Korean J. Chem. Eng., 39(7), 1729-1743 (2022)
DOI: 10.1007/s11814-022-1080-9

pISSN: 0256-1115
eISSN: 1975-7220

Analysis of heat transfer performance of ORC direct contact heat exchanger by GRA-VMD-LSSVM model using optimization

Guanfeng Zheng*, Qingtai Xiao,***, Shusheng Zhu****,*****, Hua Wang***,******,*******, Jian Geng*, Shuang Zhao*, and Junwei Huang*,†**

*Faculty of Mechanical and Electrical Engineering, Yunnan Agricultural University, Kunming, Yunnan 650201, China
**Faculty of Metallurgical and Energy Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650093, China
***State Key Laboratory of Complex Nonferrous Metal Resources Clean Utilization, Kunming University of Science and Technology, Kunming, Yunnan 650093, China
****State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming, 650201, China
*****Key Laboratory for Agro-biodiversity and Pest Control of Ministry of Education, Yunnan Agricultural University, Kunming, 650201, China
******Engineering Research Center of Metallurgical Energy Conservation and Emission Reduction, Ministry of Education, Kunming, Yunnan 650093, China
*******National & Local Joint Engineering Research Center of Energy Saving and Environmental Protection Technology in Metallurgy and Chemical Engineering Industry, Kunming, Yunnan 650093, China
(*Received 30 October 2021 • Revised 15 January 2022 • Accepted 27 January 2022*)

Abstract−The relationship between heat transfer effect and influencing factors in direct contact with evaporator is studied. It is important to optimize the design and setting of direct contact heat exchangers and further improve the heat transfer process of heat exchangers by understanding the influencing factors of direct contact evaporator. The methods used are grey correlation analysis, variational mode decomposition and least square support vector machine algorithm to calculate the experimental heat transfer coefficient. The conclusions are that grey relation analysis can find the complex relationship between different influencing factors and reduce the amount of data to improve the accuracy of prediction. Hence, the novelty of this paper is to propose a simple, efficient and accurate hybrid prediction model for VHTC prediction. The place where it goes beyond previous efforts in the literature is that different influencing factors on the heat transfer performance in the direct contact heat transfer process are considered. The results show that the prediction accuracy of the heat transfer coefficient can be improved by 7% by optimizing the data only with grey correlation analysis. The prediction accuracy can be improved up to 53% after using the hybrid model.

Keywords: Direct Contact Heat Exchanger, Volumetric Heat Transfer Coefficient, Grey Relation Analysis, Variational Mode Decomposition, Support Vector Machine

INTRODUCTION

With the demand and consumption of energy increasing day by day [1], energy conservation has become an important subject for study [2]. Metallurgical engineering is the main industry of energy consumption in China, but the actual average utilization rate is lower than 40%, of which the waste heat generated and wasted in industrial production accounts for 20-50% [3-5]. Therefore, an efficient heat transfer technology is needed to recycle and utilize low-grade waste heat resources, which can better save energy, improve energy utilization efficiency, reduce pollution and better achieve zero emissions of carbon dioxide [6,7]. The organic Rankine cycle (ORC) is an efficient technique for low temperature waste heat recovery [8]. Minea et al. investigated the feasibility of using ORC systems in waste heat power generation [9]. They showed that using ORC systems is efficient and stable. Nami et al. compared the efficiencies of ORC and Rankine in combined cooling, heating and power systems [10]. The results show that higher energy efficiency, exergy efficiency and environmental friendliness can be obtained when ORC is used. Hou et al. showed that ORC system can save energy consumption and effectively reduce carbon dioxide emissions when it is applied to industrial waste heat power generation [11]. In fact, a large amount of waste heat resources is generated from industrial production. When the heat energy in the waste heat resources is utilized, the energy utilization efficiency is greatly improved [12,13].

To improve energy efficiency and reduce heat waste, several investigations on ORC systems have been carried out [14]. The selection, characteristic parameters and influencing factors of the fluid in the ORC system have been considered, and the heat transfer performance has been analyzed. The performance of the traditional heat

†To whom correspondence should be addressed.
E-mail: oldin230@163.com

exchanger is mainly determined by the different heat transfer medium in the middle. Therefore, the direct contact heat exchanger (DCHE) has been proposed to improve the heat transfer effect of ORC and reduce the heat loss from the heat transfer to the intermediate medium [15]. Halkarn et al. evaluated and predicted the heat transfer capacity by changing the influencing factors affecting the experiment and calculating the volumetric heat transfer coefficient (VHTC) [16,17]. Continuous experiments are used to obtain the conditions which influence the best heat transfer results. A large number of data have been obtained in the experiment, which requires further processing and analysis of the obtained data. Therefore, the data is processed and analyzed to better predict VHTC and improve accuracy [18]. The heat transfer performance is determined by the different experimental conditions set in the experimental process. The influencing factors of the experimental data and VHTC data are obtained through the experiment. The prediction accuracy of VHTC can be improved by analyzing and processing the relationship between influencing factors and VHTC data. However, it is difficult to determine the impact of multiple factors on the heat transfer coefficient. The relationship between the factors and the heat transfer coefficient is complicated and difficult to explain briefly. Fortunately, machine learning is very good at dealing with the correlation between different factors in large amounts of data.

Grey relation analysis (GRA) is a kind of grey system that can analyze the influence degree and correlation degree of multiple variables on dependent variables. The correlation degree of each variable is compared, sorted and evaluated. The main factors affecting the dependent variables are determined and their trends and changes can be inferred [19]. The shortcomings of analysis methods, such as analysis of variance and range analysis, are made up and grey relation analysis is often used in objective optimization. Das et al. illustrated that the grey relational degree is used to sort the parameters to find the optimal parameter combination to obtain better surface roughness and microhardness [20]. Ren et al. carried out grey relation analysis on the changes of the five parameters of coal, and obtained the main factors affecting the thermal physical property parameters of coal through calculation [21]. Hong et al. proposed the method of combining grey relational degree and RBF neural network to predict trihalomethanes, and found that using grey relation analysis could generate a good neural network model and improve the prediction accuracy [22]. Therefore, the complex relationship between VHTC data and VHTC data influencing factors can be well dealt with by grey correlation analysis. The main factors affecting VHTC can be found by grey correlation analysis. The data quality can be improved by using grey relational degree to optimize the data. However, the process is fluctuating, nonlinear and there are many influencing factors in the direct contact heat transfer process. It is difficult to analyze the non-stationary characteristics of VHTC data and other chaotic signals.

In fact, variational mode decomposition (VMD) is applied in the further processing of VHTC data in this paper. It is a signal processing method for non-linear, non-stationary data. The original signal is decomposed by VMD into different numbers of intrinsic mode functions (IMFs), which are more stable and reliable. In recent years, VMD has become a research hotspot, the optimal center frequency and the limited bandwidth are adjusted adaptively by the number of VMD decomposition [23]. VMD subdivides the signal in frequency domain and separates the components, so that the modal mixing problem of empirical mode decomposition (EMD) is overcome. Zhang et al. proposed to improve the wind speed prediction model by VMD, and found that the prediction accuracy could be improved and the error would be smaller after the wind signal information was fully mined [24]. Xiao et al. proposed that feature signals could be extracted by VMD method, which can shorten the execution time and improve the average accuracy of feature selection [25]. Gyamerah et al. investigated predictive models for price sequences of non-stationary and non-stable bitcoin [26]. It is found that VMD has less prediction error and higher accuracy than EMD through analysis and comparison. Therefore, VHTC data can be decomposed by VMD, which can effectively separate and divide the inherent modal components. The decomposed modal components the VHTC signal more detailed and more reliable. VMD has a solid theoretical basis, which can transform strong non-smooth and nonlinear data into relatively stable subsequences. VMD can be better applied to the prediction model and improve the prediction ability of prediction model. For data processing, VMD has stronger robustness than EMD, which makes the decomposed data components more stable, reduces the influence of different trend signals on the prediction accuracy, and reduces the prediction error.

Support vector machine (SVM) is an intelligent algorithm that is often used in classification, regression and data analysis. SVM works mainly by constructing hyperplanes [27]. Yu et al. found that SVM can be used for multi-component quantitative analysis with good robustness and efficiency [28]. By applying the prediction model of waste generation rates based on SVM, Hu et al. could obtain higher estimation accuracy than when not in use [29]. Amber et al. compared five kinds of intelligent algorithms for building electricity consumption and found that using SVM model within a certain range can improve the prediction accuracy and make the error smaller [30]. A variety of factors affecting VHTC data can be analyzed by SVM. SVM has strong generalization ability and good convergence effect. However, SVM has disadvantages such as long time consuming and the need to choose appropriate kernel function, slow convergence speed and so on. To make the convergence speed faster and the operation more simple and convenient, SVM needs to be further developed [31]. In this paper, the least square support vector machine (LSSVM) is an evolution of SVM, which makes up for the disadvantage of long time consumption of SVM, improves the computing efficiency, convergence speed and comes with radial basis function (RBF) kernel function. Sun et al. showed that higher fitting degree, lower prediction error value and higher prediction accuracy can be obtained by the reference of LSSVM model [32]. Singh et al. demonstrated that greater effectiveness and accuracy can be achieved in modeling and predicting confirmed cases using LSSVM models, which can help countries to prepare and take necessary measures [33]. Han et al. found that the accuracy rate of fault diagnosis was improved because the nonlinear parameters of the chiller were processed by the LSSVM model [34]. Compared with the SVM model, the prediction time of the LSSVM model is shortened by more than half. These results encourage us to apply LSSVM to the prediction of VHTC in direct contact heat

exchangers.

The originality of the paper is that a GRA-VMD-LSSVM model based on data optimization is proposed for investigating the direct contact heat transfer system in this work. The sufficient results to justify the novelty of a high-quality journal paper include the illustration of experimental heat transfer data (i.e., VHTC data set with large sample size) and the combination and optimization of intelligent algorithms. In the open literature, to our best knowledge, the simplified relationship models between boiling flow and heat transfer were obtained with low precision in special cases. However, what is missing is that the general relationship model should be proposed to describe clearly the inter-dependent relationship with high precision. It has not been seen in the literature that the grey relational degree is used to analyze the main influencing factors in the direct contact heat exchanger. The exact, general, and quantitative relationship model between flow and heat transfer in ORC direct contact heat transfer system needs to be proposed with the help of the advanced signal processing technology and machine learning algorithm. The advantages of the methods used in this work are high-precision, fast computation speed, strong adaptability and ability to deal with nonlinear characteristics.

In this paper, the implementation is as follows. First, GRA is used to analyze the factors affecting of the volume heat transfer coefficient, and the main factors and redundant factors are found. VMD method is used to decompose the experimental data. Then, the data obtained by the two algorithms are introduced into the LSSVM model, and the LSSVM model is used to combine the data together to process the VHTC data. In addition, through the comparison of different evaluation indexes, the accuracy of the prediction model is illustrated. It is necessary for it to deal with the redundant data in big data, improve the accuracy of prediction model, and provide guidance for practical work.

For rest of this article Section 2 shows the establishment of an experimental platform for ORC direct contact heat exchanger and the acquisition of VHTC data. Section 3 provides detailed algorithms for grey relation analysis, VMD and LSSVM, and introduces the construction of a new hybrid model. The prediction ability of the optimized and unoptimized data and the prediction effect of the new hybrid model and the single model are compared in Section 4. The summary and conclusions are given in Section 5.

EXPERIMENTAL

1. Heat Transfer Apparatus

Fig. 1 shows the ORC direct contact heat exchanger test platform. The test platform is mainly composed of two circulation loops. One is the continuous phase fluid flow circulation loop in the experiment, which is mainly composed of direct contact evaporator, gear oil pump and electric heater. The other is the flow loop of the dispersed phase fluid in the experiment, which is mainly composed of direct contact evaporator, condenser and liquid storage tank. Among them, the choice of working fluid is a very important consideration. The materials used in this experiment have good environmental protection performance, no toxic substances, and have a high economy. Synthetic heat conduction oil and R245fa are two excellent materials which are widely used in industry at present. Therefore, the synthetic heat conduction oil is used as a continuous phase fluid and R245fa as a dispersed phase fluid in this experiment. R245fa is a liquid hydrocarbon fluoride, safe and reliable, colorless and transparent, whose specific gravity is greater than water and it has a lower boiling point and high steam thermal conductivity.

The heater is connected with the temperature control device, which can control the temperature and improve the safety. Working medium centrifugal pump and gear pump are installed with frequency conversion governor; they can change the flow rate and velocity in order to do a better experiment. In the process of the experiment, first, the heat transfer oil is added to the device that directly contacts the evaporator through the heat transfer oil pump, while the working medium is injected into the liquid storage tank. Then the valve of the distributed phase working medium flow circulation loop is closed, the valve of the continuous phase fluid flow

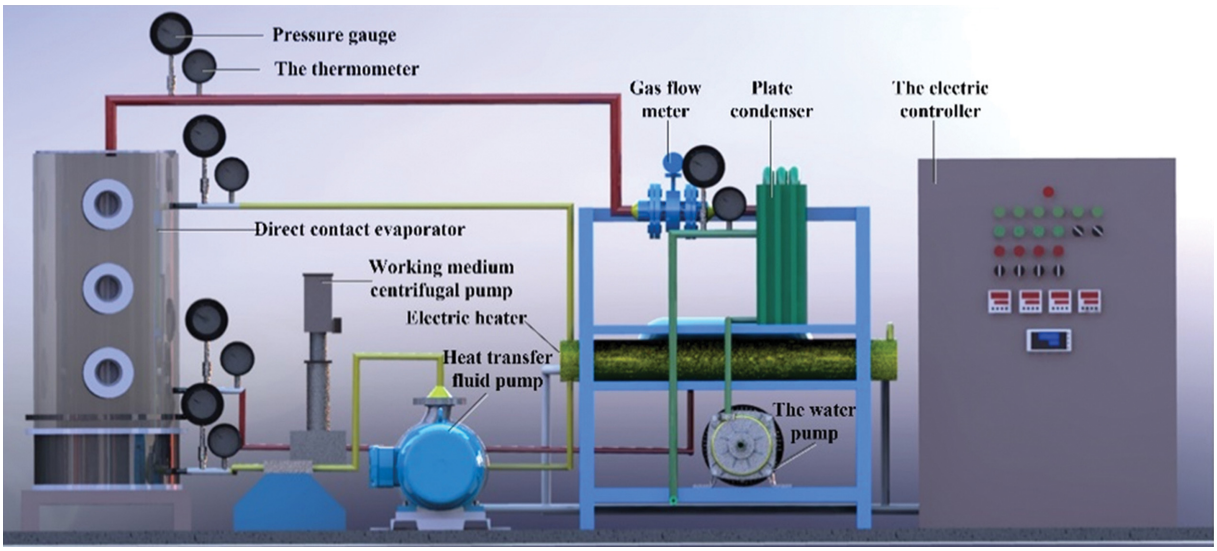

Fig. 1. Experimental setup of the ORC direct contact heat transfer system.

Table 1. Performance indicators of the main measuring instruments

Physical quantity side points	Name of instrument	Range	Accuracy
Temperature	Thermometer	0-150 °C	±0.4%
Pressure	Pressure gage	0-1 MPa for high pressure 0-0.6 MPa for low pressure	±0.4%
Mass flow of liquid	Intelligent electromagnetic flowmeter	0-250 kg/h	0.5%
Gas mass flow rate	Gas mass flow meter	10-200 kg/h	0.2%

circulation loop is opened, and the electric heater is run by the experimental operators. The heat transfer oil is heated to the experimental conditions, and then the valve of the circulating loop of the dispersed phase working medium is opened to run the centrifugal pump of the working medium. Finally, the working medium is injected into the device that directly contacts the evaporator, and the working medium and the heat transfer oil are directly contacted for heat exchange. The performance indexes of the test instrument are shown in Table 1. Affected by environmental change, measurement error, human error and other factors, the measured experimental data is inevitably uncertain, but the error range of the experiment is within the specified range. Therefore, the results obtained in this experiment are reliable and credible. When the working medium centrifugal pump is opened, it is necessary to start recording the data. Data are recorded every two minutes for each temperature, pressure and flow meter.

2. Experiment Data Obtained

Direct contact heat transfer is a process of heat transfer between gas and liquid or between liquid and liquid. Without traditional metal heat transfer surface, it is difficult to calculate the heat transfer effect of the device. Therefore, it is necessary to calculate and measure the heat transfer effect of direct contact heat exchanger through the volume heat transfer coefficient U_V. The formula is as follows [35]:

$$U_V = \frac{Q}{V \times \Delta T} \tag{1}$$

From Eq. (1), it can be seen that the volumetric heat transfer coefficient can reflect the compactness and heat transfer performance of the direct contact evaporator. The symbol Q is defined as follows [36]:

$$Q = q_m(h_{do} - h_{di}) \tag{2}$$

In the formula, the continuous phase in the experimental process is represented by subscript c; the dispersed phase in the experimental process is represented by subscript d; i is the inlet of continuous phase or dispersed phase; o is the outlet of the continuous phase or the dispersed phase. Where q_m is the mass flow rate of steam in the dispersed phase measured by the mass flow meter, kg/s. h is the total energy of the dispersed phase. The symbol ΔT is given as follows [37]:

$$\Delta T = \frac{(T_{ci} - T_{do}) - (T_{co} - T_{di})}{\ln \frac{(T_{ci} - T_{do})}{(T_{co} - T_{di})}} \tag{3}$$

where W represents the volume of the continuous phase fluid in the device, m^3; Q is the heat transfer between the working medium and the heat transfer oil in unit time, kJ/s; ΔT is the average temperature difference, °C. h_{do} is the enthalpy of the working medium flowing to the outlet, kJ/kg; h_{di} is the enthalpy of the working medium flowing from the inlet to the evaporator in direct contact, kJ/kg; T_{ci} is the temperature of the heat transfer oil when it passes through the entrance after being heated by the electric heater, °C; T_{co} is the outlet temperature of heat transfer oil, °C; T_{di} is the temperature when the working medium enters into direct contact evaporator, °C; T_{do} is the temperature at the outlet after the working medium changed into steam, °C.

The calculation of heat transfer coefficient is complicated and difficult because the traditional heat exchanger has no metal heat exchange surface. Therefore, the volumetric heat transfer coefficient is used to evaluate the heat transfer performance of the direct contact evaporator. Eq. (1), Eq. (2) and Eq. (3) can be used to calculate the volume heat transfer coefficient. In fact, there are eight influencing factors in the process of this experiment: height of continuous phase (Z), flow rate of dispersed phase (U_d), flow rate of continuous phase (U_c), ΔT, inlet temperature of continuous fluid (T_{ci}), mass flow rate of dispersed phase steam (m_g), heat transfer quantity (Q), volume of continuous phase (V). The relationship between fluid and heat transfer is very complicated. There is no specific formula that can be used all the time. In this paper, the factors affecting the heat transfer coefficient and experimental VHTC data were input into the grey relation analysis model for screening. The main factors were found and the weak factors were removed. The optimized data were input into the LSSVM model for VHTC prediction. From this point of view, the influence of main factors (i.e., summary temperature of continuous fluid and heat transfer quantity) and secondary factors (i.e., flow rate of continuous phase) on heat transfer performance was studied.

Data such as Q, V and ΔT were evaluated through experiments. Then VHTC data were calculated from these data. The data were processed and analyzed by the grey relational degree analysis model, and the processed data were provided to the new hybrid prediction model. Data optimization can reduce data volume and workload, improve computing efficiency, and reduce the complexity of data and interference from useless data. The average VHTC simplifies the information and lacks the influence of various conditions on the VHTC, which will reduce the prediction accuracy of the model. The instantaneous VHTC can evaluate the heat transfer capacity of the direct contact evaporator in real time according to the change of different influence factors, which will increase the relevant information. Combining the influencing factors with IMFs can better establish the prediction model and improve the predic-

tion accuracy of the prediction model. Therefore, the instantaneous VHTC can be used to better monitor and research the heat exchange effect of direct contact evaporator. Then the data were recorded and prepared for subsequent experiments. Through this experiment, a large amount of data was obtained. The main factors and secondary factors can be distinguished first, so as to reduce useless information and improve the weight of the main factors. The processed data was used to train and predict VHTC. The future experimental trend can be obtained. Therefore, this research can provide theoretical guidance for future work and experiments.

PROPOSED DATA OPTIMIZATION AND HYBRID SKILL

1. GRA, VMD and LSSVM

Grey relation analysis is an analysis of a constantly developing and changing system. It is mainly through the comparison of sequence to judge the geometric proximity between the factors. The correlation degree is obtained through analysis to evaluate the degree of connection between a certain thing and its related influencing factors. The greater the correlation between the influencing factors and the object, the closer the relationship is. It shows that the degree of influence of an influencing factor on the object is greater. Thus, the influence factor is the main influence factor affecting the development of the thing.

First, the VHTC data are determined as the reference sequence (x_0 sequence, i.e., the mother sequence) and the eight influencing factors of the VHTC data as the comparison sequence (x_i sequence, i.e., the sub-sequence). Because the data of various factors in the system may have different dimensions, it will be difficult to compare or to draw correct conclusions when comparing. Therefore, dimensionless processing is required. The dimensionless is defined as follows [22]:

$$y_0(k)' = \frac{y_0(s)}{y_0(1)} y_i(k)' = \frac{y_i(s)}{y_i(1)} \tag{4}$$

Then, the absolute difference between mother sequence and subsequence is calculated, as shown in Eq. (5).

$$\Delta_{oi}(s) = |y_0(s) - y_i(s)|,\ i=1, 2, \cdots, n;\ k=1, 2, \cdots, m \tag{5}$$

The correlation coefficient between the parent sequence and the subsequence is further calculated, where ρ is the resolution coefficient, as shown in Eq. (6). The smaller ρ is, the stronger the resolution is. The value of ρ is 0 - 1, usually ρ=0.5, which can get a good results.

$$\varepsilon_{oi}(t) = \frac{\Delta_{min} + \rho\Delta_{max}}{\Delta_{oi}(s) + \rho\Delta_{max}} \tag{6}$$

Finally, the final correlation degree is calculated by Eq. (7). The higher the correlation degree is, the closer the influencing factors are to VHTC. Therefore, the correlation coefficient can be used to find the main factors affecting the VHTC data, and then the hybrid VMD-LSSVM prediction model can be established.

$$\gamma_{oi} = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_{oi}(t) \tag{7}$$

where $y_0(s)$ is the new dimensionless mother sequence; $y_i(s)$ is a new dimensionless subsequence; $\Delta_{oi}(s)$ is the difference between the mother sequence and the sub-sequence; $\varepsilon_{oi}(t)$ is the correlation coefficient; γ_{oi} is the final correlation degree of the comparison sequence with the reference sequence.

VMD is an adaptive signal processing method based on EMD. It can find the optimal center frequency and limited bandwidth through exploration. The VMD method uses the non-recursive and variational mode solution mode to process the original signal, which has good anti-noise ability and non-stationary signal processing effect. The original signal frequency domain is subdivided and the components are separated through the VMD. The original signal is decomposed into k components to ensure the minimum sum of the estimated bandwidth of each mode, so that the sum of all the decomposed modes is equal to the original signal. The original signal is decomposed into k IMFs. To ensure the modal components with limited bandwidth and center frequency and minimize the sum of estimated bandwidth of each mode, the constraint condition is that the sum of all modal components is equal to the original signal. The specific variational constraint expression is as follows [38]:

$$\min_{\{u_M\},\{\omega_M\}} \left\{ \sum_M \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_M(t) \right] e^{-j w_M t} \right\|_2^2 \right\} \tag{8}$$

$$\text{s.t.} \sum_{M=1}^{k} u_M = f \tag{9}$$

where k is the number of modes that VHTC needs to decompose; $\{u_M\}$, $\{\omega_M\}$ are the Mth modal component and center frequency after VHTC decomposition respectively; $\delta(t)$ is Dirac function; $*$ is the convolution operator.

IMFs decomposed by VMD does not represent other basis functions, but the information and properties of VHTC. These details are relatively complex and detailed. Therefore, the LSSVM algorithm with great potential is introduced to solve the problem of complex and nonlinear data. Based on the principle of structure minimization, nonlinear problems and high dimensional pattern recognition problems are solved by LSSVM. Problems such as low data prediction efficiency and long training time can also be solved by LSSVM. The LSSVM algorithm comes with RBF kernel function, which can carry out nonlinear mapping, and has fewer parameters to reduce the complexity of the model. The regression expression of the LSSVM model is as follows:

$$f_x = w^T \varphi(x_k) + b \tag{10}$$

where w is the weight component of the eigenspace; b is the bias constant; $\varphi(x_k)$ is the RBF kernel function of LSSVM.

As the details of VHTC, IMFs can make VHTC be well analyzed [39]. IMFs are more stable, more symmetrical signals with waveforms that reflect the underlying vibrational characteristics of the data itself and closely surround the corresponding central frequency. To make the prediction accuracy of VHTC more accurate, more complex information needs to be transformed into simple, detailed information inputted into the LSSVM model. The original signal of VHTC is composed of eight influencing factors. If any of the eight factors change, the original signal will change. The

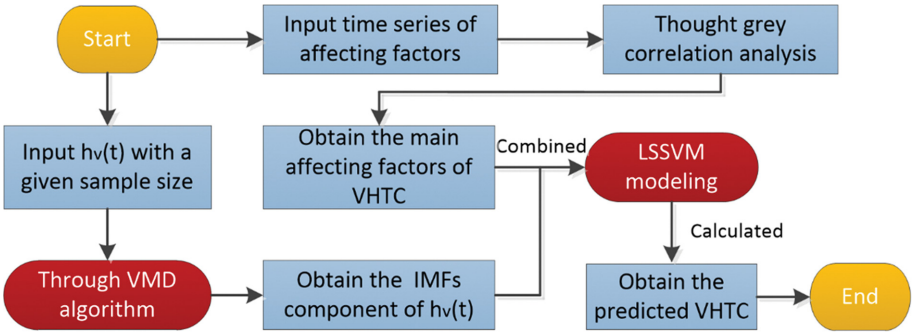

Fig. 2. Flowchart of the hybrid model combining GRA, EMD and LSSVM.

LSSVM model is a new machine learning algorithm to deal with nonlinear and complex structures. It has strong generalization ability. For a limited number of samples, the solution complexity and computation time can be significantly reduced by LSSVM. It is a simple and flexible prediction model. For modeling, prediction accuracy can be further improved by processing the data in the input LSSVM. The stability and accuracy of the prediction model can be improved when major variables are selected or unimportant variables are removed. In this way, relevant variables with low disturbance, low noise and many common points can be obtained.

2. Proposed GRA-VMD-LSSVM Model

This new hybrid prediction model is proposed in this paper, which consists of three parts. In fact, it is a hybrid solution method composed of three intelligent algorithms: GRA,VMD and LSSVM. The necessity of this method is to propose a new hybrid prediction model to strengthen the prediction of VHTC. The gap is a hybrid model that can make the experimental results more suitable for the prediction of VHTC. VHTC time series is obtained by calculating the experimental data obtained from the experiment. By using the obtained VHTC data, the research is carried out.

As shown in Fig. 2, the specific operation steps are composed of two parts. (1) The influencing factors of VHTC data are analyzed through GRA, and the main factors affecting VHTC are obtained. GRA can analyze sample data more simply, reduce the complex relationship between samples to a certain extent, find the main factors, remove the interference of non-main factors on time series, and improve data quality. At the same time, the VHTC time series is input into the VMD intelligent algorithm, and the IMFs can be obtained by VMD. The obtained IMFs have the trend characteristics of time series, which makes the information richer, more detailed, and more stable. (2) The main factor data is combined with IMFs and imported into the LSSVM model. Through regression calculation, the prediction of VHTC is obtained. These main factors and IMFS are taken as the detailed information of influencing VHTC data, and input into the LSSVM model, so as to provide more detailed data information for the LSSVM model and make the prediction of VHTC more accurate. The feasibility of the experimental method can be obtained by comparing the real value obtained from the real experiment with the predicted value obtained from the hybrid prediction model. Therefore, this method can be used to analyze the implicit complex relationship between VHTC and the influencing factors affecting VHTC in the direct contact heat transfer process. Theoretical guidance for experiments can be provided by identifying the main influencing factors. The source of this idea is to combine the advantages of GRA in dealing with the correlation between the complex influencing factors of VHTC, the advantages of adaptive data decomposition of VMD and the advantages of regression prediction of detailed information of LSSVM. In this way, data can be optimized, interference information and redundant information can be reduced, data processing and information reinforcement can be strengthened, prediction ability can be stronger and prediction results can be more reliable.

The following three problems are analyzed in this paper. (1) The number of influencing factors is analyzed. The number of different influencing factors will affect the prediction ability. By changing the number of different influencing factors, the influence on the prediction of VHTC can be found and the optimal number of influencing factors can be found to improve the prediction accuracy of the model. (2) By analyzing the different number of IMFs decomposed by VMD, the influence of the prediction model on the prediction of VHTC can be found by changing the number of IMFs generated by self-adaptation, so as to improve the prediction ability of the model. (3) By analyzing the proportion of training set, the prediction ability of VHTC prediction model is changed with different proportion of training set, so that the appropriate proportion of training set can be found and the reliability of prediction model can be improved.

When the prediction results of the test are obtained, it is necessary to compare and analyze the prediction models under each condition to understand the differences between the models under each different situation. The hybrid prediction model of GRA-VMD-LSSVM is established, and then the prediction values and training data obtained by the hybrid mode are processed. The evaluation indexes of the prediction results are obtained, such as mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE). The relevant equations are given as follows [40]:

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i-y_i\right| \tag{11}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2} \tag{12}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (13)$$

where y_i represents the real value obtained by the real experiment, and $\hat{y}_i$ represents the predicted value obtained by the prediction model. MAE, MAPE, and RMSE need to be understood in order to compare the results. The values of the three evaluation indexes range from 0 to positive infinity, but the greater the value is not the better. On the contrary, the closer the values of the three indicators are to 0, the smaller the error will be, and the closer they are to the real value, so the prediction accuracy is higher and more reliable. Pearson correlation coefficient (P), Spearman correlation coefficient (S) and Kendall correlation coefficient (K) are also introduced to evaluate the correlation of the prediction model. The formula is as follows [41]:

$$P = \frac{cov(x, y)}{\sqrt{cov(x, x)} \times cov(y, y)} \quad (14)$$

$$S = 1 - \frac{6\Sigma_i^n d_i^2}{n(n^2 - 1)} \quad (15)$$

$$K = \frac{n_c - n_d}{0.5 \times n \times (n-1)} \quad (16)$$

where x, y are the two variables for the correlation comparison, cov(x, y)is the total error of the two variables, d_i is the difference of rank, n_c is the number of the harmonious pairs, n_d is the number of the discordant pairs, and n is the size of the total data volume for correlation comparison. Among them, the size of the correlation coefficient is defined between −1 and 1. When its absolute value is approximately close to 1, it indicates that the model is more suitable. The prediction accuracy of the prediction model can be improved by using the appropriate prediction model. In this way, the correlation coefficient and error can be combined to evaluate the hybrid prediction model more accurately and reliably, making the prediction results of VHTC more accurate.

RESULTS AND DISCUSSION

1. Comparison of Data Optimizations

First, GRA, VMD and LSSVM are used to process and model the data. GRA is a quantitative description and comparison method for the development and change of the system to reflect the degree of correlation between variables. The eight influencing factors affecting VHTC are ranked, as shown in Table 2. By removing redundancy and interference data, the data can be optimized, the data quality can be improved and the prediction accuracy can be fur-

Table 2. Calculation and ranking of GRA on eight influencing factors of heat transfer coefficient

Influence factor	Grey relation analysis	Rank
Including height of continuous phase	0.824411	3
Flow rate of dispersed phase	0.736413	7
Flow rate of continuous phase	0.630079	8
LMTD	0.751438	6
Inlet temperature of continuous fluid	0.836004	2
Mass flow rate of Irradiation phase steam	0.817776	5
Heat transfer quantity	0.857505	1
Volume of continuous phase	0.824411	4

Table 3. Part of the experimental data grey correlation degree correlation coefficient data set

Z	U_d	U_c	LMTD	T_{ci}	m_g	Q_V	V
0.830733	0.774657	0.815221	0.515088	0.562222	0.558663	0.532531	0.483024
0.830733	0.774657	0.992199	0.917118	0.938687	0.930678	0.857277	0.733647
0.830733	0.774657	0.983710	0.971987	0.891159	0.951113	0.896229	0.756530
0.830733	0.774657	0.906330	0.960092	0.752241	0.904441	0.895284	0.780687
0.830733	0.774657	0.884552	0.895401	0.972400	0.907498	0.907308	0.853820
0.830733	0.774657	0.830734	0.950673	0.939838	0.949954	0.989457	0.846912
0.830733	0.774657	0.815221	0.515088	0.562222	0.558663	0.532531	0.483024
0.830733	0.774657	0.992199	0.917118	0.938687	0.930678	0.857277	0.733647
0.830733	0.774657	0.983710	0.971987	0.891159	0.951113	0.896229	0.756530
0.830733	0.774657	0.906330	0.960092	0.752241	0.904441	0.895284	0.780687
0.830733	0.774657	0.884552	0.895401	0.972400	0.907498	0.907308	0.853820
0.830733	0.774657	0.830734	0.950673	0.939838	0.949954	0.989457	0.846912
0.830733	0.774657	0.815221	0.515088	0.562222	0.558663	0.532531	0.483024
0.830733	0.774657	0.992199	0.917118	0.938687	0.930678	0.857277	0.733647
0.830733	0.774657	0.983710	0.971987	0.891159	0.951113	0.896229	0.756530

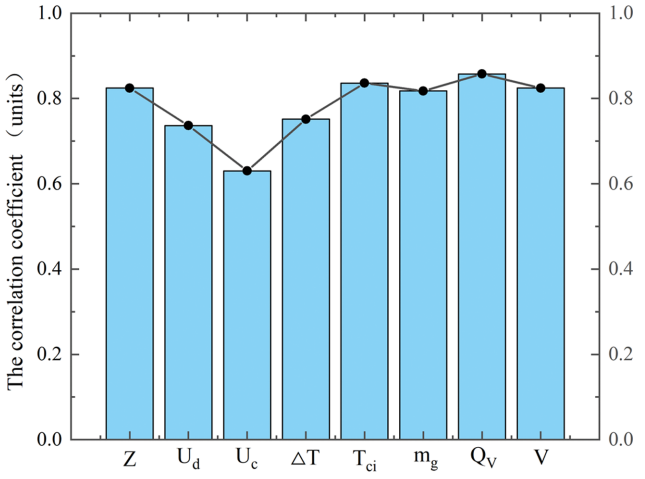

Fig. 3. The correlation coefficient fluctuation of eight influencing factors affecting VHTC is analyzed by GRA.

ther improved. It is found that the correlation coefficient between the including height of continuous phase and the volume of continuous phase is the same, and the flow rate of continuous phase is the lowest. This is consistent with the actual situation of calculating VHTC, because the experimental conditions represented by the volume of continuous phase and the including height of continuous phase are both the size of the direct contact heat exchanger. Therefore, GRA can be used to conduct a good correlation analysis of the influencing factors and find which factors are the main factors affecting VHTC data. As shown in Table 3, the correlation coefficient of each data under the condition of each influencing factor was calculated and analyzed. It is found that the correlation coefficient of most of the data is above 0.7, which indicates that the VHTC data is indeed calculated from the experimental data and proves its correlation.

As shown in Fig. 3, the fluctuation of correlation coefficient can be seen more intuitively. The correlation coefficient of continuous phase flow is small, because the continuous phase flow effect is weak in the actual experiment. In the whole circulating heat transfer process, the continuous phase flow is an indirect quantity that has little influence on the whole heat transfer process. Therefore, its correlation coefficient is small and has little influence on the volumetric heat transfer coefficient. It can be clearly found from the figure that the continuous phase height obtained by grey correlation analysis is consistent with the numerical size of the continuous phase volume. The effect of continuous phase height and continuous phase volume on the volume heat transfer coefficient is equally important. Because in the actual experiment process, the continuous phase volume changes with the continuous phase height. Thus, the higher the continuous phase height, the larger the continuous phase volume. Therefore, data with equal value in a set of grey correlation analysis can be arbitrarily removed to reduce the amount of data. Hence, according to this phenomenon, the effect of different influencing factors on the prediction ability is studied, and two factors with the same correlation coefficient are analyzed to predict the effect.

In this paper, the correlation coefficient is ranked by GRA to change the number of factors affecting VHTC. The influence of the number of influencing factors on the prediction results of VHTC is researched. As shown in Table 4, seven factors mean that any one of the two same correlation coefficients is removed under the condition of eight factors. Six factors mean that the factor with the worst correlation coefficient is removed under seven factors. Five factors refers to the removal of the dispersed phase flow rate on the basis of six factors. Four factors refers to the removal of ΔT on the basis of five factors. It is found that when one of the two same correlation coefficients is removed, the values of MAE, MAPE and RMSE will not change, indicating that the predicted results will not be affected under this condition. When the minimum correlation coefficient is removed again, the values of P, K and S become larger, while the values of MAE, MAPE and RMSE become smaller, indicating that under this condition the prediction results will be positively affected and the prediction ability will be improved. When the number of influencing factors is four or five, it is found

Table 4. The influence of the number of different influencing factors on prediction performance

Proportion of training set	Number of factors	75%	80%	85%
Pearson correlation coefficient	8 factors	0.9709	0.9724	0.9759
	7 factors	0.9709	0.9724	0.9759
	6 factors	0.9734	0.9727	0.9740
	5 factors	0.9706	0.9699	0.9699
	4 factors	0.6734	0.6836	0.6756
Kendall correlation coefficient	8 factors	0.8648	0.8662	0.8738
	7 factors	0.8648	0.8662	0.8738
	6 factors	0.8697	0.8691	0.8662
	5 factors	0.8593	0.8576	0.8534
	4 factors	0.4056	0.3817	0.3708
Spearman correlation coefficient	8 factors	0.9731	0.9740	0.9752
	7 factors	0.9731	0.9740	0.9752
	6 factors	0.9732	0.9731	0.9697
	5 factors	0.9694	0.9672	0.9663
	4 factors	0.5639	0.5317	0.5232
MAE	8 factors	0.0045	0.0035	0.0025
	7 factors	0.0045	0.0035	0.0025
	6 factors	0.0042	0.0033	0.0025
	5 factors	0.0043	0.0034	0.0026
	4 factors	0.0154	0.0123	0.6756
MAPE	8 factors	0.0165	0.0131	0.0094
	7 factors	0.0165	0.0131	0.0094
	6 factors	0.0152	0.0124	0.0092
	5 factors	0.0155	0.0126	0.0099
	4 factors	0.0520	0.0432	0.3708
RMSE	8 factors	0.1897	0.1702	0.1414
	7 factors	0.1897	0.1702	0.1414
	6 factors	0.1813	0.1688	0.1456
	5 factors	0.1907	0.1773	0.1558
	4 factors	0.5907	0.5290	0.5232

that the values of P, K and S are decreased, while the values of MAE, MAPE and RMSE are increased, indicating that when the influencing factors are removed again on the basis of six factors, important information may be removed, leading to larger errors and worse prediction effect. When the proportion of the training set is 75%, 80% and 85%, the values of MAE, MAPE and RMSE gradually decrease with the increasing proportion of the training set, indicating that the larger the proportion of the training set is, the better the training effect will be, the smaller the error will be and the more accurate the prediction effect will be. Therefore, when GRA is used, the correlation ranking of the influencing factors of VHTC can be conducted to find the major and minor factors. Then the factors with relatively low correlation and the same correlation degree are removed, so as the influence of main factors on VHTC is enhanced, the proportion of important information in the prediction model is increased, and the prediction accuracy is improved. The data can be optimized and the data quantity can be reduced to improve the data quality of the input prediction model.

VMD algorithm is used to decompose the composite signal into k IMF components, which is adaptive. VMD is used for signal decomposition, so that the problems of mode aliasing and end-point effect can be solved, and the existence of pulse, interference or noise can be analyzed accurately. As can be seen from Fig. 4, VHTC data is unstable and nonlinear composite data. The two

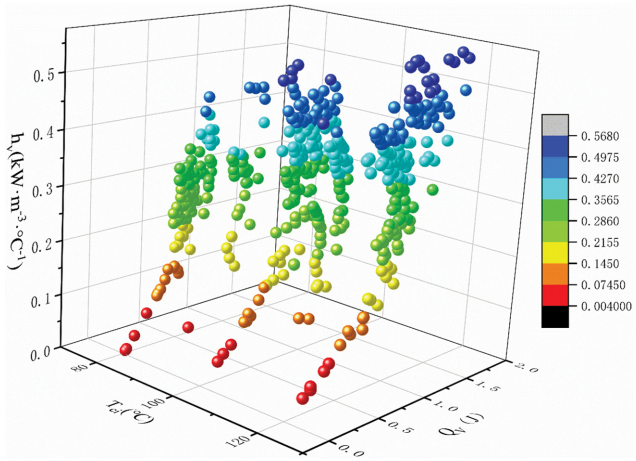

Fig. 4. VHTC data under different operating conditions.

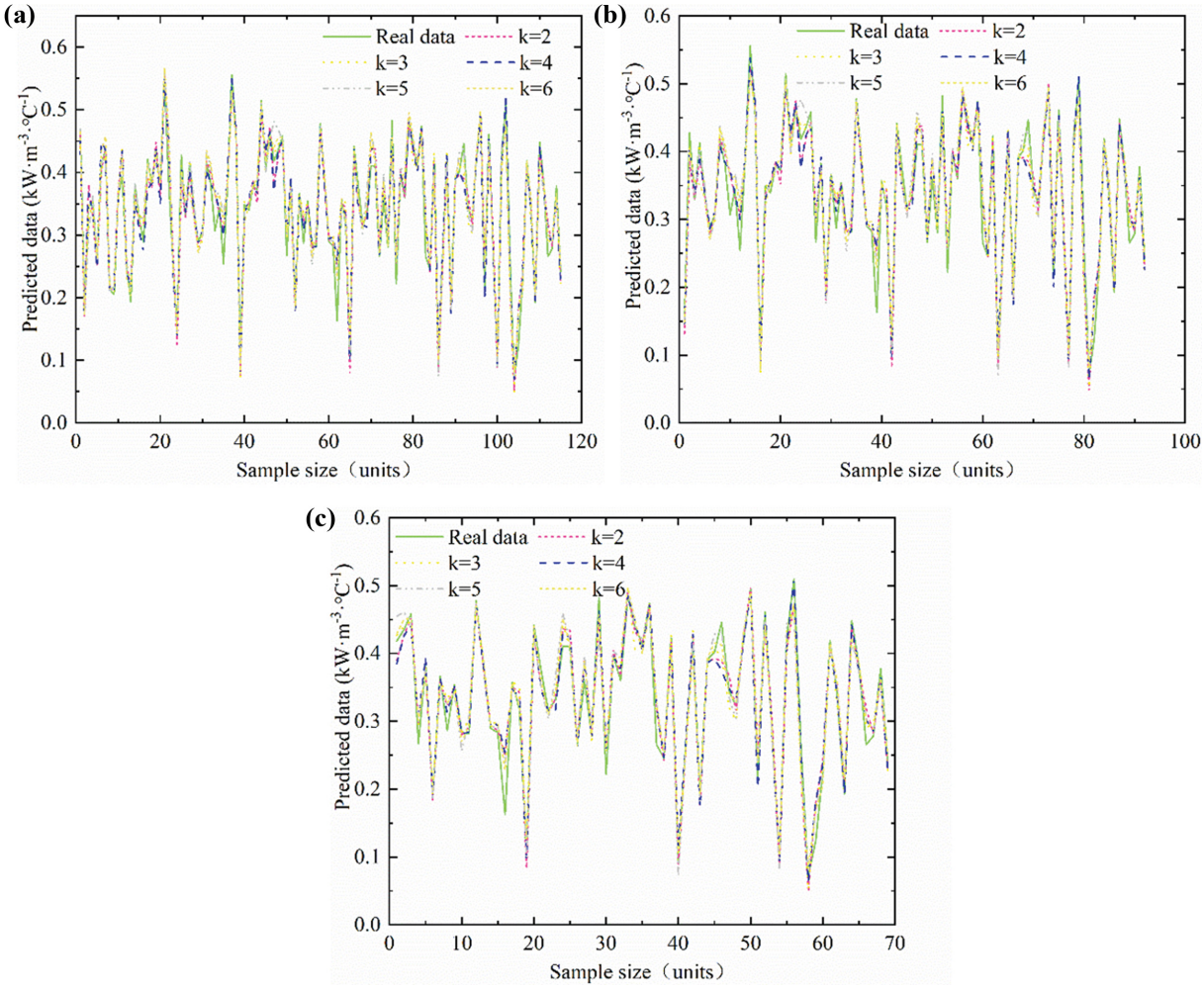

Fig. 5. Influence of VMD decomposition number on VHTC data prediction under different proportion of training sets (a) 75%, (b) 80%, (c) 85%.

factors with the largest correlation degree obtained by GRA are taken as X and Y axes, where X-axis is continuous phase inlet temperature, Y-axis is heat transfer quantity, and Z-axis is volumetric heat transfer coefficient. As can be seen from the figure, h_v increases with the increase of ΔT and Q_v, but most of the values are basically concentrated in the middle position, which is consistent with the trend of changes over time in the experiment. It can reflect that the VHTC data is nonlinear and irregular. Therefore, it is necessary to process VHTC data to obtain more stable and detailed data information.

VHTC data is decomposed by VMD. The results are shown in Fig. 5. H is the original signal of VHTC data. It can be clearly seen that the VHTC signal is volatile, unstable and irregular to follow. IMF_1-IMF_4 is the component obtained by the decomposition of VMD. The VHTC data is decomposed into signal components from high frequency to low frequency. The decomposed data is significantly more detailed, smoother, and more specific than the original VHTC data. The results show that VMD decomposition has strong adaptability and robustness, and the data can be processed and decomposed well. The problem of determining the number of VMD decompositions to be solved is difficult. Because of the different number of decomposition, the prediction performance is affected. If the number of decomposition is too much, it will lead to excess components and increase the interference information. If the number of decomposition is too small, the signal decomposition will not be thorough enough. As shown in Table 5, when the decomposition number of VMD is 4, the prediction effect of VHTC is the best. When the proportion of the training set is 75%, 80% and 85%, the values of P, K and S are all greater than 0.8. Therefore, it can be said that the forecasting model is good. When the proportion of the training set is 75%, it is found that the value of MAE decreases first and then increases with the increase of the decomposition number. When k=4, the value of MAE is the smallest. The values of MAPE and RMSE also increased first and then decreased. Therefore, it can be concluded that when k=4, the error is the smallest and the prediction accuracy is the highest. When the training set proportion was 80% and 85%, the variation trend of MAE, MAPE and RMSE was consistent with that of

Table 5. The influence of different number of VMD decomposition on model prediction performance

Proportion of training set	VMD decomposition number	75%	80%	85%
Pearson correlation coefficient	k=2	0.9746	0.9741	0.9767
	k=3	0.9762	0.9756	0.9781
	k=4	0.9771	0.9782	0.9797
	k=5	0.9721	0.9709	0.9746
	k=6	0.9759	0.9755	0.9821
Kendall correlation coefficient	k=2	0.8670	0.8619	0.8704
	k=3	0.8661	0.8686	0.8670
	k=4	0.8719	0.8767	0.8806
	k=5	0.8505	0.8438	0.8525
	k=6	0.8575	0.8548	0.8798
Spearman correlation coefficient	k=2	0.9719	0.9705	0.9731
	k=3	0.9737	0.9741	0.9737
	k=4	0.9719	0.9735	0.9749
	k=5	0.9665	0.9627	0.9662
	k=6	0.9687	0.9658	0.9770
MAE	k=2	0.0044	0.0035	0.0025
	k=3	0.0041	0.0033	0.0024
	k=4	0.0037	0.0029	0.0021
	k=5	0.0046	0.0040	0.0029
	k=6	0.0044	0.0037	0.0025
MAPE	k=2	0.0165	0.0134	0.0100
	k=3	0.0153	0.0125	0.0094
	k=4	0.0136	0.0111	0.0085
	k=5	0.0160	0.0140	0.0103
	k=6	0.0154	0.0131	0.0095
RMSE	k=2	0.1774	0.1635	0.1365
	k=3	0.1736	0.1605	0.1332
	k=4	0.1685	0.1513	0.1282
	k=5	0.1895	0.1753	0.1433
	k=6	0.1758	0.1632	0.1261

Table 6. The optimized data and IMFs for VMD-LSSVM model specific analysis

Proportion of training set	Input decomposition quantity	75%	80%	85%
Pearson correlation coefficient	IMF_1	0.9728	0.9717	0.9747
	IMF_2	0.9763	0.9750	0.9769
	IMF_3	0.9759	0.9764	0.9776
	IMF_4	0.9771	0.9728	0.9797
Kendall correlation coefficient	IMF_1	0.8654	0.8605	0.8670
	IMF_2	0.8682	0.8605	0.8687
	IMF_3	0.8664	0.8748	0.8849
	IMF_4	0.8719	0.8767	0.8806
Spearman correlation coefficient	IMF_1	0.9712	0.9694	0.9729
	IMF_2	0.9723	0.9697	0.9739
	IMF_3	0.9709	0.9728	0.9766
	IMF_4	0.9719	0.9735	0.9749
MAE	IMF_1	0.0043	0.0035	0.0025
	IMF_2	0.0040	0.0033	0.0024
	IMF_3	0.0040	0.0031	0.0023
	IMF_4	0.0037	0.0029	0.0021
MAPE	IMF_1	0.0154	0.0127	0.0094
	IMF_2	0.0146	0.0124	0.0095
	IMF_3	0.0151	0.0124	0.0096
	IMF_4	0.0136	0.0111	0.0085
RMSE	IMF_1	0.1839	0.1727	0.1443
	IMF_2	0.1716	0.1614	0.1366
	IMF_3	0.1726	0.1564	0.1333
	IMF_4	0.1685	0.1513	0.1282

the training set proportion as 75%. And with the increasing proportion of the training set, the values of MAE, MAPE and RMSE are constantly decreasing, with smaller errors and higher accuracy. When k=4, the signal is in a stable state, indicating that the decomposition of VHTC time series is optimal. Therefore, to better prediction effect, the number of VMD decomposition in this paper is determined to be 4.

2. Illustration of Model Performance

In this part, a combination of GRA, VMD and LSSVM is compared with an unoptimized single LSSVM model. As shown in Fig. 5, the influence of GRA-VMD-LSSVM hybrid model and the number of VMD decomposition on the prediction performance of VHTC is compared. The first graph takes up 75% of the training set. What is described above is the true value and the predicted value of 2 to 6 VMD decomposition. It can be found that when k=4, the gap between the true value and the predicted value is small, and the trend is basically the same. The predicted value is closer to the true value curve than the other VMD decomposition numbers. In the 0-10 area in the figure, other lines except the blue line are quite different from the green line, and the error degree is relatively high. In the 20-60 region, the blue line is more consistent with the green line, the error between the true value and the predicted value is smaller, and the prediction accuracy is higher. The second graph takes up 80% of the training set. By comparing the whole, the blue line is more consistent with the true value, and the trend is closer to the true value. The third graph describes the comparison of the true and predicted values when the training set proportion is 85%. The results verify again that when the number of VMD decomposition is 4, the prediction performance of the hybrid model is higher and the prediction results are more accurate. Therefore, after determining the decomposition number of VMD, the influence of IMFs on the prediction performance is researched.

Not all the IMF components decomposed by VMD have a positive impact on the prediction model, so it is necessary to analyze IMFs to find how many IMFs can be input to obtain the optimal prediction accuracy of the model. The influence results are shown in Table 6. For the GRA-VMD-LSSVM hybrid model, the IMFs and VHTC data are studied in combination with the main influence factors. The results show that when the prediction set accounts for 25% of the total data, the values of P, K and S are all greater than 0.8, which indicates extreme correlation. The prediction accu-

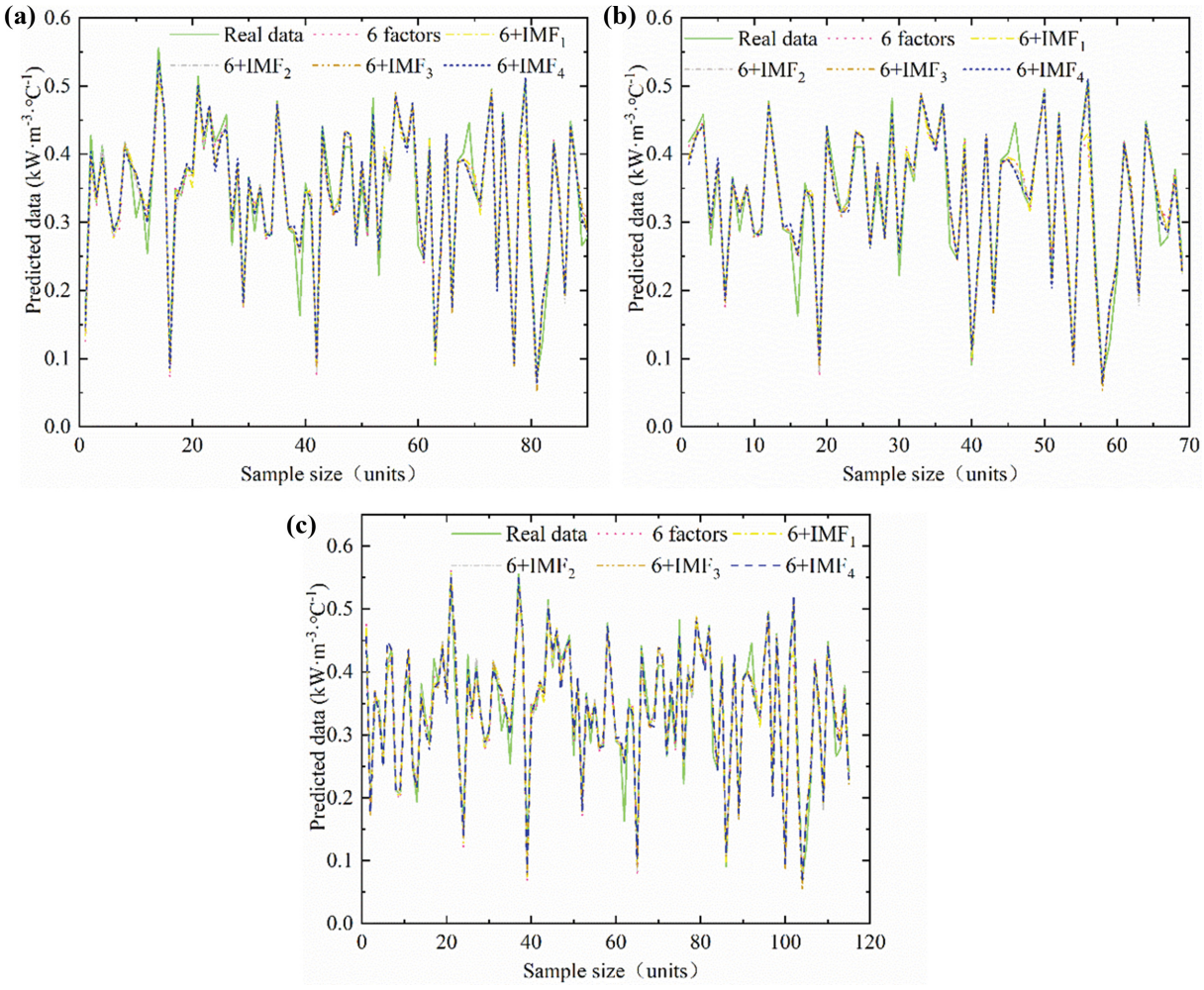

Fig. 6. Influence of different IMFs input number on the prediction performance under different proportion of training sets (a) 75%, (b) 80%, (c) 85%.

racy of the model was indeed improved by adding IMF components to the prediction model. The value of MAE decreases with the increase of the number of IMFs added into the prediction model. The values of MAE are 0.0043, 0.0040, 0.0040 and 0.0037. The values of MAPE and RMSE first decreased, then increased and finally decreased with the number of IMFs added. The values of MAPE are 0.0154, 0.0146, 0.0151 and 0.0136. The values of RMSE are 0.1839, 0.1716, 0.1726 and 0.1685. The values of MAE, MAPE, and RMSE generally declined gradually when the test set accounted for 20% and 15%. However, the final result is that when the main influencing factors and the four IMFs decomposed by VMD are all added into the prediction model, the prediction effect is the best and the error is the least. Fig. 6 shows a comparison of the predicted and true values for the six major factors with different numbers of IMFs. The comparison of the true and predicted values of the 25% test set is shown in the first figure. The blue line overlaps more closely with the green line and fits better. This suggests that six main factors plus four IMFs bring the forecast trend closer to the true value. The second and third graphs compare the real value with the predicted value when the test set accounts for 20% and 15%. It can be more clearly found that the blue line and the green line have a higher degree of fit, and the predicted result is better. The GRA-VMD-LSSVM hybrid model is summarized with evaluation indexes, as shown in Fig. 7. It can be seen more intuitively that when six major factors and four IMFs are input into the GRA-VMD-LSSVM hybrid model, the values of the three evaluation indexes are all the minimum. Therefore, this research found that the prediction effect of VHTC is the best and the prediction accuracy of model is the highest when all the six major factors and four IMFs are added into the GRA-VMD-LSSVM model.

As shown in Table 7, the optimized VMD-LSSVM hybrid model is compared with a single LSSVM model. No matter what the proportion of the training set is, the values of P, K and S are all above 0.8. It shows that the regression model has good performance. With the increasing proportion of the training set, the values of MAE, MAPE and RMSE gradually decrease. It shows that the prediction performance and accuracy are constantly improved as the error gradually decreases. When the VMD-LSSVM hybrid model was used, the MAE values were 0.0037, 0.0029 and 0.0021, respectively. Using only a single model, the MAE values were 0.0042, 0.0033 and 0.0025, respectively. It is found that the hybrid model can improve the accuracy of prediction and make better prediction. The trend of MAPE and RMSE is basically consistent with that of MAE, again proving that the predictive power of the hybrid model is higher than that of the single model.

Then, the single LSSVM model without data optimization and the GRA-VMD-LSSVM hybrid model were analyzed again, and the results are shown in Table 8. It can be clearly found from the table that although the values of P and K in the hybrid model are larger than those in the single model, the values of S decrease.

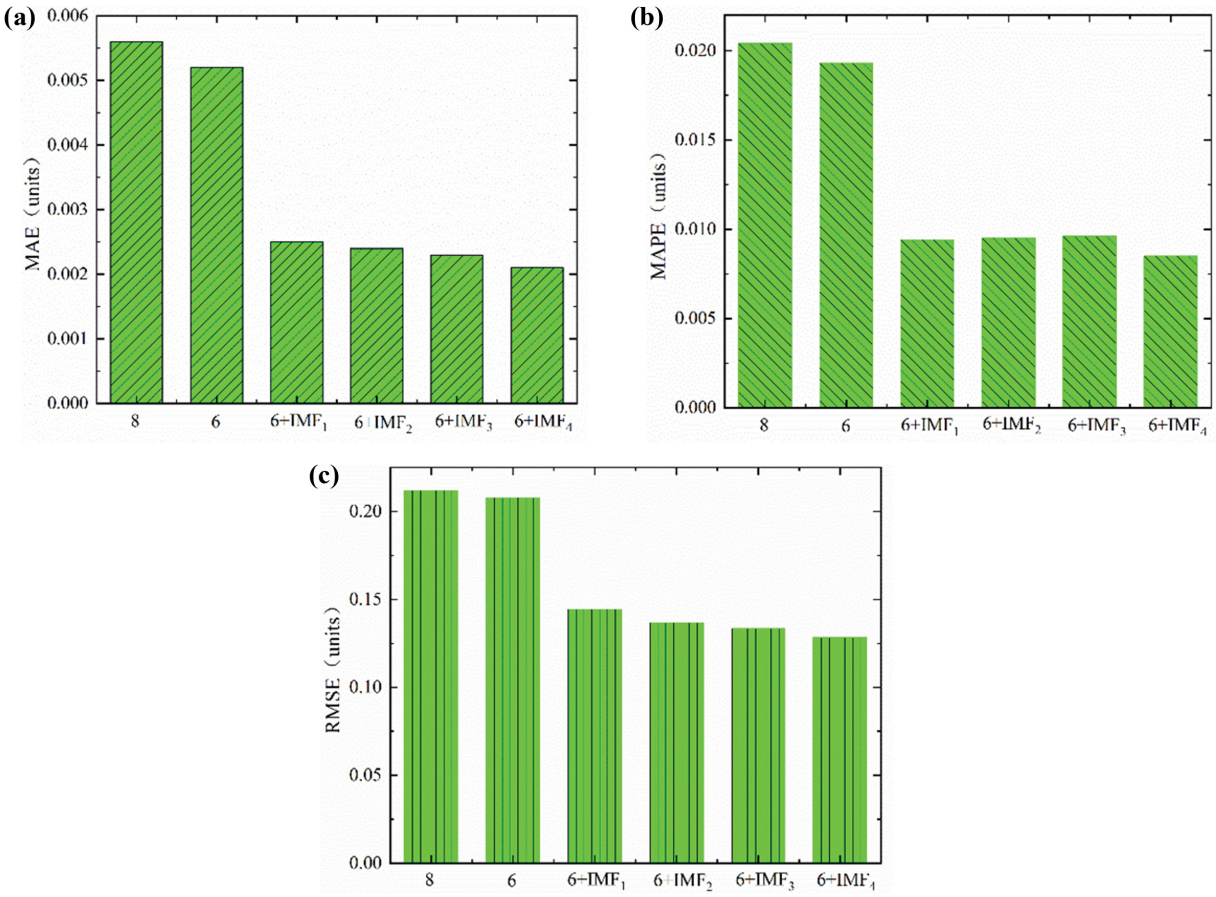

Fig. 7. Numerical analysis of different input factors (a) MAE, (b) MAPE, (c) RMSE.

Table 7. The prediction performance of the single LSSVM model and the VMD-LSSVM hybrid model were compared after data optimization

Proportion of training set	Model	75%	80%	85%
Pearson correlation coefficient	Hybrid model	0.9771	0.9782	0.9797
	Single model	0.9734	0.9727	0.9740
Kendall correlation coefficient	Hybrid model	0.8719	0.8767	0.8806
	Single model	0.8697	0.8691	0.8662
Spearman correlation coefficient	Hybrid model	0.9719	0.9735	0.9749
	Single model	0.9732	0.9731	0.9697
MAE	Hybrid model	0.0037	0.0029	0.0021
	Single model	0.0042	0.0033	0.0025
MAPE	Hybrid model	0.0136	0.0111	0.0085
	Single model	0.0152	0.0124	0.0092
RMSE	Hybrid model	0.1685	0.1513	0.1282
	Single model	0.1813	0.1688	0.1456

Table 8. Comparison of single LSSVM model and GRA-VMD-LSSVM model for prediction performance analysis

Proportion of training set	Model	75%	80%	85%
Pearson correlation coefficient	Hybrid model	0.9771	0.9782	0.9797
	Single model	0.9709	0.9724	0.9759
Kendall correlation coefficient	Hybrid model	0.8719	0.8767	0.8806
	Single model	0.8648	0.8662	0.8738
Spearman correlation coefficient	Hybrid model	0.9719	0.9735	0.9749
	Single model	0.9731	0.9740	0.9752
MAE	Hybrid model	0.0037	0.0029	0.0021
	Single model	0.0045	0.0035	0.0025
MAPE	Hybrid model	0.0136	0.0111	0.0085
	Single model	0.0165	0.0131	0.0094
RMSE	Hybrid model	0.1685	0.1513	0.1282
	Single model	0.1897	0.1702	0.1414

However, the values of P, K and S are all greater than 0.8, indicating that the model has a higher degree of correlation and a better fit with the model. Therefore, the advantages of the model cannot be compromised. The value of MAE in the hybrid model is lower than that in the single model, and the value of MAE becomes smaller as the proportion of the training set keeps increasing. It shows that with the increasing proportion of the training set, the smaller the error between the predicted value and the true value is, the higher the prediction accuracy is. The trend of MAPE and RMSE is basically consistent with that of MAE. In general, the prediction performance of the GRA-VMD-LSSVM hybrid model is better than that of the single model. The error between the actual value and the predicted value is reduced, the prediction accuracy of model is improved, and VHTC can be better predicted.

Through the above comparative study, it can be concluded that the use of GRA-VMD-LSSVM hybrid model can indeed reduce experimental data, improve data quality and improve prediction ability. In addition, Fig. 8 shows the fitting diagram of the real values obtained from the experiment and the predicted values obtained from the prediction. It can be seen that most of the predicted data are around the true value line and very few predicted values deviate from the true value line. The good performance of the GRA-VMD-LSSVM hybrid model is demonstrated. The proposed hybrid model has strong learning ability and robustness for the prediction of VHTC. The optimized combination of processed and decom-

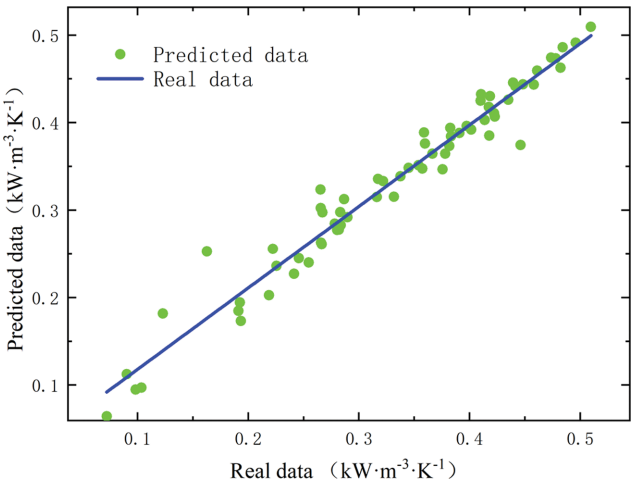

Fig. 8. Fitting results of predicted values of GRA-VMD-LSSVM and real values of VHTC.

posed data is superior to a single model.

SUMMARY AND CONCLUSIONS

A GRA-VMD-LSSVM hybrid model was proposed to predict the volumetric heat transfer coefficient of direct contact heat exchangers in a robust and efficient manner. Through finding the main factors affecting VHTC and extracting features from VHTC data, the model can predict VHTC faster and more accurately, and improve the prediction effect and generalization ability of the model effectively. By comparing with the single machine learning model, the superiority and reliability of the hybrid model are verified. The conjecture of this research is verified by comparing the predicted results with the real results. GRA-VMD-LSSVM can better predict VHTC data through detailed information and important information. It is also confirmed the influence of non-main information and the amount of VMD decomposition on the prediction accuracy. The hybrid model has better performance than the single model.

Combined with the calculation and comparative analysis of different models, the following conclusions can be drawn: (1) To find out the correlation between variables, GRA is proposed to process the factors. It was found that removing redundancy factors did improve the VHTC prediction effect. The prediction accuracy of VHTC can be improved by 7% when the redundancy factors are removed. (2) VHTC data can be converted into stable and detailed data. However, when VMD is used for signal decomposition of data, it is necessary to select k value comprehensively considering the actual situation, so as to ensure that the prediction effect of the hybrid model can achieve the best effect. The conclusion is that when the value of k is 4, the prediction ability of the model can be improved by 16% again while ensuring the effective decomposition of IMFs with different frequencies. (3) With the increasing proportion of the training set, the error between the real value and the predicted value becomes smaller and smaller, and the prediction accuracy of the model increases continuously. The values of P, K and S predicted by the hybrid model are all greater than 0.8, indicating that the predicted value is highly correlated with the real value, and the prediction performance of the hybrid model is good.

In general, the necessity of putting forward the GRA-VMD-LSSVM hybrid model is that it can effectively distinguish the main and secondary factors, weaken the nonlinear and unstable data sequence. The optimized data and extracted detailed information can make the prediction model faster and better fitting. Compared with the single model, the GRA-VMD-LSSVM hybrid model is a kind of prediction model that makes the data quality higher, more reliable and more convenient.

ACKNOWLEDGEMENTS

The authors wish to thank the referees for their numerous detailed questions and constructive criticism, which greatly improved the presentation. This work is partially supported by the National Natural Science Foundation of China (Grant Nos. 52166010, No. 51706195), Natural Science Foundation of Yunnan Province, China (No. 202101AT070202).

DECLARATION OF CONFLICTING INTEREST

The authors declare no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

NOMENCLATURE

Symbol

b : bias term
cov : covariance
d_i : difference of rank
H : random data sequence
h : enthalpy of dispersed phase [J·kg^{-1}]
h_{do} : outlet enthalpy of working medium steam [kJ/kg]
h_{di} : inlet enthalpy of working fluid [kJ/kg]
K : Kendall correlation coefficient
m_g : mass flow rate of dispersed phase steam [kg·s^{-1}]
n : number under the constant increase of experimental variables
n_c : number of the data pairs with the same consistency
n_d : number of the data pairs with different consistency
P : Pearson correlation coefficient
Q : heat transfer quantity [J]
q_m : evaporation mass flow of working medium [kg·s^{-1}]
R : residual
S : Spearman correlation coefficient
T : temperature [°C]
ΔT : logarithmic mean temperature difference [°C]
T_{ci} : inlet temperature of continuous fluid [°C]
T_{co} : outlet temperature of continuous fluid [°C]
T_{di} : inlet temperature of dispersed fluid [°C]
T_{do} : outlet temperature of dispersed fluid [°C]
U_v : VHTC [kW·m^{-3}·°C^{-1}]
$U_v(t)$: time series of VHTC [kW·m^{-3}·°C^{-1}]
$\{u_M\}$: *K*th modal component
: center frequency
V : VHTC
w^T : transposed output layer vector
x : real values in the experiment
Dirac function
y : predicted values in the experiment
y_i : value obtained by the actual experimental test
$\hat{y}_i$: predicted value obtained by the prediction model
new dimensionless mother sequence
: new dimensionless subsequence
$\varphi(x_k)$: kernel function
difference between the mother sequence and the subsequence
correlation coefficient
final correlation degree of the comparison sequence with the reference sequence
: convolution operator

Abbreviations

DCHE : direct contact heat exchanger
DCHT : direct contact heat transfer

EMD : empirical mode decomposition
GRA : grey relation analysis
IMF : intrinsic mode function
LSSVM : least squares support vector machine
MAE : mean absolute error
MAPE : mean absolute percentage error
ORC : organic Rankine cycle
RMSE : root mean square error
RBF : radial basis function
SVM : support vector machine
VHTC : volumetric heat transfer coefficient
VMD : variational mode decomposition

REFERENCES

1. M. Olfati, M. Bahiraei, S. Nazari and F. Veysi, *Energy*, **209**, 118430 (2020).
2. B. Huang, Q. Jian, L. Luo and J. Zhao, *Energy Convers. Manage.*, **138**, 38 (2017).
3. B. Jiang, D. Xia, H. Zhang, H. Pei and X. Liu, *Energy*, **208**, 118346 (2020).
4. A. Atmaca and R. Yumrutas, *Energy Convers. Manage.*, **79**, 790 (2014).
5. A. Sagastume Gutiérrez, J. B. Cogollos Martínez and C. Vandecasteele, *Appl. Therm. Eng.*, **51**, 273 (2013).
6. Y. Han and Y. Sun, *Energy*, **207**, 118172 (2020).
7. Kh. Hosseinzadeh, D. D. Ganji and F. Ommi. *J. Mol. Liq.*, **315**, 113748 (2020).
8. L. Wang, X. Bu and H. Li, *Energy*, **203**, 117809 (2020).
9. M. Vasile, *Appl. Therm. Eng.*, **69**, 143 (2014).
10. H. Nami and Anvari-Moghaddam A, *Energy*, **192**, 116634 (2020).
11. Z. Hou, X. Wei, X. Ma and X. Meng, *J. Cleaner Prod.*, **246**, 119064.1 (2020).
12. A. Giovannelli, E. M. Archilei and C. Salvini, *Energies*, **13**, 1054 (2020).
13. B. Peris, J. Navarro-Esbri and F. Moles, *Energy*, **85**, 534 (2015).
14. R. Pili, L. Martínez, C. Wieland and S. Hartmut, *Renew. Sust. Energy Rev.*, **134**, 110324 (2020).
15. H. Fu, *Chem. Eng. Process.: Process Intensification*, **149**, 107829 (2020).
16. S. Halkarni, A. Sridharan and S. Prabhu, *Int. J. Therm. Sci.*, **110**, 340 (2016).
17. Y. Zheng, H. Dong, J. Cai, J. Feng, L. Zhao, J. Liu and S. Zhang, *Appl. Therm. Eng.*, **151**, 335 (2019).
18. T. Zeng, C. Zhang, D. Hao, D. Cao, J. Chen, J. Chen and J. Li, *Energy*, **208**, 118319 (2020).
19. S. P. Sundar Singh Sivama, K. Saravananb, N. Harshavardhanaa and D. Kumarana, *Mater. Today: Proc.*, **45**, 1464 (2020).
20. B. Ppda and B. Sc, *Mater. Today: Proc.*, **28**, 568 (2020).
21. S. Ren, C. Wang, Y. Xiao, J. Deng, Y. Tian, J. Song, X. Cheng and G. Sun, *Fuel*, **260**, 116287 (2020).
22. H. Hong, Z. Zhang, A. Guo, L. Shen, H. Sun, Y. Liang, F. Wu and H. Lin, *J. Hydrol.*, **591**, 125574 (2020).
23. Y. Zhao, B. Zhang and L. Han, *Optics Commun.*, **456**, 124588 (2020).
24. Y. Zhang, G. Pan, B. Chen, J. Han, Y. Zhao and C. Zhang, *Renew. Energy*, **156**, 1373 (2020).
25. F. Xiao, D. Yang, Z. Lv, X. Guo, Z. Liu and Y. Wang, *Future Generation Computer Systems*, **110**, 1023 (2019).
26. S. Gyamerah, *J. of King Saud University - Computer and Information Sciences*, 1319 (2020).
27. A. Ms, B. Am, A. Hq, C. Ga, D. Pn and A. Cr, *Energy*, **205**, 117986 (2020).
28. Y. Yu, M. Shao, L. Jiang, Y. Ke, D. Wei, D. Zhang, M. Jiang and Y. Yang, *Optik*, **237**, 166759 (2021).
29. R. Hu, K. Chen, W. Chen, Q. Wang and H. Luo, *Waste Manage.*, **126**, 791 (2021).
30. K. Amber, R. Ahmad, M. Aslam, A. Kousar, M. Usman and M. Khan, *Energy*, **157**, 886 (2018).
31. J. Wang and J. Hu, *Energy*, **93**, 41 (2015).
32. H. Sun, Y. Wu, Z. Wu, F. Han, M. Yang and Y. Wang, *Phytochem. Lett.*, **43**, 108 (2021).
33. S. Singh, K. Parmar, S. Makkhan, J. Kaur and J. Kumar, *Chaos Solitons & Fractals*, **139**, 110086 (2020).
34. H. Han, X. Cui, Y. Fan and H. Qing, *Appl. Therm. Eng.*, **154**, 540 (2019).
35. S. Thongwik, N. Vorayos, T. Kiatsiriroat and A. Nuntaphan, *Int. Commun. Heat Mass Transf.*, **35**, 756 (2008).
36. J. Huang, J. Xu, X. Sang, H. Wang and H. Wang, *Int. J. Heat Mass Transf.*, **75**, 497 (2014).
37. A. Baqir, H. Mahood, A. Campbell and A. Griffiths, *Appl. Therm. Eng.*, **103**, 47 (2016).
38. J. Li, X. Zhang and J. Tang, *J. Appl. Geophys.*, **180**, 104127 (2020).
39. S. Bavkar, B. Iyer and S. Deosarkar, *Probl. Biocybern. Biomed. Eng.*, 41 (2020).
40. M. Alizamir, S. Kim, O. Kisi and M. Zounemat-Kermani, *Energy*, **197**, 117239 (2020).
41. C. Xiao, J. Ye, M. Rui and C. Rong, *Concurrency and Computation: Practice and Experience*, **28**, 3866 (2016).