Journal of Statistical Modelling: Theory and Applications
Vol. 6, No. 2, 2025, pp. 1-24
Yazd University Press 2025

JSMTA

Research Paper

On bias reduction for probability density function estimation based on a kernel estimator

Iman Makhdoom[1], Reza Salehi*[2]
[1]Department of Statistics, Payame Noor University, Tehran, Iran
[2]Department of Mathematics and Statistics, Sa.C., Islamic Azad University, Sanandaj, Iran

Received: February 03, 2026/ Revised: March 16, 2026/ Accepted: April 18, 2026

Abstract: The probability density function is a fundamental concept in statistics. This study focuses on estimating the probability density function using nonparametric kernel methods. Initially, the usual kernel method is introduced. Subsequently, we present the two new estimates of the probability density function, termed the biased reduced kernel estimate, the repeat of the biased reduced kernel estimate, and the proposed biased reduced kernel estimate obtained by subtracting the bias value from the kernel estimator. The paper explores theoretical properties, including the selection of the smoothness parameter, bias, variance, and mean squared error of the proposed estimator. The accuracy of the biased reduced kernel estimate is scrutinized through Monte Carlo simulations. Moreover, the mentioned methods were employed using the five real datasets. The findings reveal that the proposed biased reduced kernel method exhibits a further reduction in bias compared to the usual kernel, biased reduced kernel, and repeated biased reduced kernel methods.

Keywords: Biased reduced kernel estimation; Nonparametric density estimation; Repeat of the biased reduced kernel estimate; Smoothness parameter.
Mathematics Subject Classification (2010): 62G05, 62G07.

1 Introduction

One of the nonparametric methods for estimating the probability density function is the kernel method. In recent decades, many researchers have focused on bias reduction

*Corresponding author: `Reza.Salehi2022@iau.ac.ir`

methods in estimating the probability density function. Individuals such as Härdle (1991), Scott (1992), Wand and Jones (1995), and many others have extended and expanded it in one-dimensional and higher dimensions. Samiuddin and El-Sayyad (1990) introduced kernels that lead to bias reduction. El-Sayyad et al. (1993) proposed an estimator using certain probability identities to reduce bias. Marron and Ruppert (1992) suggested a suitable algorithm for second-degree polynomials and the cumulative beta distribution in boundary bias reduction. Ruppert and Cline (1994) introduced a method where data are transformed using an estimated cumulative distribution function (CDF) to reduce bias in kernel density estimation (KDE). This approach involves multiple iterations and can achieve convergence rates comparable to higher-order kernels. Adaptive or variable-bandwidth KDE adjusts the kernel width based on data density, allowing for more flexibility in capturing features of the underlying distribution. This method is particularly effective in multi-dimensional settings. Recent studies have proposed bias correction methods for KDE with spherical data by applying generalized jackknifing techniques. These methods improve accuracy by reducing bias, especially in higher dimensions.

Kim et al. (2003) developed a bootstrapping method for bias reduction. Mynbaev and Martins-Filho (2010) worked on relative bias reduction in the classical kernel estimator through a Lipschitz condition. Many of the bias reduction methods mentioned above yield mixed results for mixed kernel density estimators. Mynbaev et al. (2014) studied bias improvement using Epanechnikov and Gram-Charlier kernels in density estimation. Igarashi and Kakizawa (2015) investigated symmetric kernel estimators using non-negative bias correction techniques with a standard normal kernel. Slaoui (2018) presented estimators for bias reduction. Salehi and Shadrokh (2018) examined bias reduction in the estimation of the probability density function using an extrapolated geometric kernel estimator. Calonico et al. (2018) developed methods for nonparametric kernel-based density and local polynomial regression estimators, including robust bias-corrected inference and coverage error optimal bandwidth selection. These techniques enhance the reliability of KDE in practical applications. Dhaker et al. (2021) examined the issue of determining the probability density function. If it is evident from the above studies that considerable efforts have been made in recent decades to provide more accurate probability density estimates. Makhdoom and Yaghoobzadeh Shahrastani (2023) obtained the probability density and cumulative distribution functions to conduct an efficient estimation on the inverse Weibull distribution.

More recently, Withers and Nadarajah (2023) utilized Lipschitz conditions for relative bias reduction. Yoon et al. (2023) derived an optimal weight function using tools from the multidimensional calculus of variations to reduce bias in standard kernel density estimates for density ratios. This approach leads to improved estimates of prediction posteriors and information-theoretic measures. Adaptive KDE methods and generalized jackknifing Tsuruta (2024) have further improved estimator performance in high-dimensional and spherical settings. Zahnit et al. (2024) proposed bias-corrected transformed KDE for heavy-tailed distributions using the Champernowne transformation. Despite the significant progress made in the field, several important challenges remain. Many existing bias-reduction techniques rely on complex kernels or transformation procedures, which increase computational burden and complicate implementation. Additionally, some methods reduce bias at the expense of increased variance, ultimately limiting their effectiveness in improving mean squared error (MSE). Furthermore, the lack of broad empirical validation on diverse real-world datasets raises concerns about the practical utility of these techniques. To address these limitations, this paper introduces two novel bias reduced kernel estimators: (i) the repeated biased reduced Kernel (RBRK) estimator, and (ii) the proposed biased reduced Kernel (BRKp) estimator. These estimators are constructed by analytically estimating and subtracting the bias from the standard kernel estimator. Among them, the BRKp estimator demonstrates the most promising performance in terms of reducing bias , as confirmed through simulation studies and real data analysis.

The key contributions of this study are as follows: (1) the development of a new family of kernel-based bias-reduced estimators that are computationally efficient and easy to implement; (2) the derivation of theoretical properties, including bias, variance, and MSE; and (3) comprehensive empirical evaluation through Monte Carlo simulations and applications to real datasets, highlighting the effectiveness of the proposed methods-particularly BRKp-in practical settings.

One of the most common methods for assessing the accuracy of an estimator is the MSE criterion. Since MSE is decomposed into bias and variance components, reducing bias can lead to a decrease in MSE, ultimately improving the accuracy of the estimator. Therefore, in this article, our focus is on reducing the bias of symmetric kernel estimators, and we compare the results using Monte Carlo simulations with the standard normal kernel. This study also contributes to the KDE literature by offering a practical and effective bias reduction strategy that is easy to implement and performs robustly across a variety of distributions.

The rest of the paper is organized as follows. In Section 2, usual kernel estimator is considered. Bias reduction estimator is proposed in Section 3. In Section 4, a simulation study is carried out and the results are compared using a normal standard kernel. Five numerical example are done in Section 5 and finally the paper is concluded in Section 6.

2 Usual kernel estimator

The simplest method to estimate the kernel density function is as follows. Suppose $X_1, X_2, \ldots, X_n$ be a random sampling with size n from a distribution with a probability density function is unknown. In this case, the estimate of the density of a usual kernel (UK) at point x is defined as follows

$$\hat{f}_{h,n}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x-X_i}{h}). \tag{1}$$

Here, h represents a positive real number, referred to as the bandwidth parameter, and $K_h(x)$ denotes the kernel function, which must meet the following criteria:

$$a_1) \qquad \int_{-\infty}^{\infty} K(x)dx = 1,$$

$$a_2) \qquad \int_{-\infty}^{\infty} xK(x)dx = 0,$$

$$\begin{array}{ll} a_3) & K(x) = K(-x), \\ a_4) & h \to 0, nh \to \infty \ \text{ if } \ n \to \infty, \\ a_5) & \|K\|_2^2 = \int_{-\infty}^{\infty} K^2(x)dx < \infty, \\ a_6) & \mu_2(K) = \int_{-\infty}^{\infty} x^2 K(x)dx < \infty, \\ a_7) & \sup_{-\infty<x<\infty} |K(x)| \leq \infty. \end{array}$$

Also, the function f must have the following conditions

$$\begin{array}{ll} b_1) & \|f''\|_2^2 = \int_{-\infty}^{\infty} (f'')^2(x)dx < \infty, \\ b_2) & f \in C^4(R) \ \text{ and } \ \sup_x |f^{(4)}(x)| < \infty. \end{array}$$

Wand and Jones (1995) found that while the *Epanechnikov* kernel offers a slight performance advantage, all kernels are asymptotically equivalent. Consequently, kernel choice has minimal impact on the estimator's properties, particularly with large sample sizes, and the kernel's shape is relatively insensitive to the selection. The behavior of various kernels is described in Table 1.

Table 1: Efficiency of some kernels compared to the *Epanechnikov* kernel.

Kernel	Function	Efficiency
Gaussian	$K(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$	0.9512
Triangular	$K(x) = (1 - \lvert x \rvert) I_{\{\lvert x \rvert \leq 1\}}$	0.9859
Epanechnikov	$K(x) = \frac{3}{4}(1 - x^2) I_{\{\lvert x \rvert \leq 1\}}$	1
Biweight	$K(x) = \frac{15}{16}(1 - x^2) I_{\{\lvert x \rvert \leq 1\}}$	0.9939
Rectangular	$K(x) = \frac{1}{2} I_{\{\lvert x \rvert \leq 1\}}$	0.9295

As evident, the kernel estimator presented in (1) is influenced by the smoothness parameter. If h value is too small, the estimator $\hat{f}_{h,n}$ will be very rough. Therefore, it is necessary to provide a suitable method to determine the optimal value of h. Assume that the unknown density function f and the kernel function K are true under conditions b_1, a_5 and a_6, f'' is also absolutely continuous in the neighborhood of x. Based on Silverman (1986), Wand and Jones (1995), we have

$$Bias(\hat{f}_h(x)) = E(\hat{f}_h(x)) - f(x) = \frac{1}{2} h^2 f''(x) \int y^2 K(y)dy + O(h^3), \tag{2}$$

$$Var(\hat{f}_h(x)) = E(\hat{f}_h(x) - E(\hat{f}_h(x)))^2 = \frac{1}{nh} f(x) \int K^2(y)dy + O((nh)^{-1}). \tag{3}$$

So, we conclude from (2) and (3) that

$$\begin{aligned} MSE(\hat{f}_h(x)) &= Bias^2(\hat{f}_h(x)) + \text{var}(\hat{f}_h(x)) \\ &\sim \frac{h^4}{4} (f''(x))^2 \left(\int y^2 K(y)dy \right)^2 + \frac{1}{nh} f(x) \int K^2(y)dy + O(\frac{1}{n} + h^5). \end{aligned}$$

As a result, the optimal MSE will be of the order $O\left(n^{\frac{-4}{5}}\right)$with$h_{opt} = O(n^{\frac{-1}{5}})$. That is, the optimal value of the parameter h is equal to

$$h_{opt} = \left(\frac{f(x) \, \|K\|_2^2}{(f''(x))^2 (\int y^2 K(y) dy)^2 n} \right)^{\frac{1}{5}} \sim O(n^{-\frac{1}{5}}).$$

Because the relation h_{opt}depends on the unknown probability density function f and its second-order derivative, a natural way to solve this problem is to use the standard family of distributions. For example, suppose the density f belongs to the Normal family with mean $\mu$and varianceσ^2. Therefore, the optimal value of the parameter h is equal to $h_{opt} = 1.06\hat{\sigma} n^{-0.2}$, where $\hat{\sigma}$ is the sample's standard deviation.

3 Bias reduction estimator

To reduce the bias of density function, Xie and Wu (2014) proposed the biased reduced kernel (BRK) estimator an as follows

$$\bar{f}_{h,n}(x) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - X_i}{h}) - \frac{h}{2n} \int x^2 K(x) dx \sum_{i=1}^{n} K''(\frac{x - X_i}{h}). \tag{4}$$

To improve the bias reduction of estimator (4), once again we subtract its bias from the estimator. Since the bias term is unknown and inaccessible due to the presence of the unknown and inaccessible density function f, its estimated value can be used instead of f. If the kernel K is symmetric, then its odd-order moments will be zero. Consequently, we have

$$\begin{aligned}
\tilde{\bar{f}}_{h,n}(x) &= \bar{f}_{h,n}(x) - \widehat{Bias}(\bar{f}_{h,n}(x)) \\
&= \bar{f}_{h,n}(x) - \left(-\frac{h^4}{4} f^{(4)}(x) \left[\int x^2 K(x) dx \right]^2 \right) \\
&= \bar{f}_{h,n}(x) + \frac{h^4}{4} f^{(4)}(x) \left[\int x^2 K(x) dx \right]^2 \\
&= \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - X_i}{h}) - \frac{h}{2n} \int x^2 K(x) dx \\
&\quad \times \sum_{i=1}^{n} K''(\frac{x - X_i}{h}) + \frac{h^4}{4} f^{(4)}(x) \left[\int x^2 K(x) dx \right]^2 .
\end{aligned}$$

Therefore, the RBRK can be formulated as follows

$$\begin{aligned}
\tilde{\bar{f}}_{h,n}(x) &= \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - X_i}{h}) - \frac{h}{2n} \int x^2 K(x) dx \sum_{i=1}^{n} K''(\frac{x - X_i}{h}) \\
&\quad + \frac{h^4}{4} f^{(4)}(x) \left[\int x^2 K(x) dx \right]^2 .
\end{aligned} \tag{5}$$

To assess the new estimator's potential for bias and MSE reduction, we derive its bias and variance. This analysis requires conditions a_2 and b_1 to be satisfied, as well as the existence of the sixth-order derivative near x.

Theorem 3.1. *If conditions ii and iii hold, and the kernel K is symmetric with all its odd-order moments being zero, and f has a sixth-order derivative in the neighborhood of x, thus,*

$$\begin{aligned} Bias(\tilde{\tilde{f}}_{h,n}(x)) &= \frac{h^6}{8} f^{(6)}(x) \left[\int u^2 K(u)du\right]^3 + O(h^4), \quad (6) \\ Var(\tilde{\tilde{f}}_{h,n}(x)) &= \frac{1}{2nh} f(x) \left(\int u^2 K(u)du\right)^2 \int (K''(u))^2 du \\ &+ \frac{1}{8nh} f(x) \left(\int u^2 K(u)du\right)^4 \int (K^{(4)})^2(u)du + O(n^{-1}). \quad (7) \end{aligned}$$

Therefore,

$$MSE(\tilde{\tilde{f}}_{h,n}(x)) = O\left(h^8 + (nh)^{-1}\right).$$

So, the optimal MSE will be from $O(n^{\frac{-8}{9}})$ and $h = O(n^{\frac{-1}{9}})$ order.

Proof. A function can be approximated by polynomials in a neighborhood of any point in terms of its value and derivatives using Taylor's expansion. Consider first a function $f(x)$ of a single variable. Taylor's expansion for $f(x)$ about the point m is

$$f(x) = f(m) + f'(m)(x-m) + \frac{1}{2} f''(m)(x-m)^2 + \frac{1}{6} f'''(m)(x-m)^3 + R$$

where R is the remainder term that is smaller in magnitude than the preceding terms if x is sufficiently close to m. So, using the Taylor expansion and conversion $\frac{x-t}{h} = y$, we have

$$\begin{aligned} Bias(\tilde{\tilde{f}}_{h,n}(x)) =& \mathbb{E}\left(\tilde{\tilde{f}}_{h,n}(x)\right) - f(x) \\ =& \mathbb{E}\left(\bar{f}_{h,n}(x)\right) + \mathbb{E}\left(\frac{h^4}{4} f^{(4)}(x) \left[\int u^2 K(u)du\right]^2\right) - f(x) \\ =& \mathbb{E}\left(\bar{f}_{h,n}(x)\right) - f(x) + \mathbb{E}\left(\frac{h^4}{4} f^{(4)}(x) \left[\int u^2 K(u)du\right]^2\right) \\ =& -\frac{h^4}{4} f^{(4)}(x) \left[\int u^2 K(u)du\right]^2 + O(h^4) \\ &+ \mathbb{E}\left(\frac{h^4}{4} f^{(4)}(x) \left[\int u^2 K(u)du\right]^2\right) \\ =& -\frac{h^4}{4} f^{(4)}(x) \left[\int u^2 K(u)du\right]^2 + O(h^4) \\ &+ \frac{h^4}{4} \left[\int u^2 K(u)du\right]^2 \mathbb{E}\left(f^{(4)}(x)\right) \end{aligned}$$

$$
\begin{aligned}
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2 \mathbb{E}\left(\frac{1}{nh}\sum_{i=1}^{n}\frac{1}{h^4}K^{(4)}\left(\frac{x-X_i}{h}\right)\right)\\
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2 \left(\frac{1}{nh^5}\sum_{i=1}^{n}\mathbb{E}\left(K^{(4)}\left(\frac{x-X_i}{h}\right)\right)\right)\\
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2 \left(\frac{1}{nh^5}n\mathbb{E}\left(K^{(4)}\left(\frac{x-X_i}{h}\right)\right)\right)\\
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2 \left(\frac{1}{h^5}\int K^{(4)}\left(\frac{x-t}{h}\right)f(t)dt\right)\\
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2 \left(\frac{1}{h}\int K\left(u\right)f^{(4)}(x-hu)hdu\right)\\
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2\\
&\quad\times\left(\int K\left(u\right)\left[f^{(4)}(x) - huf^{(5)}(x) + \frac{(hu)^2}{2!}f^{(6)}(x)\right]du\right) + O(h^3)\\
&= A + \frac{h^4}{4}\left[\int u^2K(u)du\right]^2 \left(f^{(4)}(x) + \frac{h^2}{2!}f^{(6)}(x)\int u^2K(u)du\right) + O(h^3)\\
&= A + \frac{h^4}{4}f^{(4)}(x)\left[\int u^2K(u)du\right]^2 + \frac{h^6}{8}f^{(6)}(x)\left[\int u^2K(u)du\right]^3 + O(h^3)\\
&= -\frac{h^4}{4}f^{(4)}(x)\left[\int u^2K(u)du\right]^2 + O(h^4)\\
&\quad+\frac{h^4}{4}f^{(4)}(x)\left[\int u^2K(u)du\right]^2 + \frac{h^6}{8}f^{(6)}(x)\left[\int u^2K(u)du\right]^3 + O(h^3)\\
&= \frac{h^6}{8}f^{(6)}(x)\left[\int u^2K(u)du\right]^3 + O(h^4),
\end{aligned}
$$

in which

$$
A = -\frac{h^4}{4}f^{(4)}(x)\left[\int u^2K(u)du\right]^2 + O(h^4).
$$

Therefore, the bias value is equal to

$$
Bias\left(\tilde{\tilde{f}}_{h,n}(x)\right) = \frac{h^6}{8}f^{(6)}(x)\left[\int u^2K(u)du\right]^3 + O(h^4).
$$

We now calculate the variance of the estimator.

$$
\begin{aligned}
Var(\tilde{\tilde{f}}_h(x)) &= Var\left[\hat{f}_{n,h}(x) - \frac{h^2}{2!}\hat{f''}_h(x)\int u^2K(u)du\right.\\
&\quad\left.+\frac{h^4}{4}\hat{f}^{(4)}_{n,h}(x)\left[\int x^2K(x)dx\right]^2\right]\\
&\le 2Var\left(\hat{f}_{n,h}(x)\right) + 2Var\left(\frac{h^2}{2!}\hat{f''}_{n,h}(x)\int u^2K(u)du\right)
\end{aligned}
$$

$$
\begin{aligned}
&+2Var\left(\frac{h^4}{4}\hat{f}^{(4)}_{n,h}(x)\left[\int x^2K(x)dx\right]^2\right)\\
=\;& 2Var\left(\hat{f}_{n,h}(x)\right)+\frac{h^4}{2}\left(\int u^2K(u)du\right)^2 Var\left(\hat{f}''_{n,h}(x)\right)\\
&+\frac{h^8}{8}\left(\int u^2K(u)du\right)^4 Var\left(\hat{f}^{(4)}{}_{n,h}(x)\right), \qquad (8)
\end{aligned}
$$

in which,

$$
\begin{aligned}
Var\left(\hat{f}''_{n,h}(x)\right) &= Var\left[\frac{1}{nh^3}\sum_{i=1}^{n}K''\left(\frac{x-X_i}{h}\right)\right]\\
&= \frac{1}{nh^6}Var\left(K''\left(\frac{x-X_1}{h}\right)\right)\\
&= \frac{1}{nh^6}\left\{\int\left(K''(\frac{x-t}{h})\right)^2 f(t)dt-\left[\int K''\left(\frac{x-t}{h}\right)f(t)dt\right]^2\right\}\\
&= \frac{1}{nh^5}\int\left(K''(u)\right)^2 f(x-hu)du-\frac{1}{n}\left(f''(x)+O(h^2)\right)^2\\
&= \frac{1}{nh^5}\int\left(K''(u)\right)^2\left(f(x)-huf'(x)+O(h^2)\right)du\\
&\quad -\frac{1}{n}\left(f''(x)\right)^2+O\left(\frac{h^2}{n}\right)\\
&= \frac{f(x)}{nh^5}\int\left(K''(u)\right)^2du+O\left(\frac{1}{nh^4}\right). \qquad (9)
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
Var\left(\hat{f}_{n,h}{}^{(4)}(x)\right) &= Var\left[\frac{1}{nh^5}\sum_{i=1}^{n}K^{(4)}\left(\frac{x-X_i}{h}\right)\right]\\
&= \frac{1}{nh^{10}}Var\left(K^{(4)}\left(\frac{x-X_1}{h}\right)\right)\\
&= \frac{1}{nh^{10}}\left\{\int\left(K^{(4)}(\frac{x-t}{h})\right)^2 f(t)dt-\left[\int K^{(4)}\left(\frac{x-t}{h}\right)f(t)dt\right]^2\right\}\\
&= \frac{1}{nh^9}\int\left(K^{(4)}(u)\right)^2 f(x-hu)du-\frac{1}{n}\left(f^{(4)}(x)+O(h^4)\right)^2\\
&= \frac{1}{nh^9}\int\left(K^{(4)}(u)\right)^2\left(f(x)-huf'(x)+O(h^2)\right)du\\
&\quad -\frac{1}{n}\left(f^{(4)}(x)\right)^2+O(h^8)\\
&= \frac{f(x)}{nh^9}\int\left(K^{(4)}(u)\right)^2du+O(\frac{1}{nh^8}). \qquad (10)
\end{aligned}
$$

By substituting (9) and (10) into (8), we obtain

$$\begin{aligned} Var(\tilde{\tilde{f}}_h(x)) = & Var\left[\hat{f}_{n,h}(x) - \frac{h^2}{2!}\hat{f''}_h(x)\int u^2K(u)du + \frac{h^4}{4}\hat{f}_{n,h}{}^{(4)}(x)\left[\int x^2K(x)dx\right]^2\right] \\ & \leq b\frac{1}{2nh}f(x)\left(\int u^2K(u)du\right)^2\int (K''(u))^2du \\ & + \frac{1}{8nh}f(x)\left(\int u^2K(u)du\right)^4\int (K^{(4)}(u))^2du + O(n^{-1}) \end{aligned}$$

Therefore, it can be easily seen that the smoothness parameter is equal to $h = O(n^{\frac{-1}{9}})$. The optimal MSE will also be optimal in order$O(n^{\frac{-8}{9}})$. Hence, the proof is finished. □

To further reduce the bias of the estimator in (4), we modify the standard kernel density estimator by subtracting its bias term. Typically, this involves using a parameter ah, where a is a positive value. However, because the exact bias expression is unknown and depends on the unknown density function f, we substitute an estimated value for f. Hence,

$$\overset{\frown}{f}_{ah,n}(x) = \hat{f}_{ah,n}(x) - \widehat{Bias}(\hat{f}_{ah,n}(x)), \qquad a > 0.$$

Therefore, the BRKp estimator will be as follows

$$\begin{aligned} \overset{\frown}{f}_{ah,n}(x) &= \hat{f}_{ah,n}(x) - \widehat{Bias}(\hat{f}_{ah,n}(x)) \\ &= \hat{f}_{ah,n}(x) - \frac{(ah)^2}{2}\hat{f}''_{ah,n}(x)\int y^2K(y)dy \\ &= \frac{1}{nah}\sum_{i=1}^{n} K\left(\frac{x-X_i}{ah}\right) - \frac{ah}{2n}\int x^2K(x)dx\sum_{i=1}^{n} K''\left(\frac{x-X_i}{ah}\right). \end{aligned}$$

To determine if the new estimator can effectively reduce bias and MSE, we calculate its bias and variance. This requires conditions a_1 and b_1 to be satisfied, and the density function f to possess a fourth-order derivative near x.

Theorem 3.2. *If the conditions a_2 and b_1 are accurate and the kernel K is symmetric and all its odd moments are zero, and f has a fourth-order derivative in the neighborhood of x, then*

$$\begin{aligned} Bias(\overset{\frown}{f}_{ah,n}(x)) &= -\frac{(ah)^4}{2!2!}f^{(4)}(x)\left[\int y^2K(y)dy\right]^2 + O(h^4), \\ Var(\overset{\frown}{f}_{ah,n}(x)) &= \frac{1}{2nah}f(x)\left(\int y^2K(y)dy\right)^2\int (K'')^2(y)dy + O(n^{-1}). \end{aligned}$$

As a result

$$MSE(\overset{\frown}{f}_{ah,n}(x)) = O\left(h^8 + (nh)^{-1}\right).$$

So, the optimal MSE will be from $O(n^{\frac{-8}{9}})$ and $h = O(n^{\frac{-1}{9}})$ order (Salehi et al., 2018).

Proof. Using the Taylor expansion and conversion $\frac{x-z}{h}=y$, we have

$$\begin{aligned}
E\widehat{f}_{ah}(x) &= E\hat{f}_{ah}(x)-\frac{(ah)^2}{2!}E(\hat{f''}_{ah}(x))\int x^2K(x)dx\\
&= f(x)+\frac{(ah)^2}{2!}f''(x)\int x^2K(x)dx+O(h^4)\\
&\quad -\frac{(ah)^2}{2!}E(\hat{f''}_{ah}(x))\int x^2K(x)dx,\\
E(\hat{f''}_{ah}(x)) &= E\left[\frac{1}{anh(ah)^2}\sum_{i=1}^{n}K''\left(\frac{x-X_i}{ah}\right)\right]\\
&= \frac{1}{(ah)^3}\int K''(\frac{x-t}{ah})f(t)dt\\
&= \frac{1}{ah}\int K(\frac{x-t}{ah})f''(t)dt=\int K(u)f''(x-ahu)du\\
&= \int K(u)\left[f''(x)-ahuf'''(x)+\frac{(ahu)^2}{2!}f''''(x)-\frac{(ahu)^3}{3!}f'''''(x)\right]du\\
&\quad +O(h^4)\\
&= f''(x)+\frac{(ah)^2}{2}f''''(x)\int u^2K(u)du+O(h^4).
\end{aligned}$$

So,

$$\begin{aligned}
E\widehat{f}_{ah}(x) &= f(x)+\frac{(ah)^2}{2!}f''(x)\int x^2K(x)dx+O(h^4)\\
&\quad -\frac{(ah)^2}{2!}\int x^2K(x)dx\left[f''(x)+\frac{(ah)^2}{2!}f''''(x)\int u^2K(u)du\right]\\
&\quad +O(h^4)\\
&= f(x)-\frac{(ah)^2(ah)^2}{2!2!}f''''(x)\left[\int u^2K(u)du\right]^2+O(h^4).
\end{aligned}$$

Therefore, the bias value is equal to

$$Bias(\widehat{f}_{ah}(x))=-\frac{(ah)^4}{2!2!}f''''(x)\left[\int u^2K(u)du\right]^2+O(h^4).$$

On the other hand, the estimator variance is equal to

$$\begin{aligned}
Var(\widehat{f}_{ah}(x)) &= Var\left[\hat{f}_{ah}(x)-\frac{(ah)^2}{2!}\hat{f''}_{ah}(x)\int u^2K(u)du\right]\\
&\leq 2Var\left(\hat{f}_{ah}(x)\right)+2Var\left(\frac{(ah)^2}{2!}\hat{f''}_{ah}(x)\int u^2K(u)du\right)\\
&= 2Var\left(\hat{f}_{ah}(x)\right)+\frac{(ah)^4}{2}\left(\int u^2K(u)du\right)^2 Var\left(\hat{f''}_{ah}(x)\right). \quad (11)
\end{aligned}$$

In which,

$$\begin{aligned}
Var\left(\hat{f''}_{ah}(x)\right) &= Var\left[\frac{1}{n(ah)^3}\sum_{i=1}^{n}K''\left(\frac{x-X_i}{ah}\right)\right] \\
&= \frac{1}{n(ah)^6}Var\left(K''\left(\frac{x-X_1}{ah}\right)\right) \\
&= \frac{1}{n(ah)^6}\left\{\int\left(K''(\frac{x-t}{ah})\right)^2 f(t)dt - \left[\int K''\left(\frac{x-t}{ah}\right)f(t)dt\right]^2\right\} \\
&= \frac{1}{n(ah)^5}\int\left(K''(u)\right)^2 f(x-ahu)du - \frac{1}{n}\left(f''(x)+O(h^2)\right)^2 \\
&= \frac{1}{n(ah)^5}\int\left(K''(u)\right)^2\left(f(x)-ahuf'(x)+O(h^2)\right)du \\
&\quad -\frac{1}{n}\left(f''(x)\right)^2 + O(\frac{h^2}{n}) \\
&= \frac{f(x)}{n(ah)^5}\int\left(K''(u)\right)^2 du + O(\frac{1}{nh^4}). \qquad (12)
\end{aligned}$$

By substituting (12) into (11), we obtain

$$\begin{aligned}
Var(\widehat{f}_{ah}(x)) &\leq 2\left(\frac{f(x)}{nah}\int K^2(u)du + O(\frac{1}{nh})\right) \\
&\quad +\frac{(ah)^4}{2}\left[\int u^2K(u)du\right]^2\frac{f(x)}{n(ah)^5}\int(K''(u))^2du + O(\frac{1}{nh^4}) \\
&= \frac{f(x)}{2anh}\left[\int u^2K(u)du\right]^2\int\left(K''(u)\right)^2du + O(\frac{1}{n}).
\end{aligned}$$

Therefore, it can be easily seen that the smoothness parameter is equal to $h = O(n^{\frac{-1}{9}})$. The optimal MSE will also be optimal in order $O(n^{\frac{-8}{9}})$. The proof is completed. □

Because the BRKp estimator in (5) can produce negative values at certain boundary points x, therefore, to obtain a positive density estimator we redefine it as follows

$$\tilde{f}_{ah,n}(x) = \frac{\widehat{f}_{ah,n}(x)I_{(\widehat{f}_{ah,n}(x)>0)}}{\int\widehat{f}_{ah,n}(x)I_{(\widehat{f}_{ah,n}(x)>0)}dx}.$$

In which I_A is an indicator function on set A. The BRK estimator, as proven by (5), converges in probability to $f(x)$. Notably, the $\tilde{f}_{ah,n}$ estimator mirrors this performance and possesses similar properties, particularly as the sample size increases.

4 Simulation study

In this section, we first show using a simulation study that minimizing $MSE(\widehat{f}_{ah}(x))$ results from a reliable choice for the parameter a as well as with the help of Monte Carlo simulations, we examine and compare the resulting estimators. Therefore, we introduce any value a that minimizes the MSE of the proposed estimator as the optimal estimator for bias reduction. Because it is impossible to minimize MSE by derivation, we try to minimize it by using numerical methods. Since the amount of bias is equal to the area occupied between the actual density and the estimated density, using numerical methods such as the rectangular method and the Simpson method Burden and Faires (2010), the optimal value of a is obtained by minimizing the MSE. The data used in the simulation are generated from the standard normal distribution. The kernel K is also considered to be the standard normal density. According to Figures 7 and 8, optimal values for the parameter a in the rectangular and Simpson methods in the BRKp estimator were 0.58 and 0.69, respectively. However, since the Simpson method is more accurate, we recommend a value of 0.69 in the BRKp estimator for the optimal value a. Since the proposed estimators are based on symmetric kernels, and the optimal value of a in the simulation is also obtained under the assumption of symmetry, the optimal value of a remains constant and does not change. As shown in Figures 9 to 13, derived from real data, the estimators are presented with a fixed value of a, effectively demonstrating the robustness of our approach. Therefore, the proposed final estimates of BRKp will be as follows

$$\tilde{f}_{ah}(x)=\frac{1}{0.69nh}\sum_{i=1}^{n}K\left(\frac{x-X_i}{0.69h}\right)-\frac{0.69h}{2n}\int x^2K(x)dx\sum_{i=1}^{n}K''\left(\frac{x-X_i}{0.69h}\right). \quad (13)$$

This estimator significantly reduces the bias value compared to the UK, BRK and RBRK estimators.

In the continuation of this section, we study the amount of bias reduction of four estimators, UK, BRK, RBRK and BRKp. To investigate the effect of sample size factor on reducing bias and the accuracy of the density function estimator, we perform a simulation study as follows. We produce random samples from the standard normal symmetric distribution in sizes of 15, 30, 60, 120, 240 and 480. Each of these samples is repeated 1000 times, and the bias reduction and MSE are measured using a standard normal kernel.

Considering that the standard normal distribution is symmetric around zero, the value of x is considered equal to 0, 0.5, 1, 1.5, 2, 2.5, 3, and 3.5, and the kernel estimators are calculated and compared at these points. We also select the standard normal as a function of the K-kernel. Since the kernel is symmetric, we choose the smoothness parameter for the UK estimator equal to $h=n^{\frac{-1}{5}}$ and the BRK, RBRK and BRKp estimators equal to$h=n^{\frac{-1}{9}}$. The bias, variance, and MSE of estimators are estimated as follows

$$\begin{aligned}
Bias(\hat{f}) &= \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{f}_i-f\right),\\
Var(\hat{f}) &= \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{f}_i-\bar{\hat{f}}\right)^2, \qquad \text{with} \qquad \bar{\hat{f}}=\frac{1}{1000}\sum_{i=1}^{1000}\hat{f}_i,\\
MSE(\hat{f}) &= \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{f}_i-f\right)^2,
\end{aligned}$$

which f is the actual density $f(x)$ and $\hat{f}$ is each of the estimators $\hat{f}_h(x)$ of (1), $\bar{f}_{h,n}(x)$ (4), $\tilde{\tilde{f}}_{h,n}(x)$ (5) and $\widehat{f}_{ah,n}(x)$ (13). We assume that f is equal to the standard normal density without losing the totality. The simulation results are summarized in Tables 2-7 and Figures 1 to 6.

Table 2: Bias, variance and MSE of estimators, $n = 15$.

		Kernel estimators			
x	Criterion	UK	BRK	RBRK	BRKp
0	Bias	-0.052074	-0.024230	0.006224	-0.001255
	Variance	0.003744	0.006997	0.006991	0.013157
	MSE	0.006456	0.007584	0.007029	0.013159
0.5	Bias	-0.036816	-0.010440	0.007606	-0.000721
	Variance	0.004148	0.006409	0.006409	0.012282
	MSE	0.005504	0.006518	0.006466	0.012282
1	Bias	-0.004981	0.005709	0.009415	0.000710
	Variance	0.003763	0.006205	0.006202	0.010398
	MSE	0.003788	0.006237	0.006290	0.010398
1.5	Bias	0.019204	0.011304	0.013075	0.009175
	Variance	0.002944	0.004506	0.004509	0.007728
	MSE	0.003313	0.004634	0.004679	0.007812
2	Bias	0.023887	0.010227	-0.000336	0.003135
	Variance	0.001767	0.002317	0.002313	0.003251
	MSE	0.002338	0.002421	0.002313	0.003260
2.5	Bias	0.014595	0.003048	-0.006714	0.000115
	Variance	0.000678	0.001075	0.001075	0.001231
	MSE	0.000891	0.001084	0.001120	0.001231
3	Bias	0.007514	-0.001499	-0.008649	-0.001048
	Variance	0.000238	0.000342	0.000343	0.000354
	MSE	0.000294	0.000344	0.000417	0.000355
3.5	Bias	0.002486	-0.001983	-0.005106	-0.000549
	Variance	0.000058	0.000089	0.000089	0.000075
	MSE	0.000064	0.000093	0.000115	0.000075

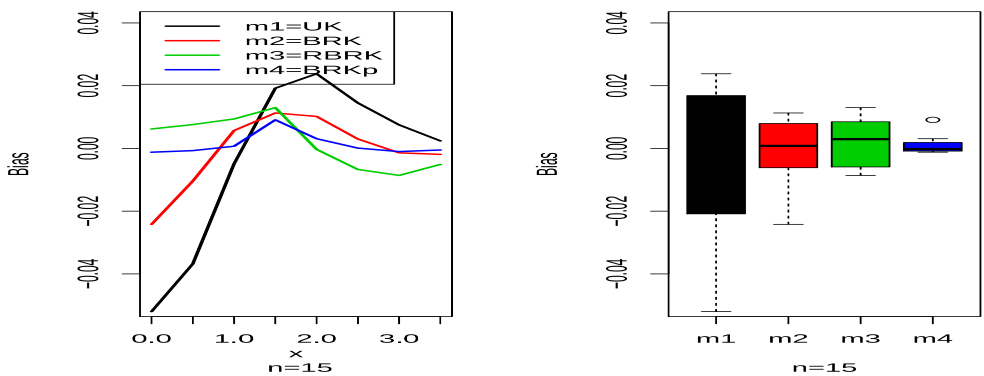

Figure 1: Linear diagram and box of bias values of different kernel density estimators with $n = 15$.

Table 3: Bias, variance and MSE of estimators, $n = 30$.

		Kernel estimators			
x	Criterion	UK	BRK	RBRK	BRKp
0	Bias	-0.063317	-0.026229	-0.012151	-0.006537
	Variance	0.004470	0.005051	0.005282	0.006930
	MSE	0.007134	0.005652	0.005421	0.006953
0.5	Bias	-0.046184	-0.017296	-0.003758	-0.002289
	Variance	0.004963	0.004930	0.004722	0.006534
	MSE	0.006604	0.004983	0.004740	0.006549
1	Bias	-0.008442	0.004502	0.004267	0.002083
	Variance	0.004296	0.004232	0.004276	0.005406
	MSE	0.004309	0.004238	0.004281	0.005407
1.5	Bias	0.016580	0.008091	0.008478	0.005722
	Variance	0.003172	0.003147	0.003116	0.004356
	MSE	0.003446	0.003194	0.003154	0.004367
2	Bias	0.020689	0.005262	0.002422	0.001235
	Variance	0.002164	0.002067	0.002059	0.002845
	MSE	0.002409	0.002086	0.002061	0.002854
2.5	Bias	0.013529	0.002376	-0.001307	0.000205
	Variance	0.001210	0.001140	0.001171	0.002279
	MSE	0.001288	0.001148	0.001183	0.002280
3	Bias	0.007194	-0.000369	-0.003470	-0.000685
	Variance	0.000664	0.000627	0.000628	0.000810
	MSE	0.000715	0.000630	0.000636	0.000811
3.5	Bias	0.002725	-0.001281	-0.003107	-0.000336
	Variance	0.000197	0.000161	0.000153	0.000268
	MSE	0.000204	0.000162	0.000158	0.000268

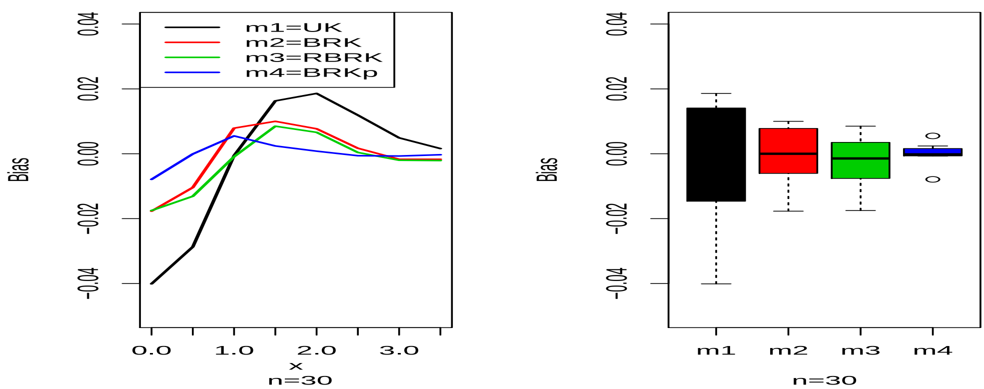

Figure 2: Linear diagram and box of bias values of different kernel density estimators with $n = 30$.

As noted earlier, we chose symmetric kernels because the standard normal kernel is both widely used and easy to implement. However, employing non-normal kernels could produce different outcomes and represents a promising direction for future research.

Table 4: Bias, variance and MSE of estimators, $n = 60$.

		Kernel estimators			
x	Criterion	UK	BRK	RBRK	BRKp
0	Bias	-0.035228	-0.012387	-0.004322	-0.002136
	Variance	0.002258	0.002474	0.002531	0.003375
	MSE	0.003245	0.002711	0.002573	0.003387
0.5	Bias	-0.025218	-0.008820	-0.001921	-0.001185
	Variance	0.002602	0.002517	0.002424	0.003276
	MSE	0.003208	0.002551	0.002438	0.003290
1	Bias	-0.004171	0.003246	0.002379	0.001374
	Variance	0.002024	0.002086	0.002128	0.002877
	MSE	0.002029	0.002090	0.002137	0.002878
1.5	Bias	0.010255	0.004519	0.005051	0.003517
	Variance	0.001527	0.001537	0.001526	0.002527
	MSE	0.001689	0.001555	0.001535	0.002535
2	Bias	0.013208	0.003112	0.001232	0.000634
	Variance	0.001058	0.001049	0.001039	0.001581
	MSE	0.001140	0.001054	0.001041	0.001586
2.5	Bias	0.008268	0.000503	-0.002348	0.000042
	Variance	0.000618	0.000582	0.000592	0.001492
	MSE	0.000719	0.000586	0.000595	0.001494
3	Bias	0.004387	-0.001139	-0.002109	-0.000302
	Variance	0.000302	0.000279	0.000276	0.000554
	MSE	0.000329	0.000280	0.000278	0.000555
3.5	Bias	0.001707	-0.000723	-0.002051	-0.000132
	Variance	0.000091	0.000084	0.000080	0.000168
	MSE	0.000095	0.000085	0.000081	0.000168

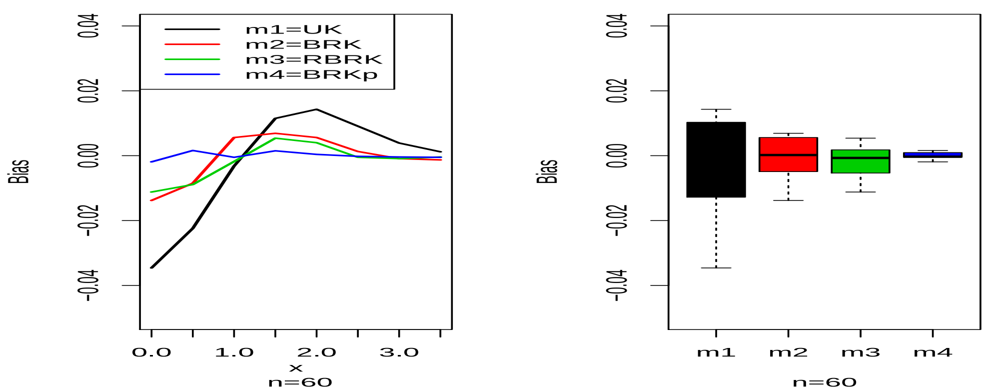

Figure 3: Linear diagram and box of bias values of different kernel density estimators with $n = 60$.

As can be seen from Tables 2-7 and the corresponding line and box plots (Figures 1-6), the performance of the UK method is weaker in terms of bias compared to the other methods. The BRK and RBRK methods exhibit relatively similar behavior, although

Table 5: Bias, variance and MSE of estimators, $n = 120$.

		Kernel estimators			
x	Criterion	UK	BRK	RBRK	BRKp
0	Bias	-0.014801	-0.004487	-0.002073	-0.001037
	Variance	0.001164	0.001265	0.001279	0.001802
	MSE	0.001551	0.001411	0.001308	0.001811
0.5	Bias	-0.010528	-0.003254	-0.001338	-0.000740
	Variance	0.001346	0.001293	0.001188	0.001854
	MSE	0.001532	0.001303	0.001197	0.001863
1	Bias	-0.002061	0.002031	0.001354	0.000827
	Variance	0.001095	0.001123	0.001149	0.001956
	MSE	0.001098	0.001125	0.001153	0.001958
1.5	Bias	0.006174	0.002850	0.003166	0.002247
	Variance	0.000883	0.000887	0.000876	0.001344
	MSE	0.000974	0.000893	0.000883	0.001347
2	Bias	0.008429	0.002160	0.000856	0.000450
	Variance	0.000568	0.000563	0.000556	0.000986
	MSE	0.000642	0.000567	0.000557	0.000988
2.5	Bias	0.005107	-0.000477	-0.001228	-0.000348
	Variance	0.000292	0.000281	0.000278	0.000722
	MSE	0.000331	0.000284	0.000279	0.000723
3	Bias	0.003127	-0.001395	-0.002340	-0.000277
	Variance	0.000130	0.000124	0.000122	0.000369
	MSE	0.000157	0.000125	0.000123	0.000370
3.5	Bias	0.001246	-0.000730	-0.002111	-0.000149
	Variance	0.000039	0.000037	0.000036	0.000226
	MSE	0.000042	0.000037	0.000037	0.000226

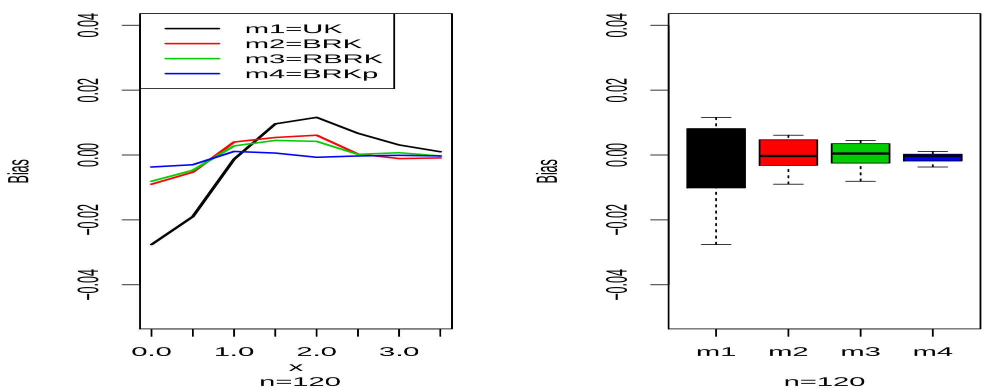

Figure 4: Linear diagram and box of bias values of different kernel density estimators with $n = 120$.

the RBRK method achieves slightly lower bias. Among all methods, the proposed BRKp method demonstrates the best performance in reducing bias. As expected, increasing the sample size leads to a noticeable reduction in bias, which is clearly

Table 6: Bias, variance and MSE of estimators, $n = 240$.

		Kernel estimators			
x	Criterion	UK	BRK	RBRK	BRKp
0	Bias	-0.021518	-0.009260	-0.008328	-0.000817
	Variance	0.000799	0.000770	0.000741	0.001600
	MSE	0.001262	0.000856	0.000810	0.001601
0.5	Bias	-0.013771	-0.004502	-0.004428	-0.000544
	Variance	0.000740	0.000700	0.000663	0.001282
	MSE	0.000929	0.000720	0.000683	0.001283
1	Bias	-0.000329	0.002509	0.001937	0.001424
	Variance	0.000582	0.000604	0.000596	0.001036
	MSE	0.000582	0.000611	0.000599	0.001038
1.5	Bias	0.008253	0.005590	0.004261	0.001569
	Variance	0.000392	0.000389	0.000332	0.000654
	MSE	0.000460	0.000421	0.000350	0.000656
2	Bias	0.008219	0.002702	0.001594	0.000849
	Variance	0.000199	0.000190	0.000154	0.000283
	MSE	0.000266	0.000198	0.000157	0.000284
2.5	Bias	0.005188	0.000181	0.000065	0.000036
	Variance	0.000072	0.000065	0.000042	0.000099
	MSE	0.000099	0.000065	0.000042	0.000099
3	Bias	0.002157	-0.000884	0.000029	-0.000137
	Variance	0.000018	0.000021	0.000012	0.000028
	MSE	0.000023	0.000022	0.000012	0.000028
3.5	Bias	0.000659	-0.000852	-0.000496	-0.000207
	Variance	0.000004	0.000004	0.000003	0.000005
	MSE	0.000004	0.000005	0.000003	0.000005

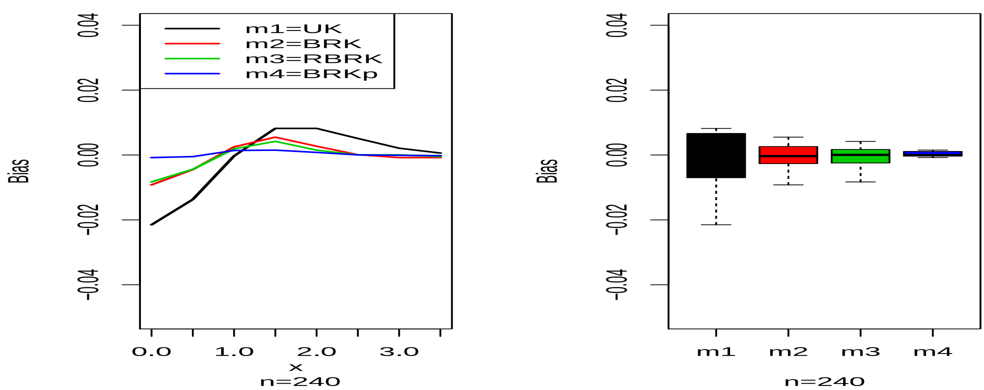

Figure 5: Linear diagram and box of bias values of different kernel density estimators with $n = 240$.

evident from the aforementioned tables and figures. Moreover, it can be observed that the BRKp method consistently provides superior bias reduction relative to the other methods, even as the sample size increases. Therefore according to Tables 2-7 and

Table 7: Bias, variance and MSE of estimators, $n = 480$.

		Kernel estimators			
x	Criterion	UK	BRK	RBRK	BRKp
0	Bias	-0.016498	-0.005508	0.004829	-0.002406
	Variance	0.000511	0.000447	0.000405	0.000746
	MSE	0.000783	0.000477	0.000428	0.000752
0.5	Bias	-0.011999	-0.003818	-0.002781	-0.000814
	Variance	0.000470	0.000411	0.000394	0.000718
	MSE	0.000614	0.000426	0.000402	0.000719
1	Bias	-0.000117	0.002201	0.002013	-0.001498
	Variance	0.000386	0.000363	0.000314	0.000535
	MSE	0.000386	0.000368	0.000318	0.000537
1.5	Bias	0.005989	0.003469	0.003161	0.001764
	Variance	0.000228	0.000247	0.000197	0.000359
	MSE	0.000264	0.000259	0.000207	0.000362
2	Bias	0.006888	0.002464	0.000956	0.000450
	Variance	0.000107	0.000103	0.000082	0.000136
	MSE	0.000154	0.000109	0.000083	0.000136
2.5	Bias	0.004140	-0.000224	0.000176	-0.000132
	Variance	0.000037	0.000040	0.000031	0.000055
	MSE	0.000054	0.000040	0.000031	0.000055
3	Bias	0.001657	-0.000612	-0.000511	-0.000148
	Variance	0.000010	0.000012	0.000009	0.000013
	MSE	0.000013	0.000012	0.000009	0.000013
3.5	Bias	0.000388	-0.000477	-0.000342	-0.000077
	Variance	0.000002	0.000002	0.000002	0.000003
	MSE	0.000002	0.000002	0.000002	0.000003

Figures 1-6, the BRKp estimator has a better performance in reducing bias than the UK, BRK and RBRK estimators.

The rectangular and Simpson's rules are two widely used numerical integration methods for approximating definite integrals. In this study, we utilized both methods to obtain the optimal value of the parameter a in the BRKp estimator. The optimal value of a was determined by minimizing the Mean Squared Error (MSE) of the BRKp estimator using these two numerical approaches.

The simulation results based on the rectangular and Simpson's rules are presented in Figures 7 and 8, respectively. According to the findings, the optimal value of a was found to be 0.58 using the rectangular rule and 0.69 using Simpson's rule. Since Simpson's method provides higher accuracy compared to the rectangular rule, the optimal value of a was set to 0.69 for subsequent analyses. Section 2 notes that the various kernels exhibit very similar performance, although it is suggested that future research explore skewed or heavy-tailed kernels. Concerning the optimal value $a = 0.69$ obtained via the Simpson method, the near-equivalence of kernel performance implies that switching kernels may alter a only marginally and will not substantially change the overall results.

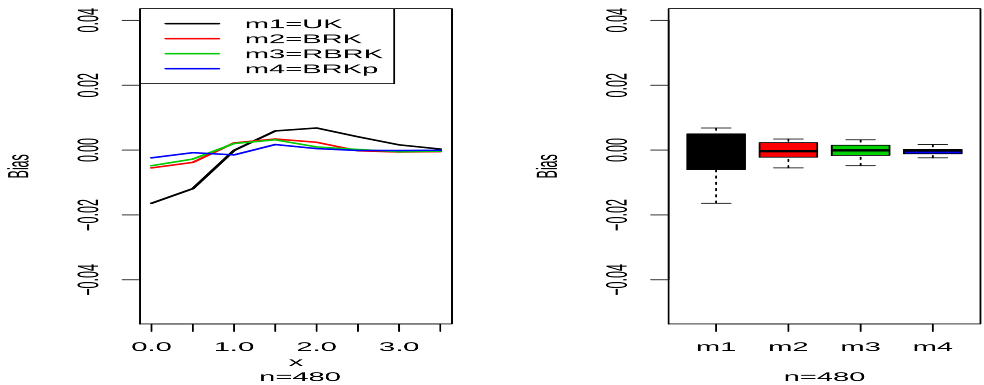

Figure 6: Linear diagram and box of bias values of different kernel density estimators with $n = 480$.

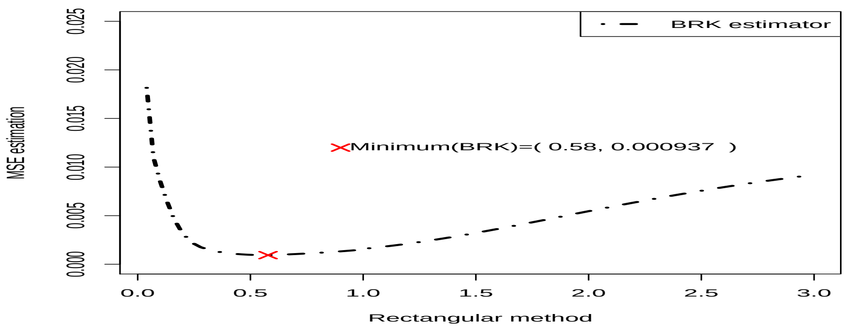

Figure 7: Graph of the MSE function as a function of the parameter using the rectangular method.

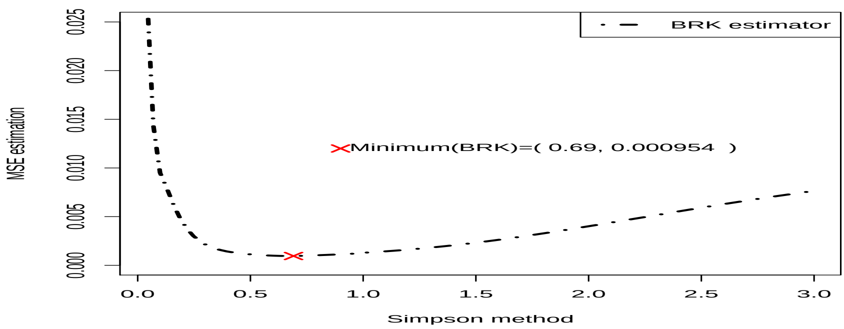

Figure 8: Graph of MSE function as a function of parameter using Simpson method.

5 Real data analysis

This section considers four sets of real data and estimates the probability density function of this data with the estimators' help. The first set contains 272 hot springs data from Farrell (1972) and Azzalini and Bowman (1990). The second set includes 62 treated patients reported in Rosner (2006) (data set 8). The third set contains 60 data showing the number of families and has been reported Rosner (2006) (data set 22). The fourth set contains 63 values measured in a physics laboratory in London on the strength of 1.5 cm glass fibers, as reported by Barreto-Souza et al. (2010). The fifth set includes 29 data from the first new deaths COVID-19 per million in the United States in June 2022, as reported by Our World in Data. Hot spring data were measured for two variables of waiting time between eruptions and eruption duration in minutes in Yellowstone National Park. Based on three estimators, UK, BRK, RBRK and BRKp, we estimate the next eruption's waiting for time density based on this data. The results are shown in Figures 9-13. As can be seen from Figure 9, the two methods UK, BRK and RBRK, work similarly, while the BRKp way is more accurate and less biased than the previous two methods. Similarly, the other three datasets' results are shown in Figures 10 to 13.

Figures 9-13 present various examples based on real datasets. These examples include data that are approximately normally distributed (Figure 12), skewed distributions (Figures 10, 11, and 13), and data with multiple modes (Figure 9). As observed across all five figures with different distribution patterns, the RBRK and BRKp methods cover a larger area of the data histograms compared to the other methods. This indicates that the proposed methods exhibit lower bias and perform more effectively in bias reduction. Moreover, they show better overlap with the data histograms compared to the `UK` and `BRK` methods. This suggests that the proposed `RBRK` and `BRKp` methods have lower bias than the other estimators and perform effectively in reducing bias.

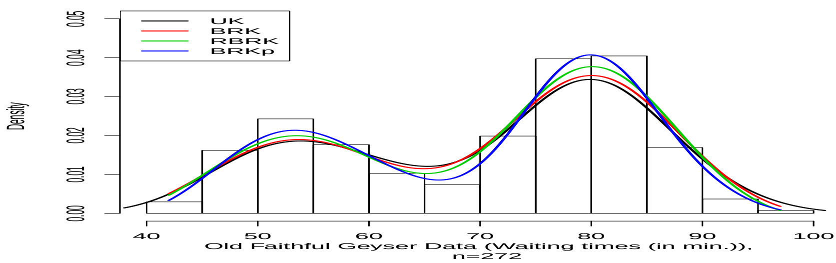

Figure 9: Histogram of real hot spring data and fitted densities.

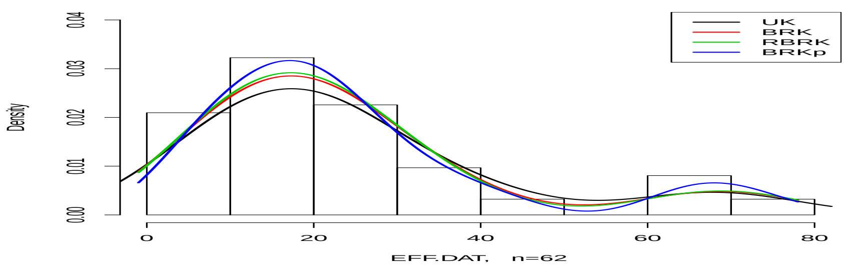

Figure 10: Histogram of eff data and fitted densities.

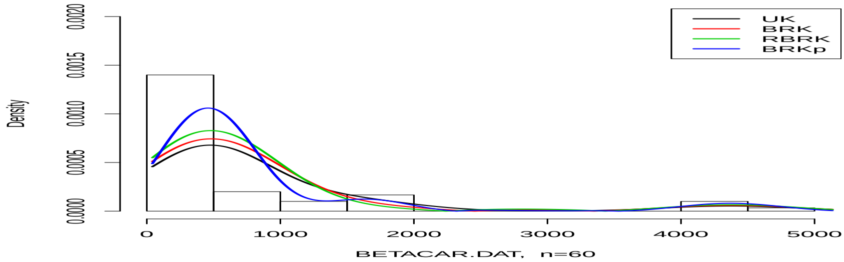

Figure 11: Histogram of Betacar data and fitted densities.

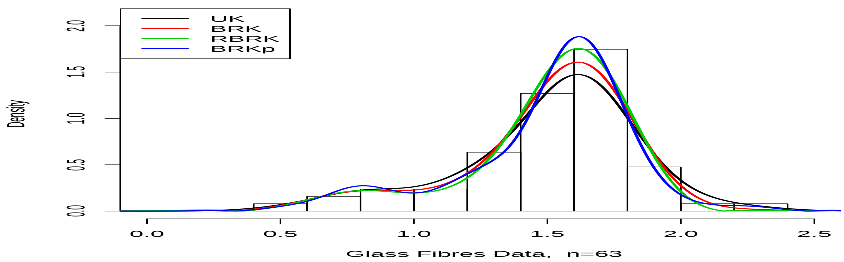

Figure 12: Histogram diagram of real data on fiberglass strength and fitted densities.

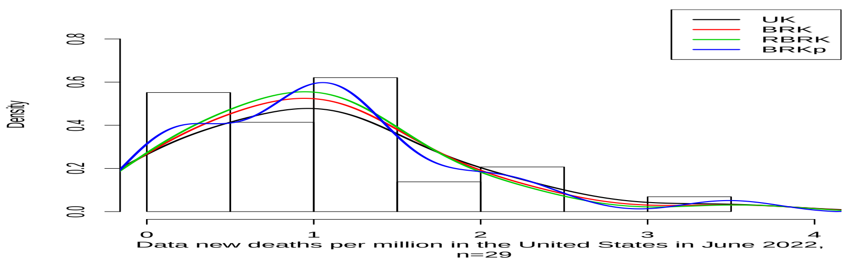

Figure 13: Histogram diagram of real data on deaths per million in the United States and fitted densities.

6 Conclusion and future research

In this paper, four UK, BRK, RBRK and BRKp methods were studied to evaluate the degree of bias. Our primary objective is to introduce new methods for kernel density estimation to reduce bias, based on the work of Xie and Wu (2014). While the MSE convergence rate of our estimator remains unchanged compared to that of Xie and Wu (2014), the bias has been significantly reduced. In the second section, we reviewed the conventional kernel density estimator, which has a bias of order $O(h^3)$ and an optimal mean squared error (MSE) of order $O(n^{-4/5})$. In the third section, we proposed two improved estimators, namely RBRK and BRKp. The BRKp estimator is designed to further reduce the bias, achieving a bias of order $O(h^4)$, with an optimal MSE of order $O(n^{-8/9})$. By modifying the smoothing parameter h to ah and determining the optimal value for the variable parameter a (through simulation), the bias of the estimator was further reduced. The validity of our claim is confirmed through the examination of Tables 1–6 and Figures 1–6 from simulations, as well as Figures 9-13 from real data, in comparison with existing studies. As observed, our proposed estimators provide better coverage of the histogram, which inherently indicates a reduction in bias. Similarly, box plots in Figures 1-6 also demonstrate bias reduction in the proposed estimators. While this paper does not address the limitations of Xie and Wu (2014) as the baseline study, and those limitations remain present, we have nevertheless introduced a kernel density estimation method with lower bias compared to existing methods, which represents a significant improvement. For future research, it is recommended to explore the kernel function K in an asymmetric form to further improve the performance of the proposed BRKp estimator.

Acknowledgement

The authors thank the anonymous reviewers and the Editor for their valuable comments and constructive suggestions, which have significantly improved the quality of this manuscript.

References

Azzalini, A. and Bowman, A.W. (1990). A look at some data on the old faithful geyser. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **39**(3):357–365.

Barreto-Souza, W., Santos, A.H. and Cordeiro, G.M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, **80**(2):159–172.

Burden, R.L. and Faires, J.D. (2010). *Numerical Analysis*. 9th edition, Cengage Learning, Brooks/Cole.

Calonico, S., Cattaneo, M.D. and Farrell, M.H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, **113**(522):767–779.

Dhaker, H., Deme, E.H. and Ciss, Y. (2021). β-divergence loss for the kernel density estimation with bias reduced. *Statistical Theory and Related Fields*, **5**(3):221–231.

El-Sayyad, G.M., Samiuddin, M. and Abdel-Ghaly, A.A. (1993). A new kernel density estimate. *Journal of Nonparametric Statistics*, **3**(1):1–11.

Farrell, R.H. (1972). On the best obtainable asymptotic rates of convergence in estimation of a density function at a point. *The Annals of Mathematical Statistics*, **43**(1): 170–180.

Härdle, W. (1991). *Smoothing Techniques: with Implementation in S*. Springer-Verlag New York.

Igarashi, G. and Kakizawa, Y. (2015). Bias corrections for some asymmetric kernel estimators. *Journal of Statistical Planning and Inference*, **159**:37–63.

Kim, C., Kim, W. and Park, B.U. (2003). Skewing and generalized jackknifing in kernel density estimation. *Communications in Statistics-Theory and Methods*, **32**(11):2153–2162.

Makhdoom, I. and Yaghoobzadeh Shahrastani, S. (2023). Inference on the probability density and cumulative distribution functions for conducting an efficient estimation on the inverse Weibull distribution. *Journal of Statistical Modelling: Theory and Applications*, **4**(2):159–169.

Marron, J.S. and Ruppert, D. (1992). Transformations to reduce boundary bias in kernel density estimation. *Journal of the Royal Statistical Society: Series B*, **56**(4):653–671.

Mynbaev, K. and Martins-Filho, C. (2010). Bias reduction in kernel density estimation via Lipschitz condition. *Journal of Nonparametric Statistics*, **22**(2):219–235.

Mynbaev, K., Nadarajah, S., Withers, C.S. and Aipenova, A.S. (2014). Improving bias in kernel density estimation. *Statistics & Probability Letters*, **94**:106–112.

Rosner, B.A. (2006). *Fundamentals of Biostatistics.* Thomson-Brooks/Cole.

Ruppert, D. and Cline, D.B. (1994). Bias reduction in kernel density estimation by smoothed empirical transformations. *The Annals of Statistics*, **22**(1):185–210.

Salehi, R. and Shadrokh, A. (2018). The bias reduction in density estimation using a geometric extrapolated kernel estimator. *Hacettepe Journal of Mathematics and Statistics*, **47**(4):1003–1021.

Salehi, R., Shadrokh, A. and Yarmohammadi, M. (2018). The relative improvement of bias reduction in density estimator using geometric extrapolated kernel. *Mathematical Researches*, **4**(2):211–228.

Samiuddin, M. and El-Sayyad, G. (1990). On nonparametric kernel density estimates. *Biometrika*, **77**(4):865–874.

Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization.* Wiley.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall.

Slaoui, Y. (2018). Bias reduction in kernel density estimation. *Journal of Nonparametric Statistics*, **30**(2):505–522.

Tsuruta, Y. (2024). Bias correction for kernel density estimation with spherical data. *Journal of Multivariate Analysis*, **203**:105338.

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

Withers, C.S. and Nadarajah, S. (2023). New classes of density estimates of low bias. *Journal of Computational and Applied Mathematics*, **429**:1115232.

Xie, X. and Wu, J. (2014). Some improvement on convergence rates of kernel density estimator. *Applied Mathematics*, **5**(11):1684–1696.

Yoon, S., Park, F., Yun, G., Kim, I. and Noh, Y.K. (2023). Variational weighting for kernel density ratios. *Advances in Neural Information Processing Systems*, **36**:5010-5027.

Zahnit, A., Sayah, A. and Yahia, D. (2024). On bias-corrected transformed kernel density estimator for heavy-tailed distributions. *Studies in Engineering and Exact Sciences*, **5**(2), e7808.