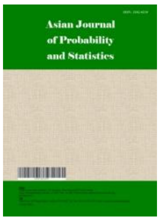



**Asian Journal of Probability and Statistics**

**18(4): 36-45, 2022; Article no.AJPAS.88947**
***ISSN: 2582-0230***

# On the Nonparametric Approach to Estimation of Non- Constant Variance Function: An Application to Nairobi Securities Exchange (NSE)

**Peter Mwangi [a*]**

[a] *KPMG Advisory Services Limited, Kenya.*

***Author's contribution***

*The sole author designed, analysed, interpreted and prepared the manuscript.*

***Article Information***

DOI: 10.9734/AJPAS/2022/v18i430455

**Open Peer Review History:**
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/88947

***Original Research Article***

***Received 27 May 2022***
***Accepted 27 July 2022***
***Published 03 August 2022***

## Abstract

Methods for estimating regression models to data in the areas showing varying variances is considered. The centre of attention is on diverse methods of evaluating varying variances. The nonparametric approach which incorporates the smoothing methods and the choice of the ideal bandwidth is discussed. Normally, the cardinal shortcoming which is of interest is the selection of the smoothing method and picking of the best bandwidth [1], Zhai, C. and Lafferty, J. [2]. The two oftenly used smoothing methods; the Gaussian Kernel and Spline are compared. The two smoothing techniques are illustrated and compared using data obtained from Nairobi Securities exchange (NSE) and found that the Gaussian Kernel outperforms the Spline smoother since it gives the best estimate of the variance.

*Keywords: Kernel smoother; spline; smoothing; bandwidth.*

## 1 Introduction

The breakdown of fixed exchange rate system has sharply increased financial risks in financial institutions. Recent finance disasters in trade portfolios like the national bank of Kenya for example, have underlined the need for accurate financial risk measures in institutions such as banks and investment firms. The nature of financial risks has changed with time and therefore the method to measure them must adapt to recent

**Corresponding author: Email: petermwangi2004.pmg@gmail.com;*

experience. It is in this context that quantitative measures have become vital in the management for both internal and external requirements parallel with others models of returns.

Due to globalization, which has resulted to a fast financial world, there is motivation to develop efficient and effective risk measures which will respond to news just like the other forecasts and must be easy to understand even when the situation is complex. Despite the simplicity of the risk, its management has remained a challenging statistical problem partly because it depends on the joint distribution of the portifolio returns which typically changes overtime. It is for this reason that we estimate the variance function of shares volume of Nairobi stock exchange market. The purpose of the paper is to provide financial managers and shareholders with a non-technical and flexible model for market-to-market reporting.

The study has a variance model that will help financial managers and shareholders in the following ways: - Information reporting, Resource allocation and Performance evaluation.

Variance function is essential for modifying performance for insecurity in business. This is important in an environment where there is large volume of trading and traders have to take on extra risk. The study of nonparametric approach to estimation of non- constant variance function provides encouragement to traders.

The way the variance function should be described in many applications is not clear. According to [3] the likelihood of drawing wrong conlusions exist if the estimating model is not correctly specified. Observation should be left uncorrelated even in a situation where the mean of the non-parametric is left unspecified [4].

## 1.1 The model

Let's consider the bivariate data $(X_i, Y_i)$ , $i = 1, \dots, n$ where the random variable $X_i$ is from uniform distribution and $Y_i = m(X_i) + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2)$. The variable X is distributed over the unit interval since it's a uniform distribution and the errors term $\varepsilon_i$ follows a standard normal distribution with mean zero (0) and variance $(\sigma^2)$. The assumption also is that the errors terms and the X-variate are uncorrelated that is and the design points are assumed to be mutually uncorrelated.

**Model 1 (Parametric regression model)**

$Y_i = f(x_i, \beta) + \varepsilon_i$ , where $f(.)$ is a known function, $\beta$ is the unknown parameter to be estimated, and the errors terms $\varepsilon_i$ are (i.i.d), such that $E[\varepsilon_i] = 0$ and $E[\varepsilon_i^2] = \sigma^2 > 0 \quad \forall_i$ for constant variance.

**Model 2 (Non-parametric regression model)**

Let $Y_i = m(X_i) + \varepsilon_i$, where $m(.)$ is unknown function to be estimated and the errors terms $\varepsilon_i$ are independent and identically distributed and satisfying and satisfying the conditions $E[\varepsilon_i] = 0$ and $E[\varepsilon_i^2] = \sigma_i^2 > 0$ (non-constant variance).

The regression curve $f(x) = 1 - x + e^{-100\left(x - \frac{1}{2}\right)^2}$ and the structures of the variance function (variance of the observations is a function of the mean).

**Case 1**

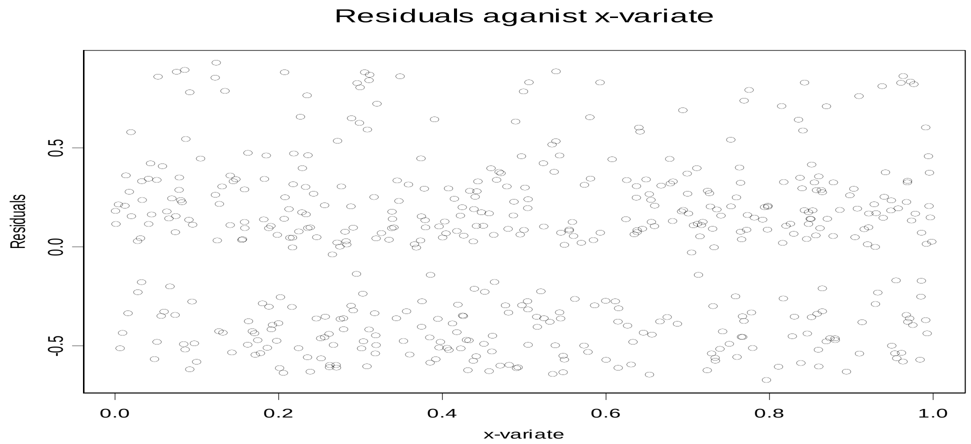
$$E[\varepsilon_i] \quad = var(Y_i) = var\big(f(x_i)\big) = \sigma^2\big(f(x_i)\big) \ or \ \sigma_i^2 = m(X_i) + 5$$

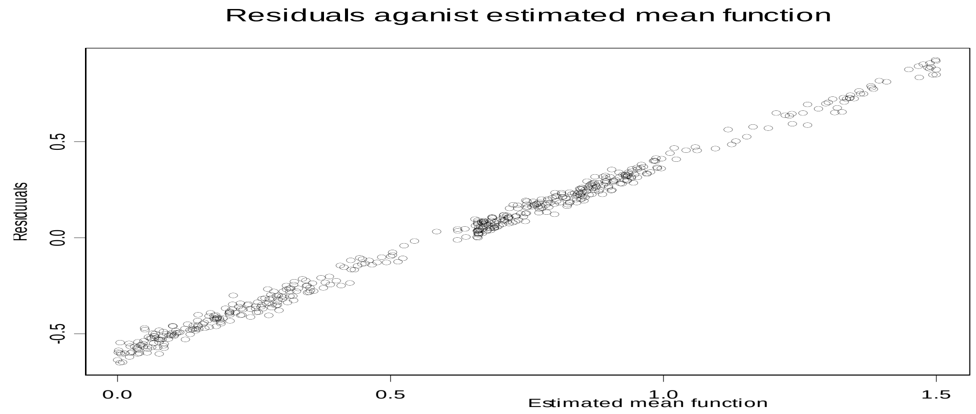
**Case 2**

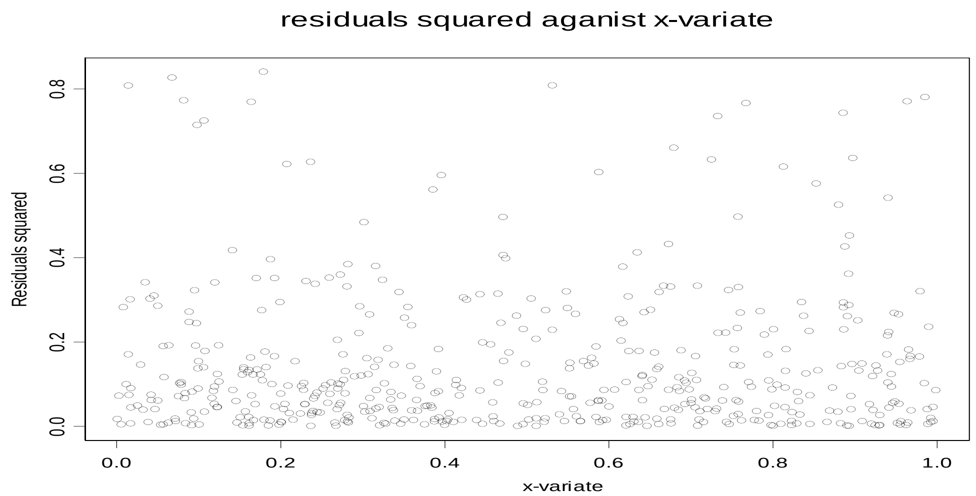
$$\sigma_i^2 = X_i^2 + 5$$

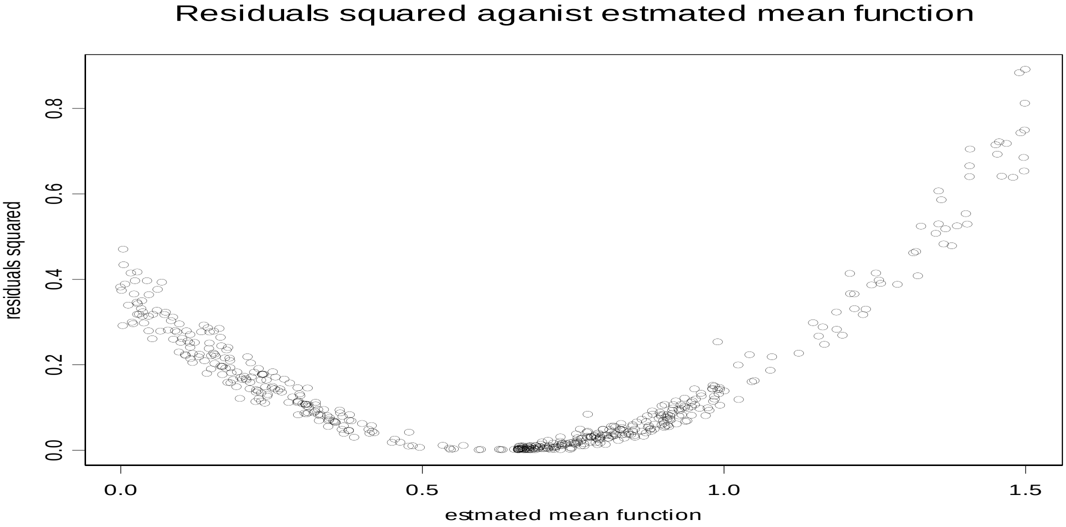
In case 1, the variance function is mean dependent, where it depends on the design points in case 2.

# Data Simulation Using Case 1 and 2 and Scatter Plots of Residuals aganist X-Variate and Residuals aganist $\hat{m}(x)$






# 2 Concept of Smoothing

Let f(.) be a smooth function at x, then it is known that the observed points at $x_i$, neighboring x must have details about the value of f(.) at x. Therefore, the estimator for f(x) can be constructed using the neighborhood mean of data adjacent to x. Levelling or smoothing of a data set $(X_i, Y_i)$, $i = 1, \ldots, n$ calls for the the estimation of the expected response curve f(.) in the regression equation where the regression curve is the function of concern Cai, Z., J ̈onsson, P., Jin, H., and Eklundh, L. [5].

In the inconsequential case in which f(.) is reliable, approximation of f(.) tends to the point of locus. In most situations and realistic research, it is improbable that the regression curve is continuous. Therefore, the assumption which is taken in consideration is that for a smooth constant function, the curve is customized as a smooth continuous function in small locality about x [6]. To investigating whether a two-dimensional scatter plot of a regression curve is provisionally constant is not easy. Therefore, an estimation of the regression curve f(.) will therefore be a point acting as a substitute of point near to the middle of this loop of outcome variables [3].

The localized mean process can be seen as the fundamental objective of smoothing. The procedure is illustrated as

$$\hat{f}(x) = n^{-1} \sum_{i=1}^{n} w_i(x) y_i,$$

Where $\{w_i(x)\}$, i = 1.,..,n represents aa arrangement of weights which may rely on the whole vector $\{x_i\}$ ,i= 1, 2,…,n. Every levelling or smoothing method to be illustrated in this work is closely coinciding with of the form $\hat{f}(x) = n^{-1} \sum_{i=1}^{n} w_i(x) y_i$. Most often the estimate $\hat{f}(x)$ is mostly called the smoother and the aftermath of smoothing process is referred to as the smooth.

The levelling parameter is the one that shapes the weights sequence $\{w_i(x)\}$, i = 1,..,n, since the smoother averages on data with different means. To greater extent the levelling parameter synchronizes the size of the locality near x implying that a local mean over too large a locality would destroy away the good with the bad [7]. Practically in most circumstances an utmost over-levelled curve would result to favored estimate $\hat{f}(x)$. Looking for a process that helps to locate the best levelling parameter that brings equilibrium between over levelling and under levelling is called the smoothing parameter selection problem [8].

# 3 Kernel and Spline Smoothers

## 3.1 Kernel function K (.)

The search for an ideal bandwidth for a given levelling parameter is one of the vital studies in the areas of estimation in statistics recently. A number of methods are available in literature even though none of them is entirely adequate. In this study we set forth to usually used methods of leveling in a data set; Kernel and spline smoothers.

1-dimensional Kernel function takes the form$K_b(x - X_i) = \frac{1}{b} K\left(\frac{x-X_i}{b}\right) = K(u)$.

It is a type of local smoother which assigns weights to the observations $X_i$. The weights decrease with the distance between the point of estimation x and $X_i$, i=1, 2, 3…n.

Various forms of kernel form of kernel exist including uniform, triangle, Epenechnikov, Bisquare and Gaussian among others. Among these kernel function, Gaussian has infinite support while all the others are bounded in [-1,1] (Carrlo, R.J and Ruppert, D 1987). In this study we use the Gaussian kernel which takes the form K(u) $= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}u^2\right]$ where $u = \frac{x-X_i}{b}$. The parameter b is called the bandwidth which determine how large neighbourhood of the target point x, is used in estimation. A large value of b (bandwidth) generates a smooth curve but with a high possibility of obscuring the interesting structures. A very small bandwidth generates a wigglier curve.

In both applied and theoretical study, the choice of the Kernels is restricted. Considering the bivariate data $(x_i, y_i)_{i=1}^n$, the smoothed value $\hat{m}(.)$ produced by a Kernel function K(.) can be given as

$$f(.) = \frac{n^{-1}\sum_{I=1}^{N} K\left(\frac{x-x_i}{b}\right)Y}{n^{-1}\sum_{i=1}^{N}\left(\frac{x-x_i}{b}\right)} \quad , \quad 0 < x, x_i < 1, \ i = 1,...,N$$

[9]

## 3.2 Cubic spline

A usual test of "loyalty to the data" for a curve f is the residuals sum of squares $\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2$if f is permitted to be any curve open in functional form. Through the interpolation of data, the distant measure can be reduced to nil by any function f. The spline levelling perspective evades this unconceivable interpolation of the data by appraising the contest between the focus to generating a curve without too much expeditious local variation (Carrlo, R.J and Ruppert, D 1987). A number of methods exist for appraising the local disparities. One of the main ways is to express the test of roughness, for example, on the first, second and subsequent derivatives. The roughness penalty which is expressed by

$$\int\left(f''(x)\right)^2 dx$$

is used here to justify local disparities. The weighted sum is therefore expressed by

$$s_r(f) = \sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2 + r\int\left(f''(x)\right)^2 dx$$

Where r expresses a levelling/smoothing parameter. The smoothing parameter $r$ represents the rate of change between deviation errors and roughness of the curve f. The problem of minimizing $s_r(.)$ on the group of all double differentiable functions on the interval $[x_{(1)}, x_{(n)}]$ with a distinctive result $\hat{m}_\lambda(x)$ which is described as the cubic spline. The data points to put into consideration are adjacent to x since y data points far away from x will have, in basically very different expected values. In obtaining smoothed curve the selection of the spline function is not enough but rather the consideration of the bandwidth b is equally very important [10].

The comparison of the Kernel and Spline leveling parameters is illustrated on generated data set. There exist diverse forms of Kernel functions in literature which are easier to use. The most important concept to consider is the correct amount of smoothing to use. The main difficult in smoothing is the choice of best bandwidth that justifies the preference to lessen the variability of the estimator and yet takes into account important little features in the affected distribution (which requires a restricted bandwidth). Every leveling method has to be accommodated by some smoothing parameter which stabilizes the degree of correctness to the data against the flatness of the best estimated curve. A choice of the smoothing parameter has to be made in practice and controls the performance of the estimators. The important thing to note in the study is that the beneficially of a nonparametric levelling method should be careful that the last result about an assessed regression curve is partially patented since even essential ideal smoothers has a certain amount of noise that leaves space for patented discernment. This therefore means that for one to make ideal decision about the about the best smoother, then one has to have a a computer with proper statistical packages. To show the optimal function we consider the Gaussian Kernel which is given by

$$\mathrm{K(x)} = \frac{1}{\sqrt{2\pi}} \exp\left[ -\frac{1}{2} x^2 \right] \qquad\qquad |x| < 1$$

Since several methods of acquiring the best bandwidth are in the literature, the selection for the best bandwidth in this study is to present several plots from simulated data set and then come up with the bandwidth which is the most outstanding amongst others. The selection criterion here has been made fast and easy by using computer software S-plus where we have come up with a program which will perform this task. In our investigation here we compare the techniques of smoothing by using data obtained from Nairobi stock exchange. The choice of the best levelling parameter is therefore arrived at by assessing the curves plotted using this program and hence choosing the one that adequately fits the data (that is the one which is optimal). In this data set, varying smoothing parameters is done until the optimal one is found. The optimal smoothing parameter that is obtained here can be applied to any data set.

# 4 Choice of Levelling/Smoothing Parameter

In non-parametric regression the trouble of resolving the extent of how much to smooth is vital. Hence in this section we will be focusing on finding a good way of choosing the smoothing parameter of various smoothing methods. The qualifications for a good bandwidth selection criterion is that it should have conceptual beneficial properties and it should relevant in use. The conceptual beneficial properties have been proposed and it assess how the estimation function is to the close to the curve.

Description of the smooth is what informs the choice of the smoothing technique [11]. If the main reason of levelling in a given set of data is to increase the ‘’wave to noise ratio” for demonstrating or to recommend a simple regression model, then a little “over levelled” curve with a distinctive chosen leveling parameter might be advisable. A rather under-levelled curve would be advisable if the attention is merely in approximating the regression curve itself. On the other hand, a well programmed and produced parameter is of great use beginning with. One of the merit of a computer generated for the Kernel smother rest in the use An advantage of automatic selection of the bandwidth for kernel smoothers is that comparison between laboratories can be made on the basis of a standardized method. Another advantage of the same lies in the application of supplementary models for exploring of more complicated regression data.

We compare two smoothing techniques using simulated data sets.

Some of the smoothing techniques are;

(i) Kernel smoothing technique
(ii) Spline smoothing technique

**Kernel smoothing technique**

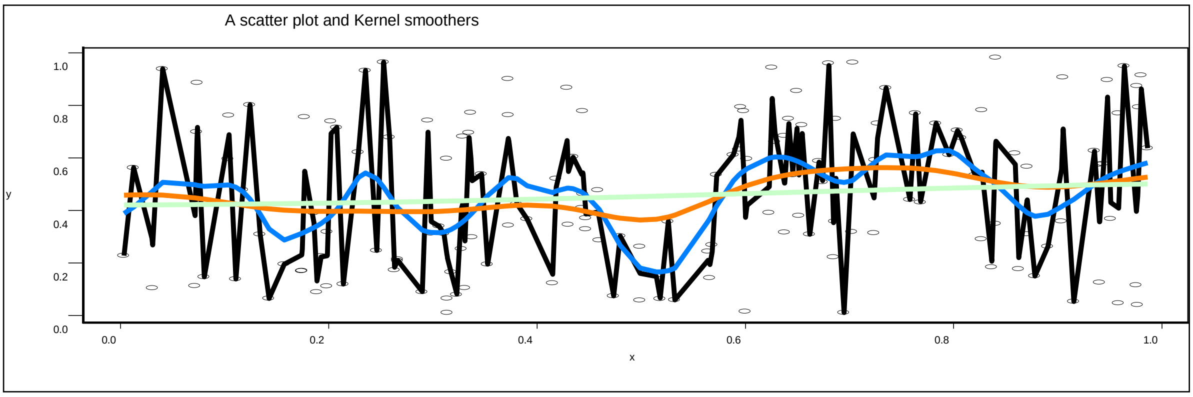



**Fig. 1. Shows 100 generated data of data points with a Kernel smoothing method with leveling parameters: b = 0.005, 0.185, 0.4 and 0.8. The blue curve whose leveling parameter b=0.185, gives a better smooth than the rest as the best smoothing parameters**

**Spline smoothing method**

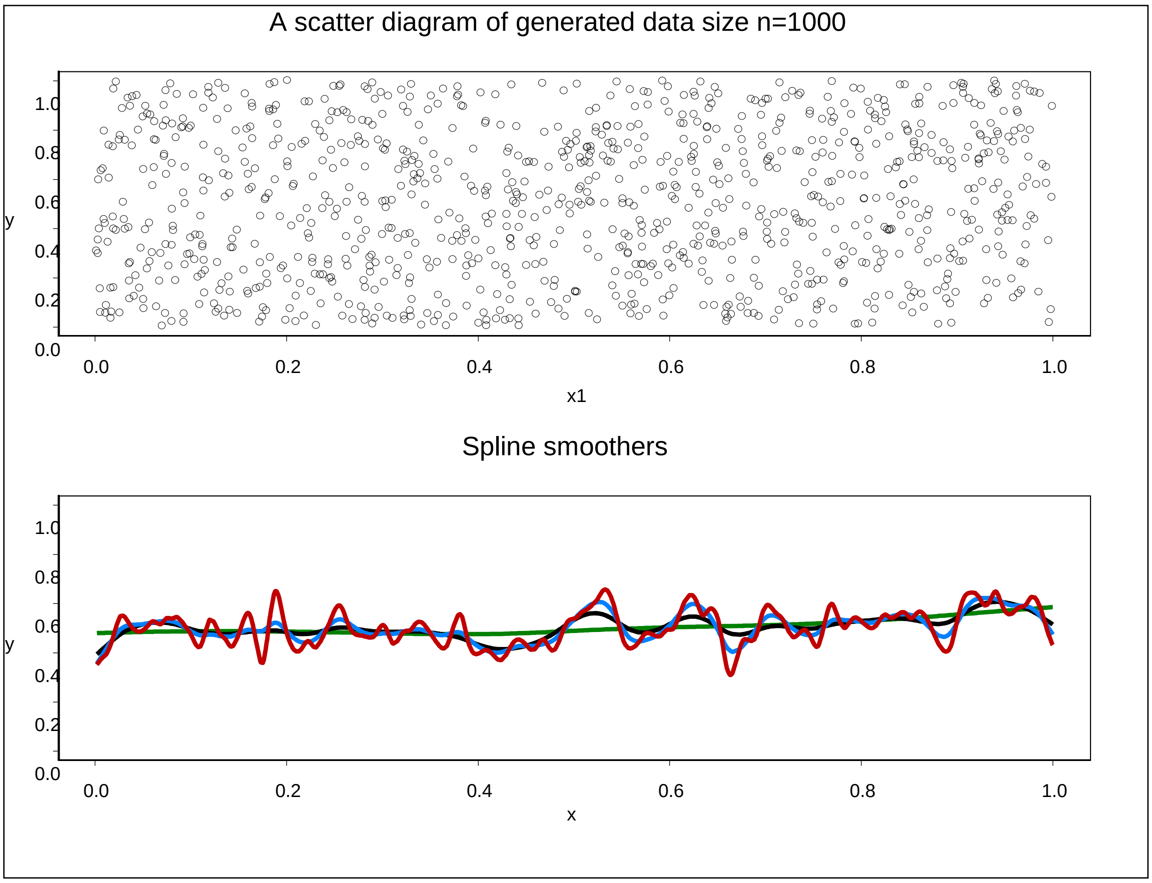



**Fig. 2. Shows 1000 generated data points with a spline levelling method with leveling parameters: $r$ = 10, 25, 40, 55. The blue wave with leveling parameter $r$ = 35 has the ideal bandwidth**

# 5 A Case Study Using Real Data Set

The assumption here is that the function of the variance is anonymous. The main purpose for approximating a regression curve in non-parametric includes;

- To forecast the worth of the response variable for observation whose information regarding the independent variable is at hand.
- to approximate the results of some independent variables on the response variable [12,13].

The variance function denoted by $Var(x_i, \hat{\beta})$ or Var$\{c_i\}$. The variance function is estimated using the residuals and the residuals are defined by $y_i - f\left(x_i, \hat{\beta}\right)$.

Then the expectation of the squared residuals gives the estimate of the variance function given by

$$E(r_i^2) = E[y_i - f(x_i, \beta)]^2 \cong V(x_i, \beta) \ [7]$$

We can also have the model in the design alone which is defined as

$$Var(y_i) = \sigma_i = Var(c_i)$$

where Var(.) is unknown

and $\{c_i\}$ is a set of i.i.d independent of $\{\varepsilon_i\}$.

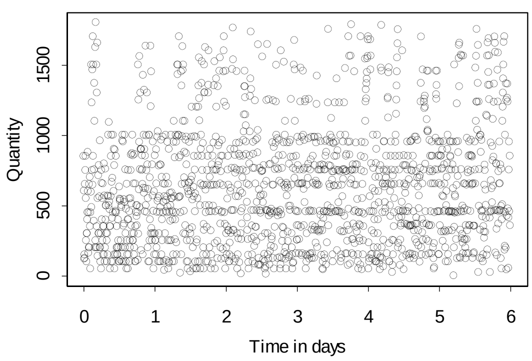
In more practical cases, to obtain a smooth environment we use a large data set with the aid of programmable computers and leveling techniques.

# 6 Example

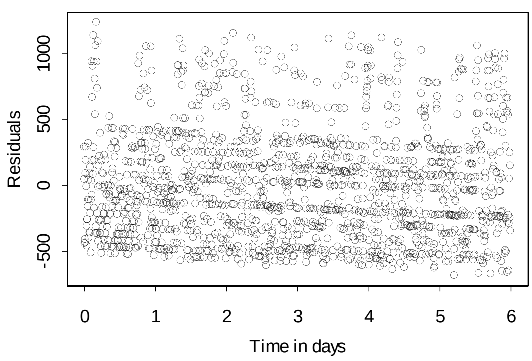
The two smoothing methods presented above are illustrated from data obtained from the Nairobi securities exchange. The share volume in consecutive months for a stretch of five years four months of Kenya commercial bank are used for finding the best smoother. The model which relates to the time ( $x_i$ ) and ( $y_i$ ) the share volume is modelled and plotted in the graphs below. Other variables include time, share volume in Kenyan shilling, residuals and residuals squared.

**Graphs of share volume against time, Residuals squared against time, Residuals squared against time and Kernel and Spline smooth of squared residuals against time**

a




b



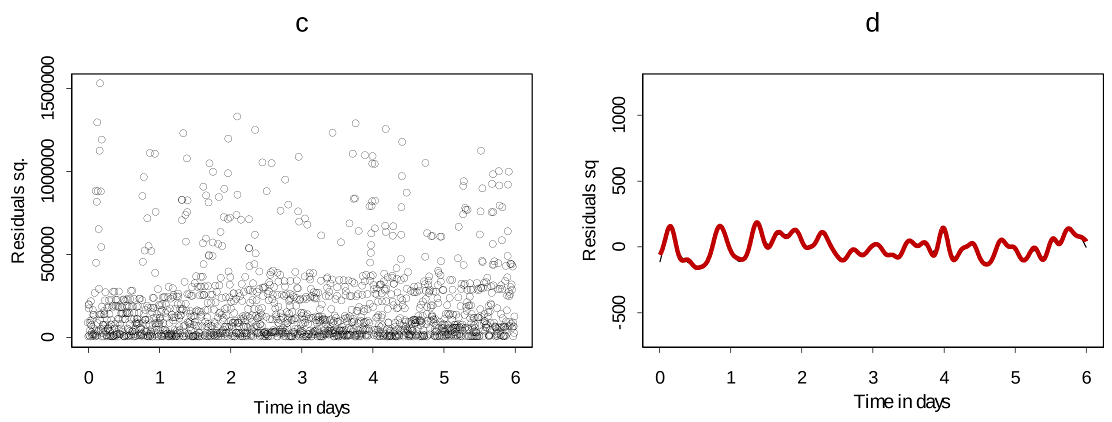



**Fig. 3. a) Shows a scatter plot of share volume on time b) Shows residuals squared on time c) Shows residuals squared against time d) Shows both kernel and spline smooth of squared residuals on time**

# 7 Conclusion

The main contribution of this study is that Kernel smoother produces the best estimate when compared to the Spline smoother. The results in Fig. 3 (d) shows that the variance of the Kernel smoother is less than that of the Spline smoother. It can also be seen from the same Fig. 3(d) that the spline has more variability around the middle and this shows that its variance is higher.

The main challenge of this study was the computational limitation of both Kernel and spline smoothers. There is a great difference along the boundaries (that is beginning and the end of the two curves). I recommend that more research be conducted to find out why there is such discrepancy on the boundaries. I also recommend that further research be conducted to explore on other data smoothing methods.

# Competing Interests

Author has declared that no competing interests exist.

# References

[1] Dennis Cook R, Sanford Weisberg. Applied regression including computing and graphics. John Wiley & Sons; 2009.

[2] Zhai C, Lafferty J. A study of smoothing methods for language models applied to ad hoc information retrieval. In ACM SIGIR Forum. ACM New York, NY, USA. 2017;51:268–276.

[3] Carrlo R.J, Ruppert D. Variance function estimation. Journal. of American Statistical Association. 1987;82:1079-1091

[4] Hardle W. Applied Nonparametric Regression, Cambridge, U.K, Cambridge University Press; 1993.

[5] Cai Z, Jonsson P, Jin H, Eklundh L. Perfor-mance of smoothing methods for reconstructing ndvi time-series and estimat-ing vegetation phenology from modis data. Remote Sensing. 2017;9(12):1271.

[6] Jobson JD, Fuller WA. Least squares estimation when the covariance matrix and the parameter vector are functionally related. Journal of the American statistical Association; 1980.

[7] Ian McLeod A. Diagnostic check for monotone spread. Computational Statistics & Data Analysis. 1998;26(4):437–443.

[8] Sanford Weisberg. Applied linear regression. John Wiley & Sons. 2005;528.

[9] Elizbar A Nadaraya. On estimating regression. Theory of Probability & Its Applications. 1964;9(1):141–142.

[10] Silverman Bernard W. Spline smoothing: the equivalent variable kernel method, The Annals of Statistics. 1984;898–916.

[11] Norman R Draper, Harry Smith. Applied regression analysis,. John Wiley & Sons. 1998;326.

[12] Brian Gray J. Applied regression including computing and graphics; 2000.

[13] Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafo MR, et al. Scanning the horizon: Towards transparent and reproducible neuroimaging research. Nature Reviews Neuroscience. 2017;18(2):115– 126.



***Peer-review history:***
*The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)*
*https://www.sdiarticle5.com/review-history/88947*