sustainability

MDPI

Article

Probabilistic Load Flow Analysis Using Nonparametric Distribution

Li Bin [1], Rashana Abbas [2], Muhammad Shahzad [2,*] and Nouman Safdar [3]

[1] School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China; direcfish163@gmail.com
[2] Department of Electrical Engineering and Technology, Muhammad Nawaz Sharif University of Engineering and Technology, Multan 66000, Pakistan; electricalengineer718@gmail.com
[3] Department of Electrical Engineering and Technology, Government College University Faisalabad, Faisalabad 38000, Pakistan; noumansafdar@gcuf.edu.pk
* Correspondence: shahzadpansota@hotmail.com

Abstract: In the pursuit of sustainable energy solutions, this research addresses the critical need for accurate probabilistic load flow (PLF) analysis in power systems. PLF analysis is an essential tool for estimating the statistical behavior of power systems under uncertainty. It plays a vital part in power system planning, operation, and dependability studies. To perform accurate PLF analysis, this article proposes a Kernel density estimation with adaptive bandwidth for probability density function (PDF) estimation of power injections from sustainable energy sources like solar and wind, reducing errors in PDF estimation. To reduce the computational burden, a Latin hypercube sampling approach was incorporated. Input random variables are modeled using kernel density estimation (KDE) in conjunction with Latin hypercube sampling (LHS) for probabilistic load flow (PLF) analysis. To test the proposed techniques, IEEE 14 and IEEE 118 bus systems are used. Two benchmark techniques, the Monte Carlo Simulation (MCS) method and Hamiltonian Monte Carlo (HMC), were set side by side for validation of results. The results illustrate that an adaptive bandwidth kernel density estimation with the Latin hypercube sampling (AKDE-LHS) method provides better performance in terms of precision and computational efficiency. The results also show that the suggested technique is more feasible in reducing errors, uncertainties, and computational time while depicting arbitrary distributions of photovoltaic and wind farms for probabilistic load flow analysis. It can be a potential solution to tackle challenges posed by sustainable energy sources in power systems.

Keywords: adaptive kernel density; probabilistic load flow; Latin hypercube sampling; HMC; Monte Carlo

check for updates

Citation: Bin, L.; Abbas, R.; Shahzad, M.; Safdar, N. Probabilistic Load Flow Analysis Using Nonparametric Distribution. *Sustainability* **2024**, *16*, 240. https://doi.org/10.3390/su16010240

Academic Editor: Miguel Carrión

Received: 30 October 2023
Revised: 14 December 2023
Accepted: 22 December 2023
Published: 27 December 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Power systems remained the contributor to environmental unsustainability, primarily due to the reliance on certain energy sources and associated environmental impacts. Efforts to transform power generation involved optimizing the efficiency of existing thermal power generation methods and increasing the adoption of sustainable and environment friendly energy sources like solar, hydroelectric, and wind energy. The uncertain behavior of this renewable energy source was a bigger problem. This problem led engineers to get creative and come up with more innovative solutions, sparking a real revolution in the power system analysis field. The field of power system analysis has undergone significant advancements over the years, and one of the key breakthroughs has been the introduction of probabilistic load flow (PLF) analysis. Various techniques [1] for handling uncertainty [2] have revolutionized the way electrical engineers assess the reliability, uncertainty, and performance of electrical grids. With the integration of renewable energy [3], the complexity of the power system increased, and deterministic load flow analysis could not handle the uncertainties associated with these sources.

The probabilistic load flow (PLF) analysis was introduced as a breakthrough in power system analysis, aiming to address the challenges of uncertainty and variability in electrical grids due to factors like load fluctuations [4] and equipment breakdowns. This concept emerged in the mid-twentieth century [5] and was initially focused on simple radial systems. The main objective of probabilistic load flow is to ascertain node voltages so as to compute the active and reactive power transfer in transmission lines while operating under specified load conditions. In response to the increasing complexity of power networks, new approaches [6] for performing probabilistic load flow analysis were developed. These approaches can be broadly categorized as parametric and nonparametric methods [7]. Parametric methods [8] use statistical parameters and assume specific distributions to represent uncertain variables, while on the other hand, data-driven approaches [9] make few assumptions about the underlying distribution of variables.

One popular method, the Monte Carlo Simulation (MCS) [10–12], involves creating random scenarios by sampling from probability distributions and conducting load flow studies. However, it has limitations, such as high processing requirements and convergence difficulties during simulations. Although the most accurate way was the MCS technique, the main issue was the computational burden. Later on, Markov Chain Monte Carlo [13–15] was also used due to its best estimation features. The main problem with Markov was the covariance matrix. If the covariance matrix is poor, errors will increase.

Analytical methods [16,17] seek to derive analytical equations for the probabilistic distribution of system variables using probability density functions, providing faster computation but limited applicability to simplified power system models. Time-Series [18] analysis relies on historical data to predict future system behavior, but it may overlook new trends and depend on data quality and availability. Sensitivity analysis investigates the effect of parameter changes on system output. It is useful for finding crucial parameters, but it does not provide a thorough probabilistic evaluation on its own. To improve understanding of power system behavior, sensitivity analysis [19] is frequently employed in conjunction with other methodologies.

Another strategy for reducing the number of scenarios required for probabilistic analysis while retaining statistical representativeness is Latin hypercube sampling (LHS). Latin hypercube sampling (LHS) [20–22] systematically samples input parameters to account for uncertainty; the quality of probabilistic data produces the best results with way less computational burden than MCS. One of the key concerns of parametric methods is the reliance on potentially incorrect assumptions about variable probability distributions, which can lead to imprecise conclusions, especially in non-standard configuration systems. Furthermore, these methods may struggle to deal with extreme events, which are increasingly due to climate change and rising energy demand fluctuations. While both parametric and nonparametric approaches have their merits and demerits, nonparametric methods, driven by data and computing technology, appear to be a promising path for the future of probabilistic load flow analysis.

Nonparametric methods, such as kernel density estimation [23] and adaptive bandwidth kernel density estimation (AKDE) [24,25], are gaining popularity due to their adaptability and ability to capture complex, real-world uncertainty patterns. They excel in handling large-scale [22,26], interconnected power systems with challenging probability distributions. Both parametric and nonparametric approaches are used to model input variables, but the proposed nonparametric estimation approach, like AKDE, shows cut-above results to examine the probability features of PV and wind energy. The probability density of the input random variables is estimated using kernel density estimation (KDE) with adaptive bandwidth. The data are sampled with Latin hypercube sampling (LHS) and compared with benchmark Monte Carlo Simulation (MCS). Hamiltonian Monte Carlo (HMC) is used for the validation of results. It accurately shows the characterized arbitrary probability distribution of embedded renewable energy. For the LHS method selected, sample and permutation procedures are applied. The modeling of PDF estimation is a key

factor in reducing error and uncertainty in the input random variables, directly improving accuracy in the output in the form of accurate estimation resulting in accurate outcomes.

The document is organized into sections. Section 2 is about the proposed concept and method for adaptive bandwidth KDE. Section 3 explains an LHS-based PLF method and a case study procedure. The results, performance analysis of the proposed method, and conclusions are discussed in Sections 5 and 6 respectively.

2. Kernel Density Estimation with Adaptive Bandwidth (AKDE)

AKDE is a statistical technique used to assess the PDF of the random variables (RVs). Unlike the traditional kernel density estimation, AKDE allows the bandwidth of the kernel to vary across different regions of the data, which makes it more suitable for datasets with varying densities.

AKDE involves estimating the density at a given point by averaging the values of the kernel function across the data points, with the kernel bandwidth adjusted to capture the local density of the data to accurately depict the PDF, reducing error in density estimation.

2.1. Kernel Density Estimation Basic Theory

Suppose there is an S dimension input variable $X = \{x_1, x_2, \ldots, x_s\}$, for each X_s (s = 1, 2,. . ., S) is a sample of random variables X_s following the PDF $f(x)$, x is the PDF variable, Kernel estimate $\hat{f}(x)$ of original $f(x)$ allocates an individual ith sample data point x_i a function $K(x_i)$, termed as the kernel function. According to KDE theory [27] (x) can be stated as

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{x_i - x}{h}\right) \quad (1)$$

where h is bandwidth, N is the number of samples, and $K(\cdot)$ is the kernel function. The Gaussian kernel is used for Equation (1), and it shows that the kernel density estimation emphasizes the calculation of the probability density based on the weighting of the kernel function. For arbitrary distributions, KDE is a better method because it only relies on samples. Equation (1) follows the property [28] that when $n \to +\infty$, $h \to 0$, and $nh \to +\infty$.

$$\text{Bias}\left(\hat{f}(x)\right) = \frac{h^2}{2} k_2 f''(x) \quad (2)$$

$$\text{var.}\left(\hat{f}(x)\right) = \frac{1}{nh} k f(x) \quad (3)$$

where Bias and variance of the kernel density estimation are shown by Bias $(\cdot)$ and var. $(\cdot)$, respectively. k_2 and k are functions of $k(\cdot)$ given as

$$k_2 = \int k(v) \cdot v^2 dv \quad (4)$$

$$k = \int k^2(v) dv \quad (5)$$

2.2. Adaptive Kernel Density Estimation

The choice of a kernel density function in statistical analysis has a minor effect on the accuracy of the estimate, while the choice of the bandwidth h is critical for accurate estimation. Instead of being a crucial factor in determining the results of the analysis, the choice of the kernel function is frequently a question of preference or convenience. Instead, the choice of data and bandwidth has a greater impact on the quality and accuracy of density estimates. Therefore, within the larger context of statistical analysis, researchers can often view the selection of kernel density function as a very minor matter. The finest bandwidth h can be obtained by decreasing the AMISE (Integrated Squared Error Mean) of KDE. The error between the actual density function $f(x)$ and its estimator $\hat{f}(x)$ can be

evaluated using the MSE (mean squared error), which estimates the difference between $\hat{f}(x)$ and $f(x)$ for each x.

$$\mathrm{MSEx}\left(\hat{f}\right) = E\left[\hat{f}(\mathrm{x}) - f(\mathrm{x})\right]^2 \tag{6}$$

The MSE for $\hat{f}(x)$ is calculated at every value of x using Equation (6), which is then transformed into Equation (7).

$$\mathrm{MSEx}\left(\hat{f}\right) = \frac{h^4}{4} {k_2}^2 f''(x)^2 + \frac{1}{nh} k f(x) \tag{7}$$

The trade-off between the Bias $(\cdot)$ and var.$(\cdot)$ of the estimator is crucial for minimizing the inaccuracy. The Bias and variance can be reduced by selecting an appropriate bandwidth h. AKDE method employs the estimated version of AMISE, calculated by expanding MISE using the Taylor series and considering only the most significant terms. The optimum bandwidth h_{op} can be determined by minimalizing the AMISE at every x.

$$\mathrm{AMISE} = \int \left[\mathrm{Bias}\ \hat{f}\ (x)^2\right] + \mathrm{var}\left[\hat{f}(\mathrm{x})\right] \mathrm{dx} \tag{8}$$

Previous work focused on minimizing Equation (6) for a fixed bandwidth h. In contrast, AKDE minimizes the AMISE at every x to reduce estimation error. The optimal bandwidth h_{op} can be obtained by solving $(d\mathrm{MSE}(x))/(dh(x)) = 0$, which occurs when MSE reaches its minimum value. h_{op} is achievable only when

$$h_{op} = h(x)\ n^{-1/5} \tag{9}$$

where $h(x) = [\frac{\kappa f(x)}{n {k_2} f''(x)^2}]^{1/5}$.

Consideration for each variable value achieves better accuracy with adaptive bandwidth approximation than the conventional KDE. The proposed method has two issues that need to be addressed. The first problem is that when the derivative of the mean squared error (MSE) concerning the bandwidth (h) is zero, it is known that it will either receive an extreme maximum value or a minimum value. But for optimum bandwidth, it must be minimal. Proof must be given to show that it will attain only the minimum value. The second problem is that Equation (7) appears to be infinite when the second derivative of the density function (f) is zero, resulting in a singular point. This contradicts the requirement in Equation (2) that the bandwidth approaches zero. Thus, solutions must be proposed to resolve this contradiction.

To address the first problem, it is shown that $d\mathrm{MSE}(\mathrm{x})/dh(x)$ for h along with fixed x is a monotonically growing function, and the minimal value of MSE is achieved at a particular bandwidth value. To address the second problem, a functional study-based proof is shown to explore the feasibility of solving the problem. A solution is then drawn out based on the modification of the cited points from the functional analysis approach. Assuming that $f''(x_0) = 0$, and given that $\hat{f}(\mathrm{x})$ is continuous and second-order derivable within a small domain covering x_0, defined as $[x_0 - \theta_1, x_0 + \theta_2]$, the interval of $\hat{f}(x)$ can be replaced with $g(x)$, where $g(x)$ follows the given optimization. The minimum of the objective function is

$$\int_{x_0-\theta_1}^{x_0+\theta_2} \left[\frac{h^4}{4} {k_2}^2 g''(x)^2 + \frac{1}{nh} k g(x)\right] dx$$

$$\left\{\begin{array}{c} f(\mathrm{x}_0 - \theta_1) = g(\mathrm{x}_0 - \theta_1), f(\mathrm{x}_0 - \theta_2) = g(x_0 - \theta_2) \\ f'(\mathrm{x}_0 - \theta_1) = g'(\mathrm{x}_0 - \theta_1), f'(\mathrm{x}_0 - \theta_2) = g'(x_0 - \theta_2) \\ g^{(2)}(\mathrm{x}_0) = 0 \end{array}\right\} \tag{10}$$

where θ_1 and θ_2 are improved boundaries of the estimation. Typically, higher-order polynomial functions are appropriate for $g(x)$ as there are five constraints in Equation (8). In the

computation, the singular points probability is very low, even zero. The efficiency of the system is not affected by the complexity of Equation (8).

3. Load Flow Based on LHS

This technique was first applied in "Uncertainty Analysis" by choosing input values for a function

$$Y = h(X) \tag{11}$$

where $X = (x_1, x_2 \ldots x_n)$, $n = 1, 2, \ldots$

These are random variables used to calculate the mean value of Y and the cumulative distribution function (CDF). The Latin hypercube sampling chooses n distinct values from each of the k variables, $x_1, x_2 \ldots, x_n$. On the basis of equal probability, the range of each variable is divided into n non-overlapping intervals. In accordance with the probability density inside the interval, one value is chosen at random from each interval. Thusly, generated n values for X_1 are randomly paired with n values for X_2, forming n triplets and so on. These n-k tuples are the same as the n k-dimensional input vectors of MCS.

Based on the proposed AKDE-LHS, PLF using nonparametric distributions could certainly be cracked using the following procedures.

1. Set the no. of variables to n-var and the no. of samples to n-sample.
2. Generate an n-sample by an n-var matrix of random numbers that are uniformly distributed between 0 and 1 using the rand function.
3. Randomly shuffle the rows of the matrix using the randomperm function to obtain a random permutation of the sample indices for each variable.
4. Calculate the probabilities of selecting each sample for each variable by dividing the shuffled indices by the total number of samples, subtracting a small random number to avoid repeated samples, and adding a small number to avoid zero probabilities.
5. Use the inverse of the standard normal CDF (norminv) to transform the probabilities into normally distributed values.
6. Scale the normal values to the desired mean and standard deviation for each variable.
7. Repeat steps 3–6 for each variable to obtain the Latin hypercube sample.
8. The resulting s matrix will be a Latin hypercube sample with n-sample samples and n-var variables, with each variable having a mean and standard deviation specified by X_{mean} and X_{sd}, respectively.

3.1. Case Study

Testing is performed on standard IEEE 14 and 118 bus systems to check the strength and accuracy of proposed AKDE-LHS techniques. For the validity of results, standard MCS, which is simulated 20,000 times, and HMC with a sample size of 1000 are used in order to provide the desired comparison of results.

PDF Estimation

Historical data of a Solar and Wind Power Plant was taken, in 2023, from the NREL (National Renewable Energy Laboratory), and these data were composed of 2006 data from western California and eastern California [29,30], with nearly 47,339 discrete non-zero dataset points and 52,559 for wind data for analysis. Three different techniques are tested to estimate the probability density of renewable energy sources, including lognormal, kernel density estimation with fixed bandwidth, and proposed AKDE with adaptive bandwidth. The proposed adaptive bandwidth is more precise in the probability density estimation (PDE) than in traditional KDE and lognormal probability density estimations, as precisely depicted in the PDF.

The measurement data-based probability density histogram is shown in Figures 1 and 2. Standard KDE used has a similar bandwidth for every sample point and yields significant errors, while there are large differences in probabilities in consecutive intervals. The proposed AKDE focuses on values that can accurately describe various features of distributions

as it can handle sharp changes in contiguous intervals. PDF histogram is distributed into ten intervals, as revealed in Figures 1 and 2, though errors for every interval display the accuracy of PDF estimates. From the historical data taken from NREL as analysis data for photovoltaic, the zeroth data have been excluded for PDF estimation to obtain expected results quickly. It can be seen from the graph that the data from (0–10) MW are excluded. The data excluded are not distorted but rather excluded for the sake of simplicity. Its inclusion has no impact on the results.

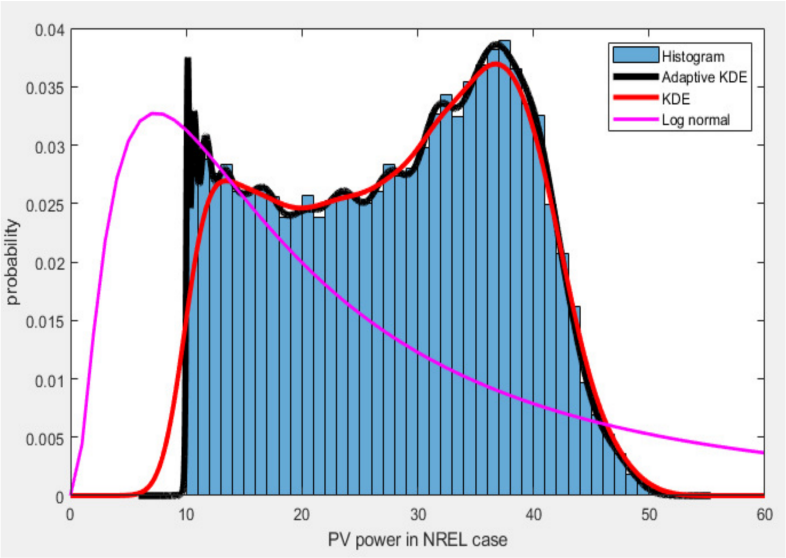


Figure 1. Photovoltaic probability estimation of NREL case.

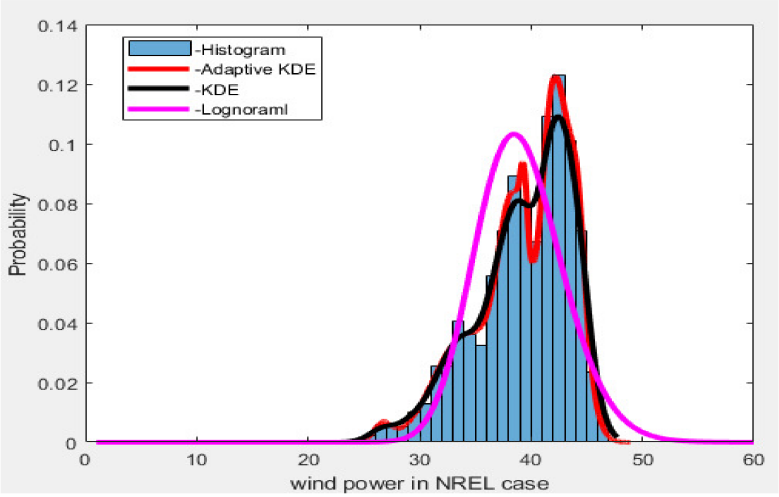


Figure 2. Wind probability estimation of NREL case.

It is essential to consider both the means and standard deviations as well as the associated errors when assessing the accuracy and reliability of these methods in statistical analysis. Table 1 provides a comparison of different statistical methods' performance in estimating various variables. The parametric method, like lognormal, has an error of 2.85021 while the error of KDE for fixed bandwidth is 2.05066 and 1.60323 for the proposed AKDE, as shown in Table 1. The differences observed between the methods can help researchers choose the most appropriate approach for their specific analysis needs, taking into account both precision and computational efficiency.

Table 1. RMS error for PDF estimation of PV and wind data.

PV Data	Lognormal	KDE	AKDE
ε_{μ}^{*} (%)	285.021	205.066	160.323
ε_{σ}^{*} (%)	99.2948	22.3299	25.1239
	Wind Data		
ε_{μ}^{*} (%)	28.6995	26.5862	8.5673
ε_{σ}^{*} (%)	320.6563	161.9865	77.0140

PDFs are revealed in Figures 1 and 2. It is found that the errors by lognormal and KDE are very large. However, errors with traditional KDE are reduced comparatively lognormal. However, the error calculated by AKDE decreased significantly. Hence, AKDE achieves a higher accuracy than conventional approaches while predicting probability distributions.

Tables 2 and 3 provide error analysis results for different approaches, with the proposed AKDE method accomplishing the highest accuracy in mean and standard deviation. The proposed method is effective for dealing with different renewable energy injection distributions and is consistent in PLF studies for arbitrary power injections, making it significant in efficiency and accuracy.

Table 2. Error analysis for 14 bus systems.

Method		MCS	AKDE-LHS	HMC	LHS Error	HMC Error
V	μ	0.010396	0.010393	0.010402	0.000288	0.000577
	σ	0.000154	0.000155	0.000146	0.006493	0.051948
	Max.	0.0106	0.0106	0.0106	0	0
δ	μ	0.103281	0.104723	−0.10092	0.013961	1.977140
	σ	0.040029	0.040512	0.040617	0.012066	0.014689
	Max.	0	0	0	0	0
P_{ij}	μ	0.393562	0.397711	0.215264	0.010542	0.453036
	σ	0.010954	0.010584	0.315917	0.033777	27.84033
	Max.	0.411965	0.412481	1.266761	0.001252	2.074923
Q_{ij}	μ	0.00156	0.001587	0.023242	0.017307	13.89871
	σ	0.001675	0.001701	0.052302	0.015522	30.22507
	Max.	0.005709	0.005723	0.162495	0.002452	27.46295

Table 3. Error analysis for 118 bus system.

Method		MCS	AKDE-LHS	HMC	LHS Error	HMC Error
V	μ	1.0405	1.0404	1.0404	9.61076×10^{-5}	9.61076×10^{-5}
	σ	0.0124	0.0123	0.0127	0.008064516	0.024193548
	max.	1.06	1.06	1.06	0	0
δ	μ	20.9045	18.7392	17.432	0.103580569	0.166112559
	σ	8.5881	11.5134	12.7897	0.340622489	0.489235104
	max.	37.6805	37.6155	38.3396	0.00172503	0.017491806
P_{ij}	μ	2.5895	3.8027	4.7794	0.468507434	0.845684495
	σ	70.5889	69.0376	67.6733	0.021976543	0.041303944
	max.	301.459	300.872	301.249	0.001945869	0.000695285
Q_{ij}	μ	0.3572	0.0684	0.568	0.808510638	0.590145577
	σ	16.5543	0.1203	15.7432	0.992733006	0.048996333
	max.	80.2127	0.5721	80.2102	0.992867713	3.11671×10^{-5}

4. PLF on IEEE-14 and IEEE-118 Bus System

The proposed methods are simulated by a PV station and a wind farm on modified 14 and 118 bus systems, respectively. These systems are considered standard for eval-

uating the system's behavior and can be modified according to one's needs for testing proposed techniques.

A comparison of different techniques is shown to prove the superiority of the proposed technique. The AKDE with LHS is known as AKDE-LHS. It is compared with the standard benchmark stochastic MCS and HMC.

4.1. IEEE-14 Bus System with Modification

On buses 4 and 9 of the IEEE 14 bus system, PV and wind farms are installed, respectively, as shown in Figure 3. LHS and HMC are simulated with an initial sample of 200, while the MCS, which is used as a reference, is simulated 20,000 times to obtain the desired results. The outcomes reveal the efficacy of the suggested method over the conventional method. The proposed technique produced almost the same results as the standard with 1000 iterations and less computational time as compared to HMC; MCS produces the same results with more iterations and a great computational burden on the system.

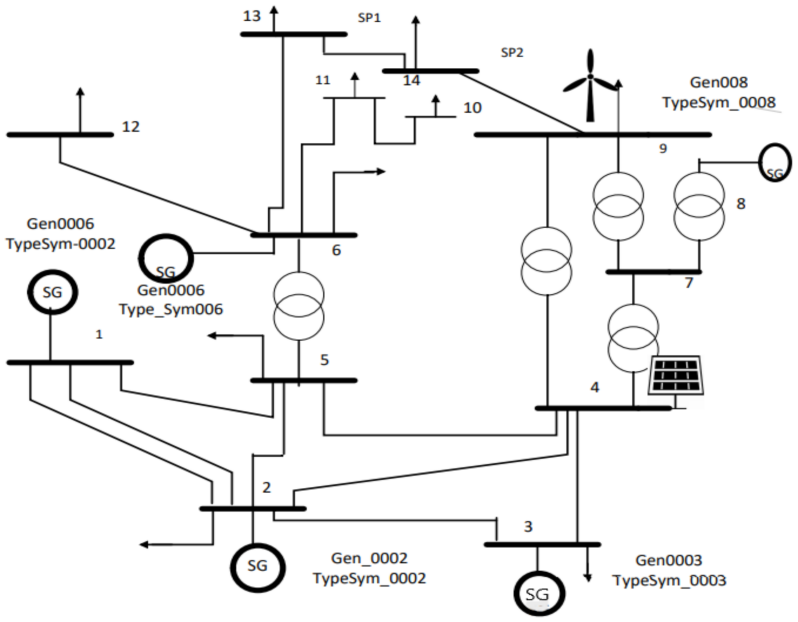


Figure 3. IEEE-14 bus system diagram.

The error indices of output variables, using Equations (9) and (10), are calculated to show the performance of other methods. Results show that AKDE-LHS achieves better accuracy in mean and standard deviation as compared to conventional MCS and HMC sampling methods.

Table 2 includes information on the errors associated with the LHS and HMC methods for each variable, with both mean and standard deviation values. The error indices show that the error by kernel density estimation method with adaptive bandwidth is around 0.00028 in the case of voltage, which is acceptable from an engineering perspective, while the HMC has an error of 0.00057, which is more than AKDE-LHS. Meanwhile, the other output variable errors are also shown in Table 2. For all variables, it is evident that all three methods yield very close estimates of the mean, with differences in the fourth decimal place. However, when it comes to the standard deviation, HMC shows a higher value than the other two methods, resulting in an error that is several times larger than that of MCS and AKDE-LHS, showing that the proposed AKDE-LHS gives almost the same result as MCS with fewer iterations that MCS generates with a 20,000 sample size, hence reducing the time required to predict the future behavior of the power system. We use a larger sample size for Monte Carlo Simulation (MCS) compared to Latin hypercube sampling (LHS), but the results are nearly the same. This is because MCS provides more accurate

answers with a large sample size, while LHS converges with accuracy more quickly, giving it an advantage over MCS in terms of efficiency.

The active power of branch 4–8/MW is shown in Figure 4. It shows that active power results are better in the case of AKDE-LHS, but in the case of HMC, the error is considerably high with the same sample size. By choosing suitable kernel functions, cross-validating the sample design to fully cover the variable space, and figuring out the ideal bandwidth, the AKDE-LHS approach can reduce mistakes in active power flow analysis. Critical parameters are identified via sensitivity analysis, and dependability is improved by model validation and Monte Carlo Simulations. With these study-specific stages, active power flow analysis utilizing AKDE-LHS becomes more accurate, leading to more reliable results in power system studies. Bus 4 voltages are shown in Figure 5, derived from a 14-bus system. It is clear that both LHS and HMC generate the same result with minor differences. The computational time for the 14-bus system gives better insight into the proposed method's computational efficiency, as shown in Figure 6.

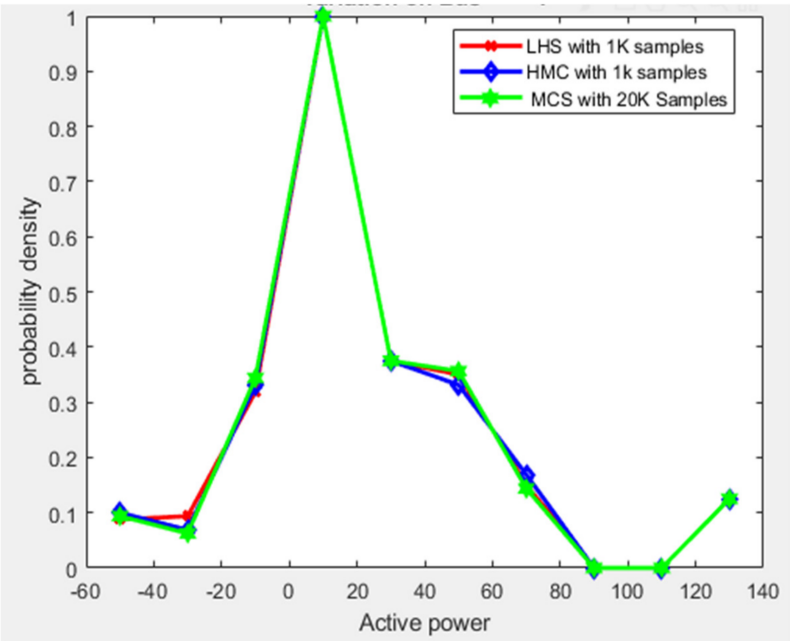


Figure 4. Active power flow of branch 4–8/MW.

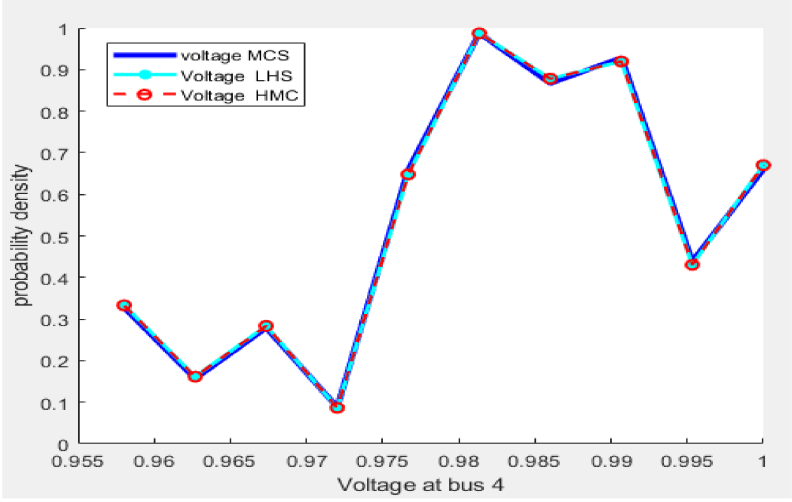


Figure 5. Voltage at bus 4.

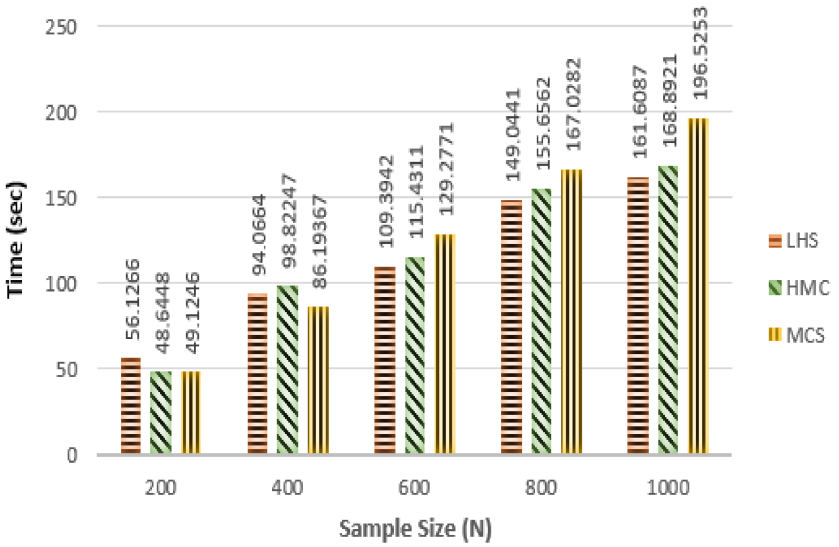


Figure 6. Computational time analysis for 14-bus system.

4.2. IEEE-118 Bus System with Modification

The proposed AKDE attains higher accuracy than conventional methods. The input varies from 24 to 34 of the IEEE 118 bus system using PV and wind farms on each bus, respectively. Figure 7 shows that the proposed technique is efficiently able to handle the variations with fewer iterations than MC.

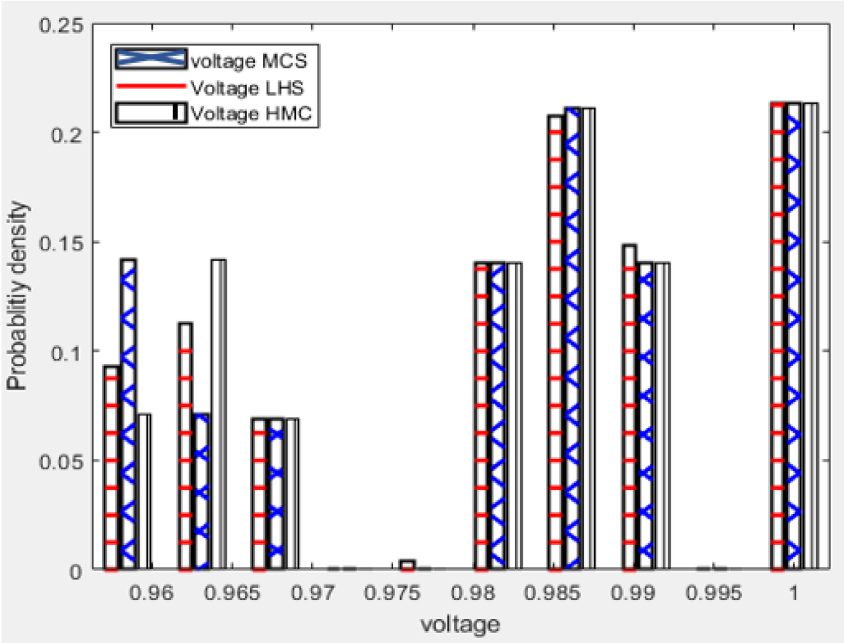


Figure 7. Voltage at bus 24.

Figure 7 reveals the probability histogram of voltage at bus 24 while Figure 8 shows the active power flow of branch P24–26. The computational time analysis of AKDE-LHS, HMC, and MCS is given in Figure 9.

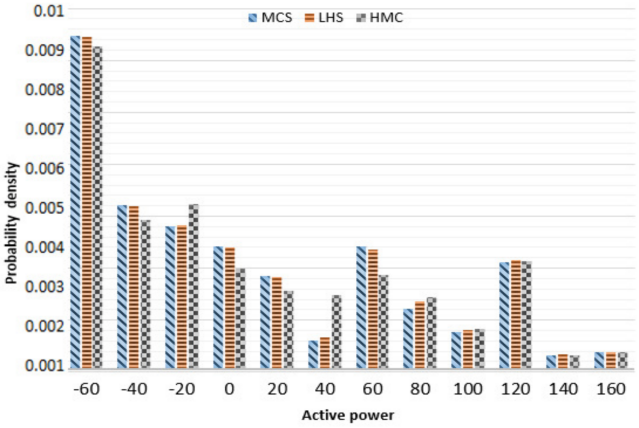


Figure 8. Output active power flow of branch 24–26/MW.

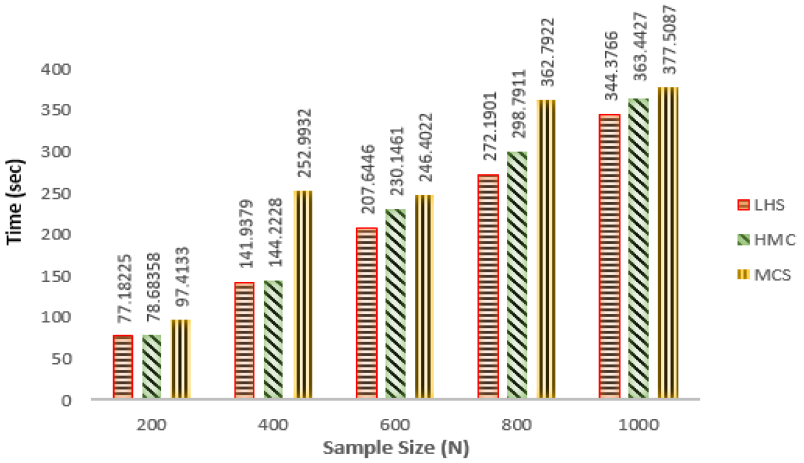


Figure 9. Computational time analysis for 118-bus system.

Error indices for the 118-bus systems are given in Table 3, which shows that the proposed technique achieves the desired results better than HMC, except in reactive power. In the other parameters, it achieved almost the same result as standard MCS, while HMC showed some errors.

5. Performance Analysis of AKDE-LHS, HMC, and MCS

To test the performance of the AKDE-LHS, HMC, and MCS methods, simulations were implemented on modified IEEE 14 and IEEE 118 test systems. The deterministic data for these systems are taken from MATPOWER test cases, and loads are modeled as Gaussian distributions. The simulations were performed using MATLAB installed on a Windows 10 machine with a 1.9-GHz Intel(R) Core(TM) i3-4030U CPU @ 1.90 GHz processor and 4 GB of RAM with a 64-bit operating system. The MCS method, with a sample size of 47,339 for PV data and 52,559 for wind data, was used as a reference for assessing the exactness of output RVs using the error guide as described in [31].

$$\varepsilon_{\mu}^{*} = \left| \frac{\mu_a - \mu_s}{\mu_a} \right| \times 100\% \tag{12}$$

$$\varepsilon_{\sigma}^{*} = \left| \frac{\sigma_a - \sigma_s}{\sigma_a} \right| \times 100\% \tag{13}$$

where * symbolizes the output RV of PLF. The output RVs are voltage V, voltage angle θ, branch active power P, and branch reactive power Q. The μ_a is the accurate mean, and σ_a

is the accurate standard deviation, which is found by MCS. μ_s symbolizes the simulated mean and σ_s is the simulated standard deviation by the proposed AKDE-LHS.

Figures 7 and 8 demonstrate that the HMC's performance is less precise than that of the AKDE-LHS, even with the same sample size. Additionally, Figures 6 and 9 reveal that the HMC's computational accuracy is not better than that of the AKDE-LHS; despite this, the AKDE-LHS does have a slightly longer computational time for certain sample sizes than HMC even with the identical number of samples, as shown in computational time analysis of the modified test bus systems with insignificant differences. Therefore, AKDE-LHS are street-ahead techniques to model input variables and offer a better computational burden.

6. Conclusions

In conclusion, the proposed techniques are evaluated on the real time data taken from NREL. The primary aim of the research is to assess the performance of the proposed technique in comparison to the conventional system when integrating a sustainable energy source into the electrical power system. To achieve this, the input variables were modeled using AKDE. PV and wind data were taken for analysis point of view. PDFs of both renewable energy sources were modeled independently and treated as input variables. The results reveal that the AKDE has better accuracy in modeling. The LHS sampling has reduced considerable computation burden. AKDE, in connection with LHS, is known as AKDE-LHS. This proposed approach was applied using MATLAB environment tools. The results show that the proposed technique is efficient enough in probability estimation to decrease the RMS error and computational burden of the system. Further analysis showed that AKDE-LHS acquired the desired results in almost 1000 samples, HMC in 1000 samples with some error, while standard MCS was obtained in 20,000 samples. The AKD-LHS method is highly accurate, outperforming the MCS and HMC in terms of accuracy and efficiency with a lesser simulation size. Moreover, the no. of simulations required for AKDE-LHS is free from network variables, indicating its potential to resolve PLF issues in power systems by renewable energy sources. Multivariable will be modeled in future work for the efficacy of the proposed technique. These techniques can be employed for the evaluation of power system reliability, economic analysis, and power system planning studies.

Author Contributions: Conceptualization, L.B.; software, R.A.; methodology, M.S.; writing and original draft preparation, N.S.; validation and proofreading, R.A.; project administration, M.S.; funding acquisition, L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Science and Technology projects from State Grid Corporation, P. R. China (Grant number 5108-202218280A-2-379-XG).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Singh, V.; Moger, T.; Jena, D. Uncertainty handling techniques in power systems: A critical review. *Electr. Power Syst. Res.* **2022**, *203*, 107633. [CrossRef]
2. Ehsan, A.; Yang, Q. State-of-the-art techniques for modelling of uncertainties in active distribution network planning: A review. *Appl. Energy* **2019**, *239*, 1509–1523. [CrossRef]
3. Ramadhani, U.H.; Shepero, M.; Munkhammar, J.; Widén, J.; Etherden, N. Review of probabilistic load flow approaches for power distribution systems with photovoltaic generation and electric vehicle charging. *Int. J. Electr. Power Energy Syst.* **2020**, *120*, 106003. [CrossRef]
4. Borkowska, B. Probabilistic load flow. *IEEE Trans. Power Appar. Syst.* **1974**, *3*, 752–759. [CrossRef]
5. Bin, L.; Shahzad, M.; Bing, Q.; Ahsan, M.; Shoukat, M.U.; Khan, H.; Fahal, N.A. The probabilistic load flow analysis by considering uncertainty with correlated loads and photovoltaic generation using Copula theory. *AIMS Energy* **2018**, *6*, 414–435. [CrossRef]
6. Wang, Y.; Zhang, N.; Kang, C.; Miao, M.; Shi, R.; Xia, Q. An efficient approach to power system uncertainty analysis with high-dimensional dependencies. *IEEE Trans. Power Syst.* **2018**, *33*, 2984–2994. [CrossRef]

7. Abbasi, A.R. Comparison parametric and non-parametric methods in probabilistic load flow studies for power distribution networks. *Electr. Eng.* **2022**, *104*, 3943–3954. [CrossRef]
8. Guo, Z.; Wei, W.; Chen, L.; Dong, Z.; Mei, S. Parametric distribution optimal power flow with variable renewable generation. *IEEE Trans. Power Syst.* **2021**, *37*, 1831–1841. [CrossRef]
9. Liu, J.; Srikantha, P. Kernel Structure Design for Data-Driven Probabilistic Load Flow Studies. *IEEE Trans. Smart Grid* **2022**, *13*, 2679–2689. [CrossRef]
10. Billinton, R.; Li, W. A Monte Carlo method for multi-area generation system reliability assessment. *IEEE Trans. Power Syst.* **1992**, *7*, 1487–1492. [CrossRef]
11. Palahalli, H.; Maffezzoni, P.; Gruosso, G. Modeling photovoltaic generation uncertainties for monte carlo method based probabilistic load flow analysis of distribution network. In Proceedings of the 2020 55th International Universities Power Engineering Conference (UPEC), Turin, Italy, 1–4 September 2020; pp. 1–6.
12. Constante-Flores, G.E.; Illindala, M.S. Data-driven probabilistic power flow analysis for a distribution system with renewable energy sources using Monte Carlo simulation. *IEEE Trans. Ind. Appl.* **2018**, *55*, 174–181. [CrossRef]
13. Speagle, J.S. A conceptual introduction to Markov chain Monte Carlo methods. *arXiv* **2019**, arXiv:1909.12313.
14. Frisoni, P.; Gozzini, F.; Vidotto, F. Markov chain Monte Carlo methods for graph refinement in spinfoam cosmology. *Class. Quantum Gravity* **2023**, *40*, 105001. [CrossRef]
15. Jones, G.L.; Qin, Q. Markov chain Monte Carlo in practice. *Annu. Rev. Stat. Its Appl.* **2022**, *9*, 557–578. [CrossRef]
16. Li, C.; Gong, Z.; Wang, Y.; Wu, W.; Ni, S. Analytical Stochastic Power Flow for Power Systems with High Penetration of Renewable Energy. In Proceedings of the 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2), Taiyuan, China, 22–24 October 2021; pp. 2484–2489.
17. Wang, C.; Liu, C.; Tang, F.; Liu, D.; Zhou, Y. A scenario-based analytical method for probabilistic load flow analysis. *Electr. Power Syst. Res.* **2020**, *181*, 106193. [CrossRef]
18. Tao, Y.; Xu, M.; Guo, Z.; Zhang, J.; Cui, Z.; Tan, Y.; Lei, J. Time series probabilistic power flow calculation of distribution network based on nonparametric kernel density estimation. In Proceedings of the 2022 International Conference on Computers, Information Processing and Advanced Education (CIPAE), Ottawa, ON, Canada, 26–28 August 2022; pp. 16–20.
19. Urquizo, J.; Calderón, C.; James, P. Using a Local Framework Combining Principal Component Regression and Monte Carlo Simulation for Uncertainty and Sensitivity Analysis of a Domestic Energy Model in Sub-City Areas. *Energies* **2017**, *10*, 1986. [CrossRef]
20. Huang, Y.; Chen, S.; Chen, Z.; Hu, W.; Huang, Q. Improved probabilistic load flow method based on D-vine copulas and Latin hypercube sampling in distribution network with multiple wind generators. *IET Gener. Transm. Distrib.* **2020**, *14*, 893–899. [CrossRef]
21. Aibin, L.; Wenyi, L. Reliability evaluation of distribution network with distributed generation based on latin hypercube sequential sampling. In Proceedings of the 2020 3rd International Conference on Electron Device and Mechanical Engineering (ICEDME), Suzhou, China, 1–3 May 2020; pp. 97–99.
22. Deng, X.; Zhang, P.; Jin, K.; He, J.; Wang, X.; Wang, Y. Probabilistic load flow method considering large-scale wind power integration. *J. Mod. Power Syst. Clean Energy* **2019**, *7*, 813–825. [CrossRef]
23. Zhang, W.; He, Y.; Yang, S. Day-ahead load probability density forecasting using monotone composite quantile regression neural network and kernel density estimation. *Electr. Power Syst. Res.* **2021**, *201*, 107551. [CrossRef]
24. He, Y.; Cao, C.; Wang, S.; Fu, H. Nonparametric probabilistic load forecasting based on quantile combination in electrical power systems. *Appl. Energy* **2022**, *322*, 119507. [CrossRef]
25. Nosratabadi, H.; Mohammadi, M.; Kargarian, A. Nonparametric probabilistic unbalanced power flow with adaptive kernel density estimator. *IEEE Trans. Smart Grid* **2018**, *10*, 3292–3300. [CrossRef]
26. Rajan, R.; Fernandez, F.M.; Yang, Y. Primary frequency control techniques for large-scale PV-integrated power systems: A review. *Renew. Sustain. Energy Rev.* **2021**, *144*, 110998. [CrossRef]
27. Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Stat.* **1962**, *33*, 1065–1076. [CrossRef]
28. Ye, A. *Nonparametric Econometrics*; Nankai University Press: Tianjin, China, 2003.
29. Obtain Solar Power Data: Western State California. Available online: https://www.nrel.gov/grid/solar-power-data.html (accessed on 30 June 2023).
30. Obtaining the Eastern Wind Dataset. Available online: http://www.nrel.gov/electricity/transmission/eastern_wind_dataset.html (accessed on 30 June 2023).
31. Yu, H.; Chung, C.; Wong, K.; Lee, H.; Zhang, J. Probabilistic load flow evaluation with hybrid latin hypercube sampling and cholesky decomposition. *IEEE Trans. Power Syst.* **2009**, *24*, 661–667. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.