Multi-Model Data Fusion to Improve an Early Warning System for Hypo-/Hyperglycemic Events

Ransford Henry Botwey – *IEEE Student Member*, Elena Daskalaki - *IEEE Member*, Peter Diem, Stavroula G. Mougiakakou - *IEEE Member*

Abstract— **Correct predictions of future blood glucose levels in individuals with Type 1 Diabetes (T1D) can be used to provide early warning of upcoming hypo-/hyperglycemic events and thus to improve the patient's safety. To increase prediction accuracy and efficiency, various approaches have been proposed which combine multiple predictors to produce superior results compared to single predictors. Three methods for model fusion are presented and comparatively assessed. Data from 23 T1D subjects under sensor-augmented pump (SAP) therapy were used in two adaptive data-driven models (an autoregressive model with output correction – cARX, and a recurrent neural network – RNN). Data fusion techniques based on i) Dempster-Shafer Evidential Theory (DST), ii) Genetic Algorithms (GA), and iii) Genetic Programming (GP) were used to merge the complimentary performances of the prediction models. The fused output is used in a warning algorithm to issue alarms of upcoming hypo-/hyperglycemic events. The fusion schemes showed improved performance with lower root mean square errors, lower time lags, and higher correlation. In the warning algorithm, median daily false alarms (DFA) of 0.25%, and 100% correct alarms (CA) were obtained for both event types. The detection times (DT) before occurrence of events were 13.0 and 12.1 min respectively for hypo-/hyperglycemic events. Compared to the cARX and RNN models, and a linear fusion of the two, the proposed fusion schemes represents a significant improvement.**

I. Introduction

Prevention of hypo- and hyperglycemic events is of high importance for individuals with Type 1 Diabetes (T1D). The use of data from Continuous Glucose Monitoring (CGM) devices along with data-driven modeling techniques have been proposed for predicting glucose profile and for generating alerts whenever these events are forecasted. Individuals with T1D are thus enabled to take appropriate actions (e.g. carbohydrate intake or suspension of insulin infusion in case of a hypoglycemic event, and additional insulin or physical activity in case of a hyperglycemic event) to prevent the onset of the event. Specifically, several data-driven approaches have been proposed for the short term prediction of future glucose concentration in T1D patients. Most of these use past record of glucose measurements from a CGM device, with or without record of insulin intake and carbohydrate intake as input to different models such as autoregressive based [1]-[6], artificial neural networks (ANN) [2], [7], [8] and support vector machines [9]. Studies have shown that the various prediction models present different advantages and disadvantages. It has also been shown that the real time prediction of future glucose concentration in T1D patients loose its accuracy as the prediction horizon (PH) is increased [2], [7], making it difficult to detect and predict hypo-/hyperglycaemic events with certainty. A comparison of three different models [2] showed quite significant differences in performance in terms of root mean square error (RMSE), correlation coefficients (CC), and sensitivity of hypo-/hyperglycemia predictions. Moreover, the influence of model-specific characteristics is seen for example in the ANN model's tendency to overestimate hypoglycemic events despite its efficient prediction of normoglycemia and hyperglycemia [2], [8]. In summary, no specific method provides 100% reliable event predictions. To surpass this problem, a number of approaches have been proposed for combining multiple predictors with the hope of obtaining results superior to that of single predictors. Proposed methods include alarm triggering of hypoglycemic events by a voting scheme fusion of different prediction models [10], and the use of Bayesian combination rules [11]. An early warning system (EWS) based on a linear fusion of two adaptive data-driven models has also been proposed [12].

The proposed methods for combining multiple glucose prediction models are simple and do not take into consideration inherent uncertainties related to the problem. Thus, more sophisticated approaches for data fusion need to be investigated. Data fusion deals with amalgamation of data from different sources where the original data alone are not enough to draw conclusions from about the state of a quantity being observed. The major potential advantage is that by integrating complementary information from different sources, an improved and more reliable performance can be achieved [13]. Fields of application of data fusion include defense applications such as automated target recognition, area surveillance, guidance and control of autonomous vehicles, remote sensing, and medical imaging and diagnosis.

Aim of this study is to develop an efficient strategy for a reliable short-term prediction of glucose concentration in individuals with T1D, with emphasis on hypoglycemia prediction. This involves the implementation and

R. H. Botwey is with the Diabetes Technology Research Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Switzerland.

E. Daskalaki is with the Diabetes Technology Research Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Switzerland.

S. G. Mougiakakou (corresponding author) is with the Diabetes Technology Research Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Murtenstrasse 50, 3010 Bern, Switzerland, Email: stavroula.mougiakakou@artorg.unibe.ch.

Peter Diem is with the Department of Endocrinology, Diabetes & Clinical Nutrition, Bern University Hospital (Inselspital), Switzerland.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

comparative assessment of advanced fusion schemes for data-driven glucose prediction models. It is expected that a combination of two or more modeling methods to take advantage of their complementary prediction performances may lead to i) accurate prediction of hypo-/hyperglycemic events, ii) less false alarms, and iii) improved detection times, leading to the enhancement of the EWS presented in [12].

II. Materials And Method

The flowchart towards the development of an EWS based on advanced data fusion techniques is presented in Fig. 1. Data from an individual with T1D under sensor augmented pump (SAP) therapy are used as input to a number of data-driven models. The models' outputs are fused and fed to a decision algorithm for the generation of warnings whenever a hypo-/hyperglycemic event is predicted.

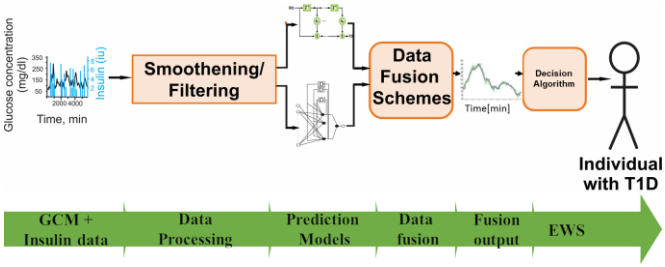

Figure 1. Flowchart of the hypo-/hyperglycemia prediction and warning generation.

A. Data

CGM data from 23 individuals with T1D under SAP therapy with the following average anthropometrics: 17 to 70 years of age, 24.2 ± 4 kg/m^2 Body Mass Index (BMI), and 7.3 ± 0.7% glycosylated hemoglobin (HbA1c) were used. All patients used Medtronic insulin pumps (Medtronic MiniMed Inc., Northridge, CA, USA) combined with a real-time CGM system under normal daily living conditions. The sensor glucose values were equally sampled every 5 minutes. Detailed presentation of the used data is given in [12]. Half of the dataset for each patient was used in training and identification of the parameters of the prediction models, and the other half was used for evaluation of the model's performance.

B. Prediction Models

The two adaptive data-driven models (an autoregressive model with output correction –cARX, and a recurrent neural network –RNN) presented in [12] were used. Each model uses sensor glucose and insulin pump data in order to predict the glucose profile in PH of 15, 30 and 45 minutes.

C. Fusion Schemes

Data fusion employs technologies from a wide range of areas including artificial intelligence, pattern recognition, and statistical estimation. For this specific problem two different technologies will be investigated for merging the outputs of the glucose prediction models. The technologies are based on the Dempster-Shafer evidential theory – DST (one method), and on evolutionary programming (two methods).

1. Dempster-Schafer Evidential Theory

The DST is an extension to probability, and is ideal for modeling uncertainties in systems; thus enabling the fusion of information (evidence) from different sources to obtain a unified statement with a higher degree of certainty [14]. Evidence has a two-dimensional mass which comprise of (1) the degree of belief of a hypothesis/set A or lower probability function– $Bel(A)$, and (2) the plausibility of A or upper probability function – $Pl(A)$. All the possible system states are contained in the frame of discernment Θ. A probability mass m is assigned to every element in 2^Θ (m: $2^\Theta \rightarrow [0,1]$), with m satisfying the conditions:

$$m(\varnothing) = 0 \quad \text{and} \quad \sum_{A \in 2^\Theta} m(A) = 1 \tag{1}$$

The belief $Bel(A)$ for a set A is defined as the sum of all probability masses of subsets of the set of interest:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

The plausibility $Pl(A)$ is the sum of all the probability masses of the sets B that intersect the set of interest A:

$$Pl(A) = 1 - \sum_{B \cap A = \phi} m(B) \tag{3}$$

The DST combination rule is used to combine evidence from two sources by means of their probability masses (m_1 and m_2). The combination rule is defined as

$$m(C) = [m_1 \oplus m_2](C) = \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{(1-k)} \tag{4}$$

where $k = \sum_{A \cap B = \phi} m_1(A) m_2(B)$ is the degree of disagreement among the sources. Fig. 2 shows the DST fusion process.

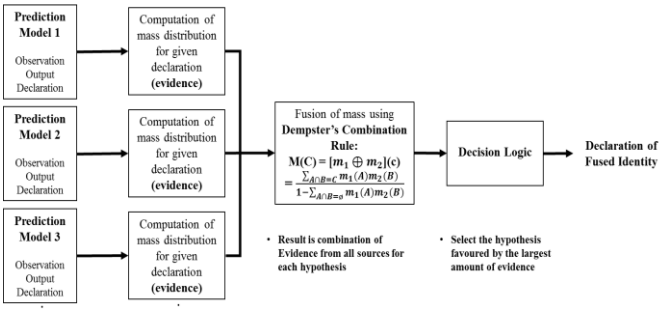

Figure 2. DST Fusion process –Fusion of evidence from three sources.

In this specific problem, Θ is obtained from the three main states hypo-, normo-, and hyperglycemia, giving a power set (2^Θ) of 2^3=8 possible states (propositions), with the two adaptive data-driven models as data sources (S_1, S_2). Each predicted data (from one of the participating prediction models) falls into one of the three states. Through the DST fusion scheme the predictions of all participating models at time t are merged into one output with a specific predicted

state (fused identity). Using a probability distribution scheme (on an interval scale from 20 to 599 mg/dL) the glucose value is calculated as predicted consensus output at time *t*.

2. Fusion using Evolutionary Methods

Evolutionary programming methods are based on the Darwinian principles of natural selection. Initially a population of possible solutions to the given problem is generated. A fitness function is defined and used to evaluate the results in each generation. Genetic operations (crossover, mutation, and reproduction) are applied to the best solutions in the population with the hope of introducing innovations and improving the solutions of succeeding generations. With evolutionary methods hybrid systems which combine two or more intelligent systems can be created. Through their typical characteristics of performing random search in the solution space, they are able to find optimal solutions, making them flexible and adaptive to many problems. Evolutionary methods have been successfully used in data fusion application [15]. In this specific problem two different algorithms were used: genetic algorithms and genetic programming.

2.1 Genetic Algorithms

Genetic Algorithms (GA) are search algorithms based on the mechanics of natural selection and natural genetics, and which combine survival of the fittest among string structures with a randomized yet structured information exchange to facilitate innovations in search and optimization problems [16]. In this fusion application, each finite-length string (chromosome) is made up of eight genes (bits) with values (alleles) from the defined alphabet (0, 1). Each individual in the population (output value from a prediction model) is represented by two chromosomes (16 bits). In addition to the 16 bits, each individual has 4 more bits, two of which code the classification of that individual's state, one bit codes the confidence/fitness value, and the last bit codes the prediction time. The initial confidence values is the probability of that individual's value corresponding to its classification. Fig. 3 shows how each predicted data is encoded into bits of strings in the GA fusion application.

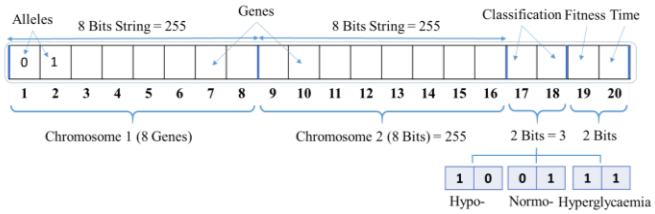

Figure 3. Encoding of an individual (predicted glucose) in the population into a 20-bit string.

Fitness function, selection Criteria, and run parameters: The fitness function enables the discovery of optimal solutions based on a well-defined equation. However, unlike in most applications, there is no specific equation for determining glucose concentrations as this is dependent on the constantly varying physiological system. Three strategies for finding optimal fusion, and thus fit solutions were implemented. These include matching/fitting of evolved solutions to; 1) Prediction with the highest confidence factor, 2) Mean of prediction model's outputs, and 3) Similarity Template – ST (which uses past glucose samples and output of prediction models). Best results were obtained with the ST using the following run parameters: 50 generations, 10x replication of predicted data, lowest fitness for selection: 0.6, and probability of mutation pm < 0.5. The fitness function based on ST is given by equations (5) and (6) below.

$$\text{Template: } T(t) = \frac{\sum_{i=1}^{m} D_i(t)}{m} \qquad (5)$$

where D_i is day *i*'s glucose data at time *t*, and *m* is the number of days.

$$\text{Fitness function: } M(t) = \frac{\sum_{i=1}^{n} S_i(t) + T(t)}{n+1} \qquad (6)$$

where $S_i(t)$ is the predicted value of prediction model *i* at time *t*, T is the template value and *n* is number of models.

2.2 Genetic Programming

Genetic programming (GP) is also based on Darwin's principle of natural selection. Pre-defined functions are randomly selected and used to automatically create programs which try to find the best solution to a given problem. In our model fusion application, 14 basic numerical and Boolean functions which form the root and nodes of the GP-Program trees were defined. Each function takes two inputs (confidence values from prediction models). The leaves of the program trees are the fused confidence values (max=1). With each prediction data pair, 20 trees of variable height (max: 10) are created. Each tree has two outputs (from left and right tree branches).

Fitness function, selection criteria, and run parameters: In addition to the fitness function definition for the GA method, a receiver operating characteristic (ROC) curve is used to ascertain how well the fused data fits to the fitness function.

D. Decision Algorithm

The fused output is fed to a decision algorithm which is designed to produce warnings using a series of rules. Details are provided in [12].

III. Results And Discussion

The performances of the three fusion schemes (DST, GA, and GP) were evaluated on the basis of the RMSE, time lag (TL), and CC for PH of 15, 30, and 45 minutes. Table I shows a comparative assessment of the performances of the cARX and RNN models with the fusion schemes, while Table II compares the original EWS with the EWSs based on the advanced fusion schemes presented in this paper (EWS-DST; EWS-GA; EWS-GP). Among the three proposed fusion schemes the evolutionary based (GA and GP) presented superior performance over the DST both for prediction and alerts generation. Besides, the fusion schemes improved on the prediction performance of the individual models, with GA and GP presenting lower TLs and RMSEs,

and higher correlation (Table 2). Furthermore, from Table 3 it can be seen that the EWS-GA and EWS-GP outperformed the initially proposed EWS in terms of daily false alarm. However the detection time in advance of hypo- and hyperglycemic events was slightly reduced. In terms of prediction accuracy an equal performance of 100% was obtained.

TABLE I. MEDIAN (5TH - 95TH PERCENTILES) PREDICTION PERFORMANCE OF THE CARX AND RNN MODELS COMPARED WITH THAT OF THE DST, GA AND GP FUSION SCHEMES

Criteria	cARX	RNN	DST-F	GA-F	GP-F
PH = 15 minutes					
TL	5.0 (0.0-9.5)	5.0 (0.0-10.0)	10.0 (0.5-10.0)	0.0 (0.0-5.0)	0.0 (0.0-9.5)
RMSE	16.8 (11.3-33.8)	11.9 (7.7-22.7)	15.3 (10.14-27.9)	8.5 (6.0-18.6)	9.2 (6.1-21.4)
CC	0.96 (0.87-0.97)	0.98 (0.95-0.99)	0.97 (0.87-0.98)	0.99 (0.96-1.00)	0.99 (0.96-1.00)
PH = 30 minutes					
TL	15.0 (10.0-24.5)	10 (5.5-15.0)	20.0 (10.0-25.0)	5.0 (0.0-14.5)	10.0 (5.0-20.0)
RMSE	27.7 (19.0-49.5)	18.9 (12.8-32.3)	26.5 (18.48-42.6)	13.8 (10.4-27.3)	17.7 (12.8-33.7)
CC	0.90 (0.71-0.93)	0.94 (0.89-0.96)	0.91 (0.70-0.95)	0.97 (0.92-0.98)	0.95 (0.88-0.97)
PH = 45 minutes					
TL	30.0 (20.5-39.5)	20.0 (10.0-25.0)	30.0 (15.5-40.0)	10.0 (0.0-20.0)	20.0 (10.0-30.0)
RMSE	37.0 (25.4-61.1)	26.1 (17.2-39.8)	34.1 (25.35-54.5)	19.2 (13.2-31.9)	25.0 (16.7-39.9)
CC	0.82 (0.58-0.88)	0.90 (0.78-0.93)	0.82 (0.60-0.90)	0.94 (0.87-0.97)	0.90 (0.80-0.94)

TABLE II. PERFORMANCE OF THE EWS AND THE EWS BASED ON DST (EWS-DST), GA (EWS-GA) AND GP (EWS-GP) FUSION SCHEMES IN MEDIAN (5TH - 95TH PERCENTILES). [CA: CORRECT ALARMS, DT: DETECTION TIME, DFA: DAILY FALSE ALARMS]

Criteria	EWS	EWS-DST	EWS-GA	EWS-GP
(A) Hypoglycemia				
CA (%)	100 (100-100)	100 (100-100)	100 (93-100)	100 (94-100)
DT (min)	16.7 (10-25)	18.4 (10-31)	13.0 (10-25)	12.3 (9-20.0)
DFA	0.8 (0.0-1.2)	1.0 (0.17-3.4)	0.25 (0-0.65)	0.17 (0-0.8)
(B) Hyperglycemia				
CA (%)	100 (95-100)	100 (92-100)	100 (93-100)	100 (92-100)
DT (min)	14.7 (5.1.-19)	11.6 (7.2-18)	12.1 (8-16.1)	12.0 (8 -16.2)
DFA	0.8 (0.0-1.4)	0.73 (0.3-1.7)	0.25 (0-0.98)	0.33 (0.0-1.2)

IV. CONCLUSION

These experiments have shown, that data fusion schemes can be used to merge the complimentary performances of different glucose prediction models, with the aim of removing uncertainties and inaccuracies associated with them. Thus, accuracy in functionality, and minimization of false alarms in the detection of hypo-/hyperglycemic events in T1D patients is feasible. From the results, it can be seen that the fusion schemes permit higher detection accuracy with much lower false alarm rates. Further investigation is needed in order to improve the detection time. Furthermore, it has to be noted that the proposed fusion schemes can be extended to include additional data-driven glucose prediction models, e.g. SVM.

Finally, the proposed EWS based on advanced fusion of glucose prediction models might be a useful tool for hypoglycemic unawareness, while it can be used as a safety module in insulin pumps and controllers for an artificial pancreas.

REFERENCES

[1] M. Eren-Oruklu, A. Cinar, D.K. Rollins, L. Quinn, "Adaptive system identification for estimating future glucose concentrations and hypoglycemia alarms," *Automatica*, vol. 48, no. 8, pp. 1892–1897, 2012.

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

[3] M. Eren-Oruklu, A. Cinar, L.Quinn, "Hypoglycemia prediction with subject-specific recursive time-series models," *Journal of Diabetes Science and Technology,* vol. 4, no. 1, pp. 25-33, Jan. 2010.

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

[5] A. Gani, A.V. Gribok, Y. Lu, W.K. Ward, R.A. Vigersky, J. Reifman, "Universal glucose models for predicting subcutaneous glucose concentration in humans," *IEEE Transactions on Information Technology in Biomedicine,* vol. 14, no. 1, pp. 157-165, Jan. 2010.

[6] D.A. Finan, F.J. Doyle, C.C. Palerm, W.C. Bevier, H.C. Zisser, L. Jovanovic, D.E. Seborg, "Experimental evaluation of a recursive model identification technique for type 1 diabetes," *Journal of diabetes science and technology,* vol. 3, no. 5, pp. 1192-202, 2009.

[7] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, C. Cobelli, "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series", *IEEE Trans. Biomed. Eng.,* vol 54, no. 5, pp. 931-937, 2007.

[8] S.M. Pappada, B.D. Cameron, P.M. Rosman, "Development of a neural network for prediction of glucose concentration in type 1 diabetes patients", *Journal of diabetes science and technology*, vol. 2, no. 5, pp. 792-801, 2008.

[9] E. Georga, V.C. Protopappas, D. Ardigo, M. Marina, I. Zavaroni, D. Polyzos, D.I. Fotiadis, "Multivariate prediction of subcutaneous glucose concentration in type 1 diabetes patients based on support vector regression", *IEEE J Biomed. & Health Informatics*, vol. 17, no. 1, pp. 71-81, 2013.

[10] E. Dassau, F. Cameron, H. Lee, B.W. Bequette, H. Zisser, L. Jovanovic, H.P. Chase, D.M. Wilson, B.A. Buckingham, F.J. Doyle, "Real-Time hypoglycemia prediction suite using continuous glucose monitoring," *Diabetes Care*, vol. 33, no. 6, pp. 1249-1254, 2010.

[11] F. Stähl, R. Johansson, E. Renard, "Bayesian combination of multiple plasma glucose predictors", *IEEE EMBC2012,* pp. 2839-44, 2012.

[12] E. Daskalaki, K. Nørgaard, T. Zűger, A. Prountzou, P. Diem, S. Mougiakakou, "An early-warning system for hypo-/hyperglycaemia events based on fusion of adaptive prediction models," *J. Diabetes Science Technology*, vol. 7, no. 3, pp. 689-698, May 2013.

[13] K.L. Crowley, Y. Demazeau, "Principles and technique for sensor data fusion," *Signal Processing*, vol. 23, no. 1-2, pp. 5-27, 1993.

[14] T.L. Wickramarathne, K. Premaratne, M.N. Murthi, "Toward efficient computation of the Dempster–Shafer Belief theoretic conditionals," *IEEE Transactions on Cybernetics*, vol. 43, no. 2, pp. 712-724, 2013.

[15] B.F. Buxton, W. B. Langdon, S.J. Barrett, "Data fusion by intelligent classifier combination," *Measurement and Control*, vol. 34, no. 8, pp. 229-234, 2001.

[16] D. E. Goldberg, "*Genetic Algorithms in Search, Optimization and Machine Learning*", Boston: Addison-Wesley, 1989 (Reprint 2006).