Data-Based Identification of Prediction Models for Glucose

J. Manuel Velasco[1] mvelascc@ucm.es
Stephan Winkler[2] Stephan.Winkler@fh-hagenberg.at
J. Ignacio Hidalgo[1] hidalgo@ucm.es
Oscar Garnica [1] ogarnica@ucm.es
Juan Lanchares[1] julandan@ucm.es
J.Manuel Colmenar[1] josemanuel.colmenar@urjc.es
Esther Maqueda[3] esthermaq@gmail.com
Marta Botella marta.botella@salud.madrid.org
Jose-Antonio Rubio joseantonio.rubio@salud.madrid.org

*Adaptive and Bioinspired Systems Group (ABSys)
Facultad de Informática
Universidad Complutense de Madrid (Spain)

[2]School of Informatics, Communications and Media
University of Applied Sciences Upper Austria

[3]Endocrinology and Nutrition Service
Hospital Virgen de la Salud, Toledo (Spain)

[4]Endocrinology and Nutrition Service
Hospital U. Príncipe Asturias, Alcalá de Henares (Spain)

ABSTRACT

Diabetes mellitus is a disease that affects to hundreds of million of people worldwide. Maintaining a good control of the disease is critical to avoid severe long-term complications. One of the main problems that arise in the (semi) automatic control of diabetes, is to get a model explaining how glycemia (glucose levels in blood) varies with insulin, food intakes and other factors, fitting the characteristics of each individual or patient. In this paper we compare genetic programming techniques with a set of clsssical identification techniques: classical simple exponential smoothing, Holt's smoothing (linear, exponential and damped), classical Holt and Winters methods and auto regressive integrated moving average modelling. We consider predictions horizons of 30, 60, 90 and 120 minutes. Experimental results shows the difficulty of predicting glucose values for more than 60 minutes and the necessity of adapt GP techniques for those dynamic enviroments.

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*Heuristic methods*; G.1.6 [**Numerical Analysis**]: Optimization—*Global optimization*

Keywords

Genetic Programming, Modeling, Diabetes

1. INTRODUCTION

Diabetes Mellitus (DM) is a disease affecting more than 300 millions people in the world. Many factors influence the appearance of Diabetes, but we can generalize saying that all patients suffer a defect in either the secretion or in the action of insulin, which is essential for the control of blood glucose levels. The result is that cells do not assimilate glucose and, as a consequence, there is a rise in blood glucose levels (or hyperglycemia).

Roughly speaking we can find two main kinds of diabetes; Type 1 (T1DM) and Type 2 Diabetes (T2DM). In T1DM the pancreatic cells do not produce insulin due to an autoinmune process. On the other hand, a patient with T2DM suffers from insulin resistance since the cells that need the glucose fail to use it properly. In all the patients mainting a good glycemic control is esentential to avoid not only short term, but also long terms complications. One of the most serious short term complication is a diabetic coma caused by a very low level of glycemia (or hypoglycaemia), which can eventually devent on the death of the patient. Among the long term problems we can mention blindness, renal failure, sores and infections in feet, damage to nerves in the body, etc.

The good new is that most of the patient with good control of the levels of glycemia have a normal life and avoid (or at least delay) the appearance of complications. However, this is not an easy task. Maintaining an approapiate control of the glucose implies measures of glucose in blood using a Continous Glucose Monitoring (CGM) system or/and Glucose Meters (GM). The patient also needs to count or estimate the amount of carbohydrates is going to eat and to have the knowledge for deciding the units of insulin he/she will need in order to remain on good glucose levels after the meals. This is the cause that a high percentage of diabetics (around 50 %) do not achieve a real control of their glucose levels. The ideal solution would be an artificial pancreas (AP) and this is the main area of research in the field.

To achieve a completely autonomous glycemic control, a control algorithm (CA) receiving information from a CGM system would be necessary. By forescating the evolution of blood glucose, using a predictive model of the response system, the CA would indicate to an insulin pump when to inject a bolus of insulin and the amount necessary. This system is usually called an AP. One of the main problems for the development of the AP is the lack of accurate models for predicting the future of the glucose. Although there are some clasical approxiamtions there is still to much to do for predictions within an horizon of more than 90 minutes.

Genetic Programming (GP) [1] has proven to be effective on other predicition problems. In particular Winkler et al. have applied

GECCO'15, July 11-15, 2015, Madrid, Spain.

http://dx.doi.org/10.1145/2739482.2768508.

symbolic regression based on GP using a structure identification framework described in [2] and [3] in other identification problems. In this paper we investigate if the mentioned methods could be succesful for glucose level predictions in humans. We will analyze GP results and compare them with those obtained with a set of classical identification techniques. In particular we have implemented classical Simple Exponential Smoothing (SES), three Holt's smoothing approaches (linear, exponential and damped), classical Holt and Winters methods and auto regressive integrated moving average (ARIMA) modelling. For a full explanation of all the classical techniques that we have employed, we refer to [4].

Experimental Results shows that GP can sometimes outperforms classical techniques for predictions on the horizon of 30 and 60 munites. Unfortunately, we can not say that we obtain good results for 90 and 120 minutes. Results show that it is very difficult to predict glucose values for horizons of more than 60 minutes. We present here some insight on this issue, by analyzing espectral properties of the data of some simulated and real patients.

The rest of the paper is organized as follows. Section 2 reviews some of the previous approaches on glucolse level prediction. Section 3 explains the methodology applied in this paper, with the description of the techniques, and patients characteristics of both in-silico and real input data. Section 4 describes the experimental setup, while 5 analyzes the experimental results. On Section 6 we propose some future work and conclude the paper.

2. RELATED WORK

The problem of predicting and modelling glucose levels has been an intensive area of research during the last 10 years. Two are the main targets of these studies. Some of them tried to predict the glucose levels with a time horizon of up to two hours, since this is usually the time step needed by the patient to be confortable after a meal. There are also some researchers that tried to identifiy 24 hours models. The utility of the last is different and is usually more effective when programming an insulin pump or when establishing an insulin profile for longer periods.

We can find in literature some approximations providing models for the average case [5]. However, there are hardly few approaches adapted to the particularities of each patient. Most of the models in the literature apply classical modeling techniques, resulting in linear equations defined profiles, or models with a limited set of inputs [6]. One exception is [7], where Heusden et al. proposed the use personalized information of the patient. The type of information is not clearly described since is aomthing mentioned as easily accesible by the specialist or the automatic system. The main problem with this approach is that it can not incorporate other important factors such as exercise or stress that clearly affect glycemias and it is only useful with linear models and.

Recently Hidalgo et al. proposed the application of evolutionary computation techniques to obtain customized models of patients, unlike most of previous approaches which obtain averaged models. The proposal is based on a kind of genetic programming based on grammars known as Grammatical Evolution (GE). The proposal has been tested with in-silico patient data and results are clearly positive. Authors present also a study of four different grammars and five objective functions[8].In the test phase the models characterized the glucose with a mean percentage average error of 13.69 %, modeling well also both hyper and hypoglycemic situations. However the application of this technique for short time predictions is still under research, since it will be necessary to adapt the grammars for dynamic enviromments.

Other personalized control approaches were presented by the main research groups on AP. Those are proposals following the clinical practice. Treatment for subjects with T1DM uses rates of basal insulin delivery, insulin to carbohydrate ratios and individual correction factors, typically from observations of the specialist. However, those models are often inaccurate, since clinical data in T1DM are not extensive enough to identify the exact models [9].

There are also some models used in artificial pancreas systems or closed loop control models: artificial pancreas systems are closed loop control systems trying to emulate the action of a pancreas [10] They are based on the assumption that it is possible to reach a good control with approximate models, provided that the model is related to the control objective [11].The main risk is hypoglycemia as a reaction to an excessive insulin administration, usually due to the lack of accurate individualized models.

We can find also Autoregressive models (AR) [12], protocols to improve the reliability of the models [10], and solutions by commercial companies Those approaches were designed to facilitate the control of diabetes but they work only with the glucometers of the company.

Evolutionary computation has a high potential to incorporate to the model factors which are difficult to quantify, in other words to collect system dynamics, allowing to obtain individualized models since they are able to provide a solution for each set of data on a single patient.

3. METHODOLOGY

As we have already mentioned the aim of this paper is to test the ability of GP for obtaining models and predictions of glucose levels in humans. We will also try to check what happens with classical techniques when applied to the same problem instances. After analyzing both options, we will be able to draw conclusions and see how to deal with future work.

The methodology of the research is as follows: First, we try both, the classic techniques and GP with data from in-silico patients insilico. We will work with data instances already studied in previous articles [13]. For each of the techniques, and for each patient obtain a model using as training data a 24 hour instance. Then, using this model try to make predictions for values of the glucose after 30, 60, 90 and 120 minutes. For the predictions we will use another 24 hour data instance generated by the simulatir. With the analysis of the results and data input will try to get some light on the difficulty of predictions on time horizons longer than 60 minutes.

Hence, in this section we describe the process of obtaining models and then we present a brief description of the different techniques used in the paper. Details about data and simulator will be given at Section 4.

3.1 Problem description

Figure 1 shows the process of optimization and modeling. Starting from a set of historical data including glucose data and of other events of the patient, measured every 15 minutes for 24 hours, a model of the data is obtained. From this mode,l and using another 24 hours dataset, the validation process for 30, 60, 90 and 120 minutes is performed.

Le us suppose that the patient is using two types of insulin, IS (Short term insulin) and IL (Long term insulin). We have registered the set of measures explained below. Given a set of measures

$$GL = \{gl_0, gl_1, \cdots, gl_n\}$$

of the glucose level of a patient, and knowing that gl_i was measured at time t_i. Given a set of intakes

$$CH = \{ch_0, ch_1, \cdots, ch_n\}$$

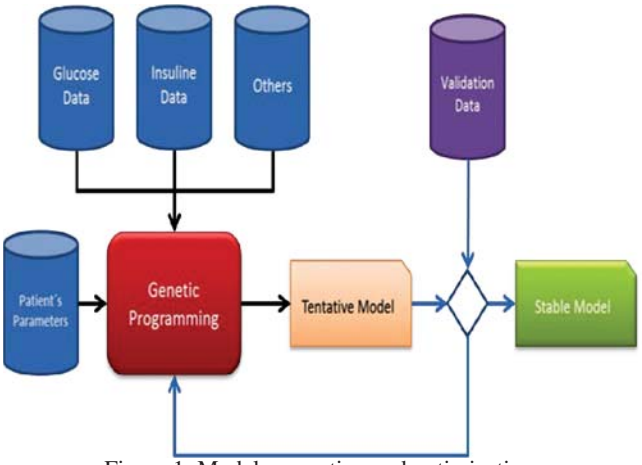


Figure 1: Model generation and optimization.

k	GL	CH	IS	IL
...	...	...	...	...
30	170.88974	0	0	0
31	171.55425	0	0	0
32	172.27976	0	3	12
33	173.05923	30	0	0
34	174.09018	0	0	0
...	...	...	...	...
40	237.54628	0	0	0
41	247.25104	20	0	0
42	250.72465	0	0	0
43	251.90543	0	0	0
...	...	...	...	...

Table 1: Portion of a 24-hours dataset for a in-silico patien (Joy Wilson).

of carbohydrates units of a patient, and knowing that ch_i was eaten at time t_i. Given a set of injections of insulin of type S

$$IS = \{is_0, is_1, \cdots, is_n\}$$

and knowing that is_i was injected at time t_i. Given a set of injections of insulin of type L

$$IL = \{il_0, il_1, \cdots, il_n\}$$

and knowing that il_i was injected at time t_i. The problem of modeling the glucose blood level of a patient, in its simplified form, can be formulated as follows: Find an expression of estimated glucose values, denoted as $\widehat{GL}$

$$\widehat{GL}(k+1) = f(\widehat{GL}, CH, IS, IL), 0 \leq k \leq N \quad (1)$$

which minimizes the fitness function F, which tries to close the gap between real and estimated glucose values:

$$F = \sum_{i=0}^{n} \sqrt{(GL(i) - \widehat{GL}(i))^2} \quad (2)$$

Where $\widehat{GL}$ corresponds to previous estimated glucose values, CH corresponds to previously ingested carbohydrates and IS and IL correspond to previously injected insulin for both types, short and long effect. It should be noted that the model will provide estimated glucose values, denoted as $\widehat{GL}$. Hence, for each time step, estimated glucose is obtained by using previous estimated glucose values and actual carbohydrates and insulin units. Therefore, the dataset should provide input values for the variables in our glucose model proposal. Table 1 shows a reduced version of a data set for one of the in-silico patients under study. Table 1 shows an example of the dataset of one of the in-silico patients, named Joy Wilson. For each time step, represented in one line of the table, k is the actual time, GL is the actual glucose level, CH is the carbohydrates units ingested, IS is the short effect insulin injected and IL is the long effect insulin injected.

Now we will briefly descrribe de set of classical techniques applied and also the features ans structure of the GP prediction tool.

3.2 Classical simple exponential smoothing

When we have a time series with no visible trend or seasonal behaviour we can employ one of the exponential methods of smoothing to forecast future values. The most basic version of them is called: *Simple Exponential Smoothing* or SES and it is shown in its general form in the Equation 3. In the instant t, We want to get a prediction with a horizon of H minutes, so we use the weighted observations during period P, previous to instant t. As we can see in this equation, we can change the weights associated with each observation varying the α parameter. In the dataset we have experimented with, we found that the best results were for values of α close to one. this way, Equation 3 becomes Equation 4 and our predictions at horizon H are just the value of the glucose at time t. So, we hope that a good forecasting technique can, at least, beat SES (we will see in the section 5 that when H = 120 this is not easy).

$$\hat{y}_{t+H|t} = \sum_{j=0}^{P-1} \alpha(1-\alpha)^j y_{t-j} \quad (3)$$

$$\alpha \approx 1 \Rightarrow \hat{y}_{t+H|t} = y_t \quad (4)$$

3.3 Holt's smoothing: linear, exponential and damped

Now, let's take into account the possibility of our glucose series is exhibiting some form of trend (constant or not). So, in Equation (5) we add a slope component to SES and, therefore, we need now two smoothing parameters, Equations (6) and (7). This is the basis of the Holt's method (1957) and we call it: HOLT_linear as the forecasts are a linear function of h.

$$\hat{y}_{t+H|t} = Level_t + H * Slope_t \quad (5)$$

$$Level_t = \alpha_0 y_t + (1-\alpha_0)(Level_{t-1} + Slope_{t-1} \quad (6)$$

$$Slope_t = \alpha_1(Level_t - Level_{t-1}) + (1-\alpha_1)Slope_{t-1} \quad (7)$$

$$\hat{y}_{t+H|t} = Level_t * Slope_t^H \quad (8)$$

$$Level_t = \alpha_0 y_t + (1-\alpha_0)(Level_{t-1} * Slope_{t-1} \quad (9)$$

$$Slope_t = \alpha_1(\frac{Level_t}{Level_{t-1}}) + (1-\alpha_1)Slope_{t-1} \quad (10)$$

$$\hat{y}_{t+H|t} = Level_t + (\phi + \phi^2 + \cdots + \phi^n)Slope_t \quad (11)$$

$$Level_t = \alpha_0 y_t + (1 - \alpha_0)(Level_{t-1} + \phi Slope_{t-1} \quad (12)$$

$$Slope_t = \alpha_1 (Level_t - Level_{t-1}) + (1 - \alpha_1)\phi Slope_{t-1} \quad (13)$$

We have also employed two variants of the Holt's method that we have called HOLT_exp and HOLT_damped. In the first, instead of adding, we multiply the level and the slope, Equations (8), (9) and (10). In the second, we have added a damping parameter ϕ (between 0 and 1), Equations (11), (12) and (13). This conservative idea was proposed by Gardner and McKenzie (1985) and, as we will see, suits well glucose time series for the highest values of H.

3.4 Classical Holt and Winters

The most general expression of the Holt's method that includes a seasonal component is the HOLT-WINTERS method. In the Equations (14), (15), (16) and (17), we show the additive version in which the seasonal variations don't depend on the level of the series. The multiplicative version would add an useless complexity to the glucose time series. As before, we employ a period P to make the calculus. Although the seasonal component practically disapears in the real patients, it has a small representation in the in-silico patients, so we decided to included it in this study.

$$\hat{y}_{t+H|t} = Level_t + HSlope_t + Seasonal_{t-P} \quad (14)$$

$$Level_t = \alpha_0 y_t + (1 - \alpha_0)(Level_{t-1} + \phi Slope_{t-1} \quad (15)$$

$$Slope_t = \alpha_1 (Level_t - Level_{t-1}) + (1 - \alpha_1)\phi Slope_{t-1} \quad (16)$$

$$Seasonal_t = \gamma(y_t + Level_{t-1} - Slope_{t-1}) + (1-\gamma)Seasonal_{t-P} \quad (17)$$

3.5 ARIMA

The last classical technique that we have experimented with is the non-seasonal auto regressive integrated moving average (ARIMA) model, Equation (18).

$$\dot{y} = c + \phi_1 \dot{y}_{t-1} + \cdots + \phi_p \dot{y}_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t \quad (18)$$

We can divide the last equation in two parts. The first part, Equation (19) is a regression of x_t based on the observations during period P, that is to say, is an autoregression. The other part of (18) is (20). This equation is a linear combination of the current error term and the q most recent past error terms during period P. This is the Moving Average part and it must not be confused with the weighted moving average that we employed in the exponential smoothing techniques. Integrated means that we have removed drift (if present) by differencing the time series.

$$y = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + e_t \quad (19)$$

$$y = c + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \quad (20)$$

3.6 Genetic Programming

We have applied symbolic regression based on GP using a structure identification framework described by Winkler (2009) and Affenzeller et al. [3]. We have used the following parameter settings for our GP test series: The mutation rate was set to 15^o %, gender specific parents selection (combining random and roulette selection) was applied as well as strict offspring selection [3] (OS, with success ratio as well as comparison factor set to 1.0); Figure 2 shows the GP cycle with OS. The functions set described in [2] (including arithmetic as well as logical functions) was used for building composite function expressions. The maximum model size was set to 50; the population size was set to 100 and the maximum selection pressure was also set to 100. We used the GP implementation in HeuristicLab [14]. For each patient we trained models with minimum time delay 90 minuted and with minimum time delay 120. We executed a GP ensemble modeling approach: For each target (i.e. for each patient and min time delay 90 and 120, resp.) we executed GP 10 times, the best 5 models (on training data) were collected and their estimated values on the test data were averaged.

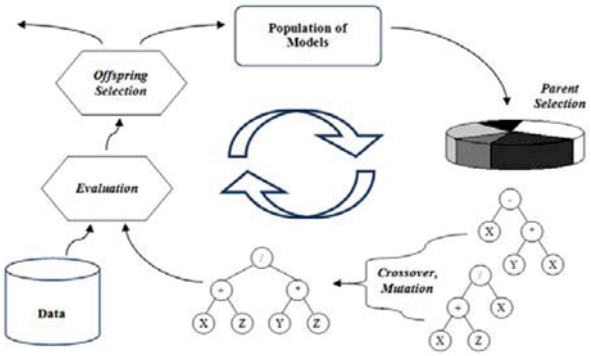


Figure 2: Genetic programming including offspring selection (Affenzeller et al. 2009).

3.7 Wavelet Power Spectrum

The Wavelet Power Spectrum is a visual way that allow us to react on certain events (like oscillations, peaks or discontinuities) that are within the time series. The power spectrum shows the energy distribution over different frequencies of a time series as a whole and the Wavelet Power Spectrum give us a similar vision but for every observation of the time series. The theory behind Wavelets is huge and complex and we suggest the reading of [15] for a deeper understanding. Here, we enumerate the basic recipe for getting a Wavelet Power Spectrum:

- First, we make a convolution of the glucose and the Morlet wavelet function with a certain width (scale) throughout the whole time series.
- We repeat the previous step for different scale values. For some scale values we need to fill the series with a zero padding. The results obtained this way are shown in the plot in a striped zone and the proper values can be seen inside a cone-like region.
- This way, for every minute of the dataset we get the wavelet coefficients at different scales.
- Last, we plot the contour of the square of the absolute values of the wavelet transform.

We use Wavelet Power Spectrum as a tool for input data analysis in the last part of the paper.

4. EXPERIMENTAL SETUP

In this section we describe the characteristics of the eight in-silico patients we deal with, as well as the configuration of each set of experiments. As we have explained, we predict a future value of

the level of glucose in the patient's blood depending on, at least, the glucose level, the carbohydrates ingested and the insulin injected. Hence we need to consider those values in our datasets.

4.1 In-silico Patients

We work with a set of 10 in-silico patients obtained with AIDA simulator (www2aida.org). More precisely, our data series represent measures taken each minute along the day. (see section 3.1 for details). This simulator although a bit old allows to obtain data series for simulated patients (in-silico patients) that can be individualized for different situations usually present in the daily clinical practique. For this purpose, the simulator offers several characterized patients. The glucose values for each patient were obtained by introducing different carbohydrates and insulin values and then running the simulator. The description of each one of the patients can be found on the website, but we replicate them here for the sake of clarity. The patients are the following: **Joy Wilson**. This woman is on three injections of short and/or intermediate acting insulin each day, with a split-evening dose. She wants to start a family, but consistently has had quite high blood glucose levels in the early afternoon.

Howard Kistler. This 45 year old man was diagnosed as having diabetes at the age of 14. He is currently on a regimen of combined short and/or intermediate acting insulin preparations four times per day. He tends to higher blood glucose values overnight but has a low blood glucose in the mid-morning.

Steven Jones. This man is a relatively newly diagnosed insulin-dependent (type 1) diabetic patient. He tends to quite high blood glucose levels in the middle of the day, despite not eating excessively.

Elizabeth Whittaker. She still often goes hypoglycemic in the middle of the day, especially between breakfast and lunch. She is on a slightly unusual regimen taking a short acting insulin preparation three times per day, with an intermediate acting preparation twice a day – at lunchtime and before bed.

Lizzy Laurence. This overweight 58 year old insulin-dependent (type 1) diabetic patient has had major problems losing weight. She is quite sensitive to insulin. In addition, she smokes and is at great risk of suffering a heart attack or stroke.

Mohammed Abdullah. This man often wakes with 'sweats' and feeling profoundly unwell in the middle of the night. However, his blood sugars are quite respectable when he gets up at 7:30 AM. In such a situation he needs to measure his blood glucose when he wakes in the middle of the night, feeling unwell. Clearly injecting so much insulin before he goes to bed isn't a good idea.

David Robins. This 18 year old insulin-dependent patient has just left home for the first time to go to the University. He isn't a very good cook and hasn't been taking good care of himself. He feels pretty awful most mornings and even going to bed early hasn't helped. He tends to quite low blood sugars in the morning, at times being at risk of going 'hypo'.

Wendy. This 50 year old insulin-dependent diabetic woman has quite high blood sugars throughout most of the day, especially after lunch. She is adamant that she can't change her diet - she attends a lot of business lunches and dinners.She also refuses to inject any more frequently than two times per day.

David Jones. This 34 year old insulin-dependent diabetic man (diagnosed as a boy aged 8) has impaired renal function as a result of diabetic nephropathy. He tends to run very high blood glucose levels overnight, which will be contributing to the appearance of his diabetic complications.

Hugh Allibaster. This 35 year old insulin-dependent diabetic man recently switched to using an insulin pen, injecting three 'shots' of short-acting insulin before breakfast, lunch, and dinner, while taking a single dose of long-acting insulin before going to bed. However, he hasn't quite gotten full control of his blood sugars, still tending towards high blood glucose levels overnight.

4.2 Real Patient

After developing all the experiments we have also used some data of a real patient. In the case of the real patient, we collected these data with the help of the patient. Actual glucose values were obtained with a continuous glucose monitoring system (CGMS) during five days. Carbohydrate units ingested are calculated based on the daily meals. Insulin injected, distinguished by insulin type, is also an information that the patient usually write down using either an application [8], a diary or the CGMS . The selected patient is a patient with a great glycemic variation, with 3 years of evolution of the diabetes ilness. [1]

5. EXPERIMENTAL RESULTS

In Figure 3 we can see the power spectrum plot which shows the energy distribution over different frequencies of a glucose time series from a real patient. The maxima of energy are detected at scales higher than 512 and 1024 min. which corresponds to cycles of 8, 16 and 24 hours and are the rhythmic biological cycles of a normal person's life. As we can see it is very difficult to detect any other frequency that can help us to forecast glucose levels. From a different point of view, Figures 5 and 6 show the same idea. On those figures we present a segment of the phase space trajectory for a in-silico and a real patient respectively. The phase space plot has three axis, one for the current observation at instant t and the other two for the observations with lags (t-l) and (t-2l). For instance Lag1 = 15, lag2 = 30 means that each point represents (t, t-15, t-30). This way we embed the trajectory of the time series in a 3D plot and we expect that recurrent behaviour, if exists, will show up. As we can see, the longer the lag, the more confused the phase space trajectory gets, which is a clear sign that explains why all the forecasting algorithms are performing so poorly for the longest prediction horizons in this study. We can also detect that, although we can hardly find some trajectories for the in-silico patient, the situation is even worst for the Real Patient (Figure 6).

Now, we compare GP with classical techniques. On Table 2 we show the mean absolute percentage error (MAPE) for 10 in-silico patient and every technique distinguishing amongst the four prediction horizons. Several conclusions arise from this table (although no complete significance statistical test are performed yet):

- In the 30 minutes horizon, all the classical techniques are better than the Simple Exponential Smoothing (SES). In the 120 minutes horizon, only the HOLT_damped technique is better than SES.
- The ARIMA model (lighter background color) is the best algorithm for every patient in the 30 minute horizon and for the majority of them in the 60 minutes horizon. Nevertheless, for 90 and 120 minutes, the best results are achieved by HOLT_damped and Genetic Programming (GP).

[1]We would like to note that on 27 December 2010, the Clinical Research Ethics Committee of the Hospital of Toledo authorized the development of the computer application for collecting data. It also noted that at the time of recruiting patients for the study of its usefulness must resubmit for approval. to this or another Ethics Committee. On 6 June 2012 he Clinical Research Ethics Committee of the Hospital of Alcalá de Henares, Madrid authorized the use of the data collected, provided that the privacy of the data is ensured and the informed consent of patients is made.

- GP obtains much better results with Elizabet Whittaker and Howard Kitsler than the classical techniques in the 90 and 120 minutes horizon. The sudden oscillations of these two patients are only modelled using GP.
- With David Robins, all techniques fall down. Robins' glucose level falls down during night in a steady way and ascends abruptly afterwards and it seems that it is going to be very difficult to model.

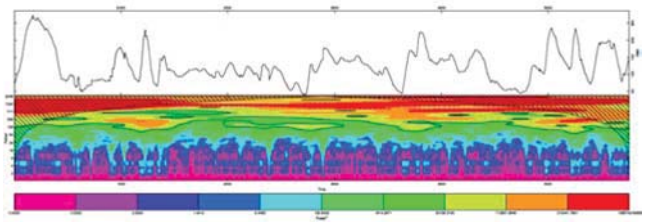


Figure 3: The power spectrum shows the energy distribution over different frequencies of a time series.

6. CONCLUSIONS

In this paper we have compared a set of classical techniques with GP. We can say, despite a deeper statistical analysis needed, that at 120 minutes the best prediction techniques are GP, HOLT damped and SES. These two last techniques really are a way to leave and recognize the difficulty of predicting. Although GP has improved its results for some patients, the high level of error in their predictions to 120 min could cause serious problems if used in an artificial pancreas.

The next step initiated in this paper is to analyze the time series data of glucose levels in terms of the predictability. We have also analyzed data in terms of the space trajectory and the power spectrum for the 10 in-silico patients and also for a real patient. We can see some correlation of the values of glucose at time t with those at time $t-30$ and $t-15$ for the in-silico patient. However we can not say the same for $t-60$, $t-90$ and $t-120$, which clearly indicates the noisy nature of the series. The situation is even worst for the case of real patients.

We can conclude that we need to combine GP techniques and improve also previous approximation with Grammatical Evolution made in previous works. We should seek for alternative grammars prepared for dynamic environments that incorporate mechanisms for generating constant and dynamic grammars.

7. ACKNOWLEDGMENT

This research is supported by the Spanish Minister of Science and Innovation (TIN2014-54806-R) whereas the development of HeuristicLab is supported by the Austrian Funding Agency FFG within the projects Heureka! (818420) and HOPL (843532).

8. REFERENCES

[1] John R Koza. *Genetic programming: on the programming of computers by means of natural selection*, volume 1. MIT press, 1992.

[2] Stephan M Winkler. *Evolutionary system identification: modern concepts and practical applications*. Universit atsverlag Rudolf TraunerTrauner, 2008.

[3] Michael Affenzeller, Stefan Wagner, Stephan Winkler, and Andreas Beham. *Genetic algorithms and genetic programming: modern concepts and practical applications*. Crc Press, 2009.

[4] Robert J Hyndman and George Athanasopoulos. Forecasting: principles and practice. *Online textbook*, 2013.

[5] A.D.A. American-Diabetes-Association. Standards of medical care in diabetes 2010. *Diabetes Care*, 33(S1):11–61, 2010.

[6] C. Cobelli, C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, and B.P. Kovatchev. Diabetes: Models, signals, and control. *Biomedical Engineering, IEEE Reviews in*, 2:54 –96, 2009.

[7] K. van Heusden, E. Dassau, H.C. Zisser, D.E. Seborg, and F.J. Doyle. Control-relevant models for glucose control using a priori patient characteristics. *IEEE Transactions on Biomedical Engineering*, 59(7):1839 –1849, july 2012.

[8] J. Ignacio Hidalgo, Esther Maqueda, José L. Risco-Martín, Alfredo Cuesta-Infante, J. Manuel Colmenar, and Javier Nobel. glucmodel: A monitoring and modeling system for chronic diseases applied to diabetes. *Journal of Biomedical Informatics*, (0):–, 2014.

[9] Daniel A. Finan, Cesar C. Palerm, Francis J. Doyle, Dale E. Seborg, Howard Zisser, Wendy C. Bevier, and L. Jovanovic. Effect of input excitation on the quality of empirical dynamic models for type 1 diabetes. *AIChE Journal*, 55(5):1135–1146, 2009.

[10] E. Dassau, H. Zisser, B. Grosman, W. Bevier, M.W. Percival, L. Jovanovic, and F.J. Doyle. Artificial pancreatic betta-cell protocol for enhanced model identification. *Diabetes*, pages A105–A106, 2009.

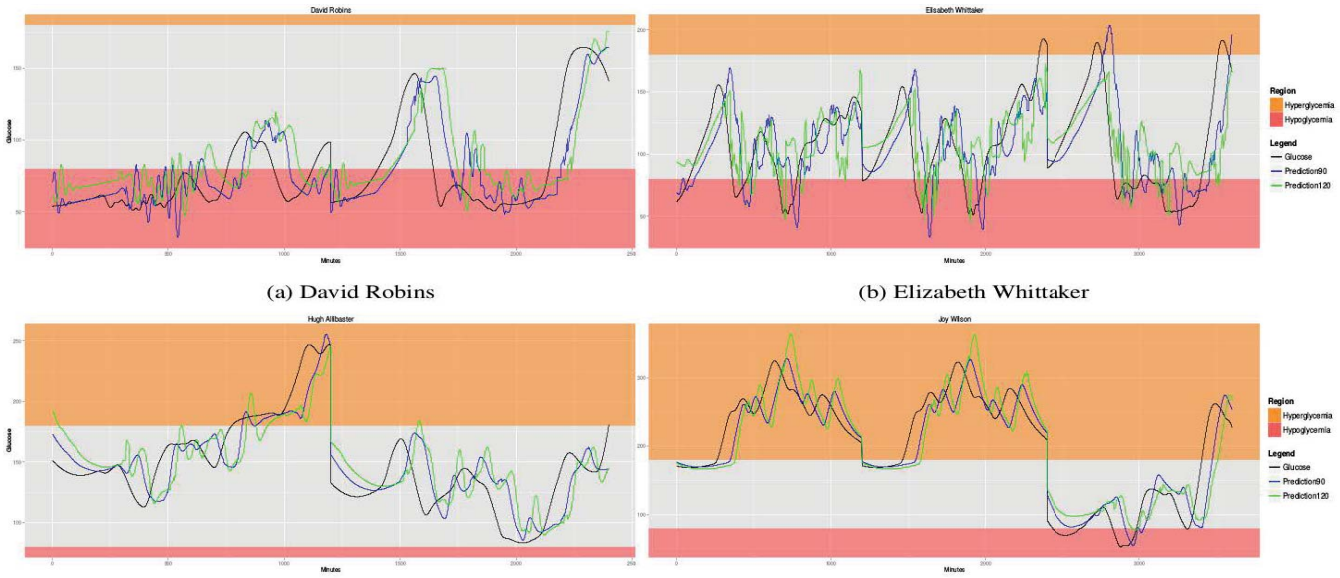
[11] M. Gevers. Identification for control: From the early achievements to the revival of experiment design. In *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05*, page 12, dec. 2005.

[12] A. Gani, A.V. Gribok, Yinghui Lu, W.K. Ward, R.A. Vigersky, and J. Reifman. Universal glucose models for predicting subcutaneous glucose concentration in humans. *Information Technology in Biomedicine, IEEE Transactions on*, 14(1):157 –165, jan. 2010.

[13] José Ignacio Hidalgo, J. Manuel Colmenar, José Luis Risco-Martín, Esther Maqueda, Marta Botella, José Antonio Rubio, Alfredo Cuesta-Infante, Oscar Garnica, and Juan Lanchares. Clarke and parkes error grid analysis of diabetic glucose models obtained with evolutionary computation. In Dirk V. Arnold and Enrique Alba, editors, *Genetic and Evolutionary Computation Conference, GECCO '14, Vancouver, BC, Canada, July 12-16, 2014, Companion Material Proceedings*, pages 1305–1312. ACM, 2014.

[14] S Wagner, G Kronberger, A Beham, M Kommenda, A Scheibenpflug, E Pitzer, S Vonolfen, M Kofler, S Winkler, V Dorfer, et al. Architecture and design of the heuristiclab optimization environment. In *Advanced Methods and Applications in Computational Intelligence*, pages 197–261. Springer, 2014.

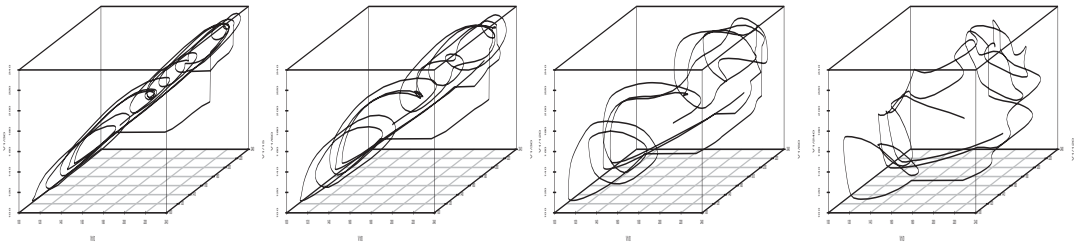
[15] Gerald Kaiser. *A Friendly Guide to Wavelets*. Birkhauser Boston Inc., Cambridge, MA, USA, 1994.



(a) David Robins (b) Elizabeth Whittaker

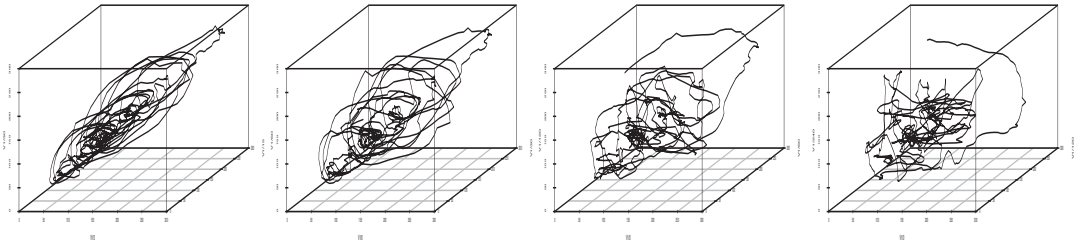
(c) Hugh Allibaster (d) Joy Wilson

Figure 4: Predictions using genetic Programming for 90 and 120 minutes. For the shake of clarity, we present here only 4 patients. Other patients have similar figures to one of those, in terms of predictions, hypo and hyperglycemias, etc..



(a) Lag1 = 15, lag2 = 30 (b) Lag1 = 30, lag2 = 60 (c) Lag1 = 60, lag2 = 120 (d) Lag1 = 120, lag2 = 240

Figure 5: Segment of the phase space trajectory for in-silico patient



(a) Lag1 = 15, lag2 = 30 (b) Lag1 = 30, lag2 = 60 (c) Lag1 = 60, lag2 = 120 (d) Lag1 = 120, lag2 = 240

Figure 6: Segment of the phase space trajectory for real patient

Table 2: Mean Absolute Percentage Error (MAPE)

(a) Prediction horizon 30 min

	SES	HOLT_linear	HOLT_damped	HOLT_exp	HOLT_WINTERS	ARIMA	GP
David Jones	4,68	2,05	2,21	2,05	2,06	1,16	3,71
David Robins	7,68	4,86	4,56	4,89	4,89	4,51	7,00
Hugh Allibaster	4,82	2,01	2,05	2,04	2,03	1,32	8,52
Lizzy Laurence	4,03	1,43	1,71	1,44	1,44	0,83	4,60
Mohammed Abdullah	5,11	2,02	2,27	2,02	2,03	1,21	3,16
Wendy Couger	5,21	1,68	2,07	1,72	1,69	1,16	4,51
Elizabeth Whittaker	11,15	5,52	5,55	5,46	5,55	5,17	5,17
Howard Kistler	8,52	4,02	4,01	4,13	4,04	3,39	2,74
Joy Wilson	5,40	1,91	2,27	1,91	1,92	1,09	2,70
Steven Jones	4,45	1,96	1,98	1,96	1,98	1,19	2,43
Average	6,11	2,75	2,87	2,76	2,76	2,10	4,45

(b) Prediction horizon 60 min

	SES	HOLT_linear	HOLT_damped	HOLT_exp	HOLT_WINTERS	ARIMA	GP
David Jones	8,41	6,97	6,18	6,94	6,99	5,91	7,79
David Robins	13,96	13,82	11,36	13,77	13,87	14,85	14,98
Hugh Allibaster	8,87	6,86	5,82	6,92	6,88	6,11	8,74
Lizzy Laurence	7,47	5,07	4,95	5,20	5,09	4,27	8,98
Mohammed Abdullah	9,23	7,03	6,40	7,23	7,05	6,18	5,81
Wendy Couger	9,81	5,89	5,88	6,10	5,91	4,97	8,41
Elizabeth Whittaker	21,47	16,93	14,85	16,07	16,97	18,68	6,09
Howard Kistler	16,24	12,32	10,67	12,68	12,36	12,42	4,88
Joy Wilson	10,27	6,63	6,47	6,63	6,65	5,40	3,99
Steven Jones	8,30	6,71	5,65	6,71	6,73	5,77	5,30
Average	11,40	8,82	7,82	8,82	8,85	8,46	7,50

(c) Prediction horizon 90 min

	SES	HOLT_linear	HOLT_damped	HOLT_exp	HOLT_WINTERS	ARIMA	GP
David Jones	11,04	13,06	9,68	12,99	13,09	13,61	12,13
David Robins	18,75	24,93	17,98	24,99	24,99	28,95	18,54
Hugh Allibaster	12,20	13,17	9,59	13,35	13,20	13,72	22,68
Lizzy Laurence	10,25	10,00	8,11	10,44	10,02	10,35	19,18
Mohammed Abdullah	12,42	13,47	10,18	14,27	13,50	14,43	11,32
Wendy Couger	13,89	11,77	9,99	12,50	11,79	11,38	14,10
Elizabeth Whittaker	30,05	30,46	25,06	29,68	30,51	36,54	8,41
Howard Kistler	22,75	23,11	17,58	24,44	23,15	24,75	10,10
Joy Wilson	14,38	13,14	10,87	13,14	13,17	12,92	8,35
Steven Jones	11,39	12,57	9,25	12,57	12,59	13,10	8,61
Average	15,71	16,57	12,83	16,84	16,60	17,97	13,34

(d) Prediction horizon 120 min

	SES	HOLT_linear	HOLT_damped	HOLT_exp	HOLT_WINTERS	ARIMA	GP
David Jones	13,07	18,94	11,97	19,00	18,96	22,88	13,14
David Robins	22,47	37,22	23,37	38,34	37,29	45,14	23,77
Hugh Allibaster	14,92	20,03	12,85	20,60	20,06	22,92	20,23
Lizzy Laurence	12,61	15,60	10,82	16,65	15,63	18,24	22,63
Mohammed Abdullah	14,95	20,49	13,13	22,38	20,52	23,55	14,52
Wendy Couger	17,42	19,16	14,25	21,06	19,19	20,31	17,82
Elizabeth Whittaker	36,69	45,31	33,62	45,90	45,36	55,95	12,69
Howard Kistler	28,09	35,03	23,82	39,46	35,07	39,87	12,68
Joy Wilson	17,61	20,41	14,88	20,41	20,44	22,57	10,19
Steven Jones	13,63	18,76	12,23	18,76	18,79	21,54	9,24
Average	19,15	25,10	17,09	26,26	25,13	29,30	15,69