A Hybrid Pelican Optimization Algorithm and Black Hole Algorithm for Kernel Semi-Parametric Fusion Modeling

Firas A. Yonis AL-Taie*1, Zakariya Yahya Algamal2, Omar Saber Qasim3

[1] Department of Mathematics, University of Mosul, Mosul, Iraq
[2] Department of Statistics and Informatics, University of Mosul, Mosul, Iraq
[3] Department of Mathematics, University of Mosul, Mosul, Iraq
Emails: firas_a.al_taie@uomosul.edu.iq; zakariya.algamal@uomosul.edu.iq; omar.saber@uomosul.edu.iq

Abstract

This paper investigates the process of selecting a hyperparameter for use in a kernel semiparametric regression model for fusion data, which is an important tool in various scientific study fields. The selection of the best model to use in advance is not a simple task, and one of the most fascinating current advances in the application is the use of hybrid metaheuristics algorithms to increase the exploration and exploitation capacity of traditional meta-heuristic algorithms. In this study, a hybrid optimization method that combines the pelican algorithm with the black hole algorithm is presented, which achieves a lower mean squared error (MSE) in comparison to other competing techniques. Data merging through the suggested hybrid metaheuristics algorithm gives superior performance in terms of computing time when compared to both the CV-method and the GCV-method. This work has practical implications for researchers and practitioners who use statistical modeling techniques in their work, especially those dealing with data merging for improved accuracy and efficiency.

Keywords: Black hole algorithm (BHA); Pelican optimization algorithm (POA); semiparametric model; kernel methods; cross-validation, data fusion.

1. Introduction

Kernel Semi-Parametric Fusion Modeling is a strategy for statistical modeling that combines parametric and non-parametric techniques for studying complicated data. To do this, a non-parametric methodology, such a kernel smoothing technique, is applied to the data before any other analysis is performed. After the data has been processed, a model with parameters is used to make estimates for the variables of interest. Unlike when utilizing either non-parametric or parametric approaches alone, the goal of kernel semi-parametric fusing models is to produce a better fit to the data. For data sets with a complicated structure that cannot be described by a single parametric model, or when the underlying distribution is unknown, this method becomes very effective. Fusion, in the context of kernel semi-parametric fusion modeling, means the integration of many modeling techniques into a single framework for a more thorough examination of the data. Finance, economics, engineering, and biology are just few of the areas that have employed this method to examine large data sets and anticipate future trends or events [1,2].

Semi-parametric regression models are of great importance in statistics as well as in econometrics, because of their ability to model events [1, 2]. Note the semi-standard regression model given:

$$y_j = x_j\alpha + f(t_j) + \varepsilon_j j = 1,2,\dots,N \qquad (1)$$

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

Where the explanatory variables are denoted by a vector $x_j = (x_{j1}, x_{j2}, \ldots, x_{jp})$, $\alpha = (\alpha_{1,}\alpha_{2,}\ldots,\alpha_p)'$ is a parameters vector that are unknown in p-dimensional, t_j are known non-random in $D \subset \mathbb{R}$, $f(t_j)$ is a smooth function which are unknown and ε_j 's are independent random error and identically distributed with mean 0 and variance σ^2. and ε_j 's independent of (x_j, t_j) [3].

A lot of semiparametric regression model techniques are based on procedures for various nonparametric regression [4]. There were various methods for appreciation α and $f(.)$. The difference in the research method or the solution method for the non-parametric procedure can be considered as the alternative method [5]. Where these differences exist, they are used to remove trends that originate from a mentioned function, and the estimator is not required for the function $f(.)$ which in some cases is called difference-based procedure [6]. Provided that $f(.)$ It can be differentiable and the t-coordinates are divergent, the effect of the function is likely to be omitted $f(.)$ by properly between the data. In model (Eq.(1)), Liu, Ning [7], , Roozbeh, Arashi [2] it relied on an estimation of the linear component and the estimation based on the differences used is optimal meaning that the value of the linear component is effective and well approximated and also the estimator of the non-parametric component is the minimum and the optimal approach to the semi-parameter where the model used the higher differences to reach the optimal capacity in setting the part "linear part" and we use the quality of different sequences [8].

This manuscript aims to address the problem of optimizing the hyper-parameters of the semi-parametric regression model by proposing hybrid algorithms and comparing them to other methods to show the efficiency of the proposed improvement. The main contribution of this study is to provide the best estimate for the semi-parametric models after the hybrid algorithm provides a perfect evaluation of the hyperparametric and thus the best estimate of the non-parametric part in the semi-parametric kernel model.

However, there are some limitations to this study. The selection of hyperparameters is focused on a specific type of model, kernel semi-parametric regression, and may not be applicable to other types of models or regression analysis. Additionally, the datasets used in the study may not be representative of other real-world datasets, and the results may not generalize to other datasets.

The essay is formatted as follows: Section 1 provides an introduction to the study. In Section 2, we explain the kernel semi-parametric regression model. Section 3 discusses the implications of the findings. The technique that has been suggested is demonstrated in the following section (Section 4). The fourth section of this manuscript consists of performing a simulation study and presenting experimental results. Finally, a concluding remark can be found in section 5.

2. Kernel semiparametric regression model

The kernel estimate method is one of the significant statistical tools, and it has a wide variety of applications in the real world [9]. It has been effectively utilized, both to univariate and multivariate situations, with positive results. Kernel estimation is a method of data smoothing that involves estimating a density function with the use of independent sample data that is dispersed in a same fashion. Exploratory data analysis and visualization are two applications that make use of the data mining approach known as kernel estimation [10].

Both parametric and nonparametric kernel estimators are wide categories that may be used to organize kernel estimators. Estimators that use parametric kernels begin with the assumption of a distribution's functional form, which reduces the complexity of the task to just estimating the distribution's parameters. It is not necessary to use a functional form in order to estimate the density using a nonparametric kernel estimator. Estimators that are nonparametric provide for greater flexibility than their parametric counterparts, but they typically require more work from a computational standpoint [11, 12].

Let $Y = \left(Y_1,\ldots,Y_N\right)^T$ is a vector of (N × 1) continuous outcomes , with Y_j being the value for subject $j = 1,\ldots,N$, and X and Z are (N × P) and (N × Q) matrices of predictors , respectively, such that Y depends on Z and X a as shown below. This is the kernel semiparametric model:

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

$$Y = X\alpha + \mathrm{H}(Z) + e \quad (2)$$

where α is a vector of regression coefficients for the linear parametric portion with a size of (P × 1), H(.) is a function with unknown smoothness, and e is a vector of independent and homoscedastic errors with a size of (N × 1) , such as $e_i \sim N\left(0,\sigma^2\right)$. The linear component of the model is denoted by the $X\ \alpha$, and it is presumed that element X includes an intercept. The "kernel component" is H(Z). The function H(.) may be found in the function space denoted by the notation G_k, , which was formed by the kernel function k(.,.). K is the (N × N) Gram matrix of k(.,.) , and it is constructed in such a way that j indicates the degree to which subjects i and Z are alike $K_{ji} = k\left(Z_j, Z_i\right)$. The procedure for estimating H () involves expressing it as a linear combination of k(.,.) in such a way that:

$$H = K\beta \quad (3)$$

The primary objective of the kernel techniques is to employ a nonlinear mapping in order to transform the data set $R=[Y_1,Y_2,...,Y_n]$ into a feature space that possesses a higher dimension. According to Cover's theorem, complex nonlinear problems in the initial low-dimensional space can be more readily addressed in the modified space since they are more easily linearly separable in the modified space [13].

3. The proposed hybridization

The optimization solution for the objective function in Eq. (3) is affected in many ways by each of the hyperparameters of the kernel type. There is no mathematically sound way for identifying the precise values that need to be utilized, and the performance of the kernel function is substantially impacted by the choice of these hyperparameters. The study on kernel semiparametric regression model thus places a significant amount of weight on the selection of such hyperparameters. The CV-method and GCV-method approach is currently the method that has the greatest amount of support in the research community for determining hyperparameters. On the other hand, it has been pointed out that it might take a very long time [14].

The natural world has served as a source of motivation for the development of several meta-heuristic algorithms. In the realm of scientific study, swarm intelligence is indeed a crucial instrument for the solution of many difficult issues[15]. There has been a lot of research done on swarm intelligence algorithms, and they have been effectively applied to a range of difficult optimization issues [16]. Pelican optimization algorithm (POA) proposed by Trojovsky and Dehghani [17] is based on the manner in which pelicans go about their hunts and the tactics they use, which is a thoughtful adaptation that has enabled these birds to become proficient hunters.

The POA is an algorithm that is built on a population, and in this algorithm, pelicans are considered to be members of this population. In algorithms that are based on populations, each member of the population represents a possible solution. Every individual in the population has a location inside the search space, and this determines the values that they suggest for the variables that make up the optimization issue. In the beginning, population members are chosen at random and then initialized in accordance with the problem's lower bound and upper bound by employing eq. (5)

$$x_{\mathrm{j,i}} = I_i + rand * (u_i - I_i),\ \mathrm{j} = 1, 2, ..., \mathrm{n},\ \mathrm{i} = 1, 2, ... \mathrm{m} \quad (4)$$

where $x_{j,i}$ is the ith variable's value as stated by the jth candidate solution, n is the population's size, m is the number of variables involved, $rand$ is a random and real number between [0,1], Ii is the ith lower bound, and ui is the ith upper bound of problem variables.

In Eq. 5, a matrix referred as the population matrix is used to identify the individuals of the population of pelicans that live in the POA. The suggested solutions to the problem variables are listed in the columns of this matrix, while the rows of this matrix each indicate a potential answer to the problem.

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

$$X = \begin{bmatrix} X_1 \\ . \\ . \\ X_2 \\ . \\ . \\ X_n \end{bmatrix}_{n\times M} = \begin{bmatrix} x_{1,1}\dots x_{1,i}\dots x_{1,M} \\ \dots \\ \dots \\ x_{j,1}\dots x_{j,i}\dots x_{i,M} \\ \dots \\ \dots \\ x_{n,1}..x_{n,i}..x_{n,M} \end{bmatrix}_{n\times M} \quad (5)$$

Where X_j is the jth pelican and X is the pelican population matrix.

One of the potential answers to the problem at hand is for the whole population of POA to assume the shape of pelicans; this is the short way to solve the problem. As a consequence of this, the potential solutions may individually be used to conduct an analysis of the objective function associated with the given problem. In the equation, there is a vector that is referred to as the objective function vector[18]. This vector is used to calculate the values that are produced for the objective function as seen in Eq. (6)

$$U = \begin{bmatrix} U_1 \\ . \\ . \\ U_2 \\ . \\ . \\ U_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} \mathrm{U}(X_1) \\ . \\ . \\ \mathrm{U}(X_2) \\ . \\ . \\ \mathrm{U}(X_3) \end{bmatrix}_{n\times 1} \quad (6)$$

Where U is the vector of the Objective function, and U_j is the value of the objective function for the jth candidate solution. In order to improve potential solutions, the POA runs simulations that model the actions and strategies that pelicans employ as they attack and pursue their prey. The simulation of this hunting method takes place in two stages:

(i) Getting closer to the prey (exploration phase).

(ii) Floating on the surface of the sea with wings (exploitation phase).

3.1 Phase 1: Getting closer to the prey (exploration phase)

The initial step of the process involves the pelicans determining the position of the prey and then moving toward the region that they have determined. The modeling of this pelican's technique leads to the scanning of the search space, which demonstrates the exploration capacity of the suggested point of attack in locating various regions of the search space. The fact that the position of the prey is chosen at random inside the search space is one of the most significant aspects of POA. This increases POA's ability for exploration in the exact search for the space that may hold viable answers to the issue. Eq. (7) provides a mathematical representation of the aforementioned ideas as well as the approach used by pelicans as they go to their hunting grounds

$$x_{j,i}^{p1} = \begin{cases} x_{j,i} + rand.(\mathrm{p}_i - r.\mathrm{x}_{j,i}), \mathrm{U}_p < U_j; \\ x_{j,i} + rand.(x_{j,i} - p_i), else, \end{cases} \quad (7)$$

r is a random integer that can be either [one or two], p_i is the position of prey in the ith dimension, and U_p is the value of its Objective function. $x_{j,i}^{p_1}$ represents the new state of the jth

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

pelican in the ith dimension based on phase 1. The value of the parameter r, which can either be 1 or 2, will be chosen at random. This parameter is chosen at random for each iteration as well as for each member of the group. It delivers additional displacement for a member whenever the value of the parameter is exactly two, which might lead that member to discover newer locations of the search space. As a result, the POA exploration power is affected by parameter r, which makes it more difficult to correctly scan the search space. If the value of the goal function is better in a new position for a pelican, then that new position will be considered acceptable. When doing this kind of updating, which is known as efficient updating, the algorithm is stopped from going to places that are not optimal. The equation used to represent this process is shown below in Eq (8).

$$X_j = \begin{cases} X_j^{p_1}, \mathrm{U}_j^{p_1} < U_j \\ X_j, else, \end{cases} \quad (8)$$

where the new status of jth pelican is denoted by $X_j^{p_1}$ as well as $U_j^{p_1}$ represents the objective function value that based on phase 1.

3.2 Phase 2: Floating on the surface of the sea with wings (exploitation phase)

Following the first step, which consists of swimming to the surface of the water, for the next part of the process, the pelicans use the water's surface to lift the fish using their wings, after which they gather the prey in the pouch in their throat. Pelicans are able to capture a greater number of fish in the region that is being assaulted as a result of this tactic. The POA is made to approach to better places in the hunt area as a result of modeling this movement of pelicans. This procedure enhances both the power of the local search and the capability of exploitation offered by POA. In order to arrive at a more optimal answer from a mathematical perspective, the algorithm will need to conduct an investigation into the spots that are located close to where the pelican is located. This hunting strategy utilized by pelicans is modeled mathematically in the form of an Eq (9).

$$x_{j,i}^{p_2} = x_{j,i} + RR.(1 - \frac{tt}{TT}).(2 * rand - 1) x_{J,I}, \quad (9)$$

where $x_j^{p_2}$ is the new status of the jth "pelican" in the ith dimension based on second phase, RR is a constant, and it's equal to 0.2, $RR \cdot (1 - \frac{t}{TT})$ is the nearness radius of $x_{j,i}$ while, tt is the iteration number, and TT is the Max-number of iterations. The value "$RR \cdot (1 - \frac{tt}{TT})$" denotes of the nearness which the members of the population are located in order to search locally near every member in an effort to converge on a more optimal solution. This parameter has an influence on the power of exploitation of the POA, which brings us that much closer to the ideal global solution. The value of this coefficient is set to be high at the beginning of the iterations, and as a consequence, a bigger region surrounding each member is taken into consideration. The "$RR \cdot (1 - \frac{tt}{TT})$" coefficient will drop as the number of replicated copies of the method rises, which will lead to lower radii for the neighborhoods of every member. Because of this, we are able to scan the region surrounding each individual in the population with steps that are both smaller and more precise, which enables the POA to converge on solutions that are closer to the global ideal (or even exactly the global optimal) based on the utilization concept.

At this point, effective updating has also been utilized to either accept or reject the newly proposed pelican posture, which is depicted in Eq. (10)

$$X_j = \begin{cases} X_j^{p_2}, \mathrm{U}_j^{p_2} < U_j; \\ X_j, else, \end{cases} \quad (10)$$

where $x_j^{p_2}$ is the new status of the jth element "pelican" and $U_j^{p_2}$ is its objective function value based on second phase.

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

One of the most recent examples of evolutionary algorithms is called the black hole algorithm (BHA). This algorithm was inspired by the phenomena of black holes in space. This method was initially presented by Hatamlou in the year 2013 [19]. When a star of enormous size dies, a space-based black hole is created as a remnant of the event. Even light is unable to get away from the gravitational pull of the black hole because its gravitational force is so strong [20].

The beginning of the procedure in BHA involves the establishment of the stars that are present in the search space. These stars represent the population. The initialization is produced by a random process. The particular star with the best available function is chosen to become the black hole, and the black hole immediately begins devouring the populations of stars that surround it. After then, each star moves in the direction of the black hole. The following is one possible formulation for this movement:

$$x_j^{t+1} = x_j^t + \theta \times (x_{BH} - x_j^t), \quad j = 1, 2, \dots \mathrm{N}_s, \tag{11}$$

where x_j^t and x_j^{t+1} represent the positions of the jth star during iterations t and $t+1$, respectively. The coordinates of the black hole in the search space are denoted by x_{BH}, a random integer inside the interval is denoted by θ, and the total number of stars is denoted by N_s (candidate solutions in the search space). It is essential to emphasize the fact that the black hole does not move since it possesses the highest objective value and, as a result, draws all of the other stars [19, 21].

After finding the movement of the stars using equation (12), if indeed the objective function value of a star is superior the value of the black hole, then the star will choose as the black hole. This happens after the movement of the stars has been determined. When stars are travelling in the direction of the black hole, there is a probability that they will pass over the singularity (border of the black hole).

In the procedure for the black hole, the radius of the event horizon, also known as the Schwarzschild radius, is determined by

$$R = \frac{f_{BH}}{\sum_{j=1}^{N_s} f_j}, \tag{12}$$

where the value of the objective function f_{BH} corresponds to the black hole, and the value of f_j corresponds to the jth. When the distance between such a candidate solution and the black hole is smaller than the threshold set by Eq. (12), the candidate in question is eliminated, and a new candidate is generated before being randomly dispersed over the search space.

The term "hybridization" refers to the process of combining many distinct metaheuristics algorithms into a single structure. As a result, a hybrid algorithm may make use of the benefits offered by the various algorithms, which would then result in performance that is more favorable than that offered by a single algorithm [22]. BHA, on the other hand, has the ability to circumvent the local optimum solution. The BHA has several characteristics with a number of different metaheuristics methods.

Hybridization is something that has been suggested in our research as a way to keep a healthy balance between exploration and exploitation, to prevent each algorithm from prematurely converging, and stay away from becoming stuck at local optimums. Inside this work, a novel hybrid algorithm, as a wrapper approach, of POA and BHA is presented in order to improve the technical specifications of our proposed algorithm and to attain faster convergence with the intention of escaping from local optima. The purpose of this work is to improve the performance of our proposed algorithm. To be more specific, BHA is integrated into POA in order to make POA more efficient by improving the exploitation and exploration capabilities of the POA algorithm. This hybridization will efficiently aid in the search to identify the most important factors associated of hyperparameters in that which is connected to the kernel semiparametric model with good prediction performance [23]. The hybridization process was carried out by creating an initial population of the Pelican swarm with random locations and also considering them as stars, then the best is calculated according to the objective function of the black hole and when the condition is satisfied, the movement will be from the Pelican algorithm equation, if the condition is of the black hole is not

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

achieved, the interstellar radius is calculated then Move according to the black hole equation until the condition is met, then return to the Pelican to complete the technique. Figure 1 presents a flowchart of our recently developed hybridization process.

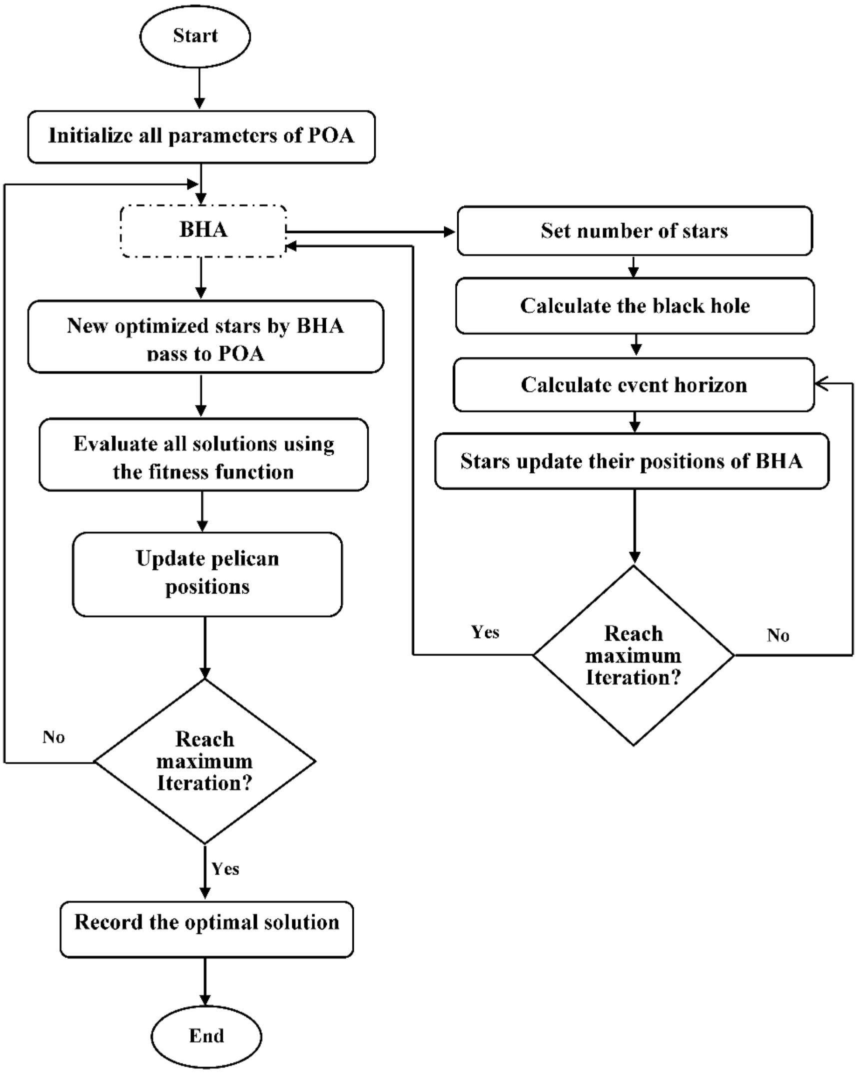

Figure 1: The flowchart of the new hybridization For Kernel Semi-Parametric Fusion Modeling .

Below, we detail the many parameters sets that can be used with our suggested technique:

a) 25 pelicans have been allotted, and $t_{max}=500$ iterations have been planned out.

b) Each pelican's location is chosen at random. The hyperparameter ρ is depicted here as the pelican in its typical position. Initial pelican locations are drawn from a uniform distribution in the interval [0.5,12].

c) a Fitness-F is characterized by the formula which is as follows:

$$\text{fitness} = \min \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \qquad (13)$$

d) Eq. (7) and (9) are used to recalculate the locations.

e) Iterate (c) and (d) until a t_{max} is achieved.

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

Pseudo-code of the Algorithm

Start
1. Input the initialization Pelican, and set number of stars
2. For each star i the fitness function is evaluated.
3. Determine the best star has the best fitness value as the Black Hole
4. Change the position of each star according to equation (11).
5. If a star arrives at a location with a cost lower than the Black Hole.
6. Calculate the event horizon radius.
7. When the distance between a star and the black hole, (the best star) is less than R, that star is collapsed and replaced by a new star that is randomly generated and distributed in the search space.
8. If the stop condition is satisfied, go to (9); otherwise, go to (6).
9. Initialization of the position of Pelicans and calculate the objective function.
10. Generate the position of the prey at random.
11. Phase 1: Moving towards prey (exploration phase).
12. Update the ith population member for pelican.
13. Phase 2: Winging on the water surface (exploitation phase).
14. Calculate new pelican of the jth dimension
15. Update the ith population member for pelican.
15. Update best candidate solution.
End

4. Experimental results

In this section, a comparison is made between our suggested hybridization, POA-BHA, and the regular POA and BHA, as well as the CV-method and the GCV-method. We set $RR=0.2$ in Eq (9). There are four different sample sizes, which are as follows: In order to increase the efficiency of the algorithms that are employed, four distinct models are being investigated. The simulation is carried out twenty-five times in total.

Four models will be used, and their regression functions will be as follows:

Model 1 [24]:

$$y_j = \sum_{i=1}^{6} x_{ji}\alpha_i + f(t_j) + \varepsilon_j \text{ , j = 1,..., N,} \qquad (14)$$

where $\alpha=(3,1,-3,2,-5,4)'$ and $f(t)=\frac{1}{3}\left[\phi(t;-3,0.81)+\phi(t;0,0.36)+\phi(t;3,0.81)\right]$, Which is a mix of normal densities for $t\in[-5,5]$ $and\ \phi(x;\mu,\sigma^2)$ is a "normal density function" with mean μ and variance σ^2. The main reason for selecting such a structure for the nonlinear part is to check the efficiency of nonparametric estimation for the wavy function. Moreover, $\varepsilon \sim N(0,\sigma^2 V)$ for which the

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

element of V are $v_{ji}=(\frac{1}{N})^{|j-i|}$. Four values of $\sigma^2=0.01$ is investigated. Produced randomly between 0 and 1 by uniform distribution, x is an explanatory variable used to make predictions.

Model 2[25]:

$$y_j = \sin\left(2\pi x_{1,j}\right) + 4\left(1+x_{2,j}\right)\left(1-x_{2,j}\right) + \frac{2x_{3,j}}{1+0.8x_{3,j}^2} + \varepsilon_j \sim N(0,0.09^2) \quad (15)$$

All of the independent variables, ranging in value from 0 to 1, are produced by a normal distribution.

Model 3:

$$y_j = \sum_{i=1}^{5} z_{ji}\alpha_i + g(u_j) + \varepsilon_j, \mathrm{j}=1,2,\ldots,\mathrm{N} \quad (16)$$

Where $\alpha=(1.5,2,3,-5,4)^T$, $u_j=\frac{(\mathrm{j}-0.5)}{N}$, and

$$g(u_j) = \sqrt{u_j(1-u_j)} * \sin\left(\frac{2.1*\pi}{u_j + 0.05}\right) \quad (17)$$

With $\varepsilon_j \sim N(0,\sigma^2)$ and $\sigma^2=0.01$.

Model 4:

$$y_j = \sum_{i=1}^{6} z_{ji}\alpha_i + g(u_j) + \varepsilon_j, \mathrm{j}=1,2,\ldots,\mathrm{N} \quad (18)$$

where $\alpha=(-1.-1,2,3,-5,4)^T$ and $g(u_j)=\sin(2\mathrm{u}_j)*\cos(5\mathrm{u}_j)$ with $\varepsilon_j \sim N_n(0,0.09)$.

By the Mean Square Error, the result of the comparison of the proposed POA-method, BHA-method, POA-BHA-method and CV-method and GCV-method in Tables 1 – 4. When we look at the findings that were acquired, we notice that POA-BHA-method demonstrated higher performance than the other algorithms. It has the lowest results when compared to the other rivals, thus this indicates that it is the most effective algorithm. This is then followed by the POA, which got the lowest MSE value, whilst the CV-method and GCV-method techniques are the least accurate ones. The suggested technique, which combines POA and BHA, has a somewhat better decrease in terms of MSE in comparison to both CV-method and GCV-method, while having just a slightly better reduction in comparison to POA and BHA. For Model 2, with n equal to 150, the reduction in MSE due to the utilization of POA-BHA was found to be 20.80% and 17.48%, respectively, when compared with the CV-method and GCV-method techniques. On the other side, when compared with BHA and POA, it was 6.31% and 5.51% respectively. With regard to the size of the sample, there is a correlation between an increase in the sample size and a reduction in the MSE values. In spite of this, the performance of the suggested approach, POA-BHA, is still superior to that of all other methods in every model.

Table 1: The Model 1's average MSE

Methods	n=30	n=50	n=100	n=150
CV-method	1.8322	1.6216	1.5212	1.3183
GCV-method	1.7965	1.5859	1.4855	1.2824
BHA-method	1.5402	1.3296	1.2292	1.0263

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

POA-method	1.4799	1.2693	1.1689	0.9658
POA-BHA-method	**1.3112**	**1.1006**	**1.0002**	**0.7972**

Table 2: The Model 2's average MSE

Methods	n=30	n=50	n=100	n=150
CV-method	2.7941	2.5835	2.4831	2.2821
GCV-method	2.7044	2.4938	2.3934	2.1905
BHA-method	2.4436	2.233	2.1326	1.9295
POA-method	2.4272	2.2166	2.1162	1.9131
POA-BHA-method	**2.3217**	**2.1111**	**2.0107**	**1.8076**

Table 3: The Model 3's average MSE

Methods	n=30	n=50	n=100	n=150
CV-method	3.8195	3.6089	3.5085	3.3054
GCV-method	3.7298	3.5192	3.4188	3.2157
BHA-method	3.469	3.2584	3.158	2.9549
POA-method	3.4526	3.242	3.1416	2.9385
POA-BHA-method	**3.3471**	**3.1365**	**3.0361**	**2.833**

Table 4: The Model 4's average MSE

Methods	n=30	n=50	n=100	n=150
CV-method	3.6854	3.4748	3.3744	3.1717
GCV-method	3.5957	3.3851	3.2847	3.0816
BHA-method	3.3349	3.1245	3.0239	2.8208
POA-method	3.3185	3.1079	3.0075	2.8044
POA-BHA-method	**3.213**	**3.0024**	**2.902**	**2.6989**

Tables 5–8 present the average amount of time required to do 25 iterations in order to further demonstrate the efficacy of the method we have designed. The time used by the algorithm should be as little as possible for optimal performance. The results of the suggested algorithm POA-BHA algorithm are presented in Tables 5–8, and they indicate that they are superior than those of the other algorithms. When compared to the other three models' methodologies, it produces the findings that are the least favorable. This was then followed by the POA, which ended up achieving the second lowest time, whilst the (CV-method and GCV-method) algorithms are the worst ones. The p-values (*) derived from the "Wilcoxon rank" sum test, which is a nonparametric statistical test, are utilized, and the significance threshold is set at 5%. It is necessary to do the statistical test in order to demonstrate that the POA-BHA offers a substantial advancement in comparison to the other algorithms. It is clear that there is a statistical difference between POA-BHA and CV-method and GCV-method for all models, and this difference can be observed in the data.

Based on the comparisons made above, one may get the conclusion that the suggested POA-BHA has promising outcomes and outperforms its other rivals, CV-method and GCV-method, by a significant margin.

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

Table 5: Model 1's computational time for each method used, is in seconds.

Methods	n=30	n=50	n=100	n=150
CV-method	88*	101*	109*	122*
GCV-method	75*	87*	96*	110*
BHA-method	67*	81*	88	103*
POA-method	60	73*	81	92
POA-BHA-method	**53**	**68**	**75**	**81**

Table 6: Model 2's computational time for each method used, is in seconds

Methods	n=30	n=50	n=100	n=150
CV-method	111*	121*	130*	143*
GCV-method	96*	108*	117*	130*
BHA-method	87	99*	111*	122*
POA-method	83	93	102*	115
POA-BHA-method	**73**	**85**	**91**	**107**

Table 7: Model 3's computational time for each method used, is in seconds.

Methods	n=30	n=50	n=100	n=150
CV-method	122*	134*	143*	156*
GCV-method	109*	121*	130*	143*
BHA-method	103*	113*	122*	135*
POA-method	94	106	115	128
POA-BHA-method	**84**	**92**	**107**	**118**

Table 8: Model 4's computational time for each method used, is in seconds.

Methods	n=30	n=50	n=100	n=150
CV-method	113*	125*	134*	147*
GCV-method	100*	112*	121*	134*
BHA-method	91*	107*	112*	128*
POA-method	85	97	106	119
POA-BHA-method	**77**	**85**	**96**	**111**

5. Conclusion

In this work, the problem of selecting an appropriate hyperparameter for the kernel semiparametric regression model is discussed. The selection of the hyperparameter required the development of a hybrid pelican optimization system as well as a black hole algorithm. The data that were acquired

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

from the simulation demonstrated that the hybridization method that was recommended had a lower mean squared error (MSE) in comparison to other competitor techniques. According to the results of the experiments and the statistical analysis, when our proposed hybridization algorithm is measured against the CV-method and GCV-method, it demonstrates superior performance in terms of the amount of time that is required. This conclusion was drawn from the findings of the experiments.

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] A. Ullah, A. T. K. Wan, H. Wang, X. Zhang, and G. Zou, "A semiparametric generalized ridge estimator and link with model averaging," *Econometric Reviews*, vol. 36, pp. 370-384, 2015.

[2] M. Roozbeh, M. Arashi, and H. A. Niroumand, "Semiparametric Ridge Regression Approach in Partially Linear Models," *Communications in Statistics - Simulation and Computation*, vol. 39, pp. 449-460, 2010.

[3] H. Emami, "Local influence for Liu estimators in semiparametric linear models," *Statistical Papers*, vol. 59, pp. 529-544, 2016.

[4] R. F. Engle, C. W. Granger, J. Rice, and A. Weiss, "Semiparametric estimates of the relation between weather and electricity sales," *Journal of the American statistical Association*, vol. 81, pp. 310-320, 1986.

[5] E. Goetghebeur and L. Ryan, "Semiparametric regression analysis of interval-censored data," *Biometrics*, vol. 56, pp. 1139-1144, 2000.

[6] Y. Qin, S. Zhang, X. Zhu, J. Zhang, and C. Zhang, "Semi-parametric optimization for missing data imputation," *Applied Intelligence*, vol. 27, pp. 79-88, 2007.

[7] X. Liu, J. Ning, X. He, B. C. Tilley, and R. Li, "Semiparametric regression modeling of the global percentile outcome," *Journal of Statistical Planning and Inference*, vol. 222, pp. 149-159, 2023.

[8] T. Zhang, "Semiparametric model building for regression models with time-varying parameters," *Journal of Econometrics*, vol. 187, pp. 189-200, 2015.

[9] E. D. D. Nkou, "Recursive kernel estimator in a semiparametric regression model," *Journal of Nonparametric Statistics*, pp. 1-27, 2022.

[10] M. Roozbeh, "Shrinkage ridge estimators in semiparametric regression models," *Journal of Multivariate Analysis*, vol. 136, pp. 56-74, 2015.

[11] O. M. Ismael, O. S. Qasim, and Z. Y. Algamal, "A new adaptive algorithm for v-support vector regression with feature selection using Harris hawks optimization algorithm," in *Journal of Physics: Conference Series*, 2021, p. 012057.

[12] O. S. Qasim and Z. Y. Algamal, "A gray wolf algorithm for feature and parameter selection of support vector classification," *International Journal of Computing Science and Mathematics*, vol. 13, pp. 93-102, 2021.

[13] H. Tamiminia, S. Homayouni, H. McNairn, and A. Safari, "A particle swarm optimized kernel-based clustering method for crop mapping from multi-temporal polarimetric L-band SAR observations," *International Journal of Applied Earth Observation and Geoinformation*, vol. 58, pp. 201-212, 2017.

[14] Israa Ezzat Salem, Mijwil, M., Alaa Wagih Abdulqader, Marwa M. Ismaeel, Anmar Alkhazraji, & Anas M. Zein Alaabdin. (2022). Introduction to The Data Mining Techniques in Cybersecurity . Mesopotamian Journal of CyberSecurity, 2022, 28–37. https://doi.org/10.58496/MJCS/2022/004

[15] A. M. Fathollahi-Fard, M. Hajiaghaei-Keshteli, and R. Tavakkoli-Moghaddam, "Red deer algorithm (RDA): a new nature-inspired meta-heuristic," *Soft Computing*, vol. 24, pp. 14637-14665, 2020.

[16] S. G. Mahmood Al-Kababchee, O. S. Qasim, and Z. Y. Algamal, "Improving penalized regression-based clustering model in big data," *Journal of Physics: Conference Series*, vol. 1897, 2021.

[17] P. Trojovsky and M. Dehghani, "Pelican Optimization Algorithm: A Novel Nature-Inspired Algorithm for Engineering Applications," *Sensors (Basel)*, vol. 22, Jan 23 2022.

[18] M. Y. Mustafa and Z. Y. Algamal, "Smoothing parameter selection in kernel nonparametric regression using bat optimization algorithm," *Journal of Physics: Conference Series*, vol. 1897, 2021.

[19] A. Hatamlou, "Black hole: A new heuristic optimization approach for data clustering," *Information sciences*, vol. 222, pp. 175-184, 2013.

[20] M. Nemati, R. Salimi, and N. Bazrkar, "Black holes algorithm: a swarm algorithm inspired of black holes for optimization problems," *IAES International Journal of Artificial Intelligence*, vol. 2, p. 143, 2013.

[21] W. Gao, X. Wang, S. Dai, and D. Chen, "Study on stability of high embankment slope based on black hole algorithm," *Environmental Earth Sciences*, vol. 75, p. 1381, 2016.

[22] F. Q. Kareem and A. M. Abdulazeez, "Ultrasound medical images classification based on deep learning algorithms: a review," *Fusion: Practice and Applications*, vol. 3, pp. 29-42, 2021.

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023

[23] C. Weng, R. B. Ghazali, S. Mustafa, A. N. Kareem, and B. A. Khalaf, "Weather forecasting for batu pahat using neural network," *Fusion: Practice and Applications*, vol. 6, pp. 64-70, 2021.

[24] E. Akdeniz, F. Akdeniz, and M. Roozbeh, "A new difference-based weighted mixed Liu estimator in partially linear models," *Statistics*, vol. 52, pp. 1309-1327, 2018.

[25] T.-H. Dang-Ha, F. M. Bianchi, and R. Olsson, "Local short term electricity load forecasting: Automatic approaches," in *2017 international joint conference on neural networks (ijcnn)*, 2017, pp. 4267-4274.

Doi : https://doi.org/10.54216/FPA.110104
Received: December 04, 2022 Accepted: March 09, 2023