PREDICTION OF BLOOD GLUCOSE CONCENTRATION AHEAD OF TIME WITH FEATURE BASED NEURAL NETWORK

S. Shanthi [1], D. Kumar [2].

[1] Department of ECE, JJ college of Engineering and Technology ,Anna University of Technology, Tiruchirappalli, Tamil Nadu, India. Email: sshanthi289@hotmail.com

[2] Dean-Reseach, Periyaar Maniammai University, Vallam, Tanjore, Tamil Nadu, India. Email:kumar_durai@yahoo.com

ABSTRACT

Diabetes has become a major health challenge affecting nearly 300 million people around the world. Complications of diabetes can be prevented by proper monitoring and regulation of glucose concentration in blood plasma. Continuous Glucose Monitoring Systems help to track the time course of blood glucose. These devices have the additional feature of giving threshold alert and predictive alert which is needed for an early warning of impending hypoglycemia. However, the accuracy of predictive alerts in currently available CGM devices is not very promising. Various algorithms have been developed in this regard by researchers. Still, a 100% accuracy has not been achieved. In our work, we have approached this prediction by training a simple neural network with the extracted features of continuous glucose monitoring sensor data time series. The data was obtained in three different ways, one set from the Self Monitoring Blood Glucose values, the second set from a diabetes resource and the third one from the patients using continuous glucose monitoring systems. A feed forward neural network with back propagation algorithm is trained with features of input patterns. The network is trained and validated to meet out the performance goal. The Root Mean Square Error between the actual glucose value and the predicted glucose value is used as the performance measure. It is observed that as the length of prediction horizon extends, the error increases. However, tracking of Hypoglycemic and Hyperglycemic trends are superior to the earlier approaches.

Keywords : *CGM, Feature Extraction, Hypo glycemia, Neural Network, Prediction, RMSE.*

1.0 INTRODUCTION

Diabetes Mellitus is a metabolic disorder characterized by high blood glucose levels. Insulin is the hormone secreted from the Pancreas, which helps in the regulation of blood glucose. The Diabetes Control and Complications Trial showed that strict glycemic control significantly reduces the short term and long term complications of Diabetes [1]. One should maintain the blood glucose within the normal range (70 – 120 mg/dL or 3.6 – 6.9 mmol/L). Lower glucose levels (< 50mg/dL) is said to be Hypoglycemia which leads to excessive thirst, sweating, seizures and diabetic coma. Higher glucose levels (> 200 mg/dL) is said to be Hyperglycemia which leads to long term vascular complications Diabetic Retinopathy, Neuropathy and Nephropathy. Therefore for proper control of glucose level, monitoring is required, thus improving the quality of life. Blood glucometers are used for measurements at discrete instances of time. While the Continuous Glucose Monitoring (CGM) devices provide a minimally invasive mechanism to measure and record patient's current glycemic states as frequently as every minute. It provides maximum information about the blood glucose variations throughout the day which facilitates diabetic patients to make optimal treatment decisions. A review of currently available CGM devices is given in [2,3]**.** The prediction of hypoglycemia is a clinically important task in the management of diabetes. Since hypoglycemia has the dangerous effects like seizure and coma, it has to be predicted well in advance and preventive measures should be taken. The studies have shown that 50% of predictive alerts generated by the currently available CGM devices are of false or missing alarms [4]**.** Research is carried out by various groups around the world in this regard. This predictive monitoring is essential for the success of artificial pancreas project promoted by the Juvenile Diabetes Research Foundation [5].

1.1 Approaches in Prediction of blood glucose – A Review.

According to Bremer and Gough, the pioneer in prediction of glucose concentration in blood, if the recent blood glucose history is not random but has an exploitable structure, it might be possible to anticipate the BG values in the near future based only on previous values [6]. They tried the prediction mathematically with an Auto Regressive Moving Average (ARMA) process for prediction horizon (PH) of 10, 20 and 30 minutes. Palerm et al., have demonstrated the effect of sampling frequency, threshold selection and prediction horizon on the sensitivity and specificity of predicting hypoglycemia [7,8] . They used estimation and prediction with a Kalman Filter and managed to get the results as 90% Sensitivity and 79% Specificity. i.e to have 79 correct predictions one has to bare with 21 false alerts. Sparacino et al., used two prediction strategies based on the description of past glucose data. One is the first order polynomial (Poly (1)) and the other one is the first order Auto Regressive AR (1) model [9]. Both methods

have time varying parameters estimated by Weighted Least Squares. In both the methods, at each sampling time, a new set of model parameters is first identified by means of weighted least squares technique. The model is then used to forecast glucose level for a given prediction horizon. Mean Square Error (MSE) and Energy of First Order Differences (ESOD) were taken as the performance metrics. The analysis is done with various forgetting factors of μ = 0.2, 0.5, 0.8 and prediction horizons of 30 minutes and 45 minutes. Theresult shows the MSE of 318 mg/dL for AR(1) model and 336 mg/dL for Poly(1) model for PH of 30 minutes and for PH of 45 minutes the MSE in the range of thousands for both models. Sparacino et al. used time lag as a parameter to assess the predictive capability of the models. Reifman et al., investigated the capabilities of data driven AR models to capture the correlations in glucose time series data [10]. For PH of 30 and 60 minutes, the Root Mean Square Error (RMSE) are 26 and 36 mg/dL respectively. Pappada *et al.,* (2008) designed various neural network models with Neuro Solutions software and an electronic diary information, for the prediction of blood glucose in various PH of 50,75,100,120,150 and 180 minutes [11]. For PH of 100 minutes he obtained a mean absolute difference (MAD) of 43 mg/dL. Predictions in hypoglycemic ranges were of lower accuracy, which may be due to smaller number of training data in that range. Gani et al., combined the predictive data driven models and the frequent blood glucose measurements [12]. By simulation, they proved that stable and accurate models for near future glycemic predictions with clinically acceptable time lags can be obtained by smoothing the raw glucose data and regularizing the model coefficients. This has to be validated for real time implementation. This group has worked with AR model of higher orders (AR (30)). For PH of 30, 60 and 90 minutes, the RMSE were 1.8, 12.6 and 28.8 mg/dL respectively. Perez-Gandia et al., implemented an artificial neural network algorithm for online glucose prediction from continuous glucose monitoring [13]. The inputs of the neural network are the values provided by CGM sensor during the last 20 minutes and the output is the prediction of the next time step. The performance of the neural network prediction model was compared with the AR model. The results (RMSE) given by the neural model are 10, 18 and 27 mg/dL for PH of 15,30 and 45 minutes respectively. Robertson et al., reviewed the various neural network approaches in blood glucose prediction and arrived with an artificial neural network (ANN) architecture that is the Elman Recurrent structure [14]. They analyzed the predictive capability of the model with various longer PHs, at nights and over 5 days by feeding additional information like food, insulin dosages for a maximum of 1 hour. The RMSE between the actual and predicted blood glucose levels was calculated and an average RMSE of 0.15 ± 0.04 SD mmol/L for the five days was obtained.

In our work, we carried out the prediction of blood glucose with a simple neural network model which is trained with the assistance of extracted features. The remaining part of the paper is given as follows. Section 2 deals with research design and methods, Section 3 gives the experimental results, and Section 4 is the discussion.

2.0 MATERIALS AND METHODS

The data for the proposed method is obtained in three ways. First the data set is formed from Self Monitoring Blood Glucose (SMBG) values of the 20 patients with One Touch Ultra® glucometer. The discrete values are then smoothened to give a continuous time course using cubic splines. The details are explained in the following section. Then, each time course is simulated 100 times using Monte Carlo simulation so as to have 2000 data sets. The second data set is obtained through diabetic resource from the glucose project of the University of California San Diago [15]. The data set gives blood glucose dynamics of patients in a hospital setting with different insulin therapy. The third data set is taken from the users of Medtronic CGMS®(Medtronic Minimed, Inc.,Minneapolis,CA).

The SMBG values are collected at 8 different instances in a 24 hour day. Patients involved are of different age groups and categories. Out of 20, four are of Type 1 Diabetes with age10 ± 4 (Mean ± Standard Deviation), ten are of middle aged with 40 ± 8, Type II Diabetes and the remaining six are in the age group of 60 ± 9 with Type II Diabetes. The parameters considered include age, weight, Glycosilated Hemoglobin (HbA1C), Insulin dosages and duration of disease. A sample data sheet is shown in Table1.

Table 1: Data Collection Sheet for SMBG

Patient ID	# 01						**Date**	22.5.2011
Name	S.Nandan						**Age**	33
HbA1C	7.6						**Weight**	72 Kg
Duration of Disease	7 Years							
Time	**12:00 AM**	**3.00 AM**	**6.00 AM**	**9:00 AM**	**12.00 Noon**	**15:00 PM**	**18:00 PM**	**21:00 PM**
INSULIN Dosage	15 Units		10 Units		8 Units			
SMBG Reading in mg/dl	96	84	95	153	145	139	170	151

2.1 Spline Interpolation

The SMBG values are applied to a cubic spline to get a 24 hour continuous glucose time series. It is a multilevel B spline model to analyze and compare patient glucose profiles based on continuous monitoring data [16]. SMBG gives glucose values at discrete instances of time t_0, t_1, $t_2 \ldots t_n$. To find the glucose value at any other instance, a continuous function $y = f(x)$ may be used. This function is used to represent $n+1$ data points with $f(x)$ passing through all $n+1$ point. Then one could find the value of 'y' at any other instance of time. This is called interpolation. A polynomial is a common choice for interpolating function. Polynomial interpolation involves finding a polynomial of order 'n' that passes through $n+1$ points. Newton's divided difference polynomial method and Lagrangian interpolation are some of the methods to obtain such a polynomial. When 'n' becomes large, the polynomial shows an oscillatory behavior. Therefore when the number of data points is more, Spline interpolation is used. The common spline interpolations are Linear, Quadratic and Cubic splines [17].

In cubic spline interpolation, a series of unique cubic polynomials are fitted between each of the data points with the stipulation that the curve obtained is continuous and appear smooth. These cubic splines can then be used to determine the rates of change and cumulative change over an interval. Robust form of cubic spline could encompass unequally spaced points. The essential idea is to fit a piece wise function of the form

$$S(x) = \begin{cases} S_1(x) & \text{if } x_1 <= x < x_2 \\ S_2(x) & \text{if } x2 <= x < x_3 \\ \vdots & \\ \vdots & \\ \vdots & \\ S_{n-1}(x) & \text{if } x_{n-1} <= x < x_n \end{cases} \quad \text{---- (1)}$$

Where 'S_i' is a third degree polynomial defined by

$$S_i(x) = a_i*(x-x_i)^3 + b_i*(x-x_i)^2 + c_i*(x-x_i) + d_i \qquad \text{for } i = 1,2,\ldots..n-1. \qquad \text{---- (2)}$$

The first and second derivatives of these $n-1$ equations are fundamental to this process [16].The cubic spline is chosen to satisfy the following properties:

1. The piecewise function $S(x)$ will interpolate all data points.
2. $S(x)$ will be continuous on the interval [x_1,x_n].
3. $S'(x)$ will be continuous on the interval [x_1,x_n].
4. $S''(x)$ will be continuous on the interval [x_1,x_n].

To make the curve smooth across the interval, the derivatives must be equal at the data points,i.e., $S'(x) = S''(x)$.

To obtain continuous glucose time series, the data collected from 20 patients in a hospital setting were fitted to the cubic polynomial function. With the Monte Carlo approach, each continuous glucose profile is in turn simulated 100 times to have 2000 different distributions of blood glucose profiles. This helps to validate our methodology in different scenarios of inter person and intra person variability of signal to noise ratio of CGM sensor signal. A sample blood glucose time series obtained with spline method is shown in Fig.1 and a continuous glucose monitoring sensor data from a continuous glucose monitoring system (CGMS) is given in Fig.2.

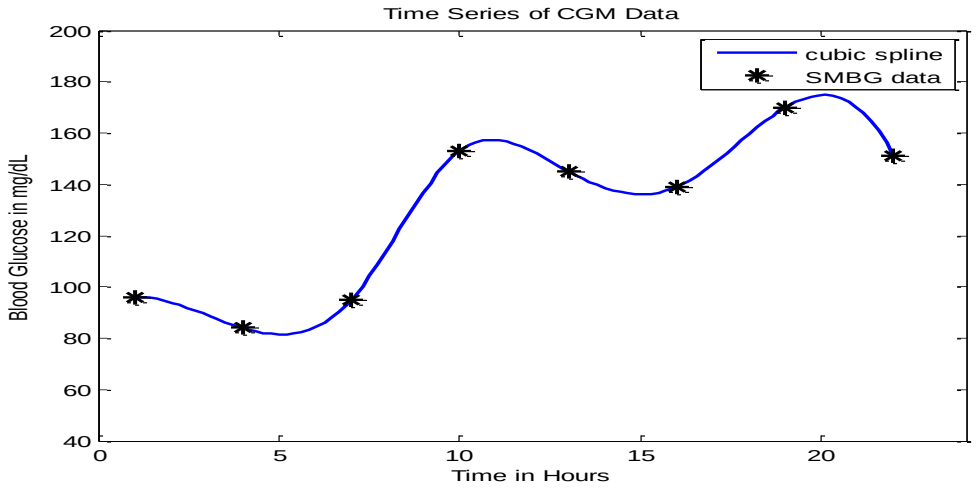

Fig.1: Blood Glucose time Series obtained with Spline Technique.

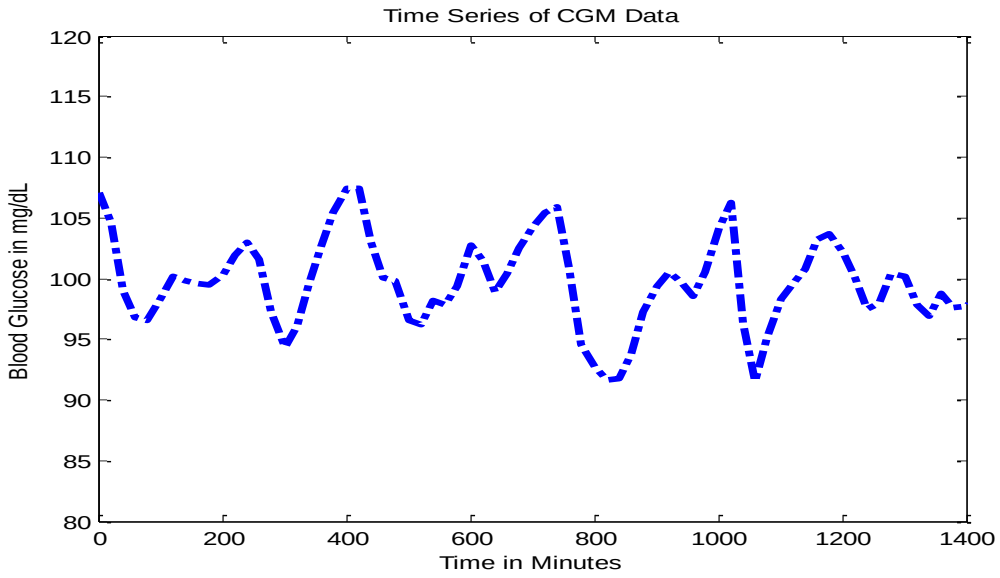

Fig.2: Continuous Glucose Monitoring Sensor Data Profile from Medtronic CGMS®.

The third set of data is obtained from Medtronic CGMS® users (25 numbers). 25% of recruited patients are of type 1 diabetes and remaining 75% are of type 2 diabetes. The patients have HbA1C value of 7-10 which reflects their average BG level in the last 3 months. The Medtronic CGM device measures interstitial fluid glucose level every 5 minutes, resulting in 288 readings in each day of continuous glucose monitoring. The data are downloaded to the personal computer through specialized software for analysis.

2.2 Need for Accuracy and Reliability in Continuous Glucose Measurements

The Interstitial fluid (ISF) Glucose Sensor are minimally invasive and amperometric type [18]**.** The continuous glucose monitoring systems assess BG fluctuations indirectly by measuring the concentration of interstitial glucose but are calibrated via self monitoring to approximate the BG. There is an average time lag of 12 minutes between arterial blood glucose and ISF glucose [19]**.** The number and timing of calibration is still a research issue [20]. Sensor performance is poor at situations like rapid excursions(sudden rise or fall) , motion artifacts and local inflammatory complications [21]. The above said factors have a direct impact on the sensor accuracy, resulting in noisy data. Hence, optimal filtering or preprocessing is required for CGM sensor data signal before sending it for further processes like generation of predictive alert or a control signal for insulin pump control.

An hybrid filtering technique for denoising of CGM sensor data with a feed forward neural network with extended Kalman filter algorithm has been developed and validated in our early work [22].

2.3 Prediction of Blood Glucose From Non linear CGM Dynamics

The dynamics of blood glucose can be described as a non linear black box model. In system identification, the non linear black box models are of two types; Non linear ARX models and the Hammerstein-Wiener model [23]. The non linear ARX model computes the output in 2 stages.

- Computes regressors from the current and past input values and the past output data. Simply, the regressors are delayed inputs and outputs.
- The non linearity estimator maps the regressors to the model output using a combination of non linear and linear functions.

The multi layer neural network is one of the non linear estimators.

$$y(t) = f(y(t-1), y(t-2), y(t-3), \ldots. y(t-n_y), u(t-1), u(t-2), u(t-3), \ldots\ldots. u(t-n_u)) \quad \text{----(3)}$$

Since the interactions between the factors for glucose metabolism are complex, multidimensional, highly non linear, chaotic, stochastically and time variant time series, the neural network model seems to be a more suitable predictor. It can model the input-output behavior of glucose metabolism without knowing the involved explicit internal processes. In fact, many neural network models have been applied for time series prediction. We have approached the prediction of BG with a simple feed forward, back propagation neural network trained with the extracted features of BG dynamics.

2.4 Network Structure

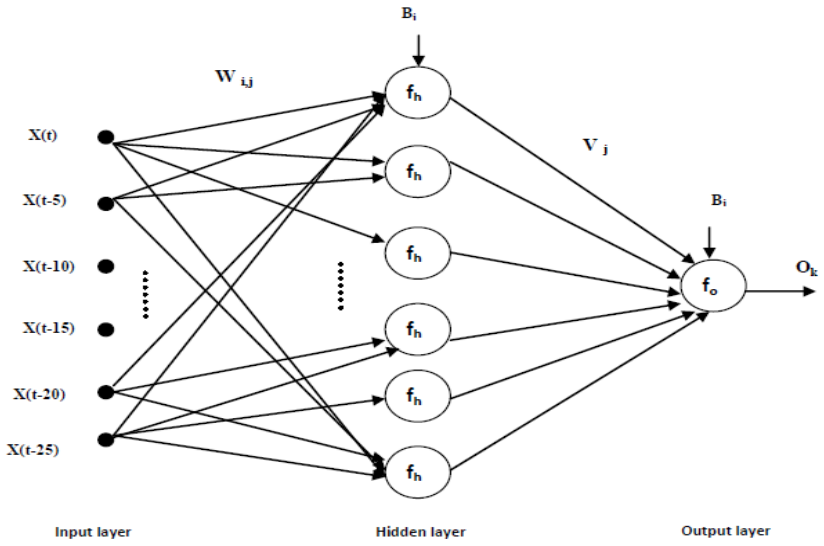

Fig. 3 : Neural Network Model

The neural network models are of varied types which differ in the architecture, learning algorithm and activation functions. Featured neural network is a network comprised of 3 layers viz., input, hidden and output layer. The input layer has six nodes for passing the inputs to the hidden layer. In turn, the hidden layer also has six nodes for processing the sum of weighted inputs with activation function (f_h). The output layer has a single node with activation function (f_o). The current and past 5 values are given as inputs to the network. i.e., a moving window of six tuples is used. Since the Medtronic CGM sensor measures the ISF glucose values once in every 5 minutes the predicted value at any instant depends on the last half an hour readings. The network is trained to capture the trend and predict the near future values. The window is optimized with false nearest neighbor method [24]. The network used for prediction of future values of CGM time series is shown in Fig.3.

2.5 Network Training

The classical back propagation adopts first order steepest descent technique as learning algorithm. Weights are modified in a direction that corresponds to the negative gradient of the error surface. Gradient is an extremely local pointer and does not point to global minimum. This hill climbing search is in zig zag motion and may stuck with a local minimum [25]. The direction may be spoiled by subsequent directions, leading to slow convergence. The classical back propagation is sensitive to parameters like learning rate and momentum rate. If the learning rate is too small, it leads to slow convergence and too large value will make the search direction jump wildly and never converge. The optimal values are difficult to find and often obtained empirically [26]. Many approaches have been given in literatures such as genetic algorithm, simulated annealing, ordinary least squares, Newton-Raphson method, and Levenberg-Marquardt algorithm [27].

2.6 Proposed Method

It is proved in the literature that inclusion of past measurements of the input variables increases the prediction accuracy of neural network model [28]. Prediction of variations in glucose metabolism is a complex process since it depends upon various internal and external factors [29]. We propose a novel feature based prediction algorithm for forecasting the blood glucose values ahead of time. Every member of the feature set contributes in training the neural network by adjusting the learning rate parameter with weighted values of extracted features. The incoming time series is represented in terms of features. The recurrent characteristics of blood glucose dynamics can be tracked with a moving window of length ‘n’ with a step size of ‘p’. For every window, the features are computed and passed to the network for learning. The standard moments of any time varying signal are mean, variance, skewness and kurtosis. Mean gives the weighted average of all expected values over an interval. Variance is a measure of how far a set of data are spread out from the mean. It is the average of squared differences from the mean. Skewness is a measure of asymmetry, while Kurtosis gives information about the peak. Approximate Entropy is a regularity statistic that quantifies the unpredictability of fluctuations in a time series. This feature set tracks the underlying variation of a dynamic signal. This helps to identify the principle components of the signal of interest. For the input CGM data ‘X’, the computations of features are given below.

Mean : $\mu = \left(\frac{1}{n}\right) \sum_{i=1}^{n} (x_i)$ ---- (4)

Variance : $\sigma^2 = \left(\frac{1}{n}\right) \sum_{i=1}^{n} (x_i - \mu)^2$ ---- (5)

Skewness : $S = E(x - \mu)^3 / \sigma^3$ ---- (6)

Kurtosis : $K = E(x - \mu)^4 / \sigma^4$ ---- (7)

Approximate Entropy: $ApEn(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r)$ ---- (8)

Where ‘Φ’ is a subsequence of length ‘*r*’, ‘*m*’ is the sequence number,’*N*’ being the length of time series. ApEn assigns a non negative number to a time series. Larger values correspond to apparent irregularity in the input pattern. AnEn measures the log likelihood that runs in patterns that are close (within ‘*r*’) for ‘*m*’ contiguous observations remain close on the next incremental comparisons [30]. ApEn is complementary to peak detection and spectral analysis in that it evaluates both dominant and sub ordinate patterns in concentration time series [31].

The above listed features are extracted from the incoming n-tuples. The feature set information is used in learning rate parameter adjustments and activation functions of hidden layer and output layer. The features assist in the excellent learning of the network in tracking the dynamics of incoming signal.

2.7 Algorithm

Step 1 : Initialize weights randomly for hidden and output layer neurons *V*, *W* .

Step 2 : For each set of inputs,

- Calculate the features $R_{i,k}$. Where i = 1,2…5. ; 'k' is the iteration.
- Compute the layers' responses.

$$Y_j^s = f_h (\sum_i W_{ij} X_i^k) \quad \text{---- (9)}$$

$$O^s = f_o (\sum_j V_j Y_j^k) \quad \text{---- (10)}$$

Where W_{ij} are the weights of the neurons between the input and hidden layer.

V_j are the weights of the neurons between the hidden and output layer.
f_h and f_o are the activation functions at the hidden and output layer respectively.

Step 3 : Compute the error : $E \leftarrow (1/2) \| d_k - O_k \|^2$. ---- (11)

Where d_k is the desired output and O_k is the actual output.

Step 4 : Compute the learning rate parameter : $\tilde{\alpha}_k = \sum a_i R_{i,k}$. ---- (12)

The coefficients 'a_i'of each of the features are estimated with Weighted Least Squares method.

Step 4 : Calculate errors δ_0 , δ_y.

$$\delta_0 = (1/2)[(d_k - O_k)* \tilde{\alpha}_k] \quad \text{---- (13)}$$

$$\delta_y = W_j^t * \delta_0 * f_y^t \quad \text{---- (14)}$$

$$\delta_y' = (1/2) [1 - y_j^2] \quad \text{---- (15)}$$

Step 5: Adjust weights of Output layer.

$$V_{k+1} \leftarrow V_k + \tilde{\alpha}_k * \delta_0 * y^t \quad \text{---- (16)}$$

Step 6: Adjust weights of Hidden layer.

$$W_{k+1} \leftarrow W_k + \tilde{\alpha}_k * \delta_y * z^t \quad \text{---- (17)}$$

- Repeat steps 2 to 6 for all the training patterns.
- If the training patterns are over in the training set, check for $E < Emax$?
  - If yes, stop training.
  - If no, assign $E \leftarrow 0$ and begin a new training cycle.

The training of neural network model involves modification of the weight vector gradually, to minimize the difference between the predicted value and the desired response. In this work, RMSE and time lag between the predicted and actual values are taken as performance measures.

3.0 EXPERIMENT

The data sets are selected where all of them have the same sampling frequency. The data set for training, validation and testing are assigned based on the dividerand function. The dividerand function is a default Matlab function used to divide the given data set into three sets, namely, training, validation and testing according to the random indices. In our application, 50% of data is used for training, 25% for validation and 25% for testing. The methodology is tested with different scenarios like training and testing of network with the same patient data, training with one patient data and testing with another patient data. This helps to check the ability of the network to account for inter individual variability of signal to noise ratio of the CGM sensor data. The intra individual variability of signal to noise ratio is tracked with feature extraction and tuning of network parameters accordingly. The input data series consists of interstitial glucose values taken once in every five minutes. A moving window of size 6-tuple is formed with the current and past five CGM values. The feature set is extracted from each window and used for the learning rate

parameter adjustment and activation functions at hidden layer and output layer. The output neuron gives the predicted glucose value for the given PH. The methodology is tested with prediction horizons of 30,45 and 60 minutes. The Root Mean Square Error (RMSE) and Time delay between the actual value and the predicted output are calculated for performance comparison with the previous recent work on prediction of CGM sensor data.

4.0 RESULTS

Fig.4 shows the predictive capability of Feature based Neural Network (FNN) with a prediction horizon of 30 minutes. A deviation was observed in the initial period of 0 to 70 minutes. However, FNN was able to catch up the dynamics of blood glucose accurately with a RMSE value of 6.3 mg/dL.

The prediction of CGM time series with a prediction horizon of 45 minutes is shown in Fig.5. It is proved from the figure that the trend of blood glucose variations are well followed by FNN with minor deviations at instances of 300 minutes and 650 minutes. That was also in the range around 4 and 6 mg/dL respectively. An overall RMSE of 9.5 mg/dL was obtained with PH = 45 minutes.

Fig.6 depicts the prediction of FNN with prediction horizon of 60 minutes. The performance measure of 14.6 mg/dL RMSE was obtained in this scenario. Although this value is higher than the earlier scenarios with PH = 30 and 45 minutes, the efficiency of FNN could be well understood from Fig.6 which shows the tracking of FNN in the same trend as the original signal.

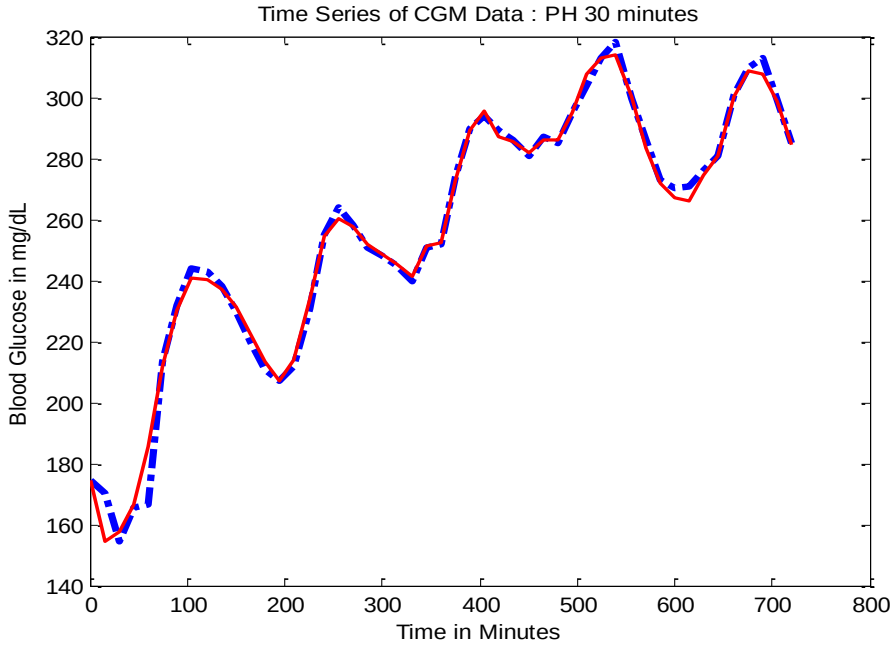

Fig.4: The predicted glucose profile with 30 minutes of PH (dashed line) overlapped with actual CGM data (smooth line).

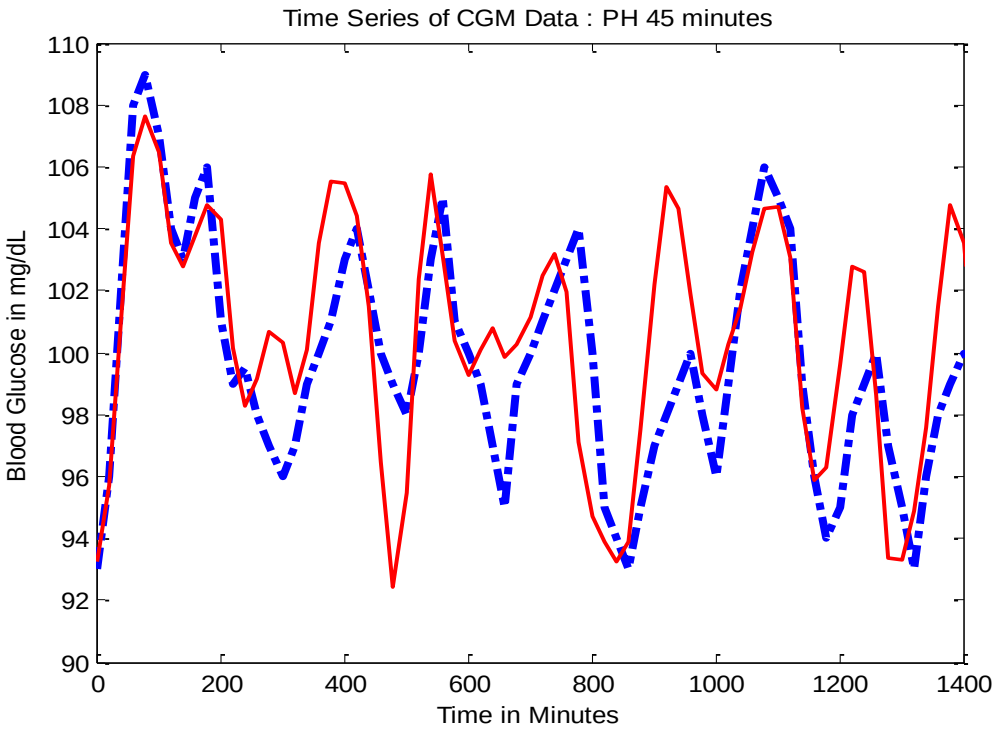

Fig.5: The predicted glucose profile with 45 minutes of PH (dashed line) overlapped with actual CGM data (smooth line).

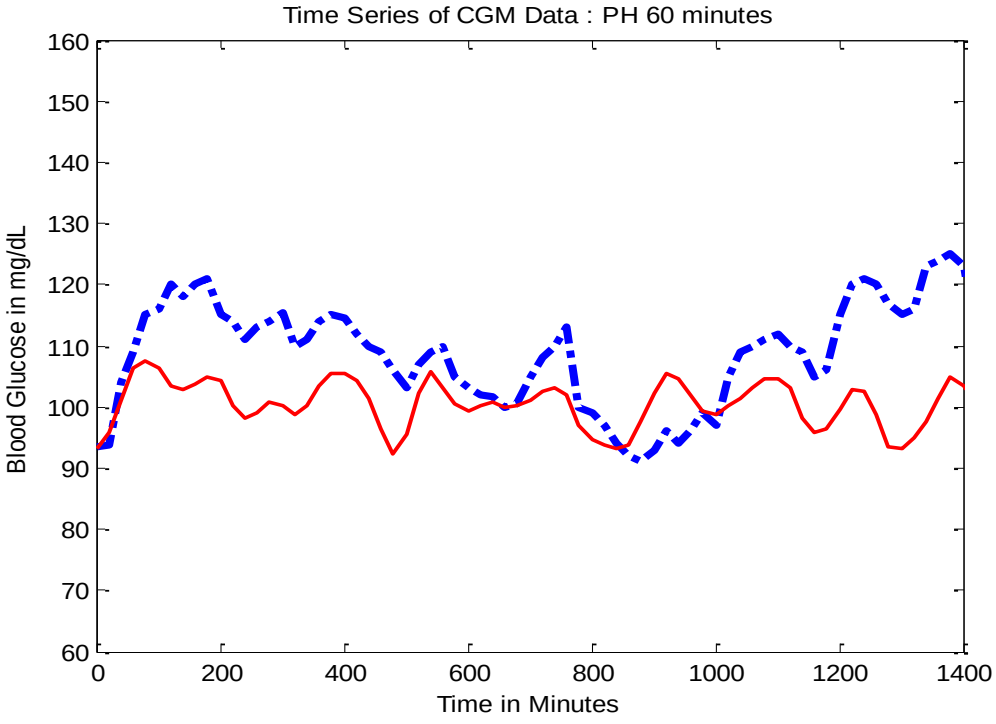

Fig.6: The predicted glucose profile with 60 minutes of PH (dashed line) overlapped with actual CGM data (smooth line).

The performance of the proposed method, Feature based Neural Network (FNN) was analyzed with all the three data sets mentioned earlier. The proposed method was compared with the recent work on prediction with neural network

trained by Levenberg-Marquardt algorithm [13]. For ease of representation, here we refer to the neural network trained by Levenberg-Marquardt algorithm as the LMNN. Table 2 gives the details of prediction error in terms of the RMSE in mg/dL and the standard deviation (SD) for both FNN and LMNN models in the prediction horizons of 30,45 and 60 minutes. In the 30 minutes prediction horizon, the prediction error of FNN was three times lower than that of LMNN in all the data sets. The mean RMSE value of FNN was 6.3 mg/dL whereas it was around 18 mg/dL for the LMNN model. For PH values of 45 and 60 minutes, the error of the LMNN was almost twice that of the FNN. In the 45 minutes PH, the RMSE of FNN was around 9.5 mg/dL while that of LMNN was 23.6 mg/dL for LMNN and in the 60 minutes PH, the RMSE of FNN was around 14.6 mg/dL while that of the LMNN model was 27.4 mg/dL.

Tables 3 and 4 show the respective average delay in minutes and the standard deviation of the rising and falling trends. For PH values of 30 minutes, the delay of the FNN was around 5 minutes, whereas it was 7 minutes for the LMNN. This delay increases with increasing PHs. For PH of 45 and 60 minutes, the delay of the LMNN was twice greater than the FNN, with LMNN: 11.7 minutes Vs. FNN: 6.3 minutes at 45 minutes PH and for PH of 60 minutes, it was LMNN: 18.2 minutes while FNN: 9.4 minutes. In both the models, the delay in the upward trend was less than that of the downward trend.

Table 2: Performance comparison with Root Mean Square Error.

Data Set	RMSE (Mean ± SD) in mg/dL for Prediction Horizons of					
	30 minutes		45 minutes		60 minutes	
	FNN	LM NN	FNN	LM NN	FNN	LM NN
1	4.0 ± 5.2	16.43 ± 7.8	9.0 ± 5.1	20.5 ± 7.3	12.6 ± 8.0	26.2 ± 4.8
2	7.8 ± 2.0	22 ± 3.4	9.8 ± 3.6	27.0 ± 6.0	14.8 ± 7.1	28 ± 2.1
3	7.5 ± 4.2	18.35 ± 6.4	10.3 ± 5.6	24 ± 8.73	15.5 ± 8.4	29.3 ± 9.6

Table 3: Time lag in upward trend.

Data Set	Time lag (Mean±SD) in minutes for Prediction Horizons of					
	30 minutes		45 minutes		60 minutes	
	FNN	LM NN	FNN	LM NN	FNN	LM NN
1	3.2 ± 2	7.0 ± 1.3	4.5 ± 3.0	8.5 ± 2.9	7.6 ± 4.1	18.3 ± 2.0
2	6.8 ± 3.3	8.0 ± 2.5	8.4 ± 2.6	11.3 ± 4.0	9.5 ± 3.2	18.8 ± 4.6
3	5.3 ± 1.5	7.28 ± 3.36	9.5 ± 4.1	12.50 ± 5.68	10.0 ± 2.5	19.3 ± 7.4

Table 4: Time lag in downward trend .

Data Set	Time lag (Mean±SD) in minutes for Prediction Horizons of					
	30 minutes		45 minutes		60 minutes	
	FNN	LM NN	FNN	LM NN	FNN	LM NN
1	8.0 ± 2.4	12.1 ± 3.0	11.1 ± 2.6	22.1± 2.3	15.7 ± 3.8	29.5 ± 4.7
2	9.4 ± 3.5	17.2 ± 5.5	12.6 ± 6.3	24.0 ± 5.7	16.6 ± 8.2	30.5 ± 6.3
3	10 ± 2.3	15.67 ± 3.8	13.1 ± 1.9	29.5 ± 6.8	15.4 ± 6.1	32.7 ± 8.8

Table 5 shows the performance evaluation of FNN in terms of the average number of iterations required for training, average mean square error (MSE) and the percentage of prediction accuracy. In general, LMNN is said to have the feature of faster training, but with a larger memory requirement. In thisstudy, the FNN took 58 iterations and the LMNN took 45 iterations. The FNN with specialized training with feature set, is able to achieve the performance goal efficiently. The MSE for the FNN was 0.000617 mg/dL while that of for the LMNN was 0.0824 mg/dL. Among the total number of predictions, 91% results are true positive whereas it was 67.9 % for the LMNN. A bench mark of performance was established by training the FNN with the specific features of incoming data set without the benefit of dimensionality reduction.

Table 5: Performance Evaluation of FNN and LMNN

	Average Number of Iterations	Average MSE	% of Correct Prediction
FNN	58	0.000617	91.435
LMNN	45	0.0824	67.891

5.0 DISCUSSION

It was observed that the RMSE in the prediction of the first data set (i.e, from SMBG values with Cubic Spline interpolation) was less compared to the other two sets in all prediction horizons. This might be due to the smoothness by cubic splines.The results in data set 2 and 3 did not differ much, since both sets are obtained from continuous monitoring systems. In PH of 30 minutes, the RMSE obtained by FNN in all the data sets was lesser than or nearer to 10 mg/dL. For PH of 45 minutes, it was close to 15mg/dL and for PH of 60 minutes it was nearer to 20 mg/dL. Thus FNN showed superior performance in all scenarios when compared with the earlier LMNN. This might be due to the training of the network with extracted features, which reflect the physiological variations of glucose metabolism.

The time lag in the predictions of the upward and downward trend were analyzed. The time lags in the downward trend were larger than that of the upward trend. This might be due to physiological reasons. i.e., the rise in BG due to ingested food is rapid and uniform. Whereas glucose disposal (fall in BG value) differs with individuals based on many factors. It had been emphasized that the features extracted from the inputs were used to track the physiological variation of the blood glucose. We also investigated the influence of sampling rate on prediction accuracy and confirmed that samples taken in five minutes interval are good enough to give accurate results. Due to lack of space, we did not present those results in this paper. The proposed method when tested under hyper and hypoglycemic ranges, produced significantly better results than the LMNN (36 % of improvement).

6.0 CONCLUSION

This work had proved the effectiveness of training the feed forward neural network with extracted features of inputs for predicting the future values of continuous glucose monitoring sensor data. The learning rate parameter derived from the extracted features, demonstrated the efficiency in computing the weights of neurons in back propagation.

This could well be understood from the performance evaluation between the FNN and the LMNN. Ninety percent of hypoglycemic occurrences were detected by the FNN even with a PH of 60 minutes which is more appreciable compared to previous works. With a simple network structure, the proposed method had exhibited promising results. 25% improvement in accuracy had been achieved when compared with LMNN method. The computational time for training was also acceptable with the high speed processors available now a day. In future, the prediction horizon could be extended to track the changing metabolism. However, a PH of 60 minutes would be sufficient for a diabetic person to take necessary action with hypoglycemia alert.

The sigmoid function used for activation in the LMNN is a globalized function. Whereas the localized feature set in the FNN increased the learning efficiency. The back propagation used in the LMNN might get trapped in local minima. When the network was unable to converge to a global minimum, the prediction accuracy would deteriorate. This could be the reason for the less promising results of the earlier approach. Exponential functions were better approximated by radial basis function neural network. However to predict the non linear dynamics of CGM time series, the specialized activation function with extracted feature set could be the best choice as proved by the results of this work.

Acknowledgement

We wish to convey our gratitude and thanks to the Diabetic Research Centre at Sushrutha Nursing Home, Tiruchirappalli for providing us the SMBG data and for valuable discussion. We also would like to extend our thanks to Medtronic Minimed CGMS® for helping us to get the continuous glucose monitoring sensor data.

References

[1] DCCT Group. "Intensive diabetes treatment and cardiovascular disease in patients with type1 diabetes", *N Eng J Medicine.;* 2005,353(25):2643-53.

[2] David C. Klonoff, "Continuous Glucose Monitoring- Roadmap to 21st century diabetes therapy" *Diabetes Care*, vol. 28, no.5, May 2005.

[3] Jay S.Skyler ,"Continuous Glucose Monitoring : An Overview of its Development", *Diabetes Technology & Therapeutics*, vol. 11, sup. 1, 2009.

[4] Atoosa Rabiee, et al.," Numerical and Clinical Accuracy of a Continuous Glucose Monitoring System during Intravenous Insulin Therapy in the Surgical and Burn Intensive Care Units*", J Diab.Sci Technol*. 2009 July; 3(4): 951–959.

[5] JDRF,Juvenile Diabetes Research Foundation, 2006: http://artificialpancreasproject.com, (Last accessed on April 30,2012)

[6] Bremer T, Gough DA. "Is blood glucose predictable from previous values? A solicitation for data". *Diabetes.* 1999;48(3):445–451.

[7] Palerm CC, Willis JP, Desemone J, Bequette BW. "Hypoglycemia prediction and detection using optimal estimation". *Diabetes Technol Ther.* 2005;7(1):3–14.

[8] Palerm CC, Bequette BW. "Hypoglycemia detection and prediction using continuous glucose monitoring–a study on hypoglycemic clamp data". *J Diabetes Sci Technol.* 2007;1(5):624–629.

[9] Sparacino G, Zanderigo F, Corazza S, Maran A, Facchinetti A, Cobelli C. "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time series*". IEEE Trans Biomed Eng*. 2007;54(5):931–937.

[10] Reifman J, Rajaraman S, Gribok A, Ward WK. "Predictive monitoring for improved management of glucose levels" . *J Diabetes Sci Technol.* 2007;1(4):478–486.

[11] Pappada SM, Camero BD, Rosman PM. "Development of a neural network for prediction of glucose concentration in type 1 diabetes patients" . *J Diabetes Sci Technol.* 2008;2(5):792–801.

[12] Gani A, Gribok AV, Rajaraman S, Ward WK, Reifman J. "Predicting subcutaneous glucose concentration in humans: data-driven glucose modeling". *IEEE Trans Biomed Eng.* 2009; 56(2):246–254.

[13] C.Perez-Gandia,A.Facchinetti,G.sparacino,C.cobelli,E.J.Gomez,M.Rigla,A.deLeiva,M.E.Hernando,"Artificial neural network algorithm for on-line glucose prediction from continuous glucose monitoring", *Diabetes Technol.Ther.* 2010,12,81-88.

[14] Robertson G, Lehman D, Sandham, Hamilton, "Blood Glucose Prediction Using Artificial Neural Networks Trained with the AIDA Diabetes Simulator: A Proof-of-Concept Pilot Study",*Journal of Electrical and Computer Engineering*, 2011, Article ID 681786,11 pages.

[15] UCSD(2008) : http://glucosecontrol.ucsd.edu (Last accessed on January 4, 2012).

[16] Hui Zheng et al., "Using A Multi-Level B-Spline Model to analyze and Compare Patient Glucose Profiles Based on Continuous Monitoring Data", *Diabetes Technol.Ther.* 2011,Vol.13,No. 6,675 - 682

[17] Fan, Jianqing & Yao, Qiwei, "*Spline Methods- Nonlinear time series: nonparametric and parametric methods*". New York,Springer-Verlog Publishing, 2005, ISBN:0-387-95170-9.

[18] N. S. Oliver, C. Toumazou, A. E. G. Cass and D. G. Johnston,' "Glucose sensors: a review of current and emerging technology ", 2009, *Diabetes Medicine* ,26, 197–210.
[19] Eda Cengiz and William V. Tamborlane , "A Tale of Two Compartments: Interstitial Versus Blood Glucose Monitoring". *Diabetes Technol Ther* ,vol.11, sup.1,2009.
[20] King.C, Anderson S.M, Breton.M, Clarke W.L, Kovatchev B.P., "Modeling of calibration effectiveness and blood to interstitial glucose dynamics as potential confounders of the accuracy of continuous glucose sensors during hyperinsulinemic clamp". *J.Diab.Sci.Technol.* ,2007,1,317-322.
[21] Gayane Voskanyan,Barry Keenan, John J.Mastrototaro and Garry M.Steil, "Putative delays in Interstitial Fluid (ISF) Glucose Kinetics Can be attributed to the Glucose Sensing systems Used to Measure them Rather Than the Delay in ISF Glucose Itself". *Jr. of Diab Science & Tech.,* vol.1,issue 5, Sep 2007.
[22] S.Shanthi, D.Kumar, "Continuous Glucose Monitoring Sensor Data: Denoising with Hybrid of Neural Network and Extended Kalman Filter Algorithm", *European Journal of Scientific Research,* Vol.66, No.2, 2011,PP:293-302.
[23] Lennart Ljung, "*System Identification: Theory for the User*", Second Edition,New Jersey,Prenticce Hall Publishing, 1999,ISBN:0-13-881640-9.
[24] Rhodes, C.; Morari, M. "The false nearest neighbors algorithm: An overview". *Computers & Chemical Engineering,* Volume 21, Supplement, 20 May 1997, Pages S1149–S1154
[25] Haykin,S, "*Neural Networks – A Comprehensive Foundation*", 1994,New York, Maxwell Macmillan International Publishing Company,ISBN:0-02-352761-7.

[26] Edward Gately,"*Neural Networks for Financial Forecasting"*, 1996, New Jersey,Wiley Publishing,ISBN:0471112127.
[27] Michael McInerney, Atam P. Dhawan. "Use of Genetic Algorithms with Back-Propagation in training of Feed-Forward Neural Networks", *Proc. of IEEE Int'l Joint Conference on Neural Networks*, Vol. 1, 1993,pp. 203- 208.
[28] V.Tresp, T.Briegel, and J.Moody,"Neural-network models for the blood glucose metabolism for a diabetic," *IEEE Transactions on Neural networks*,vol.10,1999, pp. 104-1213.
[29] Phee,H.K, Tung,W.L & Quek,C, "A Personalized approach to insulin regulation using brain-inspired neural semantic memory in diabetic glucose control". *IEEE congress on Evolutionary Computation*, Singapore, 2007, pp.2644-2651.
[30] S. M. Pincus, "Quantification of evolution from order to randomness in practical time series analysis," *Methods Enzymol.*, vol. 240, 2004, pp. 68–89.
[31] W. S. Evans, L. S. Farhy, and M. L. Johnson, "Biomathematical modeling of pulsatile hormone secretion: A historical perspective," *Methods Enzymol.*, vol. 454, 2009, pp. 345–366.

BIOGRAPHY

S.SHANTHI : Received the B.E. degree in Electronics and Communication Engineering from Madurai Kamaraj University in 1994 and M.E. degree in Communication engineering from Anna University Chennai in 2005. She is currently pursuing the Ph.D Research in area of Bio Medical Engineering in Anna University of Technology-Tiruchirappalli. Her areas of interest are Bio signal processing, closed loop control of Blood Glucose. She is in the teaching profession of nearly 11 years. And has published research papers in 3 international journals and 2 papers in international conferences.

D.KUMAR : received his PhD degree in Optical Characterization and Imaging of Biological Tissues by Simulation and Laser Reflectometry from Indian Institute of Technology, Madras in the year 2004 and done Postdoctoral Research in Bioluminescence Tomography at University of Iowa in the year of 2007. Currently he is the Head of the Department of Electronics and Communication engineering and Dean – Research at Periyar Maniammai University, Vallam, Thanjavur. He has 14 Journal Publications and 17 Conference Publications. He has conducted a conference on the topic "PRICIPAL COMPONENT ANALYSIS " with grant form AICTE and he has been as the Invited Speaker to the Indo Russia Workshop (Aug 22-27, 2009) Moscow State University, Moscow, with funding from DST, Govt. of India.