

Sensor Calibration Models for a Non-Invasive Blood Glucose Measurement Sensor

M. Stemmann, F. Ståhl, J. Lallemand, E. Renard and R. Johansson

Abstract—A calibration model was developed for a non-invasive blood glucose sensor, to determine how the blood glucose data measured by this sensor is related to blood glucose data measured with laboratory capillary finger sticks and to corrupting noise. The variability of calibration models for different patients was analyzed as well as the dynamics of the non-invasive blood glucose sensor according to reference blood glucose measurements and corrupting noise.

I. INTRODUCTION

In order to treat Diabetes Mellitus, it is unavoidable to measure the blood glucose (BG) concentration regularly. By following the changes of the BG concentration, the insulin intake, meals and exercises can be planned so as to keep the BG concentration stable. To help the patient with this task, advisory systems are developed [7], [8].

The BG concentration can be measured frequently sampled every few minutes using Continuous Glucose Measurement (CGM) systems or just a few times per day. Mostly, the measurements are done using blood samples obtained by a finger stick or by a sensor with disposable parts under the skin. However, using non-invasive methods to determine the BG concentration is more convenient for the patient, since it does not require a finger stick.

Since many years, there are a number of different non-invasive BG measurement (NIGM) sensors in development [1], [2], [3]. A non-invasive BG sensor based on spectroscopic measurements was analyzed here. Using mathematical modeling, various different parameters can be predicted from the spectroscopic data, among those BG concentration, the sensor being calibrated individually to each patient to decrease variability.

In this report, a calibration model was determined relating reference BG data and corrupting noise to the measurements of the NIGM sensor. By doing so, the effect of the reference data and the noise on the sensor data could be analyzed. Also, the difference in the calibration models for different patients gives insight into the variability of the non-invasive sensor between different patients. Moreover, using the calibration model obtained, the dynamics of the sensor

according to the reference BG concentration and corrupting noise could be analyzed.

II. METHODS

A. Data Acquisition and Blood Glucose Sensors

The BG data used in this paper was collected from 30 patients during a 3-day in-hospital period from September 2008 to June 2009 at the Centre Hospitalier Universitaire de Montpellier, associated with a European project (DIAdvisor™, [4]). The data of 23 patients had sufficient data quality to be used in this report. The in-hospital measurements allow for standardized conditions with known amounts of carbohydrates, lipids and proteins at specific times for meal intakes.

BG data were collected with the HemoCue Glucose Analyzer (HemoCue) [6] and a NIGM sensor from Ondalys. For HemoCue, between 50 and 150 data points were collected in the 3 days, while the NIGM sensor could measure a BG value every minute. The BG data measured by HemoCue were the reference data, to which the NIGM data were compared to.

The NIGM sensor measures spectroscopic data and determines three different sets of BG data using three different internal models (M1, M2 and M3). The BG estimation accuracy was improved by M1 and M2 and the robustness to patient variability was improved by M3 [5].

In order to make the BG data measured by the NIGM sensor and the HemoCue sensor comparable, the data of HemoCue were interpolated at the time points of the NIGM measurements.

B. The Calibration Model

To model the NIGM sensor, an ARMAX model [10] with an additional constant, the calibration level, added was used:

$$y_k = H_u(z^{-1}) \cdot u_k + H_w(z^{-1}) \cdot w_k + \tilde{d}, \quad (1)$$

For the purpose of the calibration model, the input signal u_k of the ARMAX model was chosen as the BG data measured by HemoCue. The output data y_k was the data of the NIGM sensor to be analyzed. The noise data w_k of the ARMAX model had to be estimated along with the parameters. An additional constant \tilde{d} represented the calibration level, which was estimated by extension of the regression vector. To simulate the calibration model the System Identification Toolbox by MATLAB [11] was used.

The properties of the transfer functions $H_u(z^{-1})$ and $H_w(z^{-1})$ can be examined using, for example, a Bode diagram to evaluate the dynamics of the NIGM sensor according to reference BG measurements and noise.

Manuscript received April 23, 2010.

M. Stemmann, F. Ståhl and R. Johansson are with the Dept. Automatic Control, Lund University, Sweden, Meike.Stemmann@control.lth.se, fredrik.stahl@control.lth.se, Rolf.Johansson@control.lth.se

J. Lallemand is with Ondalys, France, jlallemand@ondalys.fr

E. Renard is with Centre Hospitalier Universitaire de Montpellier, France, e-renard@chu-montpellier.fr

The parameters of the calibration model were estimated using Pseudo-Linear Regression [9], consisting of the following three steps:

1. The input data u_k and output data y_k were fitted to the ARMAX model using high model orders, while ignoring the noise term. Least-squares Identification was used to calculate the parameters. From those parameters, output data could be estimated.
2. The difference between the output data estimated in Step 1 and the measured output data y_k was the noise term to be estimated. If the model orders are high, the estimated noise can be assumed to be an approximation of a white noise sequence [9].
3. The estimated white noise sequence was used together with the input data u_k and the measured output data y_k to estimate the coefficients of the ARMAX model with the desired orders, using least-squares identification. From this, the estimated NIGM data \hat{y}_k could be calculated.

If the ARMAX model is estimated well, the estimated output data \hat{y}_k concurs with the measured data y_k and thus the obtained model describes the NIGM sensor.

C. Model Validation

The BG data measured by the NIGM sensor and the reference BG data from HemoCue were divided into two data sets. The first half of the data were used as identification data to determine the parameters of the calibration model. The second half were used as validation data [9] to verify the determined model by calculating the output of the model using the input data of the verification data and comparing this to the output data of the verification data. For the verification data the noise had to be determined as well. To do so, the Steps 1 and 2 of the pseudo-linear regression described above were repeated with the validation data. Instead of Step 3 however, the ARMAX model determined with the identification data was used to estimate the BG data for the validation data, using the newly calculated noise.

To evaluate the performance of the obtained model, the average absolute relative error, further called error, was used:

$$e = (1/N) \cdot \sum_{k=1}^N |100 \cdot (y_k - \hat{y}_k)/y_k| \quad (2)$$

To determine the variation of a parameter, the variance was used [12].

III. RESULTS

For a sample patient, the ARMAX model determined by pseudo-linear regression [9] as described in the previous section was:

$$\begin{aligned} & [1 - 0.35 \cdot z^{-1} - 0.22 \cdot z^{-2} - 0.16 \cdot z^{-3} - 0.01 \cdot z^{-4} - 0.07 \\ & \quad \cdot z^{-5} - 0.03 \cdot z^{-6} - 0.002 \cdot z^{-7}] \cdot y_k \\ & = [0.07 + 0.24 \cdot z^{-1} - 0.26 \cdot z^{-2} + 0.08 \\ & \quad \cdot z^{-3} - 0.05 \cdot z^{-4}] \cdot u_k + [1 - 0.06 \cdot z^{-1}] \\ & \quad \cdot w_k + 7.94 \end{aligned} \quad (3)$$

The BG data estimated with this model is shown in Fig. 1 together with the noise estimated during the Pseudo-Linear

Regression and the BG data measured by the NIGM sensor for identification and verification data. The estimated data recaptured the measured data in this case with an error of

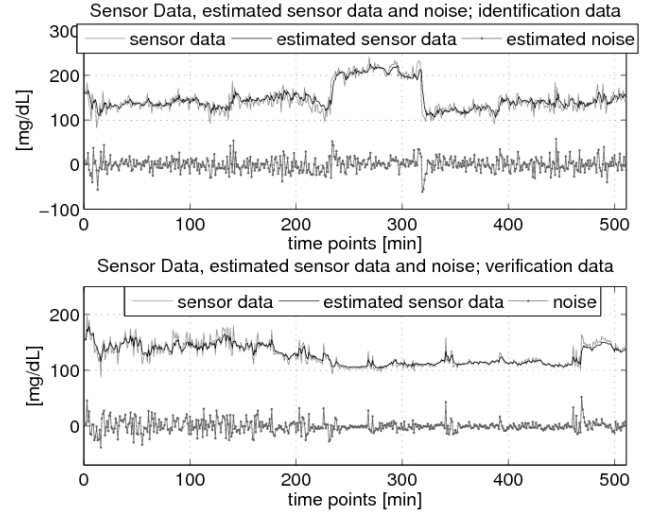


Figure 1: BG data measured by the NIGM sensor (*sensor data*), estimated with the calibration model (3) (*estimated sensor data*) and the noise w_k estimated during Pseudo-Linear Regression (*estimated noise*) for a sample patient for identification and validation data.

8.5 % for the identification data and 3.32 % for the validation data.

A calibration model was determined for all 23 patients and all combinations of the model orders na , nb and nc from 1 to 8. For each of those models, the error between the estimated BG data and the data measured by the NIGM sensor was calculated according to Eq. (2), both for the identification and the validation data. There was a model order combination for each one of the 23 patients, that had an error which was smaller than for all other model order combinations. This was the minimal error for a patient. To decide which error is smallest for each patient, the validation data was used, resulting in a smaller error for the validation data than for the identification data in general. The minimal errors for all patients were collected in a data set. The minimum, maximum, average value and variance of this

TABLE I
ERRORS [%]

M	Max		Min		Avg		Var	
	Ident	Val	Ident	Val	Ident	Val	Ident	Val
1	12.49	5.37	1.40	1.360	5.14	2.46	22.16	2.66
2	8.43	3.79	1.58	0.81	4.09	1.70	9.24	1.33
3	12.08	3.89	3.88	1.30	7.83	2.11	8.32	1.24

Maximum (Max), minimum (Min), average values (Avg) and variance (Var) of the minimal errors between the estimated data and the measured NIGM data of the calibration models over all 23 patients.

data set of minimal errors are shown in Table I for the NIGM sensor internal models M1, M2 and M3. The smallest error was achieved by the internal sensor model M2, both M1 and M3 had a larger error.

The minimum, maximum and average values of the model orders leading to the minimal error for each patient among all patients are shown in Table II together with the variance for the model orders over all patients.

The model orders of M1 and M2 varied significantly between the minimum number 1 and the maximum number 8, whereas the model orders for M3 varied much less. The variance of na for M1 and M2 was 7 % and 9 %, though it was 0.77 %, 0.87 % and 0.04 % for na , nb and nc of M1

TABLE II
MODEL ORDERS FOR MINIMAL ERROR

M	Max			Min		
	na	nb	nc	na	nb	nc
1	8	8	1	2	1	1
2	8	8	2	1	2	1
3	7	4	2	5	1	1

M	Avg			Var		
	na	nb	nc	na	nb	nc
1	6.04	6.39	1	7.59	6.34	0
2	5.74	6.04	1.70	9.02	6.50	0.22
3	5.70	2.35	1.04	0.77	0.87	0.04

Maximum (*Max*), minimum (*Min*), average values (*Avg*) and variance (*Var*) of the model orders na , nb and nc for the calibration models of all 23 patients.

respectively. The order nc had a small variance compared to na and nb .

The Bode diagram of the transfer functions $H_u(z^{-1})$ and $H_w(z^{-1})$ according to Eq. (1) for the calibration models giving the minimal error between estimated and measured NIGM data for each patient are shown in Fig. 2 and Fig. 3 respectively for M1. The Bode diagram for M2 and M3 looked similar. The magnitude diagram of the Bode diagram for $H_u(z^{-1})$ shows that the reference BG data u_k were damped over the whole frequency range and occasionally even slightly amplified for high frequencies. The noise spectrum $H_w(z^{-1})$ showed that the noise w_k was mostly amplified for low frequencies and slightly damped for higher frequencies. Furthermore, the cross-over frequency seemed to be similar for all the patients, while the gain seemed to vary.

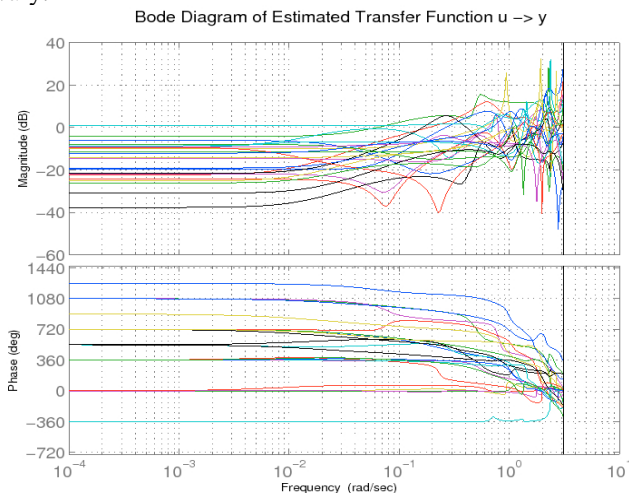


Figure 2: Bode diagram of the transfer functions $H_u(z^{-1})$ from the reference BG value to the measured output for the calibration models estimated for all 23 patients using the NIGM sensors with M1.

Figure 4 shows the Clarke-Error Grid [13] of the BG data measured by the NIGM sensor in comparison to the BG data measured by HemoCue. It can be seen, that for the point accuracy, the data points are spread mostly in regions A and

B, while they are widely spread over all regions for the rate accuracy. For the point accuracy, there were 55.9% of the data in region A and 43.23% in region B, while for the rate

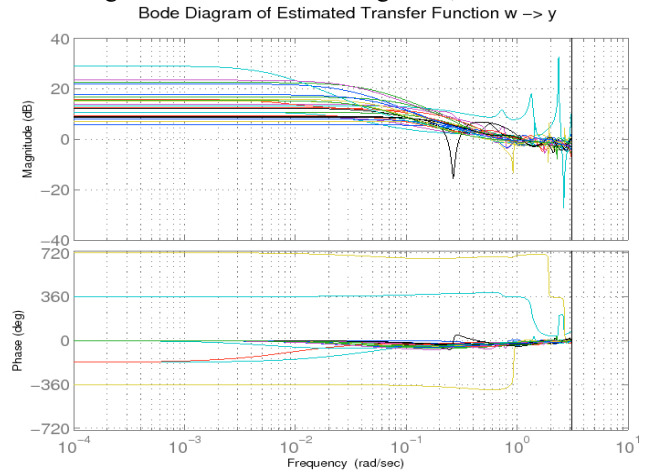


Figure 3: Noise Spectrum, estimated for all 23 patients using the NIGM sensors with M1.

accuracy only 8% of the data were in region A and 12% in region B. This can be related to the amplification of high-frequency components in Fig. 2. Due to high-frequency components in the NIGM measurements, the rate of BG change determined by the NIGM sensor did not follow the rate of BG change measured by the reference result providing HemoCue sensor.

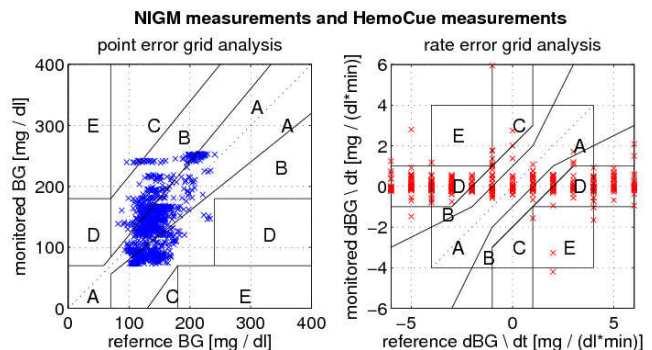


Figure 4: Clarke-Error Grid for an example patient with the internal NIGM sensor model M1.

IV. DISCUSSION

For the NIGM sensor with the sensors internal models M1, M2 and M3, a calibration model in form of an ARMAX model was developed. The BG data measured by HemoCue provided reference BG measurements (being the input) and the estimated BG data were supposed to recapture the measured NIGM sensor data (being the output). The estimated BG data and noise together with the NIGM measurements are shown in Fig. 1, while the Clarke error Grid in Fig. 4 compares the NIGM measurements to the HemoCue measurements.

By varying the model orders, the optimal calibration model giving the minimal error between the estimated and the measured data were determined for every patient. When taking the average of the minimal errors over all patients, the

NIGM sensor internal model M2 had the smallest error. Both M1 and M3 had significantly larger errors. The model orders of the optimal calibration models varied significantly among the patients for the internal sensor models M1 and M2, but less for M3. From this it seemed the model order for the calibration model depended strongly on the patient for M1 and M2. With the internal sensor model M3 however, the model orders of the calibration model seemed to be more robust with respect to the different patients. This supported the fact, that the internal sensor model M3 was designed to improve robustness, while M1 and M2 were designed to improve the accuracy of the BG measurements.

Analyzing the Bode diagram of the transfer functions from the reference BG data to the estimated blood-glucose data of the calibration model, it showed that in most cases low frequency content was damped, while high frequency content was slightly amplified. In case of the transfer function from the corrupting noise to the estimated output data, the noise was mostly amplified at low frequencies, while it was slightly damped at high frequencies.

High-frequency components and amplified noise of the NIGM measurements might have led to a poor performance of the NIGM sensor with regards to the rate of BG change. Especially the effect of the high-frequency components can be seen in the Clarke Error Grid in Fig. 4. Concerning the point accuracy 99.13% of the data were in the clinically accurate region A and in the region B of benign errors, while for the accuracy of the BG rate of change only 20% were in regions A and B, which could result from high-frequency components in the NIGM signal.

According to [14], the HemoCue BG sensor had 98% of the data within $\pm 20\%$ of BG measurements of the Yellow Springs Instrument (YSI) glucose oxidase analyzer, and therefore HemoCue was claimed to be interchangeable with YSI. However, these small inaccuracies could have been one possible source for the noise. As HemoCue in this work was used as a reference for the Ondalys sensor, no further error analysis of HemoCue was made.

Because of noise amplification and occasionally high amplification of the high-frequency content of the reference BG data compared to the low-frequency content leading to poor BG rate performance, the NIGM sensor was not expected to be available in industry presently.

V. CONCLUSION

For the non-invasive blood glucose (NIGM) sensor by Ondalys calibration models were determined for data measured with 23 different patients. The NIGM sensor used two models internally to improve accuracy of blood glucose measurements and a third one to improve robustness. The calibration models related the NIGM measurements to reference blood glucose measurement, provided by the HemoCue Glucose Analyzer [6], and to corrupting noise. The variation in the model orders of the calibration models leading to a minimal estimation error for each patient supported the robustness improvements of the third internal sensor model in comparison to better measurement accuracy of the second model.

Furthermore, the NIGM sensor damped the reference blood glucose measurement over the whole frequency range and occasionally even damped low-frequency content more than high-frequency content. The corrupting noise was amplified for low frequencies and slightly damped for high frequencies.

ACKNOWLEDGMENT

This research was supported by the European project DIAdvisorTM, FP7 IST-216592.

REFERENCES

- [1] A. Tura, A. Maran, G. Pacini, "Non-invasive glucose monitoring: Assessment of technologies and devices according to quantitative criteria", *Diabetes Research and Clinical Practice*, vol. 77, no. 1, 2007, pp. 16-40
- [2] M. A. Arnold, "Non-invasive glucose monitoring", *Current Opinion in Biotechnology*, vol. 7, no. 1, 1996, pp. 46-49
- [3] O. S. Khalil, "Non-Invasive Glucose Measurement Technologies: An Update from 1999 to the Dawn of the New Millennium", *Mary Ann Liebert*, 2004, vol. 6, no. 5, pp. 660-697
- [4] DIAdvisor: www.diadvisor.eu
- [5] Ondalys: www.ondalys.fr
- [6] HemoCue® AB Ängelholm, Sweden: www.hemocue.com
- [7] F. Ståhl, R. Johansson: "Short-Term Diabetes Blood Glucose Prediction Based On Blood Glucose Measurements" in *30th Annual Int. IEEE EMBS Conf. (EMBC 2008)*, Vancouver, BC, Canada, 2008, pp. 291-294
- [8] C. Cobelli, C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, B.P. Kovatchev, "Diabetes: Models, Signals, and Control", *IEEE Reviews in Biomedical Engineering*, 2009, vol.2, pp.54-96
- [9] R. Johansson, *System Modeling and Identification*, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [10] T. Söderström, P. Stoica, *System Identification*, London: Prentice Hall Int., 1989
- [11] L. Ljung, "System Identification Toolbox: User's Guide", The MathWorks, Inc, Natick, MA., 1988-2010
- [12] I. N. Bronstein, K. A. Semendjajew, G. Musiol, H. Mühlig, *Taschenbuch der Mathematik*, Thun und Frankfurt am Main: Verlag Harri Deutsch, 2001, 5th edition
- [13] W. L. Clarke, "The Original Clarke Error Grid Analysis (EGA)", *Diabetes Technology & Therapeutics*, 2005, vol. 7, no. 5
- [14] A. D. M. Stork, H. Kemperman, D. W. Erkelens, T. F. Veneman, "Comparison of the accuracy of the HemoCue glucose analyzer with the Yellow Springs Instrument glucose oxidase analyzer, particularly in hypoglycemia", *Eur.J Endocrinology*, 2005, no.153, pp. 275-281, 2005