. Hahn, and W. Newey, "Average and quartile effects in nonseparable panel models", Unpublished Manuscript, MIT, 2013.

[28]. J. Escanciano, D. Jacho-Chávez, and A. Lewbel, "Uniform convergence of weighted sums of nonparametric and semiparametric residuals for estimation and testing", Journal of Econometrics, 2014, 178(3), pp. 426–443.

[29]. P. Bach, H. Farbmacher, and M. Spindler, "Semiparametric count data modelling with an application to health service demand", Econometric and Data Group, University of York, UK, 2016.

[30]. P. Nie, and A. Sousa-Poza, "A fresh look at calorie-income elasticities in China", China Agricultural Economic Review, 2016, 8(1), pp. 55–80.

[31]. E. Mammen, C. Rothe, and M. Schienle, "Semiparametric estimation with generated covariates", SFB 649 Discussions Paper (Economic Risk Berlin), 2014.

[32]. J. Hambuckers, T. Kneib, R. Langrock, and A. Sohn, "A markov-switching generalized additive model for compound poisson processes, with applications to operational losses models", Georg-August-Universität Göttingen, Chair of Statistics, Humboldtallee 3, 37073 Göttingen, Germany, 2016.

[33]. W. Härdle, M. Müller, S. Sperlich, and A. Werwatz, A., "Nonparametric and semiparametric models", Springer-Verlag Berlin Heidelberg New York, 2004.

[34]. X. Tian, and X. Yu, "Using semiparametric models to study nutrition improvement and dietary change with different indices: The case of China", Food Policy, 2015, 53, pp. 67–81.

[35]. S. N. Wood, S. Goude, and S. Shaw, "Generalized additive models for large data sets", Journal of the Royal Statistical Society: Series C Applied Statistics, https://doi.org/0035-9254/15/64000, 2015.

[36]. O. Linton, and J. P. Nielsen, "A kernel method of estimating structured nonparametrically regression based on marginal integration", Biometrika, 1995, 82, pp. 93–101.

[37]. E. Severance-Lossin, and S. Sperlich, "Estimation of derivative for additive separable models", Statistics, 1999, 33, pp. 241–265.

[38]. E. Chu, A. Keshavarz, and S. Boyd, "A distributed algorithm for fitting generalized additive models", Optim Eng. 2013, 14, pp. 213–224.

[39]. T. H. Trinh, C. Thomas-Agnan, and M. Simioni, "Calorie intake and income in China: New evidence using semiparametric modelling with generalized additive models", Toulouse School of Economics, University of Toulouse Capitole, France, 2017.

[40]. C. Wei, R. Yan, S. Guo, and T. Tao, "Statistical inference on semiparametric spatial additive model", Journal of Mathematics Research, 12(2), https://doi:10.5539/jmr.v12n2p1, 2020.

[41]. X. Zheng, Y. Rong, L. Liu, and W. A. Cheng, "More accurate estimation of semiparametric logistic regression, Mathematics https://doi.org/10.3390/math9192376, 2021.

[42]. T. J. Hastie, and R. J. Tibshirani, "Varying-coefficient models", Journal of the Royal Statistical Society, Series B (Methodological), 1993, 55(4), pp. 757–796.

[43]. M. Rezapour, and K, Ksaibati, "Semi and nonparametric conditional probability density, a case study of pedestrian crashes", The Open Transportation Journal, 2021, 15, pp. 280–288.

[44]. X. Shen, I. Karibzhanov, H. Tsurumi, and S, Li, "Comparison of Bayesian and sample theory parametric and semiparametric binary response models", https://doi.org/10.34989/swp, 2022.

[45]. S. Helali, A. Masmoudi, and Y. Slaoui, "Semi-parametric estimation using Bernstein polynomial and a finite Gaussian mixture model", Entropy, https:// doi.org/10.3390/e24030315, 2022.

[46]. F. Kwasniok, "Semiparametric maximum likelihood probability density estimation", PLoS ONE 16(11), https://doi.org/10.1371/journal.pone.0259111, 2021.

[47]. F. Chai, Y. Li, X. Zhang, and Z. Chen, "Daily semiparametric GARCH model estimation using intraday high-frequency data", Symmetry, 15, 908, https:// doi.org/10.3390/sym15040908, 2023.

[48]. L. P. Damasceno, C. C. Cavalcante, T. Adalı and Z. Boukouvalas, "Independent vector analysis using semi-parametric density estimation via multivariate entropy maximization", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, 2021, pp. 3715–3719, DOI:10.1109/ICASSP39728.2021.9414839

[49]. G. Frasso, and P. H. C. Eilers, "Direct semi-parametric estimation of the state price density implied in option prices", Journal of Business and Economic Statistics, 2022, 40(3), pp.1179–1190, DOI:10.1080/07350015.2021.1906686

[50]. J. Hahn, and G. Ridder, Asymptotic variance of semiparametric estimators with generated regressors, Econometrica, 2013, 81(1), pp. 315–340.
[51]. S. J. Sheather, “Density estimation”, Statistical Science, 2004, 19(4), pp. 588–597.
[52]. I. U. Siloko, K. E. Ukhurebor, E. A. Siloko, E. Enoyoze, A. O. Bobadoye, C. C. Ishiekwene, O. O. Uddin, and W. Nwankwo, “Effects of some meteorological variables on cassava production in Edo state, Nigeria via density estimation, Scientific African, e00852, https://doi.org/10.1016/j.sciaf.2021.e00852, 2021.
[53]. T. Ruzgas, M. Lukauskas, and G. Cepkauskas, “Nonparametric multivariate density estimation: Case study of Cauchy mixture model”, Mathematics, https://doi.org/10.3390/math9212717, 2021.
[54] I. U. Siloko and E. A. Siloko, “An investigation on interdependence between rainfall and temperature in Ekpoma, Edo State, Nigeria”, Iranian (Iranica) Journal of Energy and Environment, 2023, 14(3), pp. 197–204.
[55]. J. A. Khan, and A. Akbar, “Density estimation using Gumbel kernel estimator”, Open Journal of Statistics, 2021, 11, pp. 319–328.

[56]. I. U. Siloko, K. E. Ukhurebor, E. A. Siloko, E. Enoyoze, and O. Ikpotokin, “The interactions between temperature and relative humidity: Results for Benin City, Nigeria using statistical analysis”, Current Applied Science and Technology, 2022, 22(1), pp. 1–15.
[57]. J. E. Chacón, and T. Duong, “Efficient recursive algorithms for functional base on higher order derivatives of the multivariate Gaussian density”, Statistical Computing, 2015, 25, pp. 959–974.
[58]. S. M. Somé, and C. C. Kokonendji, “Bayesian selector of adaptive bandwidth for multivariate gamma kernel estimator on $[0,\infty)^d$”, Journal of Applied Statistics, 2022, 49 (7), pp. 1692–1713.
[59]. S. Chatterjee, and B. Price, “Regression Analysis by Example”, Second Edition, New York: Wiley, 1977.
[60]. I. U. Siloko, E. A. Siloko, O. Ikpotokin, C. C. Ishiekwene, and B. A. Afere, B. A., “On asymptotic mean integrated squared error’s reduction techniques in kernel density estimation”, International Journal of Computational and Theoretical Statistics, 2019c, 6(1), pp. 89–98.