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

IEEE ACCESS 2020

Limitation and Motivation

Limitation

- LSTM networks have proven effective in forecasting across various domains, especially in blood glucose forecasting

- Blood glucose concentration time series is:
    - Highly time-varying
    - Non-linear
    - Non-stationary

- Direct application of LSTM to blood glucose prediction may impact accuracy
- Determining optimal LSTM network parameters presents a challenge

Motivation

- Multi-scale decomposition can mitigate the effects of non-stationarity on prediction results

Novelty

VMD-IPSO-LSTM for blood glucose prediction

1. Decompose CGM blood glucose time series using VMD
2. Obtain Intrinsic Mode Functions (IMFs) in different frequency bands
3. Establish LSTM prediction model for each IMF sequence
4. Optimize LSTM parameters using Improved Particle Swarm Optimization (IPSO)

Benefits

- VMD effectively reduces non-stationary changes in blood glucose concentration
- IPSO with adaptive learning strategy optimizes key LSTM parameters
- Optimization matches IMF sequence characteristics with network topology
- Improves prediction accuracy of IMF sequences

Variational Mode Decomposition (VMD)

The decomposition process of VMD can be understood as the process of finding the optimal solution to the variational problem. 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) $U_k(t)$.

$$\begin{cases} \min\limits_{\{u_k,\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\| \right\} \\ s.t. \sum_{k=1}^{K} u_k = f \end{cases} \tag{3}$$

$$\hat{u}_k^{n+1}(\omega) = \left(\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right) \times \frac{1}{1 + 2\alpha(\omega - \omega_k)^2} \tag{5}$$

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 w is updated by Equation (6).

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \tag{6}$$

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

Step 1, initialize $\{\hat{u}_k^1\}$, $\{\hat{\omega}_k^1\}$, $\hat{\lambda}^1$ and $n = 0$;

Step 2, recursively deduce $n \leftarrow n + 1$, and update u_k and ω_k; according to formula (5) and formula (6);

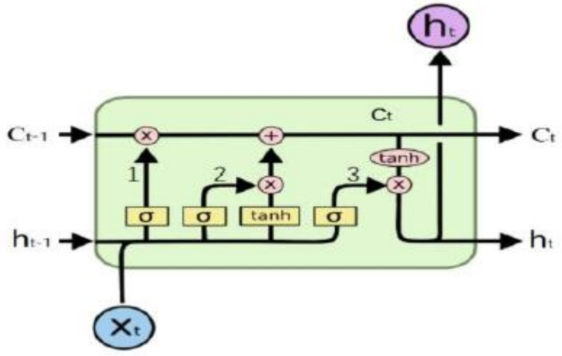
Step 3, update λ;

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + [\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)] \tag{7}$$

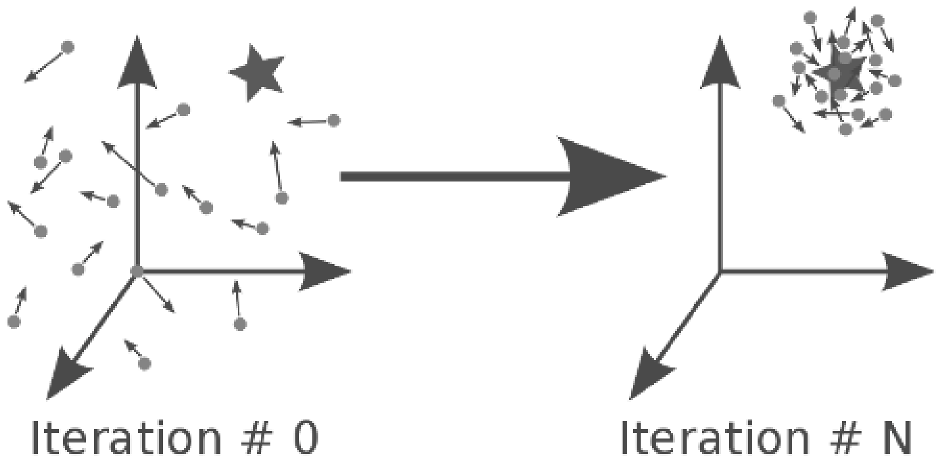
Step 4, if $\sum_k \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \left\| \hat{u}_k^n \right\|_2^2 < \varepsilon$, where ε is the discrimination accuracy, if $\varepsilon > 0$, stop the iteration, output the results, and get k modal components and their center frequencies; otherwise, return to step 2.

LSTM and PSO

- **LSTM:** is a special RNN, which learns long-term dependent information and avoids the problem of gradient disappearance



- **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.



PSO equation

- The position and velocity of the i^{th} particle are $X_{i,t}$ and $V_{i,t}$. The particle updates its position and velocity by supervising two optimal solutions.

$$V_{i,t+1} = w^* V_{i,t} + c_1^* rand^* (pbest_i - X_{i,t}) + c_2^* rand^* (gbest_t - X_{i,t}) \quad (15)$$

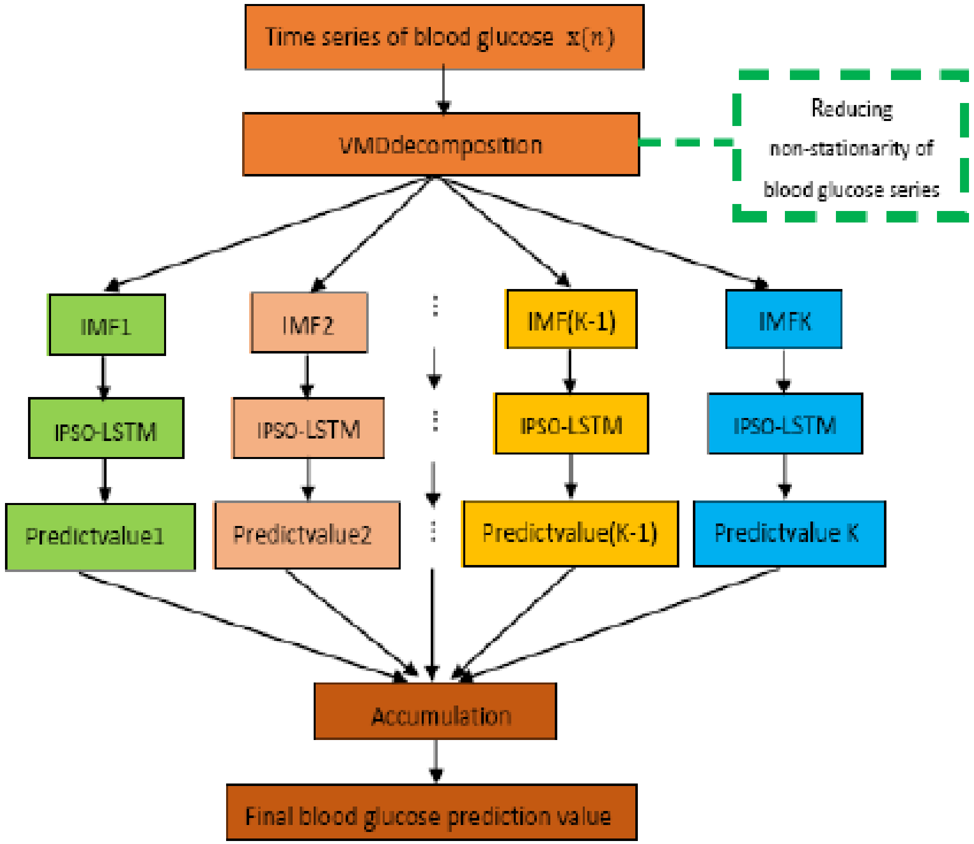
$$X_{i,t+1} = X_{i,t} + \lambda^* V_{i,t+1} \quad (16)$$

Where, personal best $pbest_i$, global best $gbest_t$

In basic PSO, fixed w will weaken the global optimization ability and slow down the convergence speed of the algorithm.

$$w = w_{\max} - (w_{\max} - w_{\min})^* \arcsin \frac{t}{t_{\max}} * \frac{2}{\pi} \quad (17)$$

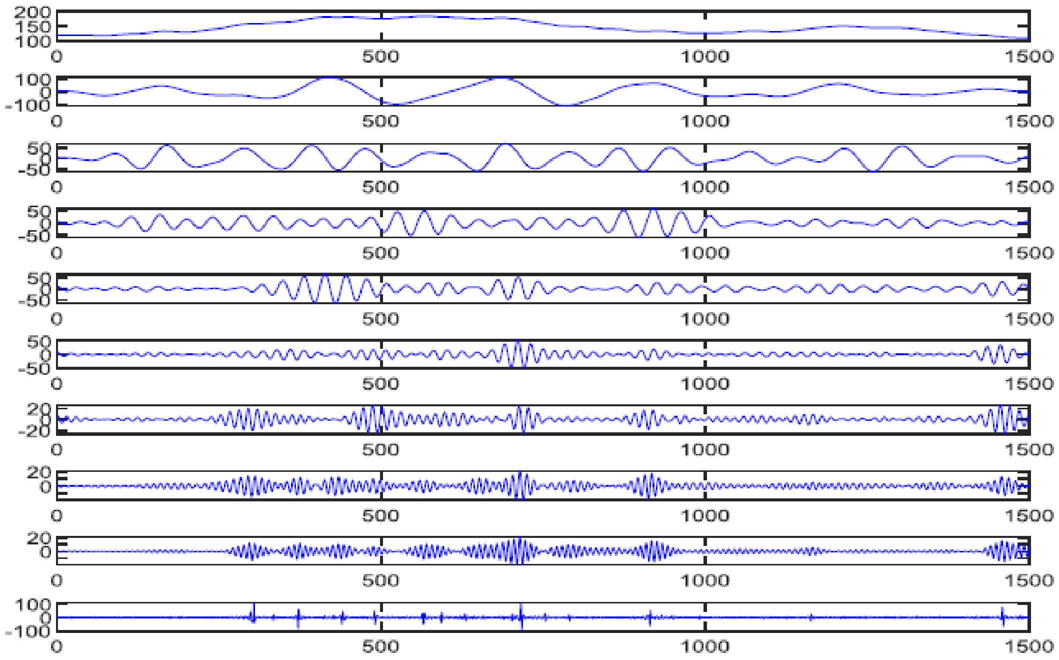
Proposed Method



Dataset and VMD example

- The data for this article comes from the RT_CGM dataset, which is freely available for study. This data set included glucose trends in 451 heterogeneous populations affected by type I diabetes.

VMD example



Forecasting with different methods

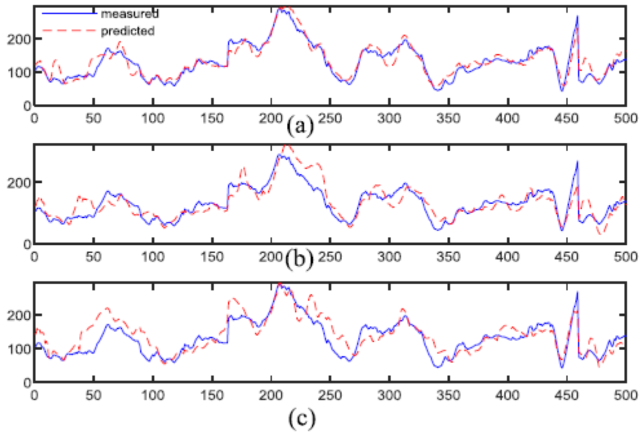


FIGURE 4. LSTM predictions at different forecasting horizons. (a) prediction (30 min horizon). (b) prediction (45 min horizon). (c) prediction (60 min horizon).

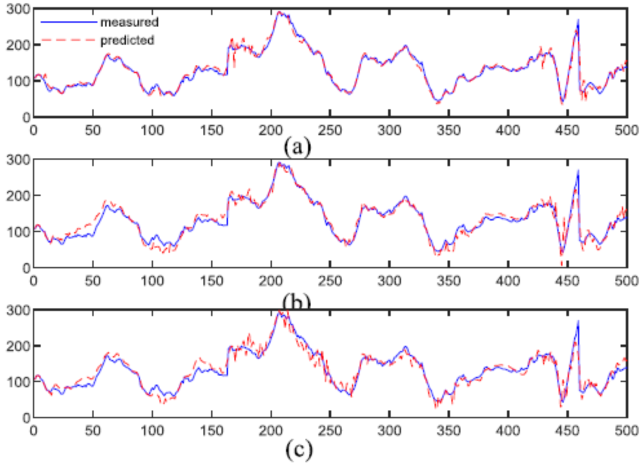


FIGURE 5. VMD-LSTM predictions at different forecasting horizons. (a) prediction (30 min horizon). (b) prediction (45 min horizon). (c) prediction (60 min horizon).

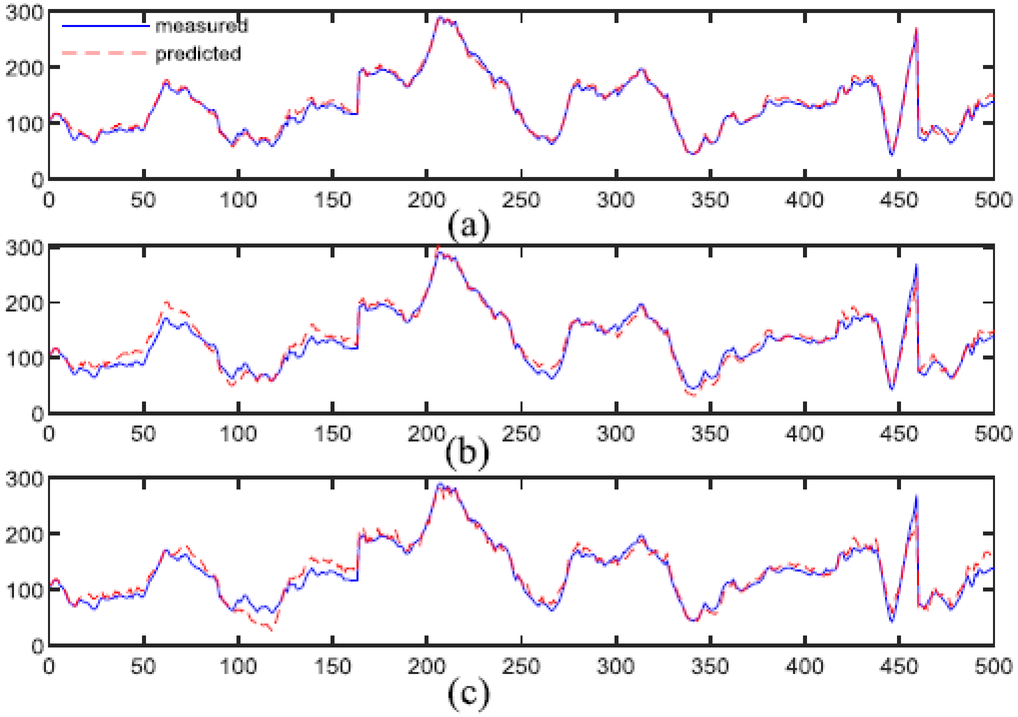


FIGURE 7. VMD-IPSO-LSTM predictions at different forecasting horizons. (a) prediction (30 min horizon). (b) prediction (45 min horizon). (c) prediction (60 min horizon).

Results

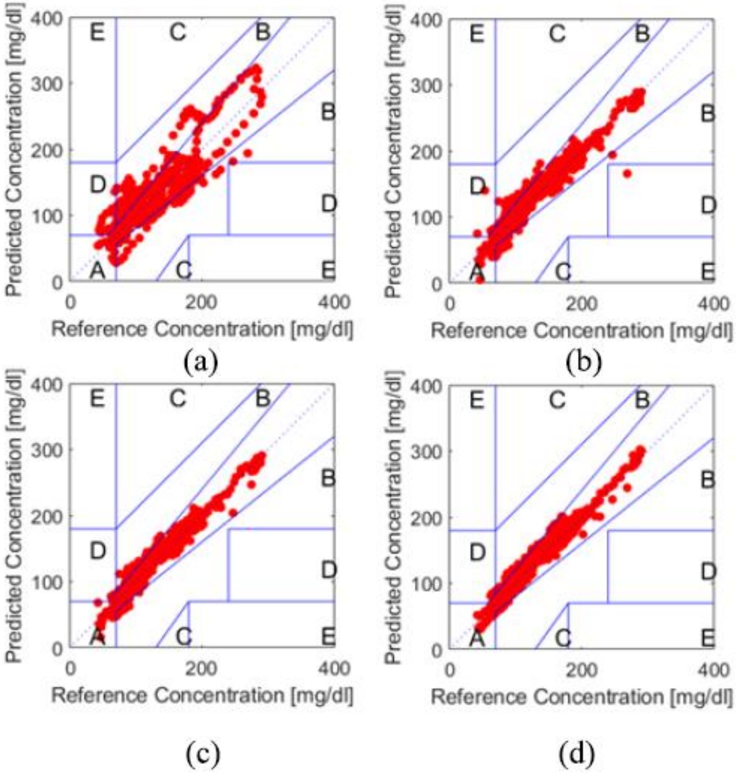


FIGURE 8. **Clarke error grid diagrams of four models (a) LSTM model with 60 min; (b) VMD-LSTM model with 60 min; (c) VMD-PSO-LSTM model with 60 min; (d) VMD-IPSO-LSTM model with 60 min.**

Model	30 min advance		45 min advance		60 min advance	
	RMSE	MAPE /%	RMSE	MAPE /%	RMSE	MAPE /%
LSTM	12.330	8.565	19.031	12.750	21.281	15.433
VMD-LSTM	5.565	3.152	8.281	5.799	9.118	6.173
VMD-PSO-LSTM	3.887	2.541	6.429	4.739	6.931	4.983
VMD-IPSO-LSTM	**3.031**	**2.219**	**5.423**	**4.006**	**5.716**	**4.149**