Anomaly Detection in the Artificial Pancreas

Avila L., Martínez E.
INGAR (CONICET-UTN)

Abstract
The integration of subcutaneous sensing and insulin delivery technologies with novel control strategies has brought closer the development of the Artificial Pancreas. Nevertheless, thought recent developments are aimed at preventing chronic complications and less patient discomfort, few works have addressed the critical issue of performance monitoring of the artificial pancreas as well as detection of abnormal functioning in any of its components. This work presents an anomaly detection monitoring tool using the widely known Clarke Error-Grid to identify functional degradation in the artificial pancreas components and guarantee safety-critical control of blood glucose levels. The effect of imperfect calibration of glucose sensors, time lag between blood glucose concentration and interstitial glucose readings, and excessive variability in glucose levels are evaluated against an expected behavior of the glucose regulation loop achieved through an optimal control policy. Results obtained evidence the feasibility of this novel use of the Clarke error grid as a comprehensive monitoring tool for the artificial pancreas.

Keywords
Artificial pancreas, Glucose Monitoring, Clarke EGA, Performance Degradation, Diabetes Management.

Introduction

Diabetes therapies involving multiple daily insulin injections and monitoring of blood glucose (BG) levels, has had some success, but remains an annoyance to the subject and often results in poor glycemic control. Insulin requirements are constantly changing due to many factors, such as the level of stress, diet, health conditions and physical activity. There is consensus that any reduction in the frequency of documented hypoglycemia or in BG variability would represent a clinically important improvement over non-automated therapies. The ideal treatment for controlling BG levels would be the use of an Artificial Pancreas (AP). The development of such technology for the treatment of insulin-dependent diabetes implies a big challenge for the scientific community, mainly because its safety-critical condition requires continuous performance monitoring. A fully automated AP, see Figure 1, would consist of a glucose sensor to monitor the BG concentration on a continuous basis with sufficient reliability and accuracy, a controller to calculate the necessary insulin infusion rates by an appropriate feedback algorithm or control policy and a infusion pump to properly supply the required amount of insulin into the blood. Wireless communication facilitates data transfer between all components in the control loop.

To achieve an optimally-controlled system, a clinically implantable AP requires well functioning and coordination of their components all the time [1]. Thought the available technology for glucose sensing, wireless systems to link pump and sensor, and miniaturization of insulin pump designs put forward a clinically applicable AP, many issues still prevent safe and optimal operation of closed-loop therapies. For example, insulin pumps are still prone to mechanical failures, being the obstruction of the infusion catheter the most common event. Likewise, implantable glucose sensors are limited by substances that can cause unstable signal output or interferences. Moreover, since most glucose sensors measure interstitial fluid (ISF) glucose rather than its concentration in plasma, a time lag between both concentrations give rise to highly correlated errors in the resulting glucose readings. On the other hand, as the control algorithm may be executed in a smartphone

device, with a wireless connection to the glucose sensor and the insulin pump, it can be affected by delays or interference in signal transmission. Besides this, it is clear that the control algorithm must operate correctly without exception, because errors can lead to severe hypo- or hyperglycemia events. Unfortunately, not only the external mechanism may give rise to faulty conditions, since in patients who experience highly variable insulin sensitivity (e.g., after physical activity), the basal rate of insulin delivery may be suddenly too high, putting the patient at risk of hypoglycemia.
A safety-critical design of the AP should guarantee proper functioning despite glycemic variability that may affect the efficacy of blood glucose regulation. These drawbacks make clear the critical role of continuous monitoring of the closed loop and early detection of any performance degradation.

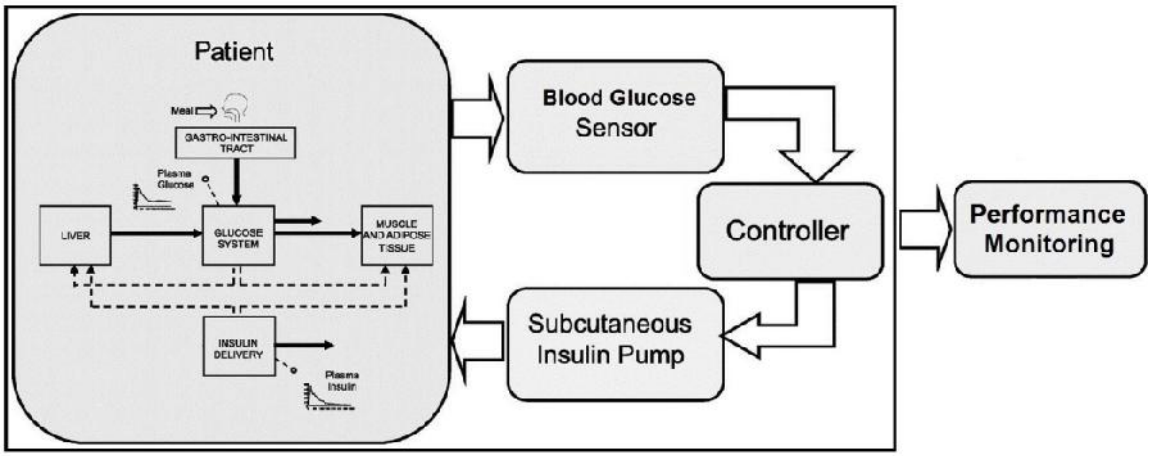

Figure 1. Diagram of the Artificial Pancreas.

Detection of performance degradation in the regulation loop can be achieved by defining an optimal expected behavior in time and comparing it progressively with real measurements obtained through Continuous Glucose Monitoring (CGM) technologies. Significant deviations from optimal behavior are evidence of changes in properly functioning of one or more components of the AP. For example, abnormal glycemic levels can be caused by misjudgments in controller tuning parameters, calibration errors of sensor device or obstructions in the catheter of the insulin pump. The key aspect to be considered is how optimal operation can be characterized in the face of glycemic variability to easily detect discrepancies between correct and faulty functioning. To this aim we focus on the novel methodology presented by De Paula et al. [2]. This policy iteration methodology combines reinforcement learning with Gaussian Processes (GP) approximation to obtain an optimal control policy π_G that guarantee low glycemic variability in Type 1 diabetic patients. Assuming an optimal control policy π_G is used to control the glucose-insulin dynamics over time, a certain reference behavior allows making a conjecture on expected BG levels. This expected system response is generated through a GP predictor using past BG levels and insulin bolus administrated. The reference optimal behavior obtained is compared at every sampling time with sensor readings to detect deviations from optimal performance in controlling glycemic variability. Significant deviations stand for anomalies in the blood regulation loop caused by ill-functioning of any of the components in the AP.

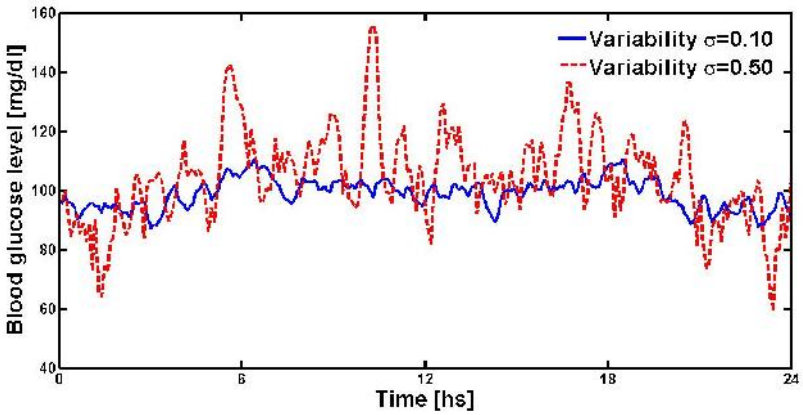

Figure 2. Simulated glycemic variability using variance parameter σ=0.10 and σ=0.50.

Methodology

The Bergman's two-compartment minimal model parameterized as described in Acikgoz and Diwekar [3] is used to model the glucose-insulin dynamics in a simulated patient with proper addition of an Ito's stochastic process to capture inter and intra-patient variability using the variance parameter σ. Reference closed-loop behavior is accomplished through a GPDP controller. GPDP is an approximate value function method that integrates reinforcement learning with Gaussian Processes (GP) for seeking an optimal control policy in the face of uncertain dynamics. The obtained control policy is compactly represented using the hyper-parameters of a GP over a wide range of physiological states. An algorithm of the GPDP using the transition dynamics also modeled as a GP and Bayesian active learning was presented by De Paula y Martínez [4]. Gaussian processes (GPs) are also particularly useful for their flexibility, facilitating accurate prediction even in the absence of strong physical models and allowing us to work within a complete Bayesian probabilistic framework.

By reducing excessive glycemic variability, patients can delay or avoid serious long term complications, including heart disease, kidney failure, blindness, and strokes [5]. To guaranteeing satisfactory performance, the AP will have to overcome conditions of metabolic disturbance and glycemic variability due to changes in diet, exercise, stress, etc. Therefore, it is essential for diabetes management software to be able to calculate glycemic excursions and to ascertain the quality of blood glucose regulation, that is, whether control is consistent or fraught with extreme variations indicating suboptimal disease management. Figure 2 depicts results obtained for a 1-day simulation of the glucose controlled dynamics using the optimal control policy π_G. Different degrees of glycemic variability were simulated through the variance parameter σ.

Current closed-loop systems are limited by suboptimal performance of the available technology. Any failure on the sensor or insulin device may easily lead to termination of closed-loop operation. The reliability of wireless communication between the components also needs to be addressed [6]. More specifically, a number of studies have concluded that prevailing technology still continues to face challenges in terms of sensitivity, stability, calibration, and the physiological time lag between plasmatic glucose and interstitial glucose concentration [7]. Kovatchev et al. [8] presented an *in silico* study describing insulin kinetics in diabetic patients facing different severe scenarios. Similarly, Breton et al. [9] performed a simulation study to assess the effect of calibration errors and time lag between plasmatic and interstitial BG, resulting from standard BG sensor monitoring.

Sensor readings affected by ill-calibration and time lag is simulated based on the work of Facchinetti et al. [10].
Clarke error grid analysis (EGA) was originally used to assess the clinical significance of differences between a glucose measurement technique being tested against venous BG reference measurements [11]. Eventually, the Clarke EGA became accepted as one of the gold standards for determining the accuracy of BG meters. The method uses a Cartesian diagram where values predicted by the technique under scrutiny are displayed on the *y*-axis, whereas the values provided by the reference method are displayed on the *x*-axis, as it is shown in Figure 3. The diagonal line represents the perfect agreement between the two readings, whereas the points below and above the 45^{o} line indicate, respectively, overestimation and underestimation of the actual glycemic values. Region A are those glucose values within 20% of the reference sensor, Region B contains points that are outside of 20% but would not necessary lead to inappropriate diabetes treatment. Region C corresponds to those points leading to unnecessary treatment, whereas Region D are those points indicating a potentially dangerous failure to detect hypoglycemia or hyperglycemia. Finally, in Region E are located those readings that would confuse treatment of hypoglycemia for hyperglycemia and vice-versa [12].

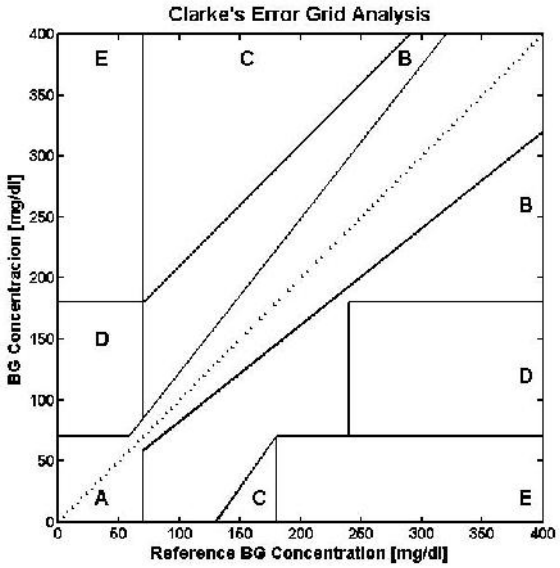

Figure 3. Classification of data according to Clarke EGA.

The novelty in this work is the use of Clarke grid as a performance monitoring tool which is able to detect abnormal glycemic levels caused by improper functioning or degradation in the components that made up an AP. The Clarke grid can be employed to visually detect deviations in BG levels from expected levels that would result from an optimal control policy for glycemic variability. Values along the *x*-axis in the grid represent predicted BG levels under proper insulin administration when the entire control loop is behaving optimally under uncertainty. Values along the *y*-axis represent BG levels obtained through a continuous monitor holding different levles of calibration errors and time lags between plasmatic glucose and interstitial fluid. Expected values are obtained through a GP prediction model, incorporating recursively control actions resulting from implementing the GPDP control policy obtaind in [4]. An autoregressive model is performed using as inputs for the prediction algorithm the BG levels for last hour as well as insulin boluses

administrated in previous time step. The predicted BG level is used as a new input to calculate the next insulin bolus to be administrated assuming the optimal control policy is implemented in the prediction horizon.

Results

To highlight the applicability of Clarke EGA grid as a monitoring tool let's consider a 1-day predicted profile of a simulated Type 1 diabetic patient, merged in order to generate a single time series consisting of 240 points with sampling time 6 min. Then, the Clarke EGA was used to assess the point accuracy of the measured sensor time series versus the corresponding predicted profile as reference for optimal blood glucose regulation. In other words, in contrast to the Kovatchev work [12], reference BG and sensor BG time series were replaced by measured and predicted time series, respectively. This allowed us to evaluate the clinical impact of ordinary glucose sensors –that is affected by time lag, poor calibration or patient variability- in diabetes management. To simulate the reference predicted profile that appear on the *x*-axis of the grid, Ito's variability parameter was set to σ=0.10, calibration error s=2% and time lag 5 min. Sensor parameters are progressively ill-adjusted to simulate performance degradation. The impact of any departure from expected behavior is depicted by the percentage of points that fall into each zone of the Clarke grid.

Table 1. Accuracy of sensor BG predictions according to Clarke EGA.

Zones	**Variability parameter s = 0.10**			
	Calibration error s = 2%		Calibration error s = 20%	
	Time lag = 5min	Time lag = 20min	Time lag=5min	Time lag=20min
Accurate[1]	57.5	53.5	32.1	29.6
Benign[2]	42.5	46.5	67.9	70.4
Inaccurate[3]	0	0	0	0
Zones	**Variability parameter s = 0.50**			
	Calibration error s = 2%		Calibration error s = 20%	
	Time lag=5min	Time lag=20min	Time lag=5min	Time lag=20min
Accurate	36.6	33.6	15.1	14.1
Benign	61.4	64.3	79.6	78.7
Inaccurate	2	2.1	5.3	7.2

Data are %

As can be seen in Figure 4, half of the point estimates for Ito variability parameter σ=0.10 fell into the A zone when the calibration error was set to 2%. Note that in Table 1all points are inside the range for benign errors (A+B zones). There are no points in the C, D or E zones. A larger amount, about 20%, of benign errors was observed when the calibration error was increased to 20%. However, no significant changes occurred when the time lag was significantly increased from 5 min to 20 min. A considerable reduction of points within zone A was observed when the variability parameter was set to σ=0.50. Though, an increase from 2% to 7.2% of points estimates accounts for inaccurate readings, no errors in the type of treatment needed were detected (points within zone E). A higher calibration error leads to a reduction, up to 50%, of points inside zone A, most moved to zone B.

[1] Accurate readings Zone = A
[2] Benign errors Zone = B
[3] Inaccurate readings Zone = Σ(C+D+E)

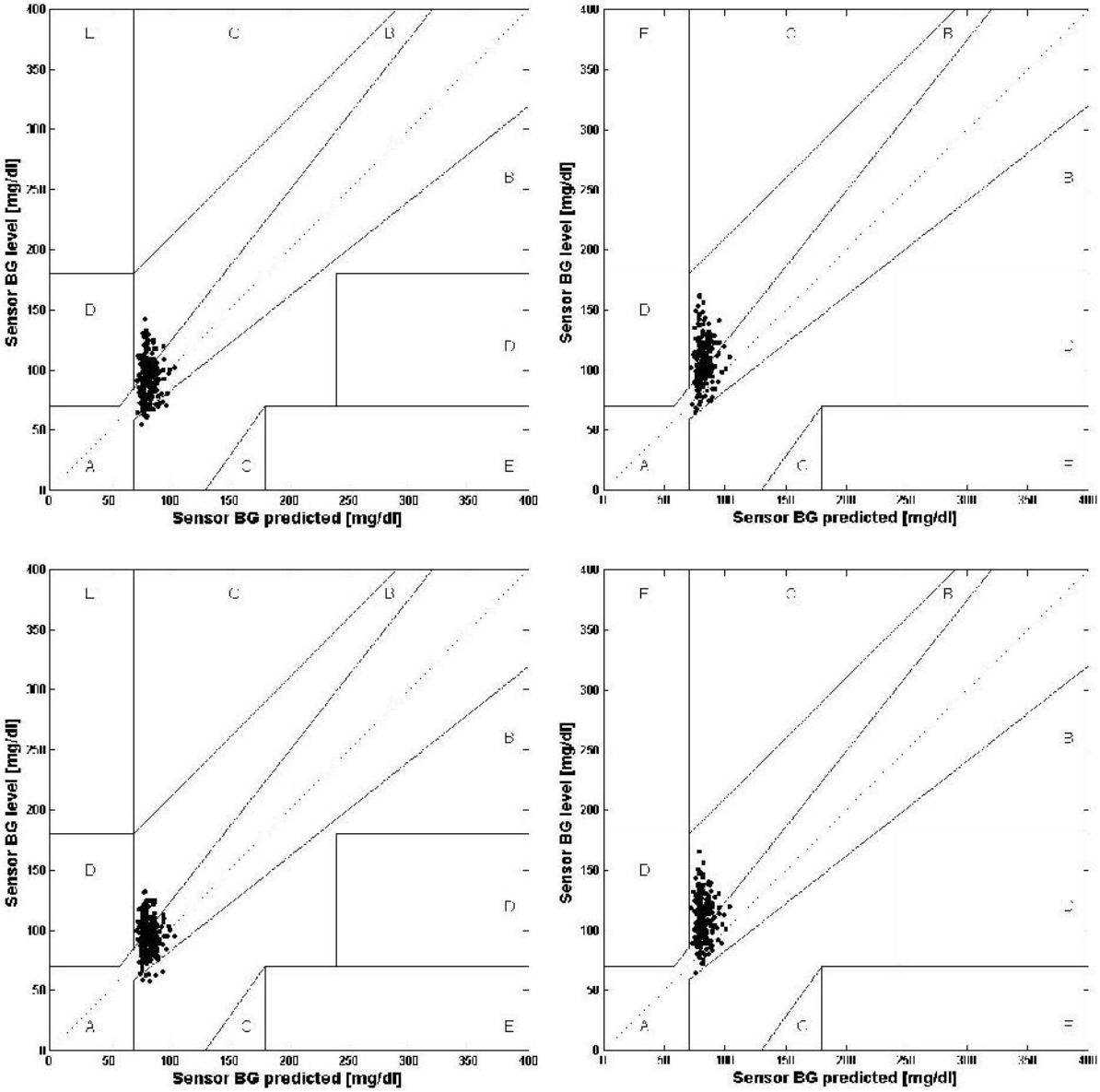

Figure 4. Clarke EGA for accuracy of continuous sensor BG predictions using variability parameter σ=0.10. Observed glucose measurements (y-axis) versus reference predicted sensor BG (x-axis). From top to bottom, point estimates for time lag=5 min and lag=20 min respectively. From left to right, point estimates for calibration error s=2% and s=20% respectively.

Clarke grids are presented in Figure 4 to evaluate continuous BG levels when sensor reading is affected by performance degradation. Values along the *x*-axis represent the expected behavior of the system when the sensor is operating under optimal conditions. This state of affairs is achieved by setting the time lag at 5 min and the calibration error at 2%, which are suited values for a properly calibrated sensor. Values on the *y*-axis correspond to sensor readings affected by delay and miscalibration. Two different settings for time lags as well as calibration errors in sensor measurements were evaluated. Firstly, when glycemic variability parameter was set to σ=0.10, the accuracy in glycemic control was reflected through a dense cloud of points, without considerable changes among grids. Proximity to zone D corresponds to an overcorrection of hyperglycemia, leading to

symptoms of hypoglycemia, due to excessive insulin administration. An adjustment of the controller tuning parameters, aims to lower the insulin bolus administrated and will guarantee a safer treatment of the patient. No evident changes occurred when the time lag was incremented from 5 min to 20 min. Nevertheless, an increment of 20% in the biased magnitude given by the glucose sensor causes a wider cloud of points, which is larger respect the *y*-axis and a drift upward the diagonal line caused by an upper mean value in the BG levels measured. No discrepancy regarding to the projection over the *x*-axis was detected, because no change in the expected behavior of the properly operating sensor was introduced. Results presented in Figure 4 and Figure 5 vividly provides evidence of the robustness of the control policy based on GPDP proposed in [4].

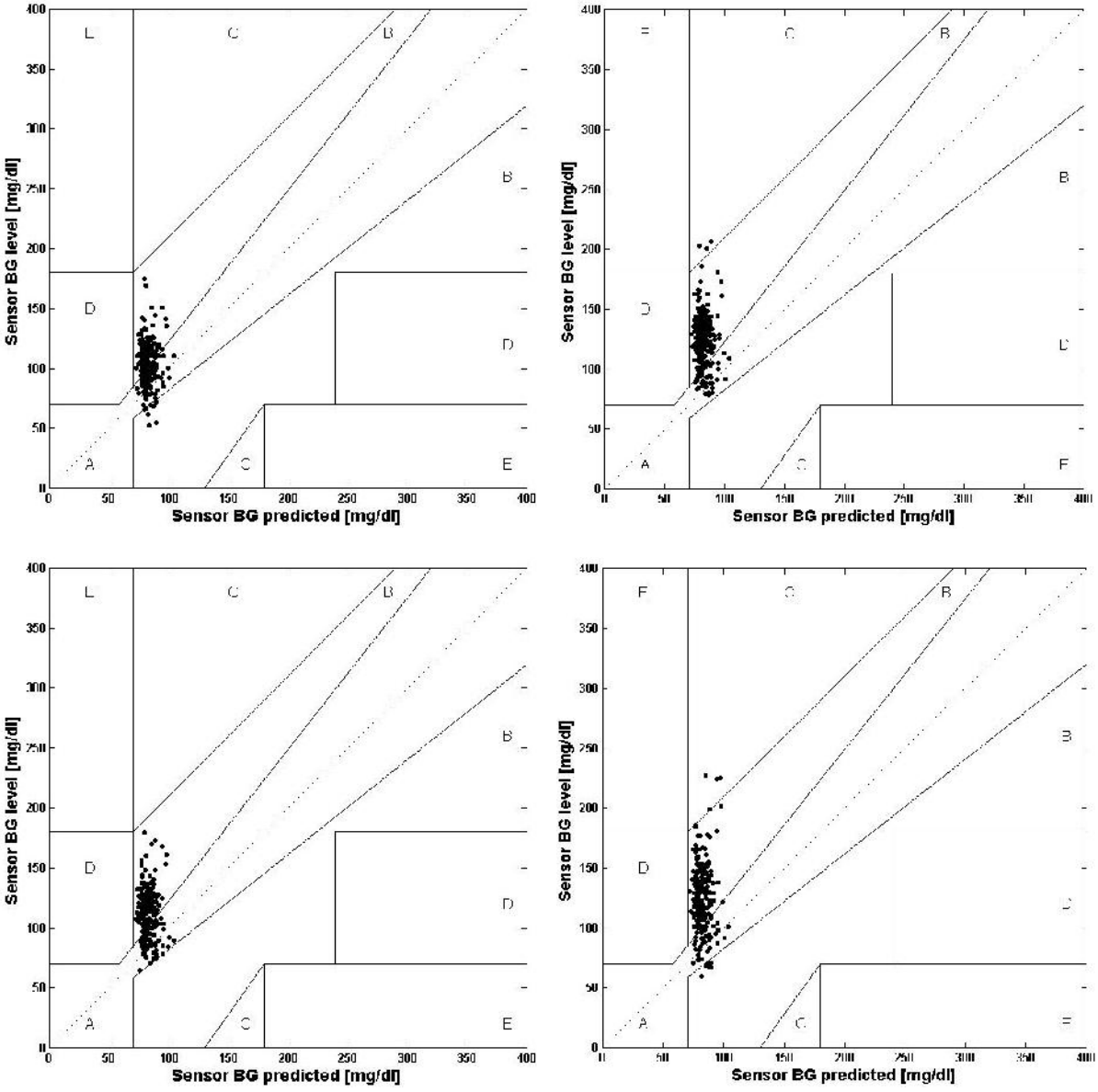

Figure 5. Clarke EGA for accuracy of continuous sensor BG predictions using variability parameter σ=0.50. Observed glucose measurements (y-axis) versus reference predicted sensor BG (x-axis). From top to bottom, point estimates for time lag=5min and lag=20min respectively. From left to right, point estimates for calibration error s=2% and s=20% respectively.

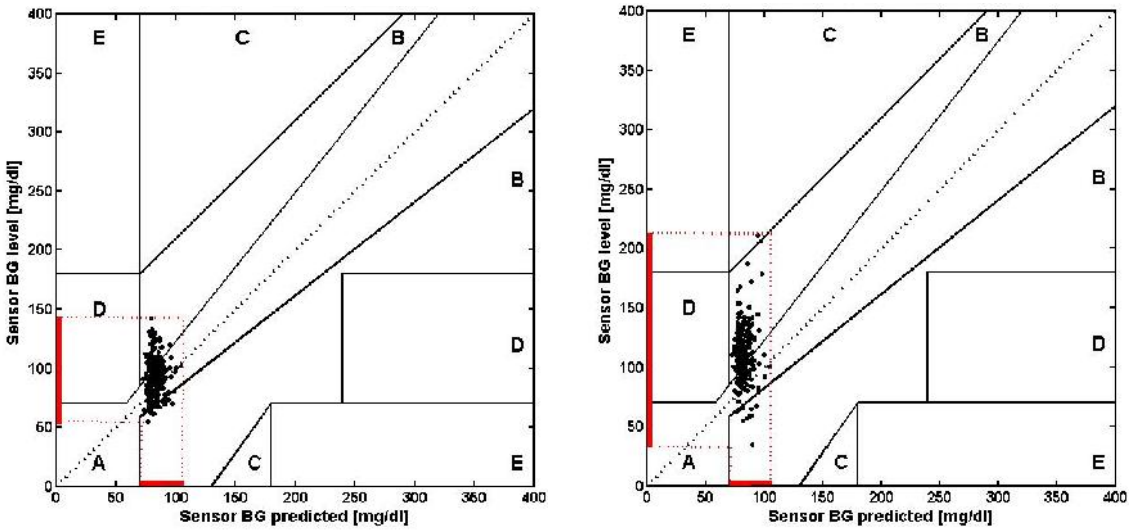

Figure 6. Projected amplitude of glycemic excursions for predicted and measured BG. Left, projected amplitude for σ=0.10, time lag= 5 min and s=2%. Right, projected amplitude for σ=0.50, time lag= 20 min and s=20%.

Grids for continuous BG levels using variability parameter σ=0.50 are presented in Figure 5. As a result of the increased variability, refer to Figure 2, the cloud of point exhibits a marked spreading with some of points falling outside of zone A+B. The upward drift from the diagonal line is more evident than in the previous case. The overcorrection of hyperglycemia, because of excessive control actions, still prevails. Due to a robust control policy, no significant faults on the treatment achieved were perceived, even when variability and sensor calibration errors were fixed to their maximum values.

Considering the amplitude of the axial projection of points as in Figure 6, excessive glycemic variability can be detected, a feasible support to other metrics of glycemic variability [13]. As maximum excursions of measured BG levels increase, also does the amplitude of the projection of points on the *y*-axis. Since *x*-axis depicts for the expected values of BG readings based on the assumption of an optimally operating environment, no changes in the respective amplitude were observed. An important hint derived from Figure 6 is that, for a glycemic system functioning near optimality, the amplitude of projected points on *y*-axis (points corresponding to measured BG levels), should be close to the amplitude of projected points on the *x*-axis (points corresponding to the optimal behavior of the artificial pancreas).

Discussion

An AP is safety-critical system which gives rise the critical issue of proper supervision of the control loop functioning and early detection of abnormal glycemic variability as well as performance deterioration in the control loop components. This work shows how the Clarke EGA tool can be employed as a performance monitoring approach to detect functional degradation in the components of an AP and guarantee adequate control of glycemic variability. A simulated model of a glucose sensor affected by different time lags and calibration errors was used to visually evaluate to what degree deviations from expected BG values have an impact in the performance grid. We emphasize the importance of using a well established method in clinical practice to assess performance degradation. A monitoring tool based on Clarke EGA is not only less time consuming for investigators, but

also easier to interpret for clinicians. On the other hand, the availability of an optimal control policy allows building a reference behavior over time which allows comparing it with current sensor readings to identify deviations from optimal loop behavior in blood glucose regulation. The combination of a widely used grid analysis tool and an expected BG signal as a reference provides a useful and comprehensive approach to monitor the performance of an AP. Current research efforts focuses in the study of the remaining components of an AP such as the control algorithm or the infusion pump using different monitoring tools to evaluate not only point accuracy in BG levels but also the dynamics involved in CGM. Some related tools are the Control-Variability Grid Analysis [14], the Rate EGA and the Continuous Glucose EGA [12].

References

[1] Jaremko J., Rorstad O.: Advances towards the implantable artificial pancreas for treatment of diabetes. Diabetes Care, 21, 444–50, 1998.
[2] De Paula M., Martinez E. Controlling blood glucose variability under uncertainty using reinforcement learning and Gaussian processes. Submitted to Applied Soft Computing, 2013.
[3] Acikgoz S., Diwekar U. Blood glucose regulation with stochastic optimal control for insulin-dependent diabetic patients. Chemical Engineering Science, 65, 3, 1227-1236, 2010.
[4] De Paula M., Martinez E. Probabilistic optimal control of blood glucose under uncertainty. Computer Aided Chemical Engineering, 30, 1357-1361, 2012.
[5] Wiley M., Bunescu R., Marling C., Shubrook J., Schwartz F. Automatic detection of excessive glycemic variability for diabetes management. Proceedings of the Tenth International Conference on Machine Learning and Applications, 148-154. IEEE Press, New York, NY, 2011.
[6] Elleri D., Dunger D., Hovorka R. Closed-loop insulin delivery for treatment of type 1 diabetes. BMC Med., 9, 120, 1-9, 2011.
[7] Klonoff D. The artificial pancreas: how sweet engineering will solve bitter problems. J. Diabetes Sci. Technol.,1, 72–81, 2007.
[8] Kovatchev B., Dalla C., Cobelli C. In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes. J. Diabetes Sci. Technol., 3, 44 -55, 2009.
[9] Breton M., Kovatchev B. Analysis, modeling, and simulation of the accuracy of continuous glucose sensors. J. Diabetes Sci. Technol., 2, 853–862, 2008
[10] Facchinetti A., Sparacino G., Cobelli C. Modeling the error of continuous glucose monitoring sensor data: critical aspects discussed through simulation studies. J. Diabetes Sci. Tech., 4, 4–14, 2010.
[11] Clarke W., Cox D., Gonder-Frederick L., Carter W., Pohl S. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. Diabetes Care, 10, 622–628, 1987.
[12] Kovatchev B., Cox D., Gonder-Frederick L., Clarke W. Evaluating the accuracy of continuous glucose-monitoring sensors: continuous glucose-error grid analysis illustrated by TheraSense Freestyle Navigator data. Diabetes Care, 27, 8, 1922–1928, 2004.
[13] Marling C., Shubrook J., Vernier S., Wiley M., Schwartz F. Characterizing blood glucose variability using new metrics with continuous glucose monitoring data. J. Diabetes Sci. Technol., 5, 871–878, 2011.
[14] Magni L., Raimondo D., Dalla Man C., Breton M., Patek S., De Nicolao G., Cobelli C., Kovatchev B. Evaluating the efficacy of closed-loop glucose regulation via control-variability grid analysis. Journal of Diabetes Science and Technology, 2, 4, 630–635, 2008.

Contact

Luis Avila. INGAR (CONICET-UTN). Santa Fe S3002 GJC. avilaol@santafe-conicet.gob.ar.
Ernesto Martínez. INGAR (CONICET-UTN). Santa Fe S3002 GJC. ecmarti@santafe-conicet.gob.ar.