Journal of Cleaner Production 250 (2020) 119492

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Influencing factors analysis and forecasting of residential energy-related CO_2 emissions utilizing optimized support vector machine

Lei Wen, Yang Cao*

Department of Economics and Management, North China Electric Power University, Baoding, 071003, Hebei, China

A R T I C L E I N F O

Article history:
Received 10 August 2019
Received in revised form
16 October 2019
Accepted 27 November 2019
Available online 28 November 2019

Handling Editor: Prof. Jiri Jaromir Klemeš

Keywords:
CO_2 emissions
Principal component analysis
Chicken swarm optimization
Support vector machine

A B S T R A C T

With the economy development and rapid urbanization, the residential usage of energy has been increasing in China, leading to more CO_2 emissions in resident sector. To predict the trend of residential energy-related CO_2 emissions accurately, it is significant to analyze the influential factors. In this paper, 18 preliminary indicators are identified by grey relational analysis to prove their correlation with CO_2 emissions firstly. To reduce the redundancy of data, 4 main components are extracted by principal component analysis as predicting input data of support vector machine (SVM). By adding chaotic mutation and nonlinear weight index, the improved chicken swarm optimization (ICSO) algorithm is proposed to optimize the parameters of SVM, hereafter referred as ICSO-SVM. Finally, the new hybrid model is applied to predict residential energy-related CO_2 emissions in Shanghai, China. The simulation results in the forecasting accuracy demonstrate that the ICSO-SVM model outperforms the compared original chicken swarm optimization model (CSO-SVM), particle swarm optimization model (PSO-SVM), genetic algorithm optimization model (GA-SVM) and basic SVM. The rigorous influencing factors analysis and the outstanding performance in predicting CO_2 emissions of ICSO-SVM model can offer relevant scholars and policy makers more breakthrough points of residential CO_2 emissions abatement.

1. Introduction

Greenhouse effect has led to tough challenges in human living environment, and the anthropogenic greenhouse gases emissions account for climate change principally, which mainly come from fossil energy combustion, taking for around 70% of the total (IPCC, 2007). In 2018, China has ranked first in the world in energy consumption, accounting for 24% of the world's, meanwhile, it takes up 27.8% of global CO_2 emissions, almost twice as much as the second emitter, the United States (BP, 2019). Though a developing country, during the thirteenth Five-Year Plan, spanning 2016 to 2020, China designs to abate carbon intensity per unit of GDP by 40–45% by 2020, while 60–65% in 2030 during 14th Five-Year Plan, respectively, on the basis of the 2005 level. Therefore, with the increasing severe challenges, it's necessary for China to formulate practical strategies for CO_2 emissions reduction.

While among all sectors in China, the energy consumption of residential sector has grown dramatically, which has more than tripled from 166.95 Mtce in 2000 to 542.09 Mtce in 2016. According to Nie and Kemp (2014), residential department in China is a principal energy consumer and the second largest only after industrial sector. With the rapid advance of people's living standards, the residential energy consumption may continue to rise, which would result in more CO_2 emissions (Nejat et al., 2015). At present, our efforts to abate CO_2 emissions have not been enough, and it is vital to arouse people's concern about reducing CO_2 emissions and guide them to live a low-carbon life (Wakiyama and Kuramochi, 2020). But the issue is that we haven't fully explored the factors that affect residential CO_2 emissions. If we can predict the emissions precisely based on these elements, not only can we verify these influencing factors, but also change the future trend of CO_2 emissions by adjusting these driving forces.

Given the China's ambitious commitments for CO_2 emissions abatement, it is of great significance to analyze the influencing factors causing residential energy-related CO_2 emissions and put forward relevant countermeasure. When analyzing the driving forces and measurement of residential CO_2 emissions in China, the study with evidence from specific Chinese cities are scarce. Most of previous researches, in views of the national level, considered the broad energy consumption data and provincial distribution, which

* Corresponding author.
E-mail address: 429641088@qq.com (Y. Cao).

https://doi.org/10.1016/j.jclepro.2019.119492
0959-6526/

ignore the detailed residential living characteristics to some extent (Ye et al., 2017; Rodrigues et al., 2019). Nevertheless, according to Shen et al. (2018), the national targets for cutting carbon emissions need to be allocated to individual provinces even cities. That is because the rapid urbanization has expedited the pace of economy and society development, making cities the core of the global economy (Lin and Benjamin, 2019). Shanghai is the largest coastal city in the east of China with prosperous economy and ultra-modern urban construction representing the future development of cities. Considering its high urbanization rate and fast rhythm of life, Shanghai's residential department may be in a phase of high energy consumption leading to amounts of CO_2 emissions. Hence, it is imperative to analyze in detail its influencing elements of residential energy-related CO_2 emissions in order to excavate superiority and weakness in how emissions could be mitigated which will be a blue print for local government and neighboring cities.

To analyze the driving forces associated with residential CO_2 emissions, various methods have been utilized by previous scholars in recent years. There are there major methods: (1) STIRPAT model (Miao, 2017); (2) the log mean divisa index (LMDI) model (Wu et al., 2019); (3) the structural decomposition analysis (SDA) model (Zhang et al., 2017); (4) quantitative methods, containing panel data (Yao et al., 2018), multiple-factor regression analysis (Li et al., 2019), quantile regression (Han et al., 2015) and so on. For instance, Wang and Zhao (2018) introduced an improved STIPAT model with panel data focusing on the regional level as well as the macro level to analyze the residential CO_2 emissions in China, and their study indicated that the household size, urbanization rate, residential consumption level and Engel Coefficient are driving factors. Through SDA at the regional level in China, Wang et al. (2019) demonstrated that the residential consumption level has a positive impact on residential CO_2 emissions and that the optimization of consumption structure would be a potential trend to achieve CO_2 emissions abatement. Besides, the grey relational analysis (Huang and Wang, 2016) and principal component analysis (Sun and Sun, 2017) are also available to explore influencing indicators, which has not been utilized in residential CO_2 emissions.

With respect to forecasting CO_2 emissions, various scholars have predicted the future trends through quantitative models (Ding et al., 2017). The multiple linear regression and multiple polynomial regression were adopted to forecast Iran's carbon dioxide emissions in 2030 under different scenarios by Hosseini et al. (2019). Song et al. (2018) applied a dynamic integrated input-output model to predict the emission peak prior to 2030 for China. However, the ability of traditional prediction methods to dispose of nonlinear sequences is not as good as artificial intelligence technology. Due to their powerful capability in dealing with nonlinear and complex problems, classical machine learning are overwhelmingly exploited in predicting CO_2 emissions. Moreover, the combination of intelligent algorithms and machine learning tools has further improved the forecasting accuracy (Zhao et al., 2018). An improved particle swarm optimization-back propagation neural network (PSO-BPNN) model was established by Sun and Xu (2016) to extract the effecting factors and CO_2 emissions in Hebei Province, China. Proposed by Sun et al. (2017), an optimized extreme learning machine (ELM) model was utilized to make a precise forecasting on CO_2 emissions, which demonstrated an improved accuracy. And SVM optimized by genetic algorithm (GA-SVM) was used to forecast CO_2 emissions of Beijing from 2016 to 2020 with scenario analysis (Li et al., 2017). To forecast carbon emissions, Qiao et al. (2019) optimized the least squares support vector machine (LSSVM) on the basis of combining lion swarm optimization and genetic algorithm which showed remarkable improvement of forecasting efficiency.

Summing up researches mentioned above, when predicting CO_2 emissions, the rationality of factors selection often was simplified or ignored by some predecessors. And the conventional influencing factors analysis methods, such as LMDI and STIPAT, can only analyze several indicators since it is difficult to add new variable in equation. For detailed consideration of more residential characteristics, this paper applies grey relational analysis to determine their relations with residential CO_2 emissions. Furthermore, to eliminate data redundancy, namely, comprehensively integrating the complex relations among driving elements, we apply principal component analysis to extract main factors which will be introduced as the inputs of forecasting model. Given the outstanding capacity of SVM in processing small sample data (Li et al., 2020), the SVM is chosen as the basic tool for CO_2 emissions forecasting. The performance of SVM model highly depends on internal parameters, so this work adopts chicken swarm optimization (CSO) algorithm to search the optimal parameters. Although intelligent optimization algorithm can observably improve the forecasting performance of machine learning tools, defects like easily trapping into local optimum, overfitting still remain when faced with specific issues (Fang et al., 2018). Hence, we proposed a improved befitting model ICSO-SVM to figure out corresponding forecasting problem.

The innovation and contribution of the present thesis is revealed as follows. First, in addition to fossil energy, when estimating the residential energy-related CO_2 emissions of Shanghai from 2000 to 2016, the local power and thermal CO_2 emission levels are taken into account to obtain a more accurate measurement of CO_2 emissions. Second, 18 pre-selection indicators, over the same time span, are identified by grey relational analysis to determine the influential factors. Third, the verified indicators are classified into 4 categories according to relevance and economic connotation: economic level, living standard, social condition and education level, then principal component analysis is applied to extract every category to simplify the input data of prediction model. Fourth, the improved chicken swarm optimization (ICSO) is created through adding logistic map and learning index to the parameters-optimized process of SVM, in a certain extent to strengthen the searching efficiency and robustness. Fifth, the new hybrid model ICSO-SVM is introduced, for the first time, to forecast CO_2 emission. Through the simulation results of forecasting residential CO_2 emissions in Shanghai, compared to other CSO-SVM, GA-SVM, PSO-SVM and SVM methods, the ICSO-SVM displays fantastic property in CO_2 emissions forecasting.

The rest of this paper is organized as follows: Section 2 introduces the methodology, including the SVM, CSO and ICSO-SVM. Section 3 introduces data analysis process and extracts the principal components. Section 4 shows the favorable performance of ICSO-SVM in CO_2 emissions forecasting, and discusses the simulation results. Section 5 provides the conclusion.

2. Methodology

2.1. Support vector machine (SVM)

Support vector machine is a widely used machine learning tool, originating from statistical learning theory. In the early development stage, it is utilized to solve problems of samples classification, then it developed into a forecasting tool according to the basic theory. For a given dataset $\{x_i, y_i\}_{i=1}^N$, N is the number of training samples; x_i and y_i are the input and output variables, respectively. The core concept is to map low-dimensional input data into high-dimensional space based on the nonlinear mapping function $\phi(x)$. And the objective is to find a function $f(x)$ as close as possible to all sample points. Then the regression equation is performed as Eq. (1):

$$f(x) = w \cdot \phi(x) + b \tag{1}$$

where w represents the weight tensor, and b is the bias.

On the principal of structural risk minimization, the regression issue can be replaced with constrained optimization problem mathematically. The optimization function and constraints are expressed at Eqs. (2) and (3):

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\delta_i + \delta_i^*)\tag{2}$$

$$\begin{cases} y_i - w\cdot\phi(x) - b \le \varepsilon + \delta_i \\ w\cdot\phi(x) + b - y_i \le \varepsilon + \delta_i^* \\ \delta_i \ge 0 \\ \delta_i^* \ge 0 \end{cases}, i = 1, 2, \cdots, N \tag{3}$$

In the above equations, C denotes the penalty factor, δ_i and δ_i^* are slack variables, ε is the loss function. Too large C will weaken the model's generalization ability, while a too small value may increase its fitting error.

Furthermore, through introducing Lagrangian multiplier and the solution of its dual problem, the nonlinear function can be presented as Eq. (4):

$$f\left(x\right) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K\left(x_i, x_j\right) + b \tag{4}$$

where α_i and α_i^* are Lagrangian multipliers, and $K(x_i, x_j)$ represents the kernel function.

Regarding the selection of kernel function, SVM will possess different performances of learning and generalization ability by choosing different kernel functions. Commonly used kernel functions are linear kernel function, polynomial kernel function and radial basis function (RBF). While the RBF has better performance than linear kernel when handle non-linear problems and has less hyper-parameters than polynomial kernel function. Hence, RBF is widely identified as an ideal function in processing complex and multidimensional samples (Dong et al., 2005; Ahmad et al., 2018). In this paper, RFB is utilized as kernel function, defined as Eq. (5):

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{5}$$

where σ denotes the width of RBF, which is wide enough to dispose finite sample.

In classical SVM, the two parameters, penalty factor C and σ in RBF, are chosen empirically or grid search method with instability and complicated calculation. To better address the issue, many intelligent algorithms are introduced into the parameters optimization. This work applied improved chicken swarm algorithm to search optimal parameters of SVM. Furthermore, forecasting mechanism of SVM is shown as Fig. 1, in this paper the input data are influencing factors, and the output is CO_2 emissions.

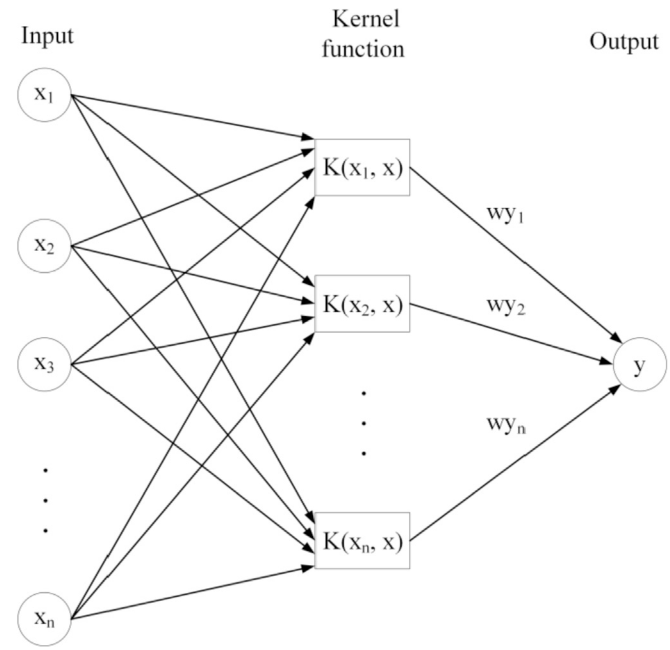

Fig. 1. The mechanism of SVM.

2.2. Chicken swarm optimization (CSO)

Proposed by Meng et al. (2014), the chicken swarm optimization, a bionic intelligent algorithm, mimics the hierarchy in the chicken swarm consisting of rooters, hens and chicks, along with their characteristic behaviors of chicken swarm. It can utilize the chickens' swarm intelligence to efficiently solve optimization problems. Specifically, each group includes one dominant rooster, several hens and chicks, and different hierarchy follows different orders of motions.

When solving optimization problems, each individual's position corresponds to a solution. Assume the following numbers: N_R, N_H, N_C and N_M, which indicate the number of roosters, of hens, of chicks, and of mother hens, respectively. In a chickens swarm of N individuals, each chicken is described by their position x_{ij}^t $(i \in [1, \cdots, N], j \in [1, \cdots, D])$ at step t, search for food in a D-dimensional space. It's significant to clarify that the optimization problems are the minimal targets in this study, thus the best N_R chicken correspond to the ones with minimal fitness values. As mentioned above, there are three types of chicken, each with its own particular equations of motion, which will be introduced in turn.

The rooters owning best fitness values can forage food in a more extensive space than that of other chickens with worse fitness values. Their movement is depicted as following:

$$x_{ij}^{t+1} = x_{ij}^t * \left[1 + randn\ \left(0, \sigma^2\right)\right] \tag{6}$$

$$\sigma^2 = \begin{cases} 1 & , \quad f_i \le f_k, \\ \exp\left(\dfrac{f_k - f_i}{|f_i| + \varepsilon}\right) & , \quad f_i > f, \end{cases} \quad k \in [1, N], k \notin i \tag{7}$$

where $randn\ (0, \sigma^2)$ is a Gaussian distribution with mean 0 and standard deviation σ^2; k is a rooster's index, randomly selected from the roosters group, and f_i is the fitness value of the corresponding x_i. In addition, ε is the smallest constant in the computer to avoid being divided by zero error.

With regard to hens, they follow the roosters in their own group. Furthermore, they also randomly trace other chickens for good food. When it comes to competing for food, the dominant hens have advantage over the more submissive ones in competing for food. These scenes can be described mathematically as follows:

$$x_{ij}^{t+1} = x_{ij}^{t} + S_1 * rand * \left(x_{r1j}^{t} - x_{ij}^{t}\right) + S_2 * rand * \left(x_{r2j}^{t} - x_{ij}^{t}\right) \quad (8)$$

$$S_1 = \exp\left(\frac{f_i - f_{r1}}{|f_i| + \varepsilon}\right) \quad (9)$$

$$S_2 = \exp(f_{r2} - f_i) \quad (10)$$

where *rand* is a random number uniformly distributed between $[0, 1]$. $r1 \in [1, \cdots, N]$ is an index of the rooter, the i^{th} hen's groupmate, while $r2 \in [1, \cdots, N]$ is an index of a chicken (rooster or hen), chosen from the chicken swarm randomly ($r1 \notin r2$).

As for chicks, they only seek for food after their mother hen. The movement updated criteria of chick is formulated below:

$$x_{ij}^{t+1} = x_{ij}^{t} + FL * \left(x_{mj}^{t} - x_{ij}^{t}\right) \quad (11)$$

$$FL = rand * 0.4 + 0.5 \quad (12)$$

where x_{mj}^{t} represents for the position of the i^{th} chick's mother, $m \in [1, N]$. And *FL* is a parameter that indicates the degree, a chick would follow its mother, chosen randomly from 0 to 2.

2.3. Improved chicken swarm optimization algorithm (ICSO)

2.3.1. Logistic map chaos optimization

Chaos, which has been universally introduced in optimization algorithm, is a common phenomenon that exists in nonlinear systems. The position of the population is generally random in a algorithm, and if they gather in one small area, the algorithm may plunge into local optimum. Featured by randomness and ergodicity in a bounded space, chaotic variables are more suitable in global optimization algorithm than random variables. In this work, every iterative optimal solution will be recorded first, then chaos map is applied to update the original optimal solutions for the diversity of population in order to avoid the stagnation of optimization. Therefore, a mutation probability p_i changing with the number of iterations is supposed to be introduced, which is described in Eq. (13) and Fig. 2. This study sets a exponential function to make it nonlinear, and this setting has a greater mutation probability, close to 100%, to avoid premature convergence in early iteration while smaller to avert that the global optimal solution may be replaced frequently when near the end. Showing excellent ergodicity and higher iterative speed, the logistic map is utilized to generate the chaotic series in population mutation. The expression of logistic map is Eq. (14):

$$p_i = p_{\min} + (p_{\max} - p_{\min}) * \exp\left(-k_1 * (i/M)^{k_2}\right) \quad (13)$$

$$X_{k+1} = \mu * X_k(1 - X_k) \quad (14)$$

where p_i indicates the i^{th} mutation probability, through experiments the $p_{\min}$ is set as 0.2 and $p_{\max}$ is 1.0, while k_1 and k_2 are set as 2 and 3, respectively; X_k denotes the k^{th} chaotic variable and k is the iteration number. In the equation, $X \in (0, 1)$, $k = 0, 2, \cdots, pop - 1$, while *pop* is the population. Has been tested in lots of experiments, under the conditions that the initial $X_0 \in (0, 1)$ and that $\mu = 4$, a chaotic sequence can be generated (Ricardo et al., 2017).

2.3.2. Chicks' movement optimization

In Eq. (11), chicks only follow their mother hens' way to search food instead of considering the rooster in their group. If mom hens got bogged down in local optimum, their chicks will fall into the same dilemma. Meanwhile, combined with the idea of particle swarm optimization algorithm, the inertia weight w is introduced in process of chicks' self-learning. The large w is beneficial in the initial stage for convergence acceleration over the whole search space, while in the end stage it needs to be small to improve search precision. Therefore, in this paper, we develop the chicks' movement by adding the two sections of rooster-learning weight R and self-learning index w, the improvement is denoted by the following equations:

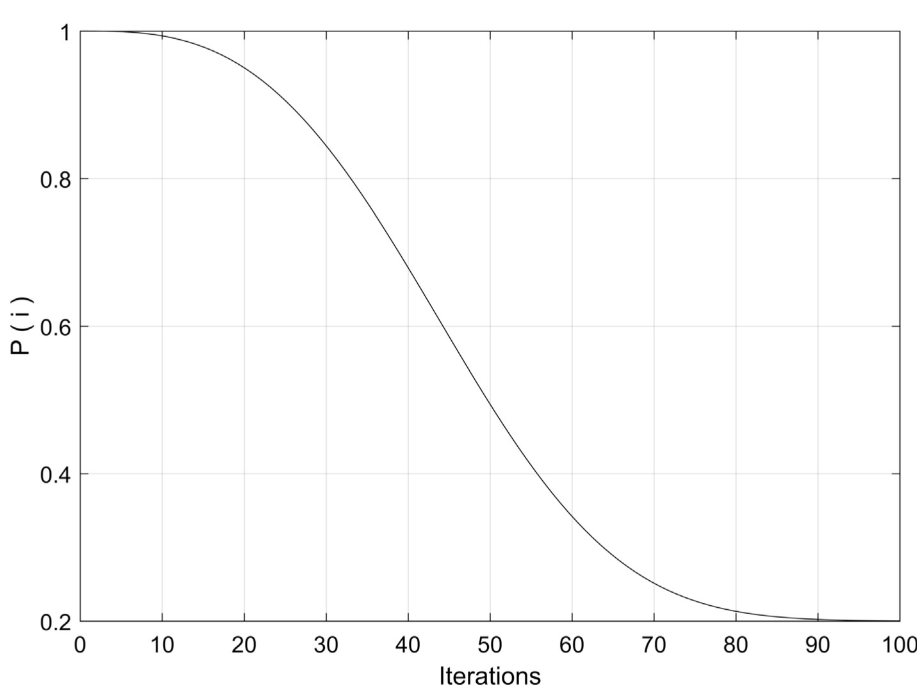

Fig. 2. The function curve of the probability of chaos mutation.

$$x_{ij}^{t+1} = w_i * x_{ij}^t + FL * \left(x_{mj}^t - x_{ij}^t \right) + R * \left(x_{rj}^t - x_{ij}^t \right) \quad (15)$$

$$R = rand * 0.4 + 0.5 \quad (16)$$

$$w_i = w_{\min} + (w_{\max} - w_{\min}) * \left(1 - (\, i/M \,)^k \right) \quad (17)$$

where the power function w_i represents the i^{th} self-learning index, *rand* is a random number from $[0, 1]$. The experimental results' setting of $w_{\min}$, $w_{\max}$ are 0.4 and 0.9, respectively, and a convex curve is expected to get for a mitigation of weight decline rate, so k is set as 2.0.

2.4. Forecasting model architecture

The forecasting model ICSO-SVM consists of three parts, whose flowchart is shown as Fig. 4. In this model, influencing factors are the input data, and CO_2 emissions are output data. Therefore, part 1 is the selection and analysis of influencing factors. The 18 preliminary indicators are identified by grey relational analysis first, then remaining 17 indicators are divided into 4 categories in the light of correlation and connotation among factors. The principal component analysis are introduced in each category to extract principal components as the input data. Part 2 is the iterative optimization procedure of improved chicken swarm algorithm using training sample data. To improve the ergodicity of CSO algorithm, this study introduces chaos mutation, namely, that the optimal solution in each iteration will be changed by the logistic mapping with a decreasing mutation probability. Eventually, the optimal solution of part 2 will be as the best parameters value of SVM in part 3 for prediction. Through comparison with other models in prediction accuracy based on testing data, the outstanding performance of ICSO-SVM in forecasting will be demonstrated. Fig. 3.

3. Data analysis

3.1. Data preprocessing and preliminary factors selection

This work estimates the residential energy-related CO_2 emissions in Shanghai and analyzes it's driving forces, and the data spanning from 2000 to 2016 years are acquired from Shanghai Statistical Yearbook and China Statistical Yearbook.

Though the data of residential energy-related CO_2 emissions isn't published directly, they can be converted by the energy consumption of residential sector in China Energy Statistical Yearbook and conversion coefficient obtained from IPCC (IPCC, 2006). However, when calculating the CO_2 emissions from the final energy consumption of residents, it would result in issues if the emission coefficients of electricity and heat are neglected (Fan et al., 2015). On the one hand, from global viewpoint, when computing the total CO_2 emissions of overall sectors, we should regard electricity and heat as secondary energy to avoid redundant computation. On the other hand, with regard to the calculation of CO_2 emissions of a specific sector such as residential sector, the residential energy-related consumption activities include cooking, lighting, heating, hot water and so on (Hirano et al., 2016). The consumption of electricity and heat should be taken into consideration because it is the significant part in residential energy consumption with no problems of double counting. Thus, residential energy consumption in Shanghai contains 12 types of energy: row coal, briquettes, coke oven gas, other gas, petrol, kerosene, diesel oil, lubricants, liquefied petroleum gas (LPG), natural gas, heat and electricity. To facilitate the description of the energy structure, these 12 energy are classified into 6 categories, namely coal (row coal and briquettes), coal gas (coke oven gas and other gas), oil (petrol, kerosene, diesel oil, lubricants, LPG), heat and electricity. But it is worth noting that during the process of calculating CO_2 emissions this paper still follows the 12 types of commercial energy in detail.

As illustrated in Fig. 5, the residential energy consumption structure in Shanghai changes dramatically from 2000 to 2016. Especially, the consumption percentage of coal descends sharply

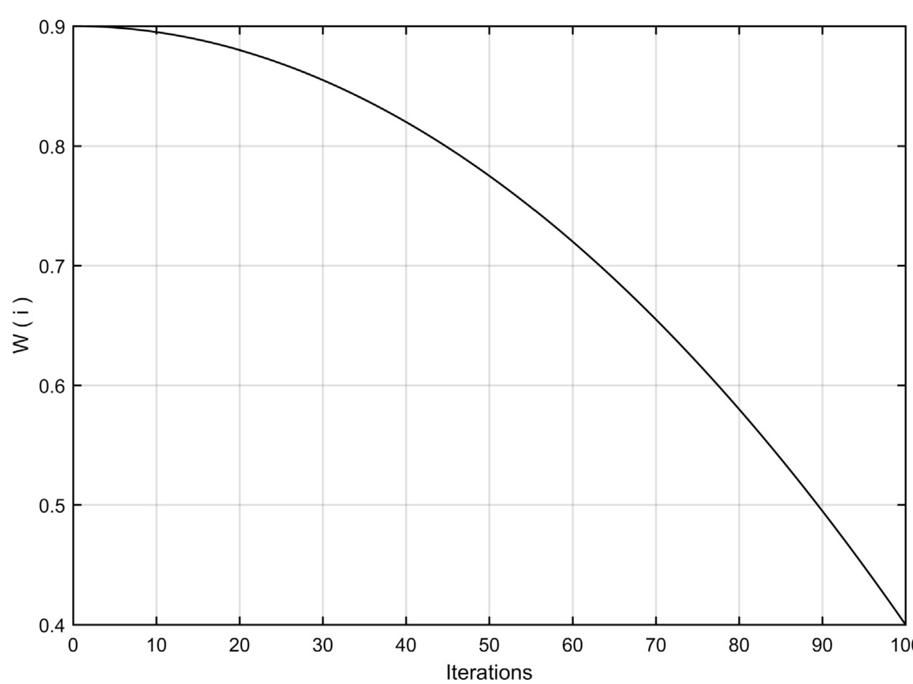

Fig. 3. The function curve of nonlinear self-learning index.

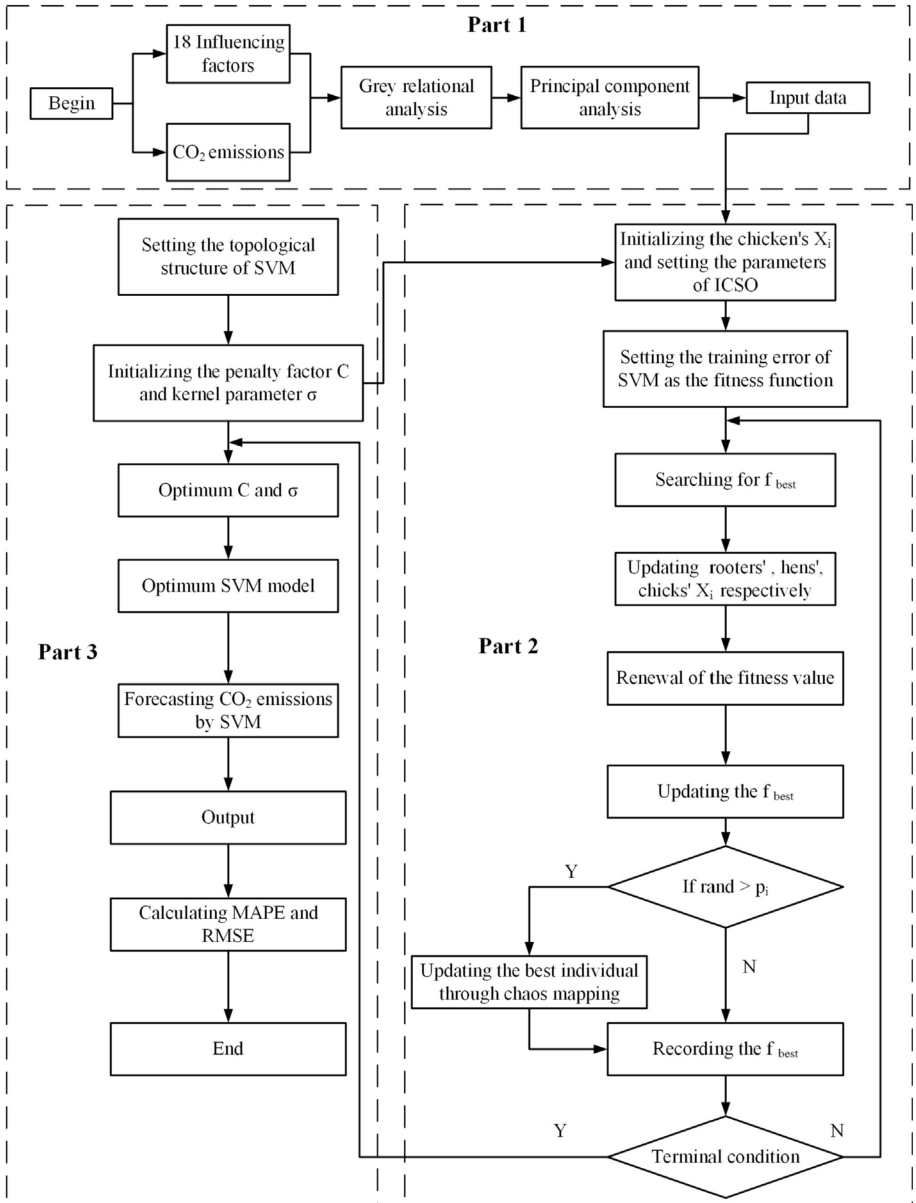

Fig. 4. The flowchart of the ICSO-SVM.

from 27.65% in 2000 to 1.44% in 2016, and coal gas has a similarly evident decline, which are connected with the fact that natural gas is becoming an alternative to them. By contrast, oil consumption has more than doubled in sixteen years to 50.55%, largely related to the popularity of vehicles. As Shanghai is located in southern China, not in the heating supply area, the residential heat consumption is little even no in recent years. While electricity always takes a big part which is one of the major residential energy consumption, ranging from around 20%–30%. Tough the residential consumption of conventional fossil energy lessens gradually, the sharply raising usage of oil, natural gas and electricity still causes the continuous augment of CO_2 emissions.

Regarding the CO_2 emissions of electricity and heat, a lot of previous literature didn't consider or just deemed unchanged on national-level overlooking the local technology level. Contributing to a more precise measurement of residential CO_2 emission in Shanghai, this article first figures out the comprehensive CO_2 emission coefficients of electricity and heat in terms of time series from 2000 to 2016, which can reveal the local CO_2 emission level during the generation of electricity and heat. Mathematically, the formulas for the CO_2 emission coefficient per unit of electricity and heat in year t are as follows, defied as $CE_{ele,\ t}$ and $CE_{heat,\ t}$ respectively.

$$CE_{ele,\ t} = \frac{\sum_{i} E_{i,\ t}^{ele} * NCV_i * CC_i * O_i}{Q_t^{ele}} \tag{19}$$

$$CE_{heat,\ t} = \frac{\sum_{i} E_{i,\ t}^{heat} * NCV_i * CC_i * O_i}{Q_t^{heat}} \tag{20}$$

where $E_{i,\ t}^{ele}$ and $E_{i,\ t}^{heat}$ represent the consumed quantity of the i energy in the process of power and heat generation in t^{th} year in Shanghai, including electricity and heat themselves; NCV_i, CC_i and O_i mean the net calorific value, carbon content, carbon oxidation

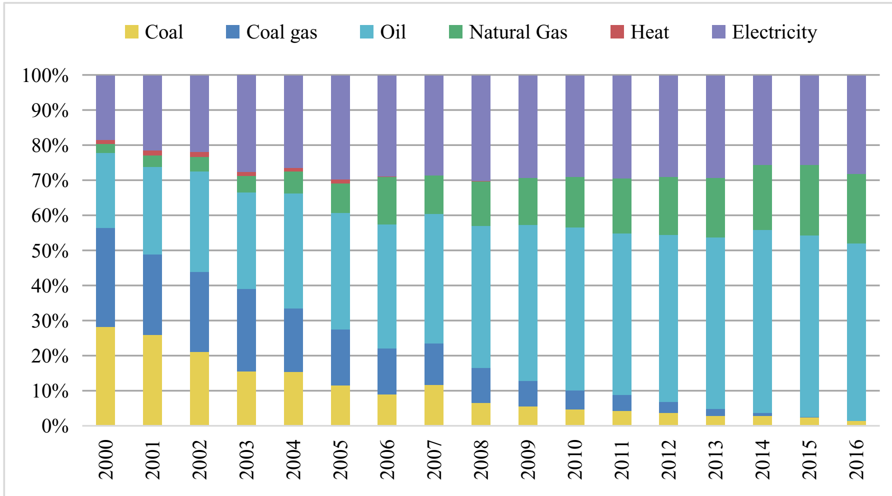

Fig. 5. Residential energy consumption structure in Shanghai from 2000 to 2016.

factor of i energy respectively, and all values are from IPCC (2006) precisely. In this work, we assume they remain constant over the selected t years, that is, the technology progress for these energies during the period has no enormous changes; Q_t^{ele} is the total power production in year t, and Q_t^{heat} denotes the heat generation.

The computing results are exhibited in Fig. 6, with the gradual development of solar power, wind power and other forms of renewable sources power generation, the integrated electricity CO_2 emission coefficient fluctuates over time, for example 7.39 tCO_2/Tce in 2011 while 6.51 tCO_2/Tce in 2012. During the same time span, that of heat has no significant change, fluctuating slightly around 2.9 tCO_2/Tce in general.

Referring to the method recommended by IPCC and the above mentioned results, the integrated computation of residential energy-related CO_2 emissions is presented by the following formula:

$$C=\sum_i C_i+\sum_j C_j=\left[\sum_i (E_i * NCV_i * CC_i * O_i)+\sum_t\sum_j E_j * CE_j\right]\times\frac{44}{12} \tag{21}$$

where C denotes the total amount of residential energy-related carbon dioxide emissions, C_i means that of fossil energy i, including row coal, briquettes, coke oven gas, other gas, petrol, kerosene, diesel oil, liquefied petroleum gas and natural gas. C_j represents that of secondary energy: electricity and heat; the 44 and 12 respectively are the molecular mass of carbon dioxide and carbon. The other indexes have the same meaning as Eqs. (19) and (20).

Shown in Fig. 7, the total residential CO_2 emissions in Shanghai exhibits an increasing trend from 10.39 Mt in 2000 to 31.83 Mt in 2016 by an average annual rate of 7.24%. Obviously, electricity-related CO_2 emissions occupy the largest share of the total, ranging from 43.59% to 60.79%. And the resident power consumption mainly goes into running household appliances, such as washing machine, air conditioner, refrigerator and so on, therefore, it makes sense to study their impact on CO_2 emissions. Ranked second is the oil-related emissions whose ratio has more than doubled, accounting for 15.67% in 2000 while 34.29% in 2016. As the raising ownership of family car, the petrol combustion puts a greater concentration of CO_2 into the atmosphere than before. As an alternative to coal and gas, the increasing consumption of natural gas has emitted more CO_2 raising from 0.16 Mt to 3.99 Mt. The analysis of CO_2 emission structure will be helpful to our future abatement research.

As for the selection of affecting factors, the above mentioned researches commonly considered only several elements, this paper selects 18 preliminary indicators, taking comprehensive consideration of more elements. Based on relativity, economic connotation,

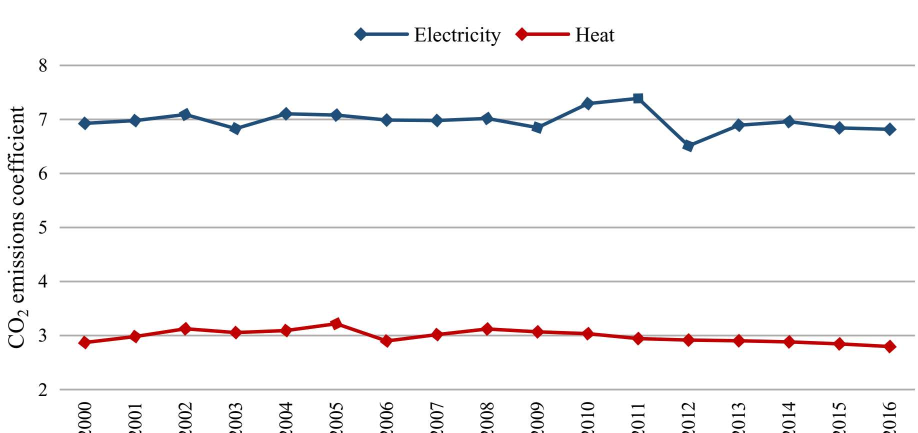

Fig. 6. Electricity and heat CO_2 coefficients in Shanghai during 2000–2016.

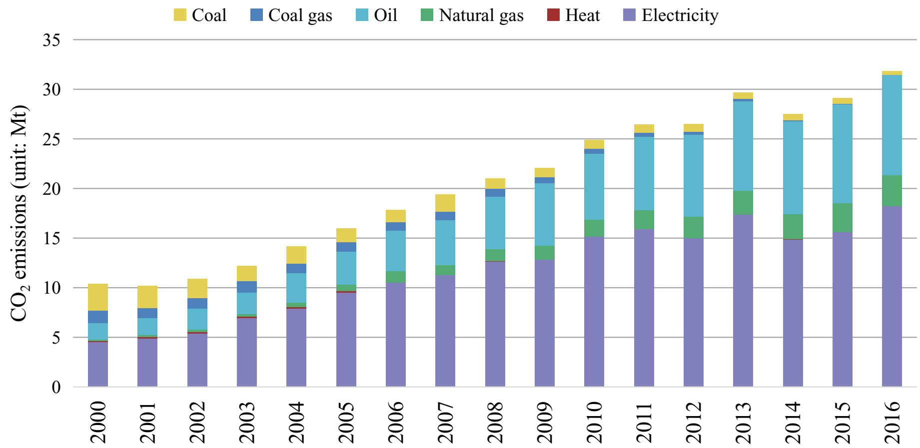

Fig. 7. Residential CO_2 emissions from different kinds of energy in Shanghai.

and the classification experience of predecessors, the 18 indicators will be classified into 4 specific categories in this article: economic level, living standard, social condition and education level. In order to facilitate the description of these factors, these preliminary factors were defined as F_{ij} in sequence, where i means the category $i \in [1, 2, 3, 4]$. The selection and classification of variables are summarized in Table 1. It is important to note that all economic indicators in this article are in constant 2000 prices, and the measurement of ownership is the holding volume every 100 households.

3.2. Grey relational analysis

Introduced in 1982, the grey theory has proven to be a practical scientific theory especially in the research of uncertain system. As a significant part of it, the grey relational analysis is applied to judge the indefinite relation between two time series by the geometrical shape comparability. The CO_2 emission is influenced synthetically by various factors, thus it's appropriate to utilize this method to analysis the correlation degree between multiple ore-selected and CO_2 emissions.

To eliminate the effect of index dimension and quantity of data, this study divides each indicator by the initial value for normalization. Then calculating the grey relation between residential energy-related CO_2 emissions and various influencing factors using Matlab 2016a, the result is shown in Fig. 8. Due to the classification of grey correlation, the value beyond 0.8 represents there is a strong correlation, while 0.6–0.8 shows there is a general relevance. Except that the grey relational degree of the F_{23} is 0.628, other indicators all shows significant relevancy. Specifically, F_{23} means the ownership of motorcycle, which only emerges weak correlation with CO_2 emissions because of the severely motorcycle ban in Shanghai a previous years. Hence, the F_{23} will not be introduced as affecting factor in the next work. Ultimately, the result reveals that there exists strong relevance between the remaining indicators and residential energy-related CO_2 emissions, which will all be taken into consideration in the later work.

3.3. Principal component analysis

As first introduced by Pearson in 1901, the principal component analysis (PCA) was applied to achieve statistical dimensionality reduction by transforming a set of potentially related variables into linearly independent ones. As a result, the transformed variables are called principal components, which can involve all information of various factors in a lower dimension. Therefore, we capture the method to extract main factors from influencing indicator group to simplify the input variables of subsequent forecasting model.

In this study, the PCA is conducted on 4 above mentioned categories in SPSS 22 respectively. Taking the living standard as example, the results of principal component analysis are presented in Table 2. First, the value of Kaiser-Meyer-Olkin (KMO) is 0.852, which is greater than 0.8, remarkably suitable to conduct principal component analysis. While 0.5 is served as the criteria to judge whether principal component analysis can be applied (Antanasijevic et al., 2014). Meanwhile the probability of bartlett's test of sphericity is 0.000, which reveals that the correlation matrix is not a unit matrix. It can be seen that the contribution rate of the first principal factor reaches up to 89.355% which can well cover all the information of this category, so in this paper the first common factor will be indicated as P_2, nominated as living standard factor.

Moreover, the results of the principal component analysis of economic level, social condition and education level can be attained

Table 1
Summary of the selection and classification of the 18 preliminary indicators.

Categories	Indicators	References	References
Economic level	F_{11}: GDP	F_{12}: per capita GDP	Wang and Zhao (2018)
	F_{13}: per capita disposable income		Wu et al. (2019)
	F_{14}: per capita consumption expenditure		Wang et al. (2019)
Living standard	F_{21}: Engel coefficient	F_{22}: ownership of air conditioner	Yuan et al. (2015)
	F_{23}: ownership of motorcycle	F_{24}: ownership of washing machine	Herrerias et al., 2017
	F_{25}: ownership of family car	F_{26}: ownership of refrigerator	Yang and Liu (2017)
	F_{27}: per capita housing space		Fan et al. (2015)
Social condition	F_{31}: permanent population	F_{32}: per capita length of roads	Miao (2017)
	F_{33}: whole social employees	F_{34}: expenditure on R&D	Zhang et al. (2015)
Education level	F_{41}: higher educated students rate		Liu et al. (2013)
	F_{42}: residential education expenditure	F_{43}: college acceptance rate	Golley and Meng (2012)

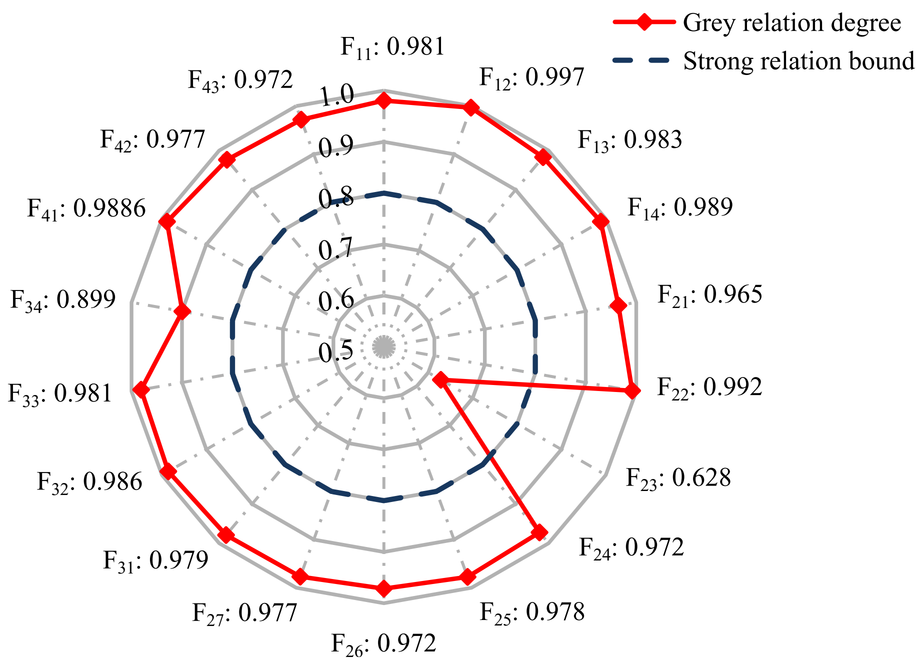

Fig. 8. Grey relational degree of 18 pre-selected indicators between CO_2.

Table 2
Principal component analysis of living standard.

Component	Total	% of variance	Cumulative %	Factor	Loading of component 1	Index	Value
1	6.255	89.355	89.355	F_{21}	−0.839	KMO	0.852
2	0.417	5.957	95.312	F_{22}	0.968	Bartlett's test	209.012
3	0.218	3.120	98.431	F_{23}	0.940	Prob.	0.000
4	0.053	0.764	99.195	F_{24}	0.978		
5	0.037	0.524	99.719	F_{25}	0.925		
6	0.014	0.206	99.925	F_{26}	0.966		
7	0.005	0.075	100	F_{27}	0.992		

from Tables 3–5. Accounting for 99.423%, P_1 is defined as economic level factor. And one principal factor taking up 89.762%, termed as $P_{3,}$ is extracted to represent the social condition. Similarly, P_4 is main element of the education level, whose accumulative contribution rate amounts to 84.479% with suitable KMO value.

Summing up, the principal component analysis is utilized to integrate the 17 preliminary elements into 4 principal components, containing economic level factor P_1, living standard factor P_2, social condition factor P_3, and education level factor P_4. As a consequence, the 4 factors will be introduced as the input variables for ICSO-SVM to forecast residential energy-related CO_2 emissions in Shanghai, which can not only involve comprehensive information of total influencing factors but also fulfill the abatement of data redundancy.

4. Results and discussion

4.1. Evaluation criterion of forecasting models

As for the forecasting accuracy analysis, mean absolute percentage error (MAPE) and root mean square error (RMSE) are introduced to verify models' performance. The equations of the two error test criterion are as follow:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \bar{y}_i}{y_i}\right| * 100\% \tag{22}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \bar{y}_i}{y_i}\right|^2} \tag{23}$$

Table 3
Principal component analysis of economic level.

Component	Total	% of variance	Cumulative %	Factor	Loading of component 1	Index	Value
1	3.977	99.423	99.423	F_{11}	0.998	KMO	0.687
2	0.015	0.385	99.809	F_{12}	0.997	Bartlett's test	204.696
3	0.007	0.169	99.977	F_{13}	0.997	Prob.	0.000
4	0.001	0.023	100	F_{14}	0.996		

Table 4
Principal component analysis of social condition.

Component	Total	% of variance	Cumulative %	Factor	Loading of component 1	Index	Value
1	3.590	89.762	89.762	F_{31}	0.982	KMO	0.749
2	0.338	8.439	98.201	F_{32}	0.874	Bartlett's test	91.365
3	0.049	1.232	99.433	F_{33}	0.970	Prob.	0.000
4	0.023	0.567	100.000	F_{34}	0.959		

Table 5
Principal component analysis of education level.

Component	Total	% of variance	Cumulative %	Factor	Loading of component 1	Index	Value
1	2.534	84.479	84.479	F_{41}	0.916	KMO	0.705
2	0.330	10.989	95.468	F_{42}	0.953	Bartlett's test	30.813
3	0.136	4.532	100.000	F_{43}	0.887	Prob.	0.000

where y_i means the actual CO_2 emissions in year i ($i = 1, 2, \cdots n$), and $\bar{y}_i$ is the corresponding predicted value.

4.2. Establishment of comparison models

To demonstrate the reasonableness of integrated influencing factors and ICSO-SVM's forecasting capability, the present study conducts the following comparison for simulation results, and the detailed steps is shown in Fig. 9. Firstly, with the 17 influencing factors and 4 principal components are chosen as the input data respectively, the validity of data dimension reduction can be proved through comparing simulation accuracy. Secondly, the ICSO-SVM is compared with the CSO-SVM to present the improvement of chicken swarm algorithm and with 3 other benchmark models, namely, PSO-SVM, GA-SVM and SVM, to verify the new model's higher forecasting accuracy.

4.3. Application of the ICSO-SVM model

As mentioned previously, ICSO is utilized to optimize the two parameter C and σ of SVM. To establish forecasting models, this paper uses historical data of residential energy-related CO_2 emissions in Shanghai from 2000 to 2012 as training data to find the best parameters of SVM, then data from 2013 to 2016 is testing data to verify ICSO-SVM's prediction performance. For the ICSO parameters initialization, the setting is presented in Table 6.

Depicted in Fig. 10, the fluctuations in average fitness of populations reveal the function of ICSO when training SVM model. What's obvious is that the fluctuation range gradually decreases with the increase of iterations, that is to say, the probability of generating chaotic sequences and the inertia weight w are tapering off to let algorithm search in a relatively small area. Particularly, it can be presented that the generation of chaos sequences does have

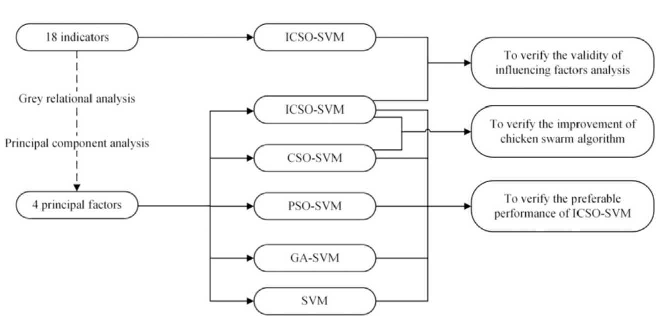

Fig. 9. The structure of comparative analysis.

Table 6
Parameters' setting of ICSO-SVM.

Parameters	Value	Parameters	Value
Maximum iterations	100	Hens percent	0.6
Population size	50	Mom hens percent	0.5
Update iterations	10	Upper bounds	100
Roosters percent	0.2	Lower bounds	0.01

the ability to increase the diversity of the population to avoid premature convergence. As results, around the 27th generation, the minimum fitness of training model is found to be 0.0441 with the best parameters searched by ICSO algorithm, while C is 77.9690 and σ is 0.1013, which will be utilized in the following testing model.

The integrated ICSO-SVM model and the other comparison models were executed by MATLAB 2016a, and the predicting results are shown as follows. First, in Table 7 and Fig. 11, the error comparison is conducted between the ICSO-SVM model with 18 indicators (18-ICSOSVM) and 4 principal factors (4-ICSO-SVM) as input data respectively. The MAPE and RMSE of 4-ICSO-SVM are 1.21% and 0.4346 separately, both lesser than those of 18-ICSO-SVM.

As illustrated in Fig. 12, the fitting curve from 2013 to 2016 of the ICSO-SVM surpasses other compared models in residential energy-related CO_2 emissions prediction. Simultaneously, Fig. 13 and Table 8 manifests the error test results. It is observed that the MAPE and RMSE value of ICSO-SVM are smaller than other models. Taken together, these results reveal that: (a) ICSO-SVM model possesses the highest forecasting accuracy with the lowest value of MAPE and RMSE in comparison with CSO-SVM, PSO-SVM, GA-SVM and SVM. (b) Based on grey relational analysis and principle component analysis, the ICSO-SVM model can highly improve the forecasting accuracy. (c) The chicken swarm algorithm does have improvement in forecasting model optimization with the introduce of chaotic mutation and learning weight.

Furthermore, based on the above research in estimation, influencing factor analysis and prediction of residential energy-related CO_2 emissions in Shanghai, we can provide some energy efficiency recommendation and policy orientation for policy makers regarding CO_2 emissions abatement.

Since 2000, Shanghai's residential energy shift to 'cleanliness' is evident that the consumption of secondary energy like electricity raises distinctly. Nevertheless, this trend in not conductive to the mitigation in CO_2 emissions, that is, the present 'clean energy' is not equal to low-carbon energy. The renewable energy (such as solar, wind, geothermal, etc) needs further development in power generation, which can decline the electric emissions coefficient accordingly and thus achieve the total carbon emission reduction.

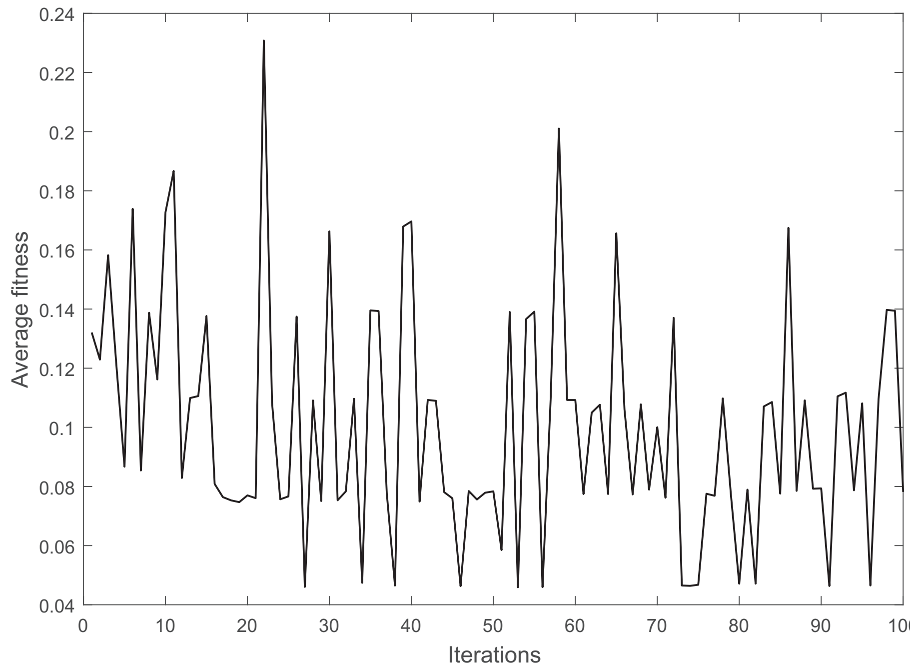

Fig. 10. The average fitness of populations in every iteration.

Table 7
Accuracy of the two compared models.

Error criterion	4-ICSO-SVM	18-CSO-SVM
MAPE (%)	1.21%	2.85%
RMSE (million tons)	0.4346	1.3872

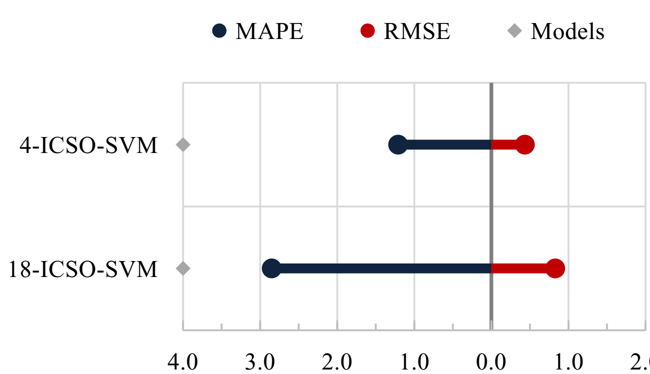

Fig. 11. Accuracy of the two models with different inputs.

With respect to the large consumption of petrol, the emerging new-energy vehicles should be promoted as the future direction, in the meanwhile, it is vital to implement the rational layout and construction of charging piles. The policy of exempting purchase tax and subsidizing for new-energy vehicles could not only conducive to the promotion of new energy vehicles, but also indirectly inhibit the residents' demand for fuel vehicles. With the increasing use of new energy vehicles, the CO_2 emissions intensity of it will gradually decrease (Herrerias et al., 2017; Xu and Lin, 2018).

Besides, enlarging investment to R&D is urgent, through technical measures can achieve energy saving in resident sector, whose potential in carbon abatement is huge. More efficient air conditioners, solar water heaters, electric cars and other should be promoted and subsidized. From Golley and Meng (2012), well-educated people possess more environmental awareness, showing be greener than average level. In some sense, they are more inclined to low-carbon consumption and more appealing to impulse people around them to save energy. Therefore, the guidance of carbon emission reduction for highly educated people should be strengthened, for instance, adopting favorable policy to advocate low-carbon mass transit so as to reduce their private car use.

Carbon market will transfer carbon cost to residents under the condition of electricity marketization while promoting carbon emission reduction in high energy consumption industries. Policymakers should focus on the sharing of carbon costs between energy-intensive industries and residential sectors to ensure the energy needs and quality of life of residents, especially low-income residents. As the blueprint of the carbon market, problems in the operation of EUETS serve as a warning to the design and arrangement of the carbon market, and the corresponding countermeasures also provide reference and experience for the improvement of the carbon market.

5. Conclusion

It's well known, as the largest CO_2 emissions country, China is accountable to develop the low-carbon economy. Nevertheless, the residential energy consumption is still gradually increasing in China, thus emitting more CO_2. Hence, researches on residential energy-related CO_2 emissions and its driving forces are vital to offer reference for the relevant agencies to formulate mitigation policy. In this study, more verified influencing factors are taken into consideration in comparison with previous studies in the field of residential CO_2 emissions, which will provide policy makers more breakthrough points to meet emissions abatement target.

Specifically, 18 preliminary indicators of residential energy-related CO_2 emissions are selected in this work, and grey relational analysis is applied to identify their correlation. Then the remaining factors that emerge strong relevance to CO_2 emissions in resident sector are introduced as influencing factors, synchronously, these indicators are classified into 4 categories according to

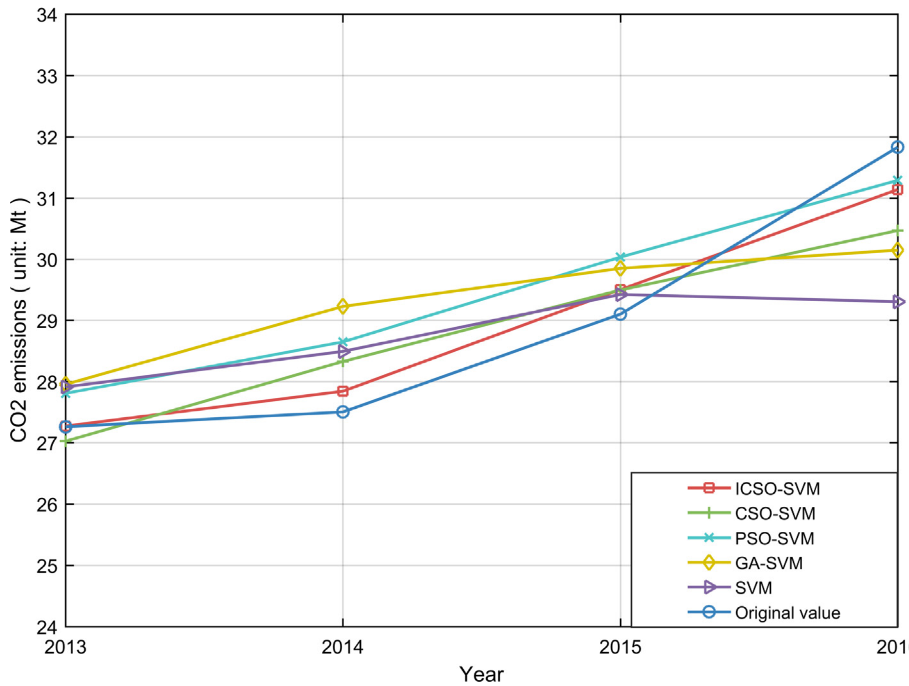

Fig. 12. Fitting curve of the three predictions and original value.

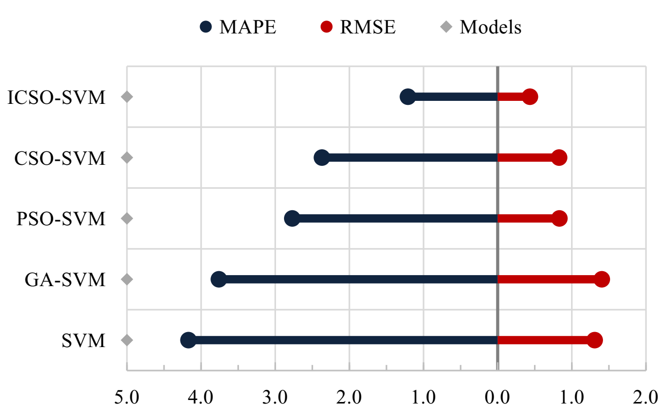

Fig. 13. Accuracy of different forecasting models.

Table 8
Accuracy of the compared models.

Error criterion	ICSO-SVM	CSO-SVM	PSO-SVM	GA-SVM	SVM
MAPE (%)	1.21%	2.37%	2.77%	3.76%	4.17%
RMSE (million tons)	0.4346	0.8282	0.8322	1.4040	1.3082

their inherent characteristics. Then principal component analysis is utilized to simplify these affecting factors. As a result, 4 components are extract from 4 categories respectively which are set as the input data of forecasting model in the following. Afterwards, this paper proposes a new hybrid ICSO-SVM model for the first time to predict CO_2 emissions. Through the testing case of Shanghai's residential sector and comparing with the forecasting results of different models, the conclusions of this article are drawn as follows:

(a) Considering local electricity and heat emissions coefficient, this study gets a more accurate measurement of residential energy-related CO_2 emissions in city level. (b) Analysis of influencing factors using grey relational analysis and principal component analysis can enhance the predicting accuracy availably. (c) The chicken swarm algorithm is improved in this work and has proved to be effective in the optimization of SVM. (d) In comparison with other methods, the newly established ICSO-SVM model is a practical and promising method in forecasting CO_2 emissions.

Although this study has developed the forecasting model in CO_2 emissions with impressive results, there still exists some issues needed further study. For instance, some potential influencing factors may have yet to be excavated. And lesser information is supposed to be lost during the dimensional reduction process. In addition, the prediction model combing more advanced intelligent algorithm is necessary to be investigated for more precise forecasting. In future research, these problems will hopefully be tackled with joint efforts.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmad, M.W., Reynolds, J., Rezgui, Y., 2018. Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees and regression trees. J. Clean. Prod. 203, 810–821. https://doi.org/10.1016/j.jclepro.2018.08.207.

Antanasijevic, D.Z., Ristic, M.D., Peric-Grujic, A.A., Pocajt, V.V., 2014. Forecasting GHG emissions using an optimized artificial neural network model based on correlation and principal component analysis. Int. J. Greenh. Gas. Con. 20, 244–253. https://doi.org/10.1016/j.ijggc.2013.11.011.

BP, 2019. Statistical review of world energy. https://www.bp.com/zh_cn/china/reports-and-publications/_bp_2019-_.html.

Ding, S., Dang, Y., Li, X., Wang, J., Zhao, K., 2017. Forecasting Chinese CO2 emissions from fuel combustion using a novel grey multivariable model. J. Clean. Prod. 162, 1527–1538. https://doi.org/10.1016/j.jclepro.2017.06.167.

Dong, B., Cao, C., Lee, S.E., 2005. Applying support vector machines to predict building energy consumption in tropical region. Energy Build. 37 (5), 545–553. https://doi.org/10.1016/j.enbuild.2004.09.009.

Fan, J., Yu, H., Wei, Y., 2015. Residential energy-related carbon emissions in urban and rural China during 1996 – 2012 : from the perspective of five end-use activities. Energy Build. 96, 201–209. https://doi.org/10.1016/j.enbuild.2015.03.026.

Fang, D., Zhang, X., Yu, Q., Chen, T., Tian, L., 2018. A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression. J. Clean. Prod. 173, 143–150. https://doi.org/10.1016/j.jclepro.2017.05.102.

Golley, J., Meng, X., 2012. Income inequality and carbon dioxide emissions: the case of Chinese urban households. Energy Econ. 34, 1864–1872. https://doi.org/10.1016/j.eneco.2012.07.025.

Han, Liyan, Xu, X., Han, Lu, 2015. Applying quantile regression and Shapley decomposition to analyzing the determinants of household embedded carbon emissions: evidence from urban China. J. Clean. Prod. 103, 219–230. https://doi.org/10.1016/j.jclepro.2014.08.078.

Herrerias, M.J., Aller, C., Ordóñez, J., 2017. Residential energy consumption: A convergence analysis across Chinese regions. Energy Econ. 62, 371–381. https://doi.org/10.1016/j.eneco.2016.06.006.

Hirano, Y., Ihara, T., Yoshida, Y., 2016. Estimating residential CO2 emissions based on daily activities and consideration of methods to reduce emissions. Build. Environ. 103, 1–8. https://doi.org/10.1016/j.buildenv.2016.02.021.

Hosseini, S.M., Saifoddin, A., Shirmohammadi, R., Aslani, A., 2019. Forecasting of CO2 emissions in Iran based on time series and regression analysis. Energy Rep. 5, 619–631. https://doi.org/10.1016/j.egyr.2019.05.004.

Huang, M., Wang, B., 2016. Factors influencing CO2 emissions in China based on grey relational analysis. Energy Sources Part A 38, 555–561. https://doi.org/10.1080/15567036.2013.802073.

IPCC (Intergovernmental Panel on Climate Change), 2007. Climate Change 2007: Mitigation: Contribution of Working Group III to the Fourth Assessment Report of the IPCC. Cambridge University Press, Cambridge, UK. https://www.ipcc.ch/report/ar4/wg3/.

IPCC, 2006. Guidelines for National Greenhouse Gas Inventories, vol. 2. https://www.ipcc-nggip.iges.or.jp/public/2006gl/chinese/vol2.html.

Li, J., Zhang, B., Shi, J., 2017. Combining a genetic algorithm and support vector machine to study the factors influencing CO2. Energies 10, 1520. https://doi.org/10.3390/en10101520.

Li, J., Zhang, D., Su, B., 2019. The impact of social awareness and lifestyles on household carbon emissions in China. Ecol. Econ. 160, 145–155. https://doi.org/10.1016/j.ecolecon.2019.02.020.

Li, L., Zhao, X., Tseng, M., Tan, R.R., 2020. Short-term wind power forecasting based on support vector machine with improved dragon fly algorithm. J. Clean. Prod. 242, 118447. https://doi.org/10.1016/j.jclepro.2019.118447.

Lin, B., Benjamin, N.I., 2019. Determinants of industrial carbon dioxide emissions growth in Shanghai : a quantile analysis. J. Clean. Prod. 217, 776–786. https://doi.org/10.1016/j.jclepro.2019.01.208.

Liu, W., Spaargaren, G., Heerink, N., Mol, A.P.J., Wang, C., 2013. Energy consumption practices of rural households in north China: basic characteristics and potential for low carbon development. Energy Policy 55, 128–138. https://doi.org/10.1016/j.enpol.2012.11.031.

Meng, X., Liu, Y., Gao, X., Zhang, H., 2014. A new bio-inspired algorithm: chicken swarm optimization. In: International Conference in Swarm Intelligence, pp. 86–94. https://doi.org/10.1007/978-3-319-11857-4_10.

Miao, L., 2017. Examining the impact factors of urban residential energy consumption and CO2 emissions in China – evidence from city-level data. Ecol. Indicat. 73, 29–37. https://doi.org/10.1016/j.ecolind.2016.09.031.

Nejat, P., Jomehzadeh, F., Mahdi, M., Gohari, M., 2015. A global review of energy consumption , CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries). Renew. Sustain. Energy Rev. 43, 843–862. https://doi.org/10.1016/j.rser.2014.11.066.

Nie, H., Kemp, R., 2014. Index decomposition analysis of residential energy consumption in China : 2002 – 2010. Appl. Energy 121, 10–19. https://doi.org/10.1016/j.apenergy.2014.01.070.

Qiao, W., Lu, H., Zhou, G., Azimi, M., Yang, Q., Tian, W., 2019. A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. J. Clean. Prod. 118612 https://doi.org/10.1016/j.jclepro.2019.118612.

Ricardo, D., Medrano-t, R.O., Denis, E., 2017. Route to chaos and some properties in the boundary crisis of a generalized logistic mapping. Physica A 486, 674–680. https://doi.org/10.1016/j.physa.2017.05.074.

Rodrigues, F.D., Behrens, P., Yuan, R., 2019. Driving forces of household carbon emissions in China : a spatial decomposition analysis. J. Clean. Prod. 233, 932–945. https://doi.org/10.1016/j.jclepro.2019.06.110.

Shen, L., Wu, Y., Lou, Y., Zeng, D., Shuai, C., Song, X., 2018. What drives the carbon emission in the Chinese cities ? – a case of pilot low carbon city of Beijing. J. Clean. Prod. 174, 343–354. https://doi.org/10.1016/j.jclepro.2017.10.333.

Song, J., Yang, W., Wang, S., Higano, Y., Fang, K., 2018. Exploring potential pathways towards fossil energy-related GHG emission peak prior to 2030 for China : an integrated input-output simulation model *. J. Clean. Prod. 178, 688–702. https://doi.org/10.1016/j.jclepro.2018.01.062.

Sun, W., Xu, Y., 2016. Using a back propagation neural network based on improved particle swarm optimization to study the influential factors of carbon dioxide emissions in Hebei Province , China. J. Clean. Prod. 112, 1282–1291. https://doi.org/10.1016/j.jclepro.2015.04.097.

Sun, W., Sun, J., 2017. Daily PM 2.5 concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm. J. Environ. Manag. 188, 144–152. https://doi.org/10.1016/j.jenvman.2016.12.011.

Sun, W., Wang, C., Zhang, C., 2017. Factor analysis and forecasting of CO2 emissions in Hebei, using extreme learning machine based on particle swarm optimization. J. Clean. Prod. 162, 1095–1101. https://doi.org/10.1016/j.jclepro.2017.06.016.

Wakiyama, T., Kuramochi, T., 2020. Scenario analysis of energy saving and CO2 emissions reduction potentials to ratchet up Japanese mitigation target in 2030 in the residential sector. Energy Policy 103, 1–15. https://doi.org/10.1016/j.enpol.2016.12.059.

Wang, C., Zhan, J., Li, Z., Zhang, F., Zhang, Y., 2019. Structural decomposition analysis of carbon emissions from residential consumption in the Beijing-Tianjin-Hebei region, China. J. Clean. Prod. 208, 1357–1364. https://doi.org/10.1016/j.jclepro.2018.09.257.

Wang, Y., Zhao, T., 2018. Panel estimation for the impacts of residential characteristic factors on CO2 emissions from residential sector in China. Atmos. Pollut. Res. 9, 595–606. https://doi.org/10.1016/j.apr.2017.12.010.

Wu, Y., Tam, V.W.Y., Shuai, C., Shen, L., Zhang, Y., Liao, S., 2019. Decoupling China's economic growth from carbon emissions: empirical studies from 30 Chinese provinces (2001–2015). Sci. Total Environ. 656, 576–588. https://doi.org/10.1016/j.scitotenv.2018.11.384.

Xu, B., Lin, B., 2018. Investigating the differences in CO_2 emissions in the transport sector across Chinese provinces: Evidence from a quantile regression model. J. Clean. Prod. 175, 109–122. https://doi.org/10.1016/j.jclepro.2017.12.022.

Yang, T., Liu, W., 2017. Inequality of household carbon emissions and its influencing factors: case study of urban China. Habitat Int. 70, 61–71. https://doi.org/10.1016/j.habitatint.2017.10.004.

Yao, X., Kou, D., Shao, S., Li, X., Wang, W., Zhang, C., 2018. Can urbanization process and carbon emission abatement be harmonious ? New evidence from China. Environ. Impact Assess. Rev. 71, 70–83. https://doi.org/10.1016/j.eiar.2018.04.005.

Ye, B., Jiang, J., Li, C., Miao, L., Tang, J., 2017. Quantification and driving force analysis of provincial-level carbon emissions in China Intergovernmental Panel on Climate Change. Appl. Energy 198, 223–238. https://doi.org/10.1016/j.apenergy.2017.04.063.

Yuan, B., Ren, S., Chen, X., 2015. The effects of urbanization, consumption ratio and consumption structure on residential indirect CO2 emissions in China: a regional comparative analysis. Appl. Energy 140, 94–106. https://doi.org/10.1016/j.apenergy.2014.11.047.

Zhang, X., Luo, L., Skitmore, M., 2015. Household carbon emission research : an analytical review of measurement, influencing factors and mitigation prospects. J. Clean. Prod. 103, 873–883. https://doi.org/10.1016/j.jclepro.2015.04.024.

Zhang, Y.J., Bian, X.J., Tan, W., Song, J., 2017. The indirect energy consumption and CO2 emission caused by household consumption in China: an analysis based on the input–output method. J. Clean. Prod. 163, 69–83. https://doi.org/10.1016/j.jclepro.2015.08.044.

Zhao, H., Huang, G., Yan, N., 2018. Forecasting energy-related CO2 emissions employing a novel SSA-LSSVM model : considering structural factors in China. Energies 1–21. https://doi.org/10.3390/en11040781.