The study proposes a smoothing method which is the arithmetic weighted value of Generalized Cross-Validation (GCV) and Unbiased Risk (UBR) methods. This study concluded that the PSM method provides the best-fit as a smoothing method, works well at autocorrelation levels ($\rho=0.2$, 0.5 and 0.8), and does not overfit time-series observations. The study recommended that the proposed smoothing is appropriate for time series observations with autocorrelation in the error term and econometrics real-life data. This study can be applied to: non – parametric regression, non – parametric forecasting, spatial, survival and econometrics observations.

Dr. Samuel Olorunfemi Adams,
Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria. Dr. Adams has well over 25 academic publications comprising of local and international journals as well as over 10 papers presented in learned conferences in and outside Nigeria.

9 786206 151890

Spline Smoothing with Autocorrelation

Samuel Olorunfemi Adams

Samuel Olorunfemi Adams

Smoothing Spline Technique For Time Series Data with Autocorrelation

LAP LAMBERT Academic Publishing

Samuel Olorunfemi Adams

Smoothing Spline Technique For Time Series Data with Autocorrelation

Samuel Olorunfemi Adams

Smoothing Spline Technique For Time Series Data with Autocorrelation

LAP LAMBERT Academic Publishing

Imprint

Cover image: www.ingimage.com

Publisher:
LAP LAMBERT Academic Publishing
is a trademark of
Dodo Books Indian Ocean Ltd. and OmniScriptum S.R.L publishing group

120 High Road, East Finchley, London, N2 9ED, United Kingdom
Str. Armeneasca 28/1, office 1, Chisinau MD-2012, Republic of Moldova, Europe
Printed at: see last page
ISBN: 978-620-6-15189-0

SMOOTHING SPLINE TECHNIQUE FOR TIME SERIES DATA WITH AUTOCORRELATION

BY

Samuel Olorunfemi Adams

ABSTRACT

Spline smoothing is a technique used to filter out noise in time series observations when predicting nonparametric regression models. Its performance depends on the choice of smoothing parameter lambda. Most of the existing smoothing methods applied to time series data tend to over-fit in the presence of autocorrelated errors. The aim of this study is to propose a smoothing method which is the arithmetic weighted value of Generalized Cross-Validation (GCV) and Unbiased Risk (UBR) methods. The objectives of the study were to: (i) determine the best-fit smoothing method for the time series observation; (ii) identify the best smoothing method that does not over-fit time-series data when autocorrelation is present in the error term; (iii) establish the optimum value of the proposed smoothing method; (iv) compare GCV, GML and UBR smoothing methods to the proposed smoothing methods in terms of sample size; and (v) test the results of simulation using real life-data.

A hybrid smoothing method of the Generalized Cross-Validation (GCV) and Unbiased Risk (UBR) was developed by adding the weighted values of Generalized Cross-Validation (GCV) and Unbiased Risk (UBR). The Proposed Smoothing Method (PSM) was compared with Generalized Maximum Likelihood (GML), GCV and UBR smoothing methods. A Monte Carlo experiment of 1,000 trials was carried out at three different sample sizes (20, 60 and 100), three levels of the autocorrelation (0.2, 05 and 0.8), and four degrees of smoothing (1, 2, 3 and 4). Real-life data on Standard International Trade Classification (SITC) export and import price indices in Nigeria between 1970 – 2018 extracted from CBN 2019 edition were also used. The four smoothing methods' performances were estimated and compared using the Predictive Mean Squared Error (PMSE) criterion.

The findings of the study revealed that:

(i) for a time series observation with autocorrelated errors,

$Adj.\ R^2(PSM\ \lambda = 0.04) = \left(1 - \frac{(1-Rsquare)\times(n-1)}{n-p}\right) = 0.9619$, provides the best-fit smoothing method for the model ;

(ii) the PSM does not over-fit data at all the autocorrelation levels considered ($\rho = 0.2, 0.5\ and\ 0.8$);

(iii) the optimum value of the PSM was at the weighted value of 0.04, with the the equation is given as $PSM\ (\lambda) = (0.04)\frac{(y-\hat{f})^T W(y-\hat{f})}{[trace(I-S\lambda)]^2} + (0.96)\frac{\frac{1}{n}\left\|W^{\frac{1}{2}}(I-S\lambda)\right\|^2}{\left[\frac{1}{n}trace\{W(I-S\lambda)\}\right]^2}$;

(iv) when there is autocorrelation in the error term, PSM performed better than the GCV, GML and UBR smoothing methods were considered at all-time series sizes (T = 20, 60 and 100);

(v) for the real-life data employed in the study, PSM proved to be the most efficient among the GCV, GML, PSM and UBR smoothing methods compared.

The study concluded that the PSM method provides the best-fit as a smoothing method, works well at autocorrelation levels (ρ=0.2, 0.5 and 0.8), and does not overfit time-series observations. The study recommended that the proposed smoothing is appropriate for time series observations with autocorrelation in the error term and econometrics real-life data. This study can be applied to; non – parametric regression, non – parametric forecasting, spatial, survival and econometrics observations.

TABLE OF CONTENTS

LIST OF TABLES

Tables **Pages**

LIST OF FIGURES

LIST OF ABBREVIATIONS

The following abbreviations and terminologies were used in this research work

ABBREVIATIONS	MEANING
T	Time Series period/Sample size
λ	Smoothing Parameter
ρ	Autocorrelation level
σ	Standard Deviation
D.S.	Degree of smoothness
AR(1)	Autoregressive process of order one
ARCH	Autoregressive Conditional Heteroscedasticity
ARMA	Autoregressive Moving Average
ASB	Average squared Bias
ASE	Average Squared Error
COCR	Cochrane and Orcutt Regression
EXPAR	Exponential Autoregressive
fMRI	Functional Magnetic Resonance Imaging
GACV	Generalized Approximate Cross-validation
GCV	Generalized Cross-validation
GFAIC	Parallel of Akaike's Information Criterion
GML	Generalized Maximum Likelihood
GLS	Generalized Least Squares
MLE	Maximum Likelihood Estimator
MSE	Mean Squared Error
PMSE	Predictive Mean Squared Error

P-SPLINE	Penalized Spline
PSM	Proposed Smoothing Method
REML	Restricted Maximum Likelihood
RGCV	Robust Generalized Cross-Validation
RMSE	Root Mean Square Error
STAR	Smooth Transition Autoregressive
TAR	Threshold Autoregressive
UBR	Unbiased Risk
WMSE	Weighted Mean Squared Error

CHAPTER ONE

GENERAL INTRODUCTION

1.0 Background to the Study

In nonparametric regression, smoothing is vital because it is utilized to eliminate noise or disturbances influences in an observation (Eubank, 1988). It can generally be applied to estimate the mean function of the non-parametric regression model (Green and Silverman, 1994), and it is likewise quite possibly the most famous method utilized for the forecast of nonparametric regression, (Hastie and Tibshirami, 1990), (Wahba, 1990). The majority of the previous research relating to spline smoothing was based on data sets that are often correlated in nature; for example, clustered, spatial and time-series data, Wahab (1990). Research has proven that correlation strongly affects the smoothing parameter's estimation, critical to smoothing Spline, (Diggle and Hutchinson, 1989) efficiency. The most well-known technique utilized in time series analysis is the classical Autoregressive Moving Average (ARMA) approach; it assumes linear dependence on past values and innovation. Stheim (1994) utilized classical linear modeling with classical data set and found that it has limitations; the increase information on deviations from the classical ARMA model is projected by spline smoothing researchers in nonlinear time series analysis.

The most popularly used nonlinear models are the smooth-transition autoregressive (STAR) models, trend estimation and forecast, Morton, Kang and Henderson (2009), edge autoregressive models, (Tong, 1978) and (1983); (Subba and Gabr, 1980), the Autoregressive Conditional Heteroscedastic models (Engle, 1982), the dramatic autoregressive models, (Haggan and Ozaki, 1981), Tong and Chan (1986); Granger and Terasvirta (1992), the irregular coefficient models, Nicholls and Quinn (1982), the bilinear models, (Subba, 1981), (Anderson and Granger, 1978);

with generalized ARCH models (Higgins and Bera, 1993) and (Bollerslev, 1986). The nonparametric spline smoothing method of estimation has been abandoned for some time. The research in curve smoothing has recently recommended a parametric methodology for its simplicity in the calculation, consistency with model assumption, and statistical convenience. The Generalized Cross-Validation parameter estimation method was extended to accommodate an ARMA time series with autocorrelation in its error term, Diggle and Hutchinson (1989). The best method of dealing with these autocorrelated error terms is temporal smoothing, which formed this research work. Smoothing techniques have been used as a major technique in the nineteenth century for the nonparametric regression empirical analysis.

Spline smoothing had been applied to time series data sets with autocorrelated error term by many researchers in the past; estimation utilizing repeated measurement data, (Hart, 1986). Regression with the presence of autocorrelation in the disturbance (Hurvich, 1990), Times series Moving average, Kohn, Ansley, and Wong (1992), fMRI time-arrangement perceptions, Carew *et al*., (2003), modelling and forecasting time series, Shen (2008), for trend estimation and prediction, Morton, Kang and Henderson (2009), fMRI time-series data revisited, Worsley and Friston (1995).

Many Spline smoothing researchers have studied modeling of time series observations with Generalized Cross Validation (GCV), Generalized Maximum Likelihood (GML) and Unbiased Risk (UBR) method. Diggle and Hutchinson (1989) extended the GCV method to estimate the smoothing parameter and autocorrelated error term. Kohn, Ansley, and Wong (1992) represented a smoothing spline by a state-space model and extended the CV, GCV and GML estimation methods to an autoregressive moving average error term. Hurvich and Zeger (1990) used a Cross-Validation method to estimate some smoothing parameters and more recently Yeudong

(2011) extended GML, GCV and UBR methods to estimate smoothing parameter when data are correlated. Almost all of these methods were developed for time series observations while some others require that the design points are equally spaced.

In this study, three existing smoothing techniques (GML, GCV and UBR) were compared with a proposed smoothing method for a time series data with autocorrelated error terms.

Nonparametric Regression Models

1.1.1 Nonlinear Regression Model

A common nonlinear regression model is written as;

$$y_i = f(\alpha_i X'_i) + \varepsilon_i \tag{1.1}$$

Where; $\alpha_i = (\alpha_1, \alpha_2, \ldots, \alpha_P)'$ is a vector of parameters to be estimated and $X_i = (X_1, \ldots, X_k)$ is a vector of predictors for the ith observations; the errors are assumed to be normally and independently identically distributed with mean 0 and constant variance σ^2. The function $f(\cdot)$, is the average value of the response Y to the predictors, it is specified in advance because as in the case of a linear regression model.

The nonparametric regression model is is written in a similar manner, but the function f is left unspecified;

$$y_i = f(X'_i) + \varepsilon_i \tag{1.2}$$

$$y_i = f(x_{i1}, x_{i2}, .. x_{ik}) + \varepsilon_i \tag{1.3}$$

The objective of nonparametric regression is to estimate the unknown regression function $f(\cdot)$ directly, rather than estimating the parameters. Most nonparametric regression methods assume that $f(\cdot)$ is an unknown smoothing function and $\varepsilon_i \sim$ normally independently distributed (0, σ^2).

An essential and important form of the general model of nonparametric simple regression, with only one predictor is given as;

$$y_i = f(x_i) + \varepsilon_i \qquad a < x_1 < \ldots < x_n < b \tag{1.4}$$

Nonparametric simple regression is usually referred to as 'scatterplot smoothing' because it's application require tracing a smooth curve through a scatterplot of y against x, Diggle and Hutchinson (1989). The simple nonparametric regression is always used because it is difficult to fit the general nonparametric regression model when there are many predictors, and it is also very difficult to display the fitted model when there are more than two predictors. More restrictive models have been developed, one such model is the additive regression model given as;

$$y_i = \beta + f_1(X_{I1}) + f_2(X_{12}) + \ldots + f_k(X_{1k}) + \varepsilon_{1i} \tag{1.5}$$

Where; the partial regression functions $f_i(\cdot)$ are assumed to be the smoothing functions, and are to be estimated from the observation. This model is significantly more exclusive than the general nonparametric regression model, but less exclusive than the linear regression model, which assumes that all of the partial-regression functions are linear, Wahba *et.al* (1995).

Variations on the additive regression model include semi parametric models, in which some of the predictors enter linearly, for example;

$$y_i = \beta + \alpha_1 x_{i1} + f_2(x_{i2}) + \ldots + f_k(x_{ik}) + \varepsilon_i \tag{1.6}$$

Nason and Silverman (1997) believed that the addictive regression model is valuable when some of the predictors are factors and some predictors enter are related, which appear as higher-dimensional terms in the model, for example;

$$y_i = \alpha + f_{12}(x_{i1}, x_{i2}) + f_3(x_{i3}) + \ldots + f_k(x_{ik}) + \varepsilon_i \tag{1.7}$$

These models extend towards generalizing nonparametric regression models and extend to Generalize Linear Models. The linear predictor of GLM is written as;

$$\eta_i = \beta + \alpha_i x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} \tag{1.8}$$

Which can also take the form, for an instant an unidentified smooth function in explanatory variables;

$$\eta_i = f(x_{i1}, x_{i2}, \ldots, x_{ik}) \tag{1.9}$$

for the most common case, or by a sum of smooth partial-regression functions

$$\eta_i = \alpha + f_1(x_{i1}) + f_2(x_{i2}) + \ldots + f_k(x_{ik}) \tag{1.10}$$

for the generalized additive model.

All of these nonparametric regression models with the projection-pursuit regression, classification and regression trees have been discussed by researchers. There are many nonparametric regression model estimation methods, but this research shall consider the classical smoothing splines method for time series observation with autocorrelation in its error term.

1.1.2 Cubic Spline Smoothing

The cubic spline is a spline consisting of piecewise third-request polynomials that go through a bunch of control focuses. The second subsidiary of every polynomial is ordinarily set to zero at the endpoints since this gives a limit condition that finishes the framework condition of m-2 conditions, this creates a purported "normal" cubic spline and prompts a straightforward tridiagonal framework that can be settled effectively to give the coefficients of the polynomials. The parameters are estimated by minimizing the residual sum of square (RSS) and a roughness penalty. A general test of “loyalty to observation" for a curve g is the residual sum of squares. If *g* is allowed to be any curve – unrestricted in functional form, then this distance test can be reduced to zero by any *g* that interpolates the observation. The curve would not be admitted because it is not exclusive and because it is has a structure oriented interpretation, Wahba (1993). Spline smoothing approach avoids impossible interpolation of the observation by evaluating the contest between the tasks of producing a good fit to the observation producing a curve without too much rapid local change. The main function of splines is for interpolation, but it can also be used for parametric and non-parametric regression modelling, the most commonly used spline smoothing technique is the cubic splines. Spline smoothing produces another technique for local polynomial regression and it is also a charming component of additives regression models.

Smoothing spline is a solution to a nonparametric regression problem with the function $g(x)$ i.e. Find $\hat{g} \in C^2\,[a, b]$ in the model (1.11) that minimizes the penalized residual sum of squares with two continuous derivatives, as given below;

$$S(g) = \sum_{i-1}^{n} \left(y_i - g(x_i)\right)^2 + \lambda \int_a^b \left(g''(x)\right)^2 dx \qquad (1.11)$$

- λ is a smoothing parameter,
- The first term in the equation is the residual sum of the square for the goodness of fit of the data
- The second term is a roughness penalty, which is large when the integrated second derivative of a regression function $g''(x)$ is also large.
- If λ approaches 0 then $g(x)$ simply interpolates the observations.
- If λ is very large, then g(x) will be selected so that $g''(x)$ is everywhere 0, which implies an overall linear least-squares fit the observations.

There are several ways to measure local variation, one could define a test of roughness based on the second derivatives. In order to explain the main ideas, the integrated squared and second derivative is mostly used, that is, the roughness; $\int_a^b \left(g''(x)\right)^2 dx$ is used to measure the local variation.

The solution based on smoothing spline for a minimum problem in equation (1.11) is known as a "natural cubic spline" with knots at $x_1, \ldots, x_n$. From this point of view, a specially structured spline interpolation which depends on a chosen value λ develops into a suitable approach of function g in a model (1.12). Let $g = (g(x_1), \ldots, g(x_n))$ be the vector of values of function g at the knot points $x_1, \ldots, x_n$. The smoothing spline estimate $\hat{g}_x$ of this vector for the fitted data $y = (y_1, \ldots, y_n)^T$ are projected by;

$$\hat{g}_\lambda = \begin{bmatrix} \hat{g}_\lambda(x_1) \\ \hat{g}_\lambda(x_2) \\ \cdot \\ \cdot \\ \cdot \\ \hat{g}_\lambda(x_n) \end{bmatrix}_{(n \times 1)} = S_\lambda \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}_{(n \times 1)} \quad \text{or } \hat{g}_\lambda = S_\lambda \tag{1.12}$$

where $\hat{g}_\lambda$ is a natural cubic spline with knots at $x_1, \ldots, x_n$ for a fixed $\lambda > 0$, and S_λ is a well-known positive-definite (symmetrical) smoother matrix which depends on λ and the knot points $x_1, \ldots, x_n$, but not on y. The function $\hat{g}_\lambda$, is the estimation of function g, it is obtained by cubic spline interpolation that depends on the condition $\hat{g}_\lambda(xi) = (\hat{g})i$, $i = 1, 2, \ldots, n$., Dursun et.al (2013), Eubank (1988), Green and Silverman (1994) and Wahba (1990).

A survey of the literature on the mean squared error properties of spline smoothing can be found in Eubank (1988). The smoothing parameter selection problem for this class of estimators has been mainly investigated by Wahba (1975), from her rich collection of results, it is imperative to mention those which are directly connected with the optimization of λ. In Wahba (1975) and Wahba (1979) convergence level of splines were considered, the first article on cross-validation was later extended to the smoothing of the log periodogram, Wahba (1975). The term

Generalized Cross-Validation (GCV) was coined by Wahba (1977) and (1985), it is a minimax type approach to the question of rates of convergence and exact bounds for the integrated squared error under a normality assumption on the error term.

1.1.3 Kernel Spline Smoothing Model

An alternative method to a design of the weight sequence $W_{ni}(y)$, is to give a detailed explanation of the weight function $Wn_i(y)$ shape by a density function with a scale parameter that changes the size and form of the weights close to y. It is appropriate to refer to this shape function as a kernel K.

Generally speaking a Kernel smoother defines a set of weights $\{W_i(y)\}_{i=1}^{n}$ for each x and defines as;

$$\hat{f}(y) = \sum_{i=1}^{n} W_i(y)x_i \tag{1.13}$$

$$\int k(U)du = 1 \tag{1.14}$$

The Kernel smoothers of the weight sequence (one-dimensional y) are given by;

$$W_{hi}(y) = \frac{K\left(\frac{y-y_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{y-y_i}{h}\right)} \tag{1.15}$$

Where; $\sum_{i=1}^{n} W_{hi}(y)x_i = 1$

The Kernel Smoother is then defined for any y as before by;

$$\hat{f}(y) = \sum_{i=1}^{n} W_{hi}(y)X_i \tag{1.16}$$

The points that are close together are similar, a kernel smoother usually defines weights that decrease in a smooth fashion as one moves away from the target point. Running mean smoothers are kernel smoothers that use a "box" kernel. A natural candidate for K is the standard Guassian density. (This is very inconvenient computationally because its never zero).

1.1.4 Penalized Spline Non-Parametric Regression Model

Penalized spline smoothing or simply P-splines is one of the most popular methods of removing noise or disturbance from data, it is generally easy to fit, it is flexible in the choice of knots and smoothing parameter. The choice of adding regression spline to a small number of inner knots with reasonable time series size, and the introduction of the penalized parameter, which affects the fitting accuracy and smoothness of the regression curve, can be traced back to O'Sullivan (1986). Penalized Spline Non-Parametric regression model originates from the model;

$$y \sim N(m(x_i), \sigma_\varepsilon^2), \text{I} = 1, 2, \ldots, \text{n} \tag{1.17}$$

Where; (x) is the smoothing function, it is an unknown regression function that needs to be estimated based on the data set $(xi,\ yi)$, $i = 1, \ldots, n$ and σ_ε^2 is the constant variance of the random deviation error of the response variable and the regression function $m(x)$. It is an easier way of estimating the penalized spline smoothing since different function may correspond to different penalized spline smoothing regression function. The best method of Selecting Penalized Non-Parametric regression is the truncated polynomial given as;

$$m(y;\alpha) = \alpha_0 + \alpha_1 y, + \cdots + \alpha_p y^p + \sum_{j=1}^{k} \alpha_{p+j} \left(y - k_j\right)_+^p \tag{1.18}$$

Where; $(y - \kappa j)_+ = \max(y - \kappa_j, 0)$ and $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{p+K})^T$ denotes a vector of regression coefficients, $\kappa_j, \ldots, \kappa$ is the specified inner knots, and p indicates the exponential order for truncated power basis.

1.1.5 Efficiency of Spline Smoothing Models for Estimating Time-Series Parameters

Spline smoothing, just like other smoothers, is an efficient solution provider to the minimization problem; it considers how the estimate performs on large observations.

Consider the model;

$$y_i = f(t_i) + \varepsilon_i, \qquad i = 1,2,\ldots,n, t_i \epsilon [0,1] \tag{1.19}$$

Where; $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)' \sim N(0, \sigma^2 W^{-1})$And σ^2 is unknown. Assume$f \varepsilon W_2^m$, where

$$W_2^m = \left\{ f: f^{(v)} absolutely\ continuous, y = 0, \ldots, m-1, \int_0^1 \left(f^m(t)\right)^2 dt < \infty \right\}$$

The spline smoothing $\hat{f}_\lambda$ refers to the minimizer of

$$minf\varepsilon W_2^m \left\{ \frac{1}{n}(Y-f)'W(y-f) + \lambda \int_0^1 \left(f^{(m)}(t)\right)^2 dt \right\}$$

Where;$y = (y_1, \ldots, y_n)'and\ f = \left(f(t_1), f(t_2), \ldots, f(t_n)\right)'$. The boundary λ controls the tradeoff between goodness-of-fit and the gauge's perfection and is frequently alluded to as the smoothing boundary, Yuedong (2011).

Let $\emptyset_v(t) \frac{t^{v-1}}{(v-1)!}, v = 1, \ldots, m$

$$Let\ R^1(s,t) = \int_0^1 (s-u)^{m-1} +$$

$\phi_v(t) = t^{v-1}/(v-1)!, v = 1, \ldots, m.$ Let $R^1(s,t) = \int_0^1 (s-u)_+^{m-1} + (t-u)_+^{m-1} du/((m-1)!)^2$, where $(x)_+ = x$ for $x \geq 0$ and $(x)_+ = 0$ otherwise.

Denote $T_{n\times m} = \{\phi_v(t_i)\}_{i=1}^{n}{}_{v=1}^{m}$ and $\sum_{n\times n} = \{R^1(t_i, t_j)\}_{i=1}^{n}{}_{j=1}^{n}$, Kimeldorf and Wahba (1971) showed that the solution has the form

$$\hat{f}(t) = \sum_{v=1}^{m} d_v \phi_v(t) + \sum_{i=1}^{n} c_i R^1(t, t_i) \tag{1.20}$$

Where $c = (c_1, \ldots, c_n)'$ and $d = (d_1, \ldots, d_m)'$ are solutions to

$$\begin{pmatrix} T'WT & T'W \\ \sum WT & \sum W + n\lambda I \end{pmatrix} \begin{pmatrix} d \\ \sum c \end{pmatrix} = \begin{pmatrix} T'Wy \\ \sum Wy \end{pmatrix} \tag{1.21}$$

Where T is given as;

$$T = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R \\ 0 \end{pmatrix}$$

1.1.6 Underfitting, Overfitting and Complexity of Model

Underfitting occurs when a statistical model cannot adequately capture the underlying structure of the data. An under-fitted model is a model where some parameters or terms that would appear in a correctly specified model are missing. Under-fitting would occur, for example, when fitting a linear model to non-linear data. Such a model will tend to have poor predictive performance. It occurs if the model or algorithm shows low variance but high bias (to contrast the opposite, overfitting from high variance and low bias). It is often a result of an excessively simple model which is not able to process the complexity of the problem. An underfitted model would ignore some important replicable (i.e., conceptually replicable in most other samples) structure in the data and thus fail to identify effects that were actually supported by the data. In this case, bias in the parameter estimators is often substantial, and the sampling variance is underestimated, both factors resulting in poor confidence interval coverage. Underfitted models tend to miss important treatment effects in experimental settings.

Overfitting is "the production of an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably. An overfitted model is a statistical model that contains more parameters than can be justified by the data. The essence of overfitting is to have unknowingly extracted some of the residual variation (i.e., the noise) as if that variation represented underlying model structure. The possibility of over-fitting exists because the criterion used for selecting the model is not the same as the criterion used to judge the suitability of a model. For example, a model might be selected by maximizing its performance on some set of training data, and yet its suitability might be determined by its ability to perform well on unseen data; then over-fitting occurs when a model begins to "memorize" training data rather than "learning" to generalize from a trend.

To lessen the chance or amount of overfitting, several techniques are available (e.g., model comparison, cross-validation, regularization, early stopping, pruning, Bayesian priors, or dropout). The basis of some techniques is either (1) to explicitly penalize overly complex models or (2) to test the model's ability to generalize by evaluating its performance on a set of data not used for training, which is assumed to approximate the typical unseen data that a model will encounter.

In model selection process, model complexity is very important, it is the number of features or terms included in a given predictive model, whether the model is linear or nonlinear. The best measure for the generalizing ability is the error of prediction of as many independent separate validation data as possible. According to Figure 1.1 the error of prediction is composed of two main contributions, the remaining interference error and the estimation error. The interference error is the systematic error (bias) due to unmodeled interference in the data, as the calibration model is not complex enough to capture all the interferences of the relationship between sensor responses and analytes. The estimation error is caused by modeling measured random noise of various kinds. The optimal prediction is obtained, when the remaining interference error and the estimation error balance each other (See arrow in figure 1.1). The effect of the prediction error increasing due to a too simple model is called underfitting whereas the effect of the increased prediction error due to a too complex model is called overfitting.

Model selection should be based not solely on goodness-of-fit, but must also consider model complexity, this is because a highly complex model can provide a good fit without necessarily bearing any interpretable relationship with the underlying process.

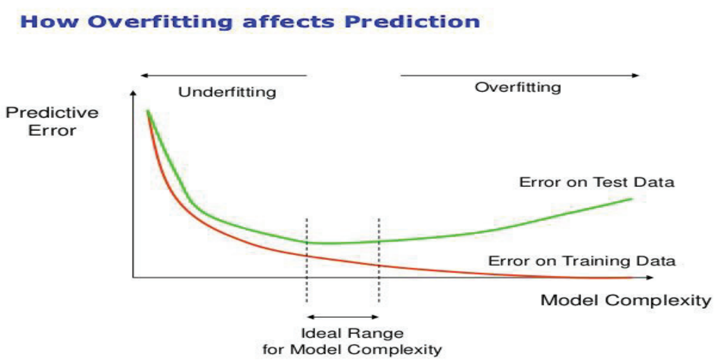


Figure 1.1: Relationship between Underfiting, Overfitting and Complexity model

1.2 Statement of the Problem

Over the last two decades, research on spline smoothing estimation methods has produced a vast amount of information and discoveries from researchers in evaluating the efficiency and performance of the existing estimation techniques when autocorrelation is present in their error terms. In this research work, a new smoothing method is proposed by taking the weighted sum of generalized cross-validation and unbiased risk. The rational behind the new estimation method is to provide a robust smoothing parameter estimation method that will alleviate the problem of overfitting models for time-series data with low, moderate and high autocorrelation level and the problem associated with the smoothing methods' performance when at different time series sample sizes.

1.3 Aim and objectives of the study

The aim of this study is to propose a smoothing method by taking the hybrid of two current spline smoothing methods, namely; the Unbiased Risk and Generalized Cross-Validation and compared its performance with three smoothing parameter selection methods, namely; Generalized Maximum Likelihood, Generalized Cross-Validation and Unbiased Risk.

The specific objectives of this study were to:

(i) determine the best-fit smoothing method for the time series observation;

(ii) identify the best smoothing method that does not over-fit time-series data when autocorrelation is present in the error term;

(iii) establish the optimum value of the proposed smoothing method;

(iv) compare using Predictive Mean Squared Error criterion, the three existing smoothing methods and the proposed smoothing method in terms of sample size; and

(v) test the performance of the proposed smoothing method using real life-data sets.

1.4 Justification of the Study

This research work was carried out to find out via Monte-Carlo simulation and real-life data the smoothing method that would be most efficient among the GCV, GML, UBR and the Proposed Smoothing Method (PSM) for measuring the degree of smoothness of a time series observation in presence of Autocorrelated error. This study shall be beneficial to the nonparametric regression and econometric researchers because it will serve as a guide in choosing the appropriate smoothing spline smoothing techniques for time-series data when autocorrelation is present in the error term. The issue of overfitting of the model caused by the existence of Autocorrelated error shall also be considered in this study.

1.5 Limitation of the study

Although many conditions were investigated in this dissertation, including real-life data and simulated data, three-time series sample sizes, three autocorrelation levels, four smoothing parameters etc., this study has some limitations. First, only the Predictive Mean Squared Error (PMSE) criterion of smoothed distributions was used to evaluate the smoothing results. Other performance criteria like the; Mean Absolute Deviation (MAD), Mean Absolute Bias (MAB), Bias, Variance, Average Squared Bias (ASB) or Mean Squared Error (MSE) and Average Squared Error (ASE) and should be considered for further research.

Second, this study evaluated the performance of different model selection strategies in spline smoothing in terms of the Predictive Mean Squared Error. Although some results were provided at the end of Chapter IV, more detailed results like the application of different degree of smoothness, standard deviation level, autocorrelation level and small sizes might lead to different conclusions about the different model supposing. All of the other estimators suggested by Dursun et al. (2013) for choosing smoothing parameters were also considered.

Third, in chapter three, the Proposed Smoothing Method was defined to consider time-series observations with autocorrelation error. The correlation structure W is approximately the same as that of the GCV estimation Method because the Proposed Smoothing Method is a hybrid of GCV and UBR, the behaviour of the PMSE is therefore slightly similar with GCV.

1.6 Definition of Relevant Terms

1. Spline: A Spline is a piece-wise polynomial with pieces defined by a sequence of knots such that the pieces join smoothly at the knots. The most common case is a linear Spline.

2. Spline Regression: This considers the problem of smoothing a scatterplot as opposed to interpolating. The approach here is to select a suitable set of knots with $k < n$ i.e. k is substantially less than n, the spline is fitted by ordinary least square. Data are fitted to a set of spline basis function with a reduced set of knots, typically by least squares, no roughness penalty is used.

3. Nonparametric Regression: is a type of regression analysis in which the predictor does not take a predetermined form but is constructed according to information derived from the data. Nonparametric regression requires larger sample sizes than regression based on parametric models because the data must supply the model structure as well as the model estimates.

4. B-Spline: Basic Spline is a function that has minimal support with respect to a given degree, smoothness and domain partition.

5. Penalized Spline: This is an intermediate estimate between regression and smoothing spline, it combines the reduces knots of regression,

6. Ridge Regression: is a method used for analyzing multiple regression data with the presence of multicollinearity. If multicollinearity occurs, least squares estimates are unbiased, but their variances are large so they may be far from the true value.

7. Kernel Regression: estimates a continuous dependent variable from a limited set of data by convolving the data points' locations with a kernel function, the kernel function specifies how to "blur" the influence of the data points so that their values can be used to predict the value for nearby locations.

8. **Predictive Mean Squared Error:** Mean squared prediction error of a smoothing or curve fitting procedure is the expected value of the square difference between the fitted values multiplied by the predictive function and values of the function.

9. Partial Spline Model: Partial smoothing splines are an important class of semi parametric regression models. In general, a partial smoothing spline model assumes the data (X_i, T_i, Y_i) given as;

$$y_i = x_i'\beta + g(t_i) + \varepsilon_i, 1, 2, \ldots, n, f \in W_m[0,1]$$

10. Autocorrelation: Autocorrelation is defined as a lag correlation of a series with itself lagged by a number of time units, Wolfgang (1994).

1.7 Organization of the Study

This thesis consists of five chapters and three appendices. Chapter one introduces the concept of Nonparametric regression model, cubic spline smoothing model, penalized spline smoothing model, non-parametric regression model and efficiency of spline smoothing models for estimating time series. The chapter also presented the problem, aim and objectives, justification of the study, study limitations, notations, and definition of relevant terminology. Chapter two reviews some relevant literature and theoretical framework on autocorrelation discussed under the following headings; the problems of autocorrelation, a typical structure of the error terms, a test of autocorrelation in a linear model, methods of the parameter of a linear model with autocorrelated error, prediction in the presence of autocorrelated, and Monte-Carlo ideas on autocorrelation. Chapter 3 addresses the theory and method used for the analysis, the Monte Carlo procedure, the model formulation, simulation result, value-generating equation in simulation, method of data generation, estimation methods used for the simulation study, namely; Generalized Cross-validation estimation method with autocorrelation structure, Generalized Maximum Likelihood estimation method with autocorrelation structure, Unbiased risks estimate method with autocorrelation and proposed smoothing method. The chapter also addressed the evaluation of the smoothing methods and sample of the simulated data. Chapter four presented the simulation results with autocorrelation in the error term and four estimation methods to real-life data. Chapter five summarizes the results obtained, the conclusion reached recommendations, contribution to knowledge and indicates possible directions for further research.

CHAPTER TWO

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.0 Introduction

This Chapter presents an empirical review of some literature related to the Generalized Cross-Validation, Generalized Maximum Likelihood and Unbiased Risk for correlational observations and autocorrelated error terms. The chapter also presents the theoretical background of Autocorrelation.

2.1 Review of Literature on GML, GCV AND UBR

Wahba (1985) used Smoothing Spline model to obtain a GML estimate for the smoothing parameter, then this estimate is compared with the GCV estimate both analytically and by Monte-Carlo method. The comparison was based on a predictive Mean Squared Error. It was discovered that GCV was somewhat better than GML for n = 64, GCV was decidedly superior for n = 128 while for n= 32, GCV was better for smaller σ^2 and the comparison close for larger σ^2.

Diggle and Hutchinson (1989) used the GCV criterion for choosing the degree of smoothing in spline regression and extended it to accommodate a time-series autocorrelated error sequence. It was demonstrated via simulation that the minimum GCV smoothing spline is an inconsistent estimator in the presence of autocorrelated error. Ignoring the moderate autocorrelation structure can seriously affect the performance of the Cross-Validation smoothing spline.

Altman (1990) used the mean squared error of kernel estimators to computed for processes with correlated errors, and the estimators were shown to be consistent when the sequence of error processes converges to a mixing sequence. The standard techniques for bandwidth selection, such as CV and GCV, are shown to perform very badly when the errors are correlated. Standard selection techniques are shown to favor under smoothing when the correlations are predominantly positive, and over smoothing when negative.

Kohn, Ansley and Wong (1992) estimated spline nonparametric regression with an unknown function observed with autocorrelated errors when the errors are modelled by an autoregressive Moving Average model. The unknown parameters were estimated by either maximum likelihood, cross-validation or generalized cross-validation. By expressing the problem in state space form O(n) algorithms were obtained to estimate the function of its derivatives and evaluate the marginal likelihood and cross-validation functions. The finite sample properties and the function estimates are evaliuated by extensive simulation study and examples were given.

Yuedong (1998) extended the GML, GCV and UBR to estimate the smoothing parameters and the correlation parameters simultaneously, when the correlation matrix is assumed to depend on a parsimonious set of parameters. GML method was recommended because it is stable and works well in all simulations. It performs better than other methods, especially when the sample size is small.

Yuedong *et al*., (2000) described three methods, GML, GCV and leaving-out-one-pair cross validation to estimate the smoothing parameters, the weighting parameter and the correlation parameter simultaneously. Based on simulated data, they concluded that the GML method has smaller Mean Squared Errors for the nonparametric functions and parameters and needs less computational time than the other methods and that it does not over fit data when the sample size is small.,

Carew *et al.,* (2003) studied smoothing splines with the degree of smoothing selected by GCV-Spline and provides a method to find an optimal smoother for an fMRI time series, to determine if GCV-Spline of fMRI time series yields unbiased variance estimates of linear regression model parameters. GCV-Spline was evaluated with a real fMRI data set and bias of the variance estimator was computed for simulated time series with autocorrelation structures derived from fMRI data. The results from the real data suggest that GCV-Spline determines appropriate amounts of smoothing. The simulations show that the variance estimates are, on average, unbiased. It demonstrates that GCV-Spline is an appropriate method for smoothing fMRI time series.

Hart and Lee (2005) described the effects of moderate levels of serial correlation on one-sided and ordinary cross-validation in the context of local linear and kernel smoothing is investigated. It is shown both theoretically and by simulation that one-sided cross-validation is much less adversely affected by correlation than in ordinary cross-validation. The former method is a reliable means of window width selection in the presence of moderate levels of serial correlation, while the latter is not. It is also shown that ordinary cross-validation is less robust to correlation.

Anna, Tong, and Yuedong (2006) considered spline smoothing of variance functions, the focus was on the selection of the smoothing parameters and three direct data-driven methods. UBR, GACV and GML were developed. In addition to guaranteed convergence, simulations show that these direct methods perform better than existing indirect UBR, GCV and GML methods. The direct UBR and GML methods perform better than the GACV method. An application to array-based comparative genomic hybridization data illustrates the usefulness of the proposed methods.

Krivobokova and Kauermann (2007) investigated the behavior of data-driven smoothing parameters, for penalized spline regression in the presence of correlated data. It was shown for other smoothing methods that mean squared error minimizers, such as GCV or the Akaike information criterion, are extremely sensitive to misspecifications of the correlation structure resulting in over- or under-fitting the data.

Yanrong, *et al.* (2010) described and studied different methods of choosing important smoothing parameter estimators. GCV was recommended as the best choice for penalized Spline Smoothing parameter selection for both computational efficiency and accuracy of the functional coefficient regression models.

Aydin and Memmedli, (2011) made a comparison between GCV and REML, it was recommended that GCV and REML are good smoothing parameter selection for small and medium sized samples.

Chen and Huang (2011) applied the smoothing spline method to fit a curve to a noisy data set, where the selection of the smoothing parameter is essential. An improved C_p criterion for spline smoothing based on Stein's unbiased risk estimate has been proposed to select the smoothing parameter. The resulting fitted curve has been shown to be superior and more stable than commonly used selection criteria and possesses the same asymptotic optimality as C_p.

Aydin, *et al.* (2013) applied most of the data-driven smoothing parameter selection methods and compared based on large and small samples sizes. The parallel of Akaike's information criterion and Generalized Cross-Validation are recommended as being the best selection criteria. For large samples the GF_{AIC} method would seem to be more appropriate while for small samples they proposed the implementation of GCV criterion.

Jansen (2015) investigates two types of results that support the use of GCV for variable selection under the assumption of sparsity. The first type of result is based on the well-established links between GCV on the one hand and Mallows's Cp and Stein Unbiased Risk Estimator on the other hand. The result states that GCV performs as well as Cp or SURE in a regularized or penalized least squares problem as an estimator of the prediction error for the penalty in the neighborhood of its optimal value.

Lukas *et al.* (2016) investigated the behavior of the optimal values of gamma and rho to identify simple practical rules to choose their optimal properties. RGCV and modified GCV perform significantly better than GCV. The performance is defined in terms of the Sobolev error, which is shown by example to be more consistent with a visual assessment of the fit than the average squared error.

Devi *et al*. (2018) discusses UBR and GCV for selecting the optimal knots in spline. The criteria for selecting the best model were based on Mean Squared Error and R-square. The simulation was performed on a spline truncated function with error generated from a Normal distribution for varied sample sizes and error variance. The results of the simulation study showed that GCV estimates the knots more accurately than UBR.

Xu and Zhou (2019) considered nonparametric regression problems and developed a model-averaging procedure for smoothing spline regression problems. Model weights were estimated using a delete-one-out cross-validation procedure to minimize the prediction error. A simulation study was performed by using a program written in R. The simulation study provides a comparison of the most well-known CV, generalized GCV, and the proposed method. The model averaging approach is straightforward to implement, and gives reliable performances in simulations.

Table 2.1.1: Summary of the Reviewed Literature and Performance of the Existing Methods in the Presence of Autocorrelation in the error term

Author	Methods	Weak method	Preferred method	Conclusion
Wahba (1985)	GCV and GML	GML	GCV performed very well	GCV was recommended for all sample size
Diggle and Hutchinson (1989)	Extension of GCV, CV and GCV	CV and GCV	Extension of GCV	Extension of GCV outperformed CV and GCV
Altman (1990)	The proposed method, CV and GCV	CV and GCV	Proposed method	GCV performed badly with autocorrelated error
Kohn et al. (1992)	GML and GCV	GML	GCV	GCV out performed other method
Yuedong (1998)	GML, GCV and UBR	GCV and UBR	GML	GML performs well in a small sample
Yuedong (2000)	CV, GCV and GML	CV and GCV	GML	GML do not over-fit for a small sample size
Carew (2003)	Optimal smoother and GCV	Optimal smoother	GCV	GCV appropriate for fMRI time series
Anna et al. (2006)	GACV, GCV, UBR and GML	GCV, GML and UBR	GACV	GACV smoothing method worked well
Krivobokoba and Kauerman (2007)	Other Smoothing methods, GCV, AIC		GCV and AIC criterion	Other Smoothing methods overfit or underfit data
Yanrong et al. (2010)	CV, GCV, GML and UBR	CV, GML and UBR	GCV	GCV best for the practicable regression coefficient model
Ayin and Memmedli (2011)	GCV and REML	None	GCV and REML	GCV and REML works well for small and medium sample
Ayin and Memmedli (2013)	Nine smoothing methods	Seven smoothing method	GFAIC and GCV	GCV and GFAIC works well for small and large sample
Jansen (2015)	GCV, UBR and SURE	UBR and SURE	GCV	GCV proceeds just as for UBR or SURE is a regularized or punished least-squares issue
Lukas *et al.* (2016)	RGCV, modified GCV and GCV	RGCV and modified GCV	GCV	Performance of GCV is more consistent
Devi et al. (2018)	UBR and GCV	UBR	GCV	GCV estimates the knots more accurately than UBR.
Xu and Zhou (2019)	(CV), (GCV), and the a model-averaging technique	CV and GCV	The model-averaging technique	The new technique is straightforward to execute and gives a solid performance in simulations.

The summary of the reviewed Literature and performance of the three classical smoothing methods in the presence of autocorrelation in the error term shows that GML over fit and over-smooth data for the sample, medium and large sample size. Yuedong, Guo and Brown, (2000), It was also observed that GCV is the most preferred spline smoothing estimation method for data with autocorrelation in the error term.

2.2 THEORETICAL FRAMEWORK

2.2.1 Autocorrelation

If the error terms are correlated in sequential order, then we have autocorrelation. Successive residual in economic time series tends to be decidedly related, Chattterjee et al. (2000). In trials, correlated perceptions might be due to an idea of the plots, the format of some total impacts through periods, bug contaminations around the surrounding plots, or a collection of nearby factors that are hindering them, (Berenlut and Web, 1974), (Williams, 1952), (Papadakis, 1937).

Autocorrelation can arise as a result of:

- Missing explanatory variables
- Incorrect specification of the numerical character of the model
- Interpolation in the statistical data
- Incorrect specification of the real random error Johnston, (1984).

The most straightforward type of the traditional linear regression model with autocorrelation disturbance is accepted to follow the first order Autoregressive (AR(1)) measure given as

$$y_t = A_0 + A_1 x_{1t} + u_t \quad (2.1)$$

Where;

$u_t = \rho u_{t-1} + \varepsilon_t |\rho| < 1$ t = 1,2,⋯, n $\quad\quad \varepsilon_t$ = N$(0, \sigma^2 I_n)$

It can be shown that $u_t = \left(0, \frac{\sigma_\varepsilon^2}{1-\rho^2}\right)$ and that $E(u_t u_{t-s}) = \rho^s \sigma_u^2$ (see equations 2.2 – 2.12) for the proof.

The outcome of using OLS estimator to model, according to (Johnston, 1984) and (Fomby et al., 1984), include;

1. The ordinary least square (OLS) estimator $\hat{\beta}_{(OLS)} = (X'X)^{-1}X'Y$ remains unbiased and consistent.
2. The variance-covariance of $\hat{\beta}$ is biased. The true variances and standard errors are underestimated, and the t and F tests are no more reliable.
3. The variances of the error term may also be seriously underestimated (biased). Thus, R^2 also becomes unreliable. Durbin and Watson (1950, 1951 and 1971) and others discussed this Autocorrelation problem extensively. Some of the properties of Autocorrelation include;

2.2.2 The Problems of Autocorrelation

The assumption of spherical disturbance of the classical regression model, namely $E(UU')=\sigma^2 I_n$ is in two folds. Firstly, the disturbance covariance at all pairs of observation points is zero. When the first holds, we have homoscedastic error terms; otherwise, we say the error terms are homoscedastic. When the second hold, we refer to the situation as a lack of serial association of Autocorrelation of the disturbance. But if the error terms of any particular period are correlated with any other error in the series, we have Autocorrelation. Thus, some, if not all, of the off-diagonal elements are nonzero.

There are several compositions of Autocorrelation in time series utilization; among the explained ones are Moving Average (MA), Autoregressive (AR) processes and the Autoregressive Moving Average (ARMA) processes.

2.2.3 Typical Structures of the Error Terms

The structure of the error terms determines the form of the CLRM with Autocorrelated error terms.

1. Autoregressive process of order 1, (AR(1)): if the significance of the disturbance value (U) in any particular time depends on its importance in the preceding period alone, the error term U is said to have a first-order autoregressive or Markov scheme, given as;

$$U_t = \rho U_{t-1} + \varepsilon_t \qquad |\rho| < 1 \qquad t = 1, 2, \cdots, n \qquad \varepsilon_t \sim N(0, \sigma^2 I_n) \tag{2.2}$$

We now establish the mean, the variance and the covariance of this Autocorrelated disturbance variable.

At period t, $U_t = \rho U_{t-1} + \varepsilon_t$ (2.3)

If we continue to perform continuous substitution of the lagged value of U, then, at period t – 1

$$U_{t-1} = \rho U_{t-2} + \varepsilon_{t-1} \tag{2.4}$$

Substituting (2.2) into (2.4)

$$U_t = \rho^2 U_{t-2} + (\rho\varepsilon_{t-1} + \varepsilon_t)$$

$$= \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2 U_{t-2} \tag{2.5}$$

At period t – 2 $U_{t-2} = \rho U_{t-3} + \varepsilon_{t-2}$ (2.6)

Substituting (2.6) into (2.5),

$$U_t = \rho^3 U_{t-3} + \rho^2 U_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t \tag{2.7}$$

If this is continued for r periods (t – r), then

$$U_t = \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2 U_{t-2} + \rho^3 U_{t-3} + \cdots + \rho^r U_{t-r} \tag{2.8}$$

Now if $r \to \infty$ the term with lagged U, $\rho^r U_{t-r} \to 0$ since $|\rho| < 1$. Thus,

$$U_t = \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2 U_{t-2} + \rho^3 U_{t-3} + \cdots = \sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r} \tag{2.9}$$

Hence,

$$E(U_t) = E(\varepsilon_t) + \rho E(\varepsilon_{t-1}) + \rho^2 E(\varepsilon_{t-2}) + \rho^3 E(\varepsilon_{t-3}) + \cdots = \sum_{r=0}^{\infty} \rho^r E(\varepsilon_{t-r}) = 0 \tag{2.10}$$

$$\mathrm{Var}(U_t) = \sigma_u^2 = E(U_t^2) - (E(U_t))^2 = E(U_t^2)$$

$$= E[\sum_{r=0}^{\infty} \rho^{r} \varepsilon_{t-r}]^2$$

$$= [\sum_{r=0}^{\infty} \rho^{2r} E(\varepsilon_{t-r})^2]$$

$$=[\sum_{r=0}^{\infty} \rho^{2r} Var(\varepsilon_{t-r})]$$

$$= \sum_{r=0}^{\infty} \rho^{2r} \sigma_{\varepsilon}^2 = \frac{\sigma_{\varepsilon}^2}{(1-\rho^2)} \qquad (2.11)$$

Therefore, despite the serial correlation (autocorrelation), the error term's variance is still homoscedastic (Chattterjee *et al*., 2000).

We now obtain the covariance of the Autocorrelated error term U as follows:

$$\text{Cov}(U_t U_{t-1}) = E(U_t U_{t-1}) - E(U_t)E(U_{t-1}) = E(U_t U_{t-1})$$

$$= E[(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \cdots)(\varepsilon_t + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \cdots)]$$

$$= E[\varepsilon_t + \rho(\varepsilon_{t-1})^2 + \rho^3(\varepsilon_{t-2}) + \cdots + cross\ product]$$

$$= \rho[E(\varepsilon_t) + E(\varepsilon_{t-1})^2 + \rho^2 E(\varepsilon_{t-2})^2 + \cdots + E(cross\ product)]$$

$$= \rho[0 + \text{Var}(\varepsilon_{t-1}) + \rho^2 Var(\varepsilon_{t-2}) + \cdots + 0]$$

$$= \rho\sigma_{\varepsilon}^2[1 + \rho^2 + \rho^4 + \rho^6 + \cdots]$$

$$= \frac{\rho\sigma_{\varepsilon}^2}{(1-\rho^2)} = \rho\sigma_u^2 \qquad (2.12)$$

Similarly,

$$\text{Cov}(U_t U_{t-2}) = E(U_t U_{t-2}) = \rho^2\sigma_u^2 \qquad (2.13)$$

In general,

$$\text{Cov}(U_t U_{t-s}) = E(U_t U_{t-s}) = \rho^s\sigma_u^2 \qquad \text{(for } s \neq t\text{)} \qquad (2.14)$$

Thus variance-covariance matrix becomes

$$E(UU') = \sigma_u^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \rho \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots\ \rho & 1 \end{bmatrix} \qquad (2.15)$$

2. Autoregressive process of order 2, AR (2): Here, the disturbance U relies on the value of error terms of the last preceding periods alone. The relationship is defined as

$U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + \varepsilon_t; \qquad |\rho_2| < 1$

$$\rho_2 + \rho_1 < 1, \rho_2 - \rho_1 < 1 \qquad t = 1, 2, \cdots, n \qquad \varepsilon_t \sim N(0, \sigma^2 I_n) \tag{2.16}$$

The variance of the error terms U for this process is

$$\sigma_u^2 = \frac{(1-\rho_2)\sigma_\varepsilon^2}{(1-\rho_2)[(1-\rho_2)^2-\rho_1^2]} \tag{2.17}$$

In general, the autoregressive of order P, i.e. AR(P), can be defined as

$$U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + \cdots + \rho_p U_{t-p} + \varepsilon_t \quad t = 1, 2, \cdots, n \quad \varepsilon_t \sim N(0, \sigma^2 I_n) \tag{2.18}$$

3. The Average Moving Process of order q, i.e. MA(q): The general MA(q) is defined as

$$U_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \cdots + \phi_p \varepsilon_{t-q} t = 1, 2, \cdots, n \quad \varepsilon_t \sim N(0, \sigma^2 I_n) \tag{2.19}$$

where ϕ is a constant less than unity. For the process, the autocorrelation coefficients are zero for all lags greater than the moving average process's order.

4. The Autoregressive, Moving Average Processes of order p and q (ARMA(p,q)):

The general autoregressive moving average ARMA (p,q) process is defined as

$$U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + \cdots + \rho_p U_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \cdots + \phi_p \varepsilon_{t-q} t = 1, 2, \cdots, n \qquad \varepsilon_t \sim N(0, \sigma^2 I_n) \tag{2.20}$$

Durbin and Watson (1971).

This formulation enables complicated processes to be represented in a suitable form of low–order polynomial.

Handling each of these error terms mentioned above requires specific approaches at the point when they happen in a direct model. Consequently, each requires various techniques for assessment and theory testing. However, efforts have been made to give a general approach to the treatment of this auto regression.

2.2.4 Test for Autocorrelation in Non-Linear Models

The importance of detecting the presence of autocorrelation in models has been widely recognized. Thus, several test techniques have been established within the last two decades, most of which have asymptotic properties. For some of these methods, their performances in finite samples are still being investigated.

Different structures of autocorrelation have given rise to various tests. However, in a reality where the mode of the error structure is not known, the problem of which one to use arises since the investigator may want to test for autocorrelation's existence without specifying the form it should take. Some tests have a generalized approach to the treatment of autocorrelation. Most of these tests require heavy computational demands and tremendous computing skills even though they may compete favourably with simpler ones and, in some cases, may perform better.

2.2.5 Methods of Parameter Estimation of Non-Linear Model with Autocorrelated Errors

The GLS and the OLS methods are the two methods that can be applied to assess the variables of the linear model when autocorrelation is present in the error term. Since the latter suffers efficiency, the former is used to improve this efficiency. However, Chipman (1979), Kramer (1980), Kleiber (2001) have discovered that the effect of OLS estimator in a linear regression having an autocorrelated disturbance depends largely on the configuration of X used. The GLS method usually requires that Ω and in specific ρ be established before the variable can be derived Formby et al. (1984). Thus, in a linear model with an autocorrelated error term

$$\hat{\beta}_{(GLS)} = (X^1\Omega^{-1}X)^{-1}X^1\Omega^{-1}Y \quad (2.21)$$

$$V(\hat{\beta}_{(GLS)}) = \sigma^2(X^1\Omega^{-1}X)^{-1} \quad (2.22)$$

Where

$$E(UU') = \sigma^2\Omega = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-2} & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-3} & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-4} & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho^{n-2} & \rho^{n-3} & \rho^{n-4} & \cdots & 1 & \rho \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & \rho & 1 \end{bmatrix}$$

And $\sigma^2 = \sigma_u^2 = \frac{\sigma_\varepsilon^2}{(1-\rho^2)}$,

And the inverse of Ω is

$$\Omega^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}$$

We now search for a suitable transformation matrix P^* as discussed above

If we consider an (n – 1) x n matrix P^* defined by

$$P^* = \begin{bmatrix} -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ 0 & 0 & -\rho & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}_{(n-1) \times n} \tag{2.23}$$

Multiplying then shows that $P^{*1}P^*$ gives a (n x n) matrix which separated from a corresponding consistent is indistinguishable with Ω^{-1} except for the primary components in the main slanting, which is ρ^2 rather than solidarity.

Now, if we consider an n x n matrix P derived from P^* through the addition of a new row to the preceding row with $\sqrt{1-\rho^2}$ amid principal situation and zero somewhere else, it is written as;

$$P = \begin{bmatrix} (1-\rho^2)^{\frac{1}{2}} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}_{(n \times n)} \quad (2.24)$$

Multiplying shows that P^1P=$(1-\rho^2)\Omega^{-1}$. The contrast among $P^*$and P; lies just in the treatment of the main example perception $P^*$is it preferable provided it loses information on the first observation? However, when n is large, the difference can be ignored, but the difference is very important in a small sample.

On the off chance that Ω or even more ρ is known, the GLS evaluation could be cultivated by applying the OLS through the change grid P* And P above. Regardless, this isn't now and again the circumstance; we resort to surveying Ω by Ω to have a potential Generalized Least Squares Estimator. This assessor becomes conceivable when ρ is replaced by an anticipated assessor ρ (Formby et al. 1984). There are a couple of various methods of dependably evaluating ρ, in any case, some of them which use either P* or P change lattice joins:

1.$\Rightarrow y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \rho u_{t-1} + \varepsilon_t$

2. $\Rightarrow \rho y_{t-1} = \rho\beta_0 + \rho\beta_1 x_{1,t-1} + \rho\beta_2 x_{2,t-1} + \cdots + \rho^2 u_{t-2} + \rho\varepsilon_{t-1}$

3. $\Rightarrow y_t - \rho y_{t-1} = \beta_0 - \rho\beta_0 + \beta_1 x_{1,t} - \rho\beta_1 x_{1,t-1} + \beta_2 x_{2,t} - \rho\beta_2 x_{2,t-1} + \cdots + \beta_k x_{k,t} - \rho\beta_k x_{k,t-1} + \rho u_{t-1} - \rho^2 u_{t-2} + \varepsilon_t - \rho\varepsilon_{t-1}\beta_0(1-\rho) + \beta_1(x_{1,t} - \rho x_{1,t}) + \beta_2(x_{2,t} - \rho x_{2,t-1}) + \cdots + \beta_k(x_{k,t} - x_{k,t-1}) + \varepsilon_t$

$$y_t^* = \beta_0^* + \beta_1 x_1^* + \beta_2 x_2^* + \cdots + \beta_k x_k^* + \varepsilon_t \quad (2.25)$$

Where

$$y_t^* = (y_t - \rho y_{t-1}), \beta_0^* = \beta_0(1-\rho), x_1^* = (x_{1,t} - \rho x_{1,t-1})$$

$x_2^* = (x_{2,t} - \rho x_{2,t-1}), x_k = (x_{k,t} - \rho x_{k,t-1}) t = 2,3, \cdots, n$

and in the matrix record;

$$P^{*} = P^{*}X\beta + P^{*}U$$

Y, X and U represent the original dependent variable, design matrix, and error vector observations.

The Cochran and Orcutt Regression (CORC) procedure is as follows;

- Start with the linear model in the presence of first-order Autocorrelated residual
- Lag the model in (1) and multiply it by ρ to obtain a new model.
- Subtract the new model in (2) from the model in (1).

  Applying these to the linear model (2 and 1)

The three statements above look like the usual multiple regression model. We only need to construct the variables that are asterisked (*) and have a value for ρ. Since ρ is unknown, Cochran and Orcutt suggested an iterative procedure for estimating β. The process is as follow:

1. Run the regression for the original model
2. Estimate the residual $\hat{u} = y_t - \hat{y}_t$ and run the regression model

   $\hat{u}_t = \rho u_{t-1} + \varepsilon_t$

   A consistent estimator of ρ is given by $\rho = \frac{\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{n} \hat{u}_{t-1}^2}$
3. The estimate of ρ obtained from 2 above is now used to construct the asterisked variables, and feasible generalized least square estimator β is obtained. From this OLS regression, the residual sum of a square is noted.
4. With the latter set of residuals, the regression model in 2 is run to obtain a new estimate of ρ which is then used to construct another stage of asterisked variables. Another feasible generalized least square estimator β is obtained. The residual sum of a square is again noted.

2.2.6 Prediction in the Presence of Autocorrelation

Parameter estimation of the regression model and prediction of the dependent variable, among many other uses, remain the two major uses of regression analysis.

Examine the linear model with AR (1) of the form

$$y_t = \beta_0 + \beta_1 X_t + U_t U_t$$

$$U_t = \rho U_{t-1} + \varepsilon_t \quad (2.26)$$

If it is desired to predict y_t at period $t+1$ given the regressor X_{t+1}, the predictor equation in (2.26) is given by

$$\hat{y}_{t+1} = \hat{\beta}_0 + \hat{\beta}_1 X_{t+1} \quad (2.27)$$

Where; $\hat{\beta}_0$ and $\hat{\beta}_1$ are estimates obtained by the OLS method. The predictor sets the disturbance term at zero, and the AR(1) process implies that; $E(X_{t+1}) = \rho U_t$, therefore, the two unknown parameters ρU_t can be estimated by

$$\hat{\rho}\hat{U}_t = \hat{\rho}(\hat{Y} - \hat{\beta}_0 - \hat{\beta}_1 \hat{X}_t) \quad (2.28)$$

Consequently, the appropriate model for predicting the y_t variable becomes

$$\hat{Y}_{t+1} = \hat{\beta}_0 + \hat{\beta}_1 X_{t+1} + \hat{\rho}\hat{U}_t \quad (2.29)$$

2.2.7 Monte Carlo Ideas on Autocorrelation

All the feasible generalized least squares estimators discussed in the earlier section for parameter estimation of the linear model with autocorrelated errors (AR(1)) are known to be asymptotically equivalent. However, the relative performance of these estimators in small samples is another matter Formby *et al.* (1984). Thus studying the finite sample properties of these estimators becomes very imperative. This seems to be very difficult analytically. However, a Monte Carlo approach is often utilized to accomplish this task.

The attention of the economist was first drawn by Cochrane and Orcutt (1949) to the fact the presence of autocorrelated error terms requires some modification of the OLS method. Their suggestion involved an autoregressive transformation of the series involved. They suggested that the quasi first differences of such series should be used. Also, they emphasized that many current formulations of economic relations are highly positively autocorrelated.

Kadiyala (1968) observed that the transformation suggested by Cochrane and Orcutt (1949) could lead to a less efficient estimator. He, therefore, suggested that the addition of one weighted observation to COCR procedure may give a better estimator practically no extra cost.

Rao and Griliches (1969) did one of the earlier Monte Carlo investigations on this study. The model used to generate the observation for the sampling experiment was given as;

$$Y_t = \beta X_t + U_t \qquad (2.30)$$

$$X_t = \lambda X_{t-1} + V_t |\lambda| < 1,$$

$$U_t = \rho U_{t-1} + \varepsilon_t |\rho| < 1, \qquad t = 1,2,\cdots,20.$$

$$E(V_t) = E(\varepsilon_t) = E(V_t V_{t-1}) = E(\varepsilon_t \varepsilon_{t-1}) = E(V_t \varepsilon_t) = 0$$

$$E(V_t^2) = \sigma_V^2, E(\varepsilon_t^2) = \sigma_\varepsilon^2$$

For any fixed values of ρ and, the experiment replicated 50-time intervals with independent $X_t$andε_t. The estimators they examined include OLS, COCR, Durbin ρ and Prais – Winstern. Their major conclusions were:

i. The ordinary least squares assessor have reduced productivity than any remaining technique computed for average and high estimations of ρ ($|\rho| \geq 0.3$).

ii. There exist an unmistakable addition from utilizing plausible summed up least squares when ρ ($|\rho| \geq 0.3$) and little misfortune from utilizing such strategies in any case.

iii. The entirety of the achievable summed up least squares strategies was not very separated from their presentation.

iv. Among the various assessors analyzed, the two-venture assessor dependent on Prais – Winstern change and the Durbin ρ was suggested in as much as the creator discovers that it is probably going to go above and beyond a great scope of boundaries than different techniques examined.

Paris and Winstein (1954) performed better in the experiment because the error terms are produced utilizing the assumption that the stochastic interaction started in the endless past.

Kramer (1980) examined Rao and Griliches (1969) work and noted some contradictions in their findings which attributed to the stochastic (random) matrix X considered in their experiment. It was said that Rao and Griliches (1969) confined their investigation to only one exogenous variable, but it is not clearly shown in their presentation whether they are estimating $\beta_1 X_t$ in the model

$$Y_t = \beta_0 + \beta_1 X_t + U_t \qquad \text{or} \qquad Y_t = \beta_1 X_t + U_t \qquad (2.31)$$

Where X_t and U_t in both cases are identical.

The effect of COCR's transformation on the estimators' efficiency when the independent variable exhibits various degrees of trend has also been studied by Maeshiro (1976).

The model considered is of the form

$$\text{Y} = \beta_0 + \beta_1 X_t + U_t \qquad \text{or} \qquad U_t = \rho U_{t-1} + \varepsilon_t \qquad (2.32)$$

Where $|\rho| < 1, E(\varepsilon_t) = E(\varepsilon_t \varepsilon_s) = 0, V(\varepsilon_t^2) = \sigma^2 I_n$ for $t \neq s$ and X_t are fixed under repeated sampling. He considered five specifications of independent variables and made an effort to investigate the estimators' asymptotic performance by considering the study using the sample size of 10, 20, 50 and 100. His results agree with that of Kadiyala (1968), except that the research did not deal with trended independent variables. Since COCR's method reduces efficiency in many cases, his study suggests that their approach should be used when an independent variable contains a reasonable trend. The sample size is not too large. However, the research could not answer the magnitudes of the efficiency of GLS over OLS and how much of the volume can be retained when the COCR is replaced by the Paris and Winstein (1954) estimator.

The work of Rao and Griliches (1969) was revisited by Spitzer (1979), but the research considered ML estimator's performance in addition to the estimators. Some of the results obtained did not agree with that of Rao and Griliches (1969). For example, unlike Rao and Griliches (1969), the ML and the non – linear estimator are superior to the two-stage methods of Durbin and Watson (1951) and Paris and Winstein (1954). Spitzer (1979) attributed this discrepancy to the fact that the ML estimator considered by Rao and Griliches (1969) is inefficient in a small sample. The jacobian term is ignored; its inclusion and n observation appear to make a significant difference.

Beach and Mackinnon (1978) derived a "Full" ML assessor which integrates the term 1 - ρ^2 in the likelihood function. They also derived an iterative procedure for the Time Series Processor (TSP) codes (White, 1978). They had a simulation study of the ML assessor through the CORC procedure using the model.

$$Y = \beta_0 + \beta_1 X_t + U_t U_t = \rho U_{t-1} + \varepsilon_t \quad (2.33)$$

Where $\varepsilon_t \sim N(0, 0.0036)$ and the explanatory variable X_t was selected to hold a full trend element as;

$$X_t = e^{(0.04t)} + w_t \text{where} w_t \sim N(0, 0.0009).$$

The sample sizes were changed from 20 to 50 in 200 during the replication process.

The discovery from the Root Mean Squared Error (RMSE) indicated that the ML estimator is very much better when compared with COCR in estimating β_1 and the difference is very remarkable for β_0 when the X's are trended. It has also been considered that a non – trending X where $X_t \sim N(0, 0.0625)$ using a sample size of 20. Their results showed that there is no virtual difference between the full information ML estimates and CORC estimates of β_1 but the full ML procedure provides dramatically better estimates of β_0. They also reported that the full ML procedure might be computationally less expensive than the CORC technique. Even though one iteration of the complete ML procedure costs slightly more than one iteration of CORC, it still requires less iteration than CORC.

Some further experiments conducted by Harvey and Mcavinchey (1978) compared the full ML procedure not just with the two-step CORC process, which was the comparison in the Beach and Mackinnon (1978) study, but also with the iteration CORC and the two-step Paris and Winstein (1954). procedure. They did not only confirm the results of Beach and Mackinnon (1978) but also brought out some important points, namely;

(i) The two-step Paris and Winstein (1954) technique is as efficient as full ML estimation for their experiments' parameter values.

(ii) The iterative CORC process is sometimes inferior to two-step CORC, especially when the explanatory variable is a time trend.

(iii) The two-step CORC process is, in turn, inferior to the two-step Paris and Winstein (1954) technique when the explanatory variable is trended.

(iv) OLS has RMSE only about 3 to 4 per cent above full ML estimation with trending data, but its relative performance deteriorates when X is a stationary random series.

Park and Mitchell (1980) provided another comparison of the estimation efficiency by considering the model

$$Y_t = X_t\beta + U_t U_t = \rho U_{t-1} + \varepsilon_t \qquad (2.34)$$

$$|\rho| < 1, \qquad t = 1,2,\cdots,n$$

$$E(\varepsilon_t) = 0, \qquad E(\varepsilon_t \varepsilon_t^1) = \sigma^2 I_n$$

With three explanatory variables, namely:

(a) one unreal series $x_t = [1, t]$

(b) one authentic trended series $x_t = [1, GNP_t]$ i.e. the yearly American GNP in Fixed currency (dollars) commenced in the year 1950

(c) an authentic untended series $x_t = [1, GNP_t]$ i.e. the yearly American GNP in Fixed currency (dollars) commenced in the year 1950.

Maeshiro (1976) applied (a) and (b), in place of (c), where quarterly quantity application commencing from 1948.

They classified the estimators according to

(i) Whether P^* or P (i.e. $(m - 1) \times m$ matrix of transformation observations.

(ii) Whether ρ is known or unknown

(iii) Whether the independent variable is trended or not.

Their results did confirm not only the findings of Beach and Mackinnon (1978) but also added the following

(i) Their results show that CORC has low performance, which is always smaller than the ordinary least square even if ρ is known or unknown.

(ii) With unknown ρ, the Prais – winstern and ML dominates the ordinary least square (OLS) and sometimes possesses greater consistency than the two-step estimators.

(iii) With ρ estimated from least-squares residuals, an iterative version of Prais – Winstern provides the best feasible estimators.

(iv) In terms of hypothesis testing, they observed that all the methods underestimate the standard errors, making the estimated coefficients emerge more significant than they are.

Nwabueze (2000) used nine different specifications of the independent variable on one – independent variable linear model with AR (1). Some of these specifications have been used in one way or the other by several authors, including Maeshiro (1976), Spitzer (1979) and Park and Mitchell (1980).

Olaomi (2004) also attempted to investigate some of these estimators' performance for a relationship between the explanatory variables and the disturbances that are significant at some levels of significance(α).

CHAPTER THREE

3.0 METHODOLOGY

3.1 Introduction

This chapter discussed the methodology applied in this research work. Section 3.2 presented Cubic Smoothing Spline Regression Model, Generalized Cross-Validation assessment technique with autocorrelation, Generalized Maximum Likelihood method of estimation with autocorrelated error, Unbiased Risk smoothing parameter selection method with autocorrelated error, Proposed Smoothing Method with autocorrelated error and proposed smoothing method Algorithm. The Monte-Carlo simulation method was presented in section 3.3, the equation used for simulation was given in section 3.4, followed by experimental design and observation generation. In section 3.6, smoothing spline assessment methods used in the study were presented with sub-headings like; and evaluation of the smoothing methods was also discussed under sub-headings like; predictive mean squared error, test for over-fitting in spline smoothing and test for goodness-of-fit smoothing methods were presented in section 3.7.

3.2 Cubic Smoothing Spline Regression Model

The most common example of the smoothing spline is the cubic spline; it is the smoothing spline's functional form and a piecewise cubic function that interpolates the dataset and ensures the smoothness of the observation. It is piecewise third-request polynomials that go through a bunch of focuses. It has a nonstop first and second subordinate with the request for the coherence of (d–1), where d is the polynomial degree. The Model with shortened force premise work b(x) changes the factors Xi by applying a premise work b(x) and fit a model utilizing these changed factors, which adds non-linearity to the model and empowers the splines to fit smoother and adaptable Non-straight capacities. The spline smoothing model is written as;

$y_i = f(t_i) + \varepsilon_i$ (3.1)

Where; y_i is the response variable, f is an unknown smoothing function, t_i is the independent/predictor variable and ε_i is zero mean autocorrelated stationary process.

The general cubic spline function is given as;

$$f(t) = at^3 + bt^2 + ct + d + \varepsilon \quad (3.2)$$

Where; a, b, c, and d = real number coefficients and $a \neq 0$, t = independent variable, ε = error term and d.f. = k-d-1 (k = number of knots and d = degree of cubic spline)

The cubic spline smoothing estimate function is $\hat{f}$ while; f refers to be the minimizer of a twice differentiable function of;

$$S(f) = \sum_{i-1}^{n} \left(y_i - \hat{f}(t_i)\right)^2 + \lambda \int_a^b \left(\hat{f}''(t)\right)^2 dt \quad (3.3)$$

Where;

- $\lambda > 0$ is a smoothing parameter,
- The initial part in equation (3.3) refers to the residual sum of the square for the integrity of the information's attack.
- The roughness penalty in the subsequent term of the equation (3.3), which is enormous when the incorporated second subsidiary of a regression function $f''(t)$ is likewise huge
- If λ moves toward 0, then $f(t)$ only interpolates the data set.
- If λ is very big, then $f(t)$ would be chosen wherefore $f''(t)$ is wherever 0, which will suggest a by and large direct least-squares fit the perceptions.

If $f(t)$ values are fixed at $f(t_1), \ldots, f(t_2)$ the roughness $\int_a^b \left(\hat{f}''(t)\right)^2 dt$ is minimized by a natural cubic spline, this solution is written as a basic function as;

$$f(t) = \beta_0 + \beta_1 f_1(t) + \ldots + \beta_{n+3} f_{n+3}(t)$$

3.2.1 Generalized Cross-Validation (GCV) Estimation Method with and Autocorrelation Structure

The term Generalized Cross-Validation (GCV) was proposed by Wahba (1977) and Craven and Wahba (1978) as a replacement of Cross-Validation (CV), it is the most popular method for choosing the complexity of statistical models. The basic principle of cross-validation is to leave the data points out one at a time and to choose the value of λ under which the missing data points are best predicted by the remainder of the data. To be precise, let g_{λ}^{-1} be the smoothing spline determined from all the information sets aside from (t_i, y_i), utilizing the worth λ for the smoothing boundary. The cross-validation decision regarding λ can then be the estimation of λ which can be written as;

$$CV(\lambda) = \frac{1}{n}\sum\{y_i - \hat{g}(t_i)\}^2 \tag{3.4}$$

Equation (3.4) is similar to the test for regression model estimation, (Cook and Weisberg, 1982). Define a matrix A (λ) by;

$$A_{ij}(\lambda) = n^{-1}g(t_i, t_j) \tag{3.5}$$

$$CV(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\frac{\{y_i - \hat{g}(t_i)\}^2}{\{1 - A_{ii}(\lambda)\}^2} \tag{3.6}$$

Craven and Wahba (1979) also proposed the application of a related test referred to as the Generalized Cross-validation, acquired from (3.6) by substituting $A_{ii}(\lambda)$ with its mean value, $n^{-1}trA(\lambda)$, this gives the score.

$$GCV(\lambda) = \frac{n^{-1}RSS(\lambda)}{(1 - n^{-1}trA(\lambda))^2} \tag{3.7}$$

Where; RSS (λ) refers to the residual sum of squares. Craven and Wahba (1979) also gave hypothetical contentions to prove that GCV ought to pick an ideal estimation of λ in the function of minimizing the mean squared error (MSE) at the design points. The forecast published practical examples bear out a good performance in Xiang and Wahba (1998). The summed-up Cross-validation technique is notable for its optimal qualities (Wahba, 1990). For any given $n \times n$, the impact matrix is given as;

$$\begin{bmatrix} \hat{f}_n, \lambda(t_1) \\ \hat{f}_n, \lambda(t_2) \\ . \\ . \\ . \\ \hat{f}_n, \lambda(t_n) \end{bmatrix} = S(\lambda)y, \text{ therefore } W_0(\lambda) \text{ can be revised as;} \tag{3.8}$$

where; $$W_0(\lambda) = \frac{\sum_{k=1}^{n}(a_{kj}y_j - y_k)^2}{(1-a_{kk})^2} \tag{3.9}$$

Generalized Cross-Validation is the changed type of Cross-Validation, a customary method for assessing the smoothing boundary. The GCV score is built by correlation with CV score which is gotten from the normal residuals by dividing them by$1 - (S_\lambda)_{ii}$. The acknowledged arrangement of GCV is to replace the documentation $1 - (S_\lambda)_{ii}$ in Cross-Validation with the mean score $1 - n^{-1}$ follow(S_λ). Consequently, by adding the residual squared and notation $\{1 - n^{-1}\ trace\ (S_\lambda)\}^2$, by the known conventional cross-approval, the GCV smoothing technique is composed numerically as;

$$GCV(\lambda) = \frac{1}{n}\frac{\sum_{k=1}^{n}\{y - f_k(x_1)\}^2}{\{1 - n^{-1}trace(S\lambda)\}^2}$$

$$GCV(\lambda) = \frac{n^{-1}\|(I - S\lambda)y\|^2}{[n^{-1}trace(I - S\lambda)]^2} \tag{3.10}$$

Where;

n = observations or data set (x_i, y_i)

λ = smoothing parameter

S_λ = refers to the ith diagonal member of the smoothing matrix

The first research on cross-validation was conducted by (Wahba, 1975), which was subsequently augmented to the log periodogram's smoothing (Wahba, 1980). The term Generalized Cross-Validation (GCV) was determined by Wahba (1977). The GCV score figured by similarity to CV score can be gotten from the normal residuals by isolating them by $1 - (S\lambda)_{ii}$. The essential plan of GCV is to supplant the components $1 - (S\lambda)_{ii}$ with the mean score $1 - n^{-1}\ tr(S\lambda)$. Consequently, adding the squared revised remaining and factor $\{1 - n^{-1}\ tr(S\lambda)\}$. Given the spline smoothing for non-parametric assessment of a relapse work in a period series setting and accepting that the reaction variable y_i are taken on the occasion t_i, for $i = 1, \ldots, n$ and that a model of the structure creates the y_i

$$y_i = f(t_i) + Z(t_i) \tag{3.11}$$

Where f(.) refers to the smoothing function and $Z(t_i)$ refers to the zero-mean, Autocorrelated stationary process. It can be said that even though t_i is specific, it is not uniformly spaced, with $t_1 < \ldots < t_n$

If the $Z(t_i)$ in (3.12) has a known correlation function, with $\mathrm{cov}\{Z(t_i), \ldots, Z(t_j)\} = \sigma^2 v_{ij}$, a normal addition of the usual smoothing spline approach amongst is to estimate f by the $\hat{f}$ which minimizes;

$$\left(y - \hat{f}\right)^T W(y - f) + \lambda \int_a^b \{f''(t)\}^2 dt \tag{3.12}$$

Amid every properly smoothed functions f, it is confirmed that $W = V^{-1} = [v_{ij}]$, $y = (y_1, \ldots, y_n)^T$ and $f = (f_{(t1)}, \ldots, f_{(tn)})^T$. It has been proven that the function $\hat{f}$ remains a natural cubic spline that has knots at the t_j. Also, if $\hat{f}$ denotes the vector with the ith element $\hat{f}(t_i)$ then there are a matrix S_λ such that $\hat{f} = S(\lambda_y)$, i.e. for fixed λ, the estimate is a direct capacity of y. This linearity proposes a nearby association between spline smoothing and bit smoothing, as shown unequivocally in (Silverman, 1984). One approach to pick the parameter denoted by λ is for the generalized cross-correlation to be minimized (Craven and Wahba, 1979). In the current setting, the common extension of this model is to limit (3.8); this gives a technique for assessing g within the sight of a realized autocorrelation structure. Concerning span assessment of g, the Bayesian assumption presented by (Silverman, 1984) extends with the connection network and V is replaced by Silverman's inverse inclining weighting lattice, which presents the posterior difference matrix, written as;

$$Var(\hat{f}) = \sigma^2 A(\lambda)V \tag{3.13}$$

The minimization of GCV (λ) as proposed by (Wahba, 1983) and (Diggle and Hutchinson, 1989) is written as;

$$GCV(\lambda) = \frac{(y-\hat{f})^T W(y-f)}{[trace(I-S\lambda)]^2} \tag{3.14}$$

Where; $(S\lambda)$ = the ith diagonal element of smoother matrix

W = V^{-1} = [v_{ij}], the correlation function

$y = (y_1, \ldots, y_n)^T$

f = $(f_{(t_1)}, \ldots, f_{(tn)})^T$

3.2.2 Generalized Maximum Likelihood (GML) Estimation Method with Autocorrelation Structure

Yuedong (1998) proposed the GML technique for correlated data that possesses single parameter for smoothing observations. However, there exist two parameters for smoothing in the case of a bivariate model that should be assessed along with the covariance boundaries. Following a comparative determination, GML is given as;

$$GML(\lambda) = \frac{y^{'}(I - S\lambda)}{[det^{+}(I - S\lambda)]^{\frac{1}{n-m}}} \tag{3.15}$$

Where; $\det^{+}(I - S\lambda)$ refers to the product of $(n - m)$ non-zero eigenvalues of $(I - S\lambda)$

Yuedong (1998) provided a Bayesian model for the GML method's general framework and can calculate a spline estimate's posterior confidence intervals. Suppose that the data are simulated via;

$$y_i = f(t_i) + \varepsilon_i, \qquad i = 1,2,\ldots,n, \qquad t_i \epsilon [0,1] \tag{3.16}$$

Where; $\in = (\in_1, \ldots, \in_m)^{I} \sim N(0, \sigma^2 W^{-1})$ which do not dependent on *f*, Model (3.12) is usually referred to as a Bayesian model, it can also be known as a hierarchical model or a mixed-effects model. This Bayesian model is similar to the model illustrated by (Wahba, 1990), though the residuals are correlated. Based on the justification of (Wahba, 1990), it can be shown that;

$$\lambda = \frac{\sigma^2}{nb}$$

$$\lim_{a\to\infty} E(f(t)/y) = \hat{f}t \text{ and } \lim_{a\to\infty} \operatorname{cov}(f/y) = \sigma^2 AW^{-1}$$

Where; F = *(F (t₁)., . . , F(tₙ))'* and $a \to \infty$ expanded prior are estimated for polynomial coefficients with degrees smaller than m.

According to Yuedong (1998), the covariance matrix W^{-1} relies on several correlations with parameter vector of τ. Interesting covariance structures refers to first-order autoregressive for time-series observation, structured symmetry or unstructured for repeated measurements, and spatial data. GML with Autocorrelation structure is therefore given by;

$$GML(\lambda) = \frac{\lambda^{1} W(I - S\lambda)}{[\det^{+} W(I - S\lambda)]^{\frac{1}{n-m}}} \tag{3.17}$$

Where;

$\det^{+}(I - S\lambda)$ = the product of the $n - m$ nonzero eigenvalues of $(I - S\lambda)$.

λ = Smoothing parameter

W = the structure of the correlation

$S\lambda$ = the smoother matrix diagonal elements

$n = n_1 + n_2$ are the pair of observations

m = number of zero eigenvalues

3.2.3 Unbiased Risk (UBR) Estimation Method with Autocorrelation Structure

Unbiased Risk is also known as the Mallow's CP criterion; it was developed by C. L. Mallows' (1973) to evaluate the regression model fit dependency on Ordinary Least Square (OLS). It is used to estimate choice situations where explanatory variables can predict a few results and locate the best model associated with subset independent variables. The more modest the estimation of the Cp, the generally exact it is, the Cp is written numerically as;

$$UBR\ (\lambda) = \frac{\|(S\lambda - I)y\|^2}{tr(I - S\lambda)} \qquad (3.18)$$

Yuedong (1998) provides the UBR technique that can be used effectively to choose a smoothing parameter for cubic spline smoothing that possesses non-Gaussian information. It was developed by using Predictive Mean Square Errors (PMSE).

The Unbiased Risk with Autocorrelation structure can be written mathematically as;

$$UBR(\lambda) = \frac{\frac{1}{n}\left\|W^{\frac{k}{2}}(I - S\lambda)y\right\|^2}{\left[\frac{1}{n}trace\left(W^{k-1}(I - S\lambda)\right)\right]^2}\ \mathrm{k} = 0, 1, 2 \qquad (3.19)$$

Where;

n = measurement/observations {xi,yi}

W = the Autocorrelation structure

λ = parameter used for smoothing

Sλ = matrix smoother of the ith diagonal member

3.2.4: Algorithm for GCV, GML and UBR

The algorithm of the GCV, GML and UBR spline smoothing selection methods are as follows.

Step 1: Given ρ and α, use the dataset (x, y) in the window to fit a curve and predict a steps ahead by linear extension.

Step 2: Calculate the mean square prediction error (MSE) for these a points.

Step 3: Advance a steps ahead.

Step 4: Repeat steps 1–3 until the end of the data set.

Step 5: Sum up all the MSEs to get the corresponding GCV, GML or UBR score for the given values of λ and ρ and α.

Step 6: Estimate λ by minimizing the GCV, GML or UBR score.

3.2.5 Proposed Smoothing Method (PSM) with Autocorrelation Structure

A smoothing spline model is usually written as:

$$y_i = f(x_i) + \varepsilon_i \tag{3.20}$$

Where; y refers to the response variable, x refers to a predictor variable, f is the Regression function and $\varepsilon_i \sim N\left(N(0, \sigma W^{-1})\right)$.

There are several options to examine whenever model (3.20) is used for non-linearity, it incorporates, observation change, addition substance items, for example, cubic spline and Spline smoothing. This research work is keen on spline smoothing because it examines non-linearity dependent on regression bend by presenting a wrinkle or twists in these crimps is created by pivot work, and the place of the turn on the fit is called hitches.

The traditional regression analysis primary purpose is to minimize the residual sum of square (RSS); the model with the minimum RSS is the preferred model. It is important to note that

Wahba (1979) proposed Cross-Validation (CV) as a technique for estimating Spline Smoothing. Instead of RSS in the customary straightforward simple regression, the residual is characterized as;

$$\varepsilon = y_i - \hat{y}_i \tag{3.21}$$

$$RSS = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.22}$$

$Y_i = f(x_i)$ refers to the observed variable and $\hat{Y}_i = f_\lambda(x_i)$ the fitted variable when many knots are introduced.

The proposed smoothing technique is the minimizer of the penalized residual sum of square (RSS) of the unbiased risk (UBR) and Generalized Cross-Validation (GCV). The unbiased risk (UBR) and generalized cross-validation (GCV) smoothing technique are modified because it has the best performance when applied to the data set correlated in nature.

$$P_k = \frac{1}{n}(y - \hat{f})^T W^g (y - f) + \lambda \int_0^1 (f''(x_i))^2 \, dx \tag{3.23}$$

The underlying articulation in the equation is the RSS for the goodness-of-fit information while the subsequent term is a rough penalty, which is huge when the integration of the second derivative, λ is the parameter used to smoothing, the regression function, $f''(x)$ is enormous when $f(x)$ isn't smooth (for example, with a slope that changes rapidly). Assuming λ approaches 0, $f''(x)$ just interpolates the information, and when λ is exceptionally large, then, at that point $f(x)$ will be chosen so that $f''(x)$ is everywhere, which infers a general linear least-squares fit for all the datasets.

Taking the Euclidean standard of (3.18) as set up by Wang (1996) and setting λ = 0, we have

$$P_k = \frac{1}{n}\left\| W^{\frac{g}{2}}(y - \hat{f}) \right\| \; g = 0,1 \tag{3.24}$$

Note that; if $f = (f(x_1), \dots, f(x_n))$ is the function vector f for knot $x_1, \dots, x_n$. The spline smoothing estimate $\hat{f}$ of the fitted observations $y = (y_1, \dots, y_n)^T$ can be written as;

$$\hat{f} = \begin{bmatrix} \hat{f}(x_1) \\ \hat{f}(x_2) \\ \cdot \\ \cdot \\ \cdot \\ \hat{f}(x_n) \end{bmatrix}_{(n \times 1)} = S_\lambda y \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}_{(n \times 1)} \quad \text{or} \quad \hat{f} = S_\lambda y \tag{3.25}$$

where $\hat{f}_\lambda$ refers to a normal cubic spline having knots at $x_1, \dots, x_n$ based on an established $\lambda > 0$, and $S\lambda$ is revealed as a matrix smoother with positive-definite (symmetrical) that relies on λ and the knot points $x_1, \dots, x_n$, but not on y. Function f_λ, the estimation of function f is derived from the cubic spline interpolation with the conditions $\hat{f}_\lambda(xi) = (\hat{f})i$, $i = 1, 2, \dots, n$.

$$P_k = \frac{1}{n^2}\left\| W^{\frac{g}{2}}(y - S\lambda y) \right\| \; g = 0,1 \tag{3.26}$$

The Generalized Cross-Validation smoothing technique is famous for its optimal qualities in the estimation smoothing techniques (Wahba 1990) and the unbiased risk (UBR) technique is usually applied when estimating the parameters of a smoothing spline with non-Gaussian observations. It also fit the time series dataset correctly and does not overfit observations, (Morris and Lieberman, 2008), (Gu, 1992) and (Wahba, et al., 1995), The spline smoothing technique that is proposed is a combination of the optimal properties and qualities of UBR and GCV.

The Generalized Cross-Validation is the minimizer of;

$$GCV(\lambda) = \frac{1}{n} \frac{\sum_{k=1}^{n} \{y - f_k(x_1)\}^2}{\{1 - n^{-1} trace(S\lambda)\}^2}$$

$$= \frac{n^{-1}\|(y - S\lambda)y\|^2}{[n^{-1}trace(I - S\lambda)]^2} = \frac{n^{-1}\|(I - S\lambda)y\|^2}{[n^{-1}trace(I - S\lambda)]^2} \quad (3.27)$$

(Diggle and Hutchinson, 1989) proposed GCV method for estimating spline smoothing function (f) when autocorrelation is present in the residual, it is written as;

$$GCV(\lambda) = \frac{(y - \hat{f})^T W(y - f)}{[trace\{1 - S\lambda\}]^2} \quad (3.28)$$

Therefore; equation (3.26) becomes;

$$P_k = \frac{1}{n}\left\|W^{\frac{g}{2}}(y - S\lambda y)\right\| = \frac{1}{n}\left\|W^{\frac{g}{2}}(I - S\lambda)y\right\| \quad (3.29)$$

The expected value (U_k) of (3.29) as established by Wang (1996) is given as;

$$U_k = \frac{1}{n}(I - S\lambda)yW^g(I - S\lambda)y - \frac{\sigma^2}{n}tr(S\lambda W^g S\lambda W^{-1})y \quad g = 0,1,2 \quad (3.30)$$

While the unbiased estimate of 3.29 is written as;

$$U_k = \frac{1}{n}y(I - S\lambda)W^g(I - S\lambda)y - \frac{\sigma^2}{n}tr(W^{g-1})S\lambda y - 2\frac{\sigma^2}{n}trW^{g-1}S\lambda \quad (3.31)$$

The minimization of $U_k$in (3.31) is known as the UBR. If σ^2 is made the subject of the formula, UBR estimator becomes;

$$\text{UBR } (\lambda) = \frac{\frac{1}{n}\left\|W^{\frac{g}{2}}(I-S\lambda)y\right\|^2}{\left[\frac{1}{n}trace\left(W^{g-1}(I-S\lambda)\right)\right]^2} \quad (3.32)$$

In this manner, an improved spline smoothing technique is proposed by adding the weighted parameters k and k − 1 with the other properties and qualities of the UBR and GCV (Adams and Ipinyomi, 2020). The combination of the two smoothing methods' quantities will result in optimal performance of smoothing methods whose model does not overfit time-series observations. The minimizer of the hybrid of (3.28) and (3.32) is the Proposed Smoothing Method given as;

PSM = (k) overfitting and optimal knot detector + $(1-k)$ best for forecasting non-Gaussian data

$$PSM\ (\lambda) = k\frac{(y-\hat{f})^T W(y-\hat{f})}{[trace(I-S\lambda)]^2} + (1-k)\frac{\frac{1}{n}\left\|W^{\frac{g}{2}}(I-S\lambda)\right\|^2}{\left[\frac{1}{n}trace\{W^{g-1}(I-S\lambda)\}\right]^2} \tag{3.33}$$

The behaviour of the minimized λ in UBR and GCV techniques under the alternate value of g = 1 as the optimum value of PSM yields;

$$PSM\ (\lambda) = k\frac{(y-\hat{f})^T W(y-\hat{f})}{[trace(I-S\lambda)]^2} + (1-k)\frac{\frac{1}{n}\left\|W^{\frac{1}{2}}(I-S\lambda)\right\|^2}{\left[\frac{1}{n}trace\{W(I-S\lambda)\}\right]^2} \tag{3.34}$$

The proposed method for estimating ***f*** is given in (3.34) subject to the condition that $0 < g < 1$ is chosen, using the algorithm in section 3.2.5.

Where;

n = number of dataset

k = weighted value, $0 < k < 1$

W = V^{-1} = Correlation Matrix for the error term

$y = (y_1, \ldots, y_n)^T$ = Smoothing function

$\hat{f} = (f(t_1) \ldots f(t_n)). y_n)^T = S_\lambda y$

Sλ= the diagonal member of the smoothing matrix

$\left\|W^{\frac{1}{2}}(I-S\lambda)y\right\|$= norm of the Euclidean vector $W^{\frac{1}{2}}(y-\hat{f})$

3.2.6 Proposed Smoothing Method (PSM) Algorithm

Step 1: Read the simulated sample data(x_i, y_i) for i = 1 – T and for each of the $i's$ determine the

Pre-selected smoothing parameters$\lambda_1, \ldots, \lambda_t$, calculate the respective set of smoothing

Spline estimates $f(\lambda) = \{\hat{f}_{\lambda 1}, \ldots, \hat{f}_{\lambda t}\}$

Step 2: For the given λ, σ and T use the data in 1 above to fit a curve and the estimate ahead by linear extension $f(x_i)$ and $\hat{f}(x_i)$

Step 3: Insert the weighted value (k) of the coefficients of GCV and UBR

Step 4: Obtain the predictive mean square error $PMSE\ (\hat{f}\lambda) = \sum_{i=1}^{t}\left[\left(f(x) - \hat{f}(xi)\right)^2\right]$ for these points

Step 5: add all values of PMSEs to get the resulting PSM value for the given λ and ρ.

Step 6: Repeat steps 1–5 for 1000 times**.**

3.3 Monte Carlo Simulation study

This part is concerned with the outcome of a Monte Carlo simulation study. This study was led to assess the achievement of the four smoothing techniques depicted in this research, for example, GML, GCV, UBR and PSM. Dataset was generated by applying a program written in R (version 3.2.3) for time-series sample sizes of; 20, 60 and 100. The experiment was replicated 1,000 for every one of the examples. The Predictive Mean Squared-Errors (PMSE), adjusted R-Square and predicted R-square was utilized to assess the smoothing techniques' quality and performance for each simulated data.

3.4 Equation used to generate the value in the data

The data generation study performed to assess and measure the performance of the four spline smoothing methods is given as;

$$y(t) = 2Sin\left(\frac{\pi}{t}\right) + \varepsilon_t \quad t = 20, 60 \text{ and } 100 \qquad (3.35)$$

Where; $\pi = 180^0$, $\varepsilon_t \sim N(0, \sigma W^{-1})$, a first-order autoregressive process AR (1) with a mean of 0, a standard deviation of 0.8, autocorrelation levels (ρ) of 0.2, 0.5 and 0.8 with a 95% confidence limit, Note that; $e_t = \rho\varepsilon_{t-1} + v_t$ and $v_t \sim N(0, \sigma^2)$

3.5 Experimental design and data generation

The experimental design adopted in this study is;

- Three-time-series samples (T) of 20, 60 and 100 were considered in the data generation
- Three autocorrelation levels were considered, i.e. ρ = 0.2, 0.5 and 0.8
- Four smoothing degrees were thought of, for example; d.s = 1, 2, 3, 4
- One standard deviation value was considered, i.e. $\sigma = 0.8$
- Dataset was simulated for 1,000 replications in each of the 3 x 3 x 4 x 1 = 36 combinations for cases T's, ρ's, λ's and σ's.

All the selected parameters the in experimental design are similar to the ones used in (Wahba, 1985).

3.6 Smoothing Spline Assessment methods used in this Study

Efforts were made in this study to examine and compare the strength of the four spline smoothing estimators, namely; Generalized Cross-Validation (GCV), Unbiased risk (UBR), Generalized Maximum Likelihood (GML) and the Proposed Smoothing Method (PSM) developed by taking the weighted hybrid of GCV and UBR.

3.7 Evaluation of the Smoothing Methods

3.7.1 Predictive Mean Square Error

A comparison was made to test the four estimation methods' effect and performance in the presence of autocorrelation error. An estimate of the four smoothing methods, the criterion, effect and performance of different autocorrelation errors of the four estimation methods (i.e.

Generalized Crossed Validation (GCV), Generalized Maximum Likelihood (GML), Proposed Smoothing Methods (PSM) ($0 < k < 1$) and Unbiased Risk (UBR)) were performed using codes written in R-console. Four different estimation methods were used i.e. GCV (V), GML (M), PSM ($0 < k < 1$) and UBR (U). Thus data generation was carried out for V, M, P and U. At the same time, the Evaluation and comparison of the Four (4) Spline Smoothing estimation methods were investigated by applying the asymptotic sampling qualities of the criterion given as; Mean Square Prediction Error (MSPE).

The Predictive Mean squared error (PMSE) of a smoothing curve or model fitting process, according to C. Mallows (1973) and C. Daniel (1973), is the difference between the expected value of the square difference of the fitted value, that is; function $\hat{f}(x_i)$ and the observed value estimate is given as the function $f(x_i)$. It is utilized to estimate the performance and attributes of smoothing methods like Cross-Validation, Generalized Cross-Validation, and Generalized Maximum Likelihood etc. The Predictive Mean Square Error (PMSE) is written numerically as;

$$PMSE(\lambda) = E\left[\sum_{i=1}^{n}\left(f(x_i) - \hat{f}(x_i)\right)^2\right] \qquad (3.36)$$

$$PMSE(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - \hat{f}(x_i)\right)^2 \qquad (3.37)$$

$$PMSE(\lambda) = \sum_{i=1}^{n}\left(E[\hat{f}(x_i)] - f(x_i)\right)^2 = \sum_{i=1}^{n} var[\hat{f}(x_i)] \qquad (3.38)$$

The Predictive Mean Square Error is usually separated into two parts; the initial part is the sum of square biases of the fitted qualities, and the other part is the number of changes in the fitted observations.

Where;

$f(x_i)$ = observed value

$\hat{f}(x_i)$ = predicted or estimated value

At each scenario of specification, for instance, say, time-series size (T) = 20**,** autocorrelation level **(**ρ) = 0.2, d.f = 1 and standard deviation (σ) = 0.8, the smoothing methods were tested and compared using the asymptotic properties of the estimators based on the PMSE criterion.

3.7.2 Test for Over-fitting in Spline Smoothing

In statistics, overfitting occurs when a model fails to fit extra information or neglect to anticipate future perceptions reliably. An overfitted model is a statistical equation that contains a greater amount of parameters than can 'be advocated by the data. If more focus goes about as bunches in the fitted model, it tends to be alluded to as an overfitted model. The embodiment of overfitting is to have unconsciously separated some lingering variety (for example the commotion) as though that variety addressed the hidden model construction.

3.7.3 Causes of over-fitting in Spline Smoothing Models

(i) Overfitting is caused by including too many predictors in the model

(ii) Overfitting can also be caused by applying a small smoothing parameter

(iii) it can be controlled by using AIC or BIC

The chance of overfitting exists because the rule utilized for choosing the model isn't equivalent to the rule used to pass judgment on the reasonableness of a model. For instance, a model may be picked by expanding its presence on some arrangement of preparing information, but its appropriateness may be dictated by its capacity to perform well on inconspicuous information. Overfitting happens when a model starts to "remember" preparing information as opposed to "learning" to sum up from a pattern.

As a limit model, if the quantity of boundaries is equivalent to or more prominent than the number of perceptions, at that point a model can consummately foresee the preparation information basically by retaining the information completely. (For a representation, see Figure 1.) Such a model, however, will bomb when making expectations.

The potential for overfitting relies not only on the measure of boundaries and dataset but additionally on the model design's relationship with the data type and the capacity of the model's residual in contrasted with the normal degree of the data disturbance. In any event, when the fitted model doesn't have an inordinate number of boundaries, it is observed that the fitted relationship will perform below on another informational collection than on the informational collection utilized for fitting (a marvel is in some cases known as shrinkage). Specifically, the estimation of the coefficient of assurance will shrivel compared with the first information.

To reduce the opportunity of overfitting, a few procedures are accessible (for example model correlation, cross-approval, regulating, early halting, pairing, Bayesian priors and withdrawal). The premise of certain methods is either (1) to expressly punish extreme complex models or (2) to test the capacity of the model, to sum up by assessing its exhibition on a bunch of information not utilized for preparing, which is accepted to surmise the common concealed information that a model will experience.

3.7.4 How to detect Overfitting

PRESS and Predicted R-square are the best and easiest ways to discover overfitted in smoothing methods and models. The result may be interpreted by simply comparing the predicted R-Square to the normal R-Square and observe if there exist is a great difference between the two test techniques. If there is a large difference between the two values, the model doesn't predict new observations and fits the true data, and there is the possibility of overfitting in the model. Overfit model has too many numbers of terms and begins to fit the random noise in the sample; it is not possible to predict random noise. The Predictive R-square is a statistical technique that determines how well a model predicts response for new observation. It is something of an in-house cooked measure, which is computed by effectively eradicating each variable from the data set, estimation of the regression model, and deciding how suitable the model forecast the removed variables. Predictive R-square is usually written mathematically as;

$$Pred.R - Square = \left(1 - \frac{Predicted\ residual\ sum\ of\ squares\ (PRESS)}{Sum\ of\ square\ total}\right) \times 100 \qquad (3.39)$$

While R-square, also known as the coefficient of determination, can be derived through;

$$R - Square = \left(1 - \frac{Sum\ of\ squares\ error}{Sum\ of\ square\ total}\right) \qquad (3.40)$$

3.7.5 Test for Goodness-of-fit for the Smoothing Methods

The goodness-of-fit of the smoothing methods explains how well the methods fit the simulated and real-life data. It also summarizes the differences between the observed value and predicted or estimated values. The Adjusted R-square was used to determine the best-fit smoothing methods. It is written mathematically as;

$$Adjusted\ R - Square = \left(1 - \frac{(1 - Rsquare) \times (n - 1)}{n - p}\right) \qquad (3.41)$$

Where; n = number of observations and p = number of parameters.

3.8 Samples of Simulated Data

A sample of the generated data when T = 20**,** ρ = 0.2, d.s. = 1 and σ = 0.8, for residual with autocorrelation is given in Table 3.1.

Table 3.1: Sample of generated data for the dependent and independent variables When; $\rho = 0.2$, T = 20, d.s. = 1 and $\sigma = 0.8$

N	X	Y
1	1.564345e-01	0.316972672
2	3.090170e-01	0.66295818
3	4.539905e-01	0.93706458
4	5.877853e-01	0.61559262
5	7.071068e-01	0.86798006
6	8.090170e-01	1.00740679
7	8.910065e-01	-0.22485289
8	9.510565e-01	0.68185480
9	9.876883e-01	1.62282915
10	1.000000e+00	0.46241107
11	9.876883e-01	0.06896351
12	9.510565e-01	1.16371538
13	8.910065e-01	1.14943203
14	8.090170e-01	-0.68363206
15	7.071068e-01	-0.87938274
16	5.877853e-01	0.21990529
17	4.539905e-01	1.82920076
18	3.090170e-01	0.88954175
19	1.564345e-01	1.16579886
20	1.224606e-01	1.30350655

CHAPTER FOUR

4.0 ANALYSIS, RESULTS AND DISCUSSIONS

4.1 INTRODUCTION

This chapter is concerned with the analysis, simulation results and discussion of finding of Generalized Cross-Validation, Generalized Maximum Likelihood, Proposed Smoothing Method (PSM) for ($0 < k < 1$) and Unbiased Risk (UBR) methods. Section 4.2 presented the simulation result of the cubic spline regression, predictive mean square errors, R-Squared, Adjusted R-Square and predicted R-Square of the four smoothing methods at three-time series periods; small (T = 20), medium (T = 60) and fairly large (T = 100), four (4) degrees of smoothing (D.S. = 1, 2, 3 and 4), considered in the presence of three autocorrelation levels; low level ($\rho = 0.2$), moderate level ($\rho = 0.5$) and high level ($\rho = 0.8$) and standard deviations level ($\sigma = 0.8$). The standard deviation was selected because other selected standard deviations showed similar properities and results. Smoothing spline curve, and box plot of the simulated results were also presented later in section 4.4. The performance of the four smoothing methods on real-life data was presented in section 4.5, while a summary of findings and discussion of the findings were presented in section 4.6 and 4.7 respectively.

4.2 ANALYSIS AND SIMULATED RESULTS

Table 4.2.1: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T=20, ρ = 0.2,* D.S. *= 1,2,3,4

Conditions	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adj. R Square	Pred. R Square
T = 20, D.S. = 1, ρ = 0.2	GML	β_0 = -0.6396 β_1 = 15.8453*** β_2= -42.6026*** β_3= 27.3335***	0.07271685	0.788134	0.774	0.7316	0.012
T = 20, D.S. = 1, ρ = 0.2	GCV	β_0 = 0.2575 β_1 = 11.9332*** β_2 =-37.9413*** β_3 = 25.7132***	0.005146929	4.938284	0.8686	0.844	0.000
T = 20, D.S. = 1, ρ = 0.2	UBR	β_0 = -0.08908 β_1 = 11.46518*** β_2 = -34.32253*** β_3 = 23.07244***	0.7095105	3.777261	0.9569	0.9488	0.067
T = 20, D.S. = 1, ρ = 0.2	PSM (K =0.01)	β_0 = -0.1310 β_1 = 11.1676*** β_2 = -31.8914*** β_3 = 22.0510***	0.702466864	3.49336	0.9662	0.9599	0.6533
T = 20, D.S. = 2, ρ = 0.2	GML	β_0 = -0.05533 β_1 = 10.87027*** β_2 = -33.50242*** β_3 = 22.66933***	0.06092801	2.638143	0.8264	0.7939	0.0011
T = 20, D.S. = 2, ρ = 0.2	GCV	β_0 = 0.05592 β_1 = 12. 17278** β_2 = -37.91667*** β_3 =26.16078***	0.0007558556	2.789043	0.7595	0.7144	0.000
T = 20, D.S. = 2, ρ = 0.2	UBR	β_0 = -0.1156 β_1 = 11.5687*** β_2 = -36.0596*** β_3 = 24.8536***	0.2uf318559	3.469432	0.7341	0.6843	0.003
T = 20, D.S. = 2, ρ = 0.2	PSM (K =0.02)	β_0 = 0.1584 β_1 = 10.0075** β_2 = -30.6810*** β_3 = 20.5930***	0.227233899	2.19728	0.849	0.8207	0.6671
T = 20, D.S. = 3, ρ = 0.2	GML	β_0 = 0.2389 β_1 = 8.6301** β_2 = -27.0257*** β_3 = 18.3359***	0.05148411	3.624478	0.8069	0.7707	0.0090
T = 20, D.S. = 3, ρ = 0.2	GCV	β_0 = -0.00695 β_1 = 13.55549** β_2 = -40.30658*** β_3 = 26.77317***	0.0002782435	3.175146	0.8461	0.8172	0.0010
T = 20, D.S. = 3, ρ = 0.2	UBR	β_0 = -0.1214 β_1 = 12.3206*** β_2 = -33.7997*** β_3 = 21.2857***	0.05151468	3.416732	0.5597	0.5324	0.0022
T = 20, D.S. = 3, ρ = 0.2	PSM (K =0.03)	β_0 = 0.3846 β_1 = 7.6419 β_2 = -27.8681* β_3 = 20.5278**	0.049977587	3.147642	0.8689	0.8443	0.5324
T = 20, D.S. = 4, ρ = 0.2	GML	β_0 = 0.5596 β_1 = 5.9713** β_2 = -23.5218* β_3 17.3125**	0.04384406	1.34165	0.6663	0.6037	0.000
T = 20, D.S. = 4, ρ = 0.2	GCV	β_0 = -0.2034 β_1 = 14.7471** β_2 = -42.272*** β_3 = 27.8714***	0.0001058604	0.956241	0.6884	0.6299	0.0009
T = 20, D.S = 4, ρ = 0.2	UBR	β_0 = -0.1656 β_1 = 9.2593 β_2 = -25.7615* β_3 = 16.5350*	0.06503716	2.59853	0.5401	0.4538	0.000
T = 20, D.S = 4, ρ = 0.2	PSM (K =0.04)	β_0 = -0.1376 β_1 = 12.0791*** β_2 = -35.2248*** β_3 = 234056***	0.062439908	0.046857	0.9678	0.9618	0.6428

Table 4.2.1 presents the summary fit result of the cubic spline regression model and the model performance criteria, namely; the PMSE, multiple, adjusted and predicted R-square based on three-time series periods *(T=20)*, four smoothing degrees *(D.S.=1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.2)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05)*. The PMSE of the four smoothing methods indicates that; the proposed smoothing method (PSM *k=0.04*) had the least PMSE of *0.046857* when *T = 20, D.S = 4, ρ=0.2*. This was closely followed by, GML with PSME of *0.788134* at *T = 20, D.S = 1, ρ = 0.2* and GCV with PSME of *0.956241* at *T = 20, D.S = 4* and$\rho = 0.2$. The result is that; the Proposed Smoothing Method performs better than the other smoothing methods at a time series size of *20* and rho = *0.2*.

The adjusted R-Square result indicated that the Proposed Smoothing Methods *(PSM = 0.04)* had the highest values of *0.9618* at, *T = 20, D.S. = 4, ρ=0.2*, which is closely followed by the *(PSM = 0.01)* of *0.9599* at *T = 20, D.S. = 1, ρ* = 0.2 and then UBR value of *0.9569* when *T = 20, D.S = 1 and ρ=0.2*. It can be inferred from the result above that; the Proposed Smoothing Method provides the best fit to the time-series observations at a time series size of 20 and rho = 0.2.

It can be seen from the result presented in Table 4.2.1 that the difference between the Multiple R-square and predictive R-square of the Proposed Smoothing Method (PSM) was the least when compared to the other smoothing methods. At *T = 20, D.S. = 1, 2, 3 and 4,*$\rho = 0.2$*,* the differences between the Multiple R-Square and predictive R-square was *0.3066, 0.1536, 0.3119, and 0.319* respectively. This indicates that the Proposed Smoothing Method (PSM) does not overfit the time series observations when the time-series size of 20 and rho = 0.2.

Table 4.2.2: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T=60, ρ = 0.2, D.S = 1,2,3,4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameters	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R – Square
T = 60, D.S. = 1, $\rho = 0.2$	GML	β_0 = 0.03431 β_1 = 9.63438*** β_2 = -29.17766*** β_3 = 19.61762***	0.00182211	2.328352	0.7789	0.7671	0.000
T = 60, D.S. = 1, $\rho = 0.2$	GCV	β_0 = -0.5186* β_1 = 14.0075*** β_2 = -38.3854*** β_3 = 25.1621***	0.0002119457	1.353605	0.7325	0.7182	0.031
T = 60, D.S = 1, $\rho = 0.2$	UBR	β_0 = -0.01773 β_1 = 10.24776*** β_2 = -30.97090*** β_3 = 20.84673***	0.05471852	2.101405	0.9301	0.9264	0.001
T = 60, D.S. = 1, $\rho = 0.2$	PSM (K =0.05)	β_0 = -0.01624 β_1 =1073462*** β_2 = -32.70862*** β_3 = 22.20427***	0.051993191	1.25185	0.9394	0.9362	0.6733
T = 60, D.S = 2, $\rho = 0.2$	GML	β_0 = -0.26110*** β_1 = 11.99961*** β_2 = -34.26321*** β_3 = 22.72655***	0.001723715	2.194483	0.9532	0.9506	0.000
T = 60, D.S = 2, $\rho = 0.2$	GCV	β_0 = -0.13304*** β_1 = 11.48356*** β_2 = -33.60079*** β_3 = 22.39321***	0.0001360985	1.123143	0.9302	0.9264	0.022
T = 60, D.S. = 2, $\rho = 0.2$	UBR	β_0 = -0.05896 β_1 = 10.8270*** β_2 = -31.12181*** β_3 = 21.39417***	0.04854293	1.887882	0.9492	0.9464	0.000
T = 60, D.S. = 2, $\rho = 0.2$	PSM (K =0.06)	β_0 = -0.18831* β_1= 12.15985*** β_2 = 35.09021*** β_3 = 23.31395***	0.04563852	1.060665	0.958	0.9558	0.221
T = 60, D.S. = 3, $\rho = 0.2$	GML	β_0 = -0.18831* β_1 = 12.15985*** β_2 = -35.09021*** β_3 = 23.31395***	0.001632201	2.094332	0.9461	0.9432	0.000
T = 60, D.S. = 3, $\rho = 0.2$	GCV	β_0 = -0.06303 β_1 = 10.78518*** β_2 = -32.70251*** β_3 = 22.20277***	8.991862e-05	2.472227	0.9436	0.9406	0.014
T = 60, D.S. = 3, $\rho = 0.2$	UBR	β_0 = -0.10848 β_1 = 11.58245*** β_2 = -34.02879*** β_3 = 22.74942***	0.04578347	1.857928	0.9474	0.9446	0.002
T = 60, D.S. = 3, $\rho = 0.2$	PSM (K =0.07)	β_0 = -0.19794* β_1 = 11.82713*** β_2 = -34.10548*** β_3 = 22.73478***	0.042584921	1.349367	0.9385	0.9352	0.7000
T = 60, D.S. = 4, $\rho = 0.2$	GML	β_0 = -0.02920 β_1 =10.77812*** β_2 = -31.19231*** β_3 = 21.57675***	0.001546984	3.615946	0.9478	0.945	0.000
T = 60, D.S. = 4, $\rho = 0.2$	GCV	β_0 = -0.14492 β_1 = 10.83452*** β_2 = -31.70387*** β_3 = 21.11344***	6.08904e-05	2.018062	0.9454	0.9425	0.022
T = 60, D.S. = 4, $\rho = 0.2$	UBR	β_0 = -0.11607 β_1 = 11.22499*** β_2 = -33.07054*** β_3 = 22.13243***	0.04321843	3.398581	0.9553	0.9529	0.000
T = 60, D.S. = 4, $\rho = 0.2$	PSM (K =0.08)	β_0 = -0.25494** β_1 = 12.44873*** β_2 = -35.69443*** β_3 = 23.63443***	0.039765827	2.416724	0.9459	0.9422	0.7752

"*", "**", "", Significant at 0.05, 0.01 and 0.001respectively***

Table 4.2.2 presents the summary fit result of the cubic spline regression model and the model performance criteria, namely; the PMSE, multiple, adjusted and predicted R-square based on time-series periods *(T = 60)*, four degrees of smoothing *(D.S. =1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.2)*.

It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05)*.

The PMSE of the four smoothing methods indicates that; the Proposed Smoothing Method (*PSM k=0.06*) had the smallest PMSE of *1.06066* when *T = 60, D.S. = 2 and ρ=0.2*. This was closely followed by, GCV with PSME of *1.12314* at *T = 60, D.S.= 2 and ρ = 0.2* and then, (*PSM k=0.05*) with PSME of *1.25185* at *T = 60, D.S. = 1* and *ρ* = 0.2. The result implies that; the proposed smoothing method outperformed the other smoothing methods at a time series size of *60* and rho = *0.2*. The adjusted R-Square result indicated that the Proposed Smoothing Methods *(PSM = 0.06)* had the highest values of *0.9558* at, *T = 60, D.S. = 2 and ρ=0.2*, which is closely followed by the UBR with *0.9529* at *T = 60, D.S. = 4, ρ* = 0.2 and then GML with value of *0.9506* when *T = 60, D.S. = 2 and ρ=0.2*. It can be inferred from the result above that; the Proposed Smoothing Method provides the best fit to the time-series observations at a time series size of 60 and rho = 0.2.

It can be seen from the result presented in Table 4.2.2 that the difference between the Multiple R-square and predictive R-square of the Proposed Smoothing Method was the least when compared to the other smoothing methods. At *T = 60, D.S. = 1, 2, 3 and 4,ρ* = 0.2, the differences between the Multiple R-Square and predictive R-square was *0.2629, 0.737, 0.6283, and 0.1707* respectively. This result shows that the Proposed Smoothing Method (PSM) does not overfit the time series observations when the time-series size of 60 and rho = 0.2.

Table 4.2.3: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T= 100, ρ = 0.2, D.S. = 1,2,3,4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R Square	Pred. R Square
T = 100, D.S. = 1, $\rho = 0.2$	GML	β_0 = -0.07252 β_1 = 11.05677*** β_2 = -32.95240*** β_3 = 22.16509***	0.001053936	2.314015	0.9385	0.9366	0.0023
T = 100, D.S. = 1, $\rho = 0.2$	GCV	β_0 = -0.13070 β_1 = 11.26376*** β_2 = -32.90224*** β_3 = 21.84165***	0.0001238005	1.334855	0.9413	0.9394	0.0011
T = 100, D.S. = 1, $\rho = 0.2$	UBR	β_0 = -0.23231** β_1 = 11.96419*** β_2-34.30740*** β_3 = 22.73138***	0.03089095	1.913073	0.9365	0.9345	0.0015
T = 100, D.S. = 1, $\rho = 0.2$	PSM (K =0.09)	β_0 = -0.28796*** β_1 = 12.53243*** β_2 =-35.51768*** β_3 = 23.47008***	0.028121907	1.057611	0.9481	0.9465	0.569
T = 100, D.S. = 2, $\rho = 0.2$	GML	β_0 = -0.29667 β_1 = 12.48217*** β_2 = -35.35595*** β_3 = 23.32564***	0.001019743	2.040446	0.9451	0.9434	0.001
T = 100, D.S. = 2, $\rho = 0.2$	GCV	β_0 = -0.22601*** β_1 = 11.93776*** β_2 = -34.13633*** β_3 = 22.56803***	6.447385e-05	0.341562	0.9481	0.9464	0.0023
T = 100, D.S. = 2, $\rho = 0.2$	UBR	β_0 = -0.13362* β_1 = 11.82477*** β_2 = -34.81465*** β_3 = 23.40127***	0.02877716	1.431244	0.9534	0.9519	0.0025
T = 100, D.S. = 2, $\rho = 0.2$	PSM (K =0.1)	β_0 = -0.14742* β_1 = 11.65768*** β_2 = -34.00987*** β_3 = 22.62938***	0.025905891	0.277435	0.9507	0.9492	0.6271
T = 100, D.S. = 3, $\rho = 0.2$	GML	β_0 = -0.14040 β_1 = 11.35370*** β_2 =-33.14178*** β_3 = 22.03136***	0.0009869954	1.990265	0.9431	0.9413	0.000
T = 100, D.S. = 3, $\rho = 0.2$	GCV	β_0 =-0.24064 β_1 = 12.29247*** β_2 = -35.12102*** β_3 = 23.25426***	6.919766e-05	0.334902	0.9381	0.9362	0.0016
T = 100, D.S. = 3, $\rho = 0.2$	UBR	β_0 = -0.1306 β_1 = 11.6022*** β_2 = -35.0624*** β_3 = 23.8722***	0.02684178	1.361115	0.8823	0.8692	0.000
T = 100, D.S. = 3, $\rho = 0.2$	PSM (K =0.11)	β_0 = -0.20346 β_1 = 11.78338*** β_2 -34.26368*** β_3 = 22.94096***	0.023896796	0.237523	0.9404	0.9385	0.6755
T = 100, D.S. = 4, $\rho = 0.2$	GML	β_0 = -0.12003* β_1 = 10.88492*** β_2 = -31.70714*** β_3 = 21.05287***	0.0009556187	1.973208	0.9300	0.9279	0.000
T = 100, D.S. = 4, $\rho = 0.2$	GCV	β_0 = -0.11858 β_1 = 11.47119*** β_2 =-33.74040*** β_3 = 22.58175***	5.411037e-05	0.332736	0.9474	0.9458	0.000
T = 100, D.S. = 4, $\rho = 0.2$	UBR	β_0 = -0.19474** β_1 = 11.64251*** β_2 = -33.84756*** β_3 = 22.51969***	0.02506713	1.337717	0.9535	0.9521	0.011
T = 100, D.S. = 4, $\rho = 0.2$	PSM (K =0.12)	β_0 = -0.20195** β_1 =-12.07546*** β_2 = -35.23636*** β_3 = 23.62322***	0.022065568	1.207069	0.9353	0.9333	0.5438

"*", "**", "", Significant at 0.05, 0.01 and 0.001respectively***

Table 4.2.3 presents the summary fit result of the cubic spline regression model and the model performance criteria, namely; the PMSE, multiple, adjusted and predicted R-square based on time-series period *(T=100)*, four degrees of smoothing *(D.S.=1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.2)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05)*.

The PMSE of the four smoothing methods indicates that; the proposed smoothing method (*PSM k=0.11*) had the smallest PMSE of 0*.237523* at *T = 100, D.S. = 3 and ρ=0.2*. This was closely followed by, *(PSM, K=0.1)* with PSME of 0*.277435* at *T = 100, D.S. = 2 and ρ = 0.2* and then, GCV with PSME of *0.334902* at *T = 100, D.S. = 3* and $\rho = 0.2$. The result implies that; the proposed smoothing method performs better than the other smoothing methods at a time series size *(T =60)* and rho = *0.2*.

The adjusted R-Square result indicated that the Unbiased Risk (UBR) had the largest values of *0.9534* at, *T = 100, D.S. = 2 and ρ=0.2*, which is closely followed by another UBR with *0.9519* at *T = 100, D.S. = 2 and*$\rho = 0.2$, then *(PSM, K=0.1)* with value of *0.9492* when *T = 100, D.S. = 2 and ρ=0.2*. It can be inferred from the result above that; the Unbiased Risk (UBR) smoothing method provides the best fit to the time-series observations at a time series size *(T = 100)* and rho = 0.2.

It can be seen from the result presented in Table 4.2.3 that the difference between the Multiple R-square and predictive R-square of the Proposed Smoothing Method (PSM) was the least when compared to the other smoothing methods. At *T = 60, D.S. = 1, 2, 3 and 4,*$\rho = 0.2$*,* the differences between the Multiple R-Square and predictive R-square was *0.3491, 0.3236, 0.2649, and 0.3915* respectively. This result shows that the Proposed Smoothing Method does not overfit the time series observations when the time series size of 60 and rho = 0.2.

Table 4.2.4: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T = 20, ρ = 0.5, D.S. = 1,2,3,4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R-Square
T = 20, D.S. = 1, $\rho = 0.5$	GML	β_0 = 0.1356 β_1 = 10.5252*** β_2 = -30.9899*** β_3 = 19.7856***	0.03760678	3.902353	0.8165	0.7821	0.1560
T = 20, D.S. = 1, $\rho = 0.5$	GCV	β_0 = 0.2845 β_1 = 8.1743* β_2 = -25.1419** β_3 = 16.4673**	0.002496273	5.700411	0.7082	0.6535	0.0017
T = 20, D.S. = 1, $\rho = 0.5$	UBR	β_0 = -0.02659 β_1 = 8.94968* β_2 = -25.78774** β_3 = 16.28941**	0.3272748	2.810875	0.6957	0.6386	0.0010
T = 20, D.S. = 1, $\rho = 0.5$	PSM (K =0.13)	β_0 = -0.04711 β_1 = 12.04046*** β_2 = -34.62610*** β_3 = 22.49415***	0.285053591	2.09983	0.8542	0.8268	0.766
T = 20, D.S. = 2, $\rho = 0.5$	GML	β_0 = -0.4148 β_1 = 13.2513*** β_2 = -35.7006*** β_3 = 22.4028***	0.03247149	2.804237	0.7833	0.7427	0.0610
T = 20, D.S. = 2, $\rho = 0.5$	GCV	β_0 = 0.2663 β_1 = 11.3367*** β_2 =-36.2679*** β_3 =24.9837***	0.001269176	3.755684	0.9301	0.917	0.000
T = 20, D.S. = 2, $\rho = 0.5$	UBR	β_0 = -0.8495** β_1 = 15.7143*** β_2 = -38.8818*** β_3 = 23.9641***	0.2982672	2.506771	0.8867	0.8654	0.016
T = 20, D.S. = 2, $\rho = 0.5$	PSM (K =0.14)	β_0 = -0.2161 β_1 = 15.1787*** β_2-45.6749*** β_3 = 31.4115***	0.256687477	2.00634	0.9054	0.8876	0.564
T = 20, D.S. = 3, $\rho = 0.5$	GML	β_0 = -0.4162 β_1 = 14.0206*** β_2 = -37.9720*** β_3 = 24.5667***	0.02821014	3.802802	0.7768	0.7349	0.000
T = 20, D.S. = 3, $\rho = 0.5$	GCV	β_0 = -0.05684 β_1 = 13.71211*** β_2 = -40.60826*** β_3 = 27.13928***	0.0003526313	3.507623	0.8258	0.7932	0.000
T = 20, D.S. = 3, $\rho = 0.5$	UBR	β_0 = -0.06938 β_1 = 10.68291*** β_2 = -32.42170*** β_3 = 21.67153***	0.2723018	3.526772	0.8449	0.8158	0.001
T = 20, D.S. = 3, $\rho = 0.5$	PSM (K =0.15)	β_0 = 0.2934 β_1 = 8.4237** β_2 = -27.5452*** β_3 = 18.8431***	0.231509425	3.51337	0.8568	0.8300	0.6711
T = 20, D.S. = 4, $\rho = 0.5$	GML	β_0 = 0.1047 β_1 = 11.7009*** β_2 = -36.7510*** β_3 = 25.1998***	0.02464804	1.47087	0.7143	0.6608	0.0009
T = 20, D.S. = 4, $\rho = 0.5$	GCV	β_0 = 0.1513 β_1 = 9.9033** β_2 = -30.5509*** β_3 = 20.5665***	0.0001254823	1.437364	0.7653	0.7213	0.020
T = 20, D.S. = 4, $\rho = 0.5$	UBR	β_0 = -0.1341 β_1 = 10.1868*** β_2 = -31.8230*** β_3 = 21.9838***	0.2490225	1.14616	0.7363	0.6868	0.0013
T = 20, D.S. = 4, $\rho = 0.5$	PSM (K =0.16)	β_0 = -0.3008 β_1 = 10.4908*** β_2 = -28.9898*** β_3 = 18.8467***	0.209198977	1.794844	0.8802	0.8577	0.742

Table 4.2.4 presents the summary fit result of the cubic spline regression model and the model performance criteria, i.e. the PMSE, multiple, adjusted and predicted R-square based on time-series period *(T=20)*, four degrees of smoothing *(D.S.=1, 2, 3 and 4)* and autocorrelation level $(\rho = 0.5)$. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05).*

The PMSE of the four smoothing methods indicates that; the Unbiased Risk (*UBR*) had the smallest PMSE of *1.14616* at $T = 20$, *D.S.* $= 4$ *and* $\rho=0.5$. This was closely followed by, *GCV* with PSME of *1.14616* at $T = 20$, *D.S.* $= 4$ *and* $\rho = 0.5$ and then, GCV with PSME of *1.437364* at $T = 100$, *D.S.* $= 4$ and$\rho = 0.5$. The result implies that; the UBR smoothing method performs better than the other smoothing methods at a time series size $(T = 100)$ and rho = *0.5*.

The adjusted R-Square result showed that the GCV had the largest values of *0.9301* at, $T = 20$, *D.S.* $= 2$ *and* $\rho=0.5$, which is closely followed by another Proposed Smoothing Method (PSM=0.14) with *0.8876* at $T = 20$, *D.S.* $= 2$ *and*$\rho = 0.5$, the UBR with value of *0.8867* when $T = 20$, *D.S.* $= 2$ *and* $\rho=0.5$. It can be inferred from the result above that; the GCV smoothing method provides the best fit to the time-series observations at a time series size $(T = 20)$ and rho = 0.5.

It can be seen from the result presented in Table 4.2.4 that the difference between the Multiple R-square and predictive R-square of the PSM was the least when compared to the other smoothing methods. At $T = 60$, *D.S.* $= 1, 2, 3$ *and* $4, \rho = 0.2$, the differences between the Multiple R-Square and predictive R-square was *0.0882, 0.3414, 0.1857, and 0.1382* respectively. This result shows that the PSM does not overfit the time series observations when the time-series size of 60 and rho = 0.2.

Table 4.2.5: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T= 60, ρ = 0.5, D.S. = 1,2,3,4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R-Square
T = 60, D.S. = 1, $\rho = 0.5$	GML	β_0 = 0.03431 β_1 = 9.63438*** β_2 = -29.17766*** β_3 =19.61762***	0.001467536	4.557857	0.736	0.7218	0.0210
T = 60, D.S. = 1, $\rho = 0.5$	GCV	β_0 = -0.5186* β_1 = 14.0075*** β_2 = -38.3854*** β_3 = 25.1621	0.0001820162	5.735483	0.7325	0.7182	0.032
T = 60, D.S. = 1, $\rho = 0.5$	UBR	β_0 = 0.2122 β_1 = 8.0745*** β_2 = -26.3317*** β_3 = 18.0531***	0.04083165	2.449087	0.6995	0.6834	0.095
T = 60, D.S. = 1, $\rho = 0.5$	PSM (K =0.17)	β_0 = -0.07609 β_1 = 11.16403*** β_2 = -32.90640*** β_3 =21.91931***	0.033921212	1.28961	0.7789	0.7671	0.412
T = 60, D.S. = 2, $\rho = 0.5$	GML	β_0 = -0.3962* β_1 = 12.4542*** β_2 = -34.5625*** β_3 = 22.7527***	0.00139338	1.300494	0.7539	0.7407	0.000
T = 60, D.S. = 2, $\rho = 0.5$	GCV	β_0 = -0.2143 β_1 = 12.8547*** β_2 = -36.6300*** β_3 = 24.2212***	0.0001434969	5.368908	0.7939	0.7828	0.0161
T = 60, D.S. = 2, $\rho = 0.5$	UBR	β_0 = 0.09892 β_1 = 9.418808*** β_2 = -30.66073*** β_3 = 21.44320***	0.0386085	1.017353	0.7670	0.7546	0.000
T = 60, D.S. = 2, $\rho = 0.5$	PSM (K =0.18)	β_0 = -0.0147 β_1 = 11.4113*** β_2 = -.33.9764*** β_3 = 22.8099***	0.031684799	1.141933	0.8192	0.8095	0.7828
T = 60, D.S. = 3, $\rho = 0.5$	GML	β_0 = -0.2384 β_1 = 12.0230*** β_2 = -34.1681*** β_3 = 22.7732***	0.001324088	4.263339	0.7329	0.7177	0.0112
T = 60, D.S. = 3, $\rho = 0.5$	GCV	β_0 = -0.4198* β_1 = 14.2173*** β_2 = 40.2033*** β_3 = 26.8429***	0.0001151162	4.218419	0.7867	0.7753	0.0078
T = 60, D.S. = 3, $\rho = 0.5$	UBR	β_0 = -0.1058 β_1 = 9.414*** β_2 = -28.1641*** β_3 = 18.7199***	0.03110486	3.611805	0.7237	0.7089	0.0001
T = 60, D.S. = 3, $\rho = 0.5$	PSM (K =0.19)	β_0 = 0.1835 β_1 = 9.4004*** β_2 = -29.3852*** β_3 = 19.8782***	0.025216809	1.289411	0.781	0.7692	0.3211
T = 60, D.S. = 4, $\rho = 0.5$	GML	β_0 = -0.2659 β_1 = 12.3001*** β_2 = -35.1093*** β_3 = 23.1599***	0.001259269	2.16225	0.7355	0.7213	0.000
T = 60, D.S. = 4, $\rho = 0.5$	GCV	β_0 = -0.2518 β_1 = 12.452*** β_2 = -36.210*** β_3 = 25.1076***	9.37482e-05	2.188967	0.7384	0.7453	0.021
T = 60, D.S. = 4, $\rho = 0.5$	UBR	β_0 = -0.3808 β_1 = 13.5769*** β_2 = -38.1391*** β_3 = 25.0224***	0.02952382	2.49091	0.7791	0.7672	0.000
T = 60, D.S. = 4, $\rho = 0.5$	PSM (K =0.20)	β_0 = -0.2591 β_1 = 11.4306*** β_2 = -32.4205*** β_3 = 21.5264***	0.023637806	1.733162	0.7879	0.7766	0.612

"*", "**", "", Significant at 0.05, 0.01 and 0.001***

Table 4.2.5 presents the summary fit result of the cubic spline regression model and the model performance criteria, i.e. the PMSE, multiple R-square, adjusted R-Square and predicted R-square based on time-series periods *(T=60)*, four degrees of smoothing *(D.S.=1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.5)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05).*

The PMSE of the four smoothing techniques indicated that; the Proposed Smoothing Method (*PSM = 0.18*) had the smallest PMSE of *0.757980* at *T = 60, D.S. = 2 and ρ=0.5*. This was closely followed by, *UBR* with PSME of *1.017353* at *T = 60, D.S. = 2 and ρ = 0.5* then, GML with PSME of *1.300494* at *T = 60, D.S. = 2* and$\rho = 0.5$. The result implies that; the Proposed Smoothing Method (*PSM = 0.18*) performs better than the other smoothing methods at a time series size *(T =60)* and rho *= 0.5*.

The adjusted R-Square result showed that the Proposed Smoothing Method (*PSM = 0.18*) had the largest values of *0.8095* at, *T = 60, D.S. = 2 and ρ=0.5*, which is closely followed by the Proposed Smoothing Method *(PSM = 0.20)* with value of *0.7879 when T = 20, D.S. = 2 and ρ=0.5,* then GCV smoothing method with *0.7828* at *T = 60, D.S. = 2 and*$\rho = 0.5$. It can be inferred from the result above that; Proposed Smoothing Method (PSM = 0.18), provides the best fit to the time-series observations at a time series size *(T = 60)* and rho = 0.5.

It can be seen from the result presented in Table 4.2.5 that the difference between the Multiple R-square and predictive R-square of the Proposed Smoothing Method was the least when compared to the other smoothing methods. At *T = 60, D.S. = 1, 2, 3 and 4,*$\rho = 0.5$, the differences between the Multiple R-Square and predictive R-square was *0.3669, 0.0364, 0.4599, and 0.1759* respectively. This result shows that the Proposed Smoothing Method does not overfit the time series observations when the time-series size of 60 and rho = 0.5.

Table 4.2.6: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T=100, ρ = 0.5, D.S. = 1 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R-Square
T = 100, D.S. = 1, $\rho = 0.5$	GML	β_0 = -0.02474 β_1 = 9.49648*** β_2 = -27.77235*** β_3 = 18.30185***	0.0009255424	2.836043	0.6955	0.686	0.086
T = 100, D.S. = 1, $\rho = 0.5$	GCV	β_0 = -0.2115 β_1 = 11.3905*** β_2 = -33.1234*** β_3 = 21.9689***	0.0001064592	4.190077	0.7958	0.7894	0.000
T = 100, D.S. = 1, $\rho = 0.5$	UBR	β_0 = -0.2296 β_1= 12.4728*** β_2 = -36.5954*** β_3 = 24.7277***	0.02981056	2.079789	0.7342	0.7259	0.001
T = 100, D.S. = 1, $\rho = 0.5$	PSM (K =0.21)	β_0 = -0.3965 β_1 = 13.6553*** β_2 = -37.8524*** β_3 = 24.6123***	0.023572699	1.012881	0.8102	0.8042	0.688
T = 100, D.S. = 2, $\rho = 0.5$	GML	β_0 = -0.1851 β_1 = 12.1911*** β_2 = 35.5182*** β_3 = 23.7436***	0.0008967007	1.334802	0.7159	0.707	0.026
T = 100, D.S. = 2, $\rho = 0.5$	GCV	β_0 = -0.2265 β_1 = 13.0928*** β_2 = 37.8117*** β_3 = 25.2492***	9.220767e-05	2.968761	0.7621	0.7546	0.0016
T = 100, D.S. = 2, $\rho = 0.5$	UBR	β_0 = -0.4555** β_1 = 13.4744*** β_2 = -38.0403*** β_3 = 25.3550***	0.02877716	1.220508	0.7631	0.7556	0.000
T = 100, D.S. = 2, $\rho = 0.5$	PSM (K =0.22)	β_0 = -0.07192 β_1 = 11.75938*** β_2= -36.36752*** β_3 = 2517009***	0.02246647	0.869461	0.8037	0.7975	0.5444
T = 100, D.S. = 3, $\rho = 0.5$	GML	β_0= -0.2037 β_1 = 12.9573*** β_2 – 37.4776*** β_3 = 25.0977***	0.0008690314	2.22264	0.7441	0.7361	0.000
T = 100, D.S. = 3, $\rho = 0.5$	GCV	β_0 = -0.1195 β_1 = 10.7767*** β_2 – 31.9788*** β_3 = 21.4264***	8.038807e-05	0.815361	0.7708	0.7636	0.092
T = 100, D.S. = 3, $\rho = 0.5$	UBR	β_0 = -0.3621* β_1 = 13.0180*** β_2 = -35.9349*** β_3 = 23.4976***	0.02778833	1.866935	0.7567	0.7491	0.000
T = 100, D.S. = 3, $\rho = 0.5$	PSM (K =0.23)	β_0 = -0.3428 β_1 = 13.3860*** β_2 = -38.8748*** β_3 = 26.2003***	0.021415503	0.720518	0.7954	0.789	0.412
T = 100, D.S. = 4, $\rho = 0.5$	GML	β_0 = -0.1604 β_1 = 11.3986*** β_2 = -32.6706*** β_3 = 21.5384***	0.000842476	1.984518	0.7532	0.7455	0.009
T = 100, D.S. = 4, $\rho = 0.5$	GCV	β_0 = -0.2635 β_1 = 12.6782*** β_2 = -36.9091*** β_3 = 24.7118***	0.0001544047	2.765412	0.7876	0.7803	0.0112
T = 100, D.S. = 4, $\rho = 0.5$	UBR	β_0 = -0.4740** β_1 = 14.5119*** β_2 = -41.2996*** β_3 = 27.7351***	0.02684178	1.815722	0.7768	0.7698	0.000
T = 100, D.S. = 4, $\rho = 0.5$	PSM (K =0.24)	β_0 =-0.08351 β_1 = 12.12945*** β_2 = -35.37438*** β_3 = 23.59913***	0.02043681	1.739182	0.8046	0.7984	0.2133

"*", "**", "", Significant at 0.05, 0.01 and 0.001***

Table 4.2.6 presents the summary fit result of the cubic spline regression model and the model performance criteria, i.e. the PMSE, multiple, adjusted and predicted R-square based on time-series periods *(T=100)*, four degrees of smoothing *(D.S.=1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.5)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05)*.

It was discovered from the PMSE of the four smoothing methods that; the Proposed Smoothing Method (*PSM = 0.23*) had the smallest PMSE of *0.720518* at *T = 100, D.S. = 3 and ρ=0.5*. This was closely followed by, GCV with PSME of *0.815361* at *T = 100, D.S. = 3 and ρ = 0.5* then, Proposed Smoothing Method (*PSM = 0.22*) with PSME of *0.869461* at *T = 100, D.S. = 2* and$\rho = 0.5$. The result implies that; the Proposed Smoothing Method (*PSM = 0.23*) performs better than the other smoothing methods at a time series size *(T =100)* and rho = *0.5*.

The adjusted R-Square result showed that the Proposed Smoothing Method (*PSM = 0.21and 0.22*) had the largest values of *0.8042 and 0.7975* at, *T = 100, D.S. = 1, 2 and ρ=0.5*, which is closely followed by the GCV smoothing method with value of *0.7894 when T = 100, D.S. = 1 and ρ = 0.5*. It can be inferred from the result above that; Proposed Smoothing Method (PSM = 0.21), provides the best fit to the time-series observations at a time series size *(T = 100)* and rho = 0.5.

It can be seen from the result presented in Table 4.2.6 that the difference between the Multiple R-square and predictive R-square of the PSM was the least when compared to the other smoothing methods. At *T = 60, D.S. = 1, 2, 3 and 4*,$\rho = 0.5$, the differences between the Multiple R-Square and predictive R-square was *0.122, 0.2531, 0.377, and 0.5913* respectively. This result shows that the PSM does not overfit the time series observations when the time-series size of 100 and rho = 0.5.

Table 4.2.7: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T=20, ρ = 0.8, D.S. = 1,2,3,4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R Square
T = 20, D.S. = 1, $\rho = 0.8$	GML	β_0 = 0.5246 β_1 = 6.3651 β_2 = -205017 β_3 = 13.1974	0.02165009	4.557857	0.5121	0.4206	0.0025
T = 20, D.S. = 1, $\rho = 0.8$	GCV	β_0 = 0.2864 β_1 = 10.1265* β_2 = -29.8584** β_3 = 18.5067**	0.002844546	5.735483	0.677	0.6164	0.021
T = 20, D.S. = 1, $\rho = 0.8$	UBR	β_0 = 0.02683 β_1 = 7.60565 β_2 = -21.53505 β_3 = 12.91283	0.6382572	2.449087	0.4789	0.3812	0.015
T = 20, D.S. = 1, $\rho = 0.8$	PSM (K =0.25)	β_0 = -0.00602 β_1 = 12.55090* β_2 = -33.67642** β_3 = 22.84042**	0.479404037	2.272232	0.7606	0.7158	0.3462
T = 20, D.S. = 2, $\rho = 0.8$	GML	β_0 = 0.4955 β_1 = 11.4249*** β_2 = -38.3034*** β_3 = 26.8238***	0.01911081	1.300494	0.8623	0.8365	0.0102
T = 20, D.S. = 2, $\rho = 0.8$	GCV	β_0 = -0.5943 β_1 = 14.4883** β_2 = -37.3956** β_3 = 22.6942**	0.001562165	5.368908	0.6388	0.5711	0.0098
T = 20, D.S. = 2, $\rho = 0.8$	UBR	β_0 = -1.2899*** β_1 = 18.429*** β_2= -42.4855*** β_3 = 25.1924***	0.6346121	1.017353	0.7878	0.748	0.0021
T = 20, D.S. = 2, $\rho = 0.8$	PSM (K =0.26)	β_0 = -0.2764 β_1 = 17.5721*** β_2 = -53.3544*** β_3 = 37.1082***	0.470019117	0.757980	0.9282	0.8959	0.6734
T = 20, D.S. = 3, $\rho = 0.8$	GML	β_0 = -0.02158 β_1 = 15.22553** β_2 = -45.24788*** β_3 = 30.27262***	0.01694719	4.263339	0.6942	0.6369	0.0131
T = 20, D.S. = 3, $\rho = 0.8$	GCV	β_0 = -0.5444 β_1 = 12.0274** β_2 = -29.1544** β_3 = 17.0444*	0.000954459	4.218419	0.6239	0.5534	0.00119
T = 20, D.S. = 3, $\rho = 0.8$	UBR	β_0 = -0.8479 β_1 = 15.9066** β_2 = -42.2310*** β_3 = 27.0919***	0.6307937	3.611805	0.6458	0.5794	0.00113
T = 20, D.S. = 3, $\rho = 0.8$	PSM (K =0.27)	β_0 = -0.05927 β_1 = 10.59759* β_2 = -32.16036** β_3 = 22.21666**	0.460737105	3.03644	0.7558	0.6725	0.4523
T = 20, D.S. = 4, $\rho = 0.8$	GML	β_0 = 0.3308 β_1 = 9.6598 β_2 = -32.6249* β_3 = 23.2359**	0.01509332	2.16225	0.5406	0.4545	0.000
T = 20, D.S. = 4, $\rho = 0.8$	GCV	β_0 = 0.5529 β_1 = 8.6756 β_2 = -30.8329 * β_3 = 21.8131**	0.0006283503	2.188967	0.6508	0.5854	0.122
T = 20, D.S. = 4, $\rho = 0.8$	UBR	β_0 = 0.02747 β_1 = 13.26481** β_2 = -42.53308*** β_3 = 30.18391***	0.6178076	2.49091	0.7091	0.6545	0.000
T = 20, D.S. = 4, $\rho = 0.8$	PSM (K =0.28)	β_0 = -0.04164 β_1 = 10.37883* β_2 = -32.14938** β_3 = 21.52423**	0.44499741	1.483121	0.7442	0.6962	0.3567

"*", "**", "", Significant at 0.05, 0.01 and 0.001***

Table 4.2.7 presents the summary fit result of the cubic spline regression model and their performance criteria, namely; the PMSE, multiple, adjusted and predicted R-square based on time series period *(T=20)*, four degrees of smoothing *(D.S. =1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.8)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05)*.

It was discovered from the PMSE of the four smoothing methods that; the Proposed Smoothing Method (*PSM = 0.26*) had the smallest PMSE of *0.757980* at *T = 20, D.S. = 2 and ρ = 0.8*. This was closely followed by, UBR with PSME of *1.017353* at *T = 20, D.S. = 2 and ρ = 0.8* then, *GML* with PSME of *1.300494* at *T = 20, λ = 2* and $\rho = 0.8$. The result implies that; the Proposed Smoothing Method (*PSM = 0.26*) performs better than the other smoothing methods at a time series size *(T =20)* and rho = *0.8*.

The adjusted R-Square result showed that the Proposed Smoothing Method (*PSM = 0.26*) had the largest values of *0.8959* at, *T = 20, D.S. = 2 and ρ=0.8*, which is followed by the GML smoothing method with value of *0.8623 when T = 20, D.S. = 2 and ρ = 0.8,* then the UBR smoothing method with adjusted R-square value of 0.7878 when *T = 20, D.S. = 2 and ρ = 0.8.* It can be inferred from the result above that; Proposed Smoothing Method (PSM = 0.26), provides the best fit to the time-series observations at a time series size *(T = 20)* and rho = 0.8.

It can be seen from the result presented in Table 4.2.7 that the difference between the Multiple R-square and predictive R-square of the PSM was the least when compared to the other smoothing methods. At *T = 20, D.S. = 1, 2, 3 and 4,* $\rho = 0.8$, the differences between the Multiple R-Square and predictive R-square was *0.3696, 0.2548, 0.3035, and 0.3875* respectively. This result shows that the PSM does not overfit the time series observations when the time-series size of 20 and rho = 0.8.

Table 4.2.8: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T= 60, ρ = 0.8, D.S. = 1,2,3,4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R Square
T = 60, D.S. = 1, $\rho = 0.8$	GML	β_0 =-0.2402 β_1 = 13.5496*** β_2 = -38.1697*** β_3 = 25.1284***	0.001198572	2.625861	0.6536	0.635	0.0033
T = 60, D.S. = 1, $\rho = 0.8$	GCV	β_0 =0.2609 β_1 = 8.0510** β_2 = -28.6189*** β_3 = 20.6836***	0.0001913283	5.817303	0.5767	0.5541	0.016
T = 60, D.S. = 1, $\rho = 0.8$	UBR	β_0 = -0.2741 β_1 = 15.1331*** β_2 = -43.9626*** β_3 = 29.5064***	0.07033239	1.118518	0.6218	0.6015	0.090
T = 60, D.S. = 1, $\rho = 0.8$	PSM (K =0.29)	β_0 = 0.07909 β_1 = 11.24005*** β_2 = -33.92403*** β_3 = 22.87027***	0.049991482	1.058281	0.6611	0.643	0.432
T = 60, D.S. = 2, $\rho = 0.8$	GML	β_0= -0.2788 β_1 = 12.2188*** β_2 = -34.2308*** β_3 = 22.8117***	0.001141677	1.027948	0.4992	0.4724	0.000
T = 60, D.S. = 2, $\rho = 0.8$	GCV	β_0 = -0.5691 β_1 = 15.7297*** β_2 = -43.8871*** β_3 = 29.3231***	0.0001502637	4.406313	0.5953	0.5736	0.011
T = 60, D.S. = 2, $\rho = 0.8$	UBR	β_0 = -0.06665 β_1 = 8.24832** β_2 = -24.62434*** β_3 = 16.32634***	0.06595468	1.230349	0.4663	0.4377	0.003
T = 60, D.S. = 2, $\rho = 0.8$	PSM (K =0.3)	β_0 = 0.3962 β_1 = 8.0227** β_2 = -26.5781*** β_3 = 18.1796***	0.046213355	0.310471	0.6181	0.5976	0.4533
T = 60, D.S. = 3, $\rho = 0.8$	GML	β_0 = -0.3228*** β_1 = 12.6622*** β_2-35.7367*** β_3 = 23.4304***	0.001088294	2.996486	0.4901	0.4628	0.017
T = 60, D.S. = 3, $\rho = 0.8$	GCV	β_0 = -0.5067 β_1 = 14.7052*** β_2 = – 40.5843*** β_3 = 26.4103***	0.0001201504	1.456264	0.6046	0.5834	0.000
T = 60, D.S. = 3, $\rho = 0.8$	UBR	β_0 = -0.2674 β_1 = 12.7413*** β_2 = -37.2583*** β_3 = 25.1483***	0.06191276	2.564013	0.5871	0.565	0.021
T = 60, D.S. = 3, $\rho = 0.5$	PSM (K =0.31)	β_0 = -0.3119 β_1 = 11.2709*** β_2 = -31.4345*** β_3 = 20.8167***	0.042757051	1.077814	0.6053	0.5842	0.3231
T = 60, D.S. = 4, $\rho = 0.8$	GML	β_0 = 0.1528 β_1 = 10.3817*** β_2 = -32.6398*** β_3 = 22.4550***	0.001038157	2.996486	0.5791	0.5565	0.000
T = 60, D.S. = 4, $\rho = 0.8$	GCV	β_0 = 0.1035 β_1 = 8.5890*** β_2 = -25.6512*** β_3 = 16.6920***	9.757026e-05	1.456264	0.5882	0.5661	0.010
T = 60, D.S. = 4, $\rho = 0.8$	UBR	β_0 = -0.5286 β_1 = 15.0744*** β_2 = -41.1956*** β_3 = 22.2068***	0.05817642	2.564013	0.5723	0.5494	0.000
T = 60, D.S. = 4, $\rho = 0.8$	PSM (K =0.32)	β_0 = 0.4314 β_1 = 8.4836** β_2 = -30.6451*** β_3 = 22.0310***	0.039591188	1.0654	0.6401	0.6208	0.433

"*", "**", "", Significant at 0.05, 0.01 and 0.001***

Table 4.2.8 presents the summary fit result of the cubic spline regression model and the model performance criteria, i.e. the PMSE, multiple R-square, adjusted R-square and predicted R-square based on time-series periods *(T=60)*, four degrees of smoothing *(D.S=1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.8)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05)*.

It was discovered from the PMSE of the four smoothing methods that; the Proposed Smoothing Method (*PSM = 0.3*) had the smallest PMSE of *0.310471* at *T = 60, D.S = 2 and ρ = 0.8*. This was followed by, *GML* with PSME of *1.027948* at *T = 60, D.S. = 2 and ρ = 0.8* then, Proposed Smoothing Method (*PSM = 0.29*) with PSME of *1.058281* at *T = 60, D.S. = 1* and $\rho = 0.8$. The result implies that; the Proposed Smoothing Method (*PSM = 0.3*) performs better than the other smoothing methods at a time series size *(T =60)* and rho = *0.8*.

The adjusted R-Square result showed that the Proposed Smoothing Method (*PSM = 0.29 and 0.32*) had the largest values of *0.643 and 0.6401* at, *T = 60, D.S. = 1 and 4 and ρ=0.8*, which is followed by GML smoothing method with adjusted R-Square value of *0.635 when T = 60, λ = D.S. and ρ=0.8.* It can be inferred from the result above that; Proposed Smoothing Method (PSM = 0.29), provides the best fit to the time-series observations at a time series size *(T = 60)* and rho = 0.8.

It can be seen from the result presented in Table 4.2.8 that the difference between the Multiple R-square and predictive R-square of the PSM was the least when compared to the other smoothing methods. At *T = 60, S.D. = 1, 2, 3 and 4*,$\rho = 0.8$, the differences between the Multiple R-Square and predictive R-square was *0.211, 0.1648, 0.2822, and 0.2071* respectively. This result shows that the PSM does not overfit the time series observations when the time-series size of 60 and rho = 0.8.

Table 4.2.9: Cubic Spline Regression Model of GML, GCV, PSM and UBR for T=100, ρ = 0.8, D.S. = 1,2,3.4 and σ = 0.8

Parameters	Smoothing Methods	β_i	Smoothing Parameter	Selection criteria			
				PMSE	R Square	Adjusted R-Square	Pred. R Square
T = 100, λ = 1, $\rho = 0.8$	GML	β_0 = 0.3593 β_1 = 8.1545*** β_2 = -27.7845*** β_3 = 19.4844***	0.0008169797	2.438085	0.5494	0.5353	0.0045
T = 100, λ = 1, $\rho = 0.8$	GCV	β_0 = -0.1129 β_1 = 11.0228*** β_2 = -32.6533*** β_3 = 21.924***	0.0001004289	4.753061	0.5767	0.5634	0.008
T = 100, λ = 1, $\rho = 0.8$	UBR	β_0 = -0.2370 β_1 = 14.1455*** β_2 = -42.2492*** β_3 = 28.9154***	0.02193897	2.841755	0.5634	0.5497	0.001
T = 100, λ = 1, $\rho = 0.8$	PSM (K =0.33)	β_0 = -0.1809 β_1 = 11.1282*** β_2 = -31.3267*** β_3 = 20.2266***	0.014732251	1.01515	0.673	0.6628	0.3165
T = 100, λ = 2, $\rho = 0.8$	GML	β_0 = 0.1120 β_1 = 7.4825** β_2 = -23.8088*** β_3 = 16.2494***	0.0007924906	0.171129	0.4371	0.4195	0.0131
T = 100, λ = 2, $\rho = 0.8$	GCV	β_0 = -0.009328 β_1 = 8.498622*** β_2 = -26.32829*** β_3 = 17.960381***	8.721889e-05	3.188995	0.4334	0.4156	0.001
T = 100, λ = 2, $\rho = 0.8$	UBR	β_0 = -0.09441 β_1 = 12.73832*** β_2 = -37.63141*** β_3 = 25.47346***	0.02123493	1.532589	0.6206	0.6088	0.022
T = 100, λ = 2, $\rho = 0.8$	PSM (K =0.34)	β_0 = -0.6479** β_1 = 14.6205*** β_2 = -39.5466*** β_3 = 25.6892***	0.014044708	0.052643	0.6313	0.6197	0.2766
T = 100, λ = 3, $\rho = 0.8$	GML	β_0 = -0.1608 β_1 = 12.6035*** β_2 = -35.4325*** β_3 = 22.9419***	0.0007689599	0.19773	0.5881	0.5753	0.000
T = 100, λ = 3, $\rho = 0.8$	GCV	β_0 = 0.07745 β_1 = 8.81596*** β_2 = -28.3936*** β_3 = 19.64520***	7.622553e-05	3.188995	0.5032	0.4877	0.000
T = 100, λ = 3, $\rho = 0.8$	UBR	β_0 = -0.1917 β_1 = 12.7727*** β_2 = -36.4885*** β_3 = 24.1191***	0.02055911	1.532589	0.5945	0.5819	0.012
T = 100, λ = 3, $\rho = 0.8$	PSM (K =0.35)	β_0 = -0.5191 β_1 = 14.5570*** β_2 = -41.0388*** β_3 = 27.1682***	0.0133901	0.85785	0.608	0.5957	0.322
T = 100, λ = 4, ρ = 0.8	GML	β_0 = 0.3072 β_1 = 6.7774*** β_2 = -24.1129*** β_3 = 17.3734***	0.0007463417	0.19754	0.5267	0.5119	0.000
T = 100, λ = 4, ρ = 0.8	GCV	β_0 = -.0.09507 β_1 = 13.05444*** β_2 = -38.07188*** β_3 = 25.36748***	0.0001467109	2.6543	0.6445	0.6333	0.013
T = 100, λ = 4, ρ = 0.8	UBR	β_0 = -0.2412 β_1 = 14.2442*** β_2 = -38.4952*** β_3 = 24.5294***	0.01991017	1.532589	0.5904	0.5776	0.071
T = 100, λ = 4, ρ = 0.8	PSM (K =0.36)	β_0 = -08702 β_1 = 10.95087*** β_2 = -33.32863*** β_3 = 22.61821***	0.012795325	0.056745	0.849	0.8207	0.4532

"*", "**", "", Significant at 0.05, 0.01 and 0.001***

Table 4.2.9 presents the summary fit result of the cubic spline regression model and the model performance criteria, i.e. the PMSE, multiple R-Square, adjusted R-Square and predicted R-square based on three-time series periods *(T=100)*, four smoothing parameters *(D.S.=1, 2, 3 and 4)* and autocorrelation level *(ρ = 0.8)*. It was revealed from the result that all the coefficients of the smoothing methods' parameters were significant at *(P-value <0.001, <0.01 and < 0.05).*

It was discovered from the PMSE of the four smoothing methods that; the Proposed Smoothing Method (*PSM = 0.34 and 0.36*) had the smallest PMSE of *0.052643 and 0.056745* at *T = 100, D.S. = 2 and 4 and ρ = 0.8*. This was followed by, *GML* with PSME of *0.1711129* at *T = 100, D.S. = 2 and ρ = 0.8*. The result implies that; the Proposed Smoothing Method (*PSM = 0.34*) performs better than the other smoothing methods at a time series size *(T =100)* and rho = *0.8*.

The adjusted R-Square result showed that the Proposed Smoothing Method (*PSM = 0.36, 0.33 and 0.34*) had the largest values of *0.8207*, 0.6628 and 0.6313 at, *T = 100, D.S. = 4, 1 and 2 and ρ=0.8.* It can be inferred from the result above that; Proposed Smoothing Method (PSM = 0.36), provides the best fit to the time-series observations at a time series size *(T = 100)* and rho = 0.8.

It can be seen from the result presented in Table 4.2.9 that the difference between the Multiple R-square and predictive R-square of the PSM was the least when compared to the other smoothing methods. At *T = 100, D.S. = 1, 2, 3 and 4,* $\rho = 0.8$, the differences between the Multiple R-Square and predictive R-square was *0.3463, 0.3547, 0.286, and 0.3958* respectively. This result shows that the PSM does not overfit the time series observations when the time-series size of 100 and rho = 0.8.

4.3 Smoothing Curves for Goodness-of-fit with Autocorrelation in the Error Term

The smoothing curve presented in Figure 4.3.1 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 20, standard deviation = 0.8 and for a small autocorrelation levels of 0.2. The observed value of the cubic spline is the (red) line, while the curve was fitted with UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML (brown). It can be seen from the curve that the smoothing parameter (λ) has been chosen by PSM (k=0.04) and has provided the best fit for the simulated time series observation.

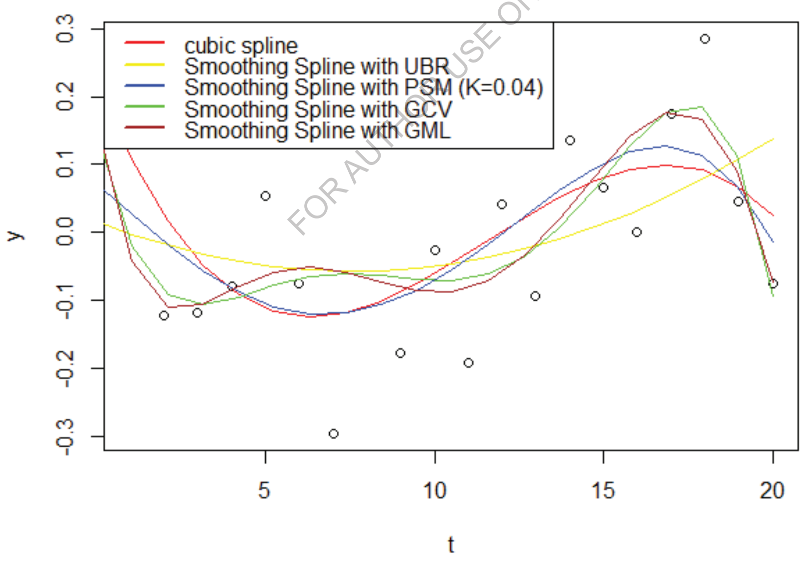


Figure 4.3.1: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 20, rho = 0.2, sigma =0.8

The smoothing curve presented in Figure 4.3.2 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 20, standard deviation = 0.8 and for a moderate autocorrelation levels of 0.5. The observed value of the cubic spline is the (red) line, and the curve was fitted with UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). It can be deduced from the curve that the smoothing parameter (λ) has been chosen by PSM (k=0.04) and has provided the best fit for the simulated time series observation.

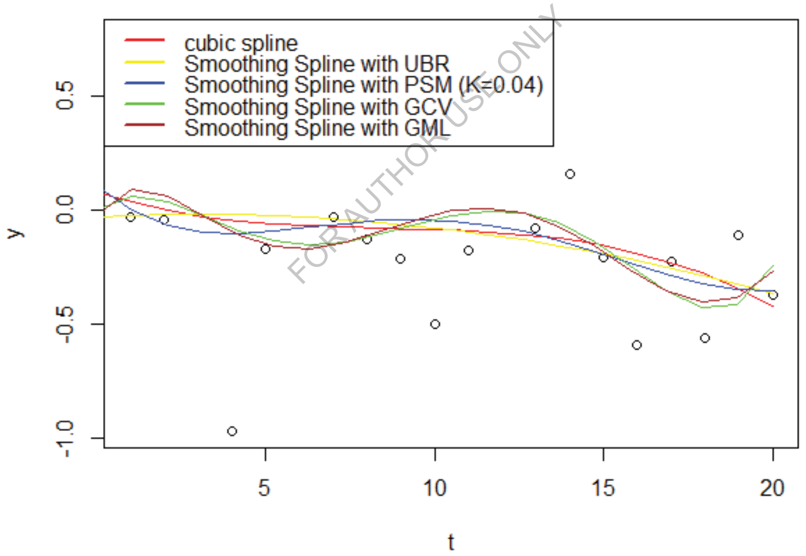


Figure 4.3.2: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 20, GCV, rho = 0.5, sigma =0.8

The smoothing curve presented shown in Figure 4.3.3 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 20, standard deviation = 0.8 and for a high autocorrelation levels of 0.8. The true function provided by the cubic spline (red) line, and four curve estimates with smoothing parameters selected by UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). The estimates by UBR, GML and GCV were wiggling indicating that the estimates were small while the estimate provided by PSM (k=0.04) was closer to the true/observed value, which indicates a good fit.

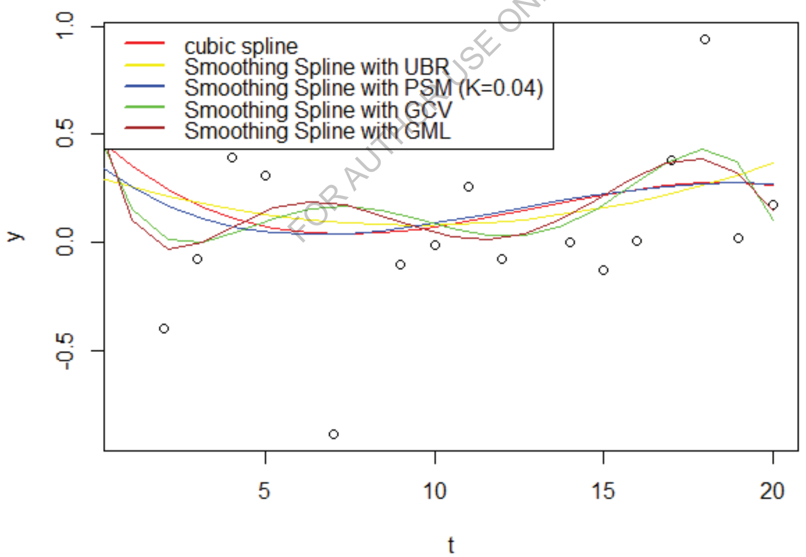


Figure 4.3.3: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 20, rho = 0.8, sigma = 0.8

The smoothing curve presented in Figure 4.3.4 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 60, standard deviation = 0.8 and for a small autocorrelation levels of 0.2. The observed value of the cubic spline is the (red) line, while the curve were fitted with UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). It can be seen from the curve that the smoothing parameter (λ) has been chosen by PSM (k=0.04) and has provided the best fit for the simulated time series observation.

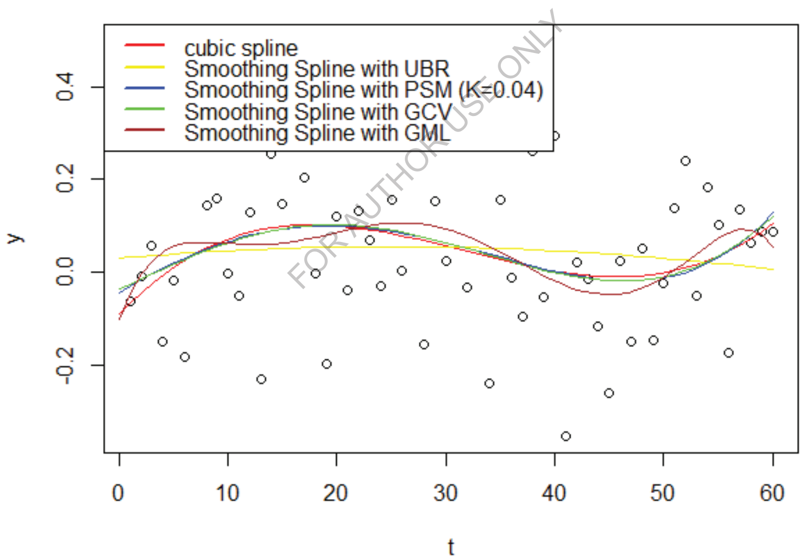


Figure 4.3.4: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 60, rho = 0.2, sigma =0.8

The smoothing curve presented in Figure 4.3.5 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 60, standard deviation = 0.8 and for a moderate autocorrelation levels of 0.5. The observed value of the cubic spline is the (red) line, and the curve was fitted with UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). It can be deduced from the curve that the smoothing parameter (λ) has been chosen by PSM (k=0.04) and has provided the best fit for the simulated time series observations.

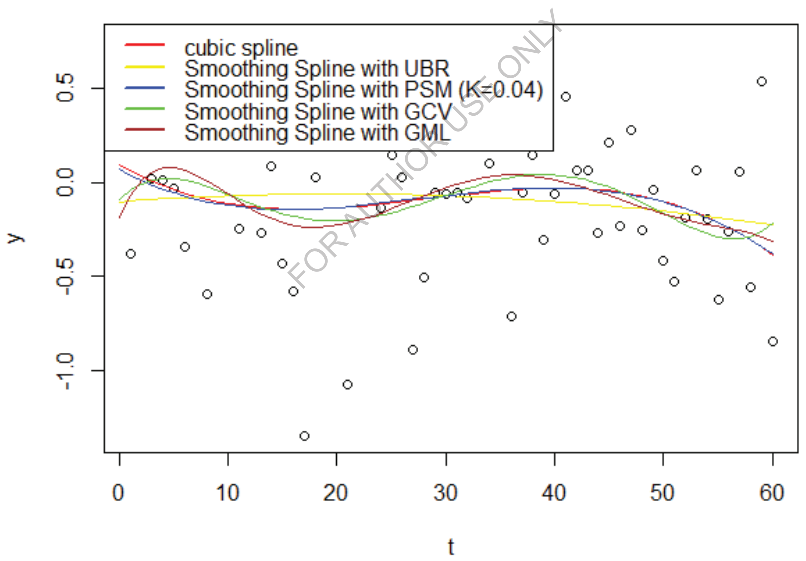


Figure 4.3.5: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 60, rho = 0.5, sigma = 0.8.

The smoothing curve presented shown in Figure 4.3.6 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 60, standard deviation = 0.8 and for a high autocorrelation levels of 0.8. The true function was provided by the cubic spline (red) line, and four curve estimates with smoothing parameters selected by UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). The estimates by UBR, GML and GCV were wiggling indicating that the estimates were small while the estimate provided by PSM (k=0.04) was closer to the true/observed value, which indicates a good fit.

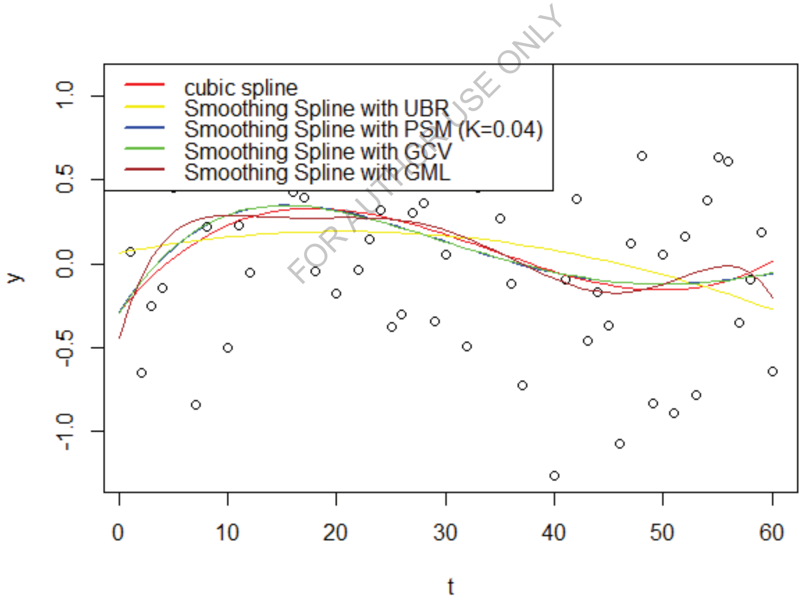


Figure 4.3.6: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 60, rho = 0.8, sigma =0.8

The smoothing curve presented in Figure 4.3.7 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 100, standard deviation = 0.8 and for a small autocorrelation levels of 0.2. The observed value of the cubic spline is the (red) line, while the curve was fitted with UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). It can be seen from the curve that the smoothing parameter (λ) has been chosen by PSM (k=0.04) and has provided the best fit for the simulated time series observations.

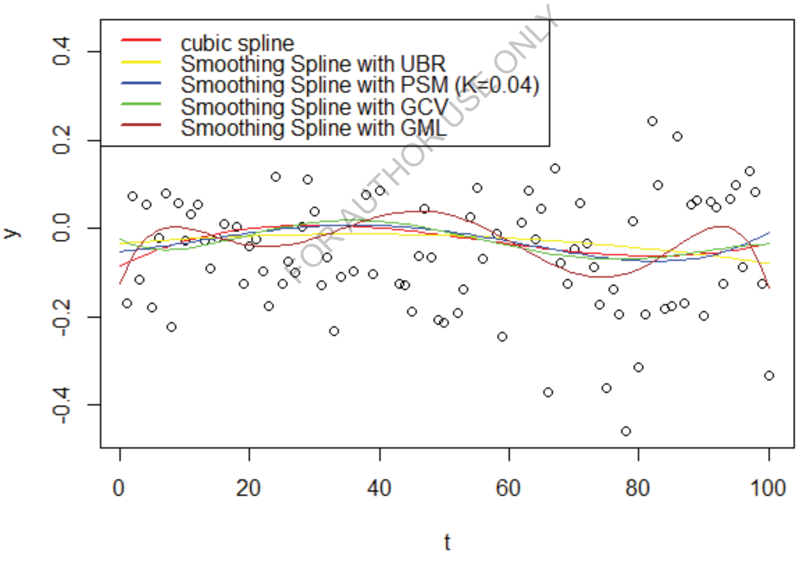


Figure 4.3.7: Plot of the observed and estimated value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 100, rho = 0.2, sigma =0.8

The smoothing curve presented in Figure 4.3.8 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 100, standard deviation = 0.8 and for a moderate autocorrelation levels of 0.5. The observed value of the cubic spline is the (red) line, and the curve was fitted with UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). It can be deduced from the curve that the smoothing parameter (λ) has been chosen by PSM (k=0.04) and has provided the best fit for the simulated time series observations.

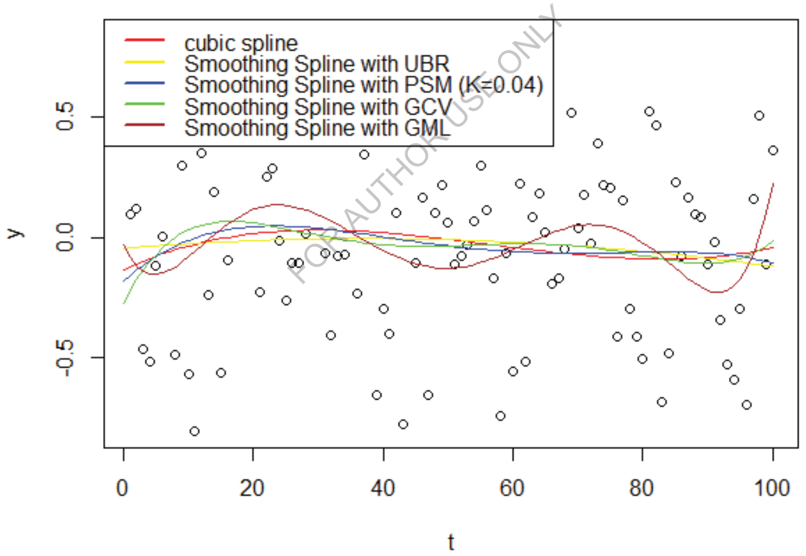


Figure 4.3.8: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 100, rho = 0.5, sigma =0.8

The smoothing curve presented shown in Figure 4.3.9 is the comparison between the four smoothing spline parameter selection methods, that is; GCV, GML, UBR and PSM (K=0.04) for period = 100, standard deviation = 0.8 and for a high autocorrelation levels of 0.8. The true function provided by the cubic spline (red) line, and four curve estimates with smoothing parameters selected by UBR (yellow), PSM (k=0.04) (blue), GCV (green) and GML(brown). The estimates by UBR, GML and GCV were wiggling indicating that the estimates were small while the estimate provided by PSM (k=0.04) was closer to the true/observed value, which indicates a good fit.

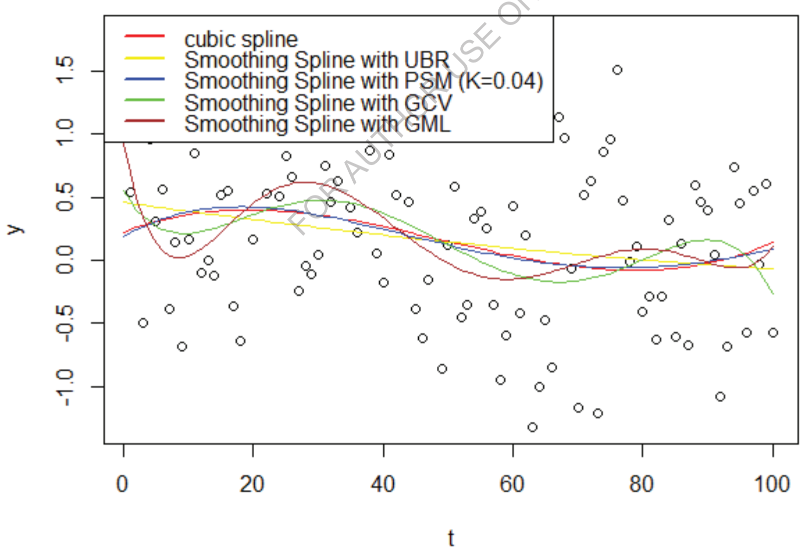


Figure 4.3.9: Plot of the observed and fitted value with smoothing parameter selected using GCV, GML, UBR and PSM (K=0.04) for T = 100, rho = 0.8, sigma =0.8

4.4: Boxplot of the simulated predictive mean squared error of GCV, GML, UBR and PSM (k=0.04)

The charts presented in figures 4.4.1 to 4.4.3 above are the boxplots of the predictive mean squared error estimates for GML, GCV, PSM and UBR when three levels of autocorrelation error exist in the time series model. The plot indicated that PSM (k = 0.04) estimates have the smallest PSMEs compared with GCV, GML and UBR from these plots.

The graphs and plots also showed that time-series observations smoothed by PSM are most efficient and consistent and provides a good fit to the time seies observations at different autocorrelation levels (ρ = 0.2, 0.5 and 0.8) and for small period (T =20). It could be inferred that the entire smoothing techniques estimate the smoothing parameters and the functions remarkably, but PSM (k = 0.04) proved to be a better estimator than GCV, GML and UBR.

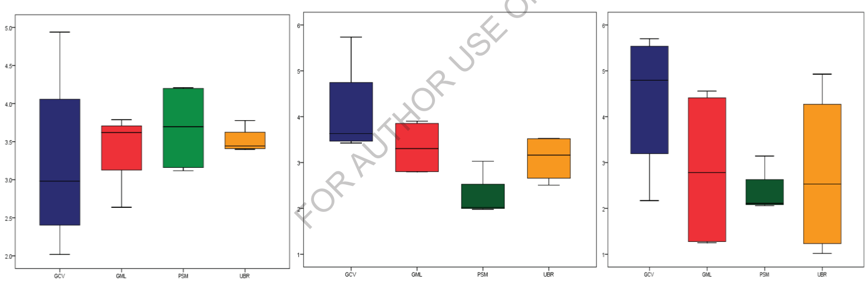


Figure 4.4.1: Box-plot of the simulated GCV, GML, UBR and PSM PMSE values for T=20, rho = 0.2 and sigma =0.8

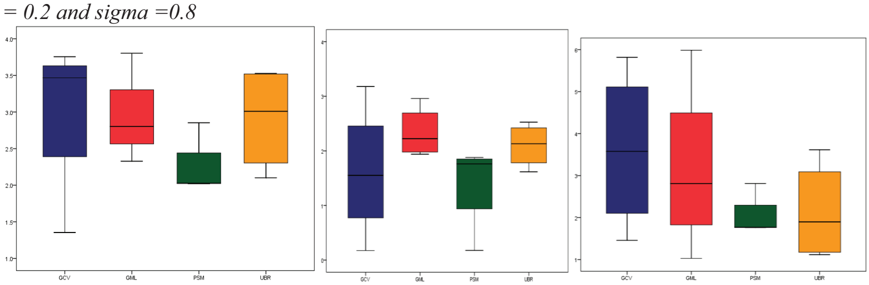


Figure 4.4.2: Box-plot of the simulated GCV, GML, UBR and PSM PMSE values for T=20, rho = 0.5 and sigma =0.8

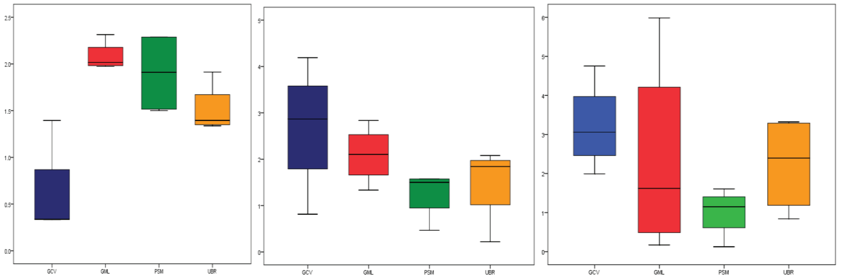


Figure 4.4.3: Box-plot of the simulated GCV, GML, PSM and UBR PMSE values for T=20, rho = 0.8 and sigma =0.8

The plots presented in figures 4.4.4 to 4.4.6 above are the boxplots of the predictive mean square error estimates for GML, GCV, PSM and UBR when three levels of autocorrelation error exist in the time series model. The plot indicated that PSM (k = 0.04) estimates have the smallest PSMEs compared with GCV, GML and UBR from these plots.

The graphs and plots also showed that time-series observations smoothed by PSM are most efficient and consistent and provides a good fit to the time seies observations at different autocorrelation levels (ρ = 0.2, 0.5 and 0.8) and for small period (T =60). It could be inferred that the entire smoothing techniques estimate the smoothing parameters and the functions remarkably, but PSM (k = 0.04) proved to be a better estimator than GCV, GML and UBR.

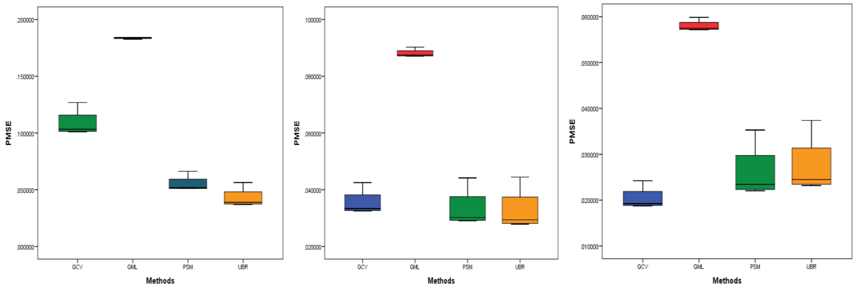


Figure 4.4.4: Box-plot of the simulated GCV, GML, PSM and UBR PMSE values for T=60, rho = 0.2 and sigma =0.8

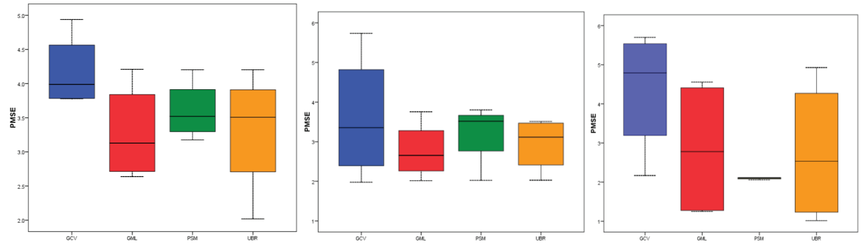


Figure 4.4.5: Box-plot of the simulated GCV, GML, PSM and UBR PMSE values for T=60, rho = 0.5 and sigma =0.8

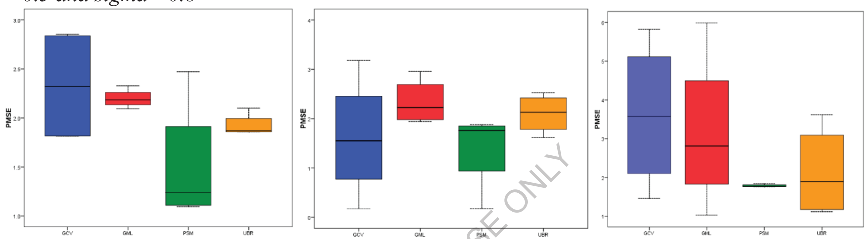


Figure 4.4.6: Box-plot of the simulated GCV, GML, PSM and UBR PMSE values for T=60, rho = 0.8 and sigma =0.8

The plots presented in figures 4.4.7 to 4.4.9 above are the boxplots of the predictive mean square error estimates for GML, GCV, PSM and UBR when three levels of autocorrelation error exist in the time series model. The plot indicated that PSM (k = 0.04) estimates have the smallest PSMEs compared with GCV, GML and UBR from these plots.

The graphs and plots also showed that time-series observations smoothed by PSM are most efficient and consistent and provides a good fit to the time seies observations at different autocorrelation levels (ρ = 0.2, 0.5 and 0.8) and for small period (T = 100). It could be inferred that the entire smoothing techniques estimate the smoothing parameters and the functions remarkably, but PSM (k = 0.04) proved to be a better estimator than GCV, GML and UBR.

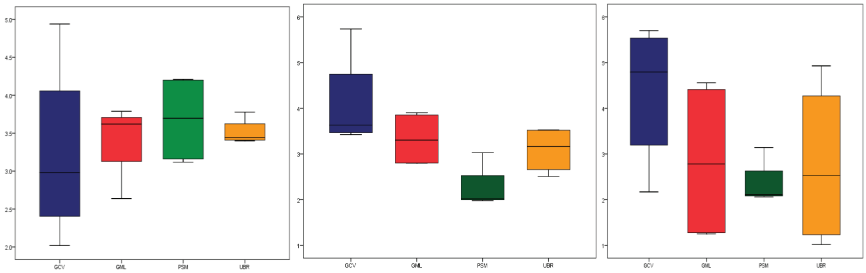


Figure 4.4.7: Box-plot of the simulated GCV, GML, UBR and PSM PMSE values for T=100, rho = 0.2 and sigma =0.8

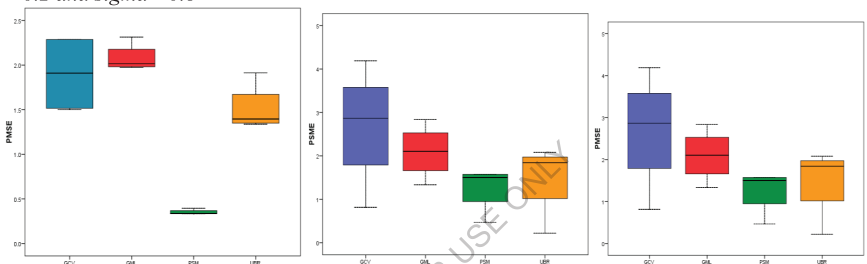


Figure 4.4.8: Box-plot of the simulated GCV, GML, UBR and PSM PMSE values for T=100, rho = 0.5 and sigma =0.8

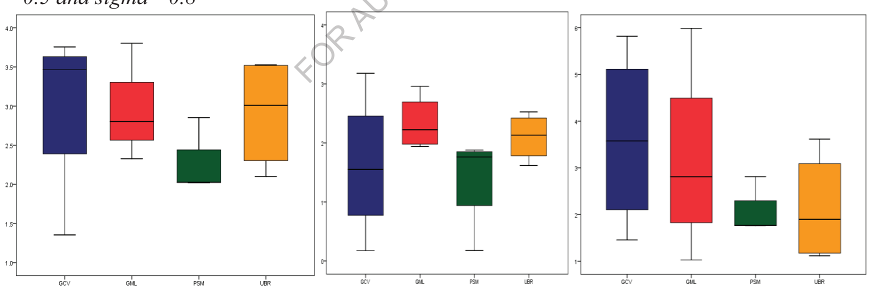


Figure 4.4.9: Box-plot of the simulated GCV, GML, PSM and UBR PMSE values for T=100, rho = 0.8 and sigma =0.8

4.5 APPLICATION OF SMOOTHING METHODS TO REAL LIFE DATA

Table 4.5.1: Test for autocorrelation for the real-life data in the presence of Autocorrelation

Box-Ljung test	
data: Residuals	
X-squared = 96.7395, df = 1,	p-value < 2.2e-16

H_0: The data are independently distributed, or the correlations in the population from which the samples are drawn are zero

H_I: The data are not independently distributed; they have serial correlation

Decision: autocorrelation exist in the model

Table 4.5.2: Stationarity test for the real-life data with smoothing parameters

Augmented Dickey-Fuller Test	
data: Residuals	
Dickey-Fuller = -3.6471,	Lag order = 5, p-value = 0.03021
alternative hypothesis: stationary	

H_0: The observation is not stationary, there exists a unit root

H_I: The observation is stationary, there is no unit root

Decision: The data is stationary, there is no unit root

Table 4.5.3: Cubic Smoothing Spline regression and Predictive Mean Square Error result for Real-life data

Smoothing Methods	β_i	Df	Selection criteria			
			PMSE	Multiple R-Square	Adjusted R Square	Predicted R Square
GML	β_0 = 9.046 β_1 = -1.275 β_2= 5.687 β_3= -9.716	209 209 209 209	42.77	0.5997	0.5940	0.102
GCV	β_0 = 9006 β_1 = -1.230 β_2= 5.677 β_3= -9728	209 209 209 209	45.78	0.6004	0.5947	0.118
UBR	β_0 = 9.162 β_1 = -1.524 β_2= 5.840 β_3= -9.987	209 209 209 209	92.80	0.6014	0.5957	0.099
PSM (K = 0.04)	β_0 = 9.187 β_1 = -1.568 β_2= 5.863 β_3= -1002	209 209 209 209	32.934	0.6021	0.5964	0.4550

Table 4.5.3 above presents the predictive mean square error of the real-life data on the standard international trade classification (SITC) export and import price indices in Nigeria between 2001 – 2018. It was discovered that the proposed smoothing Method (PSM) had the least predictive mean square error (PMSE) a confirmation that it is the preferred smoothing method for simulated and real-life data. The result also presented the multiple, adjusted and predictive R-Square. It can be inferred from the adjusted R-square of the proposed smoothing method of 59.6% that it has the best fit among the four smoothing methods.

The plot above presents the smoothing curve of the annual standard international trade classification import price index dataset in Nigeria from 1970-2018. The data used for analysis were earlier tested for stationarity and autocorrelation. As can be seen from Fig. 4.5.1,. that the proposed smoothing method with optimal smoothing parameter λ = 0.062439908 and for weighted value (k =0.04) was used to carefully analyse the residuals to try to detect disturbances or errors in the stationary part of the series. It was observed that PSM curve is very closely to the real-life data and also provides is good fit.

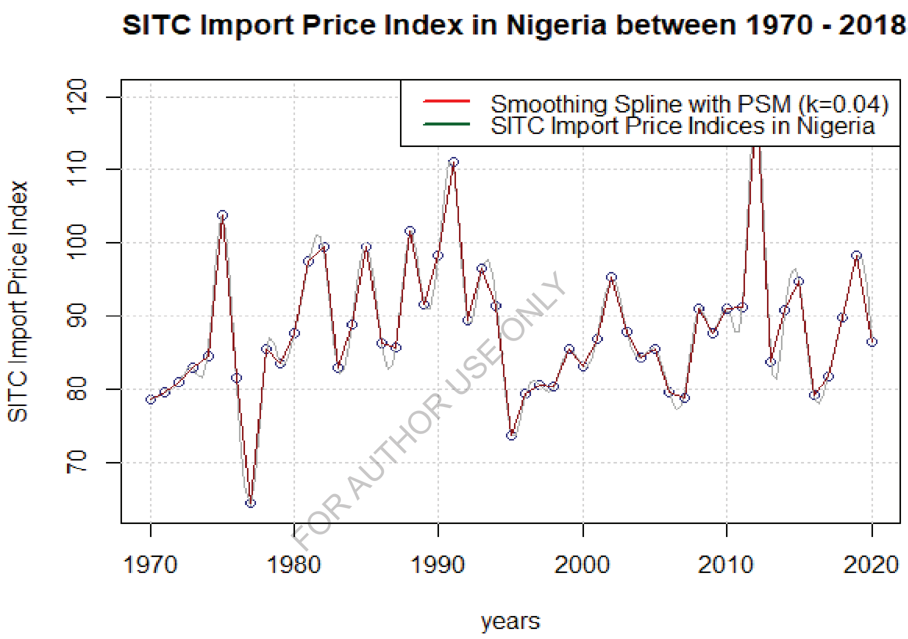


Fig 4.5.1: *Smoothing curve of SITC import price index dataset (darkblue) and fitted value (red) with Smoothing Parameters Chosen by PSM (0.04).*

The plot above presents the smoothing curve of the annual standard international trade classification export price index dataset in Nigeria from 1970-2018. The curve presented ina Fig. 4.5.2,. showd that the proposed smoothing method with optimal smoothing parameter λ = 0.062439908 and for weighted value (k =0.04) was used to smooth the residuals for disturbances or errors in the stationary part of the series. It was observed that PSM curve is very closely to the real-life data and also provides is good fit.

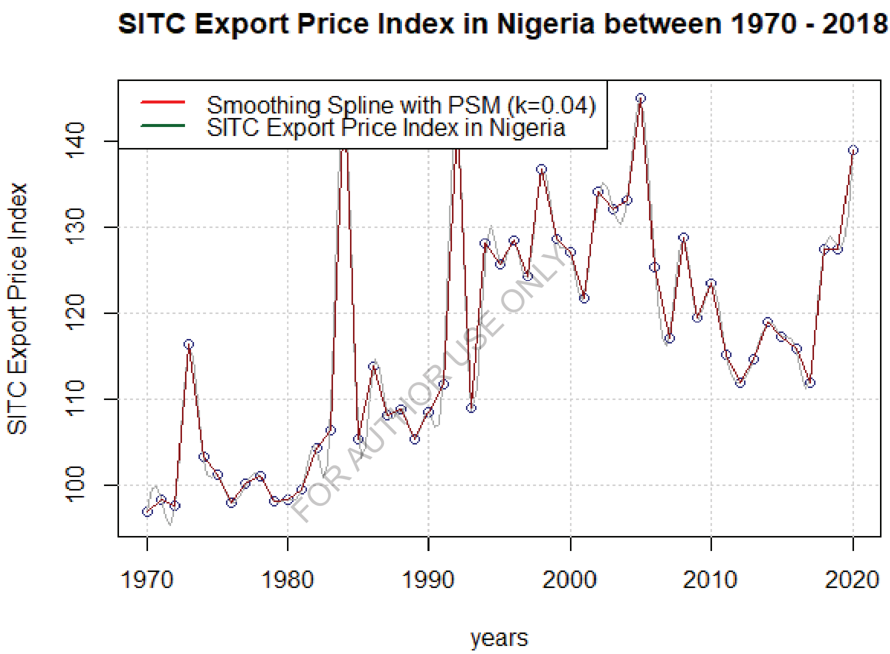


Fig 4.5.2: *Smoothing curve of SITC Export Price Index Dataset (darkblue) and Fitted (red)* with *Smoothing Parameters Chosen by PSM (0.04).*

4.6 SUMMARY OF FINDINGS

The summary of results from the cubic spline non-regression and predictive mean square error (PMSE) tables and plots showed the performances of the smoothing methods, smoothing parameters and time series sample sizes on parameters.

4.6.1 GOODNESS-OF-FIT TEST OF THE SMOOTHING METHODS

Table 4.6.1A: Summary of adjusted R-Square and rank of the smoothing methods at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100)

Time series size	*Smoothing Methods*	*Autocorrelation levels*		
		ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	GML	0.7375(2)	0.7308(2)	0.5871(4)*
	GCV	0.7575(3)	0.5649(1)	0.5816(3)
	UBR	0.6569(1)	0.7517(3)	0.4226(1)
	PSM	0.8983(4)*	0.8505(4)*	0.5711(2)
T = 60	GML	0.9015(2)	0.7254(2)	0.5317(1)
	GCV	0.8819(1)	0.7554(3)	0.5693(3)
	UBR	0.9423(3)	0.7285(1)	0.5384(2)
	PSM	0.9426(4)*	0.7806(4)*	0.6114(4)*
T = 100	GML	0.9373(2)	0.7187(1)	0.5100(1)
	GCV	0.9427(4)*	0.7720(2)	0.5249(2)
	UBR	0.9269(1)	0.7501(3)	0.5795(3)
	PSM	0.9419(3)	0.7973(4)*	0.6747(4)*

* = preferred smoothing method

Table 4.6.1B: Summary of the preferred smoothing methods based on PMSE values at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100)

Time series sizes	*Autocorrelation levels*		
	ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	PSM	PSM	GML
T = 60	PSM	PSM	PSM
T = 100	GCV	PSM	PSM
Overall Assessment	PSM, GCV	PSM	PSM, GML
Most Preferred	**PSM**	**PSM**	**PSM**

Table 4.6.1A and 4.6.1B presents the Summary of the adjusted R-Square, rank and preferred smoothing methods based on PMSE values at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100). It was discovered that PSM provides the best model fit for the time-series observations with autocorrelation in the error term at the three autocorrelation levels, that is; at ρ = 0.2 0.8 and T = 100.

4.6.2 WHEN THERE IS AUTOCORRELATION IN THE ERROR TERM AND TEST FOR OVERFITTING

The summary of results from the predictive mean square error tables and plots showed the performance of the smoothing methods, smoothing parameters, autocorrelation levels and time series sample sizes are presented in Tables 4.6.2A to 4.6.2D.

Table 4.6.2A: Summary of PMSE and rank of the smoothing methods at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100)

Time series sizes	*Smoothing Methods*	*Autocorrelation levels*		
		ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	GML	3.3503(2)	3.2623(3)	3.3739(3)
	GCV	3.6342(4)	4.3212(4)	5.1076(4)
	UBR	3.5545(3)	2.9481(2)	2.2197(2)
	PSM	2.9461(1)*	2.5398(1)*	2.2143(1)*
T = 60	GML	2.1981(4)	2.3362(4)	3.1584(3)
	GCV	1.5108(2)	2.1144(3)	3.6521(4)
	UBR	1.9263(3)	1.7922(2)	2.0416(2)
	PSM	1.1821(1)*	1.6726(1)*	1.1390(1)*
T = 100	GML	2.0795(4)	1.5945(3)	3.1319(3)
	GCV	0.5871(1)*	2.6849(4)	3.2157(4)
	UBR	1.3767(3)	1.2591(1)*	2.0272(2)
	PSM	0.6949(2)	1.3721(2)	0.8519(1)*

* = preferred smoothing method

Table 4.6.2B: Summary of the preferred smoothing methods based on PMSE values at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100)

Time series sizes	*Autocorrelation levels*		
	ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	PSM	PSM	PSM
T = 60	PSM	PSM	PSM
T = 100	GCV	UBR	PSM
Overall Assessment	PSM, GCV	PSM, UBR	PSM
Most Preferred	**PSM**	**PSM**	**PSM**

Table 4.6.2C: Summary of the predictive R-Square difference and rank of the smoothing methods at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100)

Time series sizes	***Smoothing Methods***	***Autocorrelation levels***		
		ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	GML	0.7657(4)	0.7179(2)	0.6802(4)
	GCV	0.7902(3)	0.8064(4)	0.6360(3)
	UBR	0.6954(2)	0.7899(3)	0.6318(2)
	PSM	0.2828(1)*	0.1658(1)*	0.1908(1)*
T = 60	GML	0.9066(2)	0.7315(3)	0.5546(3)
	GCV	0.9357(3)	0.7484(4)	0.5903(4)
	UBR	0.9454(4)	0.7175(2)	0.5590(2)
	PSM	0.2278(1)*	0.3350(1)*	0.2288(1)*
T = 100	GML	0.9386(3)	0.7112(2)	0.5211()
	GCV	0.9425(4)	0.7662(4)	0.5389()
	UBR	0.9302(2)	0.7577(3)	0.5869()
	PSM	0.3434(1)*	0.3144(1)*	0.3424()*

**Smoothing methods that do not overfit time series observation*

Table 4.6.2D: Summary of the smoothing methods that do not overfit time-series observations when ρ = 0.2, 05 and 0.8 and T = 20, 60 and 100

Time series sizes	***Autocorrelation levels***		
	ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	PSM	PSM	PSM
T = 60	PSM	PSM	PSM
T = 100	PSM	PSM	PSM
Overall Assessment	PSM	PSM	PSM
Most Preferred	**PSM**	**PSM**	**PSM**

Tables 4.6.2A-D, presents the summary, rank of the PMSE, predictive R-Square difference from multiple R-Square and smoothing methods that do not overfit time-series observations.

It was discovered from the summary of the predictive mean square error (PMSE) and preferred smoothing methods presented in Table 4.6.2A and 4.6.2B that the proposed smoothing method (PSM) had the least predictive mean square error at three autocorrelation levels and time series sizes except at (ρ = 0.2 and 0.5) and (T=100) where GCV and UBR proved to be more efficient than the other smoothing methods.

Table 4.6.2C and 4.6.2D present the summary of the difference between the multiple R-square and predictive R-Square. The rank of the smoothing methods at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100) were also presented in the tables. It can be inferred from the result that the proposed smoothing method (PSM) had the least difference between the multiple R-square and predictive R-Square at ρ = 0.2, 05 and 0.8 and T = 20, 60 and 100) an indication the proposed smoothing method (PSM) works well at all sample size and does not over fit time-series observations with autocorrelation present in the error term.

4.6.3 OPTIMUM VALUE OF THE PROPOSED SMOOTHING METHOD

Table 4.6.3: Summary of the optimum values of the proposed smoothing methods

Time series sizes	Smoothing degree	ρ = 0.2	ρ = 0.5	ρ = 0.8
T = 20	D.S. = 1	3.49336 (k=0.01)	2.09983 (k=0.13)	2.272321(k=0.25)
	D.S. = 2	2.19728 (k = 0.02)	2.00634 (k=0.14)	0.757980 (k=0.26)
	D.S. = 3	3.147642 (k=0.03)	3.51337 (k=0.15)	3.03644 (k = 0.27)
	D.S. = 4	0.046857 (k=0.04)	1.794844 (k=0.16)	1.483121 (k=0.28)
T = 60	D.S. = 1	1.25185 (k= 0.05)	1.28961 (k=0.17)	1.058281 (k=0.29)
	D.S. = 2	1.060665 (k=0.06)	1.141933 (k=0.18)	0.310471 (k=0.3)
	D.S. = 3	1.349367 (k=0.07)	1.289411 (k=0.19)	1.077814 (k=0.31)
	D.S. = 4	2.416724 (k=0.08)	1.733162 (k=2)	1.06543 (k = 0.32)
T = 100	D.S. = 1	1.057611 (k=0.09)	1.012881 (k=0.21)	1.01515 (k=0.33)
	D.S. = 2	0.277435 (k=0.1)	0.869461 (k=0.22)	0.052643 (k=0.34)
	D.S. = 3	0.237523 (k=0.11)	0.720518 (k=0.23)	0.85785 (k=0.35)
	D.S. = 4	1.207069 (k=0.12)	1.739182 (k=0.24)	0.056745 (k=0.36)

* = optimal value of the smoothing method

The summaries of the predictive mean square error of the proposed smoothing method are presented in Table 4.6.3. The result shows that the optimal value of the proposed smoothing method is at the predictive mean square error (PMSE) value of 0.046857. This result can be seen at the weighted value of k = 0.04 and parameters; T = 20, D.S. = 4 and ρ = 0.2.

4.6.4 EFFECT OF SAMPLE SIZE ON SMOOTHING METHOD

Table 4.6.4A: Summary of PMSE and rank of the smoothing methods at Autocorrelation levels (ρ = 0.2, 05 and 0.8) and time series sizes (T = 20, 60 and 100)

Time series size	*Smoothing Methods*	*Autocorrelation levels*		
		ρ = 0.2	***ρ = 0.5***	***ρ = 0.8***
T =20	GML	3.3503(2)	3.2623(3)	3.3739(3)
	GCV	3.6342(4)	4.3212(4)	5.1076(4)
	UBR	3.5545(3)	2.9481(2)	2.2197(2)
	PSM	2.9461(1)*	2.5398(1)*	2.2143(1)*
T = 60	GML	2.1981(4)	2.3362(4)	3.1584(3)
	GCV	1.5108(2)	2.1144(3)	3.6521(4)
	UBR	1.9263(3)	1.7922(2)	2.0416(2)
	PSM	1.1821(1)*	1.6726(1)*	1.1390(1)*
T = 100	GML	2.0795(4)	1.5945(3)	3.1319(3)
	GCV	0.5871(1)*	2.6849(4)	3.2157(4)
	UBR	1.3767(3)	1.2591(1)*	2.0272(2)
	PSM	0.6949(2)	1.3721(2)	0.8519(1)*

* = preferred smoothing method

Table 4.6.4B: Preferred Smoothing Method Based on Time Series Sizes

T	σ = 0.8			OVERALL ASSESSMENT	MOST PREFERRED
	ρ = 0.2	ρ = 0.5	ρ = 0.8		
20	PSM	PSM	PSM	PSM	**PSM**
60	PSM	PSM	PSM	PSM	**PSM**
100	PSM	PSM	PSM	PSM, GCV, UBR	**PSM**

Table 4.6.4A and 4.6.4B, presented the summary of the predictive mean square error (PMSE), the rank of the smoothing methods and the preferred smoothing Method based on the time series (T = 20, 60 and 100) when standard deviations (σ = 0.8). It was discovered from the result that PSM was the most preferred smoothing method at all-time series size (T = 20, 60 and 100).

4.6.5 REAL LIFE DATA

Table 4.6.5: Rank and preferred smoothing method for the real-life data

Method	GML	GCV	UBR	PSM
PSME (Ranks)	42.77 (2)	45.78 (3)	92.80 (4)	32.93 (1)
Preferred Smoothing method	PSM			

Table 4.6.5 presented the preferred Spline Smoothing Method for the real-life data on all SITC export and import price index in Nigeria between the years 1970 – 2018. It was discovered that PSM performed better than the GCV, GML and UBR.

4.7 DISCUSSION OF FINDINGS

The results generated from the simulation and real-life data conducted in this study has provided great insight on the smoothing method whose model produces the best fit for the time-series observations, the model whose smoothing method does not overfit data, the optimum value of the proposed smoothing method and the performance of the smoothing methods when autocorrelation is present in the error term.

4.7.1: The result on the goodness-of-fit test revealed that the proposed smoothing method had the best fit model among the competing smoothing methods on the simulated and real-life data. The proposed smoothing method's model fitted without any defection and shortcoming under cubic spline functional form with the highest adjusted R-Square of 0.9618, at T = 20, D.S. = 4, $\rho = 0.2$ and the weight value of $k = 0.04$.

4.7.2: The finding on the effect of autocorrelation in the error terms of the four smoothing methods considered in this study showed that the proposed smoothing method (PSM) works well for all levels of autocorrelation *(ρ = 0.2, 0.5 and 0.8)*. It also provided a better estimate, proved to be the most preferred smoothing method than the GML, GCV and UBR and does not overfit a time series observation with autocorrelation in the error term at a predicted R-square of *0.6218*. This result is slightly similar for GML with; Yuedong (1998) and Yuedong et al. (2000) but differs from; Hart and Wehrly (1986), Diggle and Hutchinson (1989), Altman (1990), Herrman, Gasser and Kniep (1992) and Krivobokova and Kauermann (2007).

4.7.3: The study on the optimum value of the Proposed Smoothing Method (PSM) indicated that the smoothing method performs at an optimal level when (k = 0.04) with a predictive mean square error value of 0.046857, multiple R-Square of 0.9678, Adjusted R-Square of 0.9618 and predictive R-square of 0.6428.

4.7.4: The result on the effect of sample size on the performance of the four smoothing methods shows that the proposed smoothing method is computational more efficient, consistent and works well at all sample sizes (T = 20, 60 and 100) Monte-Carlo experiment. The plots and results presented in Table 4.2.1 – 4.2.9 and figures 4.3.1 to 4.3.44 indicated that; GML, GCV and UBR showed signs of inefficiency for all the time series sizes (T = 20, 60 and 100). This finding is quite different from; Yuedong (1998) and Yuedong et al. (2000), and Kim et al. (2009).

4.7.5: The findings from the result also proved the Proposed Smoothing Method (PSM) to be more efficient among the four competing smoothing methods to real-life data. This result disagrees with the finding by Carew et al. (2003).

CHAPTER FIVE

5.0 SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 INTRODUCTION

In non – parametric regression, most of the assumptions like the error term's independence, constant variance and normally distributed data were not always met and discussed in the previous chapters. The performances of four Spline smoothing estimation methods which are: Generalized Cross-Validation, Generalized Maximum Likelihood, Unbiased Risk and the Proposed Smoothing Method, were investigated under the violated assumption of autocorrelation. The model is studied under two major subdivisions as given below:

1. There is the presence of Autocorrelation in the error term of the model.
2. For real-life data.

5.2 SUMMARY

The study was designed to compare the performance of the proposed smoothing method with three existing smoothing methods namely; Generalized Maximum Likelihood, Generalized Cross-Validation and Unbiased Risk based on the effect of sample size and method whose model do not over fit data in the presence of Autocorrelation in the error term. The criteria used to examine and compare the performance of the four spline smoothing selection methods, were the predictive mean squared error, predictive mean squared error and adjusted R-square.

Based on the foregoing, the study was motivated by the fact that the research on spline smoothing estimation methods for time series observations with low, moderate and high autocorrelation level have always resulted in overfitting models at different time series sample sizes.

In order to achieve the aim of the study, the following objectives were raised. First, the study intends to determine the best-fit smoothing method for the time series observation. Secondly, to identify the best smoothing method that does not over-fit time-series data when autocorrelation is present in the error term. Thirdly, to establish the optimum value of the proposed smoothing method, fourthly, to compare using predictive mean squared error criterion, the three existing smoothing methods and the proposed smoothing method in terms of sample size and lastly, test the performance of the proposed smoothing method using real life-data sets.

In setting out to achieve the objectives of the study, there was an extensive review of the literature to provide a theoretical basis for the study. The summary of the reviewed literature showed that, GML over fit and over smooth time series data set when the sample size is medium and large. It was also observed that GCV is the most preferred spline smoothing estimation method for time series observation with auto correlated residual.

There was also an extensive use of four spline smoothing method and three performance criteria for simulated and real-life data. At the end, the following summarizes the findings of the research.

i) PSM provides the best model fit for the time-series observations with autocorrelation in the error term at the three autocorrelation levels, that is; at $\rho = 0.2$ 0.8 and T = 100, while GCV, UBR and GML proved not to be an efficient smoothing methods..

ii) The study equally discovered that, the proposed smoothing method had the least difference between the multiple R-square and predictive R-Square at $\rho = 0.2$, 05 and 0.8 and T = 20, 60 and 100) an indication the proposed smoothing method (PSM) works well at all sample size and does not over fit time-series observations with autocorrelation present in the error term.

iii) The outcome of the study revealed that the optimal value of the proposed smoothing method is at the predictive mean square error value of 0.046857. This result can be seen at the smoothing parameter =, weighted value of k = 0.04 and parameters; T = 20, D.S. = 4 and ρ = 0.2.

iv) The study also revealed that PSM was the most preferred smoothing method at all-time series size (T = 20, 60 and 100).

v) Lastly, the study discovered that, for the real-life data on all SITC export and import price index in Nigeria between the years 1970 – 2018. PSM performed better than the GCV, GML and UBR.

5.3 CONCLUSION

In this study, smoothing spline techniques for time-series observations with auto correlated residual and real-life data were investigated for goodness-of-fit, overfitting problem and sample size.

5.3.1: Goodness-of fit test of the Smoothing methods

The result on the goodness-of-fit test revealed that the proposed smoothing method had the best fit model among the competing smoothing methods on the simulated and real-life data. The proposed smoothing method's model fitted without any defection and shortcoming under cubic spline functional form with the highest adjusted R-Square of 0.9618, at T = 20, D.S. = 4, $\rho = 0.2$ and the weight value of k = 0.04. (see 5.2.1A and 5.2.1B)

5.3.2: Test for over fitting of the smoothing method in the presence of Autocorrelation

It was discovered that the Proposed Smoothing Method (PSM) does not overfit time-series observations when autocorrelation of levels (ρ = 0.2, 0.5 and 0.8) are present in the error term. It was also the most consistent and efficient smoothing method among the four spline smoothing methods considered in this study based on time series size and performance in the presence of Autocorrelation error. The decision was taken because PSM had the least difference between the multiple coefficients of determination (R-square) and predictive R at the three Autocorrelated error levels (ρ = 0.2. 0.5 and 0.8) considered when compared to the other smoothing method especially for medium and large sample size, i.e. T = 60 and 100, (see table 5.2.2C and 5.2.2D). It was also discovered that GML and UBR were mostly affected by the Autocorrelation in residual.

5.3.3: The optimum value of the Proposed Smoothing Method

The optimum value of the Proposed Smoothing Method was at PSM (k = 0.04) with smoothing parameter of 0.062439908, PSME of 0.046857, multiple R-Square of 0.9678, adjusted R-Square of 0.9618, Predicted R-Square of 0.6428, autocorrelation level of 0.2 and at time series size = 20 (see table 4.2.1-4.29). The next in terms of performance, consistency and efficiency in the presence of Autocorrelation is Generalized Cross-Validation (GCV) then Unbiased Risk (UBR) and the least is the Generalized Maximum Likelihood (GML).

5.3.4: Effect of sample size on Smoothing Methods

The result presented in tables 5.2.4B and 5.2.4C showed that GCV and PSM are smoothing methods compared and compete favourably in the presence of Autocorrelation error and for differences in sample sizes with PSM performing slightly better at all-time series sizes (T = 20, 60 and 100). The finding of this research indicated that the Proposed Smoothing Method (PSM) performed better than the other smoothing methods at all-time series (T = 20, 60 and 100) in the existence of Autocorrelation error.

5.3.5: Result for real-life data

The finding in Table 5.2.4D for the real-life data on all SITC export and import price index in Nigeria between 1970 – 2018 **(See Appendix A2)**, indicated that PSM performed better than the GCV, GML and UBR. The simulation result under the finite sampling properties of the PMSE criterion shows that all smoothing methods were consistent and adversely affected by Autocorrelation's presence in the error term.

5.4 RECOMMENDATIONS

This work examines the performances of the classical GCV, UBR, GML and the Proposed Smoothing Method (PSM) when autocorrelation is present in the residual. The results from 1,000 replication revealed that the smoothing method's performance varies when Autocorrelation is present in the error term at two standard deviations, four smoothing parameters and different time series sample sizes. Therefore, it is recommended that;

(i) for a time series observation with autocorrelation error, PSM provides the best-fit smoothing method for the model;

(ii) PSM does not over-fit data at all the Autocorrelation levels considered ($\rho = 0.2, 0.5\ and\ 0.8$);

(iii) the optimum value of the PSM was at the weighted value of 0.04 and smoothing parameter of 0.062439908;

(iv) when there is Autocorrelation in the error term, PSM's performance was better than the competing smoothing methods considered at all-time series sizes (T = 20, 60 and 100);

(v) for econometric data, like the one employed in this study, "the all SITC export and import price index in Nigeria between 1970 – 2018, the proposed smoothing method proved to be the most efficient among the four methods compared.

5.5 CONTRIBUTION TO KNOWLEDGE

This study has contributed to knowledge by examining the effect of Autocorrelation in the error term of time series data and time series size when estimating Spline Smoothing functions, and the study has demonstrated that overfitting of data caused autocorrelation disturbance can be eliminated by applying the Proposed Smoothing Method (PSM)

The following are the specific contributions of this research:

- The study developed an new smoothing spline technique by taking the weighted hybrid of two classical techniques (i.e. GCV and UBR).
- The study developed an improved smoothing spline technique that is more consistent and efficient for time-series observations in the presence of autocorrelation error.
- The study discovered a new Spline Smoothing estimation method that does not overfit data when time-series size is moderate and large.
- This study has contributed to non-parametric Regression discourse by introducing a smoothing parameter to compete favourably with the existing classical smoothing methods like GCV, GML and UBR.

5.6 SUGGESTION FOR FURTHER STUDIES

In this study, situations in which multicollinearity and heteroscedasticity exist in the variables were not considered; future research may consider these conditions.

In this study, the exogenous variables are considered for small-time series periods, i.e. 20, 60 and 100 only; future research may be considered when the time-series sample is large, i.e. above 100.

The effect of assuming a fixed degree of smoothing for the estimation method used in this study (i.e. 1 to 4) research can be conducted for other degrees of smoothing, especially for values below one. This study only considered autocorrelation of order one in the error term; researchers may consider other levels of autocorrelation.

The researcher may also be interested in studying the smoothing methods' performances for negative Autocorrelation level and autocorrelation of order 2, 3, e.t.c. The whole exercise could be repeated using a different simulation equation, smoothing degree, performance criterion and another standard deviation for assessing the performance of the four estimators considered in this study.

REFERENCES

Adams, S.O., M.A. Zubair (2023). On Some Techniques of Selecting Spline Smoothing Parameters for Correlated Dataset with Autocorrelation Structure in the Residual. World Journal of Advanced Research and Reviews, 17(02), 68 – 78. https://doi.org/10.30574/wjarr.2023.17.2.0216

Adams, S.O., Asemota, O.J. (2023). Efficiency of the Proposed Smoothing Method over the Classical Cubic Smoothing Spline Regression Model with Autocorrelated Residual. Journal of Mathematics and Statistics Studies. 4(1), 26 – 40. https://doi.org/10.32996/jmss.2023.4.13

Adams, S.O., Zubair, M.A. & Aiyedun-Olatunde, T. G., (2022). The Role of Some Social Infrastructures Indicators on Economic Growth of Nigeria: A Causality Analysis Approach, Economic Growth and Environment Sustainability (EGNES), 1(2): 41-48. http://doi.org/10.26480/egnes.02.2022.41.48

Adams, S.O., Zubair, M.A., & Azikwe E. (2022). Artificial Neural Network Analysis of Some Selected KDD CUP 99 Dataset for Intrusion Detection, Acta Informatica Malaysia (AIM), 6(2): 50-56. http://doi.org/10.26480/aim.02.2022.50.56

Adams, S.O. & Godwin, S. (2022). Comparative Study of the Error Trend and Seasonal Exponential Smoothing and ARIMA Model Using Covid-19 Death Rate in Nigeria, International Journal of Natural Sciences Research, 10(1): 43-53. http://doi.org/10.184488/63.v10il.3031

Adams, S.O. & Bello, J.O. (2022). Modeling the Effect of Crude Oil Production and Other Factors on Nigeria Economy: An Autoregressive Distributed Lag Approach. Science Archives, 3(1): 72-79. http://dx.doi.org/10.47587/SA.2022.3109

Adams, S.O., Obaromi, A.D, Alumbugu, A.I. (2021). Goodness of Fit test of an Autocorrelated Time Series Cubic Smoothing Spline Model. Journal of the Nigerian Society of Physical Sciences. 3(3): 191-200. https://doi.org/10.46481/jnsps.2021.265

Adams, S.O., Obaromi, A.D., Rauf, I.R. (2021). Modelling the Number of Household Visit to Health Care Centres in Some Nigeria Communities Using Count Data Regression Models. Journal of Biostatistics and Epidemiology. 7(1): 36-47. https://doi.org/10.18502/jbe.v7i1.6293

Adams, S.O., Akano, R.O., Rauf, R.I. (2020). An Evaluation of the Chemical Composition of

Soft drinks in Nigeria: A Principal Component Analysis Approach. Advances in Modelling and Analysis A. 57(4):14-21. https://doi.org/10.18280/ama_a.571-403

Adams, S.O., Balogun, P.O. (2020). Panel Data Analysis on Corporate Effective Tax Rates of Some Listed Large Firms in Nigeria. Dutch Journal of Finance and Management, 4(2):1-9, em0068. ISSN (Online): 2542–4750. https://doi.org/10.21601/djfm/9345

Adams, S.O., Bamanga, M.A., Olanrewaju, S.O., Yahaya, H.U., Akano, R.O. (2020). "Modelling COVID-19 Cases in Nigeria Using Some Selected Count Regression Models". International Journal of Healthcare and Medical Sciences, 6(4):64-73, ISSN (e): 2414-2999, ISSN (p): 2415-5233. https://doi.org/10.32861/ijhms.64.57.63

Adams, S.O., Yahaya, H.U. (2020). Comparative Study of GCV-MCP Hybrid Smoothing Methods for Predicting Time Series Observations. American Journal of Theoretical and Applied Statistics, 9(5): 219-227. https://doi:10.11648/j.ajtas.20200905.15

Adams, S.O., Bamanga, M.A. (2020). Modelling and Forecasting Seasonal Behavior of Rainfall in Abuja, Nigeria; A SARIMA Approach. American Journal of Mathematics and Statistics 10(01):10–19. p-ISSN: 2162-948X, e-ISSN: 2162-8475. https://doi:10.5923/j.ajms.20201001.02

Adams, S.O., Ipinyomi, R.A. (2020). On the Efficiency of the Weighted Generalized Cross Validation and Unbiased Risk Smoothing Method for Time Series Observations with Autocorrelated Error. International Journal of Academic and Applied Research, 04(07): 70-81. ISSN: 2643 – 9603.

Adams, S.O., Bamanga, M.A., Mbusube, A. (2019). Principal Component Analysis of Nigerian Economy from 2006 – 2017. Asian Journal of Advanced Research and Reports, 5(4):1 – 9. Article no. AJARR.51196. https://doi.org/10.9734/ajarr/2019/v5i430139

Adams, S.O., Mustapha, B., Alumbugu, A.I. (2019). Seasonal Autoregressive Integrated Moving Average (SARIMA) model for the analysis of frequency of monthly rainfall in Osun state, Nigeria. Physical Science International Journal, 22(4):1-14. Article no. PSIJ.50701, ISSN: 2348-0130. https://doi.org/10.9734/psij/2019/v22i430139

Adams, S.O., Ipinyomi, R.A. (2019). A New Smoothing Method for Time Series Data in the Presence of Autocorrelated Error. Asian Journal of Probability and Statistics (AJPAS), 04(04): 1-19. Article no. AJPAS.50700, ISSN: 2582-0230. https://doi.org/10.9734/ajpas/2019/v4i430121

Adams, S.O., Ipinyomi, R.A. (2019). A Proposed Spline Smoothing Estimation Method for Time Series Observations. International Journal of Mathematics and Statistics Invention (IJMSI), 07(02): 18-25. e-ISSN: 2321 – 4767, p-ISSN: 2321 – 4759.

Adams, S.O., Yahaya, H.U., Nasiru, M.O. (2017). Smoothing Parameter Estimation of the Generalized Cross Validation and Generalized Maximum Likelihood. IOSR Journal of Mathematics, 13(1): Version V, (January – February 2017). 41 – 44. https://doi:10.9790/5728-1301054144

Adams, S.O., Ipinyomi, R.O., Yahaya, H.U. (2017). Smoothing Spline of ARMA Observations in the Presence Autocorrelation Error. European Journal of Statistics and Probability, (Online), 5(1): 1-8. P-ISSN: 2055-015, e-ISSN: 2055-0162.

Adams, S.O. Awujola., A., Alumbugu, A.I. (2014). Modeling Nigeria's Consumer Price Index Using ARIMA Modeling. International Journal of Development and Economic Sustainability, 2(2): 37 – 47. p-ISSN: 2053-2199, e-ISSN: 2053-2202

Adams, S.O., Olanrewaju, S.O., Oguntade, E.S. (2012). A comparative study of some non-parametric k-independent sample test. Zuma Journal of Pure and Applied Sciences, (9): 35-40.

Adams, S.O., Akano, R.O., Asemota, O.J. (2011). Forecasting Electricity Generation in Nigeria Using Univariate Time Series Model. European Journal of Scientific Research, 58(1): 30 – 37. ISSN 1450-216X.

Adams, S.O., Bamanga, M.A. (2010). On the Parametric F – Test and its Non – Parametric Equivalent. International Journal of Mathematical Sciences, 2(1): 117 – 120. Blackwell Educational Books.

Adams, S.O., Gayawan, E., Garba, M.K. (2009). Empirical Comparison of the Kruskal – Wallis Statistics and its Parametric Counterpart. Journal of modern Mathematics and Statistics, 3(2): 38 – 42. Medwell Journal. ISSN: 1994-5488. https://doi:jmmstat.2009.38.42.

Adji, A.R.F., Janssen, P., Sa'adah, U., Solimun, Effendi, A., Nurjannah & Amaliana, L.

Ajayi, A., Akano, R.O., Adams, S.O. (2019). Impact of Foreign Direct Investment on Employment and Unemployment Rate in Nigeria. Application of Vector Autoregression (VAR) Models (1960-2014). Asian Journal of Advanced Research and Reports, 6(1):1 – 15. Article no. AJARR.51412. https://doi:10.9734/AJARR/2019/v6i130141

Ajayi, A., Adams, S.O., Akano, R.O. (2019). Modelling Nigeria Naira Exchange Rate against some selected Country's Currencies Volatility: Application of GARCH Model. Asian Journal of Probability and Statistics, 5(1): pp.1 – 13. Article no.AJPAS.51360, ISSN: 2582-0230. https://doi:10.9734/AJPAS/2019/v5i130128

Akano, R.O., Adams, S.O. (2019). Nigeria Government's Expenditure on Economic and Social Service Development: A Canonical Correlation Analysis Approach. Archives of Current Research International., 18(04): 1 – 12. Article no. ACRI.51449, ISSN: 2278-7077. https://doi.org/10.9734/acri/2019/v18i430142

Altman, N.S. (1990). Kernel Smoothing of Data with Correlated Errors. *Journal of the American Statistical Association.* 85, 749–759.

Anna, L., Tong, T. & Yuedong, W. (2006). Smoothing Spline Estimation of Variance Function, *Journal of Computational and Graphical Statistics, 16(2), 1 – 10.*

Awujola., A., Adams, S.O., Alumbugu, A.I. (2015). Oil Exportation and Economic Growth in Nigeria. Developing Country Studies, p-ISSN: 2224-607X, e-ISSN: 2225-0565, Vol.5, No.15, 2015. www.iiste.org

Aydin, D. & Memmedli, M. (2011). Optimum Smoothing Parameter Selection for Penalized Least-squares in the Form of Linear Mixed Effect Models. *Optimization*, iFirst: 1–18.

Aydin, D., M. Memmedli, and R. E. Omay. (2013). Smoothing Parameter Selection for Nonparametric Regression using a Smoothing Spline. *European Journal of Pure and Applied Mathematics,* 6, 222–38.

Aydin, D. & Yilmaz, E. (2018). Modified Spline Regression Based on Randomly Right-Censored data: A Comparative Study. *Communications in Statistics - Simulation and Computation*, 47(9), 2587 – 2611.

Beach, C.M. & Mackinnon, J.S. (1978). A Maximum Likelihood Procedure for Regression with Autocorrelated Errors. *Econometrica*, 46, 51-58.

Bera, A.K. & Higgins, M.L. (1993). ARCH Models; Properties, Estimation and Testing. *Journal of Economic Surveys,* Wiley online library, 7(4), 305-366.

Berenlut, I. I. & Webb, G. I. (1974). Experimental Designs in the Presence of Autocorrelated Errors. *Biometrika*, 61, 427-437.

Bollerslev, T. (1986). Generalized Autoregressive Conditionally Heteroskedastics. *Journal of Econometrics*, 31(3), 307 – 327.

Cantoni, E. & Ronchetti, E. (2001). Resistant Selection of Smoothing Parameter for Smoothing

Splines. *Statistics and Computing*, 11(2), 141–146.

Carew, J. D.,Wahba, G., Xie, X., Nordheim, E.V. & Meyerand M. E. (2003). Optimal Spline Smoothing of fMRI Time Series by Generalized Cross-Validation. *NeuroImage*, 18(4), 950 – 961.

Chan, K. S. & Tong, H. (1986). On Estimating Thresholds in Autoregressive Models. *Journal of Time Series Analysis, 7*, 179 – 190.

Chatterjee, S. & Hadi, A.S. (1988). Sensitivity Analysis in Linear Regression, New York, John Wiley and Sons.

Chatterjee, S., Hadi, A.S. & Price, B. (2000). Regression Analysis by Example.Third Edition. A Wiley – Interscience Publication. John Wiley and Sons.

Chen, C.S. & Huang, H.C. (2011): An improved Cp criterion for spline smoothing. *Journal of Statistical Planning and Inference,* 144(1), 445 – 471.

Chi –Lu, C. & Van Ness, J.W (1999). Statistical Regression with Measurement Error. *London, Arnold, World Scientific, 2nd Edition.*

Chipman, J.S., (1979). The efficiency of Least Squares Estimation of Linear Trend when Residuals are Autocorrelated. *ECTRA*, 47, 115-127.

Cook, R. D. & Weisberg, S. (1982). Residuals and Influence in Regression. *Journal of the American Statistical Association*, 86, 328–332.

Craven, P. & Wahba, G. (1979). Smoothing Noisy Data with Spline Functions. *Numerical Mathematics*, 31, 377 – 403.

Cochrane, D. & Orcutt, G.H. (1949). Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms. *Journal of American Statistical Association*, 44, 32 – 61.

Daniel, C. (1973). One at a time Plans. *Journal of American Statistical Association*, 68 (342), 353 – 360.

Devi, A.R., Budiantara, I.N. & Vita-Ratnasari, V. (2018). Unbiased Risk and Cross-Validation Method for Selecting Optimal Knots in Multivariable Nonparametric Regression Spline Truncated (Case Study: The Unemployment Rate in Central Java, Indonesia, 2015). *AIP Conference Proceedings 2021.*

Durbin, J. & Watson, G.S. (1950). Testing for Serial Correlation in the Least Squares Regression I. *Biometrika*, 37, 409-428.

Durbin, J. & Watson, G.J. (1951). Testing for Serial Correlation in the Least Squares Regression II. *Biometrika*, 38, 159-177.

Durbin, J. & Watson, G.S. (1971). Test for Serial Correlation in the Least Squares Regression III. *Biometrika,* 58, 1-42.

Diggle, P.J. & Hutchinson, M.F. (1989). On Spline Smoothing with Autocorrelated Errors. *Australian Journal of Statistics*, 31, 166 –182.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the A variance of UK nation. *Econometrica*, 50, 987 – 1008.

Eilers, P.H.C & Marx, B.D. (1996). Flexible Smoothing with B-splines and penalties. *Statistical Science*, 11, 89-121.

Eubank, R. L. (1988). Spline Smoothing and Nonparametric Regression. *Marcel Dekker, Inc., Basel, New York.*

Francisco-Fernandez, M. & Opsomer, J.D. (2005). Smoothing Parameter Selection Methods for Nonparametric Regression with Spatially Correlated Errors. *Canadian Journal of Statistics*, 33(2), 279–295.

Formy, T. B., Hill, R. C. & Johnson, S. R. (1984). Advanced Econometric Methods*. Springer-Verlag,* New York, Berlin, Heidelberg, London, Paris, Tokyo.

Fuller, W.A. (1987). Measurement Error Models. New York, John Wiley and Sons.

Granger, C. & Terasvirta, T. (1992). Modelling Nonlinear Dynamic Relationships. *Oxford University Press, Oxford.*

Granger, C. & Anderson, A. P. (1978). An Introduction to Bilinear Time Series Models. Vandenhoeck & Ruprecht, *Gottingen & Zurich*.

Green, P.J. & Silverman, B.W. (1994). Nonparametric Regression and Generalized Linear Models, a Roughness Penalty Approach. *Chapman & Hall, London.*

Gu, C. (1992). Diagnostic for Nonparametric Regression Models with Additive Term. *Journal of American Statistical Association,* 87 (420), 1051 – 1058.

Haggan, V. & Ozaki, T. (1981). Modelling Nonlinear Vibrations using an Amplitude Dependent Autoregressive Time Series Model. *Biometrika*, 68, 189 – 196.

Hall, P. & Keilegom, I. (2003). Using Difference-Based Methods for Inference in Nonparametric Regression with Time Series Errors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65 (2), 443 – 456.

Hardle, W. (1991). Applied Nonparametric Regression*, Cambridge University Press, Cambridge.*

Hardle, W., Hall, P. & Marron, J. S. (1988). How far are Automatically Chosen Regression Smoothing Parameters from their Optimum? (With Discussion). *Journal of the American Statistical Association,* 83, 86 – 89.

Hart, J. D. & Wehrly, T. E. (1986). Kernel Regression Estimation using Repeated Measurements Data. *Journal of the American Statistical Association*, 81, 1080 – 1088.

Hart, J.D. (1994). Automated Kernel Smoothing of Dependent Data by Using Time Series Cross-Validation. *Journal of the Royal Statistical Society, (Series* B), 56, 529–542.

Hart, J. D. & Lee, C. (2005). Robustness of One-Sided Cross-Validation to Autocorrelation, *Journal of Multivariate Analysis*, 92, 77–96.

Hastie, T. & Tibshirami, R. (1990). Generalized Additive Models, *Chapman and Hall.* London.

Hastie, T. & Tibshirami, R. (2000). Bayesian Backfitting, (With Discussion). *Statistical Science,* 15, 193–223.

Harvey, A.C. & McAvinchey, I.D. (1978). The Small Sample Efficiency of Two–Step Estimators in Regression Models with Autoregressive Disturbances. Discussion Paper No. 78, University of British Columbia.

Herrmann, E., Gasser, T. & Kneip, A. (1992). Choice of Bandwidth for Kernel Regression When Residuals are Correlated. *Biometrika*, 79(4), 783–795.

Hurvich, C.M. (1990). The Impact of Model Selection on Inference in Linear Regression. *The American Statistician,* 44(3), 214 – 217.

Hurvich, C.M., Simonoff, J.S. & Tsai, C. (1998). Smoothing Parameter Selection in Non-Parametric Regression using an Improved Akaike Information Criterion. *Journal of Royal Statistical Society,* (Series B), 60(2), 271 – 293.

Hurvich, C. M. & Zeger, S. L. (1990). A Frequency Domain Selection Criterion for Regression with Autocorrelated Errors. *Journal of the American Statistical Association, 85*, 705 – 714.

Johnston, J. (1984). Econometric Methods. 3rd Edition, *McGraw-Hill*, New York.

Kadiyala, K.R. (1970). Testing for the independence of Regression Disturbances. *ECTRA*, 38, 71 – 117.

Kim, T., Park, B., Moon, M. & Kim, C. (2009). Using Bimodal Kernel for Inference in

Nonparametric Regression with Correlated Errors. *Journal of Multivariate Analysis,* 100 (7), 1487 – 1497.

Kimeldorf, G. & Wahba, G. (1971). Some Results on Tchebycheffian Spline Function. *Journal of Mathematical Analysis and Application,* 33(1), 82 – 95.

Kleiber, C. (2001). Finite Sample Efficiency of Ordinary Least Square in Linear Regression Models with Long Memory Disturbances. *Econometric Letter*, 72(2), 131-136.

Kramer, W. (1980). Finite Sample Efficiency of Ordinary Least Squares in the Linear Regression Model with Autocorrelated Errors. *Journal of American Statistics Association*, 75, 1005-1009.

Krivobokova T. & Kauermann, G. (2007). A Note on Penalized Spline Smoothing with Correlated Errors. *Journal of the American Statistical Association*, 102, 1328 – 1337.

Kohn, R., Ansley, C. F. and Wong, C. (1992). Nonparametric Spline Regression with Autoregressive Moving Average Errors. *Biometrika,* 79, 44 – 50.

Lee, T.M. & Solo, V. (1999). Bandwidth Selection for Local Linear Regression: A simulation study, *Computational statistics, 14, 515-532.*

Locke, C. & Spurrier, J. D. (1977). The Use of U – Statistics for Testing Normality against Alternatives with Both Tail Light. *Biometrika*, 64, 638-640.

Lukas, M.A., De Hoog, F.R. & Anderssen R.S. (2016). Practical Use of Robust GCV and Modified GCV for Spline Smoothing. *Computational Statistics*. 31(1), 269-289.

Mohammed, T., Yahaya, H.U., Adams, S.O. (2022). Modeling the Volatility for Some Selected Beverages Stock Returns in Nigeria (2012-2021): A GARCH Model Approach. Matrix, Science Mathematic (MSMK), 6(2), 41-50. https://doi.org/10.26480/msmk.02.2022.41.50

Maeshiro, A. (1976). Autoregressive Transformation, Trended Independent Variables and Autocorrelated Disturbance Terms. The *Review Economics and Statistics*, 58(4), 497-500.

Mallows, C.L. (1973). Some Comments on Cp, Technometrics, 15(4), 661 – 675.

Morton, R., Kang, E.L. & Henderson, B.L. (2009). Smoothing Spline for Trend Estimation and Prediction in Time Series. *Environmetrics*, 20(3), 249 – 259.

Morris, J.D. & Lieberman, M.G. (2008). Mallow's Cp for Selecting Best Performing Logistic

Regression Subsets. *Multiple Linear Regression Viewpoints,* 34 (1), 8-13.

Nasiru, M.O., Adams, S.O., Olanrewaju, S.O. (2017). Principal Component Analysis of Crime Data in Gwagwalada Area Command, Abuja from 1995 – 2015. American Journal of Theoretical and Applied Statistics. Vol. 6, No. 1: pp. 38-43. P-ISSN: 2326-8999 e-ISSN: 2326-9006, https://doi:10.11648/j.ajtas.20170601.15

Nason, G.P. & Silverman, B.W. (1997). Wavelet for Regression and other Statistical Problem, *Lecture Note in Statistics, Department of Statistics, University of Bristol, UK*, 1 – 36.

Neter, J. & Wasserman, W. (1974). Applied Linear Model. Richard D. Irwin, Inc

Nicholls, D. F. & Quinn, B. G. (1982). Random Coefficient Autoregressive Models, an Introduction, *Vol. 11 of Lecture Notes in Statistics, Springer-Verlag*, New York.

Nwabueze, J.C. (2000). Estimation of Parameters of Linear Regression Models with Autocorrelated Error Terms. PhD Thesis, University of Ibadan, Nigeria.

Oduntan, E.A. (2004). A Monte Carlo Study of the Problem of Multicollinearity in a Simultaneous Equation Model. An Unpublished Ph.D. Thesis Submitted to the Department of Statistics, University of Ibadan. Ibadan. Nigeria.

Olaomi J.O. (2004). Estimation of the Parameters of Single Linear Regression Model with Autocorrelated Error Terms, which are also Correlated with the Regressor. An Unpublished Ph.D. Thesis, University of Ibadan.

Opsomer J., Yuedong W. & Yang, Y. (2001). Nonparametric Regression with correlated Error, *Statistical Sciences,* 6 (2), 134 – 153.

Papadakis, J.S. (1937): Methode Statistique Pour des Experiencessur Champ. Bull. Inst. Amel. Plantes a Saloniques, 23.

Park, R.E. & Mitchell, B.M. (1980). Estimating the Autocorrelated Error Model with Trended Data. *Journal of Econometrics,* 13, 185-201.

Paris, S. J. & Winstein, C. B. (1954). Trend Estimators and Serial Correlation. Unpublished Cowles Commission, Discussion Paper, Chicago.

Pearson, E.S., Agostino, R.B.D. & Bownnan, K.O. (1977): Test for Departure from Normality: Comparison of Powers. *Biometrika*, 64, 231-246.

Ping, M., Huang, J.Z. & Zhang, N. (2015). Efficient Computation of Smoothing Splines via Adaptive Basis Sampling. *Biometrika*, 102(3), 631–645.

Rao, P. & Griliches, Z. (1969). Small Sample Properties of Several Two-Stage Regression

Methods in the Context of Autocorrelated Errors. Journal of American Statistical Association, 64, 253-272.

Rice, J.A. & Silverman, B.W. (1991). Estimating the Mean and Covariance Structure Nonparametrically when the Data are Curved. Journal of the Royal Statistical Society, 53(1), 233 – 243.

Ruppert, D, Wand, M. P. & Carroll, R J. (2003). Semiparametric Regression. *Cambridge University Press,* Cambridge.

Schimek, G M. (2000). Smoothing and Regression: Approaches, Computation, and Application. *John Willey & Sons, Inc., USA*.

Silverman, B.W. (1984). Spline Smoothing: The Equivalent Variable Kernel Method. *Annal of Statistics,* 12(3), 898 – 916.

Shen, H. (2008). On Modeling and Forecasting Time Series on Smooth Curve, *Technometrics,* 51(3), 227 – 238.

Spiegelhalter, D.J. (1977). A test for Normality Against Symmetric Alternatives. *Biometrika*, 64, 415-418.

Spitzer, J.J. (1979). Small Sample Properties of Nonlinear Least Squares and Maximum Likelihood Estimators in the Contest of Autocorrelation Errors. *JASA*, 74, 126-138.

Stheim, D. (1994). Nonlinear Time Series, a Selective Review. *Scandinavian Journal of Statistics*, 21, 97 – 130.

Subba Rao, T. (1981). On the Theory of Bilinear Time Series Models. *Journal of the Royal Statistical Society*, (Series B) 43, 244 – 255.

Subba Rao, T. & Gabr, M. M. (1980). A Test for Non-Linearity of Stationary Time Series. *Time Series Analysis*, (1), 145-158.

Tong, H. (1983). Threshold Models in Nonlinear Time Series Analysis. *Vol. 21 of Lecture Notes in Statistics, Springer-Verlag, Heidelberg.*

Tong, H. (1978). Nonlinear Time Series Analysis: A Dynamic Approach. *Oxford University Press, Oxford.*

Wang, H., Meyer M. C. & Opsomer J.D. (2013). Constrained Spline Regression in the Presence of AR(p) Errors. *Journal of Nonparametric Statistics,* 25, 809 – 827.

Wahba, G. (1975). Optimal Convergence Properties of Variable Knot Kernel and Orthogonal Series Methods for Density Estimation. *Annals of Statistics*, 3, 15 – 29.

Wahba, G. (1977). Applications of Statistics in P. Krishnaiah Edition. A Survey of Some Smoothing Problems and the Method of Generalized Cross-Validation for Solving them, *Northern Holland, Amsterdam.*

Wahba, G. (1979). Convergence Rates of Thin Plate Smoothing Splines when the Data are Noisy in T. Gasser and M. Rosenblatt. *Smoothing Techniques For Curve Estimation,* Springer-Verlag, New York.

Wahba, G. (1980). Automatic Smoothing of the Log Periodogram. Journal of the American Statistical Association, 75, 122 – 132.

Wahba, G. (1983). Bayesian Confidence Intervals for the Cross Validated Smoothing Spline. Journal of Royal Statistical Society (Series B), 45, 133-150.

Wahba (1985). A Comparison of GCV and GML for Choosing the Smoothing Parameters in the Generalized Spline Smoothing Problem. The Annals of Statistics, 4, 1378 – 1402.

Wahba, G. (1990). Spline Models for Observational Data. *CBMS-NSF Regional Conference Series in Applied Mathematics*, (59), Philadelphia: SIAM.

Wahba, G. & Wang, Y**.** (1993). Behaviour Near Zero of the Distribution of GCV Smoothing Parameter Estimates for Splines. *Statistics and Probability Letters*, (25), 105 – 111.

Wahba, G., Wang, Y., Gu, C., Klein, R., and Klein, B. (1995). Smoothing Spline ANOVA for Exponential Families. With Application to the Wisconsin Epidemiological Study of Diabetic Retinopathy. *The Annals of Statistics*, (23), 1865 – 1895.

White, K.J. (1978). A General Computer Program for Econometric Methods SHAZAM. Econometrica, 46, 239-240.

Whittaker, E.T. (1923). On a New Method of Graduation. *Proceedings of the Edinburgh Mathematical Society*, 41, 63-73.

Williams, R.M. (1952). Experimental Designs for Serially Correlated Observations. *Biometrika*, 39, 151-167.

Worsley, K.J., & Friston, K.J. (1995). Analysis of fMRI Time-Series Revisited Again. *NeuroImage,* 2:173-181.

Xiang, D. & Wahba, G. (1998). Approximate Smoothing Spline Methods for Large Data Sets in the Binary Case. *Proceedings of the 1997 ASA Joint Statistical Meetings, Biometrics Section* 94-98.

Xu, L. & Zhou, J. (2019). A Model-Averaging Approach for Smoothing Spline Regression.

Communications in Statistics - Simulation and Computation. 48(8), 2438 – 2451.

Yahaya, H. U., Oyinloye, J. S., & Adams, S. O. (2022). Modeling and Forecasting Cryptocurrency Returns and Volatility: An Application of GARCH Models. Universal Journal of Finance and Economics, 2(1), 71–90. https://doi.org/10.31586/ujfe.2022.497

Yanrong C, Tracy Z W, Haiqun L, & Yan Y. (2010). Penalized Spline Estimation for Functional Coefficient Regression Models. *Computational Statistics and Data Analysis*, 54, 891–905.

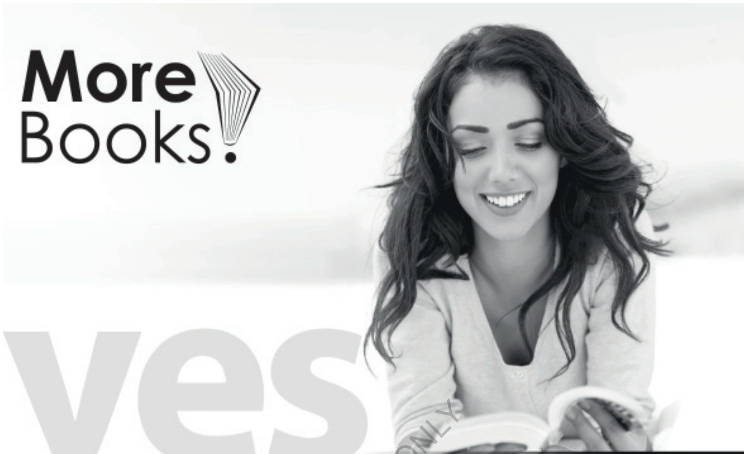
Yuedong, W., Wensheng G. & Brown, M.B. (2000). Spline Smoothing for Bivariate Data with Application to Association Between Hormones, *Statistica Sinica,* 10, 377 – 397.

Yuedong, W. & Brown, M. B. (1996). A Flexible Model for Human Circadian Rhythms, *Biometric.* (52), 588 – 596.

Yuedong, W. (1998). Smoothing Spline Models with Correlated Random Errors. *Journal of American Statistical Association.* (93) 441, 341 – 348.

Yuedong, W. (2011). Smoothing Spline: Methods and Applications. *Technometrics*, 54(3), 324-326

Zubair, M.A., Adams, S.O., Aniagolu, K.S. (2022). Economic Impact of Some Determinant Factors of Nigerian Inflation Rate. Universal Journal of Finance and Economics, 1(1): 51 – 72. https://doi.org/10.31586/ujfe2022.208


More
Books!