

Article



# Energy Consumption Estimation of the Electric Bus Based on Grey Wolf Optimization Algorithm and Support Vector Machine Regression

Wei Qin <sup>1</sup><sup>(D)</sup>, Linhong Wang <sup>1</sup>,\*, Yuhan Liu <sup>1</sup> and Cheng Xu <sup>2</sup><sup>(D)</sup>

- <sup>1</sup> School of Transportation, Jilin University, Changchun 130022, China; tanwei1718@mails.jlu.edu.cn (W.Q.); liuyh1717@mails.jlu.edu.cn (Y.L.)
- <sup>2</sup> Department of Traffic Management Engineering, Zhejiang Police College, Hangzhou 310053, China; xucheng@zjjcxy.cn
- \* Correspondence: wanglinhong@jlu.edu.cn

Abstract: Electric buses have many significant advantages, such as zero emissions and low noise and energy consumption, making them play an important role in saving the operation cost of bus companies and reducing urban traffic pollution emissions. Therefore, in recent years, many cities in the world dedicate to promoting the electrification of public transport vehicles. Whereas due to the limitation of on-board battery capacity, the driving range of electric buses is relatively short. The accurate estimation of energy consumption on the electric bus routes is the premise of conducting bus scheduling and optimizing the layout of charging facilities. This study collected the actual operation data of three electric bus routes in Meihekou City, China, and established the support vector machine regression (SVR) model by taking the state of charge (SOC), trip travel time, mean environment temperature and air-conditioning operation time as the independent variables; while the energy consumptions of the route operations served as the dependent variables. Furthermore, the grey wolf optimization (GWO) algorithm was adopted to select the optimal parameters of the proposed model. Finally, a support vector machine regression model based on the grey wolf optimization algorithm (GWO-SVR) is proposed. Three real bus lines were taken as examples to validate the model. The results show that the mean average percentage error is 14.47% and the mean average error is 0.7776. In addition, the estimation accuracy and training time of the proposed model are superior to the genetic algorithm-back propagation neural network model and grid-search support vector machine regression model.

**Keywords:** electric bus; energy consumption estimation; support vector machine; grey wolf algorithm

# 1. Introduction

Compared with traditional fuel-driven buses, electric buses can achieve zero-emissions, quiet operation and low energy consumption, which are of great significance for saving the operation costs of a bus company, reducing the driving load of drivers and improving passenger comfort. The electrification of urban bus system becomes popular worldwide in recent years. Among them, the number of electric buses in China accounts for 90% of that in the world, and China has put 400,000 electric buses into service by the end of 2020 [1,2]. Many cities in China, such as Guangzhou, Shenzhen and Zhengzhou, own more than 10,000 electric buses in operation [3,4]. Apparently, electric buses have become the bus type with the highest proportion in China.

Nevertheless, electric buses still have shortcomings, such as short driving ranges and frequent charging because of battery capacity constraints. To avoid the service interruption caused by power exhaustion of electric buses during the operation process, bus companies are required to estimate energy consumption of bus routes accurately when they formulate



Citation: Qin, W.; Wang, L.; Liu, Y.; Xu, C. Energy Consumption Estimation of the Electric Bus Based on Grey Wolf Optimization Algorithm and Support Vector Machine Regression. *Sustainability* 2021, 13, 4689. https://doi.org/ 10.3390/su13094689

Academic Editors: Dongfang Ma, Sheng Jin and Dianhai Wang

Received: 16 March 2021 Accepted: 18 April 2021 Published: 22 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/). scheduling scheme. Additionally, the energy consumption data during the operation of electric buses is necessary for the setting of charging piles and the selection of optimal vehicle type. Thus, the energy consumption estimation of electric buses routes will have an important effect on the formulation of scientific bus operation management scheme.

There are a few studies on the energy consumption estimation of electric buses, and these mostly adopt lab data. However, the energy consumption will be influenced by many external factors during the process of bus operation, creating large variation from the lab environment, which leads to the low applicability of previous estimation methods. In this regard, the establishment of an energy consumption estimation method oriented to realistic operation environments is absolutely necessary for a better service of electric buses.

### 1.1. Literature Reviews

Over the years, researchers have mainly focused on the vehicle scheduling [5–10], charging infrastructure planning and lifecycle evaluation about electric buses [11]. For example, Lajunen et al. (2016) developed a simulation models of the different powertrains in the Autonomie vehicle simulation software [12]. Wang et al. (2017) developed a modeling framework to optimize electric bus recharging schedules [13]. Chen et al. (2018) investigated the cost competitiveness of different types of charging infrastructure [14]. Jari et al. (2018) considered three different charging infrastructure scenarios for a novel cost-benefit method for the scheduling of an electric city bus fleet on a single route [15]. Lajunen (2018) concluded that high energy capacity of the battery system was crucial for the overnight charging buses to achieve adequate daily operation, whereas the battery size had a minor impact on the energy consumption and lifecycle costs of the fast-charging buses [16]. Rogge et al. (2018) provided a methodology for the cost-optimized planning of depot charging battery bus fleets and their corresponding charging infrastructure [17].

Lin et al. (2019) focused on large-scale fast charging-station planning for e-buses in the public transportation electrification process [18]. He et al. (2019) built a mixed integer linear programming model with the objective of minimizing the total cost of vehicle batteries, fast-charging stations, energy storage systems, and electricity demand charges [19]. Liu et al. (2019) conducted a quantitative economic analysis of on-route fast charging for EBs [20]. Liu et al. (2020) developed an optimization model of the ETNDP (electric transit network design problem) with the constraints of route, depot, operation and charging [21]. Ritari et al. (2020) developed a two-speed dual clutch gearbox which can reduce energy consumption by 2–3.2% [22]. Teng et al. (2020) developed a multi-objective optimization model for a single bus line operated with electric buses [23]. Yao et al. (2020) proposed a new methodology for the electric vehicle scheduling problem with multiple vehicle types in public transport based on a given multi-vehicle-type timetable [24].

However, there are a few studies concerning energy consumption estimation. Cauwer et al. (2015) studied the correlations between the kinematic parameters of the electric vehicle and its energy consumption based on real world data [25]. Wang et al. (2017) developed a spatial-temporal estimation model of the remaining energy of pure electric buses using fuzzy clustering and time series algorithms based on the vehicle location and bus routes [26]. Beckers et al. (2019) proposed a physics-based energy consumption prediction model for electric city buses. The model was derived from the first principles and supplemented by special measurements for electric buses, including dynamometer testing and coasting measurements [27]. Al-Ogaili et al. (2020) proposed an integrated model that integrated digital elevation and longitudinal dynamics models. The integrated model was used to judge the energy consumption of the bus operation, so that a better electric bus type and battery type can be selected [28]. Pamula et al. (2020) proposed a deep learning network model for estimating energy consumption on the bus route. The model used arrival times at bus stops, bus stop locations, and traffic conditions as the input. These parameters were easy to be obtained [29].

Some studies on electric vehicles also help estimate the energy consumption of EBs [30–35]. Wang et al. (2018) generated accurate driving patterns by a Markov chain

Monte Carlo simulation. The simulated driving patterns were then used to undertake an uncertainty analysis on the network impact due to EV charging [36]. Lee et al. (2012) proposed a statistical modeling approach of daily driving mission sets. The approach consisted of temporal distribution modeling and the synthesis of individual representative cycles [37]. Zhang et al. (2015) developed a systematic energy consumption estimation approach which is suitable for electric vehicle actual driving cycles. Based on comprehensive consideration of the mechanical dynamics characteristics and electric machine system of the EVs, a set of energy consumption rate estimation models were established under different operation modes from a statistical perspective [38]. Fotouhi et al. (2020) developed an energy consumption estimation model to be used in an EV fleet management system. The proposed estimator consisted of a vehicle model, a driver model, and terrain models [39].

The research findings above have important implications for formulating scientific electric bus operation management scheme. It is less than 5 years since the electric buses have been put into use on a large scale, and the research results in terms of operation management are relatively few. For the aspect of energy consumption estimation, the problems of existing research are two-fold.

Firstly, although some researchers establish energy consumption estimation model based on the actual operating data of electric buses, the data they adopted has a small range of temperature, leading to limited application scope of models. Regarding some cold regions, the environment temperature can drop to  $-4 \degree F (-20 \degree C)$  or even  $-22 \degree F (-30 \degree C)$ . Under this condition, the energy consumption of electric buses will change dramatically, and current models will be inapplicable.

Moreover, some present estimation models will take the detailed movement state of buses on the route as inputs, such as the arrival time, traffic flow state at road segments or intersections and acceleration or deceleration process. At the stage of formulating vehicle scheduling scheme, however, bus companies generally have the fundamental data (average trip travel time, average speed, etc.) only. Furthermore, lacking detailed bus movement data makes it difficult to establish the model consistent with the real environment.

## 1.2. Contributions

This study takes actual state of charge (SOC) of battery on the electric bus, trip travel time, average environment temperature and air-conditioner operation time as the independent variables, meanwhile, takes the energy consumption from route operating as the dependent variable to establish the regression model based on support vector machine, with grey wolf optimization algorithm to obtain optimal parameters of the model. The main contributions of this paper are summarized as below.

(1) An energy consumption estimation model based on actual operating data of electric buses from cold regions is established. During the collection of sample data, the lowest environment temperature is -16.4 °F and the highest is 95 °F in order to broaden the application range of the model proposed.

(2) SVR and GWO algorithms are adopted to establish regression model and optimize parameters respectively. In particular, the model proposed employs dimension-based population initialization method, which is beneficial to relieve the problem of slow convergence speed at the beginning of iteration. Compared with other intelligent algorithms, the algorithm presented in this paper has the shorter training time and smaller estimation error.

The remainder of this paper are organized as follows. Section 2 gives a general description of the collected sample data. Section 3 introduces the SVM model, GWO algorithm; and establishes GWO-SVR energy consumption model. Section 4 implements the analysis, comparison and evaluation of the estimation results. Section 5 concludes this paper and discusses the potential future direction.

# 2. Data Description

This paper establishes an energy consumption estimation model based on the actual operating data of electric buses from Meihekou City, Jilin Province, China. Meihekou City located at  $125^{\circ}15'-126^{\circ}03'$  east longitude and  $42^{\circ}08'-43^{\circ}02'$  north latitude, with urban area of about 50 square kilometers and population of 350,000, belongs to temperate continental monsoon climate. Specifically, the average annual temperature is  $41.9^{\circ}F$ , and the record high and low temperature are 96.8 °F and  $-27.4^{\circ}F$ , respectively.

Currently, the total number of bus routes in operation is 17 in urban area of Meihekou City, supported by about 100 electric buses. We collect the operation data from 3 electric bus routes (route 103, 106 and 108). Three vehicle types can be found from these three bus routes, represented as type A, type B and type C. The detailed parameters of different bus vehicle type are described in Table 1.

Туре	Battery Capacity/kWh	Driving Range */km	Vehicle Size/m <sup>3</sup>	Curb Weight/kg	Passenger Capacity/pa **
А	162.3	345	$8.5\times2.5\times3.215$	8200	60
В	199.37	375	10.5  imes 2.5  imes 3.2	11,350	76
С	202.93	385	$10.5\times2.5\times3.215$	10,550	76

Table 1. Parameters of electric buses.

Note: \* Driving range is calculated based on uniform velocity method; \*\* pa is the abbreviation of passenger.

The data was collected from 1 May 2020 to 1 February 2021. Taking advantage of the battery management system and on-board video surveillance, a variety of data can be obtained, such as the date, bus license number, the departure time and end time of each trip, route length, energy consumption, SOC and the air-conditioner operation time. We collected the operating data of 3304 bus trips in total at last. Moreover, the average environment temperature when the bus serves each trip can be calculated according to the historic weather data.

After preliminary analysis of the data, four key factors correlated with energy consumption of electric bus routes are determined, and they are SOC, trip travel time, average environment temperature and air-conditioner operation time, respectively.

## (1) State of charge (SOC)

SOC is the level of charge of an electric battery relative to its rated capacity. The units of SOC are percentage points. SOC is normally used when discussing the current state of a battery in use. According to our actual observations: The power consumption rate is slower under higher SOC while faster under lower SOC. Thus, SOC is closely related to energy consumption [40].

#### (2) Trip travel time

Generally, the trip travel time and energy consumption will increase with route length, which is usually fixed, and the trip travel time during each period has obvious difference. As a result, this paper chooses trip travel time as the input of model. The length, average trip travel time, maximum and minimum trip travel time of 3 bus routes are listed in Table 2.

Route No.	Route Length/Km	Average Trip Travel Time/Min	Maximum Trip Travel Time/Min	Minimum Trip Travel Time/Min
103	10	25	59	14
106	12	35	52	14
108	8.1	28	50	15

**Table 2.** Statistic of travel time for electric bus routes

#### (3) Average environment temperature

The internal resistance of battery will increase at low temperature, leading to the growth of its calorific value and power consumption further. While under the high tem-

perature condition, the bus will start the cooling system to lower the battery temperature, which also adds to the power consumption. Therefore, environment temperature has a great influence on the battery, and the energy consumption of electric buses can be estimated through environment temperature. Figure 1 depicts the average environment temperature of each month in Meihekou City during the period of data collection.



Figure 1. Average temperature of each month during the investigation in Meihekou City.

## (4) Air-conditioner operation time

To improve the comfort of bus carriage, the driver usually starts the air-conditioner when the temperature is at a lower or higher level. However, powerful air-conditioners will accelerate the electricity consumption. Hence, air-conditioner operation time is a key factor influencing energy consumption along the route operation.

We use the relevant data in a city of Jilin Province to estimate energy consumption for electric buses. From May 2020 to Jan 2021, 3304 pieces of valid data originating from 3 bus routes were obtained, including the fundamental data, actual running data and average environment temperature.

Table 3 shows some samples of the collected data. Totally 3304 groups of data were collected from the 3 bus lines.

Date	License Number	Mileage /Km	Energy Consump- tion/Kwh	SOC/%	Trip Travel Time/Min	Average Environment Temperature/°F	Air-Conditioner Operation Time/Min
2020/5/19	3999	10	4.0	96.4	32	45.0	0
2020/6/1	0133	8.1	3.1	66.0	24	69.8	0
2020/8/15	7696	8.1	5.2	51.6	28	80.6	28
2020/11/18	1355	12	5.7	68.7	28	44.8	0
2020/12/7	3380	20	19.1	71.6	53	12.6	52
2021/1/8	6266	12	16.1	42.8	57	-6.7	57

Table 3. Samples of the collected data in Meihekou City.

Table 4 shows the result of statistical analysis for the data (energy consumption, SOC, trip travel time, average environment temperature and air-conditioner operation time) of 3304 bus trips.

Variable	Energy Consumption/kWh	SOC/%	Trip Travel Time/Min	Average Environment Temperature/°F	Air-Conditioner Operation Time/Min
Mean	5.6	70.7	29	41.8	5
Maximum	19.1	100.0	59	94.5	59
Minimum	2.0	22.8	14	-16.4	0

Table 4. Basic statistics of collected data.

#### 3. Establishment of Estimation Model

Many factors make an impact on the electricity consumption of battery as the operating of buses. For example, road slope is also one important contributing factor to electricity consumption. The road slopes of different bus lines are also different. Besides, for a given bus line, the road slope would vary with road sections. Thus, it is hard to find an indicator to represent the road slope of a bus line. Therefore, in this study the road slope is not considered. To eliminate the influence of road slope, we use the collected data of a bus line to establish the estimation model, and the model is only applied to this bus line.

The most relevant factors to are SOC, trip travel time, average environment temperature and air-conditioner operation time, respectively. Taking these four factors as dependent variables, we established a support vector regression model based on grey wolf optimization algorithm (GWO-SVR) to realize the estimation of energy consumption for electric buses.

A support vector regression (SVR) model is set up first in this section, and then performs parameter optimization for SVR model by means of GWO algorithm. The process steps about energy consumption estimation based on established GWO-SVR model is proposed eventually.

### 3.1. Support Vector Regression Model

Support vector machine (SVM) is a binary classification model, and it is a linear classifier that looks for the separation hyperplane with maximum interval in the feature space. Additionally, SVM can be used to conduct data regression analysis. When the input variables of sample data are classification data and the output variable is label, support vector classification (SVC) can be chosen to classify data; support vector regression (SVR) is applicable to the situation when both the input and output variables are numeric data. For the issue of energy consumption estimation for electric buses, we find it conforms to the application condition of SVR through observing sample data, so this paper will process the data with SVR.

SVR is a convex optimization problem, and it is local optimal solution equals to the global optimal solution. Specifically, the local optimum does not exist. Meanwhile, the optimization objective of SVR model is structural risk minimization, which can control the desired risk of learning machine on all the sample set with guaranteeing accuracy. Hence SVR model has great generalization ability that is better to improve the universality of the energy estimation model.

It is assumed that the dependent variables of SOC, trip travel time, average environment temperature and air-conditioner operation time are  $x_1$ ,  $x_2$ ,  $x_3$  and  $x_4$  respectively, and independent variable of energy consumption is y. This condition exists high-dimensional sample:

$$D = \{(x_{11}, x_{21}, x_{31}, x_{41}, y_1), (x_{12}, x_{22}, x_{32}, x_{42}, y_2) \cdots (x_{1m}, x_{2m}, x_{3m}, x_{4m}, y_m)\}, y_j \in \mathbb{R}.$$

If we intend to estimate the energy consumption y using SVR model, it is requisite to configure an energy consumption function f(x), making the estimation result approach to the actual value of y closely. The function f(x) is shown as Equation (1).

$$f(x) = \omega^T X + b \tag{1}$$

where *X* is the set of dependent variables,  $X = [x_1, x_2, x_3, x_4]$ . Vector  $\omega$  and parameter *b* are waiting to be determined.

Assume the error threshold between f(x) and y is  $\varepsilon$ . The loss should be calculated when the absolute value of error is more than  $\varepsilon$ , then expresses it by insensitive loss function  $l_{\varepsilon}$  in Equation (2).

$$l_{\varepsilon} = \begin{cases} 0, & if|z| \le \varepsilon \\ |z| - \varepsilon, & otherwisie \end{cases}$$
(2)

where z is the calculation value of loss.

According to Equations (1) and (2), the SVR problem can be rewritten as: what the value of  $\omega$  and *b* is when loss is the minimum. The equation is:

$$\min_{\omega,b} \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^m l_{\varepsilon}(f(x_i), y_i)$$
(3)

The classifier SVM adopted by SVR is linear, but it is incapable of performing data regression through constructing a hyperplane since input variables ( $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$ ) are not linearly differentiable data sets, thereby the requirement of error threshold cannot be satisfied. In this case, this paper introduces a slack variable, and Equation (3) is rewritten as Equation (4).

$$\min_{\substack{\omega,b,\xi_i,\widehat{\xi}_i}} \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^m \left( \xi_i + \widehat{\xi}_i \right)$$
s.t.  $f(x_i) - y_i \le \varepsilon + \xi_i$ ,  
 $y_i - f(x_i) \le \varepsilon + \widehat{\xi}_i$ ,  
 $\xi_i \ge 0, \ \widehat{\xi}_i \ge 0, i = 1, 2, \dots m$ .
$$(4)$$

where *C* is penalty coefficient;  $\xi_i$  and  $\xi_i$  are slack variable factors.

To obtain the required function f(x), the extremum of Equation (4) should be solved. For the problem with objective function and constraint, it is common to solve it by Lagrange multiplier method. After introducing Lagrange multiplier  $\mu_i$  into Equation (4), the expression of Equation (5) can be obtained below.

$$L(\omega, b, \alpha, \widehat{\alpha}, \xi, \xi, \mu, \widehat{\mu}) = \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^m \left(\xi_i + \widehat{\xi}_i\right) - \sum_{i=1}^m \xi_i \mu_i - \sum_{i=1}^m \widehat{\xi}_i \widehat{\mu}_i + \sum_{i=1}^m \alpha_i (f(x_i) - y_i - \varepsilon - \xi_i) + \sum_{i=1}^m \widehat{\alpha}_i \left(f(x_i) - y_i - \varepsilon - \widehat{\xi}_i\right)$$
(5)

Let the partial derivative of  $L(\omega, b, \alpha, \widehat{\alpha}, \xi, \widehat{\xi}, \mu, \widehat{\mu})$  to  $\omega, b, \xi_i$  and  $\widehat{\xi}_i$  equal 0, and then Equation (5) will be solved as follows.

$$\omega = \sum_{i=1}^{m} \left( \widehat{\alpha}_{i} - \alpha_{i} \right) x_{i} \tag{6}$$

$$0 = \sum_{i=1}^{m} \left( \widehat{\alpha}_{i} - \alpha_{i} \right) \tag{7}$$

$$\begin{cases} C = \alpha_i + \mu_i \\ C = \widehat{\alpha}_i + \widehat{\mu}_i \end{cases}$$
(8)

Any local extreme point of convex optimization problem is global extreme point. In a similar way, local optimum is equivalent to global optimum. Transforming the problem into a convex optimization problem will be in favor of obtaining optimal estimated value of energy consumption. SVR problem after introducing Lagrange multiplier does not

belong to convex optimization but its dual function does. As a result, we transform the original problem into dual problem, and finding that the optimal solution of original problem corresponds to that of dual problem depending on the hadron duality. Putting Equations (6)–(8) into Equation (5), we can acquire the dual problem of SVR in Equation (9).

$$\max_{\alpha,\widehat{\alpha}} \sum_{i=1}^{m} y_i(\widehat{\alpha}_i - \alpha_i) - \varepsilon(\widehat{\alpha}_i - \alpha_i) - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} (\widehat{\alpha}_i - \alpha_i)(\widehat{\alpha}_j - \alpha_j) x_i^T x_j$$

$$s.t. \sum_{i=1}^{m} (\widehat{\alpha}_i - \alpha_i) = 0, 0 \le \alpha_i, \widehat{\alpha}_i \le C.$$
(9)

where  $x_i$  and  $x_j$  are different feature vector;  $x_i^T x_j$  is the inner product value of  $x_i$  and  $x_j$ .

For convex optimization problem, KKT (Karush–Kuhn–Tucher) is the necessary and sufficient condition for the existence of extreme point which is also the global extreme point. Above process needs to satisfy KKT condition, that is to say Equation (10).

$$\begin{aligned}
\alpha_{i}(f(x_{i}) - y_{i} - \varepsilon - \xi_{i}) &= 0 \\
\widehat{\alpha}_{i}\left(y_{i} - f(x_{i}) - \varepsilon - \widehat{\xi}_{i}\right) &= 0 \\
\alpha_{i}\widehat{\alpha}_{i} &= 0, \xi_{i}\widehat{\xi}_{i} &= 0 \\
(C - \alpha_{i})\xi_{i} &= 0, \left(C - \widehat{\alpha}_{i}\right)\widehat{\xi}_{i} &= 0
\end{aligned}$$
(10)

The final result of SVR is below after substituting Equation (6) into Equation (1).

$$f(x) = \sum_{i=1}^{m} \left(\widehat{\alpha}_i - \alpha_i\right) x_i^T x + b \tag{11}$$

where parameter b can be calculated by Equation (12).

$$b = y_i + \varepsilon - \sum_{i}^{m} \left( \widehat{\alpha}_i - \alpha_i \right) x_i^T x, \alpha_i \in (0, C)$$
(12)

It is difficult to obtain inner product value through mapping energy consumption, as the input variables in this paper are four-dimensional. In this regard, we will use kernel function method and solve inner product taking advantage of the function relationship between inner product and input variables directly. Kernel function method can not only simplify the calculation process, but also cut down computational complexity. If we assume that  $\phi : x \to F$  is the map from input space to a certain feature space, then:

$$\omega = \sum_{i=1}^{m} \left( \widehat{\alpha}_i - \alpha_i \right) \phi(x) \tag{13}$$

If we replace the form of feature mapping by putting Equation (13) into Equation (11), SVR can be represented as Equation (14).

$$f(x) = \sum_{i=1}^{m} \left(\widehat{\alpha}_{i} - \alpha_{i}\right) \kappa\left(x_{i}^{T}x\right) + b$$
(14)

Here  $\kappa(x_i^T x) = \phi(x_i)^T \phi(x_j)$  is kernel function. In this paper, due to the lack of prior information about energy consumption, we adopt radial basis function (RBF). It enables the objective function map to infinite dimension and its specific form is given as Equation (15).

$$\kappa(x_i, x) = \exp(-\gamma ||x_i - x||_2) \tag{15}$$

The final form of f(x) after plugging Equation (15) into Equation (14) are:

$$f(x) = \sum_{i=1}^{m} \left(\widehat{\alpha}_{i} - \alpha_{i}\right) \exp\left(-\gamma ||x_{i} - x||^{2}\right) + b$$
(16)

Penalty coefficient *C* and the parameter coefficient of kernel function  $\gamma$  are the primary parameters that need to be calibrated while structuring the SVR model. The magnitude of *C* will have a straightforward effect on the noise tolerance, which will further influence the accuracy of proposed model. Similarly, the noise tolerance is also determined by  $\gamma$ , and the model would be insensitive to noise if  $\gamma$  is large, resulting in partial sample data loss.

The determination of *C* and  $\gamma$  should depend on the sparsity of input variables. However, since the sparse degree of  $x_1$ ,  $x_2$ ,  $x_3$  and  $x_4$  is unknown, parameter optimization has become a critical problem for SVR modeling.

#### 3.2. Parameter Selection Based on Gray Wolf Optimization Algorithm

We choose grey wolf optimizer algorithm for the purpose of solving parameter optimization problem towards the SVR model above. Mirjalili et al. (2014) introduced a population intelligent optimization algorithm (grey wolf optimizer, GWO). This algorithm can perform well for the optimization calculation with few parameters. Particularly, SVR model only has 2 parameters need to be calibrated, which conforms to the parameter optimization criteria of GWO greatly.

Simple structure and few parameters that need to be adjusted characterize GWO algorithm and it has self-adaptive convergence factor and information feedback mechanism, which can balance the local optimization and global search, with a good performance in both solving precision and convergence rate. For the parameter optimization of the SVR model, GWO has become a dominant optimization algorithm in parallel with grid search and particle swarm optimization.

This paper will utilize GWO algorithm, involving social hierarchy, encircling prey, hunting, attacking prey and searching for prey, to optimize the penalty coefficient *C* and the parameter coefficient of kernel function  $\gamma$  in the SVR model.

(1) Social hierarchy

The grey wolf social hierarchy is divided into four levels, consisting of the leader wolf  $\alpha$  responsible for making decisions; the subordinate wolf  $\beta$  which is probably the best candidate of  $\alpha$ ;  $\delta$  wolf (hunters and scouts mainly) responsible for governing bottomed wolves;  $\varphi$  wolf at the bottom of the society, with the largest number, performing as substitutes. The descending order of these four levels is  $\alpha$ ,  $\beta$ ,  $\delta$  and  $\varphi$  according to top-down obedience of the wolf pack.

The wolf  $\alpha$ ,  $\beta$ ,  $\delta$  and  $\varphi$  in the energy consumption estimation model carry position information and fitness information respectively. The position information of each wolf incorporates a bivector. Furthermore, the bivector of wolf  $\alpha$ ,  $\beta$ ,  $\delta$  correspond to the top, the secondary and the third optimal of parameter *C* and  $\gamma$  respectively, as well as the position information of wolf  $\omega$  covers parameter *C* and  $\gamma$  inside the searching scope. Therefore, the resulting position information of wolf  $\alpha$  solved by GWO would be the optimal value of parameter *C* and  $\gamma$ .

Grey wolves can adjust their position according to the distance to prey, and this distance is also called fitness function, which is the only indicator to judge the merits of position information of wolf pack. After each iteration, the position information of grey wolves is taken as the parameter of SVR model. We conduct training, energy consumption estimation and result evaluation based on SVR model, regarding evaluation index as fitness function. Root mean square error (*RMSE*) is an important model judgement index. The *RMSE* of test set is taken as the fitness function, and the calculation equation is shown in the following.

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2}$$
(17)

where  $y_i$  is the actual energy consumption of test set;  $\hat{y}_i$  is the estimation of energy consumption. It will be convenient to compare wolf  $\alpha$ ,  $\beta$ ,  $\delta$  and  $\varphi$  by means of fitness function, and carry out goodness ranking about parameter *C* and  $\gamma$  again, updating the position of wolf  $\alpha$ ,  $\beta$  and  $\delta$ .

(2) Encircling prey

Grey wolves approach their prey gradually by searching. GWO prepares for the next renewal of pack position information through updating internal parameters. To avoid falling into local optimum, GWO introduces a collaborative variable, taking advantage of its randomness to solve this problem. Specifically, the mathematical model about surrounding prey is:

$$D = B \circ X_p(t) - X(t) \tag{18}$$

$$X(t+1) = X_p(t) - A \circ D \tag{19}$$

$$A = 2a \circ r_1 - a \tag{20}$$

$$B = 2r_2 \tag{21}$$

where *D* denotes the distance between the grey wolf and their prey; *t* is the current iteration times;  $X_p$  is the position vector of prey; X(t) is the present position vector of grey wolves; *A* and *B* are collaborative coefficient vector;  $\circ$  represents Hadamard product operation; *a* is a constant with regard to maximum iterations, making *A* decline from 2 to 0 linearly during the iteration process;  $r_1$  and  $r_2$  are the random variables in the range of [0, 1].

(3) Hunting

Grey wolves can identify the position information of prey and surround them. GWO utilizes collaborative vector to achieve the renewal of position information. Given that wolf  $\alpha$ ,  $\beta$  and  $\delta$  could search prey and keep the optimal  $\alpha$ ,  $\beta$  and  $\delta$  during each iteration, searching prey based on their position and updating the position of current wolves, it is the process of updating and iteration information. Thus, the mathematical model of hunting model is:

$$\begin{cases}
D_{\alpha} = B_{1} \circ X_{\alpha} - X \\
D_{\beta} = B_{2} \circ X_{\beta} - X \\
D_{\delta} = B_{3} \circ X_{\delta} - X
\end{cases}$$
(22)

$$\begin{cases} X_1 = X_{\alpha} - A_1 \circ D_{\alpha} \\ X_2 = X_{\beta} - A_2 \circ D_{\beta} \\ X_3 = X_{\delta} - A_3 \circ D_{\delta} \end{cases}$$
(23)

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{24}$$

where  $X_{\alpha}$ ,  $X_{\beta}$  and  $X_{\delta}$  denote the position vector of wolf  $\alpha$ ,  $\beta$  and  $\delta$  in the present wolf pack; X is the position vector of subordinate wolves;  $D_{\alpha}$ ,  $D_{\beta}$  and  $D_{\delta}$  indicate the distance between wolf  $\alpha$ ,  $\beta$ ,  $\delta$  and subordinate wolves respectively, which is the value of fitness calculated based on position information.

(4) Attacking prey

Grey wolves attack the prey they select. GWO found the prey by constantly updating the location information and tried to attack. From an algorithm perspective, the prey here is a local optimal solution, and attacking the prey is a process of trying to solve the local optimal solution. In the process of solving the local optimal solution, the linearity of A decreases from 2 to 0. If |A| < 1, grey wolves would make an attack to prey corresponding to obtaining the optimal solution (local optimum).

(5) Searching for prey

Grey wolves are not satisfied with the present prey and they need to look for better prey. From an algorithm perspective, the local optimal solution is not necessarily the global optimal solution. The process of finding a better prey is actually to expand the search range and find the global optimal solution. GWO takes advantage of the randomness of collaborative variable *A* and *B* to jump out of the local optimum and search global optimal solution. The hunting zone can be expanded by controlling the magnitude of *A*. We set |A| > 1 to expand searching range compulsorily and avoid getting into local optimum; meanwhile, the randomness of *B* will be beneficial for finding new solutions. *B* indicates the random weight of the effect of the wolf's position on the prey. The weight will increase with *B*, and vice versa. Note that *B* is not linear decline, and it is the feature of nonlinear descent that exerts an influence on jumping out of local optimum during the later iteration.

## 3.3. Estimation of Operating Energy Consumption Based on GWO-SVR Model

Firstly, GWO-SVR algorithm optimizes penalty coefficient *C* and the parameter coefficient of kernel function ( $\gamma$ ) by GWO and then builds SVR model to train and estimate energy consumption. Figure 2 displays the detailed steps of the estimation for operation energy consumption.



Figure 2. Flow diagram of energy consumption estimation based on GWO-SVR

# 3.3.1. Data Preprocessing

Data preprocessing includes three parts: data elimination, data set partition, and data normalization.

In order to guarantee the validity of the model, it is necessary to eliminate abnormal data according to  $3\sigma$  criterion. Exactly speaking, we assume a set of test data consists of random error only, then calculate standard deviation and partition probability interval. The data that exceed the probability interval, usually caused by gross error, are considered as abnormal data to be eliminated. By collecting and observing mass operation data in field, we find that input variables conform to the normal distribution when the sample data size is large enough. It is eligible to eliminate abnormal data and calculate sample normal characteristics. If the data is out of tolerance scope, it would be eliminated on the basis of  $3\sigma$  criterion.

After that the reserved sample data is divided into training set and test set. Training set is the key to training model. Particularly, the merits and size of training set determine the training effect, while the test set aims to check the validity of the model. In this paper, we adopt fixed partition method, which is easy to compare and analyze validity of different model subsequently. The collected data will be partitioned according to the ratio of M (training set) to N (test set). As the proportion of M increases, the training performance of the model will be better, with strong generalization and high accuracy.

After partitioning the data set, the data needs to be normalized, that is, the sample is mapped to the interval [0, 1]. Later on, the contribution of each input variable can be balanced, so that the model convergence speed will be accelerated.

## 3.3.2. Determination of Parameter Initial Value for GWO-SVR Model

There are two parts about the determination of parameter initial value for GWO-SVR model, including that for SVR model and GWO model respectively.

(1) Determination of parameter initial value for SVR model

Penalty coefficient *C*, kernel function category of SVR, and their internal parameter  $\gamma$  should be confirmed during the process of modeling. Furthermore, these parameters can control the accuracy and generalization ability of SVR model, so parameter determination

The selection of kernel function mainly depends on the number of characteristics and the size of sample data. Specifically, it is the time when the number of sample characteristics is far more than the sample size, with characteristic dimension high enough, that the data is usually linear separable and the linear kernel function can be considered; when the sample size is common and sample characteristic dimension is not high, we generally choose radial basis function (RBF). Regarding the energy consumption estimation model in this paper, it has only four characteristics and mass sample data, therefore we will take RBF as the kernel function for the proposed SVR model.

In Equation (15), there is only one unknown parameter  $\gamma$  (controlling the noise tolerance of model) inside the RBF, and different sample should select different parameter  $\gamma$ .

Due to the difficulty to carry out visible measurement about the variables SOC, trip travel time, average environment temperature, air-conditioner operation time, and the high-dimensional space characteristics of energy consumption, as well as the unavailability of applicable parameters directly, thus it is necessary to utilize GWO to select the optimal parameters.

(2) Determination of parameter initial value for GWO model

SVR model has two uncertain parameters that need to be optimized based on GWO. During the process of optimizing, the primary parameters to be set are the number of grey wolves in the population *S*, the maximum iteration times *T*, position information, hunting zone, and step length.

The number of grey wolves in the population *S* refers to the number of  $\varphi$  wolves. As this number grows, the hunting zone for the wolf pack to seek optimal solution expands gradually. In general, the more the number, the wider the hunting zone, and the convergence speed will become faster, however the computation burden will increase greatly. For the energy consumption model in this paper, the value range of *S* should be 10 to 20.

The maximum iteration *T* denotes the times that grey wolves search for prey. Iteration times should be determined according to the sample data size. To be specific, when the sample size becomes large, the iteration times will also increase, adding the burden to computation. In this regard, we value *T* between 20 and 50 for our energy consumption estimation model.

At the same time, setting the range of search parameters will improve the accuracy of the optimal parameters and search efficiency of the optimal solution, in addition, reduce the time to select the optimal parameters.

# 3.3.3. Steps for GWO-SVR Model

**Step 1.** Set the position information of wolf  $\alpha$ ,  $\beta$ ,  $\phi$  and the initial value of fitness. After setting the internal parameters of grey wolves, we should preset position information and initial value of fitness. In terms of GWO algorithm, initial population will make a big impact on the global search speed and the quality of the solution. It is common that GWO algorithm sets the position information as the random value and sets the fitness as positive infinity to initiate iteration. In this paper, Consideration of the large sample size for the energy consumption, if we adopt a randomized policy to set initial value, the convergence speed would be slower. As a result, the position information of wolf  $\alpha$ ,  $\beta$  and  $\delta$  are given based on the method of dimension estimation.

Generally, the hunting zone of penalty parameter C is given as follows.

$$C \in \left[2^{-\dim}, 2^{\dim}\right] \tag{25}$$

where dim is the dimension of the input parameter.  $x_1$ ,  $x_2$ ,  $x_3$  and  $x_4$  are input parameters in our model, so dim equals 4.  $\gamma$  can has the same hunting zone as *C*, and its population initial value can be obtained from Equation (30).

$$\gamma = \frac{1}{\dim}$$
(26)

Taking the upper quartile, lower quartile, median, and the recommended value of  $\gamma$  as the initial value, which can obtain three sets of *C* and  $\gamma$ . Making use of these three sets of data, the root-mean-square error (RMSE) which performs as the fitness function could be calculated through SVR model, and the performance of the model is better with RMSE decrease. Next, the position information will be updated according to the order of the first, second, and third optimal values of RMSE.

Step 2. Determination of the initial value of collaborative vector.

It is required to preset  $r_1$ ,  $r_2$ , and a to calculate the initial value of collaborative vector. Both  $r_1$  and  $r_2$  are random vector in the range of [0, 1]. Parameter a is defined by the following equation.

$$a = 2 - \frac{2}{T} \tag{27}$$

where T is the maximum iterations.

According to Equation (20), the coefficient vector A and B of GWO can be obtained.

**Step 3.** Update the position information and fitness of wolf  $\alpha$ ,  $\beta$  and  $\delta$ .

The position information and fitness will be recalculated based on Equations (22) and (23), then through comparison, select the smaller adaptability to update step by step.

For example, as the calculation of the fitness of grey wolf  $\alpha$ , the first thing is to calibrate the parameters of SVR model employing position information. Position information consists of parameter *C* and  $\gamma$ , which have acquired initial value through dimension calculation in Step 1. Let the initial value of *C* and  $\gamma$  as the position information of grey wolf  $\alpha$ , and then the position information as the parameters of SVR model. Next, we will train SVR model with setting  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$  in the training set as the input variables and *y* as the output variables. After this,  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$  in the test set are taken as input variables to estimate energy consumption. Finally, the error between estimated value and actual value in the test set will be analyzed and take RMSE as the fitness of this position information.

Step 4. Update iteration.

Referring to Equation (18), the position information and fitness of each grey wolf  $\varphi$  are recalculated. RMSE is an important index to evaluate the validity of the model. In particular, the smaller the RMSE, the better the model. If the RMSE of grey wolf  $\alpha$ ,  $\beta$ ,  $\delta$  is less than that for a certain grey wolf  $\varphi_1$  in the grey wolves of  $\varphi$ , replacement will be carried out. For example, the RMSE of grey wolf  $\alpha$  is smaller than that of grey wolf  $\varphi_1$ , and then the position information and fitness of grey wolf  $\varphi_1$  will be replaced by grey wolf  $\alpha$ . At the same time, grey wolf  $\beta$  and  $\delta$  exchange their position information and fitness with each other.

This step aims to achieve the optimal replacement of position information. With the continuous renewal and iteration of parameters, also for the RMSE, it is further to find parameter *C* and  $\gamma$  which are closer to the optimal value.

Then according to Equation (20), compute the value of *A* and *B* for GWO taking advantage of the randomness of parameter  $r_1$ ,  $r_2$  and *a*. Repeat step4 ~step5 and conduct iteration.

Step 5. Result of GWO algorithm.

With the continuous renewal of the position information and fitness of grey wolf  $\alpha$ ,  $\beta$  and  $\delta$  in the population, the GWO algorithm is over when iterations come to the maximum. Meanwhile, the position information of grey wolf  $\alpha$  is selected as the final model parameters of *C* and  $\gamma$ .

**Step 6.** The final training of SVR model.

*C* and  $\gamma$  are selected as the final parameters for SVR model. Like step 3,  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$  in the sample training set are input variables and *y* is output variable. Finally, the SVR training model can be obtained.

Step 7. The energy consumption estimation and evaluation based on SVR model.

Taking SOC, trip travel time, average environment temperature, and air-conditioner operation time as input variables, the estimated value of energy consumption can be calculated and renormalized. By comparing the actual value with the estimated value, we will obtain the evaluation index of this model and conduct further analysis.

### 4. Results and Analysis

# 4.1. Selection of Evaluation Index

In this paper, we choose root-mean-square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE) to evaluate the validity of the estimation model of energy consumption.

In general, RMSE could reflect the deviation degree between the estimated value from regression model and the actual value well. However, if there are individual outliers with large deviations, even if there are few outliers, it will affect the evaluation results. Therefore, it is essential to use other indicators to supplement evaluation, such as MAPE or MAE. Moreover, the estimation result will be better with MAPE or MAE approaching 0 closer.

# 4.2. Estimation Result of GWO-SVR

This paper collects the relevant data in Meihekou City, China to estimate energy consumption of electric buses, and refers to the GWO-SVR model established in Section 3.3.3. We have 3304 original data samples in total (Including three bus routes), among which 150 pieces of abnormal data are removed based on  $3\sigma$  criterion. Furthermore, the reserved data (3154 pieces) are partitioned as the ratio of 9 (training set) to 1 (test set). We filter out valid correlation variables and their value from each piece of data according to the model proposed in this paper, including independent variables SOC ( $x_1$ ), trip travel time ( $x_2$ ), average environment temperature ( $x_3$ ), air-conditioner operation time ( $x_4$ ), as well as dependent variable energy consumption value (y). Besides, the sample data is normalized by means of Equation (28).

The parameter initial value of GWO algorithm in this paper includes: the number of grey wolves in the population (S = 20), the maximum iterations (T = 50), the dimension of position vector (N = 2), crossover probability (PCR = 0.2), searching lower limit (0.01), searching upper limit (100). The position information of wolf  $\alpha$ ,  $\beta$ ,  $\delta$  and fitness are set according to step 1 in Section 3.3.3, and the detailed position information and fitness are summarized in Table 5.

Grew Wolf Name	С	$\gamma$
α	4.046875	0.25
β	8.03125	0.25
δ	12.01563	0.25

Table 5. Initial position information and fitness value

To eliminate the impact of road slope, we develop the estimation model for each bus line. The detailed energy consumption estimation results of the three bus lines are shown in Table 6:

Route	Sample Number	RMSE	MAE	MAPE
108	1676	0.9730	0.7685	14.16%
103	1105	0.9821	0.7723	14.36%
106	523	0.9998	0.7922	14.89%

Table 6. Comparison of estimation results of three bus routes.

To analyze the error, we took the operating data of a bus of route 108 in January 2021 as an example, and the comparison result between its actual value and estimated value is shown in Figure 3.



**Figure 3.** Comparisons between the actual value and estimated value of operating energy consumption from a route 108 electric bus.

GWO-SVR model error is mainly determined by the penalty factor *C* and the  $\varepsilon$  in the loss function in Equation (2). *C* influences the random error of the model. As the value of *C* increases, the estimated energy consumption will tend to the average value, leading to larger estimation errors in some samples. Additionally,  $\varepsilon$  can prevent the trend of the estimated value from shifting to the average value, so that some samples have better estimation results.

According to the step 1 in Section 3.3.3, different setting method of initial value will have great influences on model convergence speed. In view of this, we have set the initial value of position information adopting the method based on dimension setting and random setting, as well as compared convergence speed and error under the same training set and test set. The results are illustrated by Figure 4.

From Figure 4, we can find that the model begins to approximate the optimal solution after 4 iterations when the initial value of position information is set based on dimension; while the position information is set randomly, it will take multiple iterations to get close to the optimal solution because of the long distance between initial value and the optimal value. Thus, we use the GWO algorithm based on dimension setting to solve the optimum quickly and efficiently.



**Figure 4.** The impact of different setting method for position information initial value on convergence speed.

#### 4.3. Comparison of Different Energy Consumption Estimation Models

To verify the validity of the model proposed in this paper, Genetic Algorithm-Back Propagation neural network model (GA-BP) and Grid-search Support vector machine regression model (Grid search-SVR) are used to compare with GWO-SVR algorithm.

(1) GA-BP neural network model

BP (back propagation) neural network is a kind of mathematical model with nonlinear mapping capability. By means of Error Backpropagation method, it takes minimum MSE as the objective to modify the weight and threshold of the grid continuously, thus the data can be fitted with high precision; Genetic Algorithm (GA) is a heuristic search algorithm referring to biological evolution process. Primarily, GA-BP neural network uses GA to continuously modify the weight and threshold of neural network until the optimal solution is obtained, then applies BP neural network to estimate the energy consumption. It has a good ability in energy consumption estimation in the case of small sample size, but for large sample size, the storage space for chromosomes will go up exponentially. The generalization of GA-BP neural network conflicts with it learning ability. As the number of samples increases, the generalization ability gradually increases. But when it rises to a certain limit, the increase in the number of samples will decrease the generalization ability, that is, the phenomenon of "over fitting" occurs.

Applying GA-BP neural network on energy consumption estimation, the hereditary algebra is set to 100; there are 3 layers for BP neural network, among which the input layer has four nodes. The number of neurons for hidden layer is 15 and output layer has 1 node. Meanwhile, sigmoid function is selected as the activation function for input layer and hidden layer while the output function is linear.

(2) GS-SVR model

For the SVR model, grid search (GS) is a popular parameter optimization algorithm. This algorithm is essentially exhaustive search, which is trying each value of parameters within the specified range and the results are calculated, then the optimal solution is obtained by comparing the results.

As the GS algorithm are exerted on the parameter optimization of SVR, it is required to regulate the range of penalty coefficient *C* and the parameter coefficient of kernel function ( $\gamma$ ), taking RMSE as the evaluation criterion of parameter goodness. The values range of both parameters *C* and  $\gamma$  is [-10, 10] and the calculation step length is 0.05.

(3) Result comparison

This paper compares GWO-SVR with GA-BP and GS-SVR and analyzes their estimation accuracy. To ensure comparability, only the data of bus line 108 is used when estimating energy consumption based on the three above models. In addition, the ratio of training set to test set is 9:1. The optimal parameters and the estimation results of energy consumption are demonstrated by Tables 7 and 8.

ESTIMATION METHOD	Penalty Coefficient C	Kernel Function Parameter $\gamma$
GS-SVR	0.4856	1.3966
GWO-SVR	0.4683	1.3856

Table 7. The optimal value for parameters of route 108.

Table 8. Comparison of estimation results from three models of route 108.

Estimation Method	RMSE	MAE	MAPE
GA-BP	1.134	0.9998	17.03%
GS-SVR	1.0001	0.7666	14.47%
GWO-SVR	0.9730	0.7685	14.16%

As Table 8 shows, the estimation error of GA-BP neural network is big slightly, while that of SVR is relatively small which is suitable for energy consumption estimation better. The reason is that, for larger sample size, BP neural network tends to be trapped in local optimum, leading to poor generalization. By contrast, SVR is a convex optimization problem, and it is will not fall into local optimum. As a result, two SVR models have higher precision in energy consumption estimation.

Comparing GWO-SVR with GS-SVR, we can find that there is little difference between them and the MAPE of GWO-SVR is only 0.31% which is smaller than that of GS-SVR model. The main reasons causing this including: (i) algorithm accuracy error, that is to say, estimation model achieves the optimal under its own optimization algorithm (such as GS and GWO algorithm), rather than the parameter condition of SVR model. Especially, the accuracy of GS algorithm mainly determines by the hunting zone and step length of C and  $\gamma$ , whereas for GWO algorithm, iteration is the major factor and sometimes the global optimal solution still cannot be found even the iterations reaches maximum; and (ii) random error, embodied in the process of SVR model looking for support vectors, and the selected support vectors may change with parameters, causing in that the training model has different performance.

GWO-SVR is superior to GS-SVR slightly concerning accuracy, but different in training time. In this paper, we have analyzed the training time and estimation error of three models by changing sample size. The detailed results are shown in Table 9.

Estimation Method	Sample Size	Training Time/s	MAPE
	500	5	7.90%
GA-BP	1000	9	13.98%
	1500	10	17.20%
	500	99	15.12%
GS-SVR	1000	1231	14.88%
	1500	2832	14.42%
	500	13	15.03%
GWO-SVR	1000	81	14.49%
	1500	122	14.09%

Table 9. Comparison of training time and estimation error among different models

From Table 9, we see that training time grows in varying degrees for three models with the increase of sample size. Moreover, the training time of BP neural network is less than that for SVR model under the same sample size; also, the MAPE of GA-BP neural network is minor when the sample size is small, but its estimation accuracy is inferior to SVR model

in large sample size. For two SVR model, if sample size is 1000, the training times of GWO-SVR and GS-SVR are 81 s and 1321 s respectively; if sample size is 1500, the training times are 122 s and 2832 s, with a huge difference. Although there is no remarkable difference between the estimation accuracy of two models, GS-SVR is far inferior to GWO-SVR on training time.

It can be observed from Table 8 that when the sample size is 500, the effect of using BP neural network for energy consumption estimation is better. This is because, in the case of a small sample size, the BP neural network method can build a neural network model by using fewer neurons, thereby reducing the probability of generating error neurons. Therefore, the BP neural network model has high performance when the sample size is small. With the increase of the sample size, the number of neurons in the BP neural network increases, and the probability of generating error neurons also rises, which will reduce the estimation accuracy of the model.

It can also be seen from Table 8 that the gray wolf optimization algorithm has almost the same training time as other algorithms with a little sample, while the gray wolf optimization algorithm has a significant advantage in training time as the number of samples increases. That is because, compared with other algorithms, the gray wolf optimization algorithm has the advantage of reducing time complexity and speeding up the solution. When the sample size is small, its advantages are not obvious. This is because the number of iterations of the gray wolf optimization algorithm is almost the same as that of other optimization algorithms (such as the grid search optimization algorithm), the times of them to find the optimal solution is almost the same. When the sample size is large, its advantages are obvious. This is because the gray wolf optimization algorithm, having fewer iterations, converges faster than the grid search optimization algorithm and can reduce a lot of running time.

Therefore, the GWO-SVR model proposed in this paper performs well on the data processing with large sample size.

#### 5. Conclusions

Energy consumption estimation is an important groundwork for the bus company operation management. On account of collecting actual electric buses data, this paper takes SOC, trip travel time, average environment temperature and air-conditioner operation time as independent variables and establishes the estimation model for operating energy consumption based on GWO-SVR. By compared with serval typical models, conclusions can be drawn as follows.

(1) The energy consumption model proposed in this paper mainly serves for bus scheduling, so the consideration about movement state along bus line is little, such as acceleration, deceleration, stop time, etc. From the perspective of estimation results, four influence factors that we select are more reasonable, and higher estimation accuracy can be obtained.

(2) Among the collected sample data, the maximum and minimum of average environment temperatures are 95 °F (35 °C) and -16.4 °F (-26.9 °C) respectively. In this case, the model that we propose in this paper will have a better regional applicability. To eliminate the impact of road conditions, we develop the estimation model for each bus line.

(3) Compared with other estimation models, it can be found that our model not only can achieve higher estimation accuracy, but also has a fast convergence speed when the sample size is large, which will be easier to find the global optimal solution. The original SVR model cannot do well with processing large sample data, however, by adopting GWO algorithm to optimize parameters, the computation speed could be improved greatly without increasing the error.

**Author Contributions:** Conceptualization, W.Q., L.W. and Y.L.; methodology, W.Q. and L.W.; validation, W.Q., L.W. and Y.L.; investigation, Y.L. and C.X.; writing—original draft preparation, W.Q. and L.W.; writing—review and editing, L.W. and Y.L.; visualization, Y.L.; supervision, L.W.; fund-

ing acquisition, L.W. and C.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** National Natural Science Foundation of China grant number (71771062). The Ministry of Education of Humanities and Social Science Project (No. 19YJCZH200).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Conflicts of Interest: The authors declared that they have no conflict of interest to this work.

# References

- China's Electric Buses Account for 90% of the World. World Wide Web. 2018. Available online: https://tech.huanqiu.com/article/ 9CaKrnKaeNP (accessed on 9 July 2018).
- 2. The proportion of electric buses nationwide has increased to 60%. *People's Daily Online*. 2020. Available online: http://finance.sina.com.cn/roll/2020-10-26/doc-iiznezxr8107403.shtml (accessed on 26 October 2020).
- Shenzhen realizes full electric bus transportation. *People's Sunshine Network*. 2017. Available online: http://www.peopleyes.com/ nd.jsp?id=836 (accessed on 9 July 2018).
- Guangzhou has over 10,000 pure electric buses. China News Network. 2018. Available online: https://baijiahao.baidu.com/s?id= 1618374183099661501&wfr=spider&for=pc (accessed on 28 November 2018).
- 5. Bie, Y.; Xiong, X.; Yan, Y.; Qu, X. Dynamic headway control for high-frequency bus line based on speed guidance and intersection signal adjustment. *Comput. Aided Civ. Infrastruct. Eng.* **2020**, *35*, 4–25. [CrossRef]
- 6. Bie, Y.; Ji, J.; Wang, X. Optimization of Electric Bus Scheduling Considering Stochastic Volatilities in Trip Travel Time and Energy Consumption. *Comput. Aided Civ. Infrastruct. Eng.* **2021**. [CrossRef]
- Bie, Y.; Gong, X.; Liu, Z. Time of day intervals partition for bus schedule using GPS data. *Transp. Res. Part C Emerg. Technol.* 2015, 60, 443–456. [CrossRef]
- 8. Wang, S.; Zhang, W.; Bie, Y.; Wang, K.; Diabat, A. Mixed-integer second-order cone programming model for bus route clustering problem. *Transp. Res. Part C Emerg. Technol.* **2019**, *102*, 351–369. [CrossRef]
- 9. Bie, Y.; Liu, Z.; Wang, H. Integrating Bus Priority and Pre-signal Method at a Signalized Intersection: Algorithm Development and Evaluation. *J. Transp. Eng. Part A Syst.* **2020**, *146*, 04020044. [CrossRef]
- 10. Liang, S.; He, S.; Zhang, H.; Ma, M. Optimal holding time calculation algorithm to improve the reliability of high frequency bus route considering the bus capacity constraint. *Reliab. Eng. Syst. Saf.* **2021**. [CrossRef]
- 11. Zhang, L.; Zeng, Z.; Qu, X. On the Role of Battery Capacity Fading Mechanism in the Lifecycle Cost of Electric Bus Fleet. *Ieee Trans. Intell. Transp. Syst.* **2020**. [CrossRef]
- 12. Lajunen, A.; Lipman, T. Lifecycle cost assessment and carbon dioxide emissions of diesel, natural gas, hybrid electric, fuel cell hybrid and electric transit buses. *Energy* **2016**, *106*, 329–342. [CrossRef]
- 13. Wang, Y.; Huang, Y.; Xu, J.; Barclay, N. Optimal recharging scheduling for urban electric buses: A case study in Davis. *Transp. Res. Part E Logist. Transp. Rev.* **2017**, *100*, 115–132. [CrossRef]
- 14. Chen, Z.; Yin, Y.; Song, Z. A cost-competitiveness analysis of charging infrastructure for electric bus operations. *Transp. Res. Part C Emerg. Technol.* **2018**, *93*, 351–366. [CrossRef]
- Vepsalainen, J.; Baldi, F.; Lajunen, A.; Kivekas, K.; Tammi, K. Cost-Benefit Analysis of Electric Bus Fleet with Various Operation Intervals. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 1522–1527.
- 16. Lajunen, A. Lifecycle costs and charging requirements of electric buses with different charging methods. J. Clean. Prod. 2018, 172, 56–67. [CrossRef]
- 17. Rogge, M.; van der Hurk, E.; Larsen, A.; Sauer, D.U. Electric bus fleet size and mix problem with optimization of charging infrastructure. *Appl. Energy* **2018**, *211*, 282–295. [CrossRef]
- 18. Lin, Y.; Zhang, K.; Shen, Z.J.M.; Ye, B.; Miao, L. Multistage large-scale charging station planning for electric buses considering transportation network and power grid. *Transp. Res. Part C Emerg. Technol.* **2019**, 107, 423–443. [CrossRef]
- 19. He, Y.; Song, Z.; Liu, Z. Fast-charging station deployment for battery electric bus systems considering electricity demand charges. *Sustain. Cities Soc.* **2019**, *48*, 101530. [CrossRef]
- 20. Liu, Z.; Song, Z.; He, Y. Economic analysis of on-route fast charging for battery electric buses: Case study in Utah. *Transp. Res. Rec. J. Transp. Res. Board* **2019**, 2673, 119–130. [CrossRef]
- 21. Liu, Y.; Feng, X.; Ding, C.; Hua, W.; Ruan, Z. Electric transit network design by an improved artificial fish-swarm algorithm. *J. Transp. Eng. Part A Syst.* **2020**, 146, 04020071. [CrossRef]
- 22. Ritari, A.; Tammi, K.; Laitinen, H. Energy consumption and lifecycle cost analysis of electric city buses with multispeed gearboxes. *Energies* **2020**, *13*, 2117. [CrossRef]

- Teng, J.; Chen, T.; Fan, W. Integrated approach to vehicle scheduling and bus timetabling for an electric bus line. *J. Transp. Eng.* 2020, 146, 04019073. [CrossRef]
- 24. Yao, E.; Liu, T.; Lu, T.; Yang, Y. Optimization of electric vehicle scheduling with multiple vehicle types in public transport. *Sustain. Cities Soc.* **2020**, *52*, 101862. [CrossRef]
- 25. Cauwer, C.; Mierlo, J.; Coosemans, T. Energy consumption prediction for electric vehicles based on real-world data. *Energies* **2015**, *8*, 8573–8593. [CrossRef]
- Wang, L.; Yang, J.; Zhang, N.; Yang, J.; Li, Y.; He, J.; Yang, F.; Zhou, H. A spatial-temporal estimation model of residual energy for pure electric buses based on traffic performance index. *Teh. Vjesn.* 2017, 24, 1803–1811.
- 27. Beckers CJ, J.; Besselink IJ, M.; Frints JJ, M.; Nijmeijer, H. Energy Consumption Prediction for Electric City Buses. In Proceedings of the 13th ITS European Congress, Brainport, The Netherlands, 4 June 2019; pp. 3–6.
- Al-Ogaili, A.S.; Ramasamy, A.; Hashim, T.J.T.; Al-Masri, A.N.; Hoon, Y.; Jebur, M.N.; Verayiah, R.; Marsadek, M. Estimation of the energy consumption of battery driven electric buses by integrating digital elevation and longitudinal dynamic models: Malaysia as a case study. *Appl. Energy* 2020, 280, 115873. [CrossRef]
- 29. Pamua, T.; Pamua, W. Estimation of the energy consumption of battery electric buses for public transport networks using real-world data and deep learning. *Energies* **2020**, *13*, 2340. [CrossRef]
- Ma, D.; Xiao, J.; Song, X.; Ma, X.; Jin, S. A back-pressure-based model with fixed phase sequences for traffic signal optimization under oversaturated networks. *IEEE Trans. Intell. Transp. Syst.* 2020. [CrossRef]
- 31. Ma, D.; Song, X.; Li, P. Daily traffic flow forecasting through a contextual convolutional recurrent neural network modeling inter-and intra-day traffic patterns. *IEEE Trans. Intell. Transp. Syst.* **2020**. [CrossRef]
- 32. Ma, D.; Xiao, J.; Ma, X. A decentralized model predictive traffic signal control method with fixed phase sequence for urban networks. *J. Intell. Transp. Syst.* 2020. [CrossRef]
- 33. Gao, K.; Yang, Y.; Li, A.; Li, J.; Yu, B. Quantifying economic benefits from free-floating bike-sharing systems: A trip-level inference approach and city-scale analysis. *Transp. Res. Part A Policy Pract.* **2021**, 144, 89–103. [CrossRef]
- 34. Xu, Y.; Zheng, Y.; Yang, Y. On the movement simulations of electric vehicles: A behavioral model-based approach. *Appl. Energy* **2020**, *283*, 116356. [CrossRef]
- Gao, K.; Yang, Y.; Sun, L.; Qu, X. Revealing psychological inertia in mode shift behavior and its quantitative influences on commuting trips. *Transp. Res. Part F Traffic Psychol. Behav.* 2020, 71, 272–287. [CrossRef]
- 36. Wang, Y.; Infield, D. Markov Chain Monte Carlo simulation of electric vehicle use for network integration studies. *Int. J. Electr. Power Energy Syst.* **2018**, *99*, 85–94. [CrossRef]
- 37. Lee, T.-K.; Bareket, Z.; Gordon, T.; Filipi, Z.S. Stochastic modeling for studies of real-world PHEV usage: Driving schedule and daily temporal distributions. *IEEE Trans. Veh. Technol.* **2011**, *61*, 1493–1502. [CrossRef]
- Zhang, R.; Yao, E. Electric vehicles' energy consumption estimation with real driving condition data. *Transp. Res. Part D Transp. Environ.* 2015, 41, 177–187. [CrossRef]
- Fotouhi, A.; Shateri, N.; Shona Laila, D.; Auger, D.J. Electric vehicle energy consumption estimation for a fleet management system. *Int. J. Sustain. Transp.* 2020, 15, 40–54. [CrossRef]
- 40. Li, W.; Stanula, P.; Egede, P.; Kara, S.; Herrmann, C. Determining the Main Factors Influencing the Energy Consumption of Electric Vehicles in the Usage Phase. *Procedia Cirp* **2016**, *48*, 352–357. [CrossRef]