Chaotic Time Series Analysis Approach for Prediction Blood Glucose Concentration Based on Echo State Networks

Ning Li[1], Jianyong Tuo[1*], Youqing Wang[2,1]

1. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029

2. College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266000

E-mail: tuojy@mail.buct.edu.cn

Abstract: Blood glucose prediction plays a very critical role in the treatment of diabetes. With the development of continuous glucose monitoring system (CGMS), it becomes possible to know the blood glucose level at real time. In this literature, we establish a predictive model using echo state neural networks (ESN) due to its excellent performance in chaotic time series forecasting. In order to further improve the performance of the network, we optimized the ESN with leakage integral neurons and ridge regression learning algorithm. Under the same condition, the proposed method is compared with the Extreme Learning Machine and Back Propagation algorithm in terms of Root mean square error (RMSE), Time gain (TG) and the Continuous glucose-error grid analysis (CG-EGA). The experimental results demonstrate that ESN is a very suitable prediction model for blood glucose time series.

Key Words: blood glucose prediction, continuous glucose monitoring system (CGMS), echo state networks (ESN), optimized ESN, suitable prediction model.

1. INTRODUCTION

Diabetes is a chronic disease characterized by the body's inability to produce enough insulin or insulin deficiency used. According to the International Diabetes Federation (IDF), in 2015 [1], there were 415 million people with diabetes in the word, making diabetes one of the largest global health emergencies of 21st century. Specifically, Type 1 diabetes (T1D) is also increasing for β-cells abnormalities. Generally, T1D have a highly mortality rate for hyperglycemia and some complication (i.e. neuropathy, retinopathy, stroke and arterial disease). Hence, maintaining blood glucose levels in the normal range (70-150 mg/dl) is of dramatic concern in the clinical therapy. A classical medical solution [2] to patient is the use of multiple doses of insulin injections (generally 2-3 injections per day) that can lead to tight glycemic control. In recent years, with the continuous glucose monitoring system (CGMS) and insulin pump technology matures makes it possible to cure diabetic by using artificial pancreas. To successful manage blood glucose levels, the prediction algorithm plays a major role in the design of artificial pancreatic engineering.

In 1977 [3], Alibisser proposed the concept of continuous glucose monitoring system and successfully used in clinical patients. At present, multiple sensors system has been developed, which allows CGMS for several days. CGM systems are noninvasive or minimally invasive, and in many cases, the fact that they are portable that can allow patients to use in daily life. Meanwhile, CGM sensors allow the fine monitoring of glycemic concentration in a quasi "continuous" way (e.g. one every 1-5 mins) for several days (up to 14). Collection of CGM data not only enables alerting the hypo/hyperglycemia but also develop the mathematical forecasting models.

Since the introduction of CGM devices, several short-time glucose prediction methods have been proposed in recent years, including the popular time series models that use only the CGM signal as an input. In [4], Gani et al, exploited the autoregressive (AR) model based on 34 patients for thirty minutes. The results showed that the AR model is a good linear regression method for prediction. The works of Sparacino et al. were compared with [5] that proposed a first-order ARM prediction model in case of the same conditions. Obviously, a conclusion can be obtained that ARM is a more accurate prediction strategy. Palerm [6] proposed the Kalman filter method to forecast blood glucose levels based on its degree of change. They used data from CGM to predict glucose concentration, and they initialized variable prediction horizons from 1 to 30 min. For better performances, Wang et al [7] proposed a novel framework with AR model, extreme learning machine (ELM) and support vector regression (SVR) for blood glucose prediction. Experimental results showed that the adaptive-weighted algorithm achieved the best prediction performance than alone. Moreover, artificial neural networks method has also been widely used in the prediction of blood glucose. For example, Pérez-Gandía

This work is supported by National Nature Science Foundation under Grant 61374099 and Research Fund for the Taishan Scholar Project of Shandong Province of China.

et al, [8] has mentioned that if the glucose time-series are known, future blood glucose can be anticipated with feedforward neural network. Then, a recurrent neural networks approach [9] was proposed for predicting glucose concentration. The result of the proposed technique is evaluated and compared relative to that obtained from a feed forward neural networks prediction model (NNM). Even though RNN achieves good prediction, there is a problem that the algorithm is complicated to calculate.

In a recent study, Mirela Frandes et al. demonstrated that the glycemic time series has chaotic characteristics [10]. So in this article, we try to use an ESN that performs well in chaotic time prediction [11, 12]. ESN as the state-of-the-art recurrent neural network, it gains a reservoir of dynamics by training output units with a simple and fast single-stage training process. Its innovation is the computing process which can not only remember the state but also can fully stimulate the nonlinear state. Therefore, we can use simple linear regression to calculate the output weights. However, linear regression algorithms commonly used to solve practical problems tend to produce large output weights and prone to over-fitting. Therefore, based on the leakage integral neuron (Li-ESN) and ridge regression learning algorithm are used [13] to calculate the output weight. To the best of our knowledge, this is the first time that the ESN network are used for blood glucose prediction.

In this paper, ESN and optimized ESN are all first applied to blood glucose prediction. The database are obtained by real CGM of different patients. To verify the performance of ESN and optimized ESN, the comparative experiments was done with ELM and Back Propagation algorithm (BP) by use the [14] Root mean square error (RMSE), Time gain (TG) and Clarke error-grid analysis (CE-GA). We analyze the effect of prediction accuracy and time gains via behind the chart and table.

In a word, this paper has the significant contribution that is the first work to application ESN and optimized ESN on blood glucose prediction. The rest of the paper is structured as follows: Section 2 introduces ESN method and its improvements, Section 3 presents in detail that the forecasting strategy and algorithm training process, Section 4 includes the data recording and the experimental results discussion, Finally, conclusions are given in Section 5.

2. Overview of ESN and Its Variants

ESN is a new type of recurrent neural network. Due to the unique structure and simple training algorithm, ESN has been widely used in various fields. A brief description on principles and improvement of ESN is given in the following.

2.1 Standard Echo State Networks

ESN proposed by Jaeger H in [15] is one of the most plausible recurrent neural network topology to represent the human brain. An ESN consists of the input layer, the internal reserve pool and the output layer. It includes a unfixed dynamical reservoir of randomly connected neurons in hidden layer generated before training stage. The neurons of the middle unit are arbitrarily connected to each other. Figure 1 shows the network structure of the ESN [16].

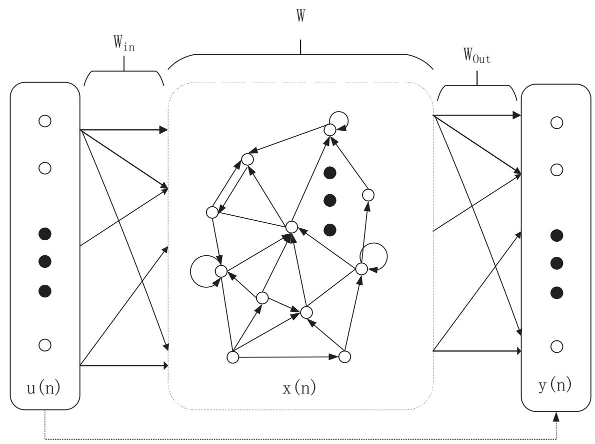

Figure 1: Echo state network structure.

First, assume that the number of neurons input layer, hidden layer and output layer are k , N , L respectively . the network can be calculated as follows at the moment n .

$$\begin{cases} u(n) = (u_1(n), u_2(n), \cdots u_k(n),) \\ x(n) = (x_1(n), x_2(n), \cdots x_N(n),) \\ y(n) = (y_1(n), y_2(n), \cdots y_L(n),) \end{cases} \tag{1}$$

Then, we need to initialize the weights W_{in} , W , W_{fy} , and reservoir of parameters ρ before the network training. The state of the neurons in the networks can be written as follows:

$$x(n+1) = f(W_{in}u(n+1) + \rho Wx(n) + W_{fy}y(n)) \tag{2}$$

$$y(n+1) = f^{out}\left(W_{out}\begin{bmatrix} u(n+1) \\ x(n+1) \\ y(n) \end{bmatrix}\right) \tag{3}$$

For nonlinear prediction problems, where $W_{fy} = 0$ $f = [f_1, f_2, \cdots f_N]^T$ are the excitation functions. The f^{out} can be hyperbolic tangent functions. In addition, only when the spectral radius ρ must be satisfied $0 \le \rho < 1$ the network is stability.

2.2 Optimized Echo State Networks

A. Leaky Integrate ESN

Li-ESN is a similar biological neuron that uses distributed time coding and information processing. That leaks incentive [17] to collect networks at an exponential rate. Now, continuous dynamic time $x(t)$ can be mathematically modeled as

$$\frac{dx(t)}{dt} = C(-\alpha x(t) + f(W_{in}u(t) + \rho W x(t)) \quad (4)$$

The parameter C is the time constant and α is the leak rate. The Euler transformation of (4) gives the discrete state equation as follows:

$$x(n+1) = (1 - C\alpha)x(n) + C(f(W_{in}u(n+1) + \rho W x(n))) \quad (5)$$

The leakage rate can be controlled to maintain the neuron state at the previous moment by leakage integral algorithm which making each neuron in the reservoir with low-pass filter and exponential smooth characteristics. In particular, when the reservoir size is reduced and asking improving the accuracy for time series prediction problem, Li-ESN is much better than the classic ESN.

B. Ridge Regression ESN

ESN trends to result in over-fitting model when only using linear regression algorithm in the case of a small number of samples. Ridge regression was proposed by Hoerl in [18] to optimize the linear regression problem. Consider the regularized model for multiple linear regressions.

$$W_{out} = (X^T X + \lambda I)^{-1} X^T y \quad (6)$$

Here λ is a variable from 0.1 to 1. In this paper, we make λ equal to 0.8. Moreover, the equation of the cost function is minimized as follows.

$$L(w, \mathrm{b}) = \mathrm{C}\sum_{j}^{T}\left(\left(wx_j\right) + b - y_{dj}\right)^2 + \|w\|^2 \quad (7)$$

3. Design of Forecasting Strategy

In this section, due to the blood glucose time series proved to have chaotic characteristics, we propose a novel prediction algorithm based on an ESN for CGM data. The model is trained using only CGM data as inputs. The flow chart for the proposed algorithm is shown in Figure 2.

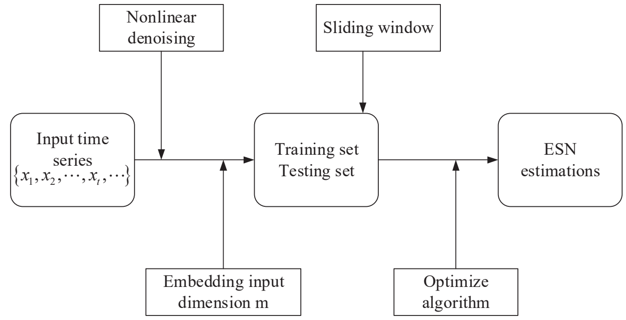

Figure 2: The blood glucose time series prediction framework.

After the data was collected, a non-linear filtering process was first performed to remove the effects of noise. In order to obtain the optimal input dimension and a more accurate prediction model, we calculated the fitting degree [19] of the related information for the original time series. Then the ESN structure was optimized by using leaky integrate neurons and optimized training algorithm. Finally, we can get an ESN estimation model.

Algorithm process of ESN:

- **Step1:** The input weights W_{in} and the parameters (N , W , SD) of reservoirs are generated fixed randomly.
- **Step2:** The pool loads the sample data and records and updates the status using leak integrator neurons, then outputs the status of savings.
- **Step3:** Calculate the output weight according to ridge regression algorithm.

4. Experimental Results and Discussion

In order to verify the effectiveness of our proposed method, we have done a simulation experiment based on MATLAB 2016. We have done a clinical trial of Artificial Pancreas (AP) in FUWAI Hospital and recorded data. The database consists of eight clinical patients equip with Deccom Seven PLUS (Sample time every five minutes) under free-living conditions. Each subject can generate 289 values daily and records continuously for three days. For each patient, we used two-thirds of his recorded blood glucose data for the training step and used the remaining part of glycemic data to test the predictions. In this paper, three performance indexes commonly used in literature were used to test the performance of the ESN and compare it with that of the reference others methods. Specific index of its mathematical expression is as follows:

- Root mean square error (RMSE, mg/dl) between predicted and target values:

$$RMSE = \sqrt{\frac{1}{K}\sum_{i=N+1}^{K}(y(i) - \hat{y}(i|i-N))} \quad (8)$$

Where K representing the length of the time series, RMSE represents the fitting of the overall time series, which was referring to the degree of fluctuation of the predicted value and the actual value of the whole. Hence, the smaller the RMSE value, the higher the prediction accuracy. The RMSE allows the error to be as same magnitude as the quantity being predicted.

- Time gain (TG, min)

$$delay = \arg\min_{j\in[0,N]} \left\{ \frac{1}{K} \sum_{i=1}^{K-N} (y(i) - \hat{y}(i+j|i+j-N))^2 \right\} \quad (9)$$

$$TG = PH - delay \bullet T_S \quad (10)$$

Here T_S is the sampling interval of the time series. The index [20] could be used to quantify the degree of delay predicted. Obviously, the TG value and the predicted track is directly proportional effect.

- The Continuous glucose-error grid analysis [21] (CG-EGA) is also widely used to estimate the prediction performance. In summary, A and B represent the clinically acceptable regions. If the predicted value is located in other regions, the prediction effect is poor.

4.1 Result

Based on the experiments, the ESN model parameters for input nodes, intermediate nodes and the sparse degree of reservoir are set to 6, 20, and 0.3, respectively. The forecasting results for 20 minutes (e.g. 4 steps in advance forecast) after by using standard ESN shown in Figure 3.

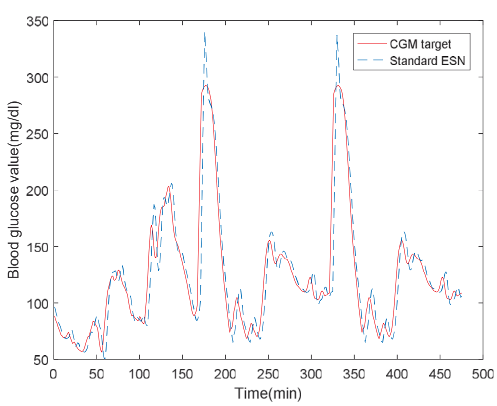

Figure 3: The glucose prediction results for patient 1 under the standard ESN (prediction horizon =20), where the real CGM data and the predicted result are represented by the red line and blue, respectively.

In the current figure, the prediction curve has been accompanied by concussion and is not smooth enough, especially in the peak position of strong vibration. This could be explained by the limited quantity of hyperglycemia in the training data. In order to get a more accurate prediction curve, we optimized the ESN and conducted experiments under the same conditions as follow in Figure 4.

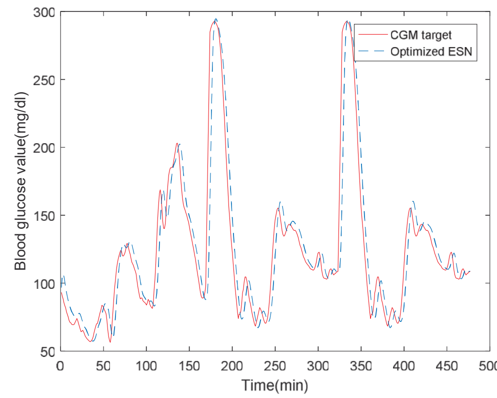

Figure 4: The glucose prediction results for patient 1 under the Optimized ESN (prediction horizon =20).

As shown in Figure 4, the prediction results tend to be smooth. Compared with the standard ESN, the optimized ESN overcomes the unexpected problem at the peak. The most likely explanation for this phenomenon is that neurons forget the state of the previous moments by controlling the leak rate.

To further assess the prediction performances of proposed approach, Figure 5 presents a comparison between the actual different prediction algorithms (e.g. BP, ELM).

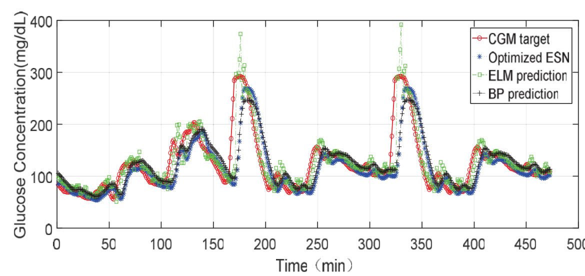

Figure 5: Prediction performance of subject 2 produced by three algorithms (prediction horizon =30).

As seen in Fig. 5 the optimized ESN strategy was compared with ELM and BP, different prediction algorithms are described by different colors respectively, the experimental result shows that neither ELM model nor the BP model forecasts as accurately as the optimized ESN model. This interpretation future emphasizes the validity of the optimized ESN.

4.2 Discussion

Obviously, it is difficult to determine which algorithm is superior from the graph. In order to make the results more clear, eight different patients were randomly chosen to compare with different algorithms. The average results of all subjects are shown in Table 1.

Table 1: The average performance comparison of different algorithms

Index \ Method	ESN	Optimized ESN	ELM	BP
RMSE[mg/dl]	16.32	15.93	19.64	28.77
TG[min]	19.5	18.5	15	10
% in A-zone	81	95	75	73

Table 1 summarize that the model prediction accuracy of the four different methods for eight subjects with evaluation index. RMSE reflects an average level of oscillations between the target and predicted values. It is noted that the larger the TG, the smaller the prediction delay. And, region A represents the confidence of the predicted value. As we know, the prediction horizon is closely related to the accuracy of the glucose concentration prediction. The prediction result is usually worse as the prediction horizon is longer. So, we have made a performance comparison that different methods are under different forecast time. The statistics results are given in Table 2.

Table 2: The performance comparison among different algorithms on different prediction horizon

Method \ Index	RMSE[mg/dl]			TG[min]			% in A-zone		
	15 min	45 min	60 min	15 min	45 min	60 min	15 min	45 min	60 min
Optimized ESN	12	36.35	52.05	15	42.5	50	96.62	74.12	69.12
ESN	15.2	37.23	59.25	15	40	47.5	94.25	73.06	46.58
ELM	17.5	49.37	109.25	13.5	40	37.5	96	61.25	48.75
BP	22.5	40.37	49	14	38.5	45	82.87	51.12	40.62

As shown in Table 2, the fluctuation of RMSE is relatively small in terms of the proposed models, among them ESN is from 15.2 to 52.05 and the range of Optimized ESN varies from 12 to 52.05. However, the four methods have a considerable effect on the TG index. For the CG-EGA index, the first three methods have similar effects and are clearly superior to than the other. In summary, the prediction accuracy of the optimized ESN is more advantageous than the other methods when using CGM data only. Although predictions will be worse with longer forecast times, the predictions of ESN and Optimized ESN are accurate over other algorithms.

The reason that is the reservoir has well non-linear approximation ability and memory ability by adjusting the adaptive parameters.

5.Conclusions

From the results that were given above previously, we can see that the ESN and its optimization prediction models succeed to predict the future glucose values. Additionally, according to the experimental results, the improved ESN has better stability than the standard ESN in peaks. Moreover, we applied different algorithms to make the comparison based on many patients and different prediction horizon. In terms of values of the Tables the similar conclusion may be drawn, which optimized neural network can achieve the best prediction results.

Due to several parameters of the model are obtained by the experiment does not have a general, so further studies will focus on parameter optimization. Besides, we put more attention on the research of personalized model.

REFERENCES

[1] K. G. Alberti, "Diabetes and the National Health Service Research and Development Program," *Diabetic Medicine A Journal of the British Diabetic Association,* vol. 11, p. 526, 1994.

[2] L. C, W. K, H. KJ, and H. SM, "The National Cholesterol Education Program - Adult Treatment Panel III, International Diabetes Federation, and World Health Organization definitions of the metabolic syndrome as predictors of incident cardiovascular disease and diabetes," *Diabetes care,* vol. 30, p. 8, 2007.

[3] A. M. Albisser, W. Lougheed, K. Perlman, and A. Bahoric, "Nonaggregating insulin solutions for long-term glucose control in experimental and human diabetes," *Diabetes,* vol. 29, pp. 241-3, 1980.

[4] A. Gani, A. V. Gribok, Y. Lu, W. K. Ward, R. A. Vigersky, and J. Reifman, "Universal glucose models for predicting subcutaneous glucose concentration in humans," *IEEE Transactions on Information Technology in Biomedicine A Publication of the IEEE Engineering in Medicine & Biology Society,* vol. 14, p. 157, 2010.

[5] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, and C. Cobelli, "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series," *IEEE transactions on bio-medical engineering,* vol. 54, p. 931, 2007.

[6] A. Facchinetti, G. Sparacino, and C. Cobelli, "Enhanced accuracy of continuous glucose monitoring by online extended kalman filtering," *Diabetes Technology & Therapeutics,* vol. 12, pp. 353-63, 2010.

[7] Y. Wang and X. Wu, "An adaptive glucose prediction method using auto-regressive (AR) model and Kalman filter," in *Ieee-Embs International Conference on Biomedical and Health Informatics*, 2012, pp. 293-296.

[8] C. Pérez-Gandía, A. Facchinetti, G. Sparacino, C. Cobelli, E. J. Gómez, M. Rigla*, et al.*, "Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring," *Diabetes Technology & Therapeutics,* vol. 12, pp. 81-88, 2010.

[9] V. Tresp and T. Briegel, "A solution for missing data in recurrent neural networks with an application to blood glucose prediction," in *International Conference on Neural Information Processing Systems*, 1997, pp. 971-977.

[10] M. Frandes, B. Timar, R. Timar, and D. Lungeanu, "Chaotic time series prediction for glucose dynamics in type 1 diabetes mellitus using regime-switching models," *Scientific Reports,* vol. 7, 2017.

[11] S. Tong, "Chaotic time series prediction based on wavelet echo state network," *Acta Physica Sinica,* vol. 61, pp. 963-963, 2012.

[12] M. Xu, M. Han, and J. Wang, "Multivariate Chaotic Time Series Prediction Using a Wavelet Diagonal Echo State Network," in *Second International Conference on Mathematics and Computers in Sciences and in Industry*, 2016, pp. 86-92.

[13] S. X. Lun, X. S. Yao, H. Y. Qi, and H. F. Hu, "A novel model of leaky integrator echo state network for time-series prediction," *Neurocomputing,* vol. 159, pp. 58-66, 2015.

[14] C. J. Willmott and K. Matsuura, "Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance," *Climate Research,* vol. 30, p. 79, 2005.

[15] H. Jaeger, "The"echo state"approach to analysing and training recurrent neural networks," *überwachtes lernen,* 2001.

[16] S. Mukherjee, E. Osuna, Girosi, and F., "Nonlinear Prediction of Chaotic Time Series using a Support Vector Machine," pp. 511-520, 2008.

[17] H. Hu and S. Lun, "Optimize state equation of leaky-integrator ESN based on extend kalman filtering," in *Chinese Automation Congress*, 2015, pp. 787-791.

[18] Hoerl, E. Arthur, Kennard, and W. Robert, "A Note on a Power Generalization of Ridge Regression," *Technometrics,* vol. 17, pp. 269-269, 1975.

[19] J. M. Mcnamara and S. R. X. Dall, "Information is a fitness enhancing resource," *Oikos,* vol. 119, pp. 231-236, 2010.

[20] A. Facchinetti, G. Sparacino, E. Trifoglio, and C. Cobelli, "A new index to optimally design and compare continuous glucose monitoring glucose prediction algorithms," *Diabetes Technology & Therapeutics,* vol. 13, p. 111, 2011.

[21] K. BP, G.-F. LA, C. DJ, and C. WL, "Evaluating the accuracy of continuous glucose-monitoring sensors: continuous glucose-error grid analysis illustrated by TheraSense Freestyle Navigator data," *Diabetes Care.,* vol. 27, pp. 1922-1928, 2004.