Blood Glucose Prediction With VMD and LSTM Optimized by Improved Particle Swarm Optimization

WENBO WANG, MENG TONG, AND MIN YU
School of Science, Wuhan University of Science and Technology, Wuhan 430081, China
Corresponding author: Wenbo Wang (wangwenbo@wust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671338 and Grant 51877161.

ABSTRACT The time series of blood glucose concentration in diabetics are time-varying, nonlinear and non-stationary. To improve the accuracy of blood glucose prediction, a short-term blood glucose prediction model (VMD-IPSO-LSTM) combining variational modal decomposition (VDM) and improved Particle swarm optimization optimizing Long short-term memory network (IPSO-LSTM) was proposed. Firstly, the time series of blood glucose concentration of patients was decomposed by using VMD method to obtain the intrinsic modal functions (IMF) of blood glucose components in different frequency bands, so as to reduce the non-stationarity of blood glucose time series. Then a prediction model was established for each blood glucose component IMF by using the long and short time memory network. Since the number of neurons, learning rate and time window length of LSTM are difficult to determine, the improved PSO algorithm is used to optimize these parameters. The optimized LSTM network was used to predict each IMF, and finally the predicted subsequence was superimposed to obtain the final prediction result. The data of 56 patients were selected as experimental data from 451 patients with diabetes mellitus. The experimental results showed that the proposed VMD-IPSO-LSTM model could achieve high prediction accuracy at 30min, 45min and 60min in advance. When predicted 60 minutes in advance, compared with the LSTM, VMD-LSTM and VMD-PSO-LSTM methods, the RMSE of proposed method decreased by 15.565, 3.402, 1.215 and the MAPE of proposed method decreased by 11.284%, 2.024%, 0.834%, and the percentage of predicted values falling into the A zone increased by 23.5%, 6.1% and 2.8% in the Clarke error grid, respectively. The improved accuracy of blood glucose prediction and longer prediction time can provide sufficient time for physicians and patients to control blood glucose concentrations and improve the effectiveness of diabetes treatment.

INDEX TERMS Blood glucose prediction, variational modal decomposition, PSO, LSTM.

I. INTRODUCTION
According to statistics, there are nearly 34.7 million diabetics in the world [1]. Diabetes has a long course of disease, which is easy to cause damage to the tissues and organs of the whole body, resulting in disease and even death [2], [3]. The artificial pancreas can effectively control and stabilize the blood glucose level of patients, and blood glucose prediction is the key to its improved control performance. However, the time variability, nonlinearity and non-stationarity of the time series of blood glucose concentration in patients greatly affect the accuracy of blood glucose prediction.

The associate editor coordinating the review of this manuscript and approving it for publication was Nuno Garcia.©

Domestic and foreign scholars have proposed various blood glucose prediction methods for data-driven models and physiological models. Wang et al. [4] improved the grey prediction model in terms of smoothness, which significantly improved the prediction accuracy of postprandial blood glucose in diabetic patients compared with AR model. Mo et al. [5] used the extreme learning machine and the regular limit learning machine for blood glucose prediction and hypoglycemia alarm. The comprehensive performance of the two machines is quite good in most cases, with good and stable specificity and acceptable sensitivity. Georga et al. [6] first used the random forest (RF) algorithm to rank the candidate feature sets (blood glucose level, insulin dose, dietary intake, physical activity, etc.), and then combined with the regression model to predict blood glucose concentration.
Shanthi [7] introduced a classical statistical method, namely the algorithm based on the Autoregressive Integrated Moving Average Model (ARIMA) model, to predict blood glucose within 30 ~ 60 minutes. Jun Yang et al. [8] improved the ARIMA method and proposed an adaptive order ARIMA method for predicting blood glucose concentration. Daskalaki et al. [9] used the glucose prediction method of machine learning to use the real-time learning of the glucose and insulin information of the recursive neural network (RNN). The model outperformed both an autoregressive (AR) model using glucose information, as well as an AR model with external insulin input (ARX). Bunescu et al. [10] proposed that using support vector regression (SVR) to predict blood glucose, by taking into account daily events, such as insulin pills and diet. Georgia et al. [11] proposed a method of combining dietary model, insulin model and exercise model to predict individual blood glucose by SVR method. Mhaskar [12] et al. proposed a deep convolutional neural network (DCNN) method to predict blood glucose, which’s performance is better than shallow network.

Recurrent Neural Network (RNN) [13] introduces the concept of timing sequence into the design of Network structure, making the network show stronger adaptability in time series data analysis. Hochreiter and Schmidhuber [14] proposed a Long Short-term Memory (LSTM) model by improving the network unit structure of RNN. By designing the control gate structure, they made up for the problems of RNN, such as gradient disappearance and gradient explosion and the lack of long-term memory ability, make the LSTM network can make full use of the long-distance time sequence information [15].

At present, LSTM neural network has been successfully applied in financial market price prediction [16], traffic flow prediction [17], ocean surface temperature prediction [18] and medical prediction [19]. The prediction method based on LSTM network has also been well applied in the prediction of blood glucose. Qingnan Sun et al. [20] used LSTM network and bidirectional LSTM network to predict the blood glucose concentration sequence, and made a comparative analysis for the prediction results of LSTM and bidirectional LSTM. Aliberti et al. [21] combined the LSTM network with the multi-patient data-driven method to predict the patient’s blood glucose concentration, and achieved a good prediction effect.

But because the time series of blood glucose concentration has strong time varying, which is a typical non-linear and non-stationary series, directly using LSTM to predict blood glucose sequence will affect the prediction accuracy to a certain extent. It is also a difficulty to determine the parameters of LSTM network. The existing research results show that multi-scale decomposition can reduce the influence of non-stationary on the prediction results. Therefore, this paper proposes a blood glucose prediction model based on various mode decomposition and Particle swarm optimization LSTM network (VMD-IPSO-LSTM). Firstly, the time series of blood glucose concentration collected by the continuous glucose monitoring (CGM) were decomposed by VMD, and the intrinsic mode function (IMF) of blood glucose component in different frequency bands was obtained, which can effectively reduce the non-stationary changes of blood glucose concentration in patients. Then, the LSTM prediction model is established for each IMF sequence, and the key parameters of the LSTM model are optimized by the improved particle swarm optimization (IPSO) algorithm of the adaptive learning strategy. The optimization of LSTM network parameters by IPSO method can match the characteristics of IMF sequence with the topological structure of network and improve the prediction accuracy of IMF sequence.

Finally, the IMFs sequence prediction results are superimposed to obtain the final blood glucose prediction, so as to realize the prediction of the patient’s future blood glucose concentration. The model in this paper was used to predict the actual data collected from diabetic patients. The experimental results verified the effectiveness of the proposed method and achieved a good prediction effect 60 minutes in advance.

II. VARIATIONAL MODAL DECOMPOSITION (VMD)

The decomposition process of VMD can be understood as the process of finding the optimal solution to the variational problem [22]. This problem can be transformed into the construction and solution of the variational problem, which involves 3 concepts: classic Wiener filtering, Hilbert transform and frequency mixing. Assume that the multi-component signal $f$ is composed of $K$ finite-band intrinsic modal components (IMF) $u_k(t)$

$$u_k(t) = A_k(t) \cos(\phi_k(t))$$

In Equation (1), phase $\phi_k(t)$ is a non-decreasing function, that is $\phi_k'(t) \geq 0$, $A_k(t)$ represents the envelope function. Assuming that the center frequency of each IMF is $\omega_k$, when the constraint condition is that the sum of each mode equals to the input signal $f$, the specific construction steps of VMD are as follows:

First perform the Hilbert transform to obtain the analytic signal of $u_k(t)$ and calculate the unilateral spectrum, then multiply by $e^{-j\omega_k t}$ to modulate the center band of $u_k(t)$ to the corresponding base band

$$\left[\left(\delta(t) + \frac{j}{\pi t}\right) u_k(t)\right] e^{-j\omega_k t}$$

Then calculate the $L^2$ norm of the demodulated signal gradient above, estimate the bandwidth of each modal signal, and obtain the constrained variational problem as:

$$\min_{\{u_k, \omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \delta(t) + \frac{j}{\pi t} \right\| u_k(t) e^{-j\omega_k t} \right\}$$

In order to find the optimal solution of the above-mentioned constrained variational problem, a Lagrangian multiplication operator $\lambda(t)$ and secondary penalty factor $\alpha$ are introduced to
transform the constrained problem into an unconstrained variational problem. In the case of ensuring the accuracy of the signal reconstruction, the Lagrangian multiplication operator \( \lambda(t) \) maintains strict constraints. The extended Lagrangian expression is as follows:

\[
L (\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^{K} \left| \partial_t \left( \left( \delta(t) + \frac{j}{\pi t} \right) u_k(t) \right) e^{-j\omega t} \right| + \left| f(t) - \sum_{k=1}^{K} u_k(t) \right|^2 + \left( \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right)
\]  

(4)

Finally, the multiplier alternating direction algorithm is used to solve the above problems, and the components and their frequencies are updated continuously. The saddle point of the unconstrained model is obtained in the end, which is the optimal solution of the original problem. All components can be obtained from equation (5) in the frequency domain:

\[
\hat{u}_k^{n+1}(\omega) = \left( \hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right) \times \frac{1}{1 + 2\alpha(\omega - \omega_k)^2}
\]  

(5)

In formula (5), \( \hat{u}_k^{n+1}(\omega), \hat{f}(\omega) \) and \( \hat{\lambda}(\omega) \) respectively represent the Fourier transform of \( u_k^{n+1}(t), f(t) \) and \( \lambda(t) \).

\( \hat{u}_k^{n+1}(\omega) \) is the result of passing the current remaining amount \( \hat{f}(\omega) - \sum_{i \neq k} u_i(\omega) \) through the Wiener filter. In the algorithm, the center frequency is re-estimated according to the center of gravity of the power spectrum of each component, and \( \omega \) is updated by Equation (6).

\[
\omega_k^{n+1} = \frac{\int_{-\infty}^{\infty} \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\hat{u}_k(\omega)|^2 d\omega}
\]  

(6)

According to the above analysis, the algorithm flow of VMD is as follows:

- Step 1, initialize \( \{\hat{u}_k\}, \{\hat{\omega}_k\}, \hat{\lambda}^1 \) and \( n = 0; \)
- Step 2, recursively deduce \( n \leftarrow n + 1 \), and update \( u_k \) and \( \omega_k \) according to formula (5) and formula (6);
- Step 3, update \( \lambda; \)

\[
\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + [\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)]
\]  

(7)

Step 4, if \( \sum_k \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|^2 / \left\| \hat{u}_k^n \right\|^2 < \varepsilon \), where \( \varepsilon \) is the discrimination accuracy, if \( \varepsilon > 0 \), stop the iteration, output the results, and get k modal components and their center frequencies; otherwise, return to step 2.

**III. LSTM PREDICTION MODEL OPTIMIZED BY IMPROVED PSO**

**A. LONG AND SHORT-TERM MEMORY Network (LSTM)**

The difference between RNN and general neural network lies in the different way of neuron connection: the information of general neural network flows in one direction from the input layer to the hidden layer to the output layer, and there is a directional cycle in the information transmission of RNN. RNN is prone to gradient disappearance and gradient explosion [24]. LSTM is a special RNN, which learns long-term dependent information and avoids the problem of gradient disappearance [21]. In the hidden layer neurons of RNN, LSTM adds a structure called memory cell to remember the past information (as shown in Figure 1), and adds three kinds of gates (input gate, forget gate, output gate) to control the use of historical information.

If the input sequence is \((x_1, x_2, \cdots, x_T)\) and the hidden layer status is \((h_1, h_2, \cdots, h_T)\), then at time \( t \) there are:

\[
f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)
\]  

(8)

\[
i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)
\]  

(9)

\[
C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_c)
\]  

(10)

\[
o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)
\]  

(11)

\[
h_t = o_t \cdot \tanh(C_t)
\]  

(12)

where \( i_t, f_t, o_t \) are input gate, forget gate and output gate respectively; \( c_t \) is the cell, \( W_f, W_i, W_C \) and \( W_o \) are weights of each layer of network respectively; \( b_f, b_i, b_c \) and \( b_o \) are threshold values of each function respectively; \( \sigma(\cdot) \) and \( \tanh(\cdot) \) are sigmoid function and tanh function respectively.

The key of LSTM is the cell state \( C \), which keeps the cell state storage at time \( t \), and the cell state storage is adjusted by forget gate \( f_t \) and input gate \( i_t \). The forget gate is for the cell to remember or forget its previous state \( C_{t-1} \); the input gate will allow or prevent the input signal from updating the unit state; the output gate is the output of the control unit state \( C \) and is transmitted to the next cell. In order to make the LSTM meet the prediction purpose, a linear regression layer needs to be added, that is:

\[
y_t = W_y h_t + b_y
\]  

(14)
where, \( y_i \) represents the output of the final prediction result; \( b_i \) is the threshold of the linear regression layer.

**B. PRINCIPLE OF PSO**

If the potential solution of the optimization problem is regarded as a particle, the particle continuously flies in space, and the position is adjusted according to its own experience and the experience of the best individual in the process of searching for the best position. The particle swarm algorithm first initializes to obtain a set of random solutions, and then iterates and finds the optimal solution by tracking the best particles in the current space. In the multi-dimensional search space, \( m \) particles form a group. In the \( t \)-th iteration, the position and velocity of the \( i \)-th particle are \( X_{i,t} \) and \( V_{i,t} \) respectively. The particle updates its position and velocity by supervising two optimal solutions. The first is the optimal solution sought by the particle itself, i.e., personal best \( pbest_i \), and the other is the optimal solution currently sought by the whole group, i.e. global best \( gbest \). When searching for these two optimal solutions, particles update their speed and new position according to the following formula:

\[
V_{i,t+1} = wV_{i,t} + c_1rand(pbest_i - X_{i,t}) + c_2rand(gbest_i - X_{i,t})
\]

\[
X_{i,t+1} = X_{i,t} + \lambda V_{i,t+1}
\]

In formula (7-8), \( w \) is the velocity inertia factor; \( c_1 \) and \( c_2 \) are the learning factors; \( Rand \) is the random number between \([0,1]\); \( \lambda \) is the velocity coefficient, usually \( \lambda = 1 \).

**C. IMPROVED PSO**

Because the global optimization ability and convergence speed of the basic PSO are limited, the nonlinear variable inertia weight is proposed to improve the performance of PSO. In basic PSO, fixed \( w \) will weaken the global optimization ability and slow down the convergence speed of the algorithm. In this paper, \( w \) is changed to the following form:

\[
w = w_{\text{max}} - (w_{\text{max}} - w_{\text{min}}) \arcsin \frac{t}{t_{\text{max}}} \cdot \frac{2}{\pi}
\]

In formula (9), \( w_{\text{max}} \) and \( w_{\text{min}} \) are the maximum and minimum values of \( w \) respectively; \( t \) is the current number of iterations; \( t_{\text{max}} \) is the maximum number of iterations. When \( t \) is small, \( w \) is close to \( w_{\text{max}} \), and \( w \) decreases slowly, which guarantees the global optimization ability of the algorithm; with \( t \) increasing, \( w \) decreases nonlinearly, and the reduction speed increases rapidly, which guarantees the local optimization ability of the algorithm, so that the algorithm can flexibly adjust the global optimization ability and local optimization ability.

**D. PROCESS OF IPSO OPTIMIZING LSTM**

The CGM time series has strong nonlinearity and non-stationarity. In order to accurately predict blood glucose concentration, this paper builds a prediction model for blood glucose concentration based on VMD and LSTM models. Since the values of some hyper parameters in the LSTM model control the network structure of the model, an improved particle swarm algorithm is used to determine the optimal hyper parameters of the LSTM in order to match the network structure with the characteristics of the glucose concentration data. Firstly, the model takes the number of hidden layer neurons, learning rate and time window size as the optimization objects of the adaptive particle swarm algorithm, and randomly initializes the position information of each particle according to the hyper parameter range; secondly, it builds the LSTM network model according to the corresponding hyper parameter values of the particle positions, and then trains the model using the training data. The specific steps of the IPSO-LSTM proposed are as follows.

**Step 1:** initialize parameters. The size of population, the number of iterations, the learning factors and the limited interval of location and speed are determined.

**Step 2:** Initialize the position and velocity of the particles. A population particle \( X_{i,O}(h_1, h_2, \varepsilon, n) \) is randomly generated, where \( h_1 \) represents the number of neurons in the first hidden layer, \( h_2 \) represents the number of neurons in the second hidden layer, \( \varepsilon \) represents the learning rate of LSTM, and \( n \) represents the size of the time window.

**Step 3:** determine the evaluation function of particles. Assign the particle \( X_{i,O} \) obtained in step 2 to the parameter of LSTM. The data are divided into training samples, validation samples and prediction samples. Input the training samples for neural network training, and after reaching the limit of iteration times, get the output value of the training samples and the output value of the verification samples of the neural network, then the fitness value \( fit_i \) of the individual \( X_i \) is defined as:

\[
fit_i = 0.5 \sum_{j=1}^{J} \frac{|y^j_i - y^j_k|}{y^j_i} \star \frac{1}{J} + 0.5 \sum_{k=1}^{K} \frac{|y^k_i - y^k_k|}{y^k_k} \star \frac{1}{K}
\]

In equation (18): \( y^j_i \) and \( y^k_i \) are the expected output values of training samples and verification samples respectively. Most of the previous studies only use the fitting error of training samples as fitness value. If the neural network is over fitted, the prediction effect of the model is not optimal. The error of validation sample directly reflects the prediction effect of the model. Therefore, the fitness function includes the fitting error of training sample and the validation error of validation sample. In this paper, the training sample error and the verification sample error are given the same weight, i.e. 0.5, and the sum of the two and the weight is used as the fitness function of the model.

**Step 4:** Calculate the fitness value corresponding to each particle position \( X_i \). The personal best and the group best are determined according to the initial particle fitness value, and the best position of each particle is taken as its historical best position.

**Step 5:** During each iteration, update the particle’s own speed and position according to equations (15) and (16) through personal best and global best; calculate the new
IV. GLUCOSE PREDICTION MODEL BASED ON VMD-IPSO-LSTM

Because the blood glucose sequence has the characteristics of randomness, strong non-linearity and non-stationary, it is difficult to effectively improve the prediction accuracy of blood glucose concentration by general prediction methods. Based on the processing of non-stationary signals by the method of variational mode decomposition and the optimization of LSTM parameters by the IPSO algorithm, this paper proposes a prediction model of blood glucose concentration based on VMD and IPSO-LSTM. The algorithm flow chart is shown in Figure 2. The blood glucose concentration prediction process based on VMD and IPSO-LSTM is as follows:

1. Use CGM to continuously collect blood glucose values of diabetic patients, obtain blood glucose time series, and normalize blood glucose data to [0, 1]; for each patient’s blood glucose time series, the first 1000 blood glucose values are selected as the training data set, and the rest 500 are used as the test data set;

2. Use VMD to decompose the non-linear and non-stationary blood glucose concentration sequence into \( K \) relatively stable IMF components (\( IMF_1, IMF_2, \ldots, IMF_K \));

3. IPSO-LSTM prediction model is established. For the \( IMF_i \) sequence, the LSTM parameters are optimized through the improved particle swarm algorithm, and the optimized LSTM is used to predict \( IMF_i \), then obtain the prediction results of each IMF component;

4. The final predicted value of blood glucose concentration of patients is the sum of predicted values of various components by IPSO-LSTM model;

5. The VMD-IPSO-LSTM prediction value is compared with the actual value and other prediction models, and the accuracy of the proposed method is verified by error index analysis.

V. RESULTS AND ANALYSIS

In the following, we will experimentally analyze the prediction accuracy of the proposed VMD-IPSO-LSTM model. The prediction algorithm is implemented in the Matlab2018a environment, and the IPSO program is written in the Matlab language in combination with the deep learning library in the Matlab2018 neural network toolbox to build the prediction model. The 125-hour CGM data of diabetic patients (1500 values in total) were selected as experimental data, where the first 1000 data were used as the training set and the remaining 500 data were used as the test set. In order to evaluate the performance of the proposed method, we compared the results obtained by the method in this paper with the results obtained by the LSTM method [21], VMD-LSTM method [27] and VMD-PSO-LSTM method. All four algorithms use the same training set and test set.

The experiment was run on a Window computer equipped with an Intel@Core™ i5-7200U CPU clocked at 2.50 GHz and 8.0 GB of physical memory.

A. SOURCE OF BLOOD GLUCOSE DATA

The data for this article comes from the RT_CGM dataset, which is freely available for study[25] (https://public.jaeb.org/directnet/study/), the sampling frequency of CGM data is 5 minutes. This data set included glucose trends in 451 heterogeneous populations affected by type I diabetes. Patients were of different ethnic origin and gender (45% males and 55% females) and were divided into three different age groups (adults over 25, adolescents aged 15–24 and children aged 8-14) [21]. The blood glucose time series of 56 patients with type I diabetes were selected and tested on the experiments. The blood glucose concentration data is randomly split the resulting data into two sub-sets for training and testing purposes, containing about 70% and 30% of the initial samples, respectively.

B. MODEL PERFORMANCE EVALUATION INDEX

In order to evaluate the performance of the model, the root mean square error (RMSE), mean absolute percent error (MAPE), and Clark error grid analysis (EGA) method [26] were selected as the performance evaluation indexes in this study. The formula for calculating the root mean square error (RMSE) is as follows:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}
\]  

(19)

FIGURE 2. The procedure of blood glucose prediction based on VMD-IPSO-LSTM.
FIGURE 3. The procedure of blood glucose prediction based on VMD-IPSO-LSTM. Decomposition results of VMD.

The calculation formula of the mean absolute percentage error MAPE is as follows:

\[
MAPE = \frac{1}{n} \sum_{i=1}^{N} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\%
\] (20)

EGA uses the Cartesian diagram principle to evaluate the accuracy of the blood glucose prediction method based on the probability that the predicted value falls in each region of A, B, C, D, and E.

C. MODEL PARAMETER SETTINGS

In the experiment, VMD was used to decompose the blood glucose sequence into 10 layers. The structure of IPSO-LSTM model is composed of input layer, LSTM layer and output layer. The loss function uses the mean square error. The training times are 1000, the training target is 0.0001, and the network model is built in the framework of MATLAB R2018a. In this model, the number of neurons, the learning rate and the time window size are set as the optimized parameters of the LSTM. At the same time, the number of particles in the particle swarm is set to 30, the maximum number of iterations is 500, the learning factor \(c_1 = c_2 = 2\), the speed inertia factor \(w = 0.8\), and the speed coefficient \(\lambda = 1\). \(w_{\text{max}} = 0.9\) and \(w_{\text{min}} = 0.1\) in the improved PSO.

D. ANALYSIS OF RESULTS

1) VMD DECOMPOSITION OF BLOOD GLUCOSE SEQUENCES

Taking a patient’s 125-hour blood glucose time series as an example, VMD decomposition is performed on the patient’s blood glucose time series, and the number of decomposition layers is selected as 10. The decomposition results are shown in Figure 3. IMF 1~IMF 10 are the intrinsic mode function components of VMD.

2) PREDICTED RESULTS

According to the experimental steps of VMD-IPSO-LSTM combination prediction model in Section IV, the blood glucose of diabetic patients is predicted, and the experimental results are compared with LSTM model [21], VMD-LSTM model [27] and VMD-PSO-LSTM model.

In the LSTM prediction model, the number of neurons in the hidden layer is 15, the learning rate is 0.005, and the time window length is 10. In the VMD-LSTM model, the decomposition layer number of VMD is 10, and the parameters of the LSTM network model are the same.

Four models were used to predict blood glucose 30, 45, and 60 minutes in advance, respectively. The comparison of predicted and true values is shown in Figure 4, 5, 6 and 7. The red line represents the predicted value, and the blue line represents the true value. Figure 4(a)~(c) are the comparisons of predicted and true blood glucose values for 30, 45 and 60 min by LSTM prediction model. Figure 5(a)~(c) are the comparisons of predicted and true blood glucose values for 30 45 and 60 min by VMD-LSTM prediction model. Figure 6(a)~(c) are the comparisons of predicted and true blood glucose values for 30, 45 and 60 min by VMD-PSO-LSTM prediction model. Figure 7(a)~(c) are the comparisons of predicted and true blood glucose values for 30, 45 and 60 min by VMD-IPSO-LSTM prediction model.

It can be seen from Figure 4, 5, 6 and 7 that the prediction effect of the LSTM prediction model is significantly worse than VMD-LSTM method, VMD-PSO-LSTM and VMD-IPSO-LSTM method. The VMD-IPSO-LSTM prediction
model uses the VMD method to separate the components of each frequency band of blood glucose signal, and then uses the network optimized by improved PSO to predict each IMF separately, which reduces the characteristic interference and non-stationarity of patients’ blood glucose signal in different scales. At the same time, improved PSO improves the stability of LSTM network and effectively improves the prediction accuracy.

It can be seen from Figure 5 and 7 that the prediction effect of VMD-IPSO-LSTM prediction model and VMD-LSTM prediction model is not significantly different when they predict 30 minutes in advance, but the prediction effect of VMD-IPSO-LSTM model and VMD-LSTM model is significantly different when they predict 45 and 60 minutes in advance. The VMD-IPSO-LSTM prediction model can well fit the patient’s blood glucose non-linear variation curve when making predictions 60 minutes in advance, which shows a great advantage compared with VMD-LSTM. This shows that better prediction effect can be obtained by using IPSO to optimize LSTM, and better parameters can be obtained by IPSO in the process of iteration to optimize LSTM.

It can be seen from Figure 6 and Figure 7 that, compared with VMD-PSO-LSTM, IPSO-VMd-LSTM has a smaller prediction error. This indicates that better prediction effect can be obtained by optimizing LSTM with IPSO, and better key parameters can be obtained in the process of IPSO optimizing LSTM. By comparing figure 5-7, it can be seen that using PSO and IPSO to optimize LSTM can effectively improve the prediction accuracy of VMD-LSTM method. However, in the prediction 45 minutes and 60 minutes in advance, the VMD-IPSO-LSTM method has significantly higher prediction accuracy than the VMD-PSO-LSTM method, which shows that IPSO algorithm has better global optimization ability, and the prediction accuracy can be further improved after IPSO optimizes LSTM.

The Clarke error grid analysis method was used to test the prediction effect, and the Clarke error grid of the patient’s nail 60 minutes in advance was obtained, as shown in Figure 8. Figures 8(a)-(d) are the Clarke error grids 60 minutes in advance for the LSTM model, VMD-LSTM model, VMD-PSO-LSTM model and VMD-IPSO-LSTM model, respectively. The results show that the predicted value of LSTM model is 71.9% in area A, 24.2% in area B and 3.9% in area D; the predicted value of VMD-LSTM model is 89.3% in area A, 9.5% in area B and 1.2% in area D; the predicted value of VMD-PSO-LSTM model is 92.6% in area A, 6.6% in area B and 0.8% in area D. The predicted value of VMD-IPSO-LSTM model is 95.4% in area A, 3.8% in area B and 0.8% in area D.

The value falling in area A has high accuracy, which means that the predicted value does not deviate from the reference value by more than 20%. The Clarke error grid of 56 patients 60 minutes in advance was analyzed. The experimental results showed that the VMD-IPSO-LSTM blood glucose
TABLE 1. Performance comparison of 4 prediction models.

| Model         | 30 min advance | 45 min advance | 60 min advance |
|---------------|----------------|----------------|----------------|
|               | RMSE           | MAPE           | RMSE           | MAPE           | RMSE           | MAPE           |
| LSTM          |                |                |                |                |                |                |
| VMD-LSTM      |                |                |                |                |                |                |
| VMD-PSO-LSTM  |                |                |                |                |                |                |
| VMD-IPSO-LSTM |                |                |                |                |                |                |
| VMD-PSO-LSTM  |                |                |                |                |                |                |

The proposed VMD-IPSO-LSTM short-term blood glucose prediction model can predict blood glucose 30min, 45min, and 60min in advance with better prediction accuracy and has better prediction ability than the LSTM prediction model, VMD-LSTM prediction model, and VMD-PSO-LSTM prediction model.

Although the method in this paper has some advantages in terms of prediction accuracy and extended glucose prediction time, there are still some limitations. Firstly, in the proposed method, the parameters of LSTM are optimized using an improved Particle Swarm Optimization algorithm in the optimization process and makes it easy to fall into the phenomenon of local optimization. We propose to use an improved PSO algorithm to optimize the key parameters of the network and improve the prediction stability of the network. The experimental results show that

1) The combination of LSTM prediction model and VMD decomposition method can effectively improve the accuracy of blood glucose concentration prediction. The VMD method can effectively solve the modal overlap problem of CGM data, reduce the non-stationary characteristics of the original series, and mine the data features in depth.

2) The improved Particle Swarm Optimization (IPSO) has better parameter finding ability, which can effectively solve the problem of difficult selection of learning parameters in LSTM and greatly improve the prediction accuracy of glucose concentration prediction model.

The proposed VMD-IPSO-LSTM short-term blood glucose prediction model can predict blood glucose 30min, 45min, and 60min in advance with better prediction accuracy and has better prediction ability than the LSTM prediction model, VMD-LSTM prediction model, and VMD-PSO-LSTM prediction model.

In particular, the blood glucose prediction model is the basis of predictive control of the artificial pancreatic system. So it is important to find an accurate blood glucose prediction model. Due to the influence of factors such as equipment and external environment, the blood glucose data collected by the continuous glucose monitoring (CGM) is non-linear and non-stationary, which seriously affects the accuracy of blood glucose prediction. In this study, the VMD decomposition method was used to stabilize the time series of blood glucose values in time-varying nonlinear and non-stationary diabetes patients. The fluctuation or trend of blood glucose values at different scales in the time series was decomposed step by step to produce a series of blood glucose components with different frequency bands and then predicted separately. Due to the influence of equipment, external environment and other factors, the blood glucose data collected by CGMs are highly non-linear and non-stationary, which seriously affects the accuracy of blood glucose prediction. The proposed method uses the VMD method to smooth the blood glucose value time series of diabetic patients. The fluctuations or trends in the time series of blood glucose values at different scales are decomposed step by step to produce a series of blood glucose components with different frequency bands, which are then predicted separately. The parameters to be optimized in the prediction process of the LSTM network are large in size, which leads to slow convergence of the algorithm in the optimization process and makes it easy to fall into the phenomenon of local optimization. We propose to use an improved PSO algorithm to optimize the key parameters of the network and improve the prediction stability of the network. The experimental results show that

1) The combination of LSTM prediction model and VMD decomposition method can effectively improve the accuracy of blood glucose concentration prediction. The VMD method can effectively solve the modal overlap problem of CGM data, reduce the non-stationary characteristics of the original series, and mine the data features in depth.

2) The improved Particle Swarm Optimization (IPSO) has better parameter finding ability, which can effectively solve the problem of difficult selection of learning parameters in LSTM and greatly improve the prediction accuracy of glucose concentration prediction model.

The proposed VMD-IPSO-LSTM short-term blood glucose prediction model can predict blood glucose 30min, 45min, and 60min in advance with better prediction accuracy and has better prediction ability than the LSTM prediction model, VMD-LSTM prediction model, and VMD-PSO-LSTM prediction model.

Although the method in this paper has some advantages in terms of prediction accuracy and extended glucose prediction time, there are still some limitations. Firstly, in the proposed method, the parameters of LSTM are optimized using an improved Particle Swarm Optimization, which leads to an increase in computation volume and an increase in running time. Short-term glucose prediction run times to be reduced. Therefore, how to integrate the idea of gradient method into
the Particle Swarm Optimization, adopt the velocity monitoring strategy to improve the optimization of LSTM parameters and reduce the running time will be an important direction for further research work. In addition, in this paper, we only rely on CGM data for short-term prediction of blood glucose concentration in the future without considering more relevant factors, such as age, gender, liver function, etc. How to combine more physical characteristics to make medium-term and long-term prediction of patients’ blood glucose concentration will also be the focus of our future research.

REFERENCES

[1] A. Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina, T. Bottis, and G. Hartvigsen, “Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes,” Artif. Intell. Med., vol. 98, pp. 109–134, Jul. 2019. doi: 10.1016/j.artmed.2019.07.007.

[2] A. Ceriello, A. Novials, A. Ortega, L. La Sala, G. Pujadas, R. Testa, A. R. Bonfigli, K. Esposito, and D. Giugliano, “Evidence that hyperglycemia after recovery from hypoglycemia worsens endothelial function and increases oxidative stress and inflammation in healthy control subjects and subjects with type 1 diabetes,” Diabetes, vol. 61, no. 11, pp. 2993–2997, 2012. doi: 10.2337/db12-0224.

[3] X. Yu, K. Turksoy, M. Rashid, J. Feng, N. Hobbs, I. Hajizadeh, S. Samadi, M. Sevil, C. Lazaro, Z. Maloney, E. Littlejohn, L. Quinn, and A. Cinar, “Model-fusion-based online glucose concentration predictions in people with type 1 diabetes,” Control Eng. Pract., vol. 71, pp. 129–141, Feb. 2018, doi: 10.1016/j.conengprac.2017.10.013.

[4] Y. Wang, F. Wei, C. Sun, and Q. Li, “The research of improved grey GM (1, 1) model to predict the postprandial glucose in type 2 diabetes,” Biomed Res. Int., vol. 2016, pp. 1–6, Jan. 2016. doi: 10.1155/2016/6837052.

[5] X. Mo, Y. Q. Wang, and X. W. Wu, “Hypoglycemia prediction using extreme learning machine (ELM) and regularized ELM,” in Proc. 25th Chin. Control Decis. Conf. (CCDC), Guiyang, China, 2013, pp. 4405–4409.

[6] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, “Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models,” Med. Biol. Eng. Comput., vol. 52, no. 12, pp. 1305–1318. Dec. 2015.

[7] S. Shanthi, “A novel approach for the prediction of glucose concentration in type 1 diabetes ahead in time through ARIMA and differential evolution,” Adv. Eng. Inform., vol. 38, pp. 4182–4186. Aug. 2017.

[8] J. Yang, L. Li, Y. Shi, and X. Xie, “An ARIMA model with adaptive orders for predicting glucose concentrations and hypoglycemia,” IEEE J. Biomed. Health Inform., vol. 23, no. 3, pp. 1251–1260. May 2019.

[9] E. Daskalaki, A. Prountzou, P. Diem, and S. G. Mougiakakou, “Real-time adaptive models for the personalized prediction of glycemic profile in type 1 diabetes patients,” Diabetes Technol. Therapeutics, vol. 14, no. 2, pp. 168–174. Feb. 2012.

[10] R. Bunescu, N. Struble, C. Marling, J. Shubrook, and F. Schwartz, “Blood glucose level prediction using physiological models and support vector regression,” in Proc. 12th Int. Conf. Mach. Learn. Appl., Dec. 2013, pp. 135–140.

[11] E. I. Georga, V. C. Protopappas, and D. I. Fotiadis, “Glucose prediction in type 1 and type 2 diabetic patients using data driven techniques,” Knowl.-Oriented Appl. Data Mining, pp. 277–296. 2011.

[12] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, “A deep learning approach to diabetic blood glucose prediction,” Frontiers Appl. Math. Statist., vol. 3, pp. 1–11. Jul. 2017.

[13] H. M. Fayek, M. Lech, and L. Cavedon, “Evaluating deep learning architectures for speech emotion recognition,” Neural Netw., vol. 92, pp. 60–68. Aug. 2017.

[14] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[15] A. Dehimi, O. Orang, and H. Showkat, “Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction,” Energy, vol. 57, pp. 382–401. Aug. 2013.

[16] T. Fischer and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions,” Eur. J. Oper. Res., vol. 270, no. 2, pp. 654–669. Oct. 2018.

[17] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, “Social LSTM: Human trajectory prediction in crowded spaces,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 961–971.

[18] Q. Zhang, H. Wang, C. Dong, G. Zhong, and X. Sun, “Prediction of sea surface temperature using Long Short-term memory,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 10, pp. 1745–1749. Oct. 2017.

[19] Y. Zhu, X. Fan, J. Wu, X. Liu, J. Shi, and C. Wang, “Predicting ICU mortality by supervised bidirectional LSTM networks,” in Proc. 1st Joint Work- shop AI Health, Organized Part Federated AI Meeting AI Health, Organized Part Federated AI Meeting. Dec. 2018, pp. 1–5.

[20] A. Aliberti, I. Pippollo, S. Terna, E. Macii, S. Di Cataldo, E. Patti, and A. Acquaviva, “A multi-patient data-driven approach to blood glucose prediction,” IEEE Access, vol. 7, pp. 69311–69325. 2019.

[21] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” IEEE Trans. Signal Process., vol. 62, no. 3, pp. 531–544. Feb. 2014.

[22] W. Hu and Z. S. Li, “A simpler and more effective particle swarm optimization algorithm,” J. Softw., vol. 18, no. 4, pp. 861–868. 2007.

[23] A. M. Rather, A. Agarwal, and V. N. Sastry, “Recurrent neural network and a hybrid model for prediction of stock returns,” Expert Syst. Appl., vol. 42, no. 6, pp. 3234–3241. Apr. 2015.

[24] Journal of Community Health Research. Diabetes Research Studies. Accessed: Oct. 2018. [Online]. Available: http://diabetes.jaeb.org/.

[25] A. M. Rather, A. Agarwal, and V. N. Sastry, “Recurrent neural network and a hybrid model for prediction of stock returns,” Expert Syst. Appl., vol. 42, no. 6, pp. 3234–3241. Apr. 2015.

[26] Journal of Community Health Research. Diabetes Research Studies. Accessed: Oct. 2018. [Online]. Available: http://diabetes.jaeb.org/.

[27] W. Hu, C. Lazaro, Z. Maloney, E. Littlejohn, L. Quinn, and A. Cinar, “Model-fusion-based online glucose concentration predictions in people with type 1 diabetes,” Control Eng. Pract., vol. 71, pp. 129–141. Feb. 2018.

[28] L. Wang, Y. Liu, T. Li, X. Xie, and C. Chang, “Short-term PV power prediction based on optimized VMD and LSTM,” IEEE Access, vol. 8, pp. 165849–165862. 2020.