Ethiopian Journal of Environmental Studies & Management 15(1): 32 – 42, 2022
ISSN:1998-0507 doi: https://ejesm.org/doi/v15i1.4
Submitted: November 10, 2021 Accepted: February 4, 2022

ON COMPARISON OF PERFORMANCE METRICS OF THE BIVARIATE KERNEL ESTIMATOR

***SILOKO, I.U.,[1] SILOKO, E.A.,[2] ENOYOZE, E.,[1] EJAKPOVI, S.U.[3] AND UKHUREBOR, K.E.[4]**
[1]Department of Mathematics and Computer Science, Edo State University Uzairue, Nigeria
[2]Department of Statistics, University of Benin, Benin City, Edo State, Nigeria
[3]Department of Mathematics, College of Education, Warri, Nigeria
[4]Department of Physics, Edo State University, Uzairue, Nigeria
*Corresponding author: siloko.israel@edouniversity.edu.ng

Abstract
The two-dimensional kernel estimators are very important because graphical presentation of data beyond three dimensional forms is oftentimes not too frequently employed in data visualizations. The frequency of the bivariate estimator in the multivariate setting is attributed to the sparseness of data that is associated with increase in dimension. The performance of bivariate kernel is reliant on the smoothing parameter and other statistical parameters. While the smoothness of the estimates generated by the kernel estimator is primarily regulated by the smoothing parameter, its performance numerically may be depended on other statistical parameters. One of the popular performance metrics in kernel estimation is the asymptotic mean integrated squared error (AMISE) whose popularity is occasioned by its mathematical tractability and the inclusion of dimension with respect to performance evaluation. The computation of the bivariate kernel AMISE besides the smoothing parameter depends on basic statistical properties such as correlation coefficient and standard deviations of the observations. This paper compares the performance of the bivariate kernel using the correlation coefficient, standard deviations and the smoothing parameter. The results of the comparison show that for bivariate observations with independent standard deviations and correlated, the AMISE values is smaller than the AMISE values computed with the smoothing parameter.

Key Words: *AMISE, Kernel, Smoothing Parameter, Standard Deviation, Correlation Coefficient*

Introduction

The estimation of bivariate data involves the application of statistical tools in the derivation of statistical properties from the observations for the purpose of predicting the behaviour of the data. Oftentimes information about bivariate observations may not be clearly presented if not properly analyzed. The act of bringing orderliness, structure and meaning to a set of bivariate observations is known as data analysis and it is of great significance in statistical fields of studies. Bivariate data analysis and estimation is wide of applicability whether the data are in qualitative or quantitative forms. Qualitative data are normally in a verbal or narrative format and are usually collected via interviews or questionnaires while quantitative data are expressed in numerical forms either small or large. There are different stages of data analysis to ensure accurate examination of the statistical properties of the observations. The stages of data analysis are from visualization, exploratory, trend and then the estimation stage. Visualization of data is neither data analysis nor a substitute for data analysis but a useful commencement point before the analysis of the data. Visualization of data is a powerful statistical tool for communication of the evaluation of findings about the observations. The process of visualization stage involves the presentation of the data graphically in order to identify the inherent qualities of the data such as the trends and patterns of the observations. The exploratory stage involves a critical examination of the observations using some statistical tools while the trend analysis stage is the re-examination of the observations over a period of time. The estimation stage of data analysis involves the actual data values been apply in the prediction future occurrences (Siloko and Siloko, 2019).

The estimation stage could involve the construction of a probability estimates using the data either from a known probability distribution or unknown probability distribution. The method of density estimation can be broadly categorized into two major approaches which are parametric estimation and nonparametric estimation. Parametric estimation assumed the observations are members of a particular distribution family and that provides basic information about the needed parameters for the estimation. On the other hand, nonparametric estimation does not make any assumption about the distribution of the observations however; the observations are subjected to statistical examination using some analytic tools. Nonparametric estimation provides a better method in data analysis owing to its ability to capture the true structure of the underlying distribution. One of the fundamental qualities of nonparametric estimation is that they are of immerse application in exploratory analysis and visualization of data (Hansen, 2019). The flexibility of nonparametric density estimators has placed the estimators in the public domain in data analysis and estimation. There are several nonparametric estimation techniques but the kernel method will be considered in this paper due to its efficiency and computational advantage using the bivariate normal kernel function. The kernel estimation technique is a nonparametric method mainly for data exploratory analysis and data visualization. Owing to the importance of the application of kernel estimator in data explorations and visualizations, novel kernel estimators are been introduced by researchers (Mugdadi and Sani, 2020; Bouezmarni *et al.,* 2020; Harfouche *et al.,* 2020; Mohammed and Jassim, 2021; Bolancé and Acuña, 2021). The kernel estimation method has indirect applications in discriminant analysis, goodness-of-fit testing, hazard rate estimation, bump-hunting, image processing, remote sensing, seismology, cosmology, intensity function estimation, and classification with regression estimation (Sheather, 2004; Simonoff, 2012; Raykar *et al.,* 2015; Silverman, 2018; Siloko *et al*., 2019a).

This paper investigates the performance of the bivariate kernel estimator using the Gaussian kernel function known as the normal kernel. The bivariate product kernel estimator is presented with its corresponding AMISE formula. The smoothing parameters and kernel estimates of the bivariate Gaussian kernel using real data examples were obtained and the kernel performances were evaluated with two versions of the AMISE as performance metrics. The aim of the paper is to examine the nature of the AMISE value considering the standard deviations of bivariate data, correlation coefficient value and the smoothing parameter.

The Mathematical Formulation

The kernel estimator is a nonparametric estimator in density estimation in which a known density function is employ in averaging the data to produce a smooth approximation (Rosenblatt, 1956; Parzen, 1962). A kernel estimate is constructed by the summation of kernel functions centred at each data point and a smoothing parameter also called bandwidth that controls the degree of smoothness of the kernel estimate. The kernel function is a standardized weighting function with its univariate estimator given as

$$\hat{f}(\mathrm{x}) = \frac{1}{nh_{\mathrm{x}}}\sum_{i=1}^{n} K\left(\frac{\mathrm{x} - X_i}{h_{\mathrm{x}}}\right), \tag{1}$$

with $K(\cdot)$ representing the kernel function, n is sample size, $h_{\mathrm{x}} > 0$ is smoothing parameter also known as bandwidth, x is range of observations to be estimated while X_i is the observations. All kernel functions must be non-negative and satisfies the kernel axioms given by

$$\begin{gathered} \int K(\mathrm{x})d\mathrm{x} = 1, \\ \int \mathrm{x}K(\mathrm{x})d\mathrm{x} = 0 \quad \text{and} \\ \int \mathrm{x}^2K(\mathrm{x})d\mathrm{x} = \mu_2(K) \neq 0. \end{gathered} \tag{2}$$

The axioms in Equation (2) have serious implication in kernel density estimation. The first condition states that the integrant of every kernel function is one; therefore, most kernel functions are usually probability density functions. The other conditions simply state that the mean of the kernel which is its average is zero however; its variance denoted by $\mu_2(K)$ is greater than zero (Scott, 2015; Siloko *et al*., 2020).

The applications of kernel estimator are mostly in multivariate case with emphasis majorly on the bivariate estimator whose density estimates can be viewed in two-dimensional form or three-dimensional form. The popularity of the bivariate kernel estimator in higher dimensional density estimation is due to the simplification of the presentation of its estimates as surface plots (familiar perspective known as wire frame) or contour plots (Silverman, 2018; Siloko *et al*., 2021). Another factor that accounts for the application of the bivariate kernel estimator is the presentation of the observations with respect to their direction. The bivariate kernel estimator is

$$\hat{f}(\mathrm{x}\,,\mathrm{y}) = \frac{1}{nh_{\mathrm{x}}h_{\mathrm{y}}}\sum_{i=1}^{n} K\left(\frac{\mathrm{x}-\mathrm{X}_i}{h_{\mathrm{x}}},\frac{\mathrm{y}-\mathrm{Y}_i}{h_{\mathrm{y}}}\right), \tag{3}$$

with $h_{\mathrm{x}} > 0$ and $h_{\mathrm{y}} > 0$ representing the smoothing parameter in X and Y axes while x and y are the ranges of observations been estimated and $K(\mathrm{x}\,,\mathrm{y})$ is the bivariate kernel function. The bivariate kernel estimator is either in a spherical or product forms in practical applications. The product approach is mostly applied in density estimation due to the differences that usually exist in the range of observations with reference to their axes and is of the form

$$\hat{f}(\mathrm{x}\,,\mathrm{y}) = \frac{1}{nh_{\mathrm{x}}h_{\mathrm{y}}}\sum_{i=1}^{n} K\left(\frac{\mathrm{x}-\mathrm{X}_i}{h_{\mathrm{x}}}\right)K\left(\frac{\mathrm{y}-\mathrm{Y}_i}{h_{\mathrm{y}}}\right). \tag{4}$$

The application of different smoothing parameter for different axes is of great significant and beneficial particularly for observations with variations in their respective axes. In kernel density estimation, the smoothing parameter plays a vital role in the estimation process; hence it is regarded and interpreted as a resolution factor when viewing observations and giving better interpretation of the structures of the observations. The performance evaluation of the kernel estimator is dependent on the smoothing parameter, therefore appropriate choices of the smoothing parameter is very imperative. The problem of smoothing parameter selection in univariate kernel is not as difficult as the multivariate kernel estimator because in multivariate estimation, there are different forms of parameterizations (Siloko *et al*., 2019b). The prominence of bandwidths in kernel density estimation especially with increase in dimensions has resulted in the introduction of new bandwidth selectors since no single method has addressed the issue of bandwidths in all situations (Siloko *et al*., 2018; Varet *et al*., 2019; Tsuruta and Sagae, 2020; Tenreiro, 2020; Bedouhene and Zougab, 2020).

The Performance Metric of Kernel Estimator

The performance of the kernel estimator is usually assessed with an objective function known as the error criterion function. Several error criteria functions exist in literature but the asymptotic mean integrated squared error (AMISE) has gained popularity in kernel estimation because of its inclusion of dimension. The AMISE has two components which are integrated variance and integrated squared bias that depend on smoothing parameters for their evaluation. The AMISE as a criterion function is given by

$$AMISE\left(\hat{f}(\mathrm{x})\right) = \int \mathrm{Variance}\left(\hat{f}(\mathrm{x})\right)d\mathrm{x} + \int \mathrm{Bias}^2\left(\hat{f}(\mathrm{x})\right)d\mathrm{x}. \tag{5}$$

The univariate AMISE of the kernel estimator can be approximated by Taylor's series expansion and the result of the approximation yields

$$AMISE\left(\hat{f}(\mathrm{x})\right) = \frac{R(K)}{nh_{\mathrm{x}}} + \frac{1}{4}h_{\mathrm{x}}^4\mu_2(K)^2R(f''), \tag{6}$$

where $R(K) = \int K^2(\mathrm{x})d\mathrm{x}$ is roughness of kernel, $\mu_2(K)^2$ is kernel variance, while $R(f'') = \int f''(\mathrm{x})^2 d\mathrm{x}$ is roughness of the unknown function been estimated. The minimization of AMISE in Equation (6) will produce the optimal bandwidth given as

$$h_{\mathrm{x-AMISE}} = \left\{\frac{R(K)}{\mu_2(\mathrm{K})^2R(f'')}\right\}^{1/5} \times n^{-1/5}. \tag{7}$$

The AMISE of the bivariate product kernel estimator is of the form

$$AMISE\left(\hat{f}(\mathrm{x},\mathrm{y})\right) = \frac{R(K)}{nh_{\mathrm{x}}h_{\mathrm{y}}} + \frac{h_{\mathrm{x}}^4}{4}\mu_2(K)^2\int\int\left(\frac{\partial^2 f}{\partial \mathrm{x}^2}\right)^2 d\mathrm{x}d\mathrm{y} + \frac{h_{\mathrm{y}}^4}{4}\mu_2(K)^2\int\int\left(\frac{\partial^2 f}{\partial \mathrm{y}^2}\right)^2 d\mathrm{x}d\mathrm{y} \tag{8}$$

In the case of the bivariate kernel estimator just as the univariate kernel, the performance of the estimator is determined by the smoothing parameter particularly the bivariate Normal kernel with the assumption of the data been independently and identically distributed. The bandwidth that minimized the AMISE of the higher dimensional kernel is given as

$$H_{AMISE} = \left[\frac{dR(K)^d}{\mu_2(K)^2\, R\left(\nabla^2 f(\mathrm{x})\right)}\right]^{\left(\frac{1}{d+4}\right)} \times n^{-\left(\frac{1}{d+4}\right)}, \tag{9}$$

with $R\left(\nabla^2 f(\mathrm{x})\right)$ denoting the roughness of the unknown function. The problem of roughness of the unknown function for different kernel functions has been addressed by several authors. In the case of the Gaussian kernel, the roughness is of the form

$$R\left(\nabla^2 f(\mathrm{x})\right) = \frac{1}{\sigma_j^{d+4}}\left\{\left(2\sqrt{\pi}\right)^{-d}\left(\frac{d}{2} + \frac{d^2}{4}\right)\right\} = \left(\frac{d(d+2)}{4\left(2\sqrt{\pi}\right)^d}\right)\sigma_j^{-(d+4)}, \tag{10}$$

where σ_{j} is the standard deviations of the observations with $j = 1, 2, \ldots, d$. On substituting Equation (10) into Equation (9), we have the bandwidth that minimizes the AMISE given as

$$H_{AMISE} = \left[\frac{dR(K)^d}{\mu_2(\mathrm{K})^2\left(\frac{d(d+2)}{4\left(2\sqrt{\pi}\right)^d}\right)\sigma_j^{-(d+4)}}\right]^{\left(\frac{1}{d+4}\right)} \times n^{-\left(\frac{1}{d+4}\right)} \tag{11}$$

The smoothing parameter that will minimise the AMISE of the bivariate Gaussian kernel for the two axes is given by

$$h_{\mathrm{x}-AMISE} = \left\{\frac{dR(K)^d}{\mu_2(K)^2\left(\frac{d(d+2)}{4\left(2\sqrt{\pi}\right)^d}\right)\sigma_{\mathrm{x}}^{-(d+4)}}\right\}^{1/6} \times n^{-1/6} \tag{12}$$

$$h_{y-AMISE} = \left\{ \frac{dR(K)^d}{\mu_2(K)^2 \left(\frac{d(d+2)}{4(2\sqrt{\pi})^d} \right) \sigma_y^{-(d+4)}} \right\}^{1/6} \times n^{-1/6}, \tag{13}$$

with σ_{x} and σ_{y} representing the standard deviations of X and Y respectively while d is dimension of the kernel function. The bandwidths derived by the application of the Gaussian kernel are generally referred to as the Normal reference rule. The AMISE of the bivariate Gaussian kernel function in the case of observations with a correlation coefficient value that is not zero and standard deviations for the two observations can be obtained from the relation given as

$$\mathrm{AMISE}\left(\hat{f}(\mathrm{x},\mathrm{y})\right) = \frac{3}{8\pi\sigma_{\mathrm{x}}\sigma_y(1-\rho^2)^{5/6}} \left(1 + \frac{\rho^2}{2}\right)^{1/3} n^{-2/3}, \tag{14}$$

where ρ is correlation coefficient of the observations (Scott, 2015). The AMISE in Equation (14) depends on the standard deviations of the observations, sample size and correlation coefficient value that measure the relationship that exist between the observations.

The Gaussian Kernel Function

The Gaussian kernel function is one of the most applied kernel functions in density estimation. The popularity of this kernel function is mainly attributed to the fact that the function is continuously differentiable and also possesses higher order derivatives. The Gaussian kernel is also widely used in most scientific fields of studies such as social and medical sciences for the analysis of data for future prediction. Another factor that accounted for the wide applicability of the Gaussian kernel function is the comprehensibility of its mathematical formulations and the resulting estimates are usually smooth unlike other distribution with ambiguous expressions. The univariate and bivariate Gaussian kernel functions are of the forms

$$K(\mathrm{x}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\mathrm{x}^2}{2}\right). \tag{15}$$

$$K(\mathrm{x},\mathrm{y}) = \frac{1}{2\pi} \exp\left(-\frac{\mathrm{x}^2+\mathrm{y}^2}{2}\right). \tag{16}$$

The Gaussian kernel with its normalization constant typically possesses an area which is under unity and is often applied in the expansions of powers of polynomials.

Result and Discussion

The AMISE values of two bivariate observations were computed using Equation (8) and Equation (14) respectively. The computation of the AMISE value in Equation (8) is independent of some statistical properties of the observations such as the standard deviations and correlation coefficient as in Equation (14). All statistical analysis and graphics were fully implemented with Mathematica Version 12 software (Wolfram Research, Inc.). Generally, the performance of the kernel estimator is depended on accurate selection of the smoothing parameter especially in the investigation of inherent features of data. A method is regarded to have outperformed other methods when it produces the least AMISE value (Siloko *et al*., 2019c).

The analysis of two-dimensional observations often employs scatterplots in their investigation to show the nature of the relationship that exist amongst the observations. Although scatterplots do show relationships that exist between variables of interest, it is only the clouded natures of the data that are usually depicted while inherent features of the observations are hidden (Wand and Jones 1995; Siloko *et al*., 2018). The significant role of unveilings the attributes of observations is one of the advantages of kernel estimates, a function that scatterplots cannot perform because the attention of the observer is only on the surface view of the observations. Kernel estimates has the ability to display vital statistical qualities which portrays hidden structures of observations under investigation.

The first data set investigated is the ages at marriage for 100 couples that applied for marriage licenses in Cumberland County, Pennsylvania USA, which is made up of two variables the ages of husbands and their wives at marriage (Sabine and Brian, 2004). The scatterplots of the data is in Figure 1, and it shows a significant relationship between the variables implying that the ages of husbands at marriage are highly positively correlated with ages of their wives at marriage. The analysis of these data addresses the issue of differences in the ages of husbands and wives. However; on a general observation of the data, wives are younger at marriage. Also noted and illustrated from the scatterplots of these variables, is the tendency for younger women to marry younger men and vice-versa. Figure 2 is the kernel estimates of husbands' ages and wife's ages with the Gaussian kernel function. The smoothing parameters for husbands' ages and wife's ages in this case are $h_x = 5.71403$ and $h_y = 5.10798$ respectively. The bivariate kernel estimate of these data clearly depicts the data being unimodal and this indicates the ages at which husbands and their wives were more likely to get married. The unimodal property of the kernel estimates are distinctly centred between ages 20 and ages 29 and this simply indicates that the probability of getting married at these ages is high for both men and women. As observed from the kernel estimates, marriages are usually more in the twenty's and the probability of getting married tends to reduce downward between ages 30 and 40. A critical examination of the kernel estimates also revealed that between ages 46 and ages 56, there is the probability of contracting marriages and this may probably be for widows and widowers or late decisions makers.

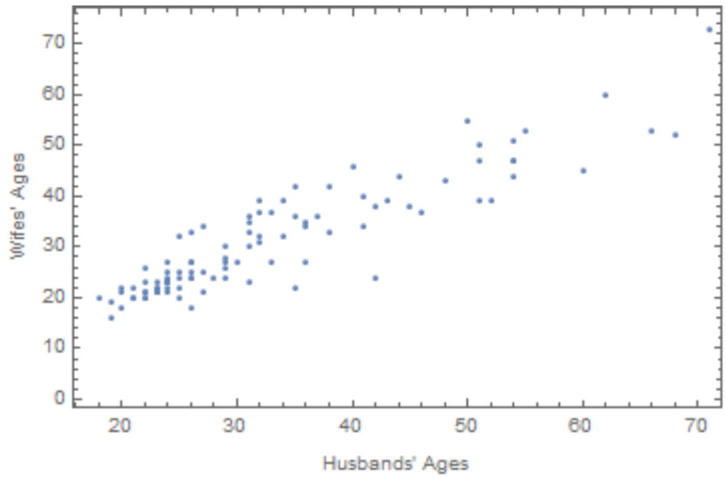

Fig. 1: Scatterplot of Husbands' and Wife's Ages at Marriage Data

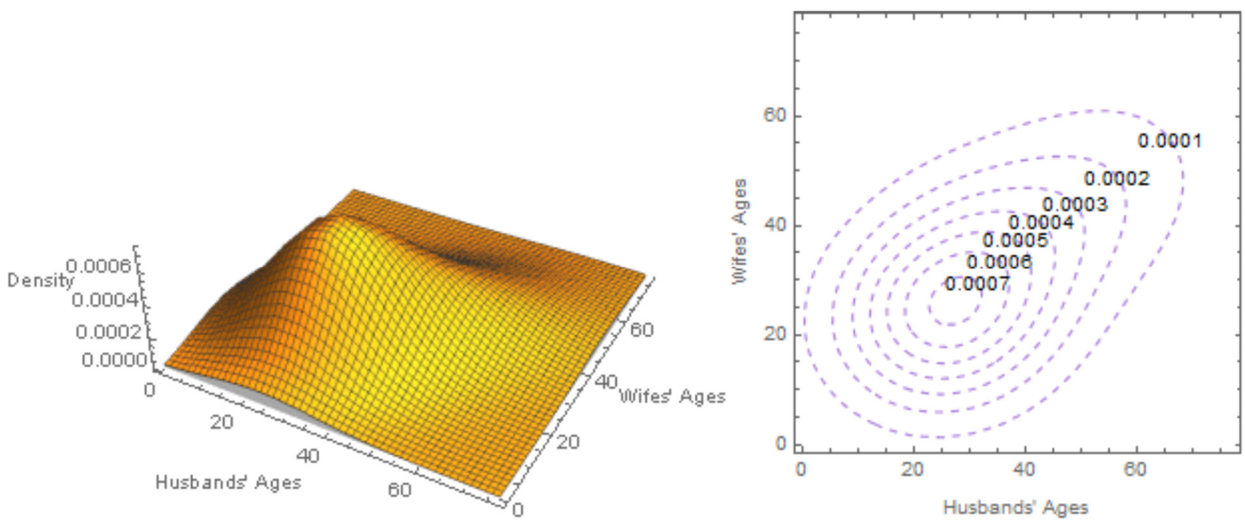

Fig. 2: Kernel Estimates (Surface plots and Contour plots) of the Ages at Marriage Data

The second data investigated is the Bunyaruguru Volcanic Field data found in Western Uganda (Bailey and Gatrell, 1995). The data comprises the location of Craters of 120 volcanoes two centers of craters which are represented by first center and second centre respectively. The scatterplots of the data is in Figure 3 and the scatterplot displayed evidence of strong relationship between the first centre and the second centre with correlation coefficient value of 0.8143976. It is evidently clear that the two centres are highly positively correlated. One of the qualities of the crater data as depicted in the kernel estimate of Figure 4 is the bimodality of the data but this unique characteristic is not apparent in the scatterplot. This demonstration of the bimodal nature of the data supports the claim that bivariate kernel estimates are very useful in structure identification in kernel density estimation. Generally, the construction of a kernel estimate is hinged on the selection of the appropriate bandwidth and the bandwidths for the crater data with respect to the first and second centres are $h_{\mathrm{x}} = 258.350$ and $h_{\mathrm{y}} = 350.740$ respectively.

The statistical properties employed in computation of the AMISE values in Table 1 and Table 2 are bandwidths denoted by h_{x} and h_{y} for the two dimensions, standard deviations of the data denoted by σ_{x} and σ_y and correlation coefficient values of the observations denoted by ρ. The AMISE values of Table 1 are larger than the AMISE values of Table 2 for the data investigated.

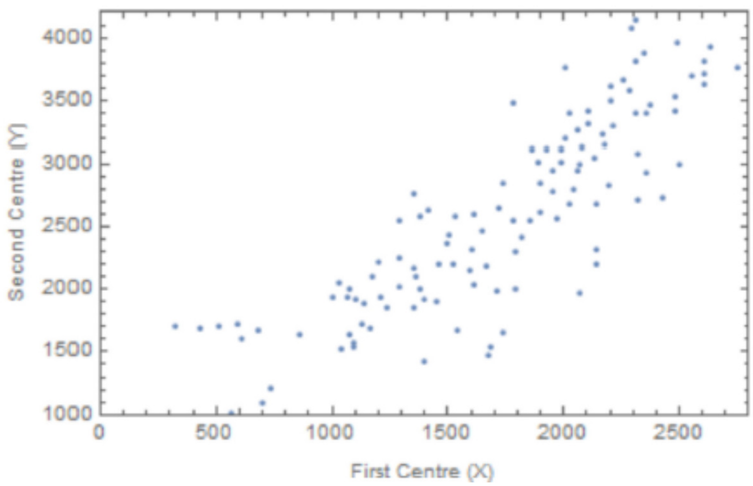

Fig. 3: Scatterplot of Crater Data of Bunyaruguru Volcanic Fields

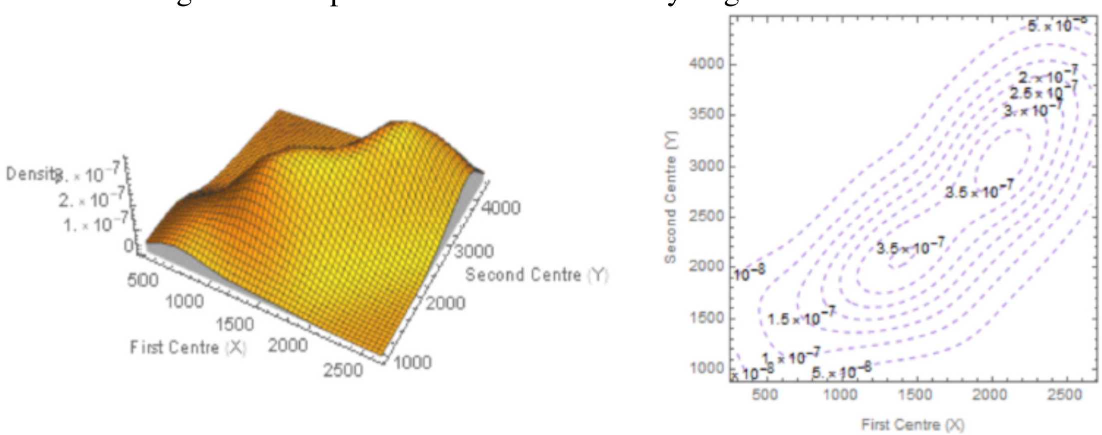

Fig. 4: Kernel Estimates (Surface plots and Contour plots) of the Volcanic Crater Data

In Table 1, the AMISE is computed using the bandwidths for the different data while Table 2 is computed by applying the standard deviations and correlation coefficients of the bivariate data.

Table 1: Bandwidths and AMISE for the Data Examined

Data Size	$\boldsymbol{h_{\mathrm{x}}}$	$\boldsymbol{h_{\mathrm{y}}}$	$\boldsymbol{AMISE}$
$n = 100$	5.71403	5.10798	0.0047331285945
$n = 120$	258.350	350.740	0.0042856861857

Table 2: Standard Deviations, ρ and AMISE for the Data Examined

Data Size	$\boldsymbol{\sigma_x}$	$\boldsymbol{\sigma_y}$	$\boldsymbol{\rho}$	$\boldsymbol{AMISE}$
$n = 100$	12.3105	11.0048	0.9122960	0.0002033513458
$n = 120$	573.771	778.960	0.8143976	0.0000000299137

It can be observed in Table 1 and Table 2 that the AMISE values decreases with increase in the sample sizes and that emphasized the benefits associated with large sample size in nonparametric estimation. The computed AMISE with the standard deviations and correlation coefficient values for the bivariate normal data is primarily dependent on the standard deviations of the data with less effect on the correlation coefficient value. The AMISE value tends to decrease with increase in the value of the standard deviations and vice-versa. Hence, bivariate data that are correlated with large standard deviations, the AMISE value is usually minimal unlike data with smaller values of standard deviations.

Conclusion

The degree of smoothness of kernel estimate and performance of kernel estimator is determined by the magnitude of the smoothing parameter. However, in the bivariate kernel that bridges the univariate and other higher dimensional kernel, the performance with respect to error criterion function is dependent on other statistical factors such as standard deviations of the data and correlations value of the observations. The performance evaluation using the AMISE as the performance metric shows that with the standard deviations of the bivariate data and correlation value, the AMISE is minimal in comparison with the AMISE values of the bandwidths. Again, the standard deviations of the bivariate data are of immerse contribution in determination of the AMISE value, with large standard deviations the smaller the AMISE value and vice-versa.

References

Bailey, T.C. and Gatrell, A.C. (1995). Interactive spatial data analysis. Longman, Harlow.

Bedouhene, K. and Zougab, N. (2020). A Bayesian procedure for bandwidth selection in circular kernel density estimation. *Monte Carlo Methods and Applications*, 26(1): 69–82.

Bolancé, C. and Acuña, C.A. (2021). A new kernel estimator of copulas based on beta quantile transformations. *Mathematics* 9: 1078. https://doi.org/10.3390/math9101078.

Bouezmarni, T., Bellengem, S. and Rabbi, Y. (2020). Nonparametric beta kernel estimator for long and short memory time series. *The Canadian Journal of Statistics*, 48: 582–595.

Hansen, B.E. (2019). Econometrics. Department of Economics, University of Wisconsin.

Harfouche, L., Zougab, N. and Adjabi, S. (2020). Multivariate generalized gamma kernel density estimators and application to non-negative data. *International Journal of Computing Science and Mathematics*, 11(2): 137–157.

Mohammed, A.D. and Jassim, N.H. (2021). A comparison study of nonparametric kernel estimators with gaussian weight function. *Journal of Physics: Conference Series*, 1818 012058.

Mugdadi, A. and Sani, R. (2020). The linear combination of kernels in the estimation of cumulative distribution functions. *Applications and Applied Mathematics*, 15(2): 875–893.

Parzen, E. (1962). On the estimation of a probability density function and the mode. *Annals of Mathematical Statistics*, 33: 1065–1076.

Raykar, V. C., Duraiswami, R. and Zhao, L. H. (2015). Fast computation of kernel estimators. *Journal of Computational and Graphical Statistics*, 19(1): 205–220.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27: 832–837.

Sabine, L. and Brian S.E. (2004). A handbook of statistical analyses using SPSS. Chapman & Hall/CRC, New York.

Sheather, S.J. (2004). Density estimation. *Statistical Science*, 19(4): 588–597.

Siloko, I.U. and Siloko, E.A. (2019). Density estimation and data analysis using the kernel method. Proceedings of 56th Annual National Conference of Mathematical Association of Nigeria, 218–226.

Siloko, I.U., Ikpotokin, O. and Siloko, E.A. (2019a). A key note on performance of smoothing parameterizations in kernel density estimation. *Tanzania Journal of Science*, 45(1): 1–8.

Siloko, I.U., Ikpotokin, O., Oyegue, F.O., Ishiekwene, C.C. and Afere, B.A., (2019b). A note on application of kernel derivatives in density estimation with the univariate Case. *Journal of Statistics and Management Systems,* 22(3): 415–423.

Siloko, I.U., Ishiekwene, C.C. and Oyegue, F.O. (2018). New gradient methods for bandwidth selection in bivariate kernel density estimation. *Mathematics and Statistics*, 6(1): 1–8.

Siloko, I.U., Siloko, E.A., Ikpotokin, O., Ishiekwene, C.C. and Afere, B.A. (2019c). On asymptotic mean integrated squared error's reduction techniques in kernel density estimation. *International Journal of Computational and Theoretical Statistics*, 6(1): 89–98.

Siloko, I.U., Ukhurebor, K.E., Siloko, E.A., Enoyoze, E. and Ikpotokin. O. (2021). The interactions between temperature and relative humidity: results for Benin City, Nigeria using statistical analysis. *Current Applied Science and Technology*, 22(1): 1–15.

Silverman, B.W. (2018). Density estimation for statistics and data analysis. Routledge, New York.

Simonoff, J.S. (2012). Smoothing methods in statistics. Springer Science & Business Media, New York.

Tenreiro, C. (2020). Bandwidth selection for kernel density estimation: A Hermite series-based direct plug-in approach. *Journal of Statistical Computation and Simulation,* 90(18): 3433–3453.

Tsuruta, Y. and Sagae, M. (2020). Theoretical properties of bandwidth selectors for kernel density estimation on the circle. *Annals of the Institute of Statistical Mathematics,* 72: 511–530.

Varet, S., Lacour, C., Massart, P. and Rivoirard, V. (2019). Numerical performance of penalized comparison to overfitting for multivariate kernel density estimation. https://hal.inria.fr/hal-02002275/

Wand, M.P. and Jones, M.C. (1995). Kernel Smoothing. Chapman and Hall, London.