Glucose Level Prediction Based on Data Driven Method

Min Qian
Department of Industrial Engineering
Tsinghua University
Beijing, China
E-mail: qianmin.tsinghua@outlook.com

Yan-Fu Li
Department of Industrial Engineering
Tsinghua University
Beijing, China
E-mail: liyanfu@tsinghua.edu.cn

***Abstract*—Diabetes is a chronic disease affecting a large number of human population worldwide. Accurate prediction of blood glucose plays an important role for diabetic patients to control the blood glucose in the normal range. In this paper, we use four popular data driven prediction methods for multi-steps ahead prediction with only the historical glucose values as input. Moreover, experiments are carried out to gain insight of the forecast delay phenomenon in the prediction. The reasons leading to the prediction delay are investigated, with the aim to improve the practical value of blood glucose prediction.**

Keywords- blood glucose prediction; forecast delay; data driven prediction method

I. INTRODUCTION

Diabetes is one of the most common chronic diseases, which is mainly caused by the lack of insulin secretion (type I diabetes), or insulin can't work (type 2 diabetes). According to the International Diabetes Federation (IDF), in 2015, 415 million people worldwide suffered from diabetes, while it is estimated that by 2035 this number will rise to 592 million[1]. Diabetes can cause many serious complications such as blindness, uremia, cardiovascular and cerebrovascular diseases, nephropathy, endangering the patient lives. There is no immediate cure method for diabetes so it is vital to monitor the blood glucose (BG) level and control it in the normal range. The importance of BG forecast for patients is due to the fact that after therapeutic measure like insulin is administered, sufficient time is needed for its full effect to take place in the human body.

In recent years, with the development of Continue Glucose Monitoring (CGM) system[2], it becomes more practical and powerful for data driven methods to be used on the BG prediction. CGM systems are usually noninvasive or minimally invasive. It can be used in patient daily life and provide large amounts of real-time glucose data for forecasting. Hence, since the introduction of CGM devices, several multi-steps ahead BG prediction methods have been proposed in the literatures, including the popular time series and machine learning forecast models that use only the historical glucose values as input [3, 4].

BG forecast methods has been reported by many different studies with primary focus on forecast accuracy. However, it is observed that in many reported results, there exists a delay when multi-step ahead forecast is performed. Forecast delay has a direct impact on practical application, but most studies did not focus on it in the analysis and some studies used a subjective method to evaluate it. It is apparent that there is lack of a comprehensive study, which provides further understanding on forecast delay.

The aim of this paper is to introduce four popular BG forecast methods and gain insight of the forecast delay phenomenon in the prediction. Forecasts are carried out using only historical BG value of diabetes patients, which obtained from CGM system sampling every 3 minutes. Prediction horizon (PH) is 30 minutes (10 steps) which is equal to onset time of insulin.

The main contribution of this paper is to verify that all kinds of BG prediction methods produce prediction delay. The experiments prove that one of the important reasons for the delay of prediction is the data noise. The increase of noise amplitude level or noise frequency will lead to the decline of prediction performance.

This paper is organized as follows. Section II introduces four typical methods of the BG prediction. Section III is the analysis of the forecast delay. Next, Section IV presents the impact of data noise on prediction performance. Finally, Section V provides conclusions and future work.

II. DATA DRIVEN METHOD OF BLOOD GLUCOSE PREDICTION

A. Autoregressive Integrated Moving Average (ARIMA)

The ARIMA method is one of the most common model to fit or forecast time series data. It is also widely used in BG prediction in some literatures [5, 6]. ARIMA models are denoted with ARIMA(p,d,q) where each parameter refers to the order of each corresponding component of the ARIMA model. p for the order of order of autoregressive (AR) model, d is the order of differencing (I) and q is the order of moving-average (MA) model.

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

The AR(p) model can be noted as:

$$x_t = \sum_{i-1}^{p} \varphi_i x_{t-i} + c + \varepsilon_t \quad (1)$$

where φ_i ... φ_p are the model parameters, c is constant, and ε_t is error.

The MA(q) model can be noted as:

$$x_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (2)$$

Where μ is the mean, ε_t ... ε_{t-q} are error terms, and θ_i ... θ_p are the model parameters.

The I(d), is a differencing process applied to the time-series data to eliminate trend and seasonality in order to obtain stationary data for modelling. A first order differencing I(1) can be noted as:

$$x_t^{'} = x_t - x_{t-1} \quad (3)$$

Hence, the overall ARIMA(p,d,q) model is shown as:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right)(1-L)^d x_t = \left(1 + \sum_{i=1}^{p} \theta_i L^i\right)\varepsilon_t \quad (4)$$

Where p, d, q represent the order of AR, I and MA component, ϕ_i is the parameter for AR component, θ_i is the parameter for the MA component and L is the lag operator.

B. Multi-Level Perceptron Neural Network (MLP NN)

The application of artificial neural networks in forecasting boomed after the development of back-propagation algorithm by Rumelhart et al [7]. And Perez-Gandia et al. (2010) used an artificial neural network model to perform the BG forecasting based on only historical blood glucose readings [8].

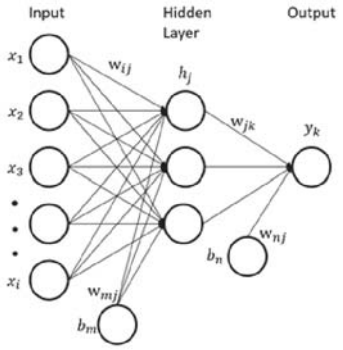


Figure 1. Basic MLP feedforward neural network

In general, a multi layered perceptron feedforward neural network can be modeled as illustrated in Fig. 1. The basic model consists of an input layer, a hidden layer and an output layer. The mathematical model of the neural network can be described as the following:

$$\text{Input Nodes} = x_1, x_2, \cdots, x_i \quad (5)$$

Input to Hidden Node (Neuron),

$$v_j = w_{mj} b_m + \sum_i w_{ij} x_i \text{ , for all j} \quad (6)$$

Output from Hidden Node (Neuron),

$$h_j = f_{act}\left(v_j\right) \quad (7)$$

Input to Output Node (Neuron),

$$s_k = w_{nj} b_n + \sum_j w_{jk} h_j \text{ ,for all } k \quad (8)$$

Output from Output Node (Neuron),

$$\hat{x}_{t+PH} = y_k = f_{act}(s_k) \quad (9)$$

Where i represents the number of input nodes, j represents the number of neurons in the hidden layer and k represents the number of output nodes. b_m and b_n represent the bias nodes in the input layer and the hidden layer respectively, where m and n indicate the number of bias nodes. Typically, only 1 bias node exists for each input and hidden layer. f_{act} represents the activation function used which can be sigmoid, arctan, linear, binary etc. The activation function provides the neural network with the capability to perform non-linear modelling. The number of hidden layer is flexible and the process of weighted sum and activation is repeated for each hidden layers. For point estimation purpose, only 1 output node is needed whereas for objectives related to confidence interval, 2 output nodes are needed to provide the upper and lower limits.

C. Long Short Term Memory Recurrent Neural Network (LSTM RNN)

Compared with MLP NN, the greatest feature of LSTM RNN is that it has a "Memory". It has the function of memorizing previously input data and making predictions based on the previous state and new input data, and that is why RNN is suit to BG prediction [9].

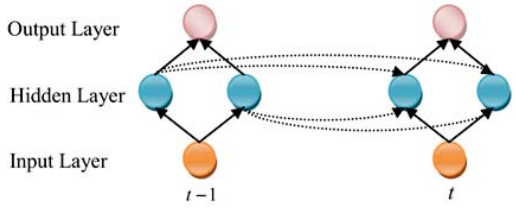


Figure 2. Basic structure of RNN

As shown in Fig. 2, the structure of the RNN across a time can be described as a deep network with one layer per time step, which can be trained across time steps using backpropagation. To overcome the problem of gradient vanishing or exploding, a long short-term memory (LSTM) architecture that involves a memory cell was constructed. The memory cell replaces hidden neurons used in traditional RNNs to build a hidden layer. At time t, recurrent neurons are input with information not just from the previous layer x_t but also from themselves of the previous position h_{t-1}. Consequently, the output y_t is influenced not only by the current input information but also by the information at time $t-1$. Mathematically, these processes can be simply described using the following transition function:

$$h_t = f\left(w_{hx}x_t + w_{hh}h_{t-1} + b_n\right) \tag{10}$$

$$\hat{x}_{t+PH} = y_t = f\left(w_{yh}h_t + b_y\right) \tag{11}$$

Where w_{hx} is the matrix of conventional weights between an input layer and a hidden layer, and w_{hh} is the matrix between a hidden layer and itself at adjacent time steps, $f_{act}(\cdot)$ is an activation function. The vectors b_n and b_y are bias parameters which allow each node to learn an offset.

D. Radial Basis Function Neural Network (RBF NN)

RBF NN is a feed-forward neural network with the strengths of unique optimal approximation (overcoming local minimum problem), simple training, fast learning convergence et al. It has been proved that RBF NN can approximate arbitrary continuous nonlinear network with arbitrary accuracy. BG of diabetes patient is highly nonlinear. Hence, it's practical and suitable for RBF NN to be used on the BG prediction [10].

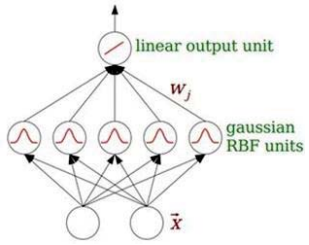


Figure 3. Basic structure of RBF NN

The structure of RBF NN is similar with MLP NN, including input layer, hidden layer and output layer. The connection weight between the input layer and the hidden layer is 1. Unlike the MLP NN mentioned earlier, RBF usually has one and only one hidden layer which contains many RBF neurons, through the RBF to map a low dimensional nonlinear separable input into a high - dimensional linearly separable space. The RBF of the hidden layer node makes a local response to the input. When the input is close to the central range of the base function, the hidden layer node will produce a large output. When it is far away from the central point, the output will show an exponential decay. Finally, the output is the linear weighted sum of the output of the hidden layer neurons. RBF is a non-negative real value function of radial symmetry of the center point, whose value only depends on the distance from the center point such as Gaussian Function.

$$\phi\left(\|x-\mu_i\|\right) = \exp\left(-\frac{\|x-\mu_i\|^2}{2\sigma_i^2}\right) \tag{12}$$

Where μ_i is the Gaussian function center of the *i-th* node in the hidden layer, σ_i is the width parameter/variance of the node.

All the four BG prediction models take the historical blood glucose values measured within the last 10 unit time (30 minutes) as input data, and the target output is the BG level after 30 minutes (PH=10), so the output value of the model is the point estimate of the BG level of patients 30 minutes after the current moment.

In addition to the methods mentioned above, there are many other methods for BG prediction, such as extreme learning machine, Kaltzman filter, regularized learning et al, which are not described in detail here.

III. Analysis of Forecast Delay

The objective of this paper is to understand the factors related to delay in blood glucose forecasting which has been ignored by most blood glucose forecasting studies.

A. Forecast Delay Assessment

In this study, two main parameters are used for the evaluation of BG forecasting, which are accuracy and delay. Accuracy is measured using root-mean-squared error (RMSE) as shown in Eq.(13). On the other hand, delay is measured by the number of back shift in predicted values that best matches the real values (in terms of RMSE):

$$RMSE = \sqrt{\frac{\sum_{j=1}^{q}\left(\hat{x}_j - x_j\right)^2}{q}} \tag{13}$$

$$delay, \min(RMSE) = \sqrt{\frac{\sum_{j=1}^{q}\left(\hat{x}_j - x_j\right)^2}{q}} \tag{14}$$

Where delay is given by the value which results in the minimum RMSE. For example, Fig. 4 shows the prediction result of patient #21 using the MLP NN method, the RMSE is 0.7905 mmol/L and delay is 9 steps calculated by Eq.(13) and Eq.(14).

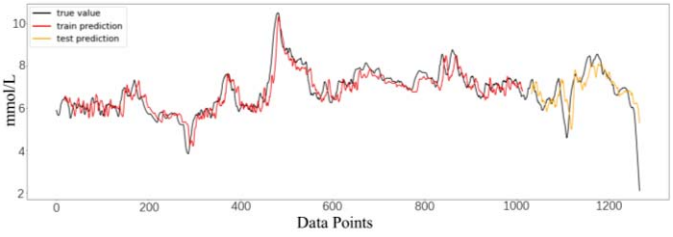


Figure 4. Prediction plot of patient #21 using MLP NN model

B. Forecast of Actual Blood Glucose Data

A test of comparison is performed by running the different forecast methods on 20 different patients' blood glucose data. Since neural network forecast methods are only evaluated on test performance (excluding train performance), the ARIMA related methods will also only be evaluated on the same section of data to maintain a comparable evaluation.

Fig. 5 shows the result of the forecasts. It is observed that the forecast performance of different patients' data have large difference, while performance of different forecast method on a single patient is relatively closer although some differences do exist. This is an indication that there is a large impact of input data on forecast performance.

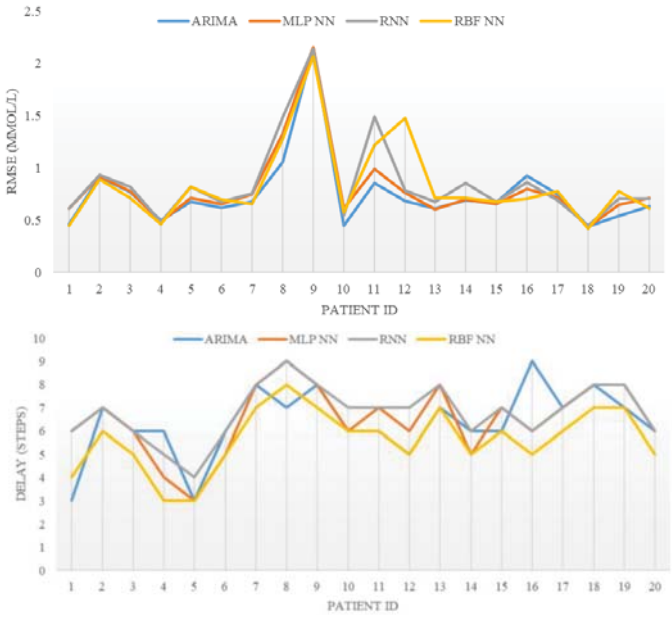


Figure 5. Forecast Results of 20 Patients with Different Forecast Methods

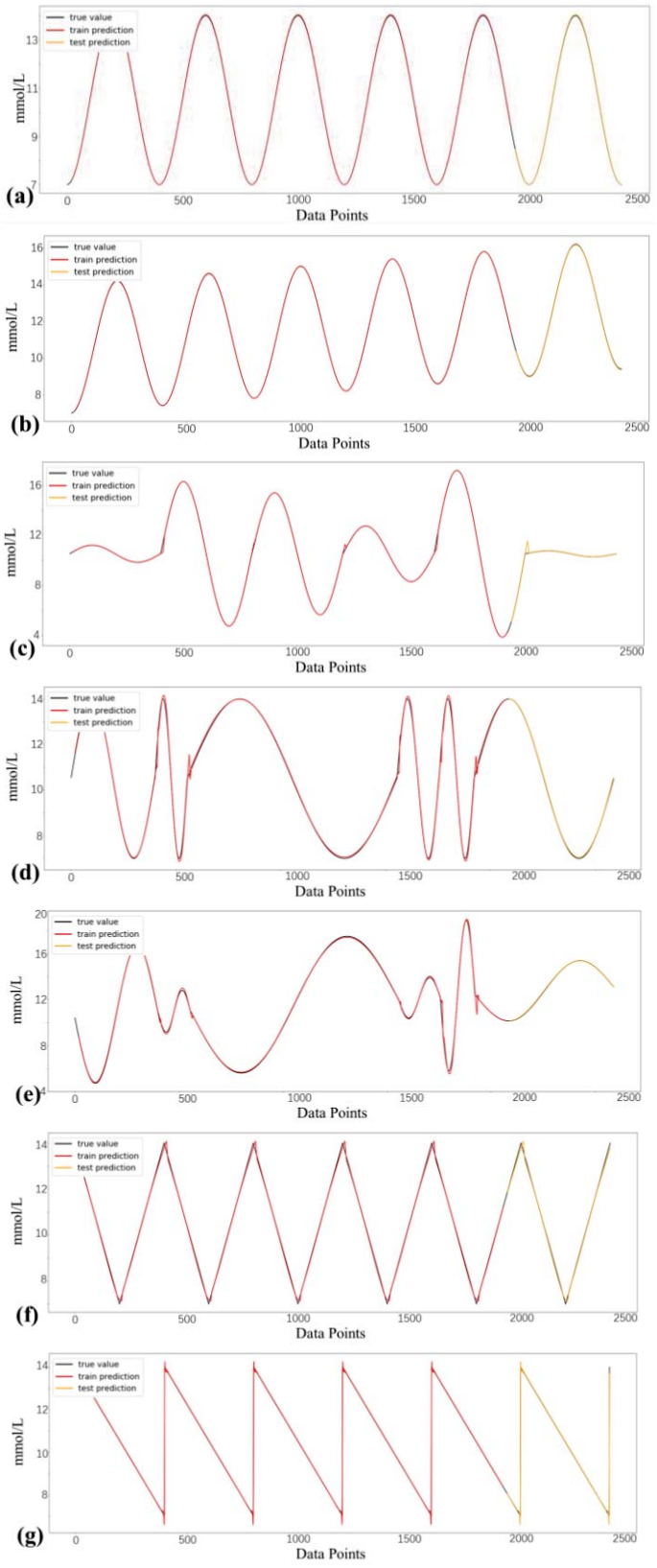
C. *Forecast of Simulated Blood Glucose Data*

In order to have a controlled experiment about the impact of input data on forecast delay, the BG data is represented by several simulated curves, each with different data characteristics. The simulation curve is based on simple curves such as sinusoid, triangular wave and sawtooth wave to mimic the fluctuations (peaks and valleys) in actual BG data.

TABLE I. SIMULATED CURVES FOR FORECASTING

Simulated Curve	Characteristic
a	Sine Curve
b	Sine Curve + Trend
c	Sine Curve + Varying Amplitude
d	Sine Curve + Varying Period
e	Sine Curve + Varying Amplitude + Varying Period + Trend
f	Triangular Curve
g	Sawtooth Curve
h	Sine Curve + Noise

According to the Fig. 6, for the simulated curve 1-7, MLP NN can provide almost perfect prediction without prediction delay. Only at the inflection point and the extreme point where the amplitude or cycle of the curve changes suddenly, some delays and inaccurate predictions can be observed. But the predicted results can be quickly restored to the state matching the actual curve without delay. It indicates that the above BG prediction method can well predict smooth curves, but it cannot well predict if the curve exits many sharp-change points. Due to many fluctuations in the actual patients' BG curve, the overall smoothness of the curve is not very well which may lead to the prediction delay.



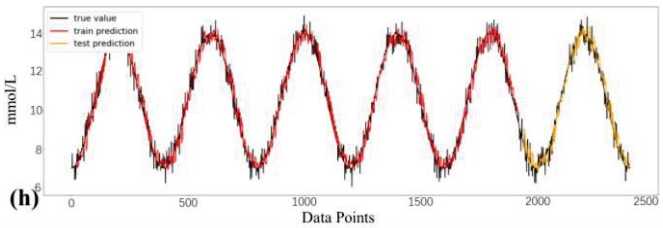


Figure 6. Forecast plot of simulated curve

In addition, for the simulated curve 8, when noise is added to the sine curve, an interesting phenomenon is observed. The whole prediction curve begin to show delay (the delay of curve 8 is 5 steps), which is similar to the result when the patient's actual BG data is predicted. It partly proves that noise in the BG data may be a crucial reason for the forecast delay.

IV. Impact of Data Noise on Prediction Performance

For verifying the influence of noise, it is necessary to conduct repeated experiments on the actual patients' blood glucose data. Reducing the noise of data, then evaluate whether the prediction delay are really improved. We use centered moving average (CMA) method to smooth the real BG data with varying degree of smoothing strength. With increased number of smoothed steps, the smoothing magnitude increased and the noise of the data is decreased.

TABLE II. Forecast Results of Smoothed Data using CMA

Input data	ARIMA		MLP NN	
	RMSE	*Delay*	*RMSE*	*Delay*
Patient 1	0.3460	9	0.3229	8
Patient 1 with CMA5	0.2874	7	0.2857	7
Patient 1 with CMA15	0.1695	4	0.1687	4
Patient 1 with CMA25	0.1110	3	0.1370	3
Patient 1 with CMA35	0.0785	3	0.0781	2

From the Table II, it is clear that as the noise is reduced, there is a significant improvement in the forecast accuracy (RMSE) and the forecast delay for both ARIMA and MLP ANN method. This is a strong validation to the inference drawn from the simulation curve experiment.

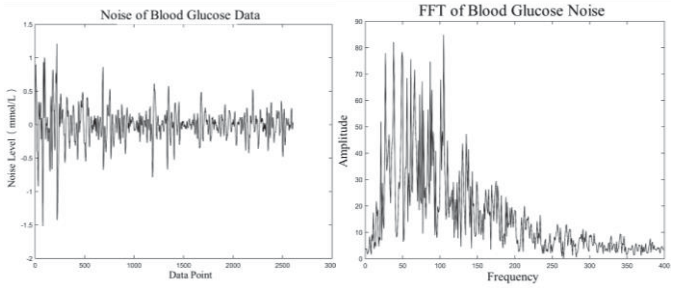


Figure 7. Noise pattern of Patient #21 BG Data and FFT plot

As it is shown that noise in BG data might contribute significantly to forecast delay, hence the noise of input data should be further analyzed to find additional insights. The noise of patients' BG data is acquired by taking the difference between the original data and the 25 steps center moving average (CMA25) smoothed data. A fast Fourier transform (FFT) is performed with the noise pattern. The FFT plot shows the dominant frequency of noise pattern as illustrated in Fig. 7 for Patient #21. It is found that instead of a dominant peak at a certain frequency, multiple high peaks exist within a certain range. Due to different length of data for each patient, the peak frequency found from FFT needs to be converted to period for practical reference.

From the results of FFT, the main repeating period is approximately 20 to 40 steps. As each data is sampled every 3 minutes, this translates to 60 to 120 minute's period of noise oscillation. Based on this time period, it is more likely that the oscillation is generated by patient's physiological process instead of the noise in CGM's electronic system, because electronic noise typically occurs at a much faster rate.

Due to the complexity of the noise in actual data, a design of experiment (DOE) is performed using data with base sine curve (Fig. 6) to evaluate the characteristic of noise on forecast performance. The evaluation is based on two noise characteristics, noise level and period. These two factors are varied in magnitude, so we provide a two-dimensional matrix table for evaluation of these two factors on forecast accuracy and forecast delay.

TABLE III. DOE of Noise Period and Level on Forecast Performance

Noise Level / Noise Period	+/- 0.5		+/-0.7		+/- 1.0	
	RMSE	*Delay*	*RMSE*	*Delay*	*RMSE*	*Delay*
5	0.2715	1	0.3433	2	0.4288	5
10	0.2408	0	0.2767	1	0.3343	3
15	0.2285	0	0.2361	1	0.2998	2
20	0.2152	0	0.2276	1	0.2604	2

From the results shown in Table III, it is clear that increased noise level or decreased noise period will result in worse performance in terms of forecast accuracy and delay. A no interaction two-factor analysis of variance (ANOVA) is performed to confirm the significance of both factors on forecast performance. Table IV and V show that both noise level and frequency cause significant impact on forecast accuracy as well as forecast delay. The p-value obtained is much less than 0.05.

TABLE IV. ANOVA on Noise Frequency and Level on RMSE

Source of Variation	SS	df	MS	F	P-value	F crit
Noise Period	0.02	3	0.007	12.14	0.006	4.76
Noise Level	0.02	2	0.009	14.39	0.005	5.14
Error	0.004	6	0.0006			
Total	0.04	11				

TABLE V. ANOVA on Noise Frequency and Level on Delay

Source of Variation	SS	df	MS	F	P-value	F crit
Noise Period	5.67	3	1.89	6.18	0.03	4.76
Noise Level	15.5	2	7.75	25.36	0.001	5.14
Error	1.83	6	0.31			
Total	23	11				

V. Conclusion

In this paper we introduce four common methods of BG forecast and use these methods to explore the delay of multi-

steps ahead BG prediction. We find the noise of BG data contributes significantly to forecast performance. Further experimental investigation shows that the increased amplitude and frequency of noise will lead to the increase of prediction delay and the decrease of prediction accuracy. It is more likely the oscillation of BG data is generated by patients' physiological process like hormonal regulation instead of the noise in CGM's electronic system.

It is recommended that further studies should investigate the source of noise (e.g. biological fluctuation, temperature, CGM system mechanical stress etc.) to evaluate if a model can be established to compensate for the prediction error caused by noise in data. Another possible approach is to obtain longer period of data for each patient in order to have many repeating periods which can be used for smoothing and filtering.

REFERENCES

[1] S. Oviedo, J. Vehi, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for T1DM patients," International Journal for Numerical Methods in Biomedical Engineering, vol. 33, no. 6, Jun 2017, Art. no. e2833.

[2] D. C. Klonoff, "Continuous glucose monitoring - Roadmap for 21st century diabetes therapy," Diabetes Care, vol. 28, no. 5, pp. 1231-1239, May 2005.

[3] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, and C. Cobelli, "Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration," Ieee Transactions on Biomedical Engineering, vol. 59, no. 6, pp. 1550-1560, Jun 2012.

[4] 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 sensior time-series," IEEE Transactions on Biomedical Engineering, vol. 54, no. 5, pp. 931-937, May 2007.

[5] J. Reifman, S. Rajaraman, A. Gribok, and W. K. Ward, "Predictive monitoring for improved management of glucose levels," Journal of diabetes science and technology, vol. 1, no. 4, pp. 478-86, 2007-Jul 2007.

[6] A. Gani, A. V. Gribok, S. Rajaraman, W. K. Ward, and J. Reifman, "Predicting subcutaneous glucose concentration in humans: data-driven glucose modeling," IEEE Transactions on Biomedical Engineering, vol. 56, no. 2, pp. 246-254, Feb 2009.

[7] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, no. 6088, pp. 533-536, Oct 9 1986.

[8] C. Perez-Gandia et al., "Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring," Diabetes Technology & Therapeutics, vol. 12, no. 1, pp. 81-88, Jan 2010.

[9] F. Allam, Z. Nossair, H. Gomma, I. Ibrahim, and M. Abdelsalam, "A recurrent neural network approach for predicting glucose concentration in type-1 diabetic patients," in Engineering Applications of Neural Networks, Pt I, vol. 363, L. Iliadis and C. Jayne, Eds. IFIP Advances in Information and Communication Technology, 2011, pp. 254-259.

[10] G. Baghdadi, A. M. Nasrabadi, and IEEE, "Controlling blood glucose levels in diabetics by neural network predictor," in 2007 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vols 1-16. Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 3216-3219.