Blood Glucose Data Processing for Automated Diagnosis

EUGEN IANCU[*], IONELA IANCU[**], MARIA MOȚA[***]
[*]Department of Automation
University of Craiova
Decebal Street, no. 107, 200440 Craiova
Eugen.Iancu@automation.ucv.ro, http://www.ace.ucv.ro
[**] Department of Physiology, Medical Informatics and Biostatistics
University of Medicine and Pharmacy of Craiova
iancu@automation.ucv.ro, http://www.umfcv.ro
Petru Rares Street, no. 4, 200349 Craiova
[** *]Department of Diabetes and Nutrition Diseases
University of Medicine and Pharmacy of Craiova
Petru Rares Street, no. 4, 200349 Craiova
mota@rdslink.ro, http://www.umfcv.ro
ROMANIA

Abstract: - The actual protocols used in diagnosis and management of the diabetes mellitus include the classical clinical trials and the physicians' experience, but they do not account by the dynamics of the blood glucose and insulin. So, it is natural to have many diabetes patients with poor control of blood glucose values. The introduction in the medical practice of the blood glucose continuous monitoring systems has made possible the automated analyse of blood glucose dynamics. Along this paper the authors present algorithms for automatic diagnosis in the diabetic patients monitoring with applications, especially in the intensive care units and telemedicine. We have focused on the statistical analysis methods in order to detect the reliable characteristics, useful in the identification of standard aspects or stable patterns for each type and stage of the complex and long-term evolution of the disease that is diabetes mellitus. Examples of the frequency range of blood glucose dynamics of normal subjects and subjects with diabetes are presented with the help of Wigner-Ville distribution. The spectral analysis reveals the frequency band edge and offers the basic information to correct determination of Nyquist sample period. These findings may have significant clinical implication in diagnosis of the diabetes mellitus, in blood glucose monitoring and the management of the diabetes therapy.

Key-Words: - Statistical analysis, Diabetes mellitus, Continuous glucose monitoring, Probability distribution function, Periodogram, Correlation

1 Introduction

The World Health Organization has estimated a rapid increase in the number of people affected by mellitus diabetes from 173 million in 1973 to 300 million in 2025. Studies regarding the increase of morbidity through diabetes, especially at children or youngsters and grave complications, permanently affective, of this disease have allowed the anticipation of high costs at the society level.

As an answer to all of these findings important programs and projects for the study of the diabetes and particularly for curing it have been financed in the last 5 years. From these, we mention:

- ADICO (Advanced Insulin Infusion Using a Control Loop)-European Union co financed project for the study of insulin treatment belonging to an automated pancreas system.
- CLINICIP – European project, consortium–type, co financed by the European Union through IST Programme FP6 for the study of the automated (closed loop) insulin treatment for patients in a critical state [1].
- DIRECNET – Diabetes research in children network.
- JDRF – Juvenile Diabetes Research Foundation International in collaboration to the Department of Defence of the USA and the American Space Agency (NASA).

Research in this area will have in the nearby future a major impact in the medical environment. They belong to the modern tendency to automate and to introduce the informatics in the human medicine [2].

The realisation of the artificial pancreas is presently a fundamental high priority research area with a strong interdisciplinary character (physiology, diabetology, automated systems engineering). The objective consisted in the building of a mathematical model of the blood glucose control system as correctly as possible and the development on its base of evolved control algorithms in an automated state, adapted to the real-life situations encountered in medical practice [3].

2 Problem Formulation

Our team's purpose was to develop new, competitive algorithms for automatic diagnosis in the diabetic patients monitoring with applications, especially in the intensive care units and telemedicine.

Automatic diagnosis in such a medical field requires the introducing of a proper method or a group of mathematical methods capable to achieve moment to moment the following objectives:

- to identify the type of diabetes
- to detect the pathological component in the early stages of the disease
- to quantify the risk level of the metabolic disorders
- to reveal the trend in the pathological state evolution
- to estimate the response to the treatments

In the beginning, we have focused on the statistical analysis methods in order to detect the reliable characteristics, useful in the identification of standard aspects or stable patterns for each type and stage of the complex and long-term evolution of the disease that is diabetes mellitus.

3 Experimental Lot

For this study we have selected 18 adult subjects (10 female and 8 male), patients with insulin dependent mellitus diabetes and 3 healthy humans. 16 patients underwent treatment with rapid and semi-lent types of insulin, at different times of the day, according to the classic method of treatment and clinically supervised. Patients maintain a satisfactory or poorly control of the blood glucose concentration for a long period of time. Two patients have received a proper dosage of insulin by a new device called "insulin pump". This offers a continuous basal rate of insulin and facilitates the administration of bolus insulin related to meals, exercise or other particular states. These patients maintain a very good control over the blood glucose concentration for a long period of time.

The blood glucose was recorded to each patient at five minute intervals, continuously for three days, using the *Real-Time Guardian Continuous Glucose Monitoring System* (CGMS) [4] in unrestrained conditions. Each patient had a normal life, with usual meals and activities at work and at home. The continuous blood glucose records represent for this study time-series of the blood glucose concentration. The following figures present the blood glucose representation for 24 hours. For exemplification we choose the following individual cases:

- Patients (P1, P2, P3) with insulin dependent diabetes (type I) under intermittent treatment with insulin injections. The CGMS displays high variability of the glucose values as an expression of an insufficient control of diabetes (Fig. 1, Fig.2 and Fig. 3).
- Patient (P4) with insulin dependent diabetes under insulin treatment administrated by insulin pump. The CGMS displays a less variability of glucose values, expression for an improved control of diabetes (Fig. 4).
- Healthy subject (P5) with normal food administration and activity. The CGMS displays a low variability of the glucose values, expression of an efficiently blood glucose control (Fig. 5).

4 Mathematical Methods

The following statistical methods have been utilised for recordings from the experimental lot: normal distribution, histogram, probability density function, periodogram and correlation function [5], [6] and [7]. The normal distribution, also called the Gaussian distribution, is an important family of continuous probability distributions, applicable in many fields. Each member of the family may be defined by two parameters: the mean μ and the standard deviation σ.

4.1 Statistical parameters

For the time series presented in Fig. 1, 2, 3, 4 and 5 have been calculated the mean and the standard deviation (Table 1). Also, using statistical prediction, it was possible to calculate the confidence intervals for 95% and 99% of blood glucose values.

Table 1.

Statistical parameters		P1	P4	P5
Mean of blood glucose values - μ (mg/dl)		199.31	101.28	86.56
Standard deviation - σ		81.02	17.29	6.14
Confidence interval for 95%	Minimum	189.91	99.27	85.85
	Maximum	208.71	103.28	87.27
Confidence interval for 99%	Minimum	186.93	98.63	85.62
	Maximum	211.69	103.92	87.50

The statistical findings reveal a very large distribution for diabetes patient P1, comparatively with P4 patient (with insulin pump) and P5 patient (normal subject).

4.2 Histogram

In statistics, a histogram is a graphical display of a table that shows what proportion of cases fall into each of several or many specified categories. In a mathematical sense, a histogram is a mapping that counts the number of observations that fall into various disjoint categories (known as *bins*). There is no "best" number of bins, and different bin sizes can reveal different features of the data. In the present paper the authors propose the usage of histograms in order characterise the time-series coming from the continuous glucose monitoring system.

Also, due the fact that we are dealing with measurements that can be corrupted by errors, we have calculated a Gaussian distribution for these data. The Gaussian distribution, also called the normal distribution, is an important family of continuous probability distributions, applicable in many fields. The importance of the normal distribution as a model of quantitative phenomena in the natural science is due to the central limit theorem. Many physiological measurements can be approximated well by the normal distribution. In addition, the normal distribution maximizes information entropy among all distributions with known mean and variance, which makes it the natural choice of underlying distribution for data summarized in terms of sample mean and variance.

- For the diabetes patients under intermittent treatment with insulin injections (P1, P2 and P3), the CGMS displays high variability of the glucose values as an expression of an insufficient control of diabetes (Fig. 1, Fig. 2 and Fig. 3). This aspect is revealed in a statistical manner by the corresponding histograms (Fig. 6, Fig. 7 and Fig. 8).
- For the patient with insulin dependent diabetes under insulin treatment administrated by insulin pump (P4), the CGMS displays a less variability of glucose values, expression for an improved control of diabetes (Fig. 4). This aspect is revealed in a statistical manner by the corresponding histogram (Fig. 9).
- For the healthy human (P5) with normal food administration and activity, the CGMS displays a low variability of the glucose values, expression of an efficiently blood glucose control (Fig. 5). This aspect is revealed in a statistical manner by the corresponding histogram (Fig. 10).

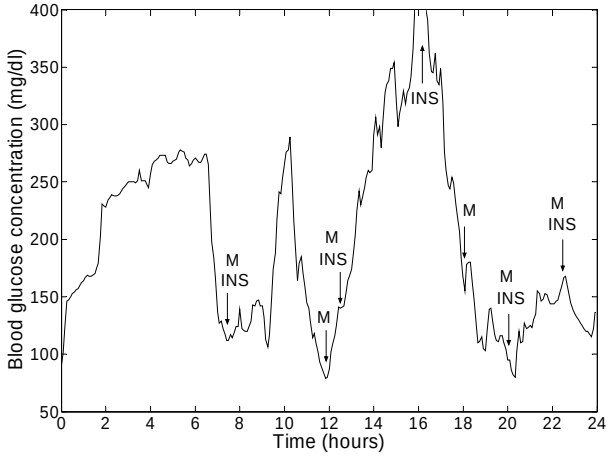

Fig. 1. Time evolution of the glucose concentration for the P1 patient.
INS – insulin treatment, M – meal.

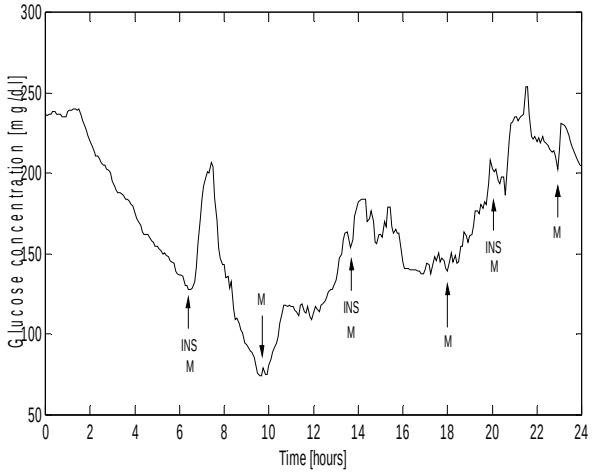

Fig. 2. Time evolution of the glucose concentration for the P2 patient.
INS – insulin treatment, M – meal.

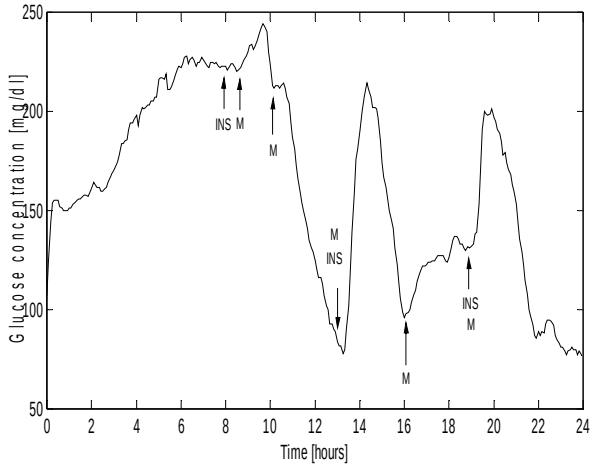

Fig. 3. Time evolution of the glucose concentration for the P3 patient.
INS – insulin treatment, M – meal.

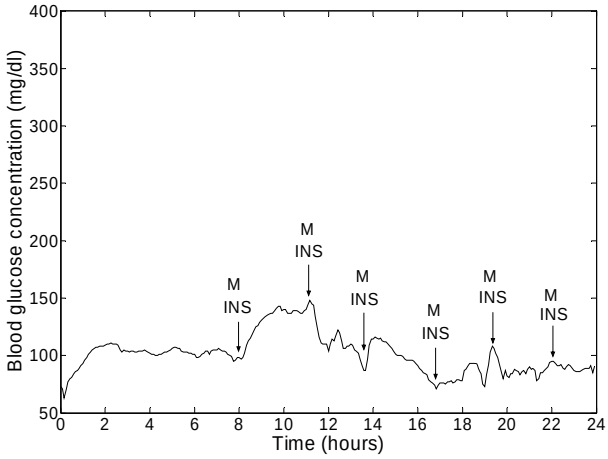

Fig. 4. Time evolution of the glucose concentration for the P4 patient.
INS – insulin treatment, M – meal.

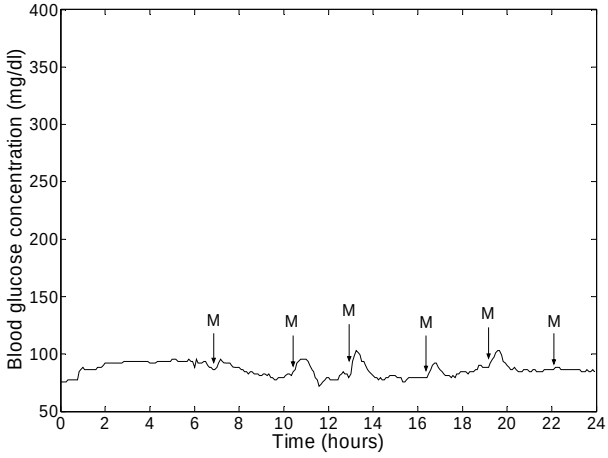

Fig. 5. Time evolution of the glucose concentration for the P5 patient.
M – meal.

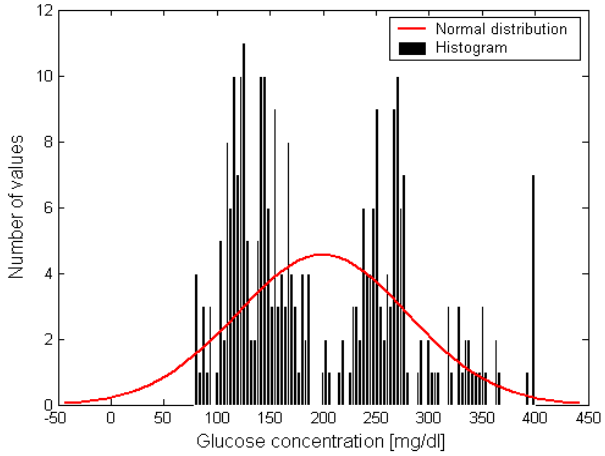

Fig. 6. Histogram of the glucose concentration for the P1 patient.

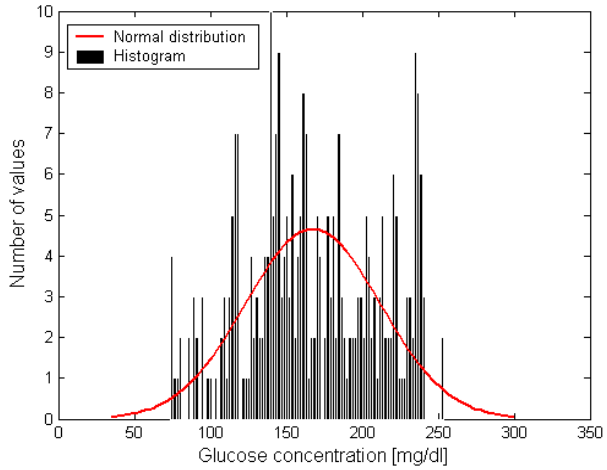

Fig. 7. Histogram of the glucose concentration for the P2 patient.

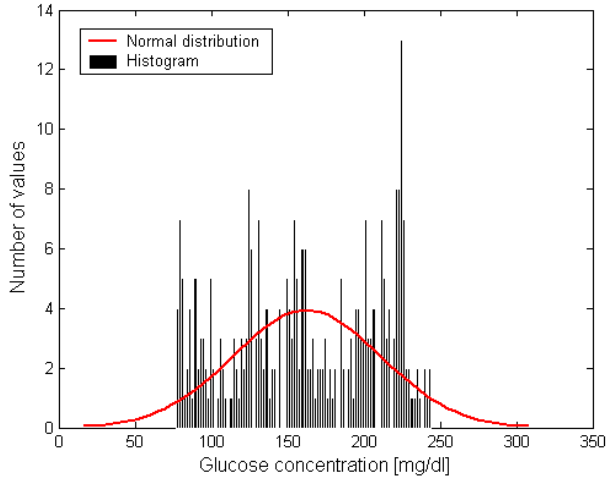

Fig. 8. Histogram of the glucose concentration for the P3 patient.

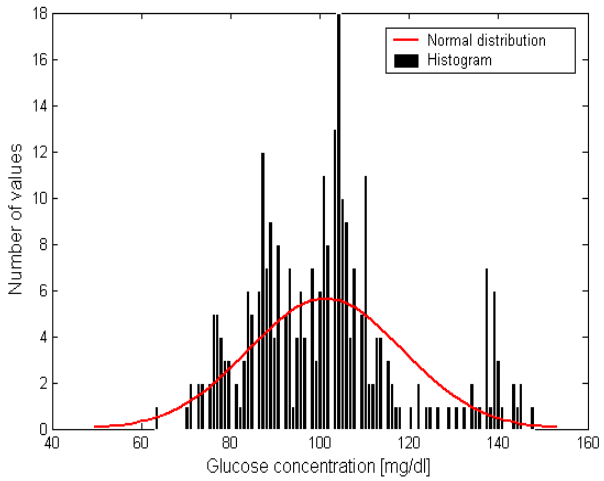

Fig. 9. Histogram of the glucose concentration for the P4 patient (insulin pump)

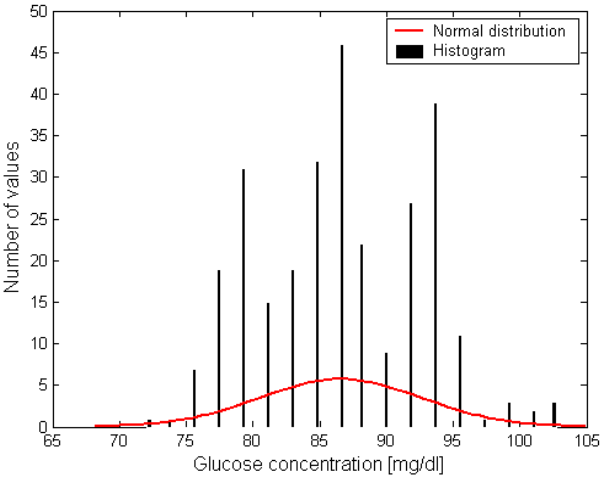

Fig. 10. Histogram of the glucose concentration for the P5 subject.

4.3 Probability distribution

The probability distribution functions (PDF) for the time series of the blood glucose are represented for selected patients in Fig. 11, 12 and 13. These graphics display clearly the differences between P1, P4 and P5.

- For P1 patient the curve show a very large distribution with reduced amplitude for the PDF.
- For P4 patient the curve show a mean distribution as a consequence of the protocol of the insulin administration by the insulin pump.
- For P5 patient the PDF is much more concentrated and reveal the high quality of the physiological blood glucose control system.

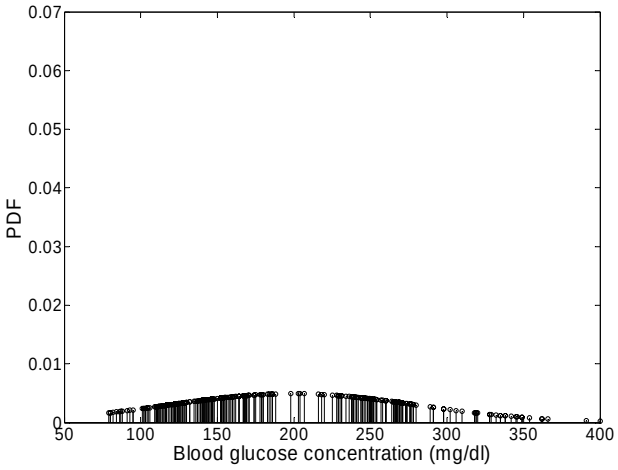

Fig.11. Probability distribution function (PDF) for the patient P1.

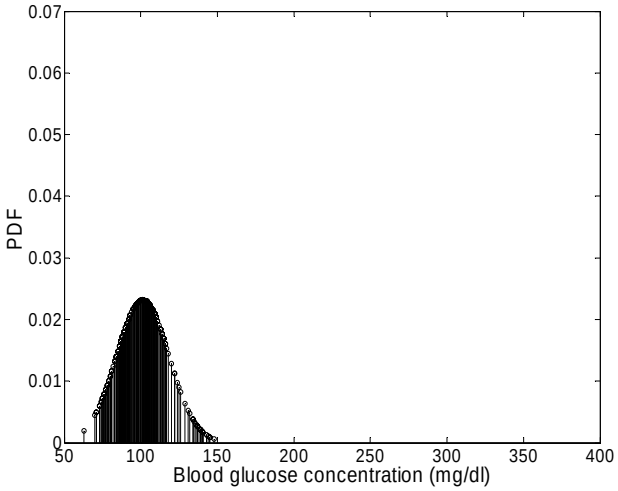

Fig.12. Probability distribution function (PDF) for the patient P4 (insulin pump).

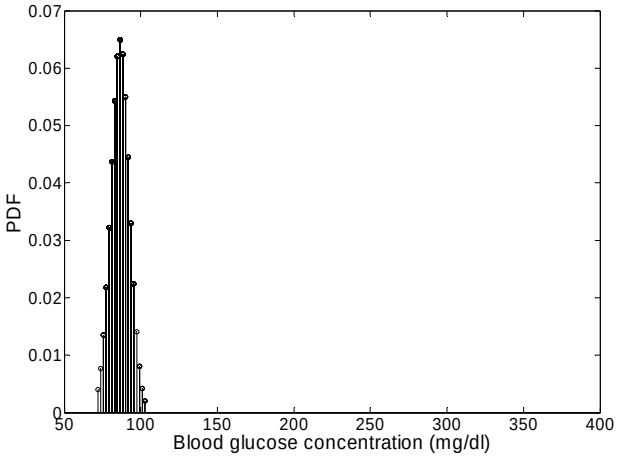

Fig.13. Probability distribution function (PDF) for a normal subject P5.

4.4 Periodogram

The periodogram method [5], [6], [8], used in signal processing, was applied in our study to the probability distribution function to display the estimated spectral power density.

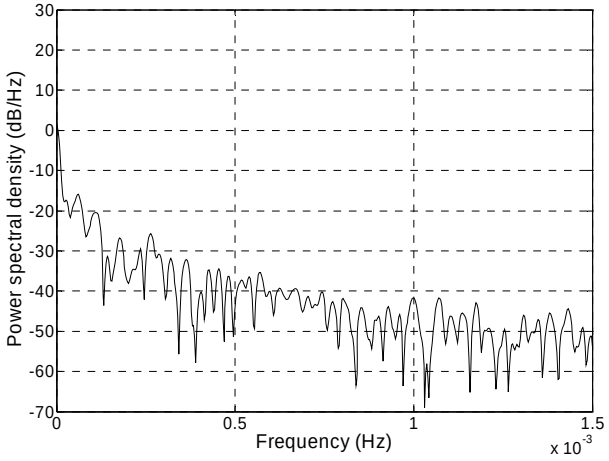

Fig.14. Periodogram estimated for probability distribution function for the P1 patient.

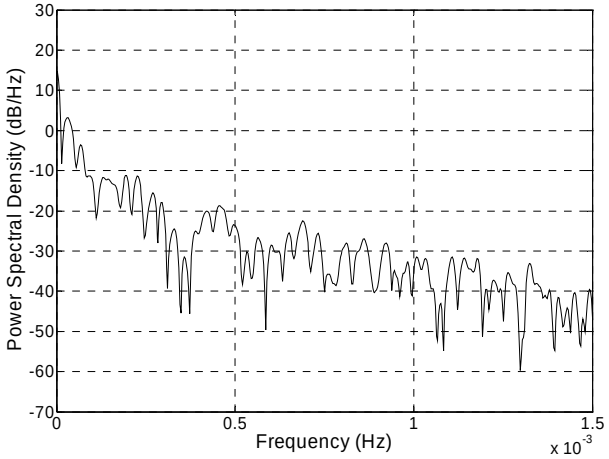

Fig.15. Periodogram estimated for probability distribution function for the P4 patient.

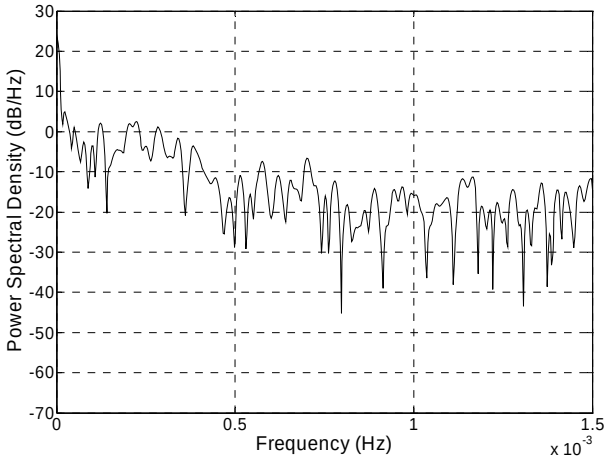

Fig. 16. Periodogram estimated for probability distribution function for the P5 patient.

For P1 patient (Fig. 14) the magnitude start around 0 (dB/Hz), for P4 patient (Fig. 15) the magnitude start around 14 (dB/Hz) and for P5 patient (Fig. 16) the magnitude start around 24 (dB/Hz). All graphics reveal the same frequency range, but the greater attenuation at P1 and a less attenuation to the healthy subject P5.

4.5 Correlation function

In Fig. 17, it is represented the correlation functions calculated for the time series of blood glucose concentration and for lags which vary from 0 to 24 hours. The graphics express the influence of the "history" over the present. We can notice the long periods during which these influences exist.

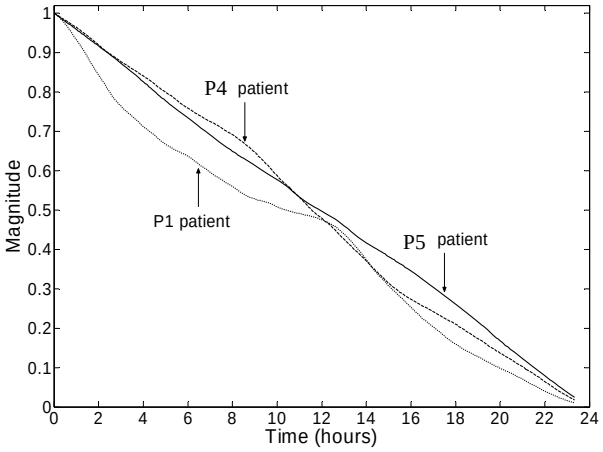

Fig. 17. Correlation functions for the analysed patients during 24 hours.

These statistical findings permit the clear discrimination between the patients type P1, P4 and P5 and their classification in accordance to the risk levels. Also, these methods can be used to determinate the thresholds for the risk levels and to calculate a match between the parameters of the mathematical model and the parameters from each patients.

4.6 Time-frequency Analysis

Many times, we are interested in details of the temporal process represented through signal $x(t)$ in certain time intervals as well as details such as the spectral density in certain frequency bands. In these cases, the analysis needs to be done very specifically, in time intervals or in the frequency bands that we are interested in. Such situations can occur, for example, in the analysis of the vocal signal, in recordings of electrocardiogram signals, heart sounds, seismic signals and so on.

The need for a combined time-frequency representation stemmed from the inadequacy of either time domain or frequency domain analysis to fully describe the nature of non-stationary signals.

A time frequency distribution of a signal provides information about how the spectral content of the signal evolves with time, thus providing an ideal tool to dissect, analyse and interpret non-stationary signals. This is performed by mapping a one dimensional signal in the time domain, into a two dimensional time-frequency representation of the signal. A variety of methods for obtaining the energy density of a function, simultaneously in the time and frequency have been devised, most notably the short time Fourier transform, the wavelet transform and the Wigner-Ville distribution [6].

Obviously, for the extraction of a window from a $x(t)$ signal we can use the "windowing" process with a function $w(t)$. The "windowed" signal:

$$x_w(t) = x(t)w(t) \tag{1}$$

is defined on a desired time interval [9], [10]. The Wigner-Ville Distribution (*WVD*) has the mathematical expression:

$$WVD(t, f) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} x_w(t+\frac{\tau}{2})x_w{}^*(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau \tag{2}$$

For a time-series x(*n*), the expression of the discrete-time Wigner-Ville Distribution, *WD*(*n*,*f*) is:

$$WD(n, f) = 2\sum_{k=-\infty}^{\infty} h_N^2(k)x(n+k)\cdot x^*(n-k)\cdot e^{-j4\pi fk} \tag{3}$$

where $h_N(k)$ is a data-window, which performs a frequency smoothing. While Fourier spectra are periodic with period equal to the sampling rate, *WD*(*n*,*f*) is periodic in frequency with period equal to half the sampling rate. This may cause aliasing, which can be removed either by oversampling, or by using the corresponding analytic signal. The distribution is negatively affected by important cross-terms, which limit its practical use. Cross-terms may be adequately reduced smoothing the distribution over time. The resulting smoothed Wigner-Ville, *SWD*(*n*,*f*) is:

$$SWD(n, f) = \sum_{m=-\infty}^{\infty} w(m)\cdot \cdot \sum_{k=-\infty}^{\infty} h_N^2(k)x(n+k+m)\cdot x^*(n-k+m)e^{-j4\pi fk} \tag{4}$$

The continuous blood glucose records represent for this study time-series of the blood glucose concentration. The following figures present the Wigner-Ville distribution of the glucose concentration, during 24 hours, for three patients with insulin dependent diabetes (P1, P2 and P3 with insulin injections), one diabetes patient under treatment administrated by insulin pump (P4) and one healthy subject (P5).

The time is indicated, indirectly through the number of the samples, between 12:00 am and 12:00 pm.

The colours show the amplitudes of the spectral components according to the following pattern: the amplitude decreases from red to blue.

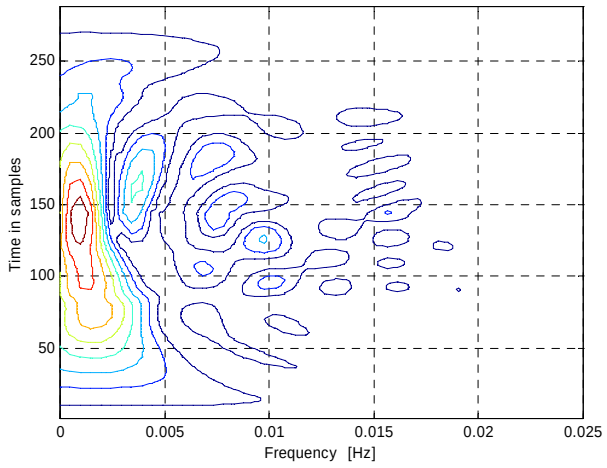

Fig. 18. Wigner-Ville distribution of the glucose concentration for the P1 patient. The maximum frequency: 0.018 Hz.

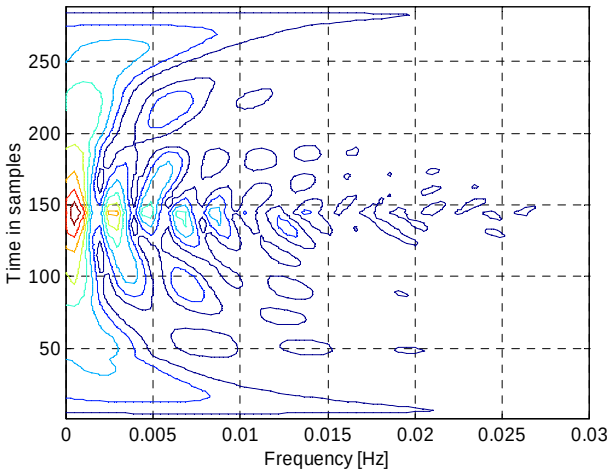

Fig. 19. Wigner-Ville distribution of the glucose concentration for the P2 patient. The maximum frequency: 0.026 Hz.

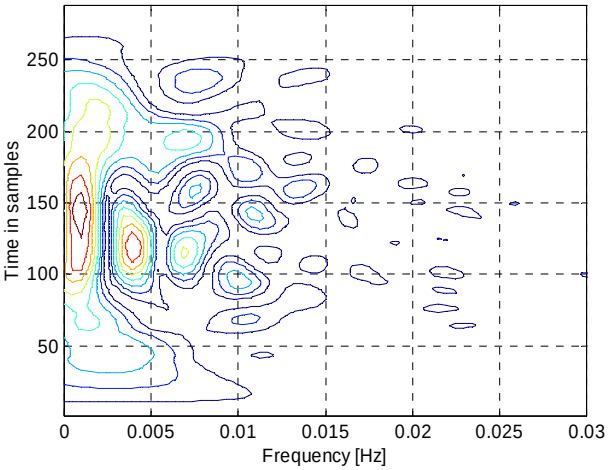

Fig. 20. Wigner-Ville distribution of the glucose concentration for the P3 patient. The maximum frequency: 0.026 Hz.

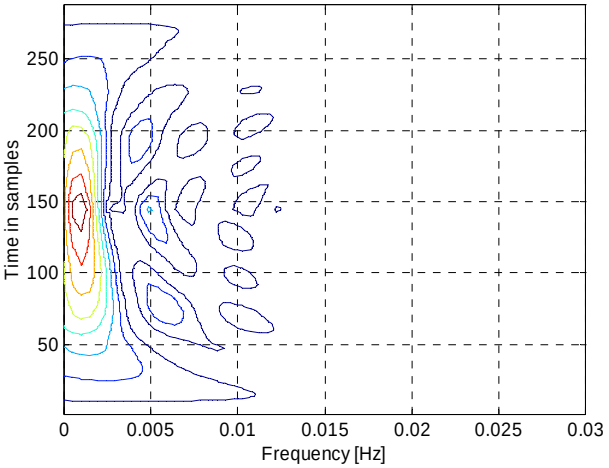

Fig. 21. Wigner-Ville distribution of the glucose concentration for the P4 patient (insulin pump). The maximum frequency: 0.0125 Hz.

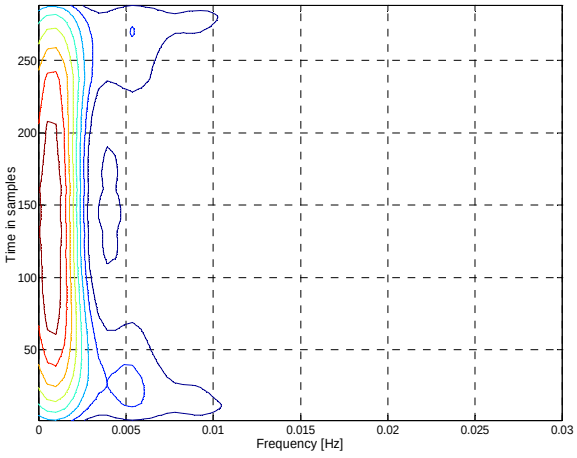

Fig. 22. Wigner-Ville distribution of the glucose concentration for the P5 (healthy subject). The maximum frequency: 0.01 Hz.

The patients display extremely diverse blood glucose dynamics and their characteristics are time varying in our methods. The Wigner-Ville distribution (WVD) reveal the next frequency bands:

- Low frequency band (0 - 0.003 Hz) is common for all the individual cases. These oscillations are continuous over 24 hours with a high amplitude around the middle of the day and reduced amplitude in the night.
- High frequency oscillations (0.003 – 0.01 Hz) with a reduced amplitude seem to be specifically in the interval 8 pm to 4 am (during the night).
- High frequency oscillations (0.003 – 0.025 Hz) were displayed only at diabetes patients during 6 am at 6 pm (during the day) in an irregular manner. These oscillations look like islets in Wigner-Ville distributions and tend to group around the middle of the day.

In this moment the experimental studies have not investigated yet the correlation between these frequency bands and the physiological parameters of the blood glucose control system.

The Wigner-Ville Distribution (WVD) reveal the frequency band edge and offer the possibility to estimate the correct Nyquist sample period (Table 2).

The histograms and the WVD diagrams show a great variability of the glucose values at diabetics patients with classical insulin treatment. Consequently, the frequency spectrum indicates the maximum frequency towards 0.026 Hz.

At diabetes patients with treatment by insulin pump, the maximum frequency is less than 0.0125 Hz. The ideal blood glucose control, corresponding to healthy human, generates a maximum frequency of 0.01 Hz.

By the Nyquist criteria, the sample period for the glucose time-series study, in this case, must be less or equal to 50 seconds. In the medical practice, the diabetes patients are frequently in the state of blood glucose control failures. Similarly, by the Nyquist criteria, the sample period for the glucose time-series study, in this case, must be less or equal to 20 seconds. These values are considerably shorter than the sample rate which are utilised in medical practice.

David Gough and collaborators [8] have found by the power spectrum estimated method a Nyquist sample period value of about 9 minutes for diabetic subjects and 7 minutes for non diabetic, corresponding to a frequency continuous band edge around 10^{-3} Hz. Paolo Magni and collaborators [11]

using Bayesian estimation have found a necessary sampling value of 10 minutes for blood glucose signal.

Table 2

	P1	P4	P5
Maximum frequency [Hz]	0.026	0.0125	0.01
Sample period proposed [Seconds]	20	40	50
Sample period recommended by other authors [Seconds]	420	-	600

We consider that the use of these new sample periods assure the accuracy of the data acquisition and reveal the all significant components. These findings may have significant clinical implications in diagnosis of the diabetes mellitus, in blood glucose monitoring and the management of the diabetes therapy.

5 Conclusion

The diabetes mellitus is a disease with serious social implications through the large number of people affected, complications and high costs that it involves. The World Health Organization has estimated a rapid increase in the number of people affected by mellitus diabetes from 173 million in 1973 to 300 million in 2025.

The complex and highly non-stationary nature [12], [13], [14], [15] of the blood glucose time series, especially in diabetic patients and the permanent influence of the external perturbations (meal, sleep, exercise, other treatments etc.) require a complex series of mathematical study methods.

The patients display extremely diverse blood glucose dynamics and their characteristics are time varying in our methods. This aspect corresponds to large variability of blood glucose observed frequently in diabetes patient's records. We suggest that this situation is the consequence of poor efficiency of the physiological blood glucose control system and/or insufficient management of the diabetes.

However, statistical methods have the advantage to be accurate and robust, simple enough to be implemented with low costs. We have proven that these methods offer an indispensable foundation for our investigation and represent a reliable support for diagnosis.

Our findings present the time-frequency characterisation of blood glucose dynamics for insulin-dependent diabetes patients and healthy human subjects and suggest the patterns of diabetes patients under classical insulin treatment, insulin pump treatment and non-diabetes subjects.

These method offer essential information about the diabetes stage at unconscious patients (coma, critical stages, etc.), uncooperant patients (psychical disease, early children, old patients), negligent patients that cannot hold a trust diary for meal, activities and treatment or cannot respect the alimentation and treatment imposed.

These methods, also, makes possible the automated diagnosis and monitoring of patients in intensive care units and in telemedicine.

Acknowledgments

This paper is part of the projects:

- *The development of automated systems for blood-glucose control at insulin-dependent patients. The implementation of evolved algorithms for computer assisted monitoring and diagnosis*, at the University of Medicine and Pharmacy of Craiova and was supported by National University Research Council, Romania.

- *Algorithms for fault detection and isolation in dynamic systems. Development of analytical methods for diagnose assisted by computer. Applications for the study of the physiological systems*, at the University of Craiova and was supported by National University Research Council, Romania.

References

[1] * * * CLINICIP-*Publishable Final Activity Report*, www.clinicip.org/fileadmin/clinicip, January 2008

[2] Athena Makroglou, Jiaxu Li, Yang Kuang, Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview, *Applied Numerical Mathematics* vol. 56, 2006, pp. 559–573, www.elsevier.com/locate.

[3] I. Iancu, E. Iancu, *Modeling and Simulation in Physiology*, Ed. Universitaria Craiova, 2003, ISBN 973-8043-406-5, pp.146-158.

[4] * * * MiniMed^{R} Manual – *Continous Glucose Monitoring System*, www.minimed.com, U.S.A.

[5] * * * (1999), *Statistical Signal Processing*, MathWorks Tutorial.

[6] A. Mertins, *Signal analysis. In Wavelets, Filter banks, Time-frequency transforms and applications*, John Wiley & Sons, 1999.

[7] Manfred Doepp, Gabriele Edelmann, Frequency Distribution Analysis of Biological Data: Clinical Evaluation, *WSEAS Transactions on biology and biomedicine*, Issue 2, volume 1, April, 2004, pp. 215-222.

[8] David A. Gough, Kenneth Kreutz-Delgado and Troy M. Bremer, Frequency characterization of blood glucose dynamics, Journal-Annals of Biomedical Engineering, Springer Netherlands, ISSN - 0090-6964 (Print) 1573-9686 (Online), Volume 31, Number 1 / January, 2003.

[9] Branko Markoski, Djordje Babic, Polynomial-based Filters in Bandpass Interpolation and Sampling Rate Conversión, *WSEAS Transactions on signal processing*, Issue 2, volume 3, February 2007, pp. 269-280.

[10] Milan Stork, Daniel Mayer, Josef Hrusak, High Resolution Analog-Digital Converter and Digital Polynomial Filter Applications, *WSEAS Transactions on signal processing*, Issue 3, volume 2, March 2006, pp. 351-356

[11] Paolo Magni, G. Sparacino, Riccardo Bellazzi, Claudio Cobelli, Reduced sampling schedule for the glucose minimal model: importance of Bayesian estimation, *Am J Physiol Endocrinol Metab* 290, 2006, E177-E184.

[12] Boris P Kovatchev, William L Clarke, Marc Breton, Kenneth Brayman, Anthony McCall, Quantifying Temporal Glucose Variability in Diabetes via Continuous Glucose Monitoring: Mathematical Methods and Clinical Application, *Diabetes Technology & Therapeutics* , Dec. 7 (6) 2005, pp.849-862.

[13] Reinhard Zick, Bettina Petersen, Mario Richter, Cornelia Haug, Comparison of Continuous Blood Glucose Measurement with Conventional Documentation of Hypoglycemia in Patients with Type 2 Diabetes on Multiple Daily Insulin Injection Therapy, *Diabetes Technology & Therapeutics*, Volume 9, Number 6, 2007.

[14] Eve Van Cauter, Kenneth S. Polonsky, André J. Scheen, Roles of circadian rhythmicity and sleep in human glucose regulation, *Endocrine Reviews* 18 (5): 716-738, Copyright © 1997 by The Endocrine Society.

[15] Rebecca J. Troisi, Catherine C. Cowie, Maureen I. Harris, Diurnal variation in fasting plasma glucose, *JAMA*, 284, 2000, pp. 3157-3159.