Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Enhancing Kernel Estimators Using the Hyperbolic Secant Kernel for Nonparametric Regression: Applications to Simulation and Real Data

Ola A. Elsamadony[1], S. E. Abu-Youssef [2]
1Department of Mathematics, Higher Institute for Advanced Administrative Sciences and Computers, Beheira, Egypt
2 Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr City (11884), Cairo, Egypt
Email:1ola.ali192@yahoo.com, 2abuyoussef@azhar.edu.eg

ARTICLE INFO

Received: 31 Dec 2024

Revised: 20 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

In this study, we rely on adaptive kernel estimation to improve the non-parametric estimation of the probability density function (*pdf*) using the hyperbolic secant kernel (*HSK*). Previous research has demonstrated that adaptive kernel estimators with diverse and different bandwidths yield superior performance. This paper introduces an enhancement to the hyperbolic secant kernel estimator (HSKE) through the use of Quartile Deviation (QD), Coefficient of Variation (CV), and Variance-to-Mean Ratio (VMR). These proposed methods have also been applied to nonparametric Nadaraya-Watson (NW) regression. The mean squared error (MSE) is a measure used to evaluate the performance the new estimators that have been suggested. A lower MSE indicates a more accurate estimator or model, reflecting its effectiveness in making precise predictions. The simulation study results showed that very positive results, demonstrating that our modification of HSKE shows good performance in all cases. The two real data sets are showed improvement in the regression model when using the new methods.

Keywords: Non-parametric, Adaptive Kernel estimates, Estimation, Regression, Mean squared error, Bandwidth, Hyperbolic Secant, standard deviation, mean.

1. INTRODUCTION AND MOTIVATIONS

One of the most researched problems in statistics is estimating an underlying distribution from data. In general, the methods used to estimate *pdf* can be sorted into two principal types: parametric and non-parametric methods. A major non-parametric method is the kernel-based approach, commonly known as Kernel Density Estimation (KDE). In econometrics, the KDE technique is sometimes referred to as the Parzen-Rosenblatt window approach. It's a method built around the histogram methodology. This method is vital tool in statistics, machine learning, signal processing and other fields and that's because its flexibility and ability to represent complex data distributions without relying on a specific parametric model, for more details see [6], [10], [15] and [25]. Rosenblatt (1956) and Parzen (1962) laid the theoretical foundations for KDE, see [13] and [14]. The bandwidth parameter (b) is crucial to the quality of KDE. It determines the width of the kernel used to smooth the data. It is also crucial to control the balance between bias and variance in the density estimate. The correct choice of b is essential to achieve a good density estimate that accurately reflects the basic distribution of the data. There are two types of band width in the kernel estimator, either fixed or variable. Fixed-bandwidth KDE involves using a constant bandwidth parameter b is applied across all data points. The KDE is expressed as:

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

where n represents the number of observations, X_i be the individual data points and b stands for the fixed bandwidth.

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The fixed bandwidth is simple, easy to implement, and computationally efficient, and it gives a smoothing effect to the estimation; the choice of bandwidth directly affects the smoothness of the density estimate: A small bandwidth leads to a less smooth and more sensitive estimate and may capture more data noise, while a large bandwidth leads to a smoother and more general estimate and may over smooth and obscure details of the data distribution.

There exist several methods for selecting a fixed bandwidth in (KDE) such as, plug-in methods: it is aim to estimate the optimal bandwidth by minimizing an approximation of the mean integrated squared error (MISE) [18] and [23], cross-validation: The goal of cross-validation techniques is to determine the best bandwidth by evaluating the prediction error performance of various bandwidths. These approaches, which are typically more data-driven and aim to minimize the estimation error by evaluating the performance of the density estimate on unseen data. An overview of a few popular techniques for cross-validation: Least-Squares Cross-Validation (LSCV), Leave-One-Out Cross Validation (LOOCV) and Biased Cross-Validation for more details see, [15], [19] and [21].

Another method is rule-of-thumb method (Silverman's Rule of Thumb), which is particularly effective for unimodal and symmetric distributions see [20]. The formula for Silverman's Rule of Thumb is as follows:

$$b = \left(\frac{4\,\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}}, \quad (2)$$

where b is the bandwidth, σ expresses the standard deviation and n is the number of data points. Another formula for Silverman's bandwidth is:

$$b = 0.9 \min\left(\sigma, \frac{IQR}{1.34}\right) n^{\frac{-1}{5}}, \quad (3)$$

where, The interquartile range is represented by IQR.

Adaptive kernel density estimation (AKDE) enhances and extends traditional kernel density estimation (KDE) by adjusting the bandwidth of the kernel function according to the local data density, thus improving the efficiency and accuracy of density estimations. AKDE is more flexible than standard KDE and can handle multimodal distributions and varying data densities more effectively.

This adaptive approach allows for more accurate and detailed estimation of the underlying data distribution, which is particularly useful when dealing with complex data sets where the distribution is not known or is too complex to be modeled with parametric methods. The bandwidth (smoothing parameter) is fixed across the entire data range, which may lead to over-smoothing in regions with high data density and low smoothing in Areas with low data density. AKDE employs a different bandwidth (b_i) than standard KDE, which has a fixed bandwidth, for more details see [16], [22] and [4].

The basic KDE is expressed as:

$$\hat{f}(x) = \frac{1}{nb}\sum_{i=1}^{n} K\left(\frac{x - X_i}{b}\right), \qquad x \in \mathbb{R}, \quad (4)$$

where n is the sample size (the number of observations or data points), while b denotes the bandwidth, serving as a smoothing parameter. The kernel function is represented by K, and X_i denotes the data points.

Hence, we can define the adaptive Kernel Density Estimation as the following:

$$\hat{f}(x) = \frac{1}{nb_i}\sum_{i=1}^{n} K\left(\frac{x - X_i}{b_i}\right), \qquad x \in \mathbb{R}, \quad (5)$$

where $K(x)$ refers to kernel function and b_i is the new bandwidth.

Abramson (1982) [1] proposed an adaptive method to vary the bandwidth according to the local density of the data, rather than using a fixed bandwidth for all data points. This approach improves the performance of kernel density estimators, especially when dealing with varying local density. The adaptive bandwidth b_i for a data point X_i is expressed as:

$$b_i = \frac{b}{\sqrt{\tilde{f}(X_i)}},$$

where $f(X_i)$ expresses a pilot estimate of the density at X_i.

In Silverman [20], especially in the context of adaptive kernel density estimation, he presents an algorithm that improves a basic kernel estimator (such as the one proposed by Abramson) by incorporating local bandwidth modulation based on local density.

Here's how the algorithm typically works, in four steps:

1. Compute the pilot density estimate $\tilde{f}(X_i)$, which is used to determine the bandwidth variation for the final adaptive estimate.
2. Determine the local bandwidth factor λ_i, which is defined as:

$$\lambda_{gi} = \left[\frac{\tilde{f}(X_i)}{g}\right]^{-\alpha},$$

where α is the sensitivity parameter and the value of $\alpha \in [0,1]$. In (1982), Abramson selected $\alpha = 0.5$; since this value yields good and effective forecasting results. Estimation using the geometric mean (g) of $\tilde{f}(X_i)$ is called (GM.KDE)

3. Define the new bandwidth $b_i = \lambda_i * b$.
4. Define the adaptive kernel estimator as shown in Equation (5).

In 2010, Demir and Toktami¸s [9], [11] modified and improve the kernel estimator by using the arithmetic mean (M.KDE) rather than the geometric mean to compute λ_i. It can be expressed as follows:

$$\lambda_{\bar{x}i} = \left[\frac{\tilde{f}(X_i)}{\bar{x}}\right]^{-\alpha}.$$

In 2014, the kernel estimator was modified to use the range rather than the arithmetic or geometric mean to compute λ_i, as described in [24]. The local bandwidth factor is expressed as follows:

$$\lambda_{Ri} = \left[\frac{\tilde{f}(X_i)}{R}\right]^{-\alpha}.$$

In 2019, the kernel estimator was modified to use the median (MED.KDE) instead of the arithmetic, geometric mean and range to compute λ_i, as described in [9]. The local bandwidth factor is expressed as follows:

$$\lambda_{Medi} = \left[\frac{\tilde{f}(X_i)}{Med}\right]^{-\alpha}.$$

In 2021, a new approach was proposed to an adaptive kernel estimator [7]. Four statistical techniques were used, which are as follows: Interquartile Range (IQR.KDE), Standard Deviation (SD), Mean Absolute Deviation (MAD), and Median Absolute Deviation (MeAD) instead of the geometric mean, an arithmetic mean, median and range.

$$\lambda_i = \left[\frac{\tilde{f}(X_i)}{IQR}\right]^{-\alpha}, \left[\frac{\tilde{f}(X_i)}{MAD}\right]^{-\alpha}, \left[\frac{\tilde{f}(X_i)}{MeAD}\right]^{-\alpha} \text{ or } \left[\frac{\tilde{f}(X_i)}{SD}\right]^{-\alpha}$$

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

This research contains four sections. Section 2 improves and proposes other adaptive kernel estimation methods for application to HSKE. Section 3 test the performance of the improved HSKE and compares it with the logistic kernel estimator based on simulated data. In section 4 the adaptive Nadaraya-Watson (NW) regression kernel methods. In section 5 real data set used to evaluate the performance of our proposed methods for regression model. Finally, Section 6 provides our observations and conclusions.

2 NEW PROPOSED OF ADAPTIVE KERNEL DENSITY ESTIMATOR

We present a new approach to the adaptive kernel estimator that employs three different statistical techniques: quartile deviation (QD.KDE), coefficient of variation (CV.KDE), and variance rate (VR.KDE).

We apply these methods to the hyperbolic secant kernel estimator (HSKE), then we can define the adaptive kernel estimator by the hyperbolic secant kernel function as

$$\hat{f}(x) = \frac{1}{nb_i}\frac{1}{\pi}\sum_{i=1}^{n} \operatorname{sech}\left(\frac{x-X_i}{b_i}\right), \tag{6}$$

where, $b_i = \lambda_i * b$ and we are calculated (b) by the rule in equation 3.

The First Statistical Technique (QD.KDE).

The local bandwidth factor is expressed as:

$$\lambda_{QDi} = \left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha}, \tag{7}$$

where QD is the quartile deviation, also known as the semi-interquartile range is calculated as

$$QD = \frac{Q_3 - Q_1}{2}.$$

Here, Q_1 represents the lower (first) quartile and represents the number that falls halfway between the bottom number and the middle number, while Q_3 represents the upper (third) quartile, which represents the number falling halfway between the middle number and the top number. From here, the new bandwidth is given by:

$$b_i = \left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha} b.$$

The adaptive kernel estimator become:

$$\hat{f}(x) = \frac{1}{n\,\pi\left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha} b}\sum_{i=1}^{n} \operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha} b}\right). \tag{8}$$

The Second Statistical Technique (CV.KDE).

The local bandwidth factor is expressed as:

$$\lambda_{CV_i} = \left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha}, \tag{9}$$

Coefficient of variation is symbolized by (CV). It compares the standard deviation (σ) of a data set to its mean (μ), and it is often expressed as a percentage. Mathematically, it is represented as: $CV = \frac{\sigma}{\mu}$. From here, the new bandwidth is given by:

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$b_i = \left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha} b.$$

The adaptive kernel estimator become:

$$\hat{f}(x) = \frac{1}{n\,\pi\left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha} b} \sum_{i=1}^{n} \operatorname{sech}\left(\frac{x - X_i}{\left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha} b}\right). \qquad (10)$$

The Third Statistical Technique (VR.KDE).

The local bandwidth factor is expressed as:

$$\lambda_{VM_i} = \left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha}, \qquad (11)$$

where VM, or variance rate, is calculated as the ratio of the variance σ^2 to the mean, it is represented as: $VR = \frac{\sigma^2}{\mu}$.

From here, the new bandwidth is given by:

$$b_i = \left[\frac{f(X_i)}{VR}\right]^{-\alpha} b.$$

The adaptive kernel estimator become:

$$\hat{f}(x) = \frac{1}{n\,\pi\left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha} b} \sum_{i=1}^{n} \operatorname{sech}\left(\frac{x - X_i}{\left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha} b}\right). \qquad (12)$$

3 SIMULATION OF ADAPTIVE KERNEL DENSITY ESTIMATOR

We were conducted to evaluate the effectiveness and performance of the estimators by using two simulated data. The first simulation study, samples were drawn from an exponential distribution with a parameter of 2, and the analysis was performed for various sample sizes (100, 200, 300, 400, and 500). Each experiment was repeated 8000 times using Mathematica software.

We calculated (*MSE*) to evaluate the performance of fixed bandwidth KDE and AKDE methods (including M.KDE, GM.KDE, HM.KDE, MED.KDE and IQR.KDE), and the new proposed methods (QD.KDE, CV.KDE, VR.KDE) using *HSK*.

The table 1 shows that the MSE decreases with increasing sample size, we are observed that all adaptive kernel estimators perform better than the fixed bandwidth kernel density estimation, and the new proposed estimators outperforms them as they have the lowest MSE, especially the CV.KDE method.

Table 1 indicates that the greater the sample size, the greater the MSE reduction ratio.

Table 1: MSE values of fixed bandwidth kernel estimation and new proposed methods of kernel estimation by using *HSK*.

n	Fixed KDE	M.KDE	GM.KDE	HM.KDE	MED.KDE	IQR.KDE	QD.KDE	VR.KDE	CV.KDE
100	0.326412	0.324113	0.323058	0.322158	0.321809	0.317598	0.305648	0.307011	0.301721

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

200	0.289226	0.28809	0.28757	0.287148	0.28671	0.283905	0.277077	0.276849	0.265034
300	0.277713	0.276926	0.276571	0.276289	0.275886	0.273601	0.268605	0.268426	0.25385
400	0.263248	0.262565	0.262293	0.262077	0.261738	0.259874	0.255897	0.25511	0.241002
500	0.2397	0.239238	0.239046	0.238897	0.238602	0.236642	0.233295	0.232831	0.218894

The second simulation study, samples were drawn from a normal distribution with mean 4 and variance 6. The analysis was conducted for various sample sizes (100, 200, 300, 400, and 500). Each experiment was repeated 8000 times using Mathematica software.

We calculated (*MSE*) to evaluate the performance of logistic kernel estimator and hyperbolic secant kernel estimator in two cases: fixed bandwidth and the proposed method using (CV).

Table 2 shows that the MSE decreases with increasing sample size. Additionally, the hyperbolic kernel estimator outperforms the logistic kernel estimator in both scenarios. Furthermore, the new method using (CV) is more effective than the fixed bandwidth method for both types of estimators.

Table 2: Comparison between HSKDE and logistic kernel density estimator using fixed bandwidth and the new propose method (CV.KDE).

n	**Kernel**	**Fixed Method**	**Proposed Method (CV)**
100	Hyperbolic Secant	0.000704107	0.000656121
	Logistic	0.00163183	0.00154305
200	Hyperbolic Secant	0.0000453472	0.0000398762
	Logistic	0.000110101	0.0000994557
300	Hyperbolic Secant	$7.48021*10^{-6}$	$6.229*10^{-6}$
	Logistic	0.0000183295	0.000015857
400	Hyperbolic Secant	$1.5156*10^{-6}$	$1.17625*10^{-6}$
	Logistic	$3.72749*10^{-6}$	$3.05155*10^{-6}$
500	Hyperbolic Secant	$4.86315*10^{-7}$	$3.56578*10^{-7}$

	Logistic	$1.19769*10^{-6}$	$9.37819*10^{-7}$

4 THE ADAPTIVE NADARAYA-WATSON (NW) KERNEL METHODS FOR NONPARAMETRIC REGRESSION

Regression analysis is one of the primary topics in applied statistics. This is because this area of study is widely used in engineering, economics, medicine, and agriculture. Finding a functional relationship between the components of two sets, X and Y, where X stands for the independent variable and Y for the dependent variable, is the main goal.

Determining this function presents challenges, especially in establishing the functional form of the relationship from a set of Y observations for unobserved X values. These issues can be resolved through a variety of methods. The two classifications of approaches that are now in use are parametric and non-parametric.

Assuming that the data has a defined functional form, parametric regression is a statistical method for modeling the relationship between one or more independent variables (predictors) and a dependent variable (response).

The data is used to estimate the model's parameters. While non-parametric regression can be useful for studying relationships in data when the underlying functional form is unknown or cannot be easily defined, making it an appropriate approach for modeling complex phenomena like disease outbreaks, for more details see [2] and [8].

We have random variables X_i and $Y_i \in$R, the regression formula is:

$$Y_i = \beta(X_i) + \varepsilon_i, \qquad i = 1, 2, 3, \cdots, n, \tag{13}$$

where ε_i are observation errors and $\beta(X_i)$ the unknown regression function which can estimated by the following equation:

$$\hat{\beta}(x) = \frac{\sum_{i=1}^{n} y_i \; K_b(x - X_i)}{\sum_{i=1}^{n} K_b(x - X_i)}. \tag{14}$$

where,

$\hat{\beta}(x)$: The estimated value of the regression function.

$K(.)$: The kernel function.

b: The bandwidth, which controls the smoothness of the estimate.

X_i: The values of the independent variable.

y_i: The values of the dependent variable.

n: The number of observations.

The NW estimator was introduced by researchers Nadaraya and Watson in 1964 as a non-linear approximation to regression models that rely on empirical data [12]. The NW kernel estimator can be fixed or variable, and it is dependent on the smoothing parameter h (bandwidth).

Traditional kernel regression with fixed bandwidth sometimes faces some challenges, including over smoothing: in regions where data points are dense, using a large bandwidth can hide the fine details of the relationship, and undersmoothing: in sparse regions, a small bandwidth can lead to noisy estimates.

Adaptive kernel regression addresses these problems by allowing the bandwidth to vary according to the density of the data. Dense regions use a smaller bandwidth to obtain finer accuracy, while sparse regions use a larger bandwidth to ensure smoother estimates.

For more details see, [9], [17] and [4]. The adaptive NW kernel estimator can be written as follow:

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

$$\beta_{ANW}(x)=\frac{\sum_{i=1}^{n}\frac{y_i}{b_i}K\left(\frac{x-X_i}{b_i}\right)}{\sum_{i=1}^{n}\frac{1}{b_i}K\left(\frac{x-X_i}{b_i}\right)}, \tag{15}$$

where, $b_i = b * \lambda_i$.

As outlined in Section (2), various improvement methods were proposed. In this section, we implement both the previously discussed techniques and the newly proposed methods to enhance the performance of the Nadaraya-Watson regression model.

The adaptive *NW* kernel estimators are given by the hyperbolic kernel function, respectively, as follows:

$$\hat{\beta}_{ANW}(x)=\frac{\sum_{i=1}^{n}\frac{y_i}{\left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha}b}\operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha}b}\right)}{\sum_{i=1}^{n}\frac{1}{\left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha}b}\operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{QD}\right]^{-\alpha}b}\right)}. \tag{16}$$

$$\hat{\beta}_{ANW}(x)=\frac{\sum_{i=1}^{n}\frac{y_i}{\left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha}b}\operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha}b}\right)}{\sum_{i=1}^{n}\frac{1}{\left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha}b}\operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{CV}\right]^{-\alpha}b}\right)}. \tag{17}$$

$$\hat{\beta}_{ANW}(x)=\frac{\sum_{i=1}^{n}\frac{y_i}{\left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha}b}\operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha}b}\right)}{\sum_{i=1}^{n}\frac{1}{\left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha}b}\operatorname{sech}\left(\frac{x-X_i}{\left[\frac{\tilde{f}(X_i)}{VR}\right]^{-\alpha}b}\right)}. \tag{18}$$

5 REAL DATA FOR THE ADAPTIVE NADARAYA-WATSON (NW) KERNEL METHODS

In this section, two real data sets used to evaluate the performance of our proposed methods for regression model.

The first data set: the data on daytime eruptions of Old Faithful Geyser in Yellowstone National Park from August 1–4, 1978, comprises observations for the first 32 eruptions, out of a total of 52. The variables of interest are *x*, representing the duration of an eruption (in minutes), and *y*, which indicates the interval until the next eruption (also in minutes). It has been suggested that a non-parametric regression can effectively model the relationship between variable *x* and variable *y*, allowing for successful predictions of *y* based on values of *x*, see [3].

x = 1.70, 1.70, 1.70, 1.80, 1.80, 1.80, 1.80, 1.90, 1.90, 1.90, 2.00, 2.30, 2.30, 2.50, 3.10, 3.20, 3.40, 3.40, 3.50, 3.50, 3.60, 3.70, 3.70, 3.70, 3.70, 3.80, 3.80, 3.80, 3.80, 3.80, 3.90, 3.90.

y = 55.00, 58.00, 56.00, 42.00, 51.00, 51.00, 45.00, 53.00, 49.00, 51.00, 51.00, 50.00, 53.00, 66.00, 57.00, 79.00, 75.00, 86.00, 80.00, 82.00, 86.00, 69.00, 79.00, 73.00, 67.00, 60.00, 86.00, 72.00, 75.00, 75.00, 74.00, 80.00.

The second data set: the data represents COVID-19 mortality rates in Italy over 59 days, recorded from February 27 to April 27, 2020, as described in [5].

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The data are as follows:
4.571,7.201,3.606,8.479,11.41,8.961,10.919,10.908,6.503,18.474,11.01,17.337,16.561,13.226,15.137,8.697,15.787,13.333,11.822,14.242,11.273,14.33,16.046,11.95,10.282,11.775,10.138,9.037,12.396,10.644,8.646,8.905,8.906,7.407,7.445,7.214,6.194,4.64,5.452,5.073,4.416,4.859,4.408,4.639,3.148,4.04,4.253,4.011,3.564,3.827,3.134,2.78,2.881,3.341,2.686,2.814,2.508,2.45,1.518.

The proposed methods estimators were more accurate than all the previous classical methods based on MSE criterion. From the real data results, note that Tables 3 and 4 compares between the FNW, previous classical (M.NW, HM.NW, GM.NW) and proposed methods (QD.NW, VR.NW, CV.NW) by using the hyperbolic secant.

The new proposed methods estimators were more accurate and efficient than all the previous classical methods based on MSE criterion. The results revealed that the CV.NW method outperformed other approaches by achieving the lowest Mean Squared Error (MSE).

The graphs of the Nadaraya-Watson regression using *HSK*, obtained for the two real data through both the classical and proposed methods, with *b* determined using Silverman's thumb rule, are presented in Figure 1 and 4 respectively.

Figures 2 and 3 represent the graphs between the classical and proposed methods at b= (0.1, 0.6) for the first real data while figure 5 represent the graph between the classical and proposed methods at b=(0.4) for the second real data.

Table 3: MSE values of fixed bandwidth NW kernel regression and new proposed methods by using *HSK* for the first data set.

b	FNW	M.NW	GM.NW	HM.NW	QD.NW	VR.NW	CV.NW
Silverman's Rule of Thumb	44.1778	44.1755	44.1101	44.0440	38.2788	25.7906	18.3926
0.1	31.3482	31.2445	31.1228	30.9974	21.1509	18.3774	18.3630
0.4	44.2149	44.2145	44.1481	44.0815	38.3224	25.8252	18.3936
0.6	53.0510	53.3468	53.1631	52.9766	41.2575	28.4328	18.7006
0.9	77.2038	77.8992	77.7270	77.5518	42.3493	28.9046	19.4995

Table 4: MSE values of fixed bandwidth NW kernel regression and new proposed methods by using *HSK* for the second data set.

b	FNW	M.NW	GM.NW	HM.NW	QD.NW	VR.NW	CV.NW
Silverman's Rule of Thumb	5.6328	5.5785	5.3951	5.1902	3.9487	3.1837	2.6123
0.4	0.4523	0.24617	0.17025	0.1069	0.03248	0.00616	0.0000109
0.7	1.6855	1.2854	1.10122	0.8932	0.52404	0.13937	0.00122
1	2.32067	2.04879	1.90211	1.71715	1.35987	0.56395	0.01488

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

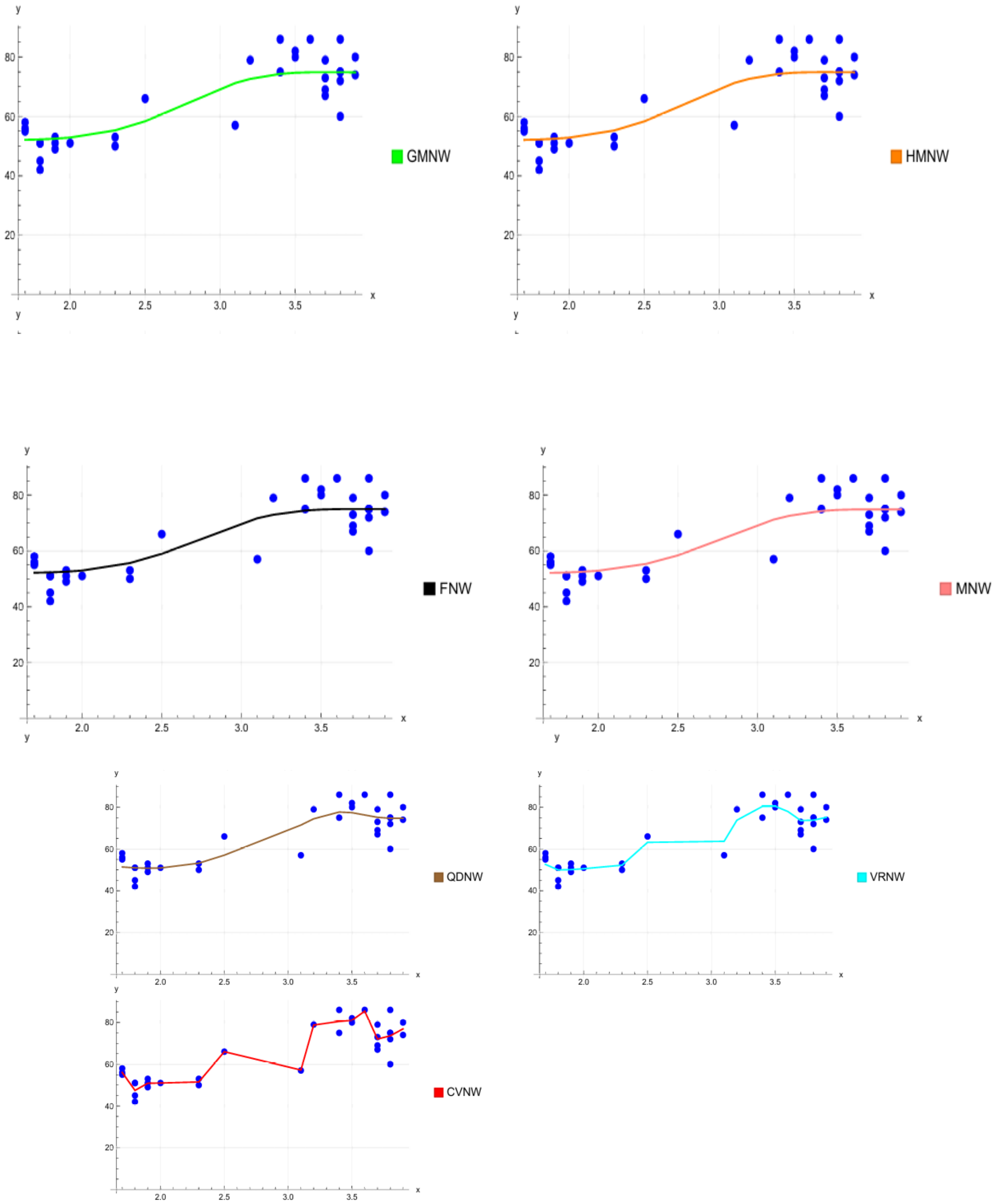


Figure 1: Classical and Proposed methods at b is Silverman's Rule of Thumb for the first data set.

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

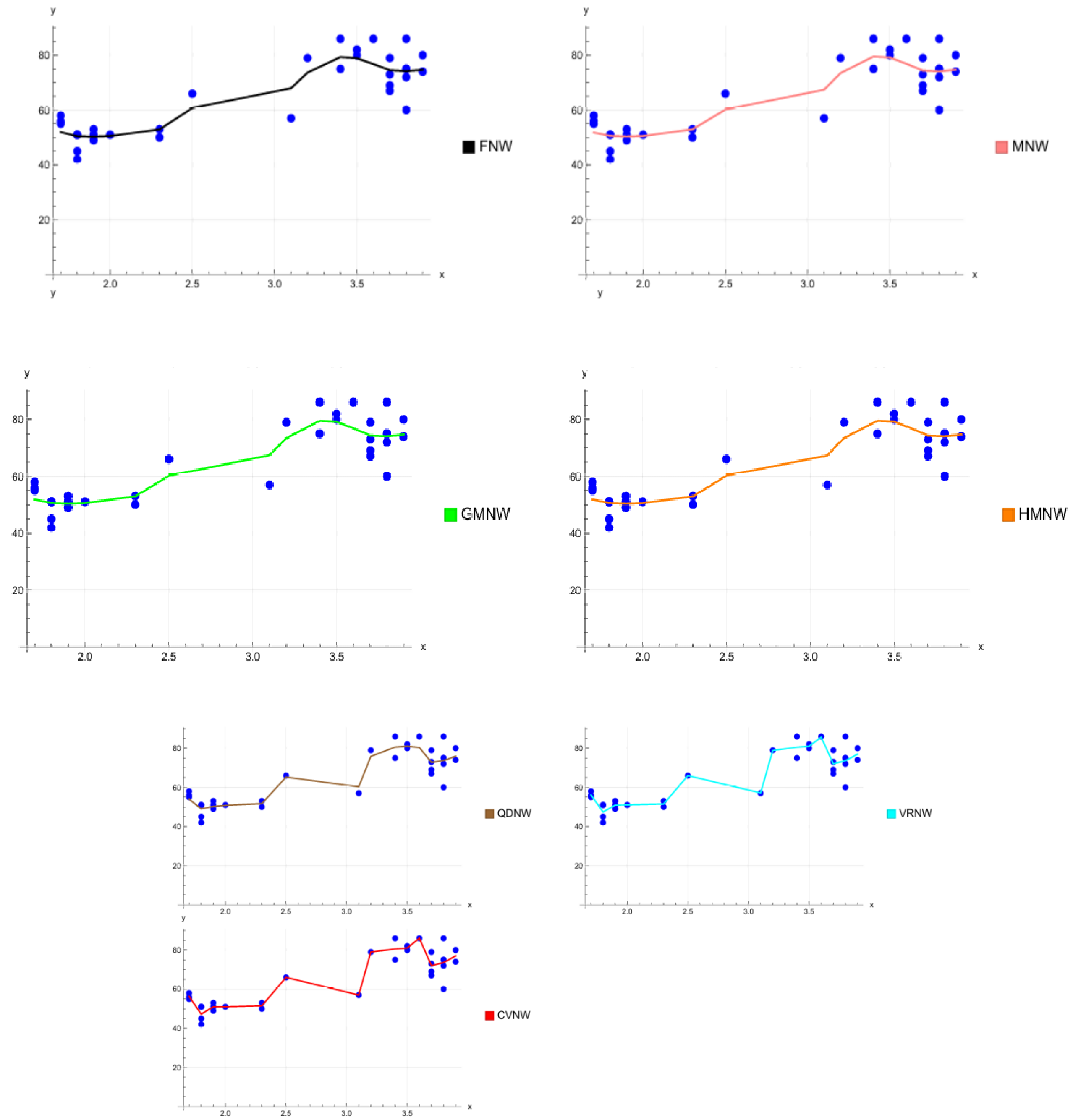


Figure 2: Classical and Proposed methods at b =0.1 for the first data set.

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

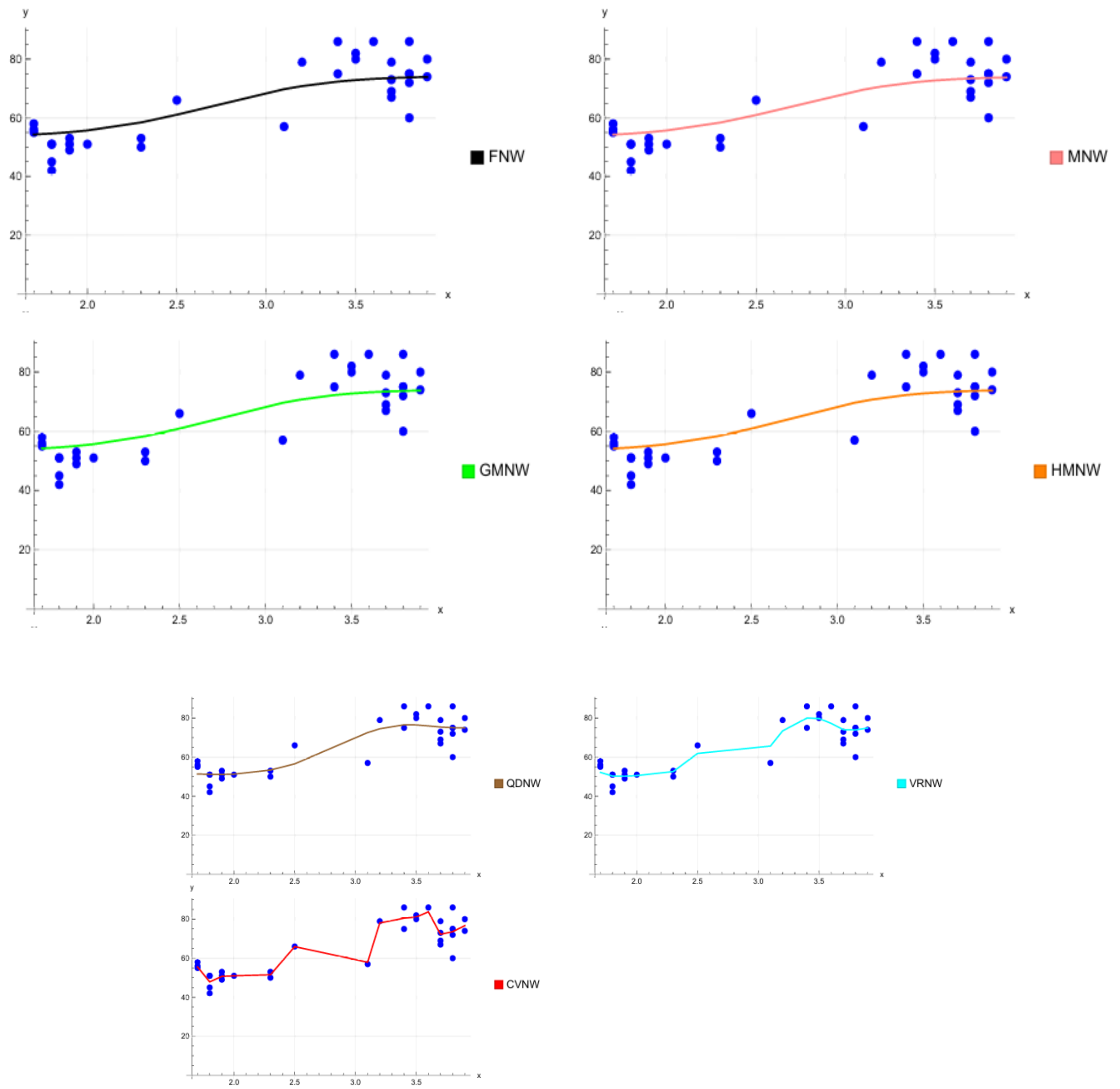


Figure 3: Classical and Proposed methods at b =0.6 for the first data set.

Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

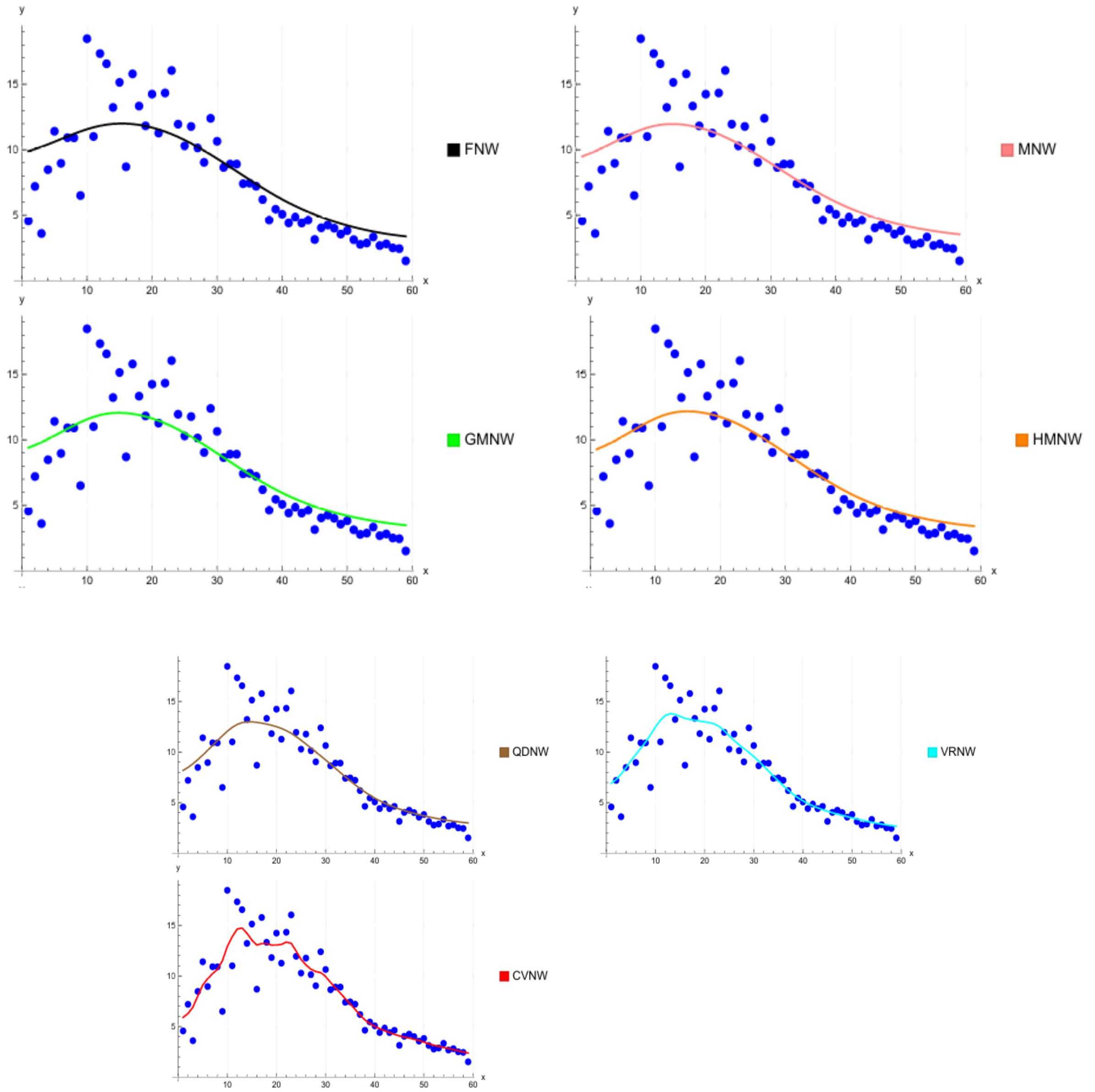
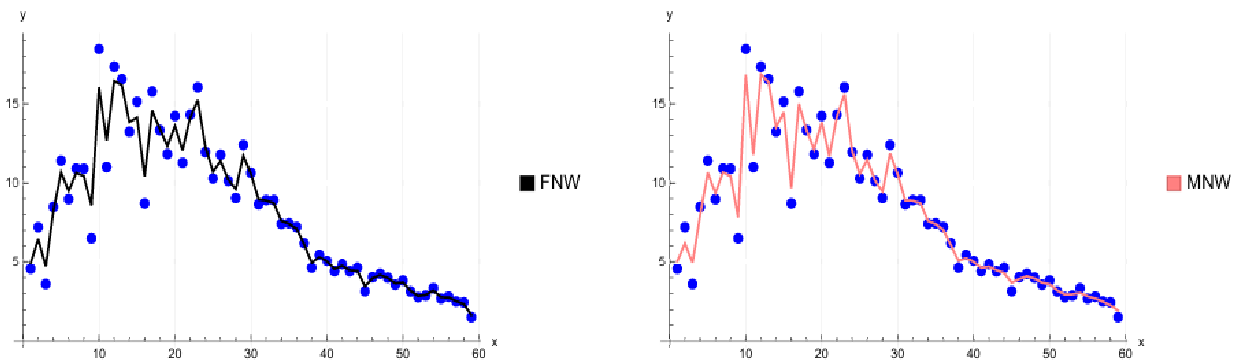


Figure 4: Classical and Proposed methods at b is Silverman's Rule of Thumb for the second data set.



Journal of Information Systems Engineering and Management
2025, 10(46s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

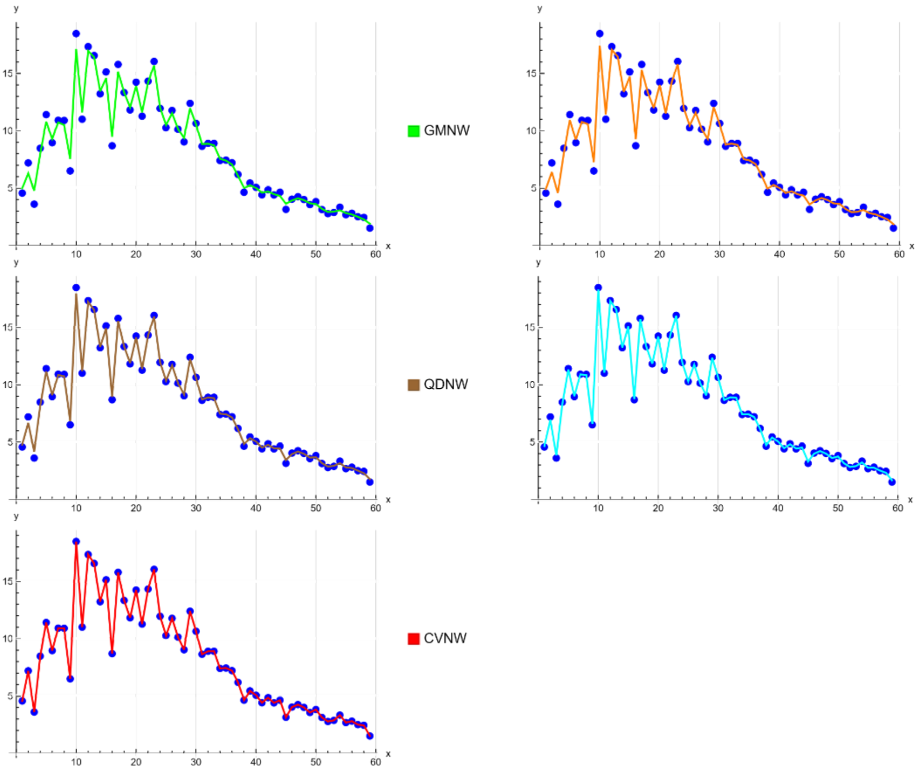


Figure 5: Classical and Proposed methods at b =0.4 for the second data set.

6 CONCLUSION

In this paper, we enhance HSKE by employing the adaptive kernel estimation technique. The newly proposed estimators for (QD, CV, VR) KDE demonstrate superior reliability compared to alternative methods, as indicated by MSE criteria across various simulations. The newly proposed estimator of CV.KDE was more efficient than any of the other methods in both simulations due to its low MSE values. In both simulations, the MSE decreases as the sample size increases. Moreover, we observed that the HSKE is more efficient than the logistic kernel estimation in fixed bandwidth case and the adaptive method CV. The result of the above study is that newly adaptive kernel estimation methods significantly enhance performance of the hyperbolic secant kernel estimator. We also noticed that when applying the proposed methods to nonparametric regression, it achieves high efficiency compared to other previous methods in studies, and it is clear that the best method was the CVNW method.

REFRENCES

[1] Ian S Abramson. On bandwidth variation in kernel estimates-a square root law. *The annals of Statistics*, pages 1217–1223, 1982.

[2] Zaid Abd-Alrazak Al-Homsi. *Comparison among some Nonparametric Methods in Estimating Regression Models*. PhD thesis, Zarqa University, 2017.

[3] Taha Hussein Ali, Heyam Abd Al-Majeed Hayawi, and Delshad Shaker Ismael Botani. Estimation of the bandwidth parameter in nadaraya-watson kernel non-parametric regression based on universal threshold level. *Communications in Statistics-Simulation and Computation*, 52(4):1476–1489, 2023.

[4] Khulood Hamed Aljuhani and LI Al Turk. Modification of the adaptive nadaraya-watson kernel regression estimator. *Scientific Research and Essays*, 9(22):966–971, 2014.

[5] Hisham M Almongy, Ehab M Almetwally, Hassan M Aljohani, Abdulaziz S Alghamdi, and EH Hafez. A new extended rayleigh distribution with applications of covid-19 data. *Results in Physics*, 23:104012, 2021.

[6] Hassan S Bakouch, Christophe Chesneau, and Ola A Elsamadony. The gumbel kernel for estimating the probability density function with application to hydrology data. *Journal of Data, Information and Management*, 3:261–269, 2021.

[7] Hazhar TA Blbas and Wasfi T Kahwachi. A comparison between new modification of adaptive nadaraya-watson kernel and classical adaptive nadaraya-watson kernel methods in nonparametric regression. *Cihan University-Erbil Scientific Journal*, 5(2):32–37, 2021.

[8] Mohammed A Dakhil and Jassim N Hussain. A comparative study of nonparametric kernel estimators with gaussian weight function. In *Journal of Physics: Conference Series*, volume 1818, page 012058. IOP Publishing, 2021.

[9] Serdar Demir and Oniz Toktami¸s. On the adaptive nadaraya-watson kernel regression estima-¨ tors. *Hacettepe Journal of Mathematics and Statistics*, 39(3):429–437, 2010.

[10] Artur Gramacki. *Nonparametric kernel density estimation and its computational aspects*, volume 37. Springer, 2018.

[11] Degui Li and Runze Li. Local composite quantile regression smoothing for harris recurrent markov processes. *Journal of econometrics*, 194(1):44–56, 2016.

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

[13] Emanuel Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.

[14] Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the national Academy of Sciences*, 42(1):43–47, 1956.

[15] Mats Rudemo. Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, pages 65–78, 1982.

[16] Raid B Salh, Hazem I El Shekh Ahmed, and Iyad M Alhoubi. On improvement of hazard rate function estimation using adaptive kernel estimates. *IUG Journal of Natural Studies*, 25(3), 2017.

[17] William R Schucany. Adaptive bandwidth choice for kernel regression. *Journal of the American Statistical Association*, 90(430):535–540, 1995.

[18] David W Scott, Richard A Tapia, and James R Thompson. Kernel density estimation revisited. Nonlinear Analysis: Theory, Methods & Applications, 1(4):339–372, 1977.

[19] David W Scott and George R Terrell. Biased and unbiased cross-validation in density estimation. Journal of the american Statistical association, 82(400):1131–1146, 1987.

[20] Bernard W Silverman. Density estimation for statistics and data analysis. Routledge, 2018.

[21] Charles J Stone. An asymptotically optimal window selection rule for kernel density estimates. The Annals of Statistics, pages 1285–1297, 1984.

[22] Philippe Van Kerm. Adaptive kernel density estimation. The Stata Journal, 3(2):148–156, 2003.

[23] Matt P Wand and M Chris Jones. Kernel smoothing. CRC press, 1994.

[24] Geoffrey S Watson. Smooth regression analysis. Sankhy¯a: The Indian Journal of Statistics, Series A, pages 359–372, 1964.

[25] Adriano Z Zambom and Ronaldo Dias. A review of kernel density estimation with applications to econometrics. International Econometric Review, 5(1):20–42, 2013.