Linear Modeling and Prediction in Diabetes Physiology

Cescon, Marzia

2011

Document Version:
Publisher's PDF, also known as Version of record

Linear Modeling and Prediction in Diabetes Physiology

Marzia Cescon

Department of Automatic Control
Lund University
Lund, June 2011

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

ISSN 0280–5316
ISRN LUTFD2/TFRT--3250--SE

Printed in Sweden,
Lund University, Lund 2011

Abstract

Diabetes Mellitus is a chronic disease characterized by the inability of the organism to autonomously regulate the blood glucose level due to insulin deficiency or resistance, leading to serious health damages. The therapy is essentially based on insulin injections and depends strongly on patient daily decisions, being mainly based upon empirical experience and rules of thumb. The development of a prediction engine capable of personalized on-the-spot decision making concerning the most adequate choice of insulin delivery, meal intake and exercise would therefore be a valuable initiative towards an improved management of the desease.

This thesis presents work on data-driven glucose metabolism modeling and short-term, that is, up to 120 minutes, blood-glucose prediction in Type 1 Diabetes Mellitus (T1DM) subjects.

In order to address model-based control for blood glucose regulation, low-order, individualized, data-driven, stable, physiological relevant models were identified from a population of 9 T1DM patients data. Model structures include: autoregressive moving average with exogenous inputs (ARMAX) models and state-space models.

ARMAX multi-step-ahead predictors were estimated by means of least-squares estimation; next regularization of the autoregressive coefficients was introduced. ARMAX-based predictors and zero-order hold were computed to allow comparison.

Finally, preliminary results on subspace-based multi-step-ahead multivariate predictors is presented.

Acknowledgments

First of all I would like to thank my supervisor Rolf Johansson for having introduced me to graduate studies and for the opportunity he gave me to join the DIAdvisor project. I'm truly indebted with Anders Rantzer and Per Hagander for their help and encouragment during the last couple of months. Financial support was provided by the European FP7 IST-216592 within the project DIAdvisor ^{TM}. All the DIAdvisor collegues have been excellent sources of insightful discussions at many levels. In particular, their expertise in various fields other than control theory allowed me to have a broader view on many topics and is therefore highly appreciated. Further, they were a boost in pursuing the same objectives: keep on the good work! I would like to thank the persons that proofread the manuscript and provided me with feedback on the subject matter: Rolf Johansson, Dawn Tilbury, Per Hagander, Fredrik Ståhl and Aivar Sootla.

Esteem and gratitude go to Fredrik Ståhl with whom I have been so lucky to spend the good days and the bad days in DIAdvisor over the past almost three years, both at the Department and around Europe for the various project-related meetings. During the last academic year I have had the pleasure to share the office with Dawn Tilbury: not only a skilled researcher and expert in many fields, but also an excellent and inspiring office mate!

The Department of Automatic Control provides a great environment for PhD studies and I'm honoured to be part of it. The secretaries Eva Schildt, Britt-Marie Mårtensson, Eva Westin and Ingrid Nilsson helped me with various administrative issues but most importantly

kept an eye on me as my mum would have done! Anders Blomdell and Leif Andersson deserves credit for computer support, Leif is also acknowledged for the typesetting of this thesis. I owe my gratitude to the Ph.D. fellows Anna, Meike, Isolde, Maria, Daria, Mikael, Karl M., Erik and Aivar for their moral support! Becoming homeless almost immediately after my arrival in Lund brought its benefits: Rolf and Siv Braun! I have really enjoyed being your daughter even for only three months. I acknowledge the SK RAN triathles, in particular Angela, Emil, Esther, Frida, Johannes, Marianne, Matthew and Rickard for being my support system over the past year. Last, I would like to thank my family for being supportive in all I have been doing and for making me feel your love despite the distance!

Contents

1

Introduction

1.1 Context and motivation

Diabetes Mellitus is a chronic disease of disordered glucose metabolism due to defects in either insulin secretion from the pancreatic β-cells or insulin action. Type-1 diabetes (T1DM), also called insulin-dependent diabetes mellitus (IDDM) is characterized by no production of insulin what so ever, whereas type-2 diabetes is caused by decreased sensitivity of the tissues to the metabolic effect of insulin. The basic effect of insulin lack or insulin resistance is to prevent the efficient uptake and utilization of glucose by most cells of the body, resulting in abnormally high blood sugar levels (hyperglycemia). Sustained hyperglycemia is associated with acute ketoacidosis, nephropaty, rethinopaty, neuropathy and damages to the cardio-vascular system [Williams and Pickup, 1992].

With estimated 285 million affected people worldwide in 2010, diabetes mellitus is one of the most widespread diseases and causes 4 million deaths per year, ranking fifth by cause-specific mortality in most high-income countries, after communicable diseases, cardiovascular diseases, cancer and injury, being undoubtedly one of the most challenging health problems in the 21st century [IDF, 2011]. Healthcare expenditures on diabetes were estimated to account for 11.6% of the total healthcare expenditure in the world in 2010. Beside direct costs, indirect costs associated with lost workdays, restricted activity

days, lower productivity at work, premature mortality and permanent disability reached 50% of the direct costs.

Both from quality of life and economic perspectives it is then critical for diabetes patients to regulate their blood glucose tightly, keeping its level within the target range, i.e., 70-140 [mg/dL] [The American Diabetes Association, 2010], through intensive insulin therapy. The strategy comprises test of blood glucose levels at least four times a day, taking insulin with every meal by injections or subcutaneous infusions by a pump and patient assistance by health care team.

Although the standard tools in diabetes care improved significantly during the last decades, including the availability of self-monitoring of blood glucose (SMBG), insulin pen injectors, pumps and insulin analogues, insulin therapy remains one of the most difficult to manage. As a matter of fact, the treatments still strongly depend on the patient's daily decisions about insulin delivery adaptations. Many factors have to be considered in this decision process: health status, current blood glucose level, blood glucose target, insulin sensitivity, diet, meal composition, foreseen activities and individual experience of insulin effects on blood glucose level. Meanwhile, failure in management of insulin therapy has significant impact on short-, medium- and long-term prospects.

The task is non-trivial, for this reason several attempts to the development of a biomedical device performing close-loop control of blood glucose have been made in the past and are still undergoing nowadays, within the so-called Artificial Pancreas Project [AP, 2011]. However, many patients would rather prefer to maintain control over their own blood glucose management and would therefore benefit from some sort of decision support. The availability of a blood glucose predictor that would inform the patient on the near future blood glucose and offer advice on how to modulate insulin therapy in relation to food intake and out-of-target glucose deviations would therefore be highly valuable.

The DIAdvisor ™ project

Against this background, the European FP7-IST research project DIAdvisor ™ [DIAdvisor, 2011] aims at developing a personalized blood glucose predicting system and an advisory control system which can be used on the spot to assist the users in different daily situations, predicting hyperglycemic deviations, following meals and stressful events,

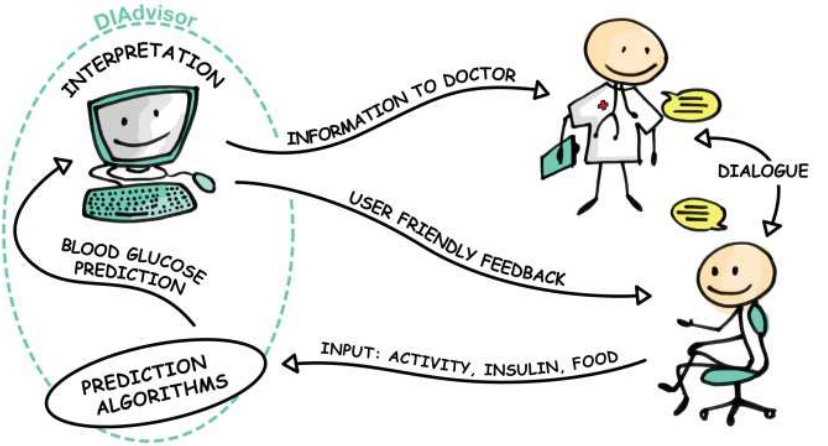

Figure 1.1 DIAdvisor^{TM} modeling and prediction [DIAdvisor, 2011]

and giving them advices about how to adjust their treatments. The DIAdvisor ^{TM} tool would constitute a mobile short-term blood glucose predictor and treatment advisor. Figure 1.1 clarifies the concept. The predictor system would need user input concerning patient characteristics (e.g., insulin sensitivity or resistance), patient condition (e.g., fasting, meal time, rest or physical activity, illness, stress), therapeutic mode (type of insulin delivery route, type of insulin preparations), time and size of meals, inputs from non-invasive glucose sensors, wearable vital signs sensors and blood glucose meter measurements, and is expected to produce short-term blood glucose predictions to be graphically shown to the patients and suggestions to the user from a decision support module. The users would have the opportunity to accept or reject the advice, thereby assuring safety.

1.2 Statement of the problem

The development of a safe predictive and advisory system, such as the DIAdvisor^{TM} tool, would require patient-specific dynamical models of the glucose metabolism able to describe the blood glucose evolution based on the most significant inputs, namely, meal carbohydrates, exhogenously injected insulin and possibly energy expenditure due to physical activity [Spurr *et al.*, 1988] to be used in a model predictive control set-up [Maciejovski, 2002]. Characteristics required in such models include a good trade-off between simplicity and accuracy, stability, qualitative correct responses to inputs and predictive capabilities with particular emphasis on hypo- and hyperglycemia detection. In addition, predictors targeting blood glucose forecasting need to be developed in parallel, as well.

Motivated by the above, this thesis addresses the questions of how to identify personalized models from individual patient data and how to provide individualized short-term blood glucose predictors.

1.3 Contribution of the thesis

- Exploitation of a unique database
- Individual-specific data-based models of the glucose-insulin interaction. Since the final goal motivating the modeling effort was the development of a model-based predictive controller for diabetes management, model characteristics are:
  - stability;
  - a good tradeoff between simplicity and accuracy;
  - the impulse responses of the transfer functions from insulin and meal intake to blood glucose are physiologically sensible;
  - whiteness of the residuals;
- Individual-specific short-term blood glucose predictors.

1.4 Publications

The thesis is based on the following publications:

Cescon, M. (2011): "Adaptive subspace-based prediction of T1DM glycemia." In *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC2011) (Submitted)*. Orlando, FL.

Cescon, M. and R. Johansson (2009): "Glycemic trend prediction using empirical model identification." In *Proc. of the 48th IEEE Conference on Decision and Control (CDC2009)*, pp. 3501–3506. Shanghai, P.R. China.

Cescon, M. and R. Johansson (2010): "Multi-step-ahead multivariate predictors: a comparative analysis." In *Proc. of the 49th IEEE Conference on Decision and Control (CDC2010)*, pp. 2837–2842. Atlanta, USA.

Cescon, M. and R. Johansson (2011): "On data-driven multistep subspace-based linear predictors." In *Proc. of the 18th IFAC World Congress*. Milano, Italy.

Cescon, M., F. Ståhl, R. Johansson, and M. Landin-Olsson (2009a): "Short-term diabetes blood glucose prediction based on blood glucose measurements." In *Proc. of the 2nd International Conference on Advanced Technologies and Treatments for Diabetes (ATTD2009)*, pp. A121. Athens, Greece.

Cescon, M., F. Ståhl, M. Landin-Olsson, and R. Johansson (2009b): "Subspace-based model identification of diabetic blood glucose dynamics." In *Proc. of the 15th Symposium on System Identification (SYSID2009)*, pp. 233–238. Saint-Malo, France.

Ståhl, F., M. Cescon, R. Johansson, and E. Renard (2009): "Infinite horizon prediction of postprandial breakfast plasma glucose excursion." In *Proc. of the 9th Annual Diabetes Technology Meeting (DTM2009)*, pp. A163 San Francisco, CA.

Other relevant publication:

Cescon, M., I. Dressler, R. Johansson, and A. Robertsson (2009): "Subspace-based identification of compliance dynamics of parallel kinematic manipulator." In *2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 1028–1033. Singapore.

2

Background

2.1 Blood glucose regulation

Glucose is the primary substrate for energy in the insulin-independent tissues (IIT), i.e., central nervous system and red blood cells, therefore it is essential for the body that an adequate amount of glucose is always present in the bloodstream. In a healthy subject, blood glucose concentration is maintained within relatively narrow limits at around 90 [mg/dL] by a highly complex neuro-hormonal control system [Williams and Pickup, 1992] that guarantees the balance between glucose entering the bloodstream after liver gluconeogenesis and intestinal absorption following meals and glucose uptake from the peripheral tissues. When the glucose in blood rises to a high concentration typically after a food intake, insulin is secreted by the β-cells of the pancreas, causing most of the glucose absorbed to be stored almost immediately in the liver in the form of glycogen. Then, between meals, when food is not available, the blood glucose concentration begins to fall, causing a rapid decrease in insulin secretion, which in turn stimulates glucagone secretion by the α-cells of the pancreas. Glucagone functions in the opposite direction of insulin, promoting the breakdown of the liver glycogen back to glucose, which is then released into the bloodstream preventing glucose concentration from falling too low. These actions maintain plasma glucose within tight limits, assuring a constant supply of fuel to the IIT.

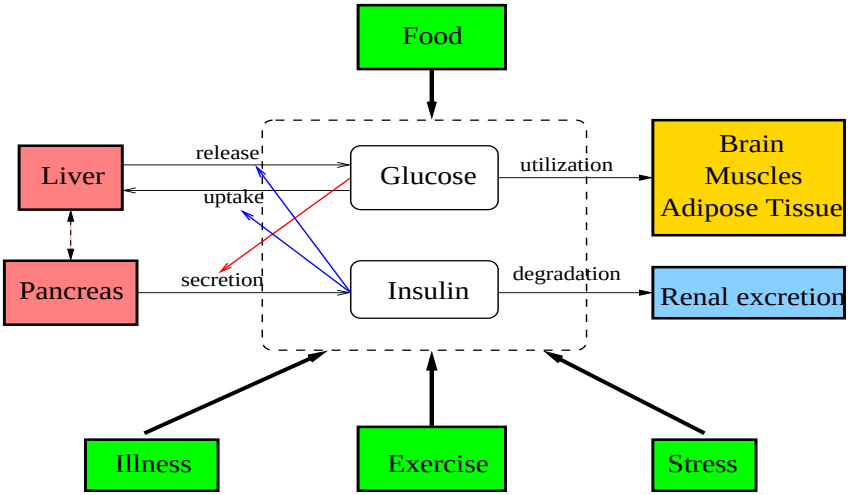

Figure 2.1 Glucose-insulin control system in the non-diabetic subject. 1. Increased blood glucose stimulates insulin secretion from the pancreas; 2. Insulin promotes liver uptake, storage and use of glucose; 3. Glucose is released back from the liver during fasting conditions

The glucose and insulin systems interact by feedback control signals: if a glucose rise occures after a meal, the β-cells secrete more insulin in response to increased plasma glucose concentration and in turn insulin signaling promotes glucose uptake by the liver and utilization by the muscles and various organs, inhibiting glucose production thereby bringing the plasma glucose level back toward a steady-state value. These control interactions are usually referred to as insulin sensitivity and β-cell responsitivity. Figure 2.1 shows the glucose-insulin control system for the non-diabetic subject.

2.2 Diabetes Mellitus

When glucose metabolism is impaired by either lack of insulin secretion or decreased sensitivity of the tissues to insulin, blood glucose concentration starts fluctuating widely, leading to a syndrome called

Diabetes Mellitus. In type I diabetes, also called insulin-dependent diabetes mellitus (IDDM), there is an absolute deficiency of insulin secretion, which is due to β-cell destruction. The basic effect of insulin lack or insulin resistance to glucose metabolism is the prevention of efficient uptake and utilization of glucose by most cells of the body. As a result, blood glucose concentration increases (hyperglycemia), cell utilization of glucose falls and consequently utilization of fat and proteins for energy increases causing release of free fatty acids, cholesterol and phospholipids in the plasma. This has multiple effects throughout the body associated with damage, dysfunction and failure of various organs. Sustained hyperglycemia is associated with acute ketoacidosis, nephropaty, rethinopaty, neuropathy and damages to the cardiovascular system [Williams and Pickup, 1992].

Because insulin deficiency defines T1DM, insulin replacement is the hallmark of therapy. Focusing on tight blood glucose targets, the philosophy of insulin replacement is to mimic with injections the insulin secretion pattern in the non-diabetic person. In the non-diabetic subjects, insulin is secreted into the portal circulation at two rates: a slow basal secretion throughout the 24 hours and an augmented rate at meal times. The basal insulin concentration is sufficient to keep a constant glucose concentration during fasting conditions and the prandial insulin doses should enhance an increased glucose uptake during and after meals. Intensive insulin therapy has been strongly promoted during the last decade, following the results of the major Diabetes Control and Complications Trial (DCCT) [The Diabetes Control and Complications Trial Research Group, 1993] and follow-up Epidemiology of Diabetes Interventions and Complications (EDIC) [The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Study Research Group, 2005] studies in order to keep blood glucose levels as close to normal as possible (i.e., 70-140 [mg/dL]). However, insulin therapy may risk potentially severe induced hypoglycemia, resulting from too high levels of insulin, leading to loss of consciousness, coma and eventually death.

At present, the most commonly used index for assessment of long-term control is provided by a component of the glycated haemoglobin, namely, the HbA_{1c}, which is a measure of the integrated blood glucose control over the 2-3 months preceding the test and is about 4-6% in healthy subjects [Williams and Pickup, 1992].

2.3 Inherent challenges in T1DM modeling

Despite significant efforts devoted to the problem of blood glucose regulation in type 1 diabetic patients over the last several decades (see e.g. [Cobelli *et al.*, 2009] for a comprehensive review), many inherent challenges that must be overcome still remain. At the most basic level, the disease can be viewed as a process having one output, namely, glucose concentration in plasma, and two inputs, namely, meal carbohydrates and administered insulin. The first and perhaps most crucial challenge to overcome in modeling is that of poor data excitation: often the inputs are simultaneous and in the same ratio, the so-called insulin-to-carbohydrate ratio, precluding the possibility of distinguishing their relative effects. It is still an open issue how to strongly excite the system in order to obtain data meaningful for identification purposes preserving at the same time the patients from the risk of serious clinical events. Second, the most widespread way of treating diabetes comprises a series of impulse-like control actions, i.e., insulin injections and food intakes, applied several times during the day at irregular sampling instants, typically at wake up, meal and bed times, the decisions being based on scarce assays of the controlled variable, i.e., blood glucose. This rises the problem of non-uniformly sampled and infrequent data and, since the signals in play interact in the bloodstream, introduces assumptions on the subcutaneous-to-intravenous insulin absorption and gastro-intestinal carbohydrates absorption dynamics.

In the last ten years, advances in sensor technology saw the advent of continuous glucose monitors (CGM), systems capable of measuring glucose concentration frequently (e.g., every 5 minutes) for several days, providing the patient with well-sampled data in real time. However, it is important to stress that together with the benefits, they introduce yet another limitation. Indeed, those devices measure glucose concentration in the interstitium and not in plasma. Interstitial glucose (IG) fluctuations are related to BG presumably via diffusion process [Steil *et al.*, 2005], [Keenan *et al.*, 2009]. This leads to a number of issues, including distortion (which incorporates a time lag) and calibration errors, and necessitates the development of methods for their mitigation. In particular, it is necessary to consider that, since the BG to IG kinetics acts as a low-pass filter, the frequency content of interstitial glucose is different from that of blood glucose [Breton *et al.*,

2008], [Miller and Strange, 2007]. On average, glucose levels in the interstitial fluid lag the glucose levels in capillary blood by 14 minutes. This is a physiological phenomenon that can vary from one person to another.

As for the inputs, when taking into account the appearance of insulin in the bloodstream from subcutaneous delivery and that of glucose in plasma after a meal, new time-lags and dynamics are introduced; further, subcutaneous insulin infusion involves degradation at the site of delivery. In addition, meals must be recorded by the patients, and the actual amount of carbohydrates must be estimated, a process that is prone to errors. Also, in practice, the combination of simple and complex carbohydrates, fats and proteins can affect the glucose absorption in the digestive system. Unrepresented inputs, such as stress and illness constitute another challenge to diabetes modeling. Furthermore, it is a well known fact that physical activity, apart from having a glycemia lowering effect due to utilization of glucose by the muscle cells, enhances insulin sensitivity, playing a substantial role in the picture, but the magnitude and duration of such effects are hard to consider. Another important aspect is the degree of variability of the overall system dynamics over the day (the so-called "dawn phenomenon", for instance, is characterized by increased insulin resistance during the morning hours [Williams and Pickup, 1992]).

A priori knowledge of the diabetes process indicates two fundamental properties that should be satisfied by any model:

- the gain associated with the insulin input should be negative (i.e., an increase in insulin results in a decrease in glucose concentration)

- the gain associated with the meal input should be positive (i.e., a meal results in an increase in glucose concentration)

that a sound and valid model need to show. However, the values of the above mentioned gains are related to age of the subject, disease duration, BMI, insulin sensitivity, β-cells responsitivity and probably many more unknown factors so that it is not clear how to take them into account in the modeling process.

2.4 Literature survey

In diabetes research and therapy, modeling of the glucose-insulin control system has received significant attention for more than 40 years [Cobelli *et al.*, 2009]. Several types of models serving different purposes were proposed, most of these efforts being first-principles based descriptions of the physiological relationships associated with T1DM, and only to a lesser extent mathematical modeling by means of system identification.

Physiological models

The first pioneering work describing the effects on plasma glucose and insulin of an oral or intravenous administration of glucose was that of Bolie [1961], later modified by Ackerman and McGucking [1964] in order to provide a model of the glucose metabolism during an oral glucose tolerance test (OGTT). Greater attention was received by the so-called minimal model [Bergman *et al.*, 1979], [Bergman *et al.*, 1981], developed for the specific purpose of quantifying pancreatic responsiveness and insulin sensitivity during an intravenous glucose-tolerance test (IVGTT) in non-diabetic individuals. The model consists of three differential equations describing plasma glucose and plasma insulin in a remote compartment, accounting for neither the dynamics of subcutaneous insulin infusion nor the dynamics of gut glucose absorption from a carbohydrate meal. Glucose-insulin pharmacokinetics/pharmacodynamics in non-diabetic subjects was described by a 19-state model developed by Sorensen [1985], the major shortcoming of this model being the failure in capturing the hyperglycemic events characteristic of type 1 diabetes [Lynch and Bequette, 2002]. A good trade-off between simplicity and accuracy is represented by the model in [Hovorka *et al.*, 2004]. The model inputs are the rate of subcutaneously infused fast acting insulin, meal carbohydrate amount and time of ingestion and its outputs are plasma glucose and insulin concentrations. The model was later revisited, in that the absorption kinetics associated with subcutaneous insulin delivery was modified [Wilinska *et al.*, 2005], the model replacing the original subcutaneous insulin subsystem consisting in two parallel fast and slow channels for insulin absorption as well as a degree of insulin degradation at the injection site. The state-of-the-art meal simulation model of the glucose-insulin system for non-diabetics,

FDA-approved to substitute animal trials proposed by Dalla Man and co-workers [2007] used a sophisticated triple tracer method to estimate important meal-related quantities such as the rates of appearance of glucose in the blood from the meal, endogenous glucose production, utilization of glucose, and insulin secretion. Reviews of physiological diabetes models include that of Nucci and Cobelli [2000], who specifically examined several models of subcutaneous-to-intravenous insulin kinetics, Makroglou and co-workers [2006] presenting an overview of existing software packages specific to diabetes modeling and finally Cobelli and co-workers [2009], discussing the main contribution to both modeling and control in diabetes from the early 1960s.

Predictive models

Although seemingly simple in concept, the problem of glucose prediction in an active individual has to date proved intractable. Currently, continuous glucose monitoring (CGM) devices are the available technology able to provide high/low glucose alarms when certain preset threshold levels have been crossed and to deliver early-warnings of events that are likely to occur if the current trend continues. To date many studies have investigated the possibility of predicting blood glucose concentration for the purpose of regulating glucose intervention, most of this research being based on data generated by a simulation model (e.g., [Hovorka *et al.*, 2004], [Roy and Parker, 2006], [Dalla Man *et al.*, 2007], [Percival *et al.*, 2008]). Originally developed by Bremer and Gough [1999] the idea of T1DM CGM time-series analysis has been further pursued by Sparacino et al. [2007] and Gani et al. [2009] to predict near-future glucose concentration from its past history. However, the limited accuracy and the lack of exploitation of the dynamic interplay between previously injected insulin, meal intake and eventually exercise reduce or even eliminate the clinical benefits of the approach. Bremer and Gough [1999] used 10-min data from ambulatory T1DM patients to identify autoregressive (AR) models. They explored 10, 20 and 30 minutes prediction horizons, and report that the 10-min predictions are accurate and that for certain data, 20-min or 30-min predictions may also be acceptable for a limited set of data only. They provided no quantification of the accuracy of the model predictions. Hovorka and his group [2004] performed experiments with ten T1DM patients under clinical conditions, using their physiological model to

make predictions of 15-min glucose data up to 60 minutes into the future. The glucose was measured intravenously, but delayed by 30 min to mimic subcutaneous measurement. The model parameters were recursively estimated using a Bayesian method. Sparacino et al. [2007] collected 48 h of continuous (3-min) glucose data from 28 T1DM subjects. The data in their study were for ambulatory conditions, but were filtered to remove noise spikes. In their retrospective analysis, they recursively identified simple polynomial and AR models from these time-series data. They investigated prediction horizons of 10 and 15 steps (i.e., 30 and 45 min).
In [Reifman *et al.*, 2007] autoregressive models were identified from 5-days long glucose time-series belonging to a population of 5 in-hospital subjects. Performances were evaluated on 30- and 60 minutes prediction horizons. Eren-Oruklu et al. [2009] proposed a recursive AR and ARMA model identification strategy with an adjustable forgetting factor for healthy and type-2 diabetics. An ARX model with a nonlinear forgetting factor scaled according to the glucose range was considered in [Castillo-Estrada *et al.,* 2010a], [Castillo-Estrada *et al.,* 2010b], and a 45-minute prediction horizon showed good results. Finally, in [Gani *et al.*, 2010] it was asserted that a universal data-driven model identified from a CGMS time-series of a patient could be used to make near-future glucose concentration predictions for other patients without any model customization procedure. They used regularization techniques to filter data from 34 subjects, then, using the filtered data, they develop auto-regressive models of order 30 to the purpose of making short-term, 30-min-ahead predictions.

3

Experimental Conditions and Clinical Data Acquisition

In the framework of DIAdvisor^{TM} [DIAdvisor, 2011], a European FP7-IST research project on diabetic blood-glucose prediction and improved blood-glucose management in patients, acquisition of bioclinical data linked or potentially involved in blood glucose control from insulin-treated diabetic patients was accomplished in a series of experiment sessions. The investigations focused on a population of basal-bolus regimen treated subjects, either as combination of multiple daily insulin injections (MDI) including long-acting and fast-acting analogues or as continuous subcutaneous insulin infusion (CSII) of fast-acting analogues from a pump. The clinical study was performed in the Clinical Investigation Center CIC-CHU in Montpellier, France. Collected data include: specific patient parameters (e.g., gender = male, age = 43 years old, BMI = 23.7, weight = 67 kg), characteristics related to diabetes (e.g., disease duration = 10 years, insulin delivery = external pump), associated health conditions and therapies, food intakes and administered insulin doses registered in a logbook, capillary glucose strips, interstitial glucose levels, plasma glucose and plasma insulin concentration from drawn blood samples as well as vital signs. Table 3.2 lists the data collected and their description.

3.1 Equipment

For the whole duration of the study, the subjects were equipped with state-of-the-art devices provided by the DIAdvisor^{TM} Consortium complying with the study protocol as explained in the following subsections.

HemoCue^{TM} Glucose Analyzer

The HemoCue^{TM} Glucose Analyzer [Hemocue, 2011] is a blood glucose meter based on a glucose dehydrogenase method and consisting of a pocket size handheld analyzer and a unique disposable microcuvette. The analyzer was factory calibrated and no calibration is needed between cuvette batches. The device was used by each patient as a reference glucose meter, assessing plasma glucose levels from finger-stick samples.

Abbott Freestyle Navigator^{TM}

The Abbott Freestyle Navigator^{TM} [Abbott, 2011] is a Continuous Glucose Monitoring System (CGMS) consisting of an amperometric electrochemical sensor placed under the skin, a wireless transmitter connected to the sensor and a wireless receiver collecting the sensor signals. The subcutaneous sensor was inserted about 5 mm into the subcutaneous tissues and could be worn for up to five days before replacement. Calibration against capillary glucose was required at specific times, namely, 10, 12, 24 and 72 hours after a 10-hour initialization period. Using the WIRED ENZYME^{TM} technology, the sensor converted glucose concentration to electrical current. Once every minute the transmitter sent the estimate of the inter-

sitial glucose concentration to the receiver, which displayed the final values once every 10 minutes.

VivoMetrics Clinical LifeShirt®

The VivoMetrics Clinical LifeShirt® [Grossman, 2004] is non-invasive and fully-integrated physiologic data monitoring system consisting of a light-weight shirt with embedded vital signs sensors and a receiver/ transmitter. In this trial heart rate, respiration rate, skin temperature and body movements were recorded. To measure pulmonary function, sensors were worn around the chest and the abdomen. A three-lead single-channel ECG measured the heart rate and a three-axis accelerometer recorded the patient posture and activity level.

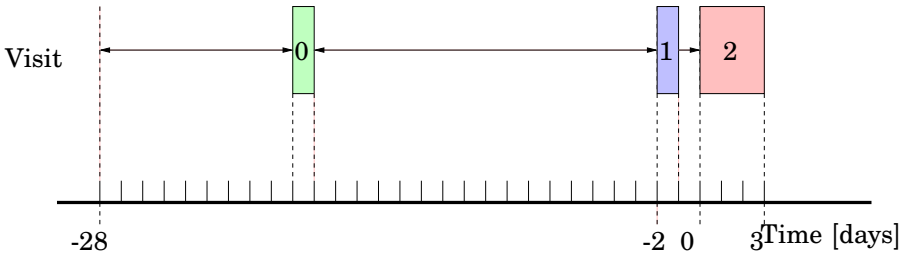

Figure 3.1 DIAdvisor^{TM} Data Acquisition Trial. *visit 0* Screening visit (green box), *visit 1* sensor initialization visit (blue box), *visit 2* in-hospital tests (red box)

3.2 Study protocol and experiments

The clinical study consisted of three visits: *visit 0* for patients screening, *visit 1* for sensors initialization and *visit 2* for 75-hours in-hospital tests. Figure 3.1 gives the flow chart of the trial.

Visit 0

Prior to any study-related procedure, the purpose of this visit was to perform a screening examination of the patient, the outcome of it being recorded in the clinician's sheet.

Visit 1

Within 4 weeks of *visit 0* the patient was admitted at the clinic to initialize the Abbott Freestyle Navigator^{TM} device. The sensor was inserted subcutaneously into the patient's skin and calibrated against capillary glucose by a nurse. Although glucose sensing is available after a 10-hours initialization period following sensor calibration, optimal sensor signal stability and accuracy is obtained after 24 hours. For this reason, the subject was hospitalized at best 48 hours after sensor insertion, so to begin the tests with a well-calibrated device.

Visit 2

During the whole 3-days-long visit, the patient was permanently equipped with the Abbott Freestyle Navigator^{TM} [Abbott, 2011] and the

VivoMetrics Clinical LifeShirt® [Grossman, 2004] devices. Standard meals for breakfast (8:00 am), lunch (1:00 pm) and dinner (7:00pm) were served, the amount of administered carbohydrates being 42, 70 and 70 grams, respectively. Blood samples were collected by nurses to measure plasma glucose and plasma insulin concentrations: every hour during day, every 2 hours during night, every 15 minutes after meals for 2 hours. A specific sampling scheduled was adopted after breakfasts: 30 min before, mealtime, 10, 20, 30, 60, 90, 120, 150, 180, 240, 300 min after, for a total of 37 blood samples per day. No specific intervention on usual diabetes treatment was scheduled during the period. The patients decided their insulin needs according to the HemoCue^{TM} Glucose Analyzer [Hemocue, 2011] measurements as usually in activities of daily life.

Table 3.1 Patient info

Patient ID	Insulin Therapy	HBA1c [%]	BMI
102	MDI	6.5	26.5
103	CSII	9.1	23.7
104	MDI	7.6	20
105	CSII	7.8	24.1
106	CSII	7.8	21.2
107	CSII	8.9	25.3
115	MDI	8.5	19.7
120	MDI	9	22.4
130	MDI	8.8	29.4

3.3 Patients selection criteria

A total of 30 diabetic subjects, male and female adults, were included in the study to allow availability of data from a wide spectrum of patients under basal-bolus regimen. Among these, a population of 9

patients was chosen, the selection criteria being the quantity of data collected ($>$ 80% of the expected), no sensor failures and logbook correctly filled in. Exclusion criteria were data partially collected due to malfunctioning of the devices, laboratory results not available and/or patients not observant in annotating insulin/meal intakes. Table 3.1 reports the characteristics of the selected patients. Figure 3.2 presents

Table 3.2 List of collected data samples and description

Data Description	Units	Device
Patient parameters	-	Clinician's sheet
Insulin Therapy	-	Clinician's sheet
Interstitial glucose	[mg/dL]	Navigator^{TM}
Intermediate signal	[a.u.]	Navigator^{TM}
Carbohydrates	[g]	Patient's logbook
Lipids	[g]	Patient's logbook
Proteins	[g]	Patient's logbook
Fast-acting insulin	[IU]	Patient's logbook
Slow-acting insulin	[IU]	Patient's logbook
Hyperglycemia-correction	[IU]	Patient's logbook
Plasma glucose	[mg/dL]	Blood sample
Capillary glucose	[mg/dL]	HemoCue^{TM}
Basal plasma insulin	[mIU/L]	Blood sample
Bolus plasma insulin	[mIU/L]	Blood sample
Total plasma insulin	[mIU/L]	Blood sample
Heart rate	[beats/min]	LifeShirt®
Respiration rate	[breaths/min]	LifeShirt®
Activity level	[a.u.]	LifeShirt®
Skin temperature	[°C]	LifeShirt®

[a.u.] arbitrary units

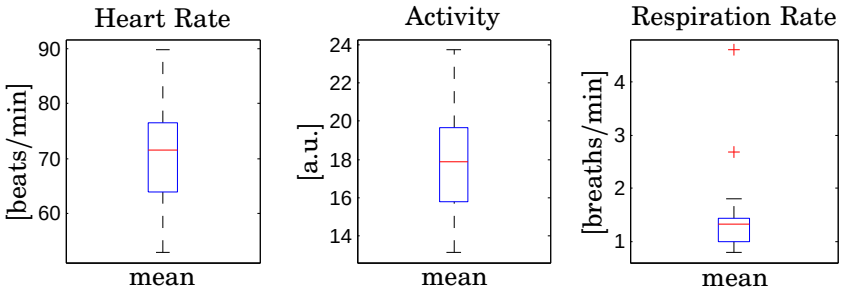

Figure 3.2 Vital signs mean values: *Left* Heart Rate [beats/min], *Center* Activity Level [a.u.], *Right* Respiration rate [breaths/min]. Each box presents results over the considered population. The central mark is the median, the edges of the box are the 25th and 75th percentiles.

boxplots of the vital signs mean value over the selected population. Figures 3.5 and 3.6 display the recorded data for a representative CSII pump patient, whereas Figs. 3.3 and 3.4 show a representative MDI subject records.

Throughout the thesis, we will present results for these two representative subjects, referring the reader to the appendix for the population study.

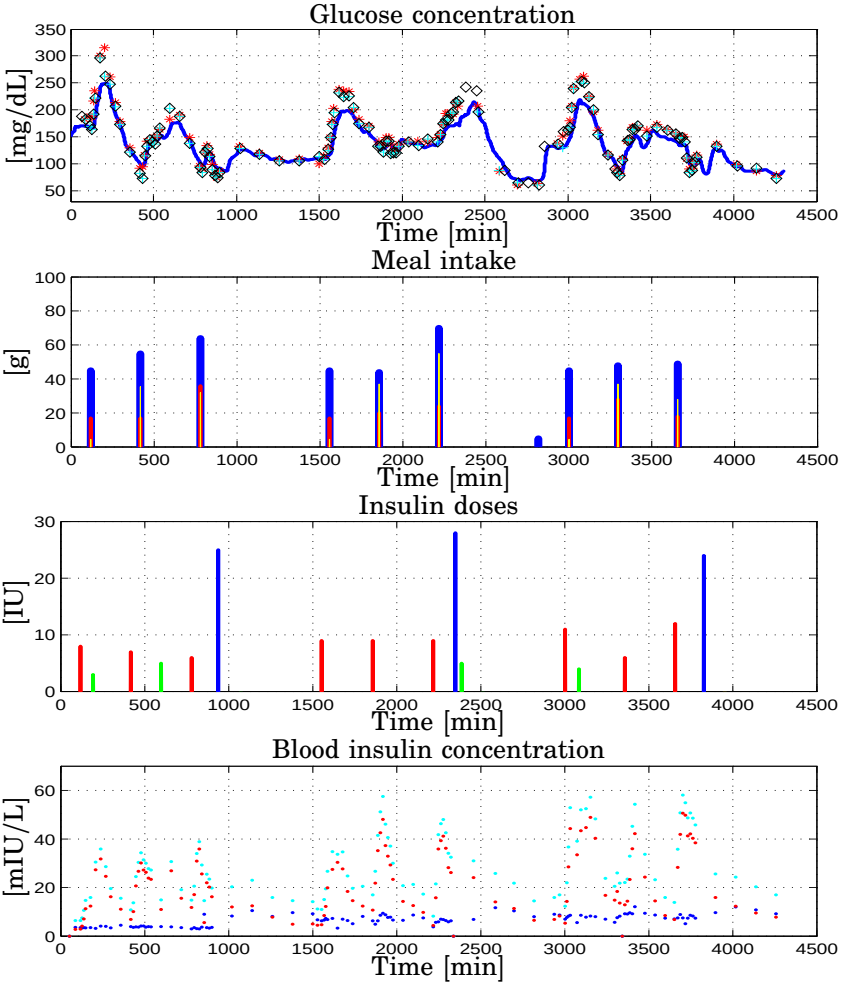

Figure 3.3 Patient CHU0102 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

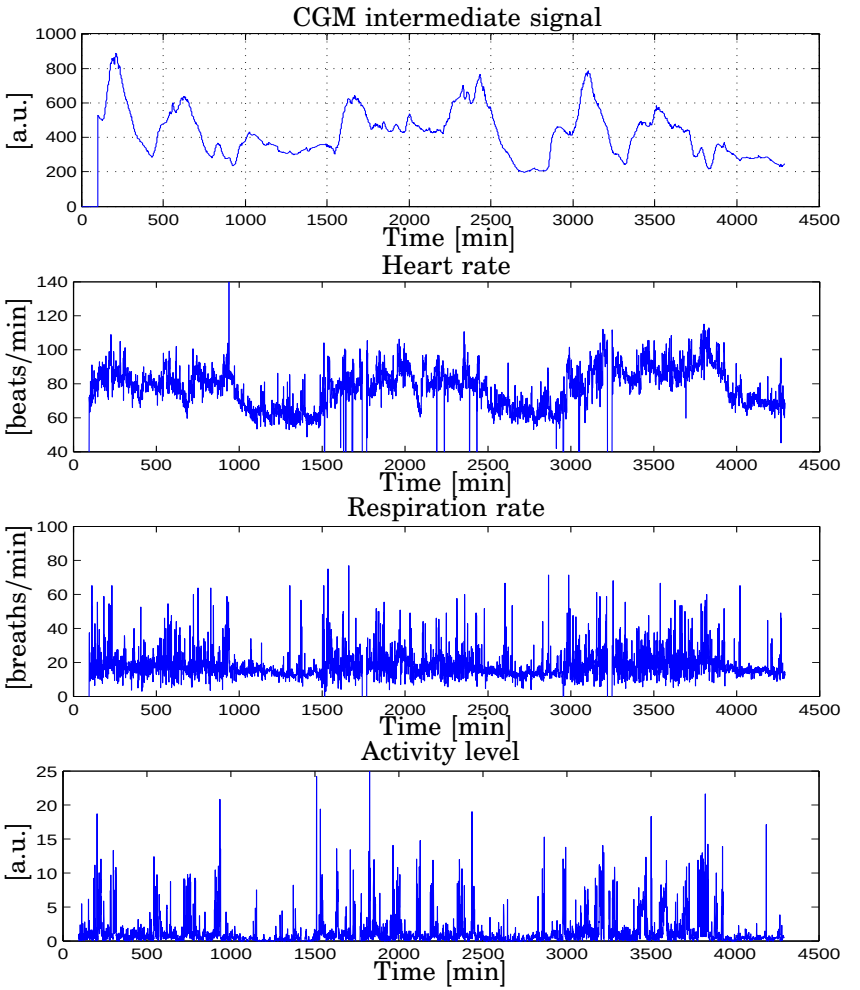

Figure 3.4 Patient CHU0102 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [beats/min]; *Lower Center* Respiration rate [breaths/min]; *Bottom* Activity level [a.u.]

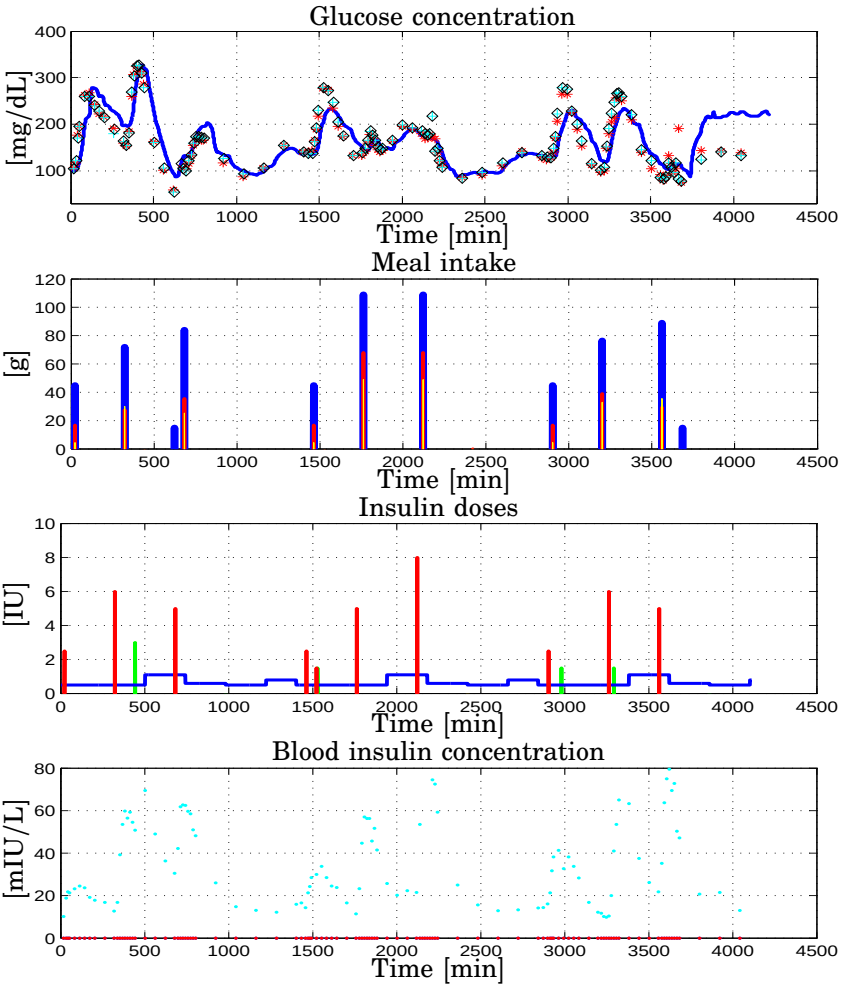

Figure 3.5 Patient CHU0107 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

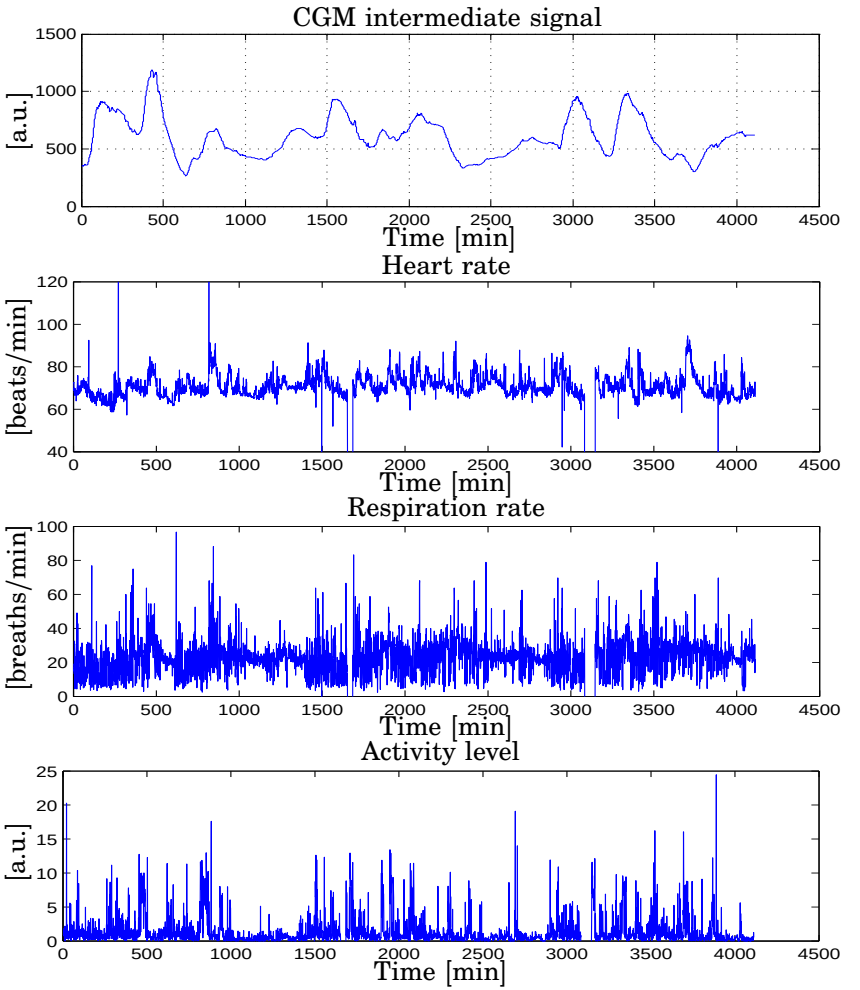

Figure 3.6 Patient CHU0107 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [beats/min]; *Lower Center* Respiration rate [breaths/min]; *Bottom* Activity level [a.u.]

4

Modeling of the Gluco-regulatory System

As already mentioned, the physiology of glucose metabolism in diabetes can be thought of as having one output, i.e., glucose level in the bloodstream, and two main inputs, i.e., carbohydrate intake and administered insulin [Dalla Man *et al.*, 2006]. Further, given that physical activity has been proven to decrease plasma glucose levels due to increased glucose uptake by the exercising muscles [Williams and Pickup, 1992], the effect of exercise, i.e., increased heart rate, respiration rate and body movements, is therefore to be regarded as additional input.

Based on current knowledge of the overall physiological model, the following subsystems have to be considered (Fig. 4.1):

- the glucose subsystem (GS), describing glucose kinetics after intestinal absorption following a food intake;
- the insulin subsystem (IS), accounting for the pharmacokinetics of the exogenously administered insulin;
- the physical activity and energy expenditure subsystem (EES), measuring the rate of physical activity intensity;
- the glucose-insulin interaction (GII).

This chapter will be concerned with data-driven modeling of the GII subsystem, exploiting compartment models from the literature to de-

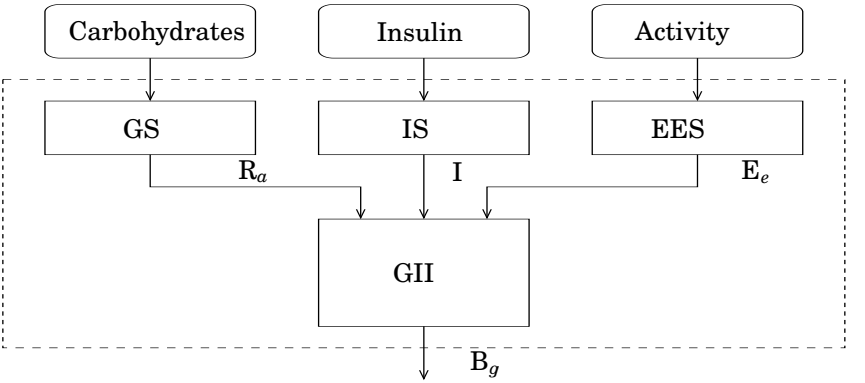

Figure 4.1 Overview of the physiological model describing diabetic blood glucose dynamics

scribe the GS and IS subsystems, and simply using the squared heart rate and accelerometer data to account for the energy expenditure.

4.1 Input modeling

Glucose Intestinal Absorption Modeling

Glucose transit through the stomach and upper small intestine was described by a nonlinear chain of three compartments (Fig. 4.2), where the first two compartments represent the stomach (solid and liquid phase) and the third one depicts the intestine. Model equations are [Dalla Man *et al.*, 2006] :

$$\begin{aligned}
\dot{q}_{sto1}(t) &= -k_{21} \cdot q_{sto1}(t) + D \cdot \delta(t), & q_{sto1}(0) &= 0 \\
\dot{q}_{sto2}(t) &= -k_{empt} \cdot q_{sto2}(t) + k_{21} \cdot q_{sto1}(t), & q_{sto2}(0) &= 0 \\
\dot{q}_{gut}(t) &= -k_{abs} \cdot q_{gut}(t) + k_{empt} \cdot q_{sto2}(t), & q_{gut}(0) &= 0 \\
R_a(t) &= \frac{f \cdot k_{abs} \cdot q_{gut}(t)}{m_b}, & R_a(0) &= 0
\end{aligned} \tag{4.1}$$

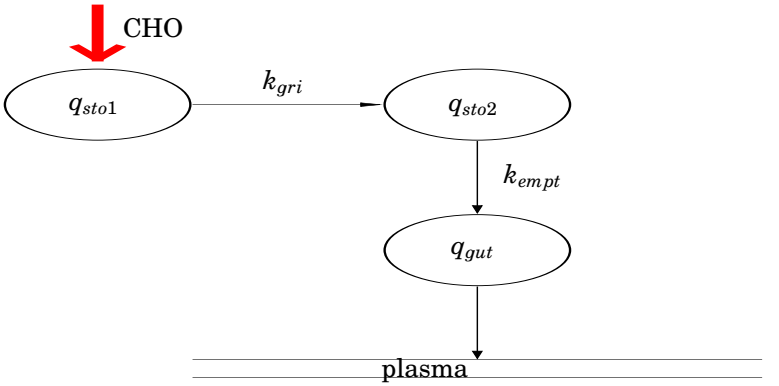

Figure 4.2 Glucose intestinal absorption model. The ingested carbohydrates (CHO) transit through the stomach represented by compartments q_{sto1} and q_{sto2} and the intestine represented by q_{gut} before reaching the plasma circulation

where q_{sto1} $[mg]$ and q_{sto2} $[mg]$ are the amounts of carbohydrates in the stomach (solid and liquid phase, respectively), D $[mg]$ is the amount of ingested carbohydrates, q_{gut} $[mg]$ is the carbohydrate mass in the intestine, k_{21} is the rate of grinding, k_{empt} the rate of gastric emptying, m_b $[kg]$ the subject's body weight, k_{abs} the rate of absorption and f the fraction of intestinal absorption that actually appears in plasma. The rate of gastric emptying was a non-linear function of the amount of carbohydrates in the stomach q_{sto} according to the following relationship:

$$k_{empt}(q_{sto}) = k_{min} + \frac{k_{max} - k_{min}}{2} \cdot \\ \cdot \{\tanh[\alpha(q_{sto} - b \cdot D)] - \tanh[\beta((q_{sto} - c \cdot D)] + 2\} \tag{4.2}$$

with

$$\begin{aligned} q_{sto}(t) &= q_{sto1}(t) + q_{sto2}(t) \\ \alpha &= \frac{5}{2 \cdot D \cdot (1-b)} \\ \beta &= \frac{5}{2 \cdot D \cdot c} \end{aligned} \tag{4.3}$$

Table 4.1 reports mean population values for the parameters appearing in 4.1, 4.2, 4.3.

Table 4.1 Parameter values in the glucose intestinal absorption modeling

Parameter	Value	Measurement Unit
k_{max}	0.0558	[min^{-1}]
k_{min}	0.0080	[min^{-1}]
k_{abs}	0.0568	[min^{-1}]
b	0.82	dimensionless
d	0.01	dimensionless
f	0.9	dimensionless

Insulin Kinetics

The insulin flow $s(t)$ from the subcutaneous compartments enters the bloodstream and is degraded in the liver and in the periphery (Fig. 4.3) according to the following model equations:

$$\begin{aligned} \dot{I}_p(t) &= -(m_2 + m_4)I_p(t) + m_1 I_l(t) + s(t), \qquad I_p(0) = I_{pb} \\ \dot{I}_l(t) &= -(m_1 + m_3)I_l(t) + m_2 I_p(t), \qquad I_l(0) = I_{lb} \\ I(t) &= \frac{I_p(t)}{V_i} \end{aligned} \tag{4.4}$$

where I_p and I_l are insulin masses in plasma and liver, respectively, I accounts for the total plasma insulin concentration, V_i [L/kg] the distribution volume of insulin, the suffix b standing for the basal states, m_2, m_3, m_4 [min^{-1}] are rate parameters:

$$\begin{aligned} m_2 &= \frac{3}{5}\frac{I_{CL}}{H_{E_b}(V_i m_b)} \\ m_3 &= m_1 \frac{H_{E_b}}{1 - H_{E_b}} \\ m_4 &= \frac{2}{5}\frac{I_{CL}}{V_i m_b} \end{aligned} \tag{4.5}$$

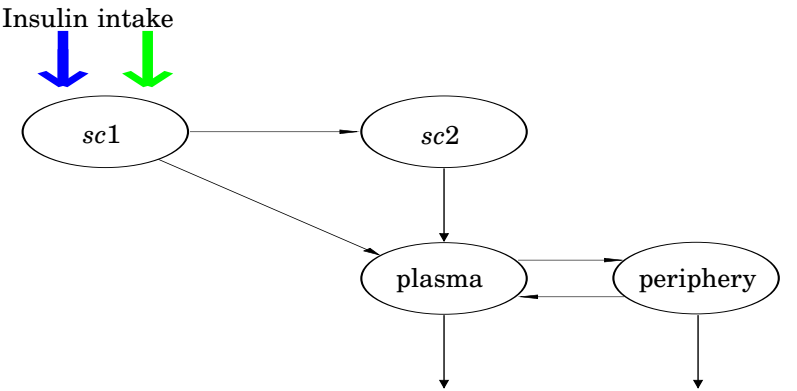

Figure 4.3 Insulin kinetics model. It accounts for both slow- and fast-acting insulin. Compartments $sc1$ and $sc2$ represents the subcutaneous insulin infusion module.

where H_{E_b} [dimensionless] is the basal hepatic insulin extraction, while I_{CL} [L/min] is the insulin clearance. The subcutaneous insulin infusion model in the diabetic subject is represented by two compartments:

$$\begin{aligned} \dot{I}_{sc1}(t) &= -(k_d + k_{a_1})I_{sc1}(t) + I_{ir}(t), & I_{sc1}(0) &= I_{sc1ss} \\ \dot{I}_{sc2}(t) &= k_d I_1(t) - k_{a_2} I_{sc2}(t), & I_{sc2}(0) &= I_{sc2ss} \\ s(t) &= k_{a1} I_1(t) + k_{a_2} I_{sc2}(t) \end{aligned} \tag{4.6}$$

with I_{sc1}, I_{sc2} [pmol/kg] the amount of nonmonomeric and monomeric insulin in the subcutaneous space, respectively, k_d [min^{-1}] the rate constant of insulin dissociation, k_{a_1} [min^{-1}] and k_{a_2} [min^{-1}] the rate constants of nonmonomeric and monomeric insulin absorption, respectively (different model parameters may account for different insulin analogues), and I_{ir} [pmol/kg/min] is the exogenous insulin infusion rate for a CSII patient, or the fast- and slow-acting insulin boluses in an MDI subject. The basal value of insulin in the subcutaneous com-

partments, i.e., I_{sc1ss} and I_{sc2ss} depend on the basal I_{ir_b} according to:

$$\begin{aligned} I_{sc1ss} &= \frac{I_{ir_b}}{k_d + k_{a1}} \\ I_{sc2ss} &= \frac{k_d}{k_{a2}} \cdot I_{sc1ss} \end{aligned} \tag{4.7}$$

Model parameters are given in Table 4.2.
Figures 4.4 and 4.5 show the filtered inputs for the representative patients.

A mismatch between actual total blood insulin concentration and that obtained with the above mentioned model was noticed and will be discussed in detail in Chapter 6.

Table 4.2 Parameter values for the subcutaneous insulin infusion model

Parameter	Fast insulin	Slow insulin	Unit
k_{a1}	0.004	0.0002	$[\text{min}^{-1}]$
k_{a2}	0.0182	0.00091	$[\text{min}^{-1}]$
k_d	0.0164	0.00164	$[\text{min}^{-1}]$
m_1	0.1766	0.1766	$[\text{min}^{-1}]$
V_i	0.05	0.05	[L/kg]
CL	1.1069	1.1069	$[\text{min}^{-1}]$
HE_b	0.6	0.6	dimensionless

4.2 Data analysis

Data analysis was performed in the following order [Johansson, 1993]:

- autospectrum of inputs

$$S_{uu}(i\omega) = \mathcal{F}\{\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} u(t)u^*(t-\tau)dt\}$$

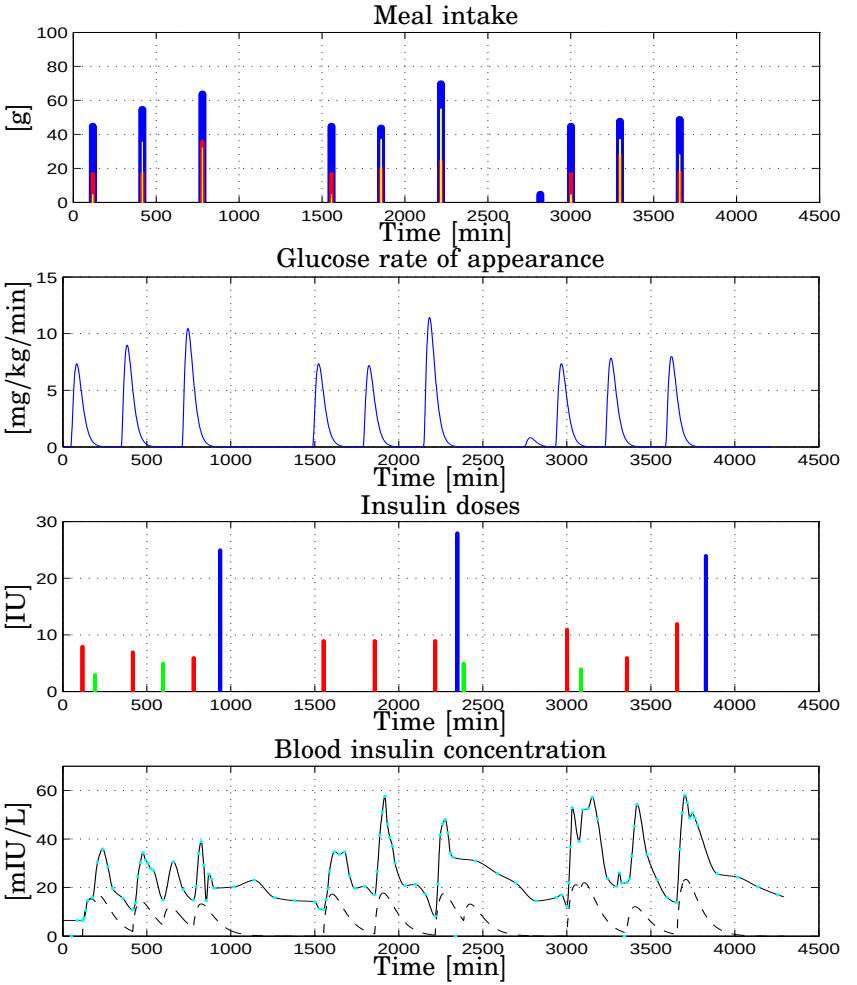

Figure 4.4 Patient CHU0102 data vs. Time [min]. *Top* Meal intake: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal; *Lower Center* Insulin doses: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

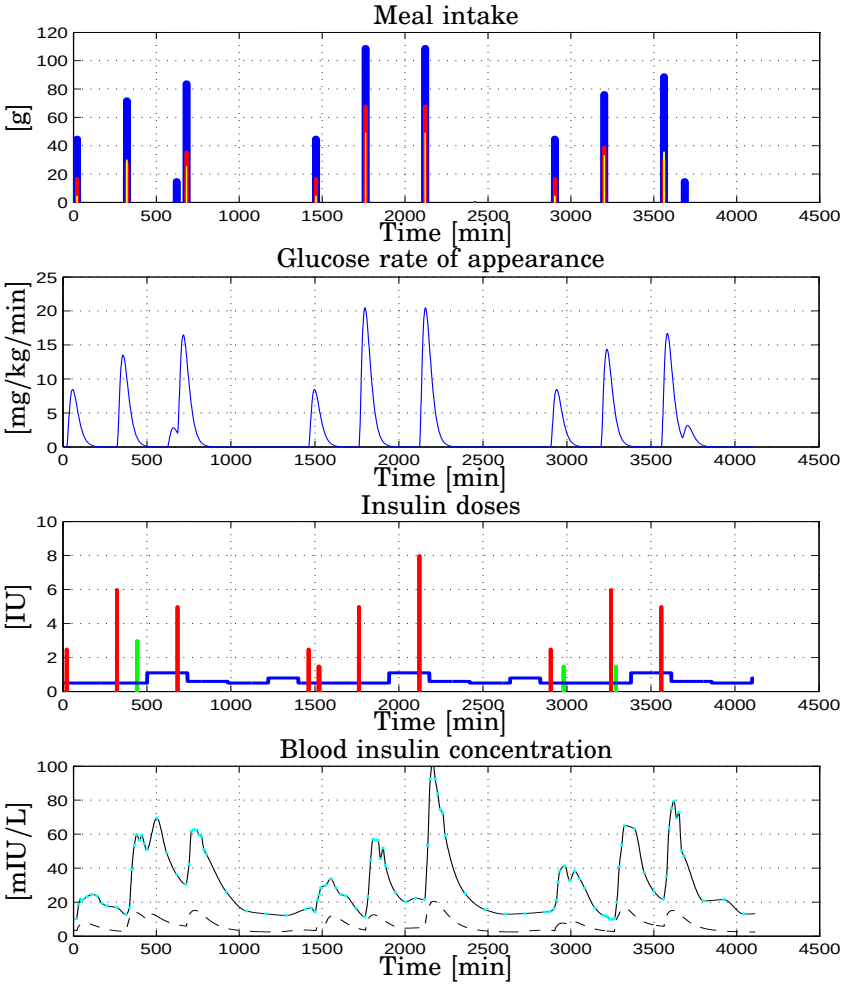

Figure 4.5 Patient CHU0107 data vs. Time [min]. *Top* Meal intake: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal; *Lower Center* Insulin doses: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

- cross spectrum between inputs and output

$$S_{uy}(i\omega) = \mathcal{F}\{\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} u(t)y^*(t-\tau)dt\}$$

- quadratic coherence spectrum beween inputs and output

$$\gamma^2_{uy}(\omega) = \frac{\mid S_{uy}(i\omega) \mid^2}{S_{uu}(i\omega)S_{yy}(i\omega)}$$

Data pre-processing

For purposes of model identification, removal of the mean value of the data series was done as part of standard data pre-processing [Ljung, 1999]. In addition, originally non-uniformly sampled, plasma glucose concentration and plasma total insulin concentration from laboratory results were linearly interpolated and uniformly resampled, the resampling period being 1 minute.

4.3 Problem formulation

Given the inputs:

- interpolated total plasma insulin concentration I [mIU/L] from drawn blood samples;
- plasma glucose rate of appearance R_a [mg/kg/min] after carbohydrate intestinal absorption;
- squared heart rate H_r [(beats/min)^2];
- squared acceleration a [a.u.];

and the output:

- interpolated blood glucose B_g [mg/dL] from drawn blood samples

the objective was to find an individual-specific and physiological relevant model of the glucose-insulin interaction for each of the subjects in the selected population.
Minimum requirements on the model were:

- stability;
- white residuals;
- qualitative correct blood glucose responses to
  - 1 [IU] fast-acting insulin;
  - 10 [g] carbohydrates;

Additional requirement on the model were:

- FIT $\geq$ 50% on 60-minutes-ahead model-based prediction on validation data;
- VAF $\geq$ 50% on 60-minutes-ahead model-based prediction on validation data.

4.4 Model estimation

The approach considered for modeling was system identification of discrete-time, time invariant linear models [Ljung, 1999]. First, models from plasma glucose rate of appearance R_a and total plasma insulin I to blood glucose B_g were estimated, next, heart rate H_r and acceleration a were added as additional inputs to evaluate their impact on blood glucose dynamics (see Fig. 4.6). The data belonging to each of the selected patients records was equally divided into two parts: the first one for the calibration procedure of obtaining the optimal model structure and model parameters, and the second one for validation of the chosen configuration.

Assuming that input-output data $\{u_k, y_k\}$, $k = 1, ..., N$ were available, model structures describing the GII dynamical system were:

- autoregressive moving average with exogenous inputs (ARMAX) model

$$A(z^{-1})y_k = z^{-k_1}B_1(z^{-1})u_k^1 + \cdots + z^{-k_m}B_m(z^{-1})u_k^m + C(z^{-1})w_k \quad (4.8)$$

where

$$A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a} \quad (4.9)$$

$$B_i(z^{-1}) = b_{0,i} + b_{1,i} z^{-1} + \cdots + b_{n_{bm,i}} z^{-n_{bm}} \quad (4.10)$$

$$C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_{n_c} z^{-n_c} \quad (4.11)$$

n_a, n_{bi}, n_c integers representing the orders of the polynomials, $u_k^1, \cdots, u_k^m$ are the inputs, y_k is the output, w_k denotes the coloured noise and z^{-1} is the backward shift operator;

- state-space model in innovation form

$$\begin{cases} x_{k+1} = \mathcal{A}x_k + \mathcal{B}u_k + \mathcal{K}e_k \\ y_k = \mathcal{C}x_k + \mathcal{D}u_k + e_k \end{cases} \tag{4.12}$$

denoting with n the dimension of the state-space, m the number of inputs, $\mathcal{A} \in \mathbb{R}^{n\times n}$, $\mathcal{B} \in \mathbb{R}^{n\times m}$, $\mathcal{C} \in \mathbb{R}^{1\times n}$, $\mathcal{D} \in \mathbb{R}^{1\times m}$, $\mathcal{K} \in \mathbb{R}^{n\times 1}$ and $\{e_k\}$ the noise process.

For each of the calibration dataset different methods were used for the estimation of the model parameters:

- prediction-error identification methods (PEM) [Ljung, 1999] for identification of the ARMAX structure
- subspace-based methods, namely:
  - N4SID [Van Overschee and De Moor, 1994];
  - PO-MOESP [Verhaegen, 1994];
  - PBSID [Chiuso, 2005], [Chiuso, 2007];

  for the identification of the state-space model.

Regularization

Following the approach presented in [Gani *et al.*, 2009], a smoothness constraint was imposed on the least-squares solution of the estimated autoregressive coefficients in the ARMAX model according to:

$$\hat{\theta} = \arg \min_{\theta} || \, y_{k+\tau|k} - \phi_k^{\mathsf{T}}\theta \, ||^2 + \rho^2 \, || \, \Delta\theta \, ||^2 \tag{4.13}$$

where Δ is a block-diagonal differential operator chosen to impose smoothness on the AR coefficients and ρ is the regularization coefficient.

The identification procedure is outlined in Algorithm 1. Throughout the work Matlab® System Identification Toolbox [MathWorks, 2011] and the SMI Toolbox [Haverkamp and Verhaegen, 1997] were used.

Algorithm 1 *Procedure for system identification*

ARMAX:

- for $1 \leq n_a \leq 10$, $1 \leq n_{bm,i} \leq 3$, $1 \leq n_c \leq 10$, $1 \leq n_{ki} \leq 3$
  - estimate model
  - compute Akaike Final Prediction Error
- rank the models according to their FPE in increasing order [Ljung, 1999]

N4SID:

- choose the method for the estimation of the state-space
  - CVA [Larimore, 1990]
  - MOESP [Verhaegen, 1994]
- set the past horizon $p = 120$ and the future horizon $f = 60$
- set the model order $1 \leq n \leq 10$
- select the model order $\bar{n}$
- estimate models with order $\bar{n} - 1$, $\bar{n}$, $\bar{n} + 1$

MOESP:

- set the past and future horizons $s = 30$
- select the model order $\bar{n}$
- estimate models with order $\bar{n} - 1$, $\bar{n}$, $\bar{n} + 1$

PBSID:

- set the past horizon $p = 60$, and the future horizon $f = 20$
- select the model order $\bar{n}$
- estimate models with order $\bar{n} - 1$, $\bar{n}$, $\bar{n} + 1$

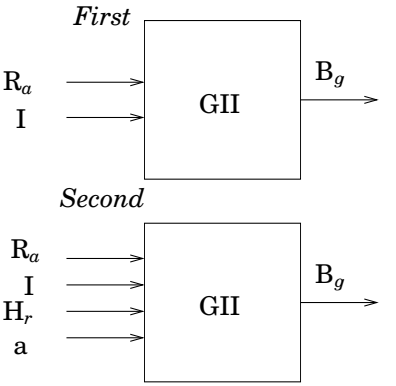

Figure 4.6 Model estimation procedure: *First* from R_a and I to B_g; *Second* from R_a, I, H_r and a to B_g

Model Evaluation and Selection Criteria

The system identification procedure provided a plethora of models for each of the subjects in the population. However, to the purpose of model-based controller design, it suffices to select one model per patient only. Necessary requirements on a model suitable for inclusion and exploitation in the DIAdvisor^{TM} tool were:

- stability;
- white residuals;
- physiologically sensible responses to insulin and food intake, i.e., blood glucose concentration should decrease in response to insulin and increase in response to food intake.

In particular, in order to assess whether the model showed correct responses to inputs, the simulated blood glucose reactions to a

- 1 [IU] fast insulin injection
- 10 [g] carbohydrates intake

were compared, the appearance of insulin in blood after subcutaneous injection being obtained with Eqs. (4.4), (4.6), (4.7) and the parameters

listed in Table 4.2, and similarly, the glucose rate of appearance in plasma after an oral glucose ingestion being calculated using Eqs. (4.1) and parameters in Table 4.1.

Each of the estimated models, i.e. the ARMAX models ranked in ascending FPE, the N4SID, MOESP and PBSID models ranked in ascending model order, were evaluated according to the diagram in Fig.4.7. When a requirement was not fulfilled, a model with higher FPE in the case of ARMAX models or higher order in the case of the state-space models, was taken for evaluation.

Those models passing the tests depicted in Fig. 4.7 were compared on the basis of their prediction performances on 30, 60, 90, 120 minutes ahead prediction. In particular, the performances of model-based predictors obtained with the Matlab® System Identification Toolbox [MathWorks, 2011] command *predict.m* were evaluated according to:

- Percentage FIT

$$FIT = \left(1 - \frac{|\, y_k - \hat{y}_k \,|}{|\, y_k - \bar{y}_k \,|}\right) \times 100\%$$

- prediction error variance

$$E\{(y_k - \hat{y}_k)(y_k - \hat{y}_k)^\intercal\}$$

- Percentage Variance Accounted For (VAF)

$$VAF = 1 - \frac{E[(y_k - \hat{y}_k)(y_k - \hat{y}_k)^\intercal]}{E[y_k y_k^\intercal]} \times 100\%$$

and compared to those achieved with the zero-order hold (ZOH) $\hat{y}_{k+\tau|k} = y_k$, with $\tau = 30, 60, 90, 120$.

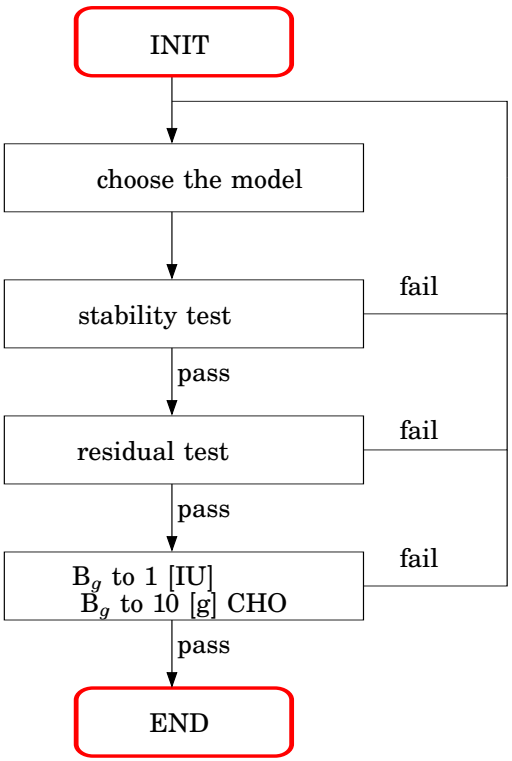

Figure 4.7 Diagram for model evaluation.

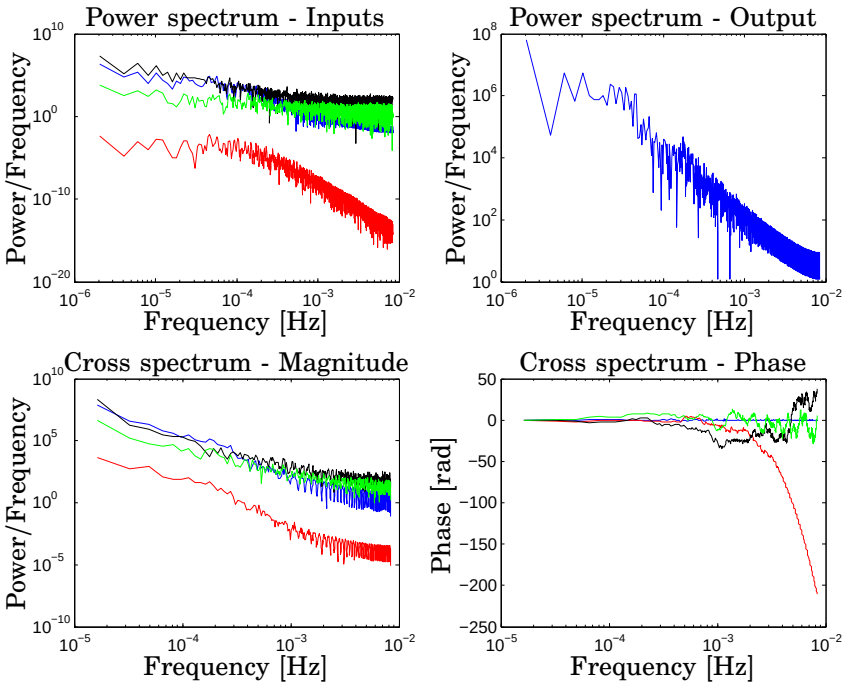

Figure 4.8 Patient CHU0102 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2$ (Hz)] (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

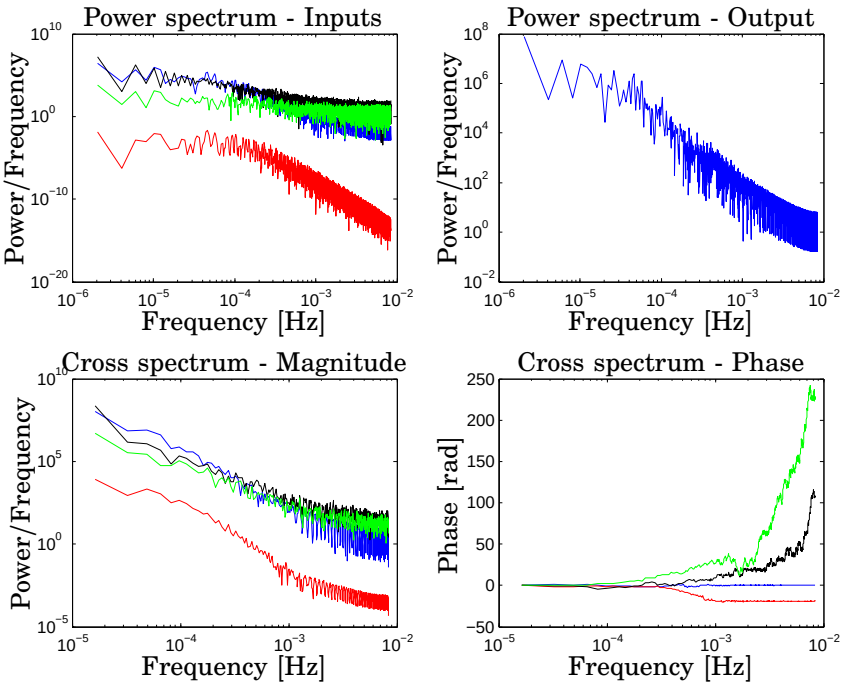

Figure 4.9 Patient CHU0107 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2\ (\text{Hz})]$ (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

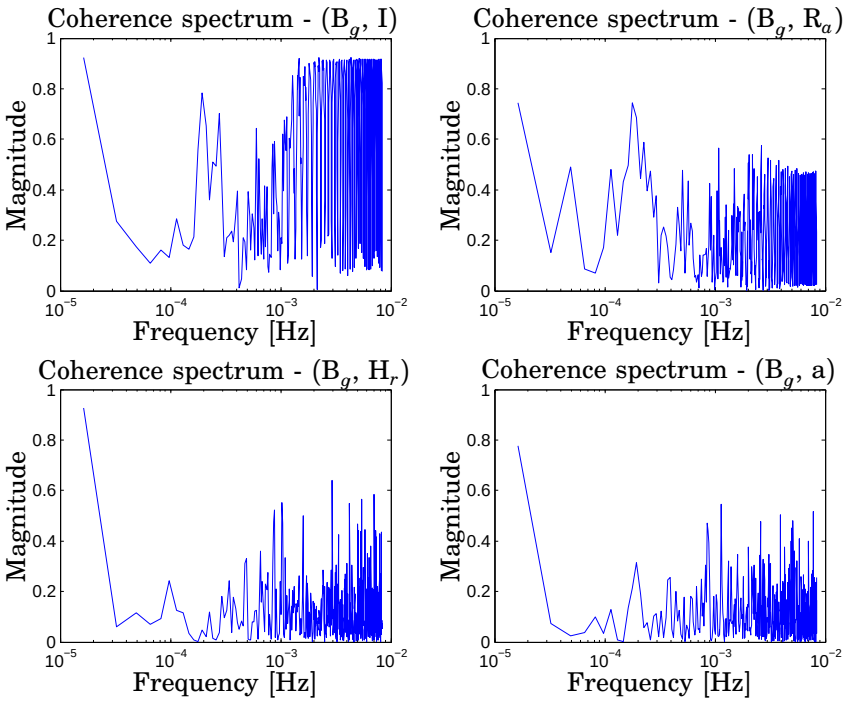

Figure 4.10 Patient CHU0102. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

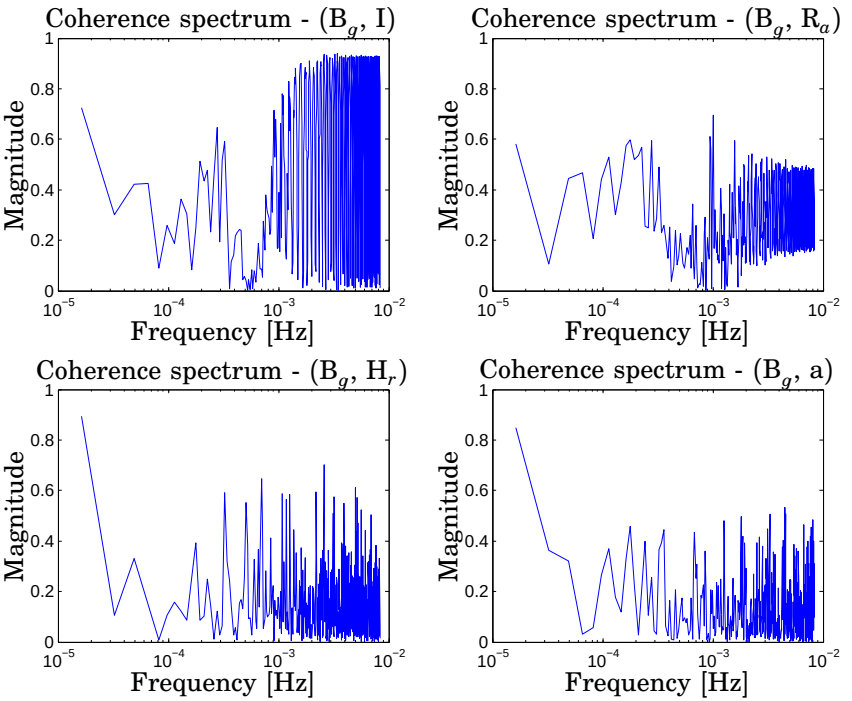

Figure 4.11 Patient CHU0107. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

4.5 Results

Data analysis

The autospectra (power spectra) of inputs and output showing the frequency contents of the signals investigated are reported in Figs. 4.8 and 4.9 for the representative patients 102 and 107, respectively and in Appendix B for the remaining patients in the selected population. The coherence spectrum between the inputs and the controlled variable are shown in Figs. 4.10 and 4.11 and in Appendix C.

Models

Detailed results out of the procedure outlined in Fig. 4.7 are given in Tables 4.3, 4.4 and 4.5. The methods failing to provide a model complying with the criteria in Sec. 4.4 are listed along with the tests that were not met. Pole-zero diagrams of the models selected for patients 102 and 107 as well as residual analysis are given in Figs. 4.13 and 4.16, respectively. Figures 4.14 and 4.17 report step and impulse responses as well as the model output to 1 [IU] of insulin and 10 [g] carbohydrates. Modeling results over the selected population are presented in Appendix D.

Performances on short-term predictions, i.e., up to 120 minutes are displayed in Figs. 4.18 and 4.19, while comparisons of the model-based predictors with the projection of the current glucose value in the future, i.e., the ZOH, are reported in Figs. 4.20 and 4.21 and quantitatively in Tables 4.6, 4.7, 4.8 (notice that the table were edited from Matlab output, the number of decimal places not representing numerical accuracy or significance). Appendix E deals with the remaining subjects while the boxplots in Figs. 4.22 show mean population performances. Finally, adding the squared heart rate and squared acceleration as additional inputs to the existing models resulted in Figs. 4.23 and 4.24.

Table 4.3 Model evaluation: fulfilled/ not fulfilled requirements for the discarded models. Patient 102, 103, 104

Patient ID	method	n	stability	residuals	B_g to I, R_a
102	N4SID	3	$\surd$	x	
		4	$\surd$	x	
		5	$\surd$	x	
	PBSID	3	x		
		4	x		
		5	x		
	MOESP	3	$\surd$	x	
		4	$\surd$	x	
		5	$\surd$	x	
103	N4SID	2	$\surd$	x	
		3	$\surd$	x	
		4	$\surd$	$\surd$	x
	PBSID	4	x		
		5	x		
		6	x		
	MOESP	3	$\surd$	x	
		4	$\surd$	x	
		5	$\surd$	x	
104	N4SID	3	$\surd$	x	
		4	$\surd$	$\surd$	x
		5	$\surd$	$\surd$	x
	PBSID	3	x		
		4	$\surd$	x	
		5	x		
	MOESP	5	$\surd$	x	
		6	x		
		7	$\surd$	$\surd$	x

$\surd$ = pass; x= fail

Table 4.4 Model evaluation: fulfilled/ not fulfilled requirements for the discarded models. Patient 105, 106, 107

Patient ID	method	n	stability	residuals	B_g to I, R_a
105	N4SID	2	√	√	x
		3	√	√	x
		4	√	√	x
	PBSID	2	x		
		3	√	√	x
		4	√	√	x
	MOESP	3	√	x	
		4	√	x	
		5	√	x	
106	N4SID	2	√	x	
		3	√	x	
		4	√	x	
	PBSID	3	√	x	
		4	x		
		5	√	x	
	MOESP	5	√	x	
		6	√	x	
		7	√	x	
107	N4SID	2	√	x	
		3	√	x	
		4	√	√	x
	PBSID	2	√	x	
		3	√	x	
		4	x		
	MOESP	4	√	x	
		5	√	x	
		6	√	x	

√ = pass; x= fail

Table 4.5 Model evaluation: fulfilled/ not fulfilled requirements for the discarded models. Patient 115, 120, 130

Patient ID	method	n	stability	residuals	B_g to I, R_a
115	N4SID	3	√	x	
		4	√	x	
		5	√	√	x
	PBSID	3	x		
		4	x		
		5	x		
	MOESP	5	√	x	
		6	√	x	
		7	√	x	
120	N4SID	2	√	√	x
		3	√	√	x
		4	√	√	x
	PBSID	2	√	x	
		3	√	x	
		4	√	x	
	MOESP	5	√	x	
		6	√	x	
		7	√	x	
130	N4SID	2	√	√	x
		3	√	√	x
		4	√	√	x
	PBSID	2	x		
		3	x		
		4	√	x	
	MOESP	3	x		
		4	√	√	x
		5	x		

√ = pass; x= fail

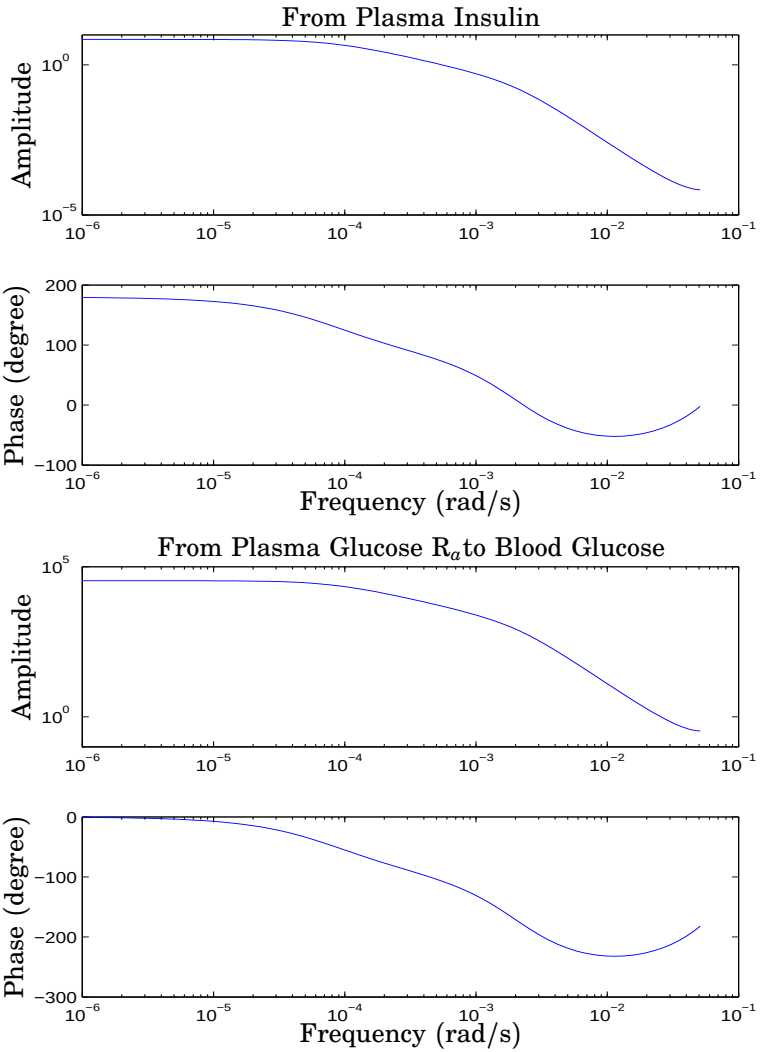

Figure 4.12 Patient CHU0102. ARMAX 3rd-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

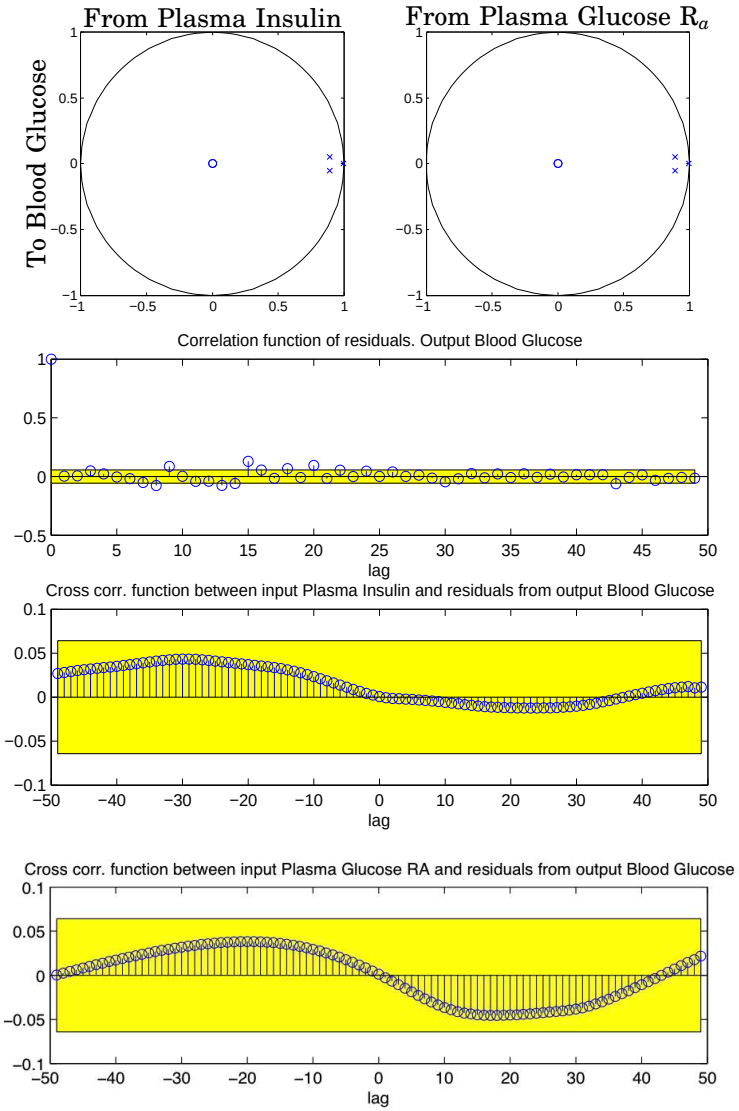

Figure 4.13 Patient 102. ARMAX 3rd-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data.

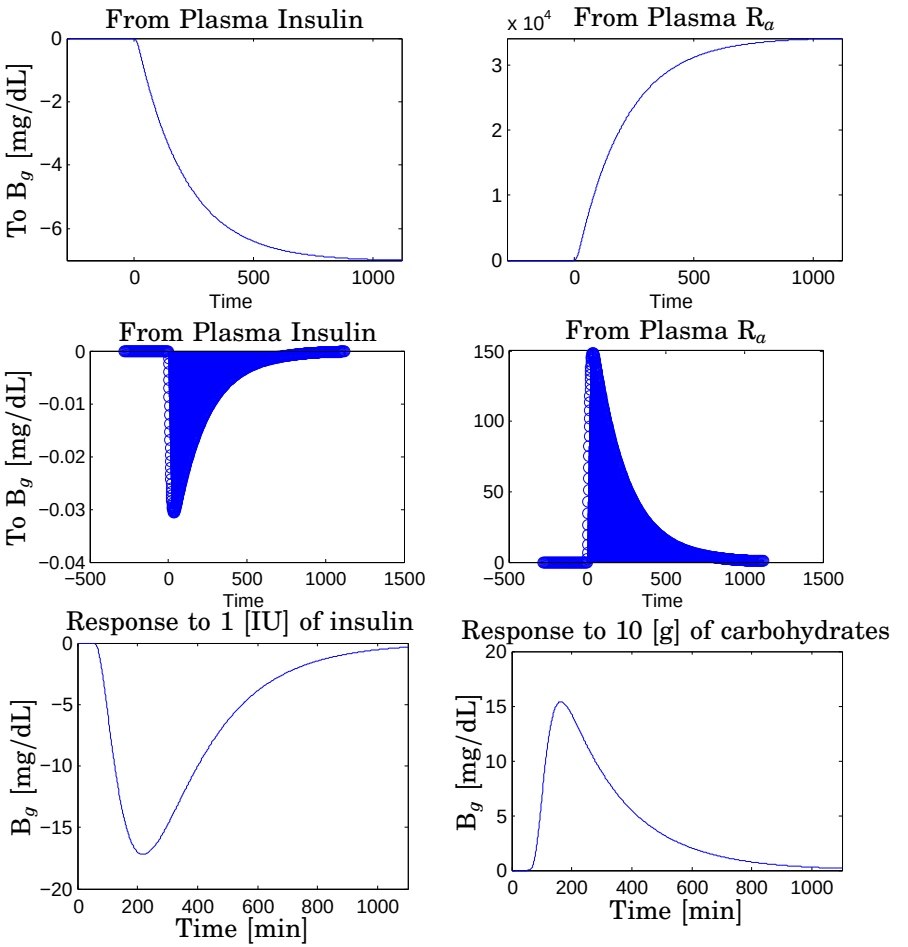

Figure 4.14 Patient CHU0102. ARMAX 3rd-order model. *Top* Step Responses; *Center* Impulse responses; *Bottom Left* Blood glucose response to 1[IU] of fast-acting insulin; *Bottom Right* Blood glucose response to 10[g] of carbohydrates

Table 4.6 Model-based predictor performance evaluation. Percentage FIT [%] vs. Prediction Horizon [min] on validation data without vital signs.

Patient ID	predictor	30[min]	60[min]	90[min]	120[min]
102	ARMAX	73.5358	49.4739	32.8066	20.7979
	ZOH	60.6483	32.5733	9.0154	-10.1487
103	ARMAX	65.0763	33.7277	19.7100	10.3026
	ZOH	50.6296	20.8109	-1.3747	-21.3287
104	ARMAX	76.0065	54.1321	42.0296	34.4656
	ZOH	52.1479	19.1685	-4.9330	-22.7172
105	ARMAX	57.1568	27.7924	13.4519	7.2640
	ZOH	47.2140	14.5802	-6.1496	-18.9310
106	ARMAX	54.7189	8.3002	-25.1534	-48.6235
	ZOH	44.0662	9.7080	-8.2996	-15.4517
107	ARMAX	68.2954	48.3223	34.0031	26.5711
	ZOH	52.2875	16.1428	-10.6265	-29.7901
115	ARMAX	79.7602	59.3186	42.1829	32.4400
	ZOH	63.6523	34.7217	12.3847	-3.4325
120	ARMAX	76.0582	49.3294	30.8174	21.5344
	ZOH	58.4849	27.0355	4.3175	-12.5778
130	ARMAX	63.1140	39.9365	20.9482	9.3143
	ZOH	58.8014	29.0776	6.3448	-9.5646

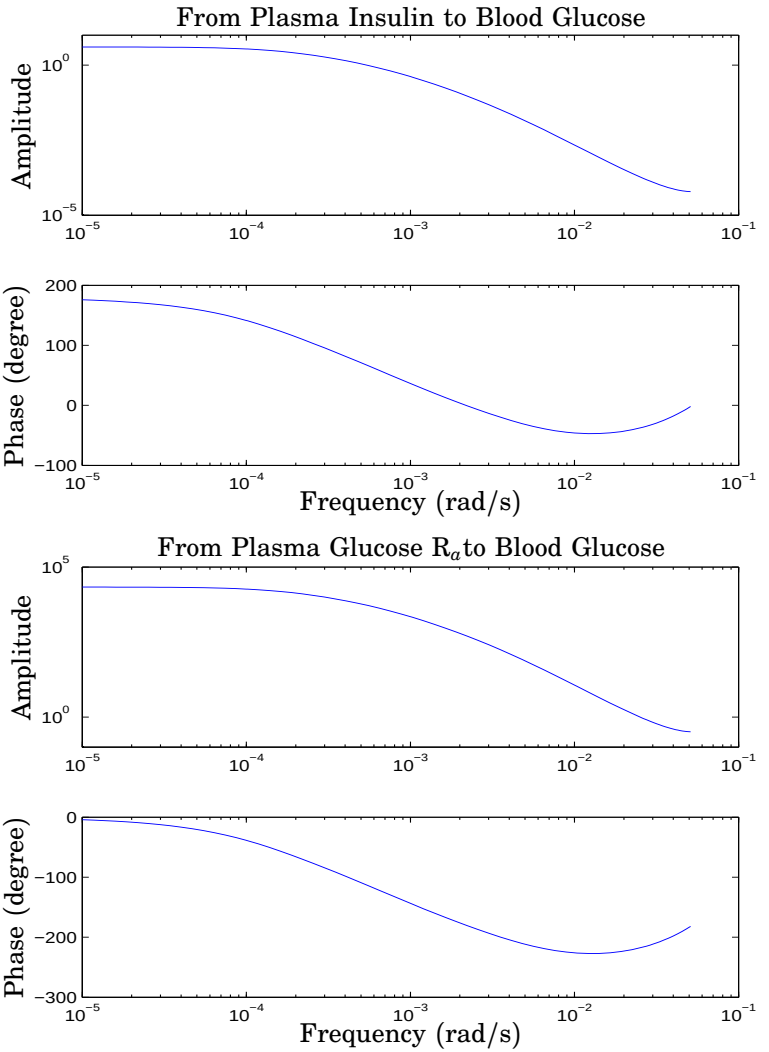

Figure 4.15 Patient CHU0107. ARMAX 3rd-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

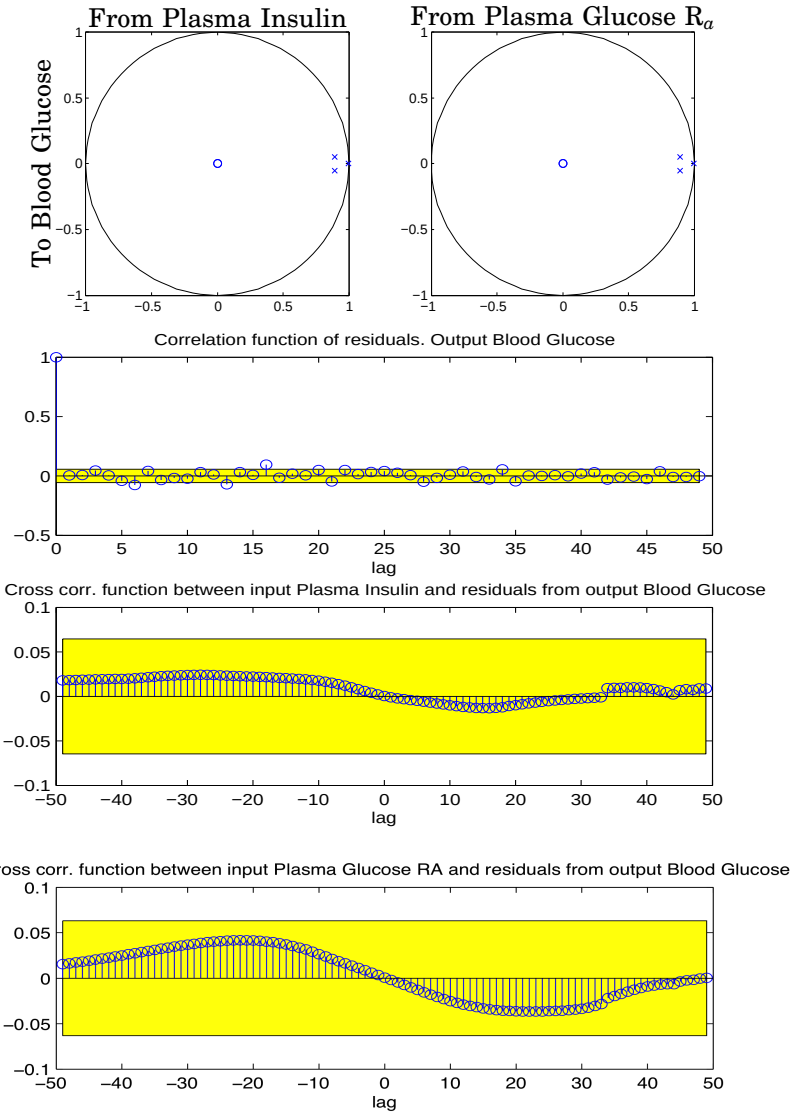

Figure 4.16 Patient 107. ARMAX 3rd-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

Table 4.7 Model-based predictor performance evaluation. Prediction Error Variance $10^3[(\mathrm{mg/dL})^2]$ vs. Prediction Horizon [min] on validation data without vital signs.

Patient ID	predictor	30[min]	60[min]	90[min]	120[min]
102	ARMAX	0.1788	0.6462	1.1344	1.5641
	ZOH	0.3981	1.1676	2.1253	3.1139
103	ARMAX	0.1654	0.5896	0.8474	1.0292
	ZOH	0.3323	0.8545	1.3988	2.0007
104	ARMAX	0.2635	0.9593	1.5290	1.9525
	ZOH	1.0513	2.9976	5.0470	6.8949
105	ARMAX	0.2842	0.8028	1.1445	1.2995
	ZOH	0.4306	1.1233	1.7245	2.1461
106	ARMAX	0.5172	2.1103	3.9098	5.4791
	ZOH	0.7973	2.0763	2.9834	3.3862
107	ARMAX	0.2433	0.6391	1.0343	1.2715
	ZOH	0.5535	1.7088	2.9718	4.0885
115	ARMAX	0.07313	0.2834	0.5559	0.7294
	ZOH	0.2423	0.7786	1.3951	1.9321
120	ARMAX	0.2020	0.8952	1.6533	2.1069
	ZOH	0.6141	1.8968	3.2605	4.5099
130	ARMAX	0.7108	1.8847	3.2648	4.2967
	ZOH	0.8868	2.6279	4.5825	6.2708

Table 4.8 Model-based predictor performance evaluation. Percentage VAF [%] vs. Prediction Horizon [min] on validation data without vital signs.

Patient ID	predictor	30[min]	60[min]	90[min]	120[min]
102	ARMAX	93.0610	74.9189	55.9707	39.2934
	ZOH	84.5496	54.6813	17.5108	-20.8571
103	ARMAX	87.8640	56.7519	37.8332	24.4987
	ZOH	75.6258	37.3160	-2.6105	-46.7695
104	ARMAX	94.2616	79.1115	66.7055	57.4847
	ZOH	77.1072	34.7272	-9.8982	-50.1352
105	ARMAX	81.6979	48.3000	26.2977	16.3122
	ZOH	72.2686	27.6622	-11.0544	-38.2070
106	ARMAX	79.7178	17.2474	-53.3144	-114.8501
	ZOH	68.7340	18.5836	-16.9867	-32.7826
107	ARMAX	90.0062	73.7458	57.5140	47.7695
	ZOH	77.2643	29.8063	-22.0801	-67.9497
115	ARMAX	96.0871	84.8332	70.2528	60.9726
	ZOH	87.0342	58.3385	25.3581	-3.3786
120	ARMAX	94.3313	74.8741	53.5968	40.8661
	ZOH	82.7652	46.7650	8.4905	-26.5775
130	ARMAX	86.3956	63.9289	37.5156	17.7668
	ZOH	83.0280	49.7050	12.2986	-20.0131

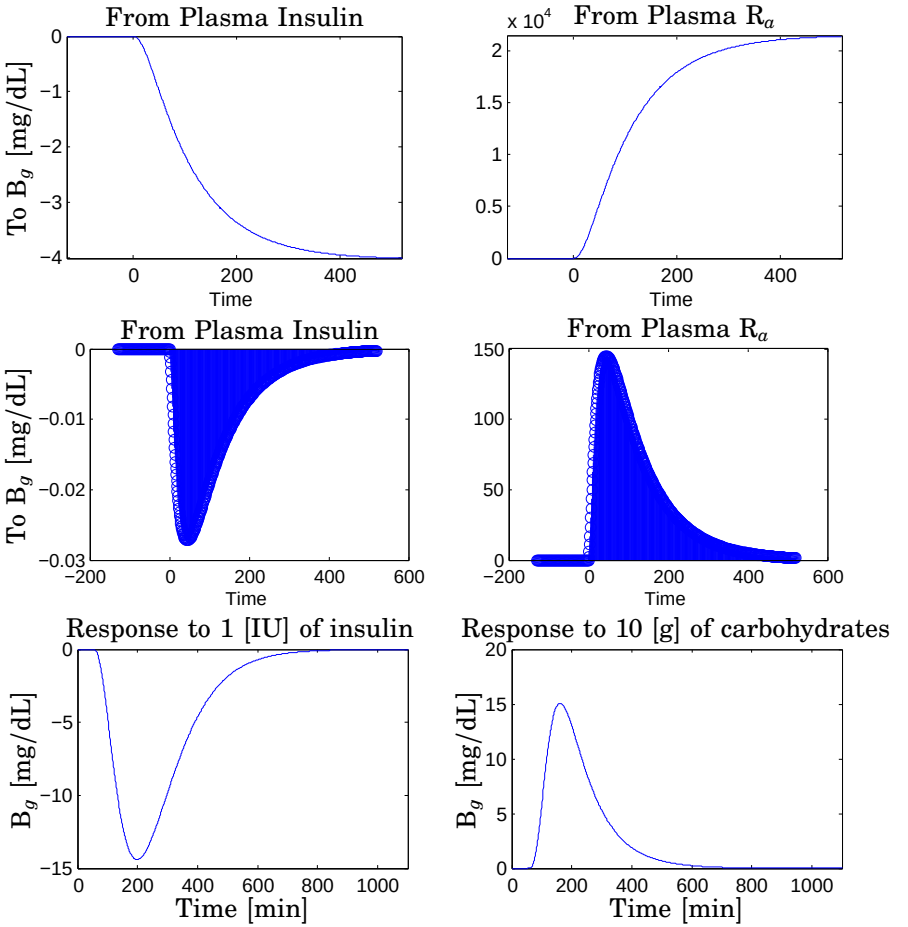

Figure 4.17 Patient CHU0107. ARMAX 3rd-order model. *Top* Step Responses; *Center* Impulse responses; *Bottom Left* Blood glucose response to 1[IU] of fast-acting insulin; *Bottom Right* Blood glucose response to 10[g] of carbohydrates

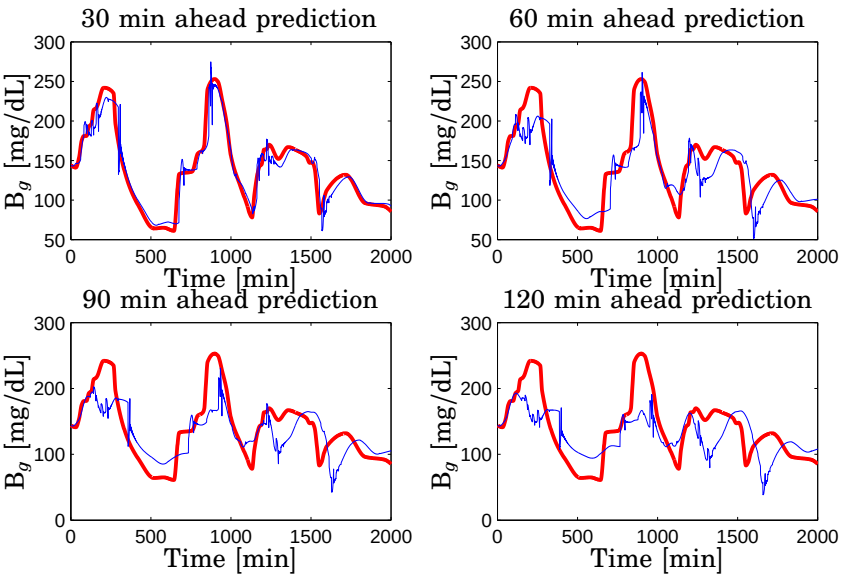

Figure 4.18 Patient CHU0102. Evaluation on validation data: 3rd-order ARMAX-based predictor (thin) and measured plasma glucose (thick) [mg/dL] vs. time [min]. Top left 30-minutes ahead; Top right 60-minutes ahead; Bottom left 90-minutes ahead; Bottom right 120-minutes ahead prediction

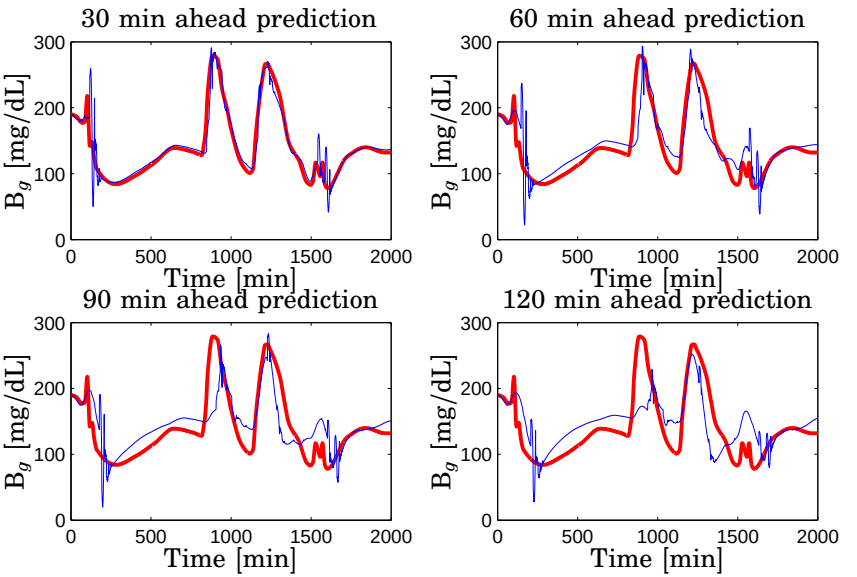

Figure 4.19 Patient CHU0107. Evaluation on validation data: 3rd-order ARMAX-based predictor (thin) and measured plasma glucose (thick) [mg/dL] vs. time [min]. Top left 30-minutes ahead; Top right 60-minutes ahead; Bottom left 90-minutes ahead; Bottom right 120-minutes ahead prediction

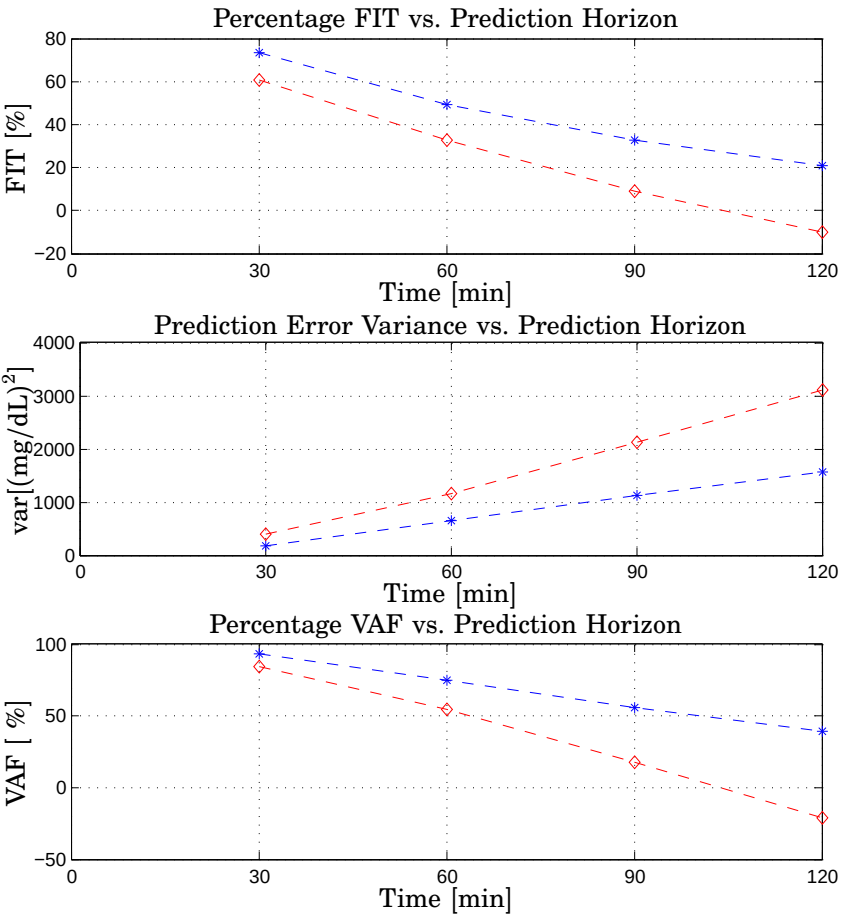

Figure 4.20 Patient CHU0102. 3rd-order ARMAX-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

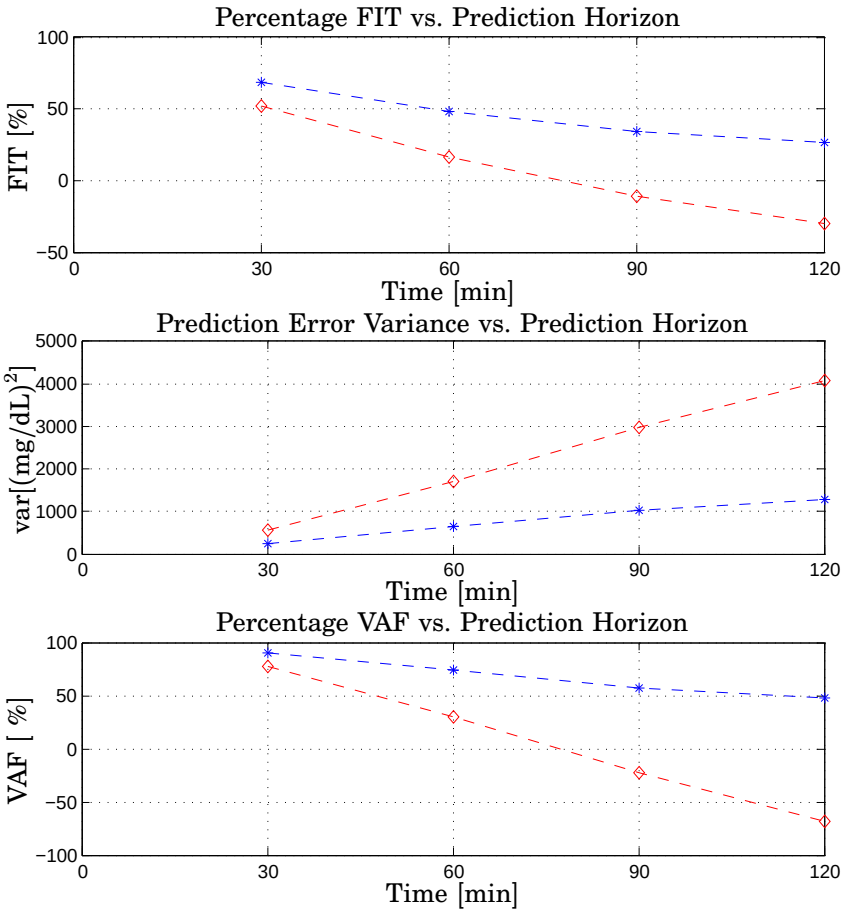

Figure 4.21 Patient CHU0107. 3rd-order ARMAX-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

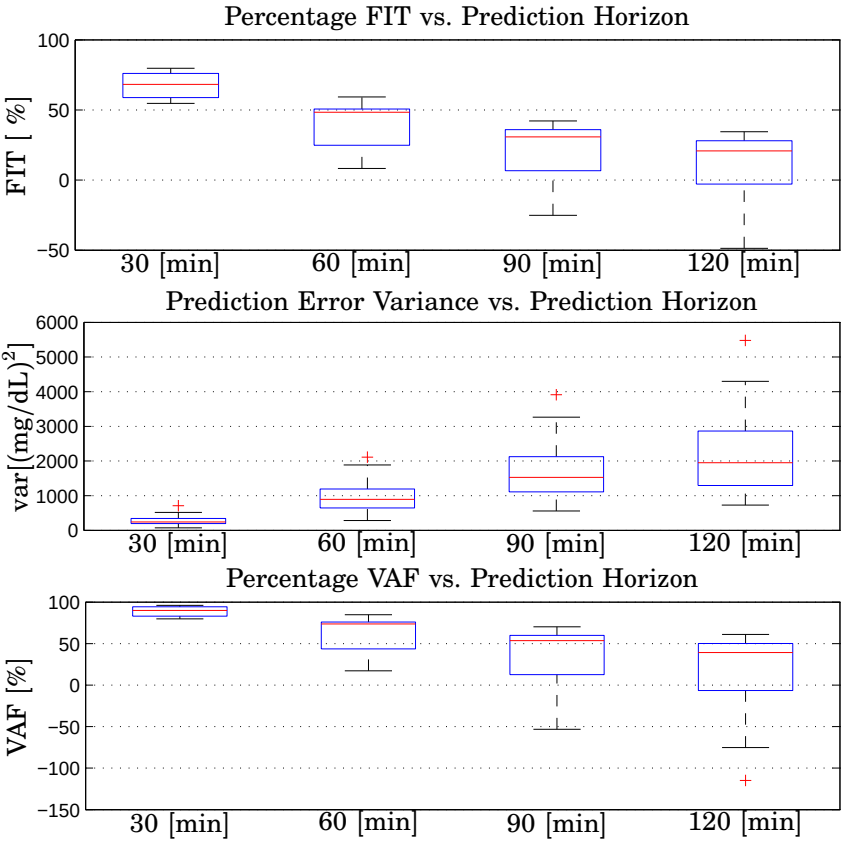

Figure 4.22 Model-based predictor performance evaluation. Population results on validation data. *Top* Percentage FIT vs. Prediction Horizon [min]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$ vs. Prediction Horizon [min]; *Bottom* Percentage VAF vs. Prediction Horizon [min]. Each box presents results over the population considered. The central mark is the median, the edges of the box are the 25th and the 75th percentiles.

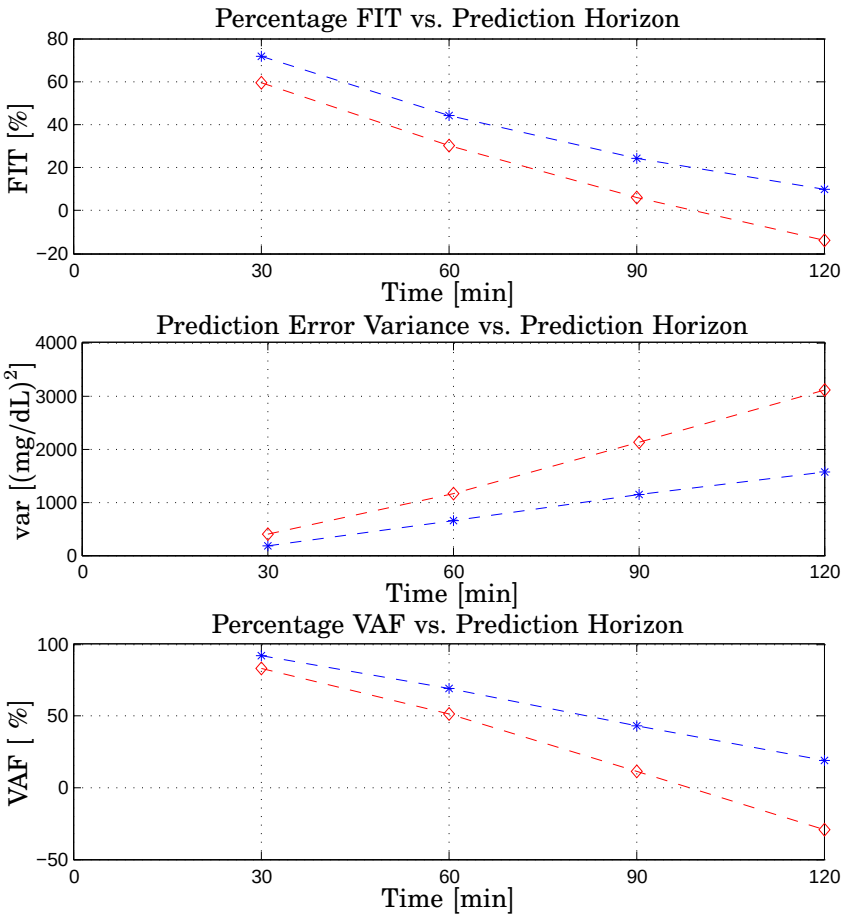

Figure 4.23 Patient CHU0102. 3rd-order ARMAX-based predictor using vital signs (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

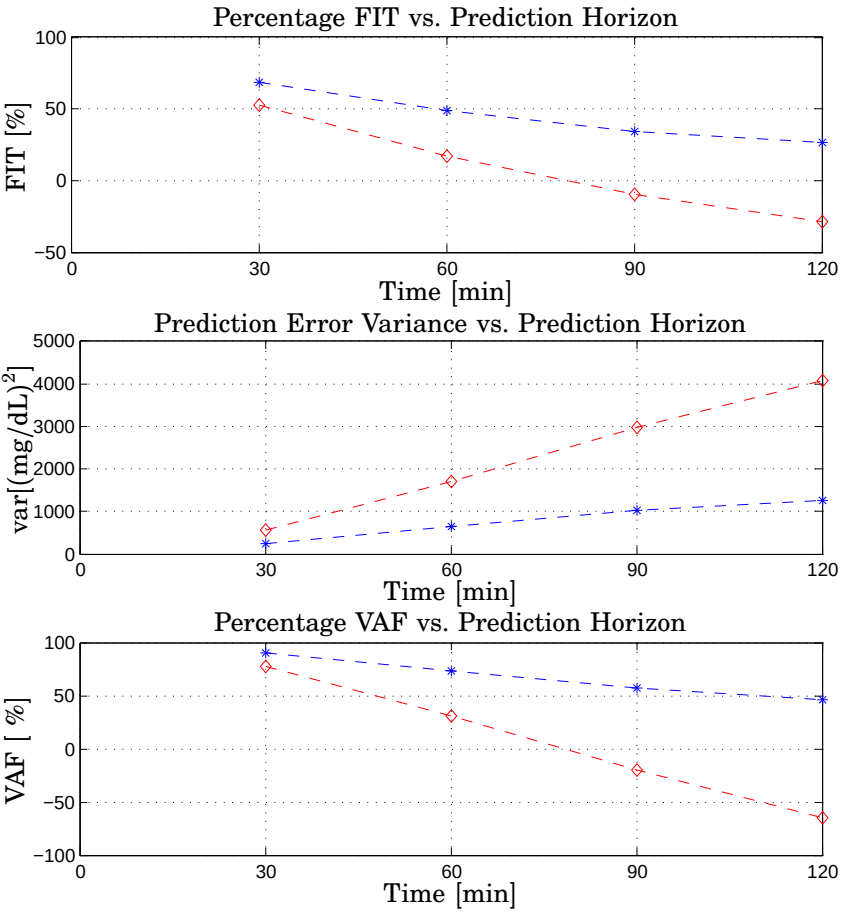

Figure 4.24 Patient CHU0107. 3rd-order ARMAX-based predictor using vital signs (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

5

Linear Prediction Strategies

The most widespread approach to the design of a linear predictor for discrete-time systems is based on a finite-dimensional model, either from first principles or identified from data ([Ljung, 1999], [Söderström and Stoica, 1989]). Along this line, the identification of multi-step-ahead optimal predictors of certain observed, i.e., available to measurements, variables have been analyzed in the past by several authors (e.g., [Åström, 1980], [Anderson and Moore, 1979], [Holst, 1977], [Mosca *et al.*, 1989], [Weiss, 1991], [Chevillon, 2007]).

This chapter will be focused on the direct identification of multi-step-ahead predictors and their application to the problem of blood glucose prediction in diabetic patients.

Preliminaries

Let $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^l$ be input and output, respectively, of an unknown discrete-time time-invariant system $\mathcal{S}_n(F, G)$ described by a general form multivariate model:

$$y_k = F(q^{-1}, \theta)u_k + G(q^{-1}, \theta)w_k \tag{5.1}$$

where θ is the parameter vector, $F(q^{-1}, \theta)$ and $G(q^{-1}, \theta)$ are transfer functions of dimension $(l \times m)$ and $(l \times l)$ respectively, q^{-1} represents the backward shift operator and $\{w_k\}$ is a white-noise sequence, independent of the past history of the joint input-output process

denoted by $z_k = [u_k^T y_k^T]^T$. Without loss of generality we assume that $\mathcal{S}_n(F, G)$ is minimal in the sense that it cannot be described by a state-space model of order less than n. All through this chapter, we shall assume z_k a purely non-deterministic (p.n.d.) process with spectral properties such that the determinant of the spectral density matrix $S_{zz}(e^{j\omega})$ has no zero on the unit circle [Hannan and Poskitt, 1988].

In the sequel we shall be concerned with two different structures for the transfer functions $F(q^{-1}, \theta)$ and $G(q^{-1}, \theta)$. First, consider the full polynomial model [Åström, 1970]:

$$A(q^{-1})y_k = B(q^{-1})u_k + C(q^{-1})w_k \tag{5.2}$$

where $A(q^{-1}) \in \mathbb{R}^{l\times l}$, $B(q^{-1}) \in \mathbb{R}^{l\times m}$ and $C(q^{-1}) \in \mathbb{R}^{l\times m}$ are the following matrix polynomials:

$$\begin{aligned}
A(q^{-1}) &= I_{l\times l} + A_1 q^{-1} + \cdots + A_{n_a} q^{-n_a}, && A_1, \cdots, A_{n_a} \in \mathbb{R}^{l\times l} \\
B(q^{-1}) &= B_1 q^{-1} + \cdots + B_{n_b} q^{-n_b}, && B_1, \cdots, B_{n_b} \in \mathbb{R}^{l\times m} \\
C(q^{-1}) &= I_{l\times m} + C_1 q^{-1} + \cdots + C_{n_c} q^{-n_c}, && C_1, \cdots, C_{n_c} \in \mathbb{R}^{l\times m}
\end{aligned}$$

This relates to (5.1) according to

$$F(q^{-1}) = A^{-1}(q^{-1})B(q^{-1}), \quad G(q^{-1}) = A^{-1}(q^{-1})C(q^{-1}) \tag{5.3}$$

Next, consider the state-space model in innovation form:

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + Ke_k \\ y_k = Cx_k + Du_k + e_k \end{cases} \tag{5.4}$$

with state vector $x_k \in \mathbb{R}^n$ and zero-mean white noise innovation process $e_k \in \mathbb{R}^l$. The transfer functions $F(q^{-1})$ and $G(q^{-1})$ are expressed by:

$$F(q^{-1}) = C(qI - A)^{-1}B, \quad G(q^{-1}) = C(qI - A)^{-1}K + I$$

Let us define $\bar{A} = A - KC$ for future reference.

5.1 Data-driven multistep subspace-based linear predictors

Notation

The available data sequences $\{u_k\}$, $\{y_k\}$ and the innovation process $\{e_k\}$ will be organized in Hankel matrices and will be denoted by uppercase letters. Subscript indices $[\iota, \kappa]$ of a matrix will be used to indicate the argument of the upper-left and the lower-left element, respectively, with the meaning of the spanned interval of time, e.g., $U_{[t_1,t_1+t_2]}$ will contain in the first column the inputs from time t_1 to time $t_1 + t_2$. Accordingly, data records of finite length N will be represented by the block rows of the block Hankel data matrices, e.g.,

$$U_k := [u_k \quad u_{k+1} \quad \cdots \quad u_{k+N-1}]$$

$$Y_k := [y_k \quad y_{k+1} \quad \cdots \quad y_{k+N-1}]$$

In the sequel, the orthogonal projection of the rows of a given matrix A onto the row space of a given matrix B will be denoted by $\hat{E}\{A \mid B\}$, whereas the symbol $\hat{E}_{\|C}\{A \mid B\}$ will denote the oblique projection of the row space of A onto the row space of B along the row space of C, the projection operator being $\hat{E}\{\cdot\}$.

Throughout the thesis k shall denote the current time instant in the identification problem, k_0 shall be the initial time from which the data are collected, so that $k - k_0$ is the past horizon denoted by p in the identification problem, T shall be such that $T - k$ represents the future horizon denoted by f, the two integers p, f being the number of block rows in the processed Hankel data matrices with $p \geq f$ while t shall denote the sample index in the prediction problem, $t > T + N$. Finally, the number of steps in the look-ahead horizon will be denoted by τ.

Statement of the problem

Let the finite sequences $\{u_k\}$ and $\{y_k\}$, $k = k_0, \cdots, T + N$ be the measured input and the corresponding output, respectively, of the system (5.1). Our aim is to find a multistep estimator of the future, i.e., not yet observed, output sequence $\{y_k\}, k = t+1, \cdots, t+\tau$ by means of linear combinations of the joint input-output up to time t.

State space form case

Consider the block Hankel matrices:

$$U^P := U_{[k_0,k)} \in \mathbb{R}^{p\cdot m\times N} \tag{5.5}$$
$$= \begin{bmatrix} u_{k_0} & u_{k_0+1} & \cdots & u_{k_0+N-1} \\ u_{k_0+1} & u_{k_0+2} & \cdots & u_{k_0+N} \\ \vdots & \vdots & \cdots & \vdots \\ u_{k-1} & u_k & \cdots & u_{k+N-2} \end{bmatrix} \in \mathbb{R}^{p\cdot m\times N}$$

and

$$U^F := U_{[k,T)} \in \mathbb{R}^{f\cdot m\times N} \tag{5.6}$$
$$= \begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+N-1} \\ u_{k+1} & u_{k+2} & \cdots & u_{k+N} \\ \vdots & \vdots & \cdots & \vdots \\ u_{T-1} & u_T & \cdots & u_{T+N-2} \end{bmatrix} \in \mathbb{R}^{f\cdot m\times N}$$

called the past and future input data matrices, respectively. Similarly, consider

$$Y^P := Y_{[k_0,k)} \in \mathbb{R}^{p\cdot l\times N} \tag{5.7}$$

$$Y^F := Y_{[k,T)} \in \mathbb{R}^{f\cdot l\times N} \tag{5.8}$$

past and future output data matrices, respectively. The integer p is chosen so to satisfy the following condition:

$$p \geq \max\{n,\tau\} \tag{5.9}$$

in order for the system 5.4 to be observable and to guarantee predictions up to the largest future horizon we wish to investigate.

As first step in the identification problem we will look for an estimator of the known future data:

$$Y_{k+d}, \qquad 1 \leq d \leq f+1 \tag{5.10}$$

Mathematically, this can be formulated as the following least-squares problem ([Van Overschee and De Moor, 1996]):

$$\hat{\gamma}, \hat{\lambda} = \underset{\substack{\gamma \in \mathbb{R}^{lp \times (l+m)p} \\ \lambda \in \mathbb{R}^{lp \times mf}}}{\operatorname{argmin}} \; || \, Y_{k+d} - [\, \gamma \quad \lambda \,] \, [\, Z^P \quad U^F \,]^T \, ||_F^2 \tag{5.11}$$

where $|| \cdot ||_F$ stands for the Frobenius norm of a matrix.

Geometrically, it can be interpreted as the orthogonal projection of Y_{k+d} onto $[\, Z^P \quad U^F \,]^T$ [Van Overschee and De Moor, 1994], i.e.,

$$\hat{Y}_{k+d} = \hat{E}[Y_{k+d} \mid [\, Z^P \quad U^F \,]^T] \tag{5.12}$$

As a matter of fact, the orthogonal projection 5.12 corresponds to the sum of two oblique projections [Katayama and Picci, 1999, Lemma 1]:

$$\begin{aligned} \hat{Y}_{k+d} &= \hat{E}_{\| U^F}[Y_d \mid Z^P] + \hat{E}_{\| Z^P}[Y_d \mid U^F] && (5.13) \\ &= \hat{\gamma} Z^P + \hat{\lambda} U^F && (5.14) \end{aligned}$$

Now, taking conditional expectation up to time k to eliminate the effect of yet unknown inputs, the output predictor for each vector Y_{k+d} as linear combination of past input and output amounts to computing an oblique projection, i.e.,

$$\begin{aligned} \hat{Y}_{k+d|k} &= \hat{E}_{\| U^F}[Y_d \mid Z^P] && (5.15) \\ &\simeq O_{k+d} X_{k+d} && (5.16) \end{aligned}$$

where O_{k+d} is the extended observability matrix and X_{k+d} is the state sequence. Stacking all the predictors on top of each other, the sought matrix of multi-step ahead predictors is obtained:

$$\hat{Y}^F = \begin{bmatrix} \hat{Y}_{k+1|k}^T & \hat{Y}_{k+2|k}^T & \cdots & \hat{Y}_{k+f-1|k}^T \end{bmatrix}^T \tag{5.17}$$

$\hat{Y}^F$ will be calculated as follows:

$$\hat{Y}^F = \hat{\gamma} Z^P \tag{5.18}$$

Algorithm 2 summarizes the implementation details. In the sequel the approach illustrated above applying (5.12) and (5.13) will be referred

Algorithm 2 *Projection*

Input: $u, y, p = f$
Output: $\hat{\gamma}$

- Construct Hankel matrices U^P, U^F, Y^P, Y^F
- Perform an LQ decomposition on $[\, Y^P \quad U^P \quad U^F \quad Y^F \,]^T$

$$\begin{bmatrix} U^P \\ U^F \\ U^F \\ Y^F \end{bmatrix} = \begin{bmatrix} \mathscr{L}_{11} & 0 & 0 \\ \mathscr{L}_{21} & \mathscr{L}_{22} & 0 \\ \mathscr{L}_{31} & \mathscr{L}_{32} & \mathscr{L}_{33} \end{bmatrix} \begin{bmatrix} \mathscr{Q}_1{}^T \\ \mathscr{Q}_2{}^T \\ \mathscr{Q}_3{}^T \end{bmatrix}$$

- Compute $\hat{\gamma} = \mathscr{L}_{31}\mathscr{L}_{11}{}^{\dagger}$

to as projection-based.
Let us now consider the innovation model (5.4) in predictor form:

$$\begin{cases} x_{k+1} = \bar{A}x_k + Bu_k + Ky_k \\ \hat{y}_{k|k-1} = Cx_k \end{cases} \tag{5.19}$$

Recall that by iteration of (5.19) it is possible to express the output data at time $k+d$, $1 \leq d \leq f+1$ according to:

$$\begin{aligned} Y_{k+d} &= C\bar{A}^d X_k + \sum_{h=1}^{d} C\bar{A}^{h-1}(KY_{k+d-h} + BU_{k+d-h}) + \\ &+ E_{k+d} \\ &= \bar{O}_d \bar{A}^p X_p + \Xi Z^P + \Psi Z^F + E_{k+d} \end{aligned} \tag{5.20}$$

where the first term depends on the initial conditions of the state, the second term depends upon past input-output data and the third on future input-output data. Stacking all the future data sequences on top of each other and assuming $\bar{A}^p$ small, i.e., disregarding the effects of the unknown initial states for sufficiently large p, we obtain the

$$\bar{O}_d = \begin{bmatrix} C \\ C\bar{A} \\ C\bar{A}^2 \\ \vdots \\ C\bar{A}^{d-1} \end{bmatrix}, \Psi = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ C[B \quad K] & 0 & \cdots & \cdots & \cdots & 0 \\ C\bar{A}[B \quad K] & C[B \quad K] & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ C\bar{A}^{f-2}[B \quad K] & \cdots & \cdots & \cdots & C[B \quad K] & 0 \end{bmatrix}, \quad (5.24)$$

$$\Xi = \begin{bmatrix} C\bar{A}^{p-1}[B \quad K] & C\bar{A}^{p-2}[B \quad K] & \cdots & \cdots & \cdots & C[B \quad K] \\ 0 & C\bar{A}^{p-1}[B \quad K] & \cdots & \cdots & \cdots & C\bar{A}[B \quad K] \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & C\bar{A}^{p-1}[B \quad K] & \cdots & C\bar{A}^{f-1}[B \quad K] \end{bmatrix} \quad (5.25)$$

matrix relations:

$$[\, Y_k^T \quad Y_{k+1}^T \quad \cdots \quad Y_T^T \,]^T = \Xi Z^P + \bar{\Psi} Z^F + E_k \quad (5.21)$$

Matrices $\bar{O}_d$, Ψ and Ξ are given in (5.24) and (5.25). Solving the least-squares problem:

$$Y_k = \Xi_0 Z^p + E_k \quad (5.22)$$

$$\hat{\Xi}_0 = \underset{\Xi_0}{\operatorname{argmin}} \; || \, Y_k - \Xi_0 Z^p \, ||_F^2 \quad (5.23)$$

where $|| \cdot ||_F$ stands for the Frobenius norm of a matrix, the Markov parameters of the system (5.4) are obtained [Chiuso, 2007]).

Using the estimated coefficients from (5.26)

$$\hat{\Xi}_0 = \left[\hat{C}\hat{\bar{A}}^{p-1}[\hat{B} \quad \hat{K}] \quad \hat{C}\hat{\bar{A}}^{p-2}[\hat{B} \quad \hat{K}] \quad \cdots \quad \cdots \quad \hat{C}[\hat{B} \quad \hat{K}] \right] \quad (5.26)$$

and the recipes in (5.27)

$$\begin{aligned} \Gamma_i &= \hat{\Xi}_i + \sum_{j=0}^{i-1} \hat{C}\hat{\bar{A}}^{i-j-1}\hat{K}\Gamma_j \\ \Lambda_i &= \hat{C}\hat{\bar{A}}^{i-1}\hat{B} + \sum_{j=1}^{i-1} \hat{C}\hat{\bar{A}}^{i-j-1}\hat{K}\Lambda_j \end{aligned} \quad (5.27)$$

with $\Gamma_0 = \hat{\Xi}_0$, $\Lambda_1 = \hat{C}\hat{B}$, the multistep predictors are given by

$$\begin{aligned}\begin{bmatrix} \hat{Y}_k \\ \hat{Y}_{k+1} \\ \vdots \\ \hat{Y}_{k+f-1} \end{bmatrix} &= \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \vdots \\ \Gamma_{f-1} \end{bmatrix} Z^P + \begin{bmatrix} \Lambda_1 & 0 & \cdots & 0 \\ \Lambda_2 & \Lambda_1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \Lambda_{f-1} & \Lambda_{f-2} & \cdots & \Lambda_1 \end{bmatrix} U^F \\ &= \hat{\Gamma} Z^P + \hat{\Lambda} U^F \end{aligned} \tag{5.28}$$

Taking conditional expectation up to time k, the output predictors are expressed by:

$$\hat{Y}^F = \hat{\Gamma} Z^P \tag{5.29}$$

This approach will be referred to as pbsid-based. Implementation details are provided in Algorithm 3.

Algorithm 3 *pbsid*

Input: u, y, p, f
Output: $\hat{\Gamma}$

- Construct Hankel matrices U^P, U^F, Y^P, Y^F
- Perform an LQ decomposition on $[\, Y^P \quad U^P \quad Y_k \,]^T$

$$\begin{bmatrix} U^P \\ Y^P \\ Y_k \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} {Q_1}^T \\ {Q_2}^T \end{bmatrix}$$

- Estimate the matrix that contains the Markov parameters

$$\hat{\Xi}_0 = L_{21} L_{11}^{-1}$$

- Extract the Markov parameters and construct the predictor matrices Γ_i according to Eq. (5.27)
- Compute $\hat{\Gamma}$ stacking all the Γ_i on top of each other

Once the operator $\hat{\Theta}$, $\hat{\gamma}$ and $\hat{\Gamma}$ have been estimated, they can be applied to new, still generated by the same underlying mechanisms, data to forecast the actual future.

We mention here that in this application, the knowledge of the estimated future inputs is available from simulation models. Then, the full form of the predictors, i.e., (5.14), (5.28) can be used.

5.2 Adaptive subspace-based prediction

For an on-line implementation of the subspace predictors, the predictor coefficients should be recomputed again each time new data becomes available, i.e., at each new time step. This can be accomplished by solving (5.23) at every sample time by means of a "new", in the sense that it contains new data, LQ decomposition [Golub and Van Loan, 1996]. Consider, for instance, the following:

$$\begin{bmatrix} Z_{[k-k_0,k)} \\ Y_{k+h} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} \tag{5.30}$$

The solution of the regression problem (5.23) can be computed as

$$\hat{\Xi}_0 = L_{21} L_{11}^{-1} \tag{5.31}$$

Recursively updating the estimate $\hat{\Xi}_0$ can be approached applying Givens QR Method [Golub and Van Loan, 1996] as in [M. Lovera and Verhaegen, 2000]. Let the blocks L_{11} and L_{21} at time $k-1$ be denoted by $L_{11}(k-1)$ and $L_{21}(k-1)$, respectively. When new input-output data are available at time k the vector $[z_p^T \quad y_k^T]^T$ can be appended to $[L_{11}(k-1)^T \quad L_{21}(k-1)^T]^T$; next, by applying a sequence of orthogonal Givens rotations $\mathcal{G}$ the matrix L can be made lower triangular, i.e., updated according to:

$$\left[\begin{array}{c|c} \sqrt{\lambda} L_{11}(k-1) & z_p \\ \hline \sqrt{\lambda} L_{21}(k-1) & y_k \end{array}\right] \mathcal{G} = \left[\begin{array}{c|c} L_{11}(k) & 0 \\ \hline L_{21}(k) & \tilde{y}_k \end{array}\right] \tag{5.32}$$

where a forgetting factor $\lambda \in [0.95, 1)$ was used to discount old data.

5.3 Results

The different estimation strategies were compared with respect to prediction capability on validation data. In particular, the performances were evaluated on different prediction horizons, namely, 30, 60, 90 and 120 minutes. Each of the data sequences available were divided into two equal parts of which the first half were used for identification and the second half for validation. As far as the least-squares identification of the ARX and regularized ARX predictor coefficients is considered, the parameters n_a and n_b were both set to 3. The regularization coefficient ρ was set to 40. The choice of the horizons in *pbsid* methods, instead, was $p = 120$, $f = 30, 60, 90, 120$, respectively, the choice of p being determined by the maximum value of the prediction horizon. Performances on short-term predictions are displayed in Figs. 5.1, 5.3, 5.2 and 5.4 as far as patient 102 and patient 107 are concerned and in Appendix F for the remaining patients. Comparisons of the model-based predictors with the projection of the current glucose value in the future, i.e., the ZOH, are reported in Figs. 5.5, 5.6 for patient 102 and 5.7, 5.8 for patient 107 and and in Appendix F for the remaining patients. Quantitatively the performances are shown in Tables 5.1, 5.2, 5.3 (notice that the tables are edited from Matlab output, the number of decimal places not representing numerical accuracy or significance) while the boxplots in Figs. 5.9 and 5.10 show mean population performances.

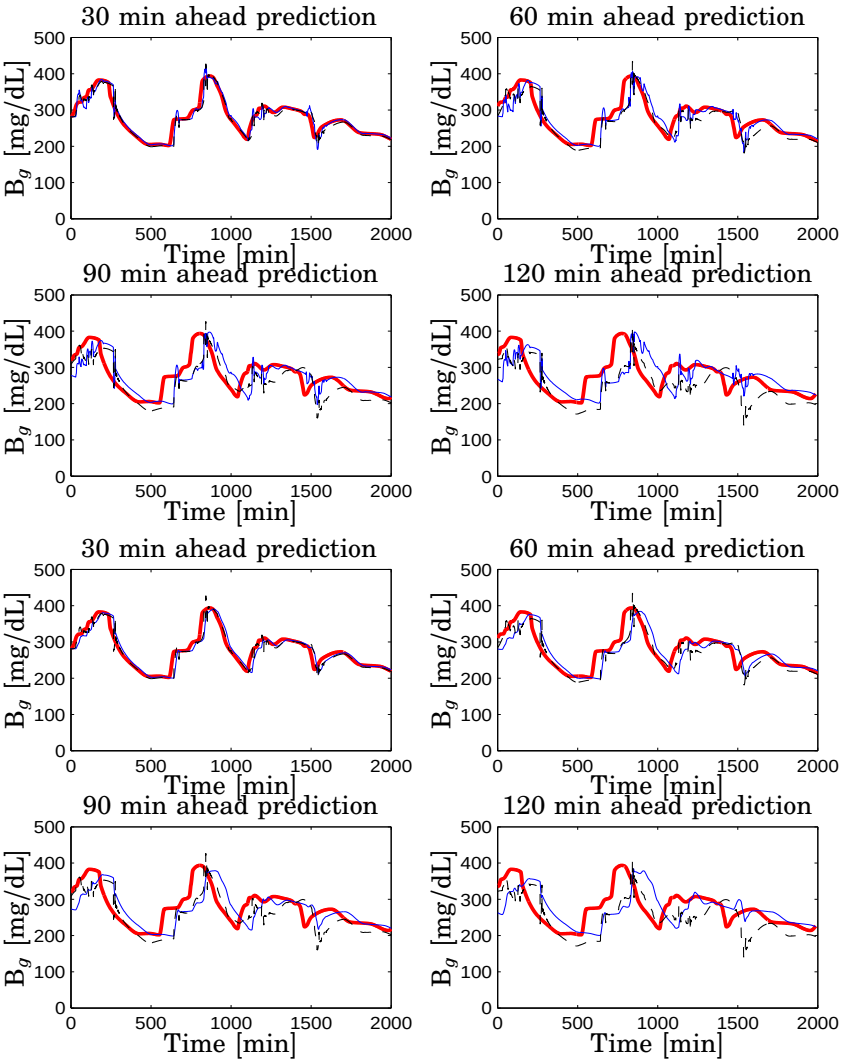

Figure 5.1 Patient CHU0102. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual plasma glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

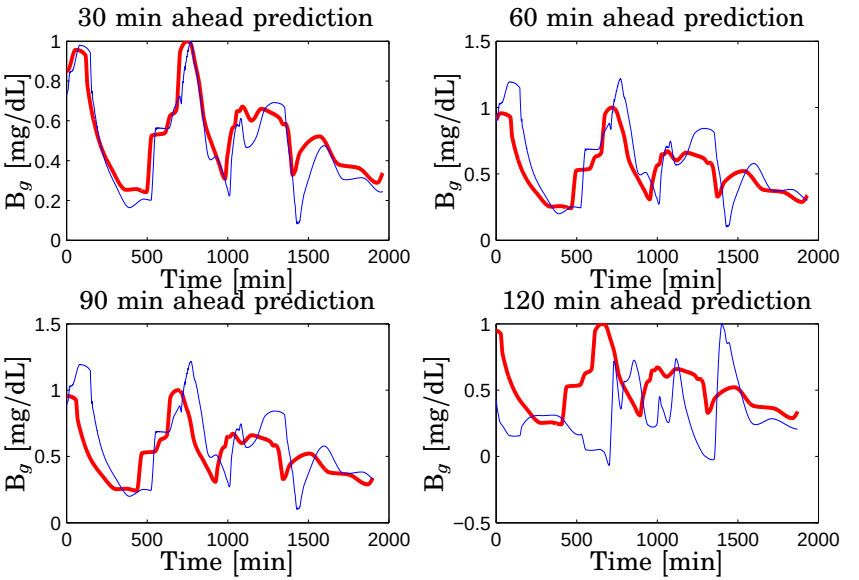

Figure 5.2 Patient CHU0102. Evaluation on validation data. PBSID-based predictor (thin) and measured plasma glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120- minutes ahead prediction. Notice that data was normalized.

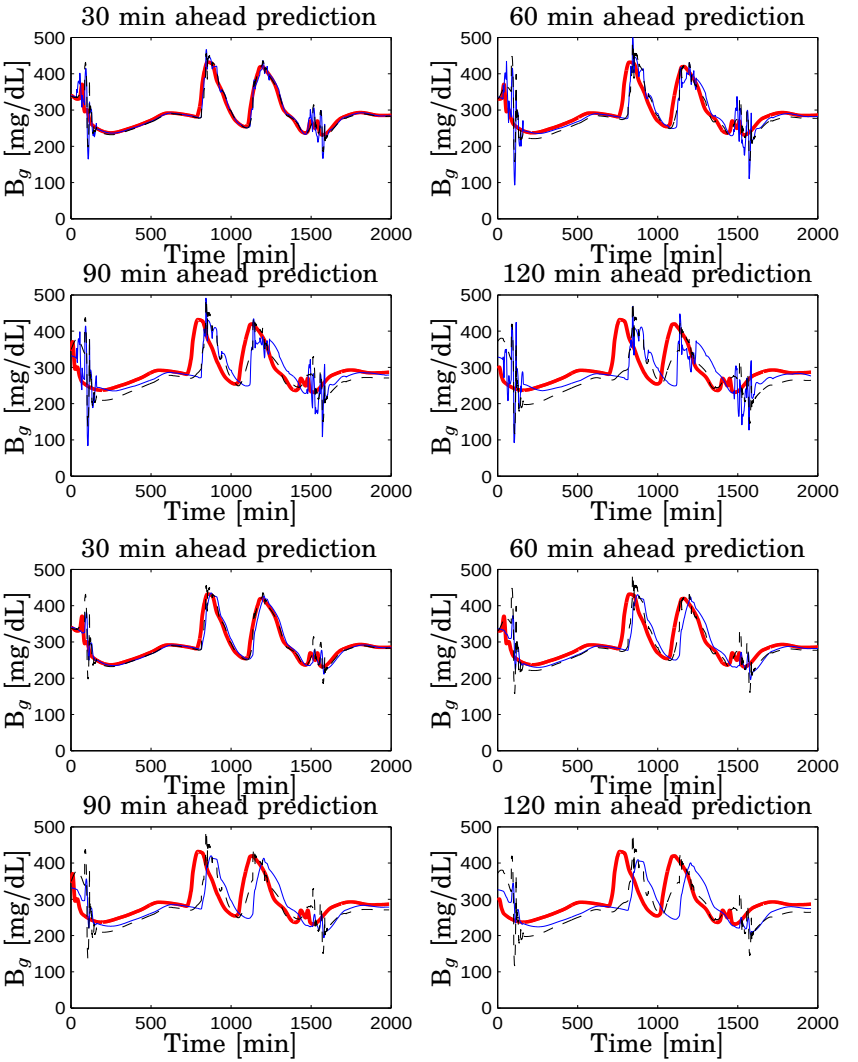

Figure 5.3 Patient CHU0107. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual plasma glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

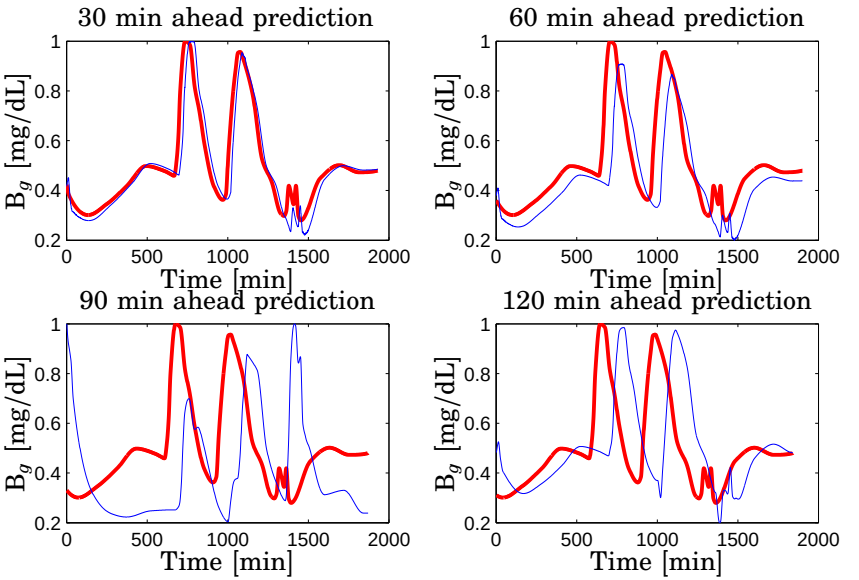

Figure 5.4 Patient CHU0107. Evaluation on validation data. PBSID-based predictor (thin) and measured plasma glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction . Notice that data was normalized.

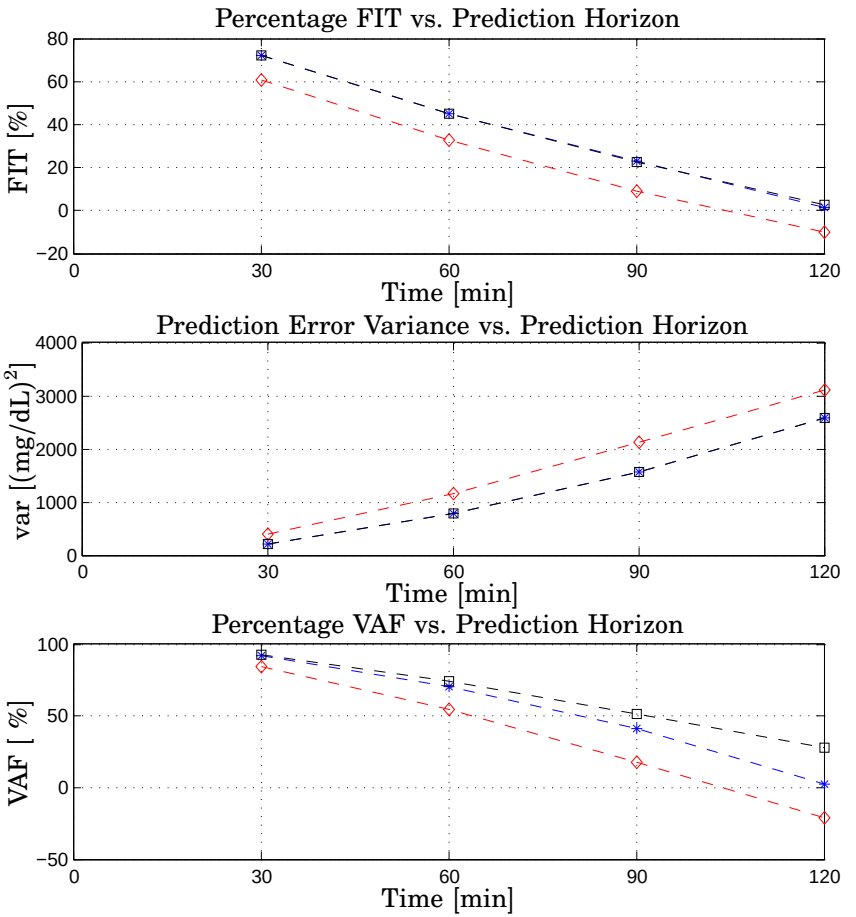

Figure 5.5 Patient CHU0102. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

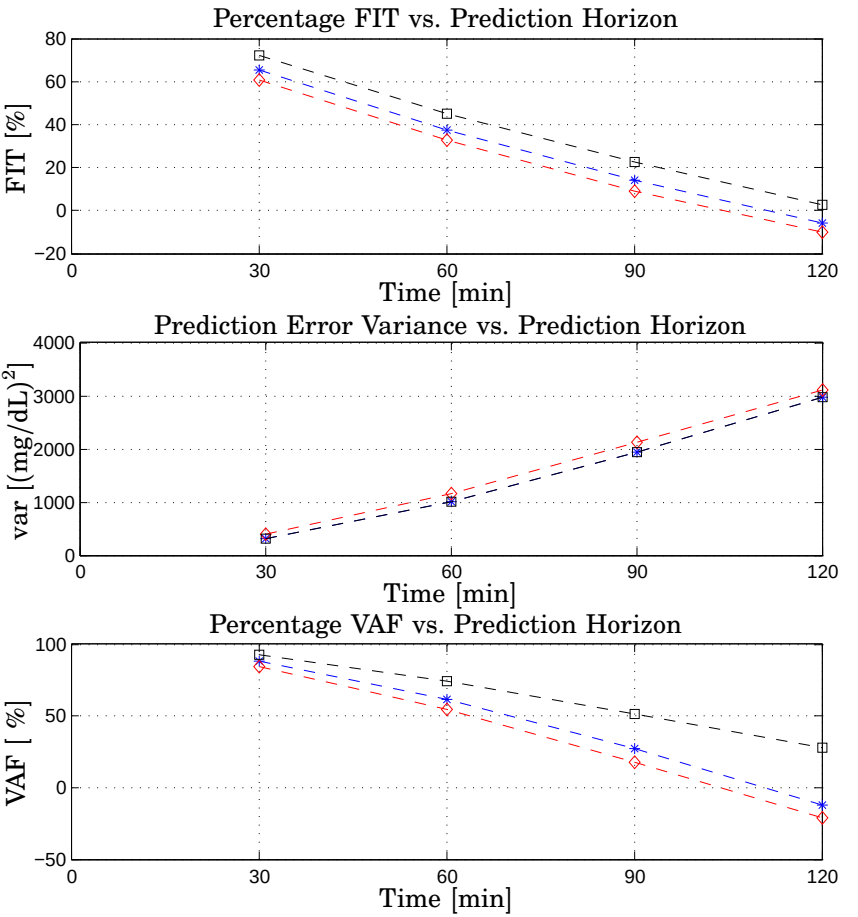

Figure 5.6 Patient CHU0102. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

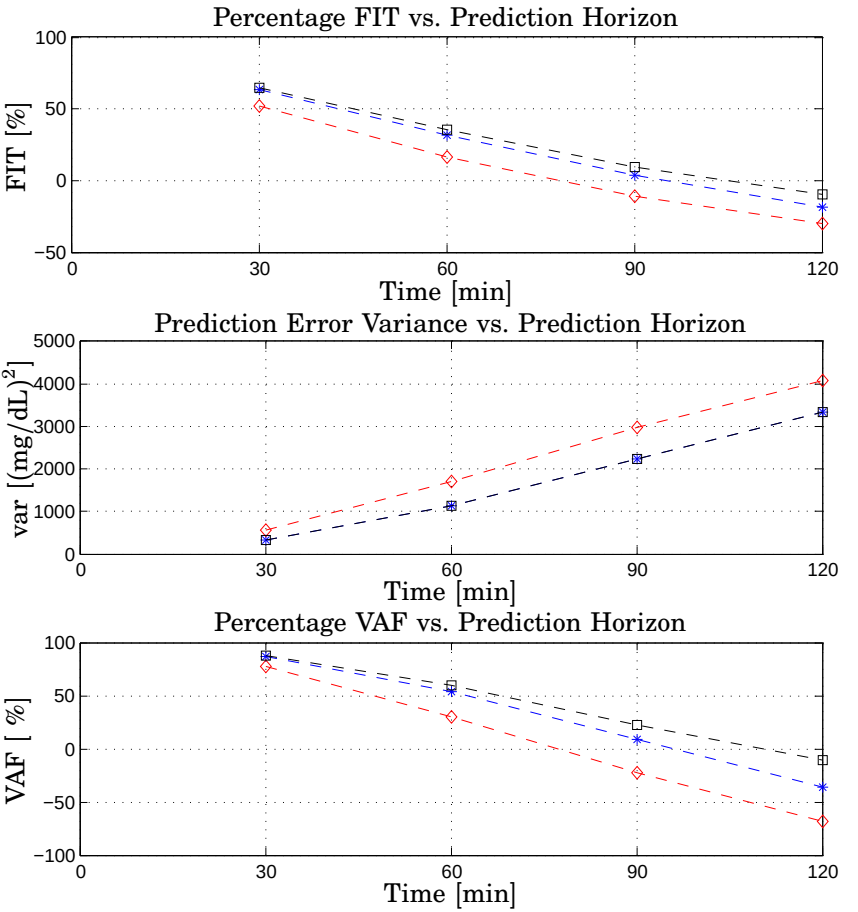

Figure 5.7 Patient CHU0107. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[\mathrm{mg/dL}]^2$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

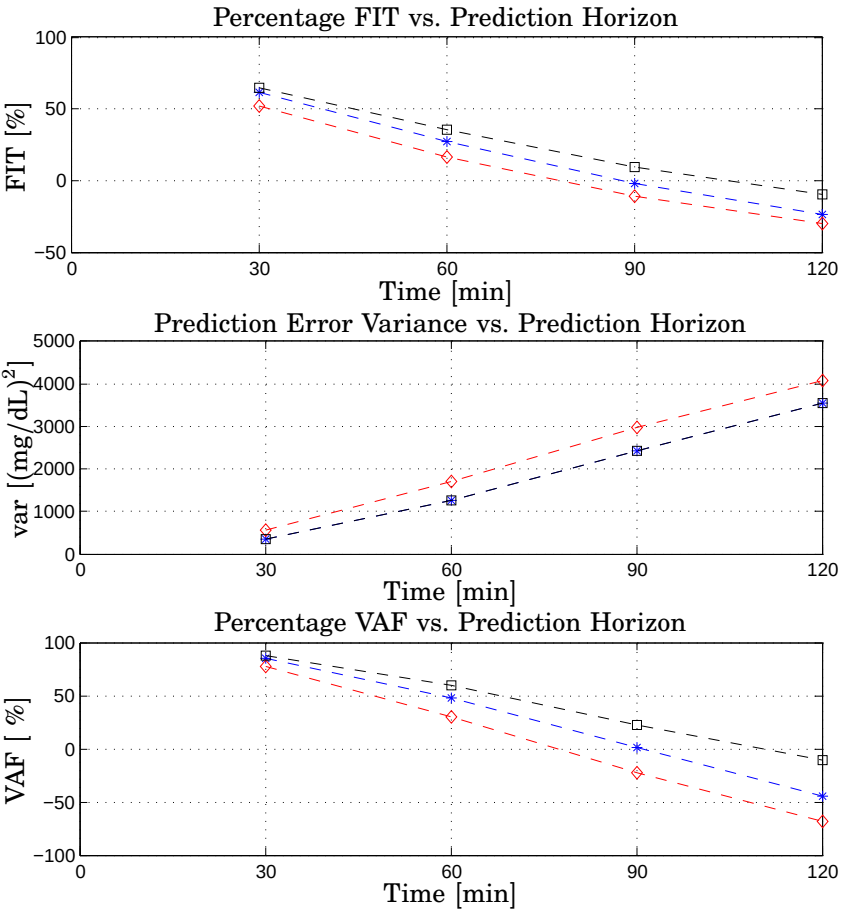

Figure 5.8 Patient CHU0107. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

Table 5.1 Predictor performance evaluation: comparison between 3rd-order ARX-based, regularized 3rd-order ARX-based and ZOH. Percentage FIT [%] vs. Prediction Horizon [min] on validation data without vital signs.

Patient ID	predictor	30[min]	60[min]	90[min]	120[min]
102	ARX	72.0738	45.2518	23.0036	1.3685
	REG	65.4577	37.3522	14.1688	-5.8967
	ZOH	60.6483	32.5733	9.0154	-10.1487
103	ARX	62.8013	26.1761	3.9531	-11.7326
	REG	55.6851	25.4080	7.0058	-7.8284
	ZOH	50.6296	20.8109	-1.3747	-21.3287
104	ARX	76.5251	45.5433	21.2761	4.2151
	REG	62.7411	30.6934	8.6147	-4.9927
	ZOH	52.1479	19.1685	-4.9330	-22.7172
105	ARX	52.3288	10.3134	-0.5414	-19.0002
	REG	44.1775	1.8077	-16.8959	-24.6196
	ZOH	47.2140	14.5802	-6.1496	-18.9310
106	ARX	58.9561	14.4690	-20.4642	-40.1471
	REG	56.2513	15.6128	-13.0854	-34.0590
	ZOH	44.0662	9.7080	-8.2996	-15.4517
107	ARX	63.5151	31.3665	3.4278	-18.4043
	REG	61.6291	27.0496	-1.7872	-23.3984
	ZOH	52.2875	16.1428	-10.6265	-29.7901
115	ARX	75.5011	46.5069	15.3315	-9.3268
	REG	64.5460	31.7013	2.7374	-17.5544
	ZOH	63.6523	34.7217	12.3847	-3.4325
120	ARX	73.9229	42.3381	17.7735	2.5326
	REG	66.2138	35.6965	13.4499	-1.4624
	ZOH	58.4849	27.0355	4.3175	-12.5778
130	ARX	64.7631	34.3644	8.6831	-7.5523
	REG	64.0173	34.4791	8.9728	-8.9190
	ZOH	58.8014	29.0776	6.3448	-9.5646

Table 5.2 Predictor performance evaluation: comparison between 3rd-order ARX-based, regularized 3rd-order ARX and ZOH. Prediction Error Variance $1.0e+03[(\mathrm{mg/dL})^2]$ vs. Prediction Horizon [min] on validation data without vital signs.

Patient ID	predictor	30[min]	60[min]	90[min]	120[min]
102	ARX	0.2039	0.7879	1.5720	2.5875
	REG	0.3066	1.0142	1.9458	2.9806
	ZOH	0.3981	1.1676	2.1253	3.1139
103	ARX	0.1804	0.7037	1.1953	1.5976
	REG	0.2603	0.7300	1.1326	1.4978
	ZOH	0.3323	0.8545	1.3988	2.0007
104	ARX	0.2535	1.3565	2.8313	4.1754
	REG	0.6398	2.2014	3.8170	5.0028
	ZOH	1.0513	2.9976	5.0470	6.8949
105	ARX	0.3492	1.2418	1.6122	2.2878
	REG	0.4712	1.4569	2.1054	2.4369
	ZOH	0.4306	1.1233	1.7245	2.1461
106	ARX	0.4155	1.8881	3.8203	5.2270
	REG	0.4902	1.8478	3.3637	4.7946
	ZOH	0.7973	2.0763	2.9834	3.3862
107	ARX	0.3238	1.1377	2.2369	3.3429
	REG	0.3560	1.2637	2.4333	3.5478
	ZOH	0.5535	1.7088	2.9718	4.0885
115	ARX	0.0984	0.4118	0.9226	1.4153
	REG	0.2083	0.6956	1.2786	1.7421
	ZOH	0.2423	0.7786	1.3951	1.9321
120	ARX	0.2458	1.2124	2.4797	3.5184
	REG	0.4096	1.5034	2.7480	3.8207
	ZOH	0.6141	1.8968	3.2605	4.5099
130	ARX	0.6436	2.1529	4.0061	5.4105
	REG	0.6675	2.1219	3.9200	5.4545
	ZOH	0.8868	2.6279	4.5825	6.2708

Table 5.3 Predictor performance evaluation: comparison between 3rd-order ARX-based, regularized 3rd-order ARX and ZOH. Percentage VAF [%] vs. Prediction Horizon [min] on validation data without vital signs.

Patient ID	predictor	30[min]	60[min]	90[min]	120[min]
102	ARX	92.2062	70.2658	40.9557	2.7404
	REG	88.2775	61.7241	26.9164	-12.0372
	ZOH	84.5496	54.6813	17.5108	-20.8571
103	ARX	86.8641	49.4839	14.8852	-16.0313
	REG	81.0465	47.5924	19.3499	-8.7841
	ZOH	75.6258	37.3160	-2.6105	-46.7695
104	ARX	94.5367	70.8960	39.9349	11.7428
	REG	86.2130	52.7679	19.0245	-5.7465
	ZOH	77.1072	34.7272	-9.8982	-50.1352
105	ARX	77.5966	21.2595	-0.7588	-40.9952
	REG	69.7671	7.6213	-31.5854	-50.1872
	ZOH	72.2686	27.6622	-11.0544	-38.2070
106	ARX	83.9405	27.7217	-45.0513	-95.5528
	REG	81.0527	29.2656	-27.7140	-79.3786
	ZOH	68.7340	18.5836	-16.9867	-32.7826
107	ARX	86.7048	53.4476	8.9546	-36.3918
	REG	85.3843	48.2915	0.9607	-44.7506
	ZOH	77.2643	29.8063	-22.0801	-67.9497
115	ARX	94.1385	73.1836	35.4365	-2.4380
	REG	87.5883	54.7063	10.5225	-26.0917
	ZOH	87.0342	58.3385	25.3581	-3.3786
120	ARX	93.2002	66.7512	32.4208	5.2581
	REG	88.6705	58.7715	25.1090	-2.8830
	ZOH	82.7652	46.7650	8.4905	-26.5775
130	ARX	87.7113	58.2864	21.8351	-4.5036
	REG	87.2550	58.8866	23.5145	-5.3532
	ZOH	83.0280	49.7050	12.2986	-20.0131

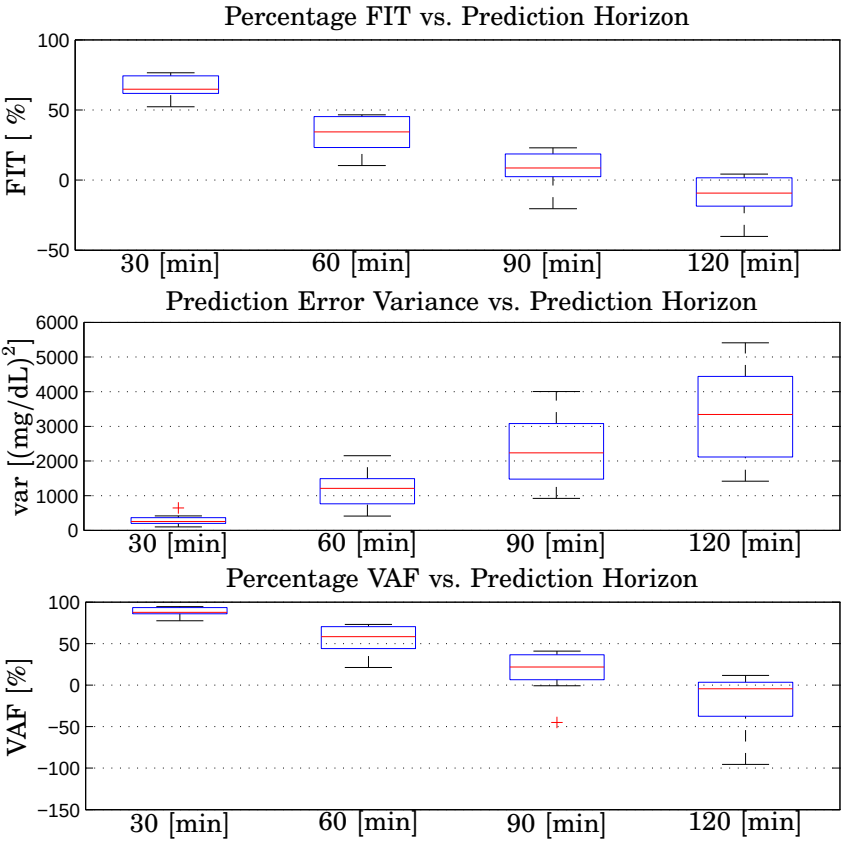

Figure 5.9 3rd-order ARX-based predictor performance evaluation. Population results on validation data. *Top* Percentage FIT vs. Prediction Horizon [min]; *Center* Prediction Error Variance $[\mathrm{mg/dL}]^2$ vs. Prediction Horizon [min]; *Bottom* Percentage VAF vs. Prediction Horizon [min]. Each box presents results over the population considered. The central mark is the median, the edges of the box are the 25th and the 75th percentiles.

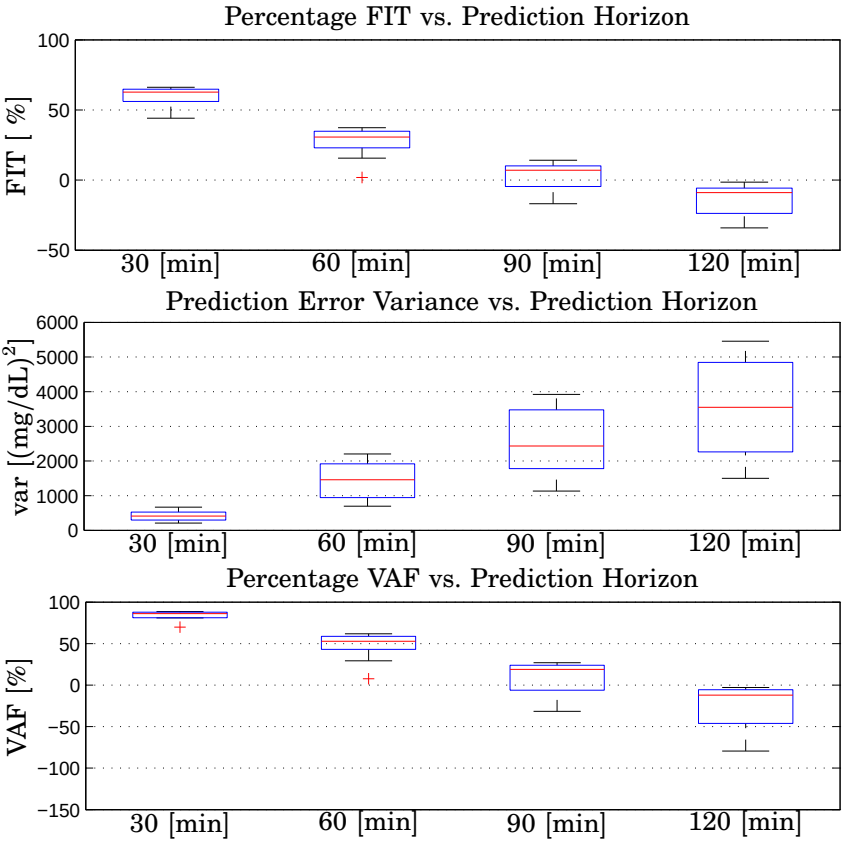

Figure 5.10 Regularized 3rd-order ARX-based predictor performance evaluation. Population results on validation data. *Top* Percentage FIT vs. Prediction Horizon [min]; *Center* Prediction Error Variance $[\mathrm{mg/dL}]^2$ vs. Prediction Horizon [min]; *Bottom* Percentage VAF vs. Prediction Horizon [min]. Each box presents results over the population considered. The central mark is the median, the edges of the box are the 25th and the 75th percentiles.

6

Discussion

6.1 Result of experiments

This thesis exploited a unique dataset which has being collected within a major European project [DIAdvisor, 2011]. Albeit being in a controlled environment, the subjects participating in the study experienced both hypo- and hyperglycemic events, proving glucose regulation hard to accomplish. The glucose curves showed diurnal variations and wide excursions over the 75 hours. When hypoglycemia occurred, this situation being common to all the patients, extra carbohydrates apart from the standard meals were given. Similarly, hyperglycemia was treated with extra insulin intakes. As it is common practice, the majority of the subjects bolused just before being served the meal. Since meal intake and the insulin injections have opposite effects on the blood glucose level, the estimation of the contribution from each of these inputs is difficult when both are active at the same time.

As for the sensor signals, the FreeStyle Navigator^{TM} continuous glucose monitor traces resulted often in poor agreement with the blood glucose reference obtained from the laboratory analysis. In particular, important offsets were present. The system measures glucose in the interstitial fluid (ISF), i.e., in between the body cells. Movements of nutrients, oxygen and glucose from the blood into the cells happen across the ISF; therefore, during times of rapid change in blood glucose, e.g. after eating, dosing insulin, or exercising, differences in glucose

measurement between interstitial fluid and finger-stick measurements are expected to be observed. In the examined trial, however, it was noted that in 40-50% of the traces the sensor was accurate at low glucose levels but inaccurate at high levels, a fact that cannot be explained by the plasma-to-interstitium dynamics only and may be explained by poor/difficult device calibration. The manufacturer also made available the intermediate signal calculated by the system at each minute. Whether or not this may overcome the limitations introduced by the physiological interstitium-to-plasma dynamics it is still unclear, since a model of how the signal relates to plasma glucose is missing. The VivoMetrics Clinical LifeShirt® [Grossman, 2004] was not worn by the subjects at all times, so various segments of the vital signs were not represented in the database.

The rich collection of blood samples provided reference plasma glucose values and plasma insulin values not available elsewhere and was therefore exploited to the purposes of the thesis.

6.2 Modeling

Over the last decades, models describing the insulin-to-glucose system dynamics were developed for the purpose of simulation and glycemic control, the approach being physiology based [Dalla Man *et al.*, 2007], [Roy and Parker, 2006], [Makroglou *et al.*, 2006]. Recently the problem of identifying such a model has been tackled from a data-driven perspective mainly using simulated data from models in the literature [Palerm *et al.*, 2006], [Finan *et al.*, 2006]. Indeed, fitting actual T1DM subject data to the models has been treated to a much less extent (e.g., [Ståhl and Johansson, 2008], [Ståhl and Johansson, 2009], [Finan *et al.*, 2007]) given the difficulties in gathering appropriate patient records.

The autospectra (power spectra) showing the frequency contents of the signals investigated and the coherence spectra between the inputs and the controlled variable were calculated [Johansson,1993]. Recall that a coherence spectrum can be interpreted as a correlation analysis (or signal-to-noise analysis) made for each frequency. A large absolute value close to 1 indicates that the input and output are correlated. A coherence value of 0.5 denotes that half of the output variation may be explained by variations in the stimulus input. As predictable, the

data collected offered poor model input excitation despite the careful selection of the subjects, because of the correlation between food intake and consequent insulin injection.

Throughout the work, the glucose flux in the bloodstream after intestinal absorption and the total, i.e., fast-acting and slow-acting, insulin flux in the bloodstream were considered as input variables. As far as the glucose absorption modeling is concerned, it is a well known fact that not only the size of the meal but also the composition of the meal affects the digestion dynamics (see e.g. [Brouns *et al.*, 2005]). Unfortunately, detailed quantitative information on meal composition was not available in the data set considered for this thesis. In absence of such information, all sources of carbohydrates were assumed to be equal. Population mean values reported in the literature were used for the parameters appearing in the meal model equations 4.1, 4.2, 4.3 and in the insulin kinetics model equations 4.4, 4.6, 4.7, thus disregarding the inter-personal variability. This explains the mismatch between the actual plasma insulin concentration profile and that obtained with the simulation model (Eqs. 4.4, 4.6, 4.7). Given the frequently drawn blood samples, it was decided to use the actual (interpolated and uniformly resampled) insulin assays for the identification of both models and predictors, and to use the physiological insulin kinetics model at a later stage, to test the blood glucose response to 1 $[IU]$ of fast acting insulin.

Individual-specific models of low-complexity were identified from the collected data. Estimated model structures included autoregressive with exogenous inputs (ARMAX) models and state-space models. As far as the ARMAX structure is concerned, identification of the model parameters was accomplished by minimization of a quadratic prediction error criterion using the Matlab® System Identification Toolbox routine *armax.m*. The range of the orders n_a, n_b, n_c was empirically set to the interval $[1:10]$, while that of the inputs-output delays was $[1:3]$. The Akaike FPE was calculated for each of the configurations and used subsequently to rank the models according to their increasing FPE[Ljung, 1999]. Regarding the subspace identification techniques, two parameters having substantial influence on the quality of the resulting model needed to be chosen, namely the lengths of past and future horizons, representing the dimension of certain Hankel matrices constructed with the data. There are no simple rules for choosing

them, however, the knowledge of the application, that is blood glucose prediction up to 120-minutes ahead, provided an initial guess. The parameters were then tuned empirically: $p = f = 120$ in the N4SID algorithm [Van Overschee and De Moor, 1994], $s = 30$ in the PO-MOESP algorithm [Verhaegen, 1994] and $p = 60, f = 20$ in PBSID [Chiuso, 2007].

The estimated models were tested according to various criteria. Stability was the first requirement that a model needed to fulfill. Residual tests with the purpose of finding remaining correlations which indicate whether the model order is adequate were carried out. With adequate model order, the residual process is white only and of sufficiently small magnitude. The residual autocorrelation and cross correlation between the prediction errors and the input tests needed to give significant (99% confidence) validation with respect to changes of sign, independence of residuals, normality, and independence between residuals and input in order for the tset to be passed. Finally, qualitatively correct responses to inputs were guaranteed. From a quantitative point of view, according to clinicians and their experience gained from clinical trials, the average lowering effect of 1IU of fast insulin falls within 25-60 [mg/dL], with peak time $60 - 240$[min], depending upon the subject's resistance or sensitivity to insulin, whereas an ingestion of 10 [g] pure dextrose makes the blood glucose rising 15 [mg/dL], in 20 minutes at best. However, these requirements seemed hard to achieve and were not fulfilled by all models.

Tables 4.3, 4.4 and 4.5 report for each of the patients the identification methods leading to models not meeting one or more of the requirements in 4.4. When a criterion was not satisfied, the corresponding model was disregarded. Overall, from the tables it emerges that the main difficulties encountered while carrying out the modeling task were assuring white residuals and estimating physiological correct inputs to output transfer functions.

Last, as far as physical activity is concerned, given the results shown in this work, it may be concluded that it does not improve significantly the predictions. One explanation could be that during the in-hospital tests, the intensity of such exercise was very low.

6.3 Prediction

In addition to model-based predictors, short-term, i.e., 30, 60, 90, 120 minutes ahead, direct predictor identification was pursued. Predictor coefficients were identified from patient data. Third-order ARX predictors were estimated by means of least-squares estimation. Smoothing the least-squares solution was accomplished by introducing a regularization coefficient ρ: when $\rho = 0$ no regularization is performed, as ρ increases the predictor coefficients are constrained resulting in a smoother predicted profile. As far as the subspace-based linear multivariate predictor is concerned, predictor coefficients were directly obtained from input-output data, with no need of the determination of the model structure. From an implementation point of view, the approach is attractive, amounting only to LQ decompositions of appropriately organized input-output Hankel matrices. The choice of two parameters was required to the user: the length of the past horizon p and the length of the future horizon f. In particular, parameter p representing the size of the past Hankel data matrices has strong connections with model order, and was given by the user from knowledge or intuitions on the system dynamics. Moreover, the lower bound on p was represented by the maximum prediction horizon one wish to investigate. Experience from simulations suggested that for prediction a small number of samples ahead it suffices to choose p double of the expected model order. It may be expected, however, that the longer the past horizon, the better the performances. Instead, the investigation carried out suggested instead the contrary. Even if apparently less crucial also the choice of the future horizon matters and affects the quality of the estimation.

The quality of the predictors developed was assessed by mathematical metrics in order to quantify the error between the predicted blood glucose profile vs. the actual ones. Specifically, predictions were evaluated with respect to

- FIT %
- prediction error variance [mg/dL]
- VAF %

and qualitative assessments concerning glucose-trends detection. Indeed, in diabetes management, the perhaps most important feature

for a predictor is the ability of capturing hypoglycemias and hyperglycemias, rather than being correct in the normo-glycemic range. The performances were compared to those achieved with the zero-order-hold and to those obtained with a third-order ARMAX model. The 3rd-order ARX predictor outperformed the ZOH with respect to all metrics and for all patients, whereas the regularized ARX failed with patient 105, probably because of a bad choice of the regularization coefficient. Moreover, the ARX predictor and its regularized version failed when compared to the ARMAX-based predictors for patients 104 and 115.

7

Conclusions and Future Work

7.1 Conclusions

This thesis deals with linear modeling and short-term prediction in diabetes physiology. Specifically, data-driven techniques were investigated to the purpose of the DIAdvisor^{TM} tool application [DIAdvisor, 2011] and evaluated for type 1 diabetes mellitus records belonging to a population of 9 subjects in hospital conditions.

Modeling

An individual-specific, physiological relevant model of the glucose-insulin interaction subsystem was identified from each of the subjects data using prediction error methods and subspace-based methods. Inputs to the models were:

- interpolated total plasma insulin concentration I [mIU/L] from drawn blood samples;
- plasma glucose rate of appearance R_a [mg/kg/min] after carbohydrate intestinal absorption;
- squared heart rate H_r [(beats/min)^2];
- squared acceleration a [a.u.];

and the output:

- interpolated blood glucose B_g [mg/dL] from drawn blood samples

ARMAX models of order in the range [3 : 6] satisfied all the criteria required, specifically:

- stability;
- white residuals;
- physiologically sensible responses to 1 [IU] of insulin and 10 [g] of carbohydrates

and were therefore selected for inclusion in the advisory tool. However, the additional requirements on model-based predictor performances were met only partially. Indeed, whereas a value of VAF $\geq 50\%$ on 60-minutes-ahead model-based prediction on validation data was achieved by all the models except the 5th-order ARMAX model for patient 105 and the 6th-order ARMAX model for patient 106, a value of FIT $\geq 50\%$ on 60-minutes-ahead model-based prediction on validation data was achieved by the 3rd-order ARMAX model identified from patient 104 data and the 6th-order ARMAX model identified from patient 115 data.

Prediction

Individual-specific short-term blood glucose predictors were identified from Type 1 Diabetes Mellitus subject data. Predictors coefficients were directly identified from the input-output data, without the intermediate model identification step. Inputs to the models were:

- interpolated total plasma insulin concentration I [mIU/L] from drawn blood samples;
- plasma glucose rate of appearance R_a [mg/kg/min] after carbohydrate intestinal absorption;

and the output:

- interpolated blood glucose B_g [mg/dL] from drawn blood samples

the structures investigated were third-order ARX-based predictors and their regularized version, and subspace-based multivariate predictors. Across the population, the performances of the proposed predictors

were superior to those achieved by projecting the last blood glucose value into the future, i.e., ZOH. However, the goal of a FIT value $\geq 50\%$ on 60-minutes-ahead prediction on validation data was not accomplished, while a VAF value $\geq 50\%$ on 60-minutes-ahead prediction on validation data was reached.

7.2 Future work

Experiment design

Poor model input excitation and input signal correlation were reported giving rise to the issue of ill-conditioning of the estimates. Whereas the importance of persistency of excitation was well recognized in the consortium, it posed problems in the ethical approval of the experimental protocol. Thus, further work is needed to investigate optimal experimental conditions and protocols in order to obtain data suitable for identification purposes without contributing to higher patient risk.

Modeling

Several topics call for attention:

First, in the meal simulation model, the development of new compartments accounting also for proteins, fat and fiber content of a meal is needed. A feasible alternative to that would be introducing the glycemic index G_i of each mixed meal in the already existing meal models from the literature (e.g. [Dalla Man *et al.*, 2006], [Dalla Man *et al.*, 2007]). Using the patient's diary and nutrition tables it is possible to estimate the G_i of each food intake. Hence, considering that lower G_i nutrients prolong the glucose rate of appearance in plasma, the glucose intestinal absorption model could be modified, e.g., making

the parameters b, c, k_{max} and k_{abs} functions of the G_i of the meal:

$$\begin{aligned} b &= 0.089 \cdot \log(GI) + 0.43 \\ c &= 0.3 \cdot \log(GI) - 0.61 \\ k_{max} &= 0.02 \cdot \frac{GI}{30} - 0.013 \\ k_{abs} &= 0.013 \cdot e^{0.02 \cdot GI} \end{aligned} \tag{7.1}$$

In addition, since patient's annotation on meal quantity and content are not always reliable, the problem of meal estimation to cope with forgotten meals and mistaken information needs to be addressed.

As far as the insulin modeling is concerned, grey-box identification of the parameters describing insulin pharmacokinetics/pharmacodynamics could be pursued, in order to fit the physiological models (e.g. [Wilinska *et al.*, 2005]) to the patients data.

A first-principle-based model quantifying the impact of physical activity on blood glucose excursion deserves future studies as well.

Further improvements of the state-of-the-art in identification, e.g., the ability of handling non-uniformly sampled data, reduction of sensitivity to initial conditions and automatic selections of model parameters, would be valuable. Most importantly, new identification methods tailored to the diabetes application need to be developed: constrained optimization could be used, for instance, to obtain physiologically correct responses to inputs. In doing so, it may be helpful to relate patient data such as BMI, insulin sensitivity and/or resistance to the insulin and glucose impulse responses.

Hybrid [Paoletti *et al.*, 2007] or Linear-Parameter-Varying (LPV) models [Bamieh and Giarre, 2002], [Verdult and Verhaegen, 2002] may capture the circadian variation of glucose in a better way than linear models, therefore this investigation is left to future work.

Prediction

Given the application, it is crucial that the predictors are correct in forecasting hypo- and hyperglycemic excursions, rather than being precise in the euglycemic range. A cost function similar to that presented in [Guerra *et al.*, 2011] could be an appropriate choice to penalize out-of-range glucose deviations.

More simulation work and theoretical research are needed to fully understand the properties of the presented multi-step predictors and to realize whether or not a safe and efficient employment in the DIAdvisor^{TM} is worth pursuing. In particular, future work will be devoted to the analysis of

- Statistical Errors
- Consistency
- Robustness
- Choice of user parameters
- Small sample size

Another topic deserving attention is that of adaptive PBSID-based multi-step multivariate prediction. Indeed, when the underlying true, yet unknown, model is changing over time an adaptive scheme may constitute a valuable step towards the development of an advisory tool capable of informing the patient at any time about the evolution of glycemia.

Last, it would be interesting to challenge linear model identification methods on data collected from patients in ambulatory conditions.

A

Patient data

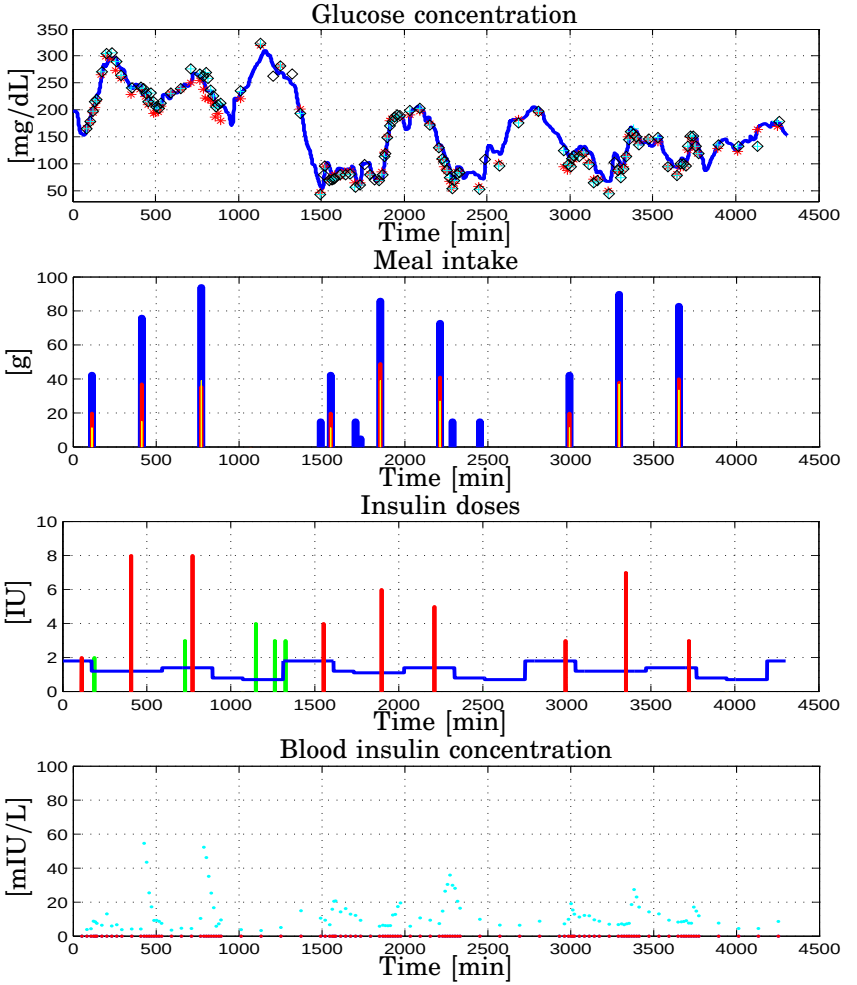

Figure A.1 Patient CHU0103 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

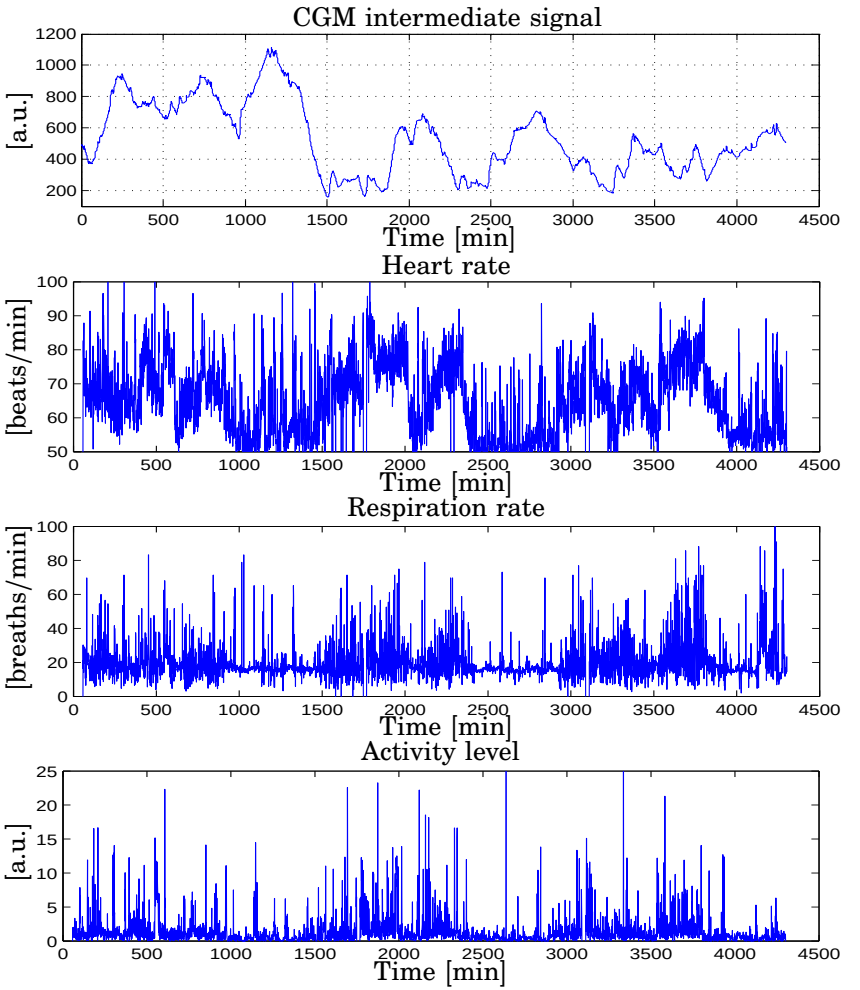

Figure A.2 Patient CHU0103 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

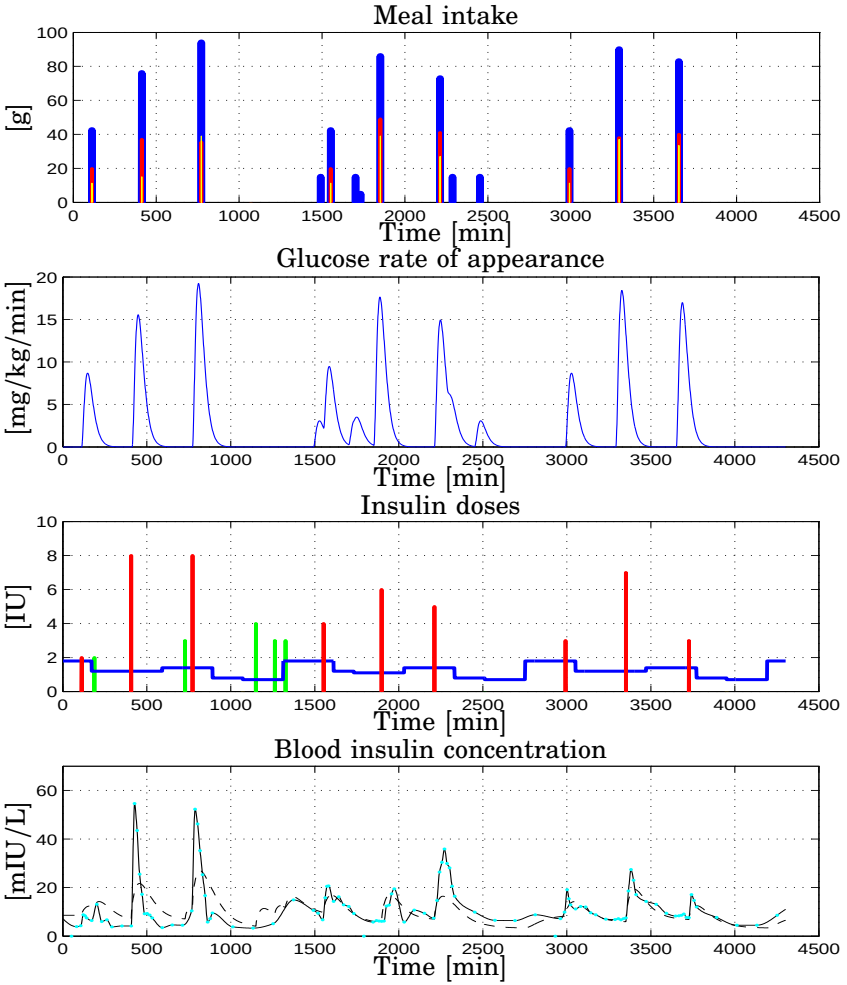

Figure A.3 Patient CHU0103 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood sample (cyan), from physiological model (dashed black)

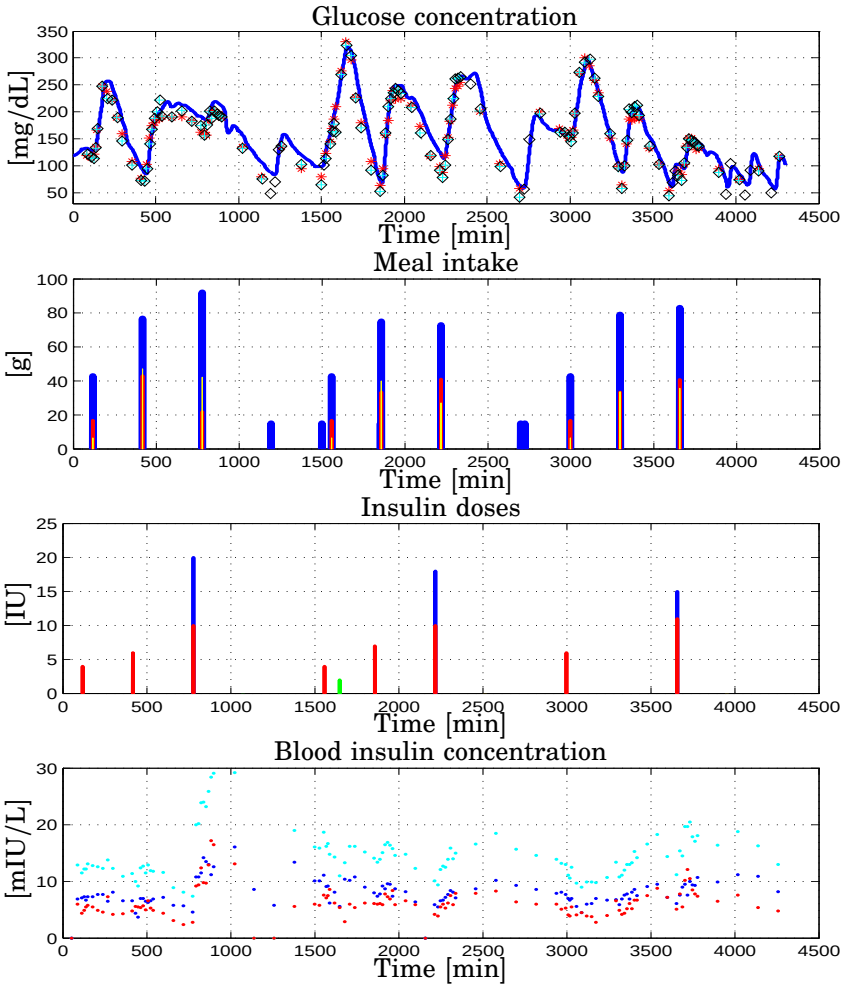

Figure A.4 Patient CHU0104 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

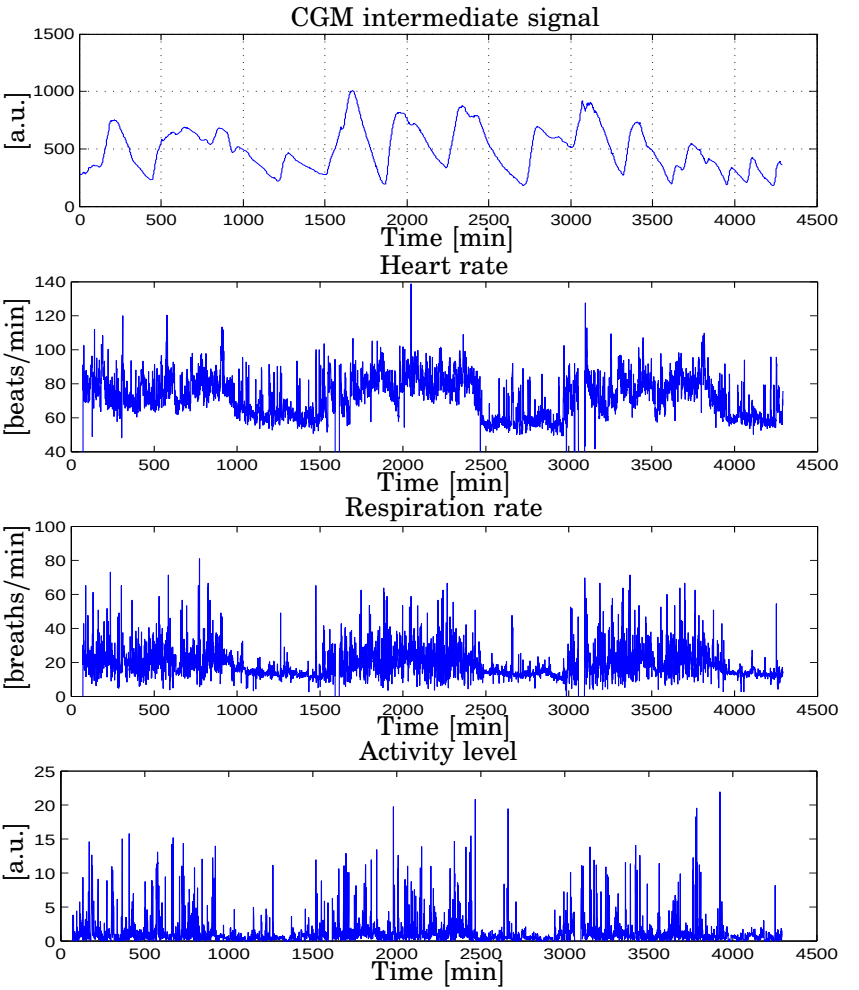

Figure A.5 Patient CHU0104 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

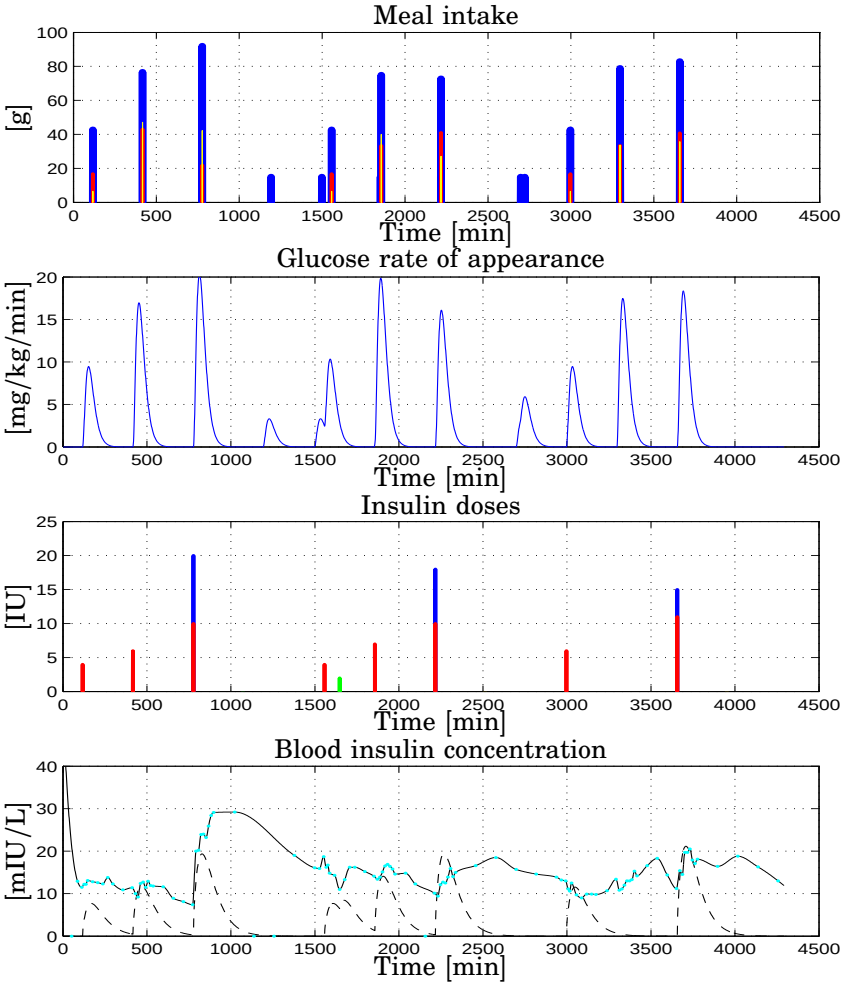

Figure A.6 Patient CHU0104 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

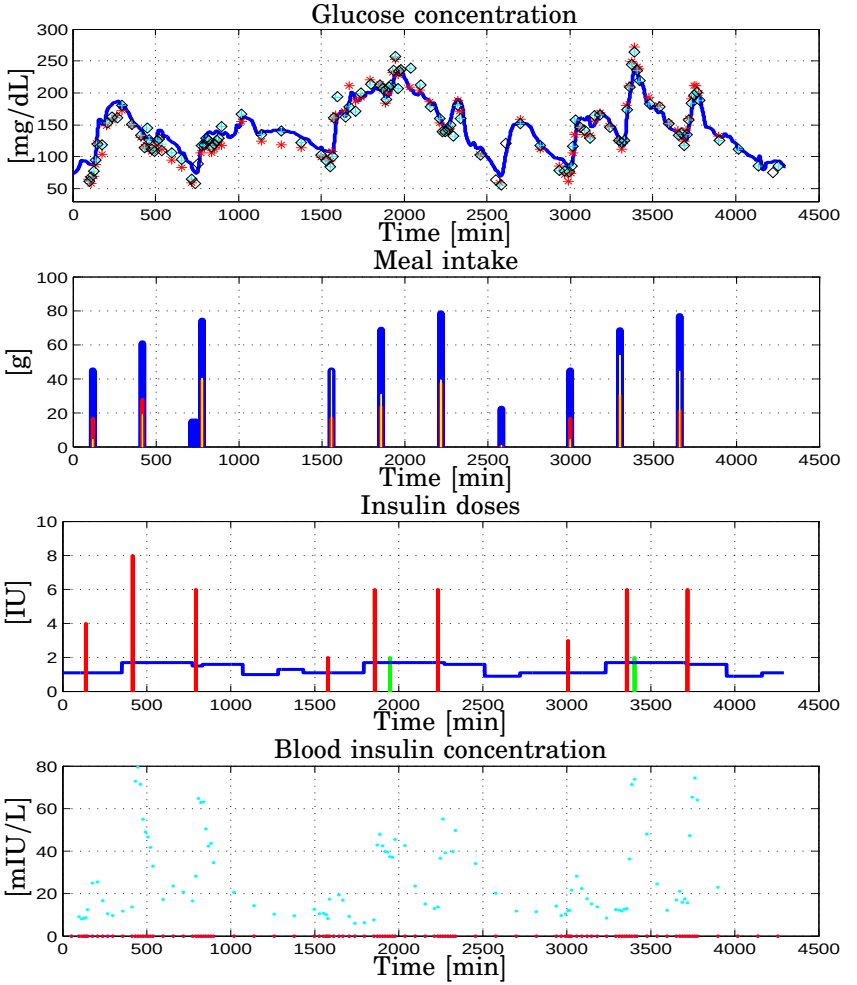

Figure A.7 Patient CHU0105 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

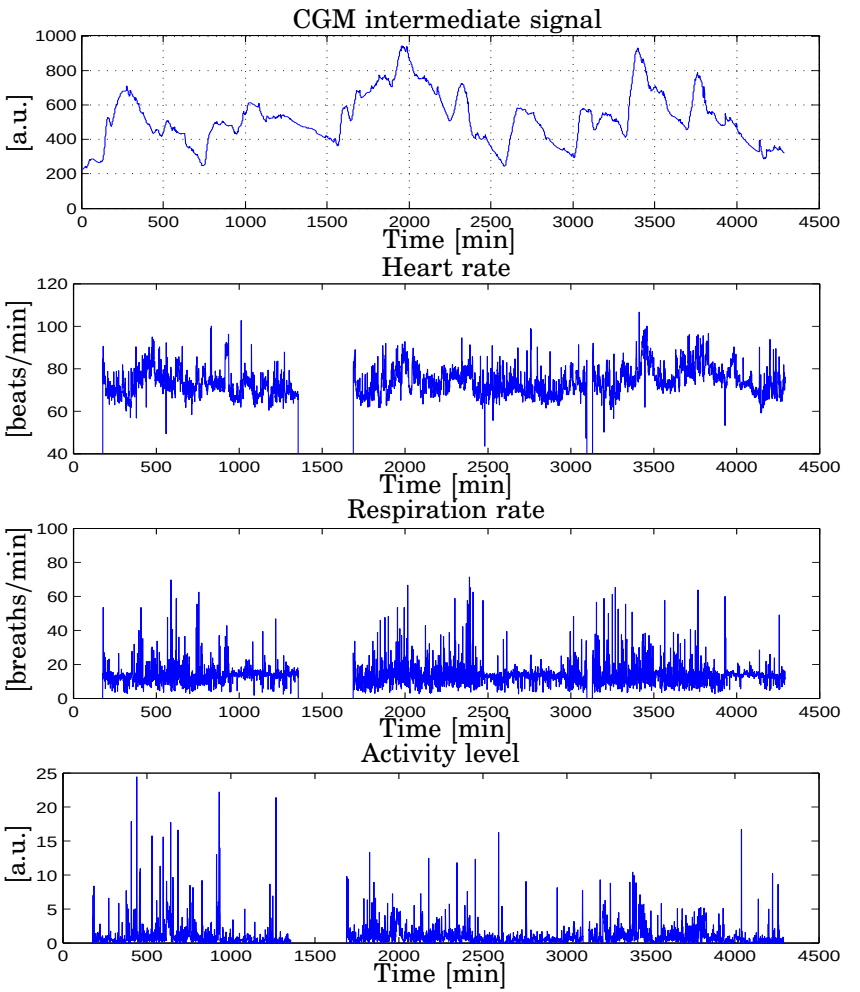

Figure A.8 Patient CHU0105 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

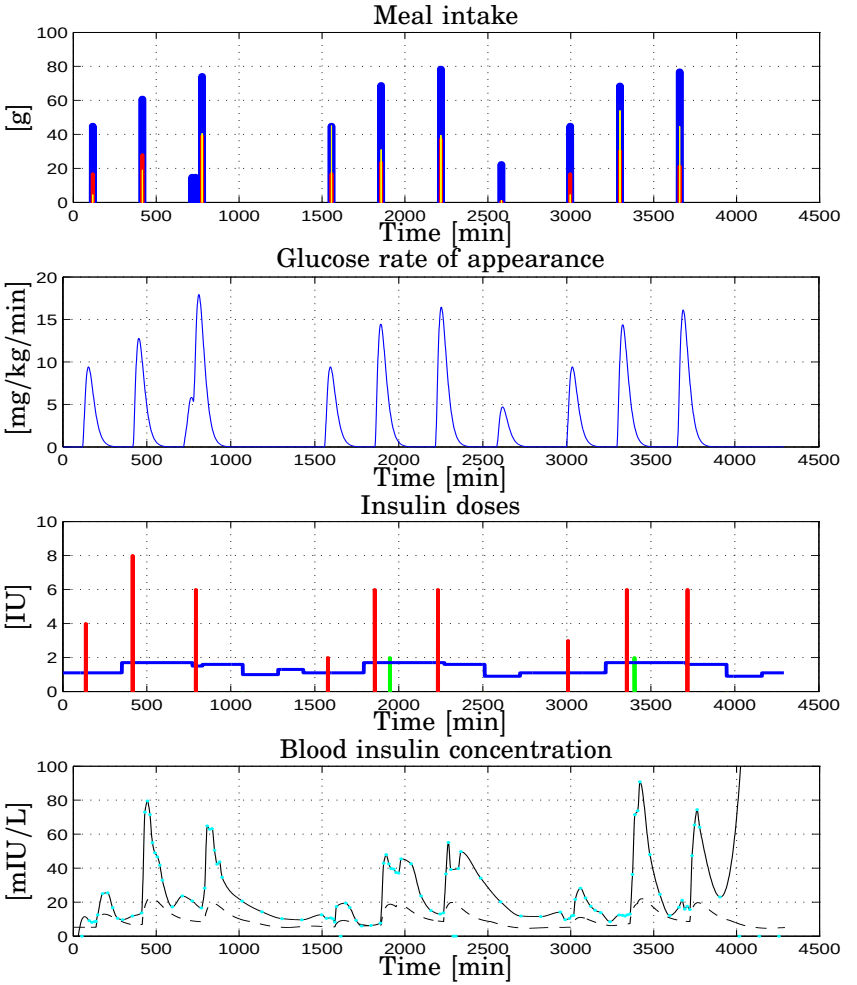

Figure A.9 Patient CHU0105 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

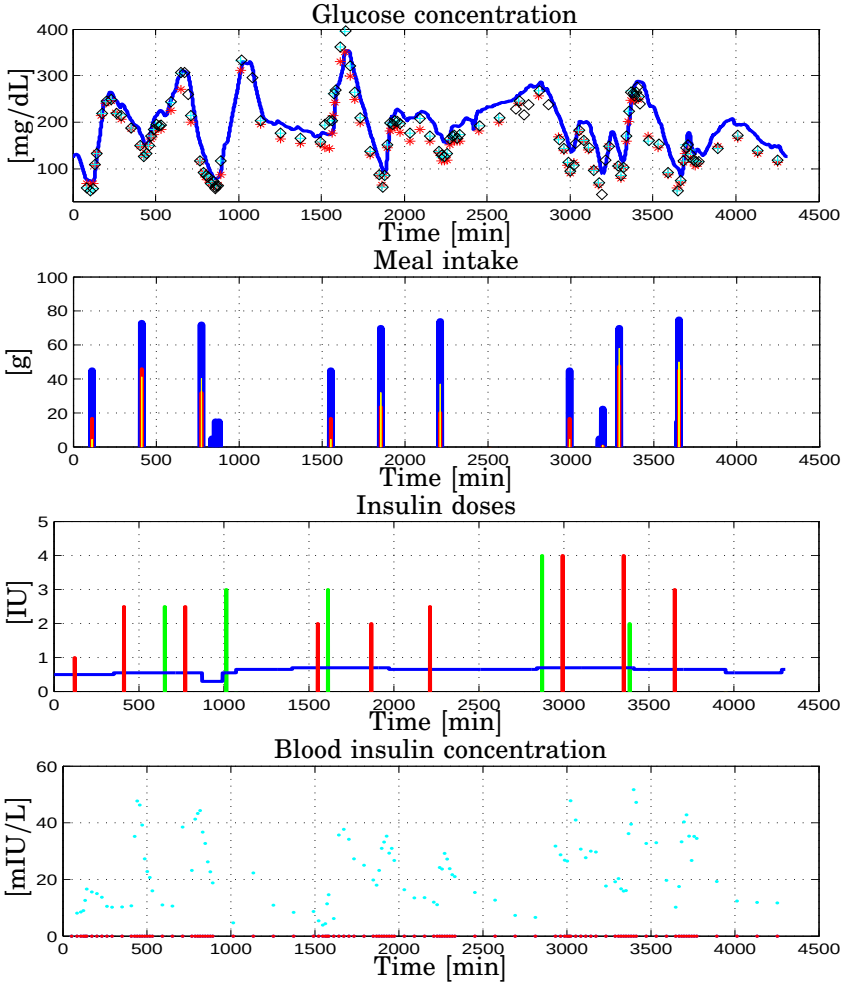

Figure A.10 Patient CHU0106 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

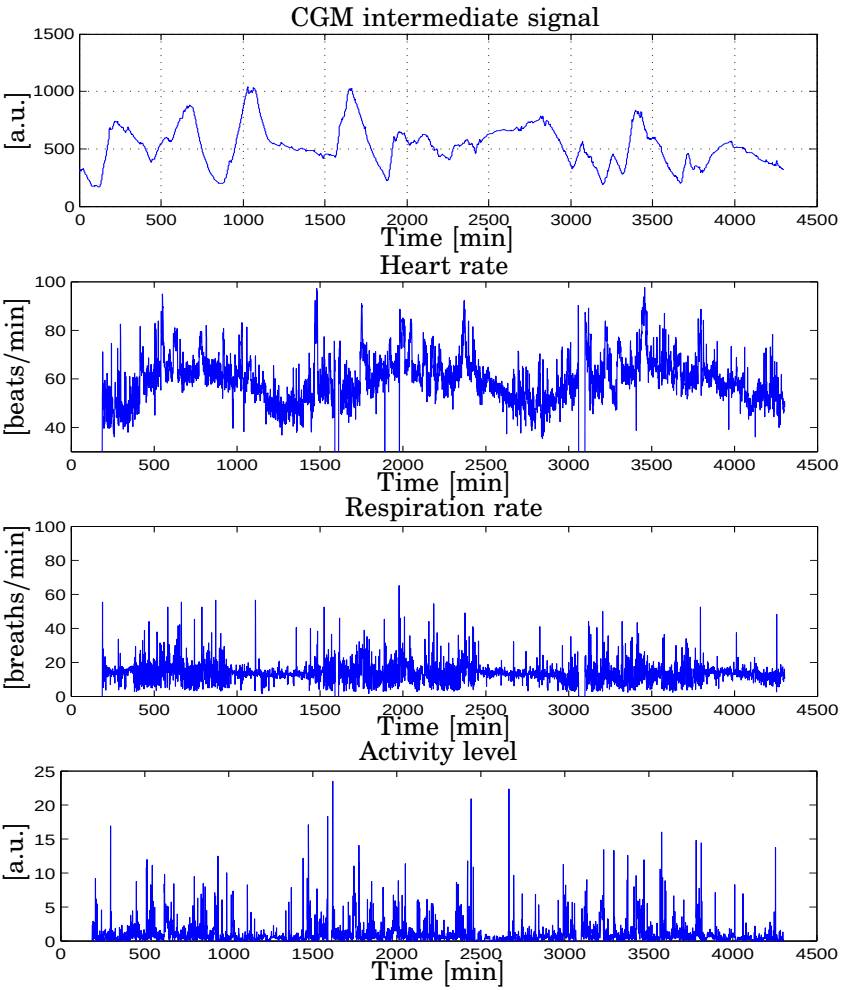

Figure A.11 Patient CHU0106 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

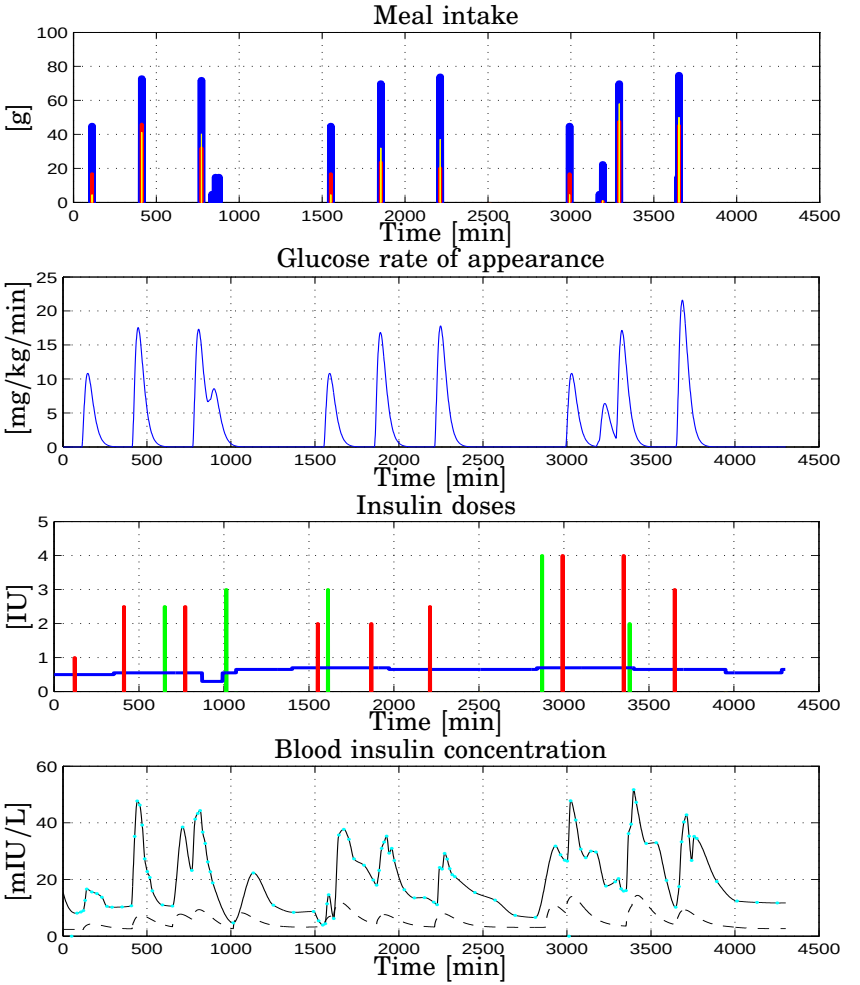

Figure A.12 Patient CHU0106 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

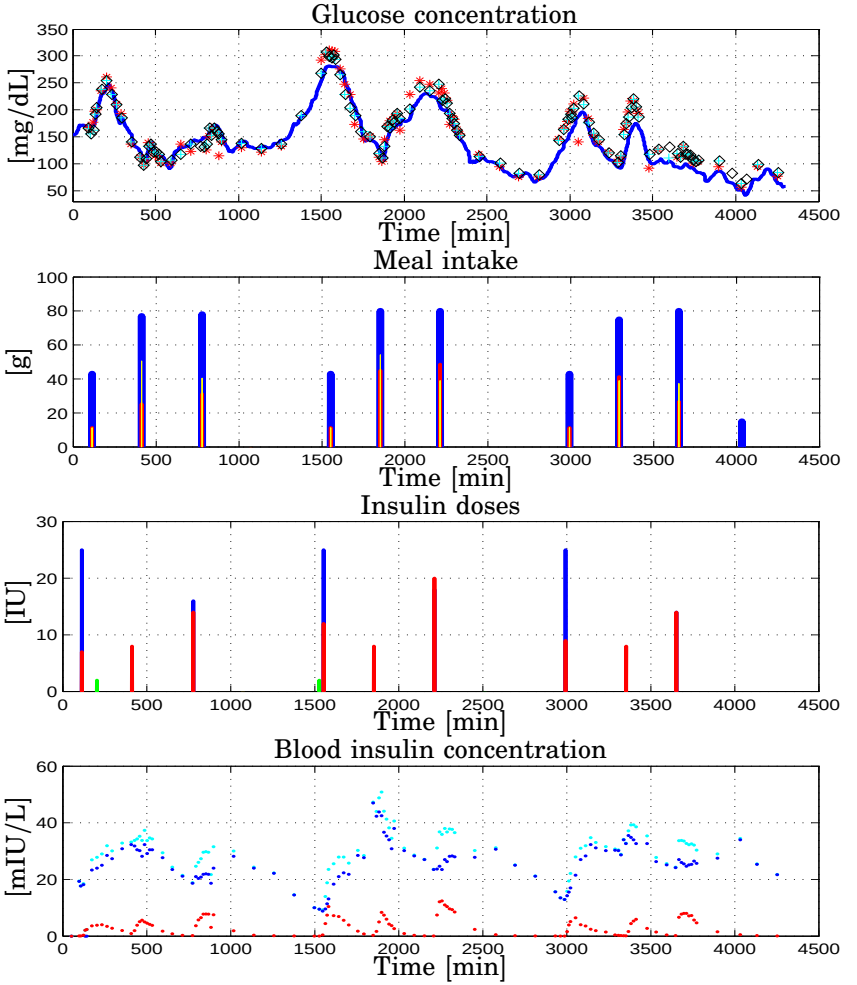

Figure A.13 Patient CHU0115 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

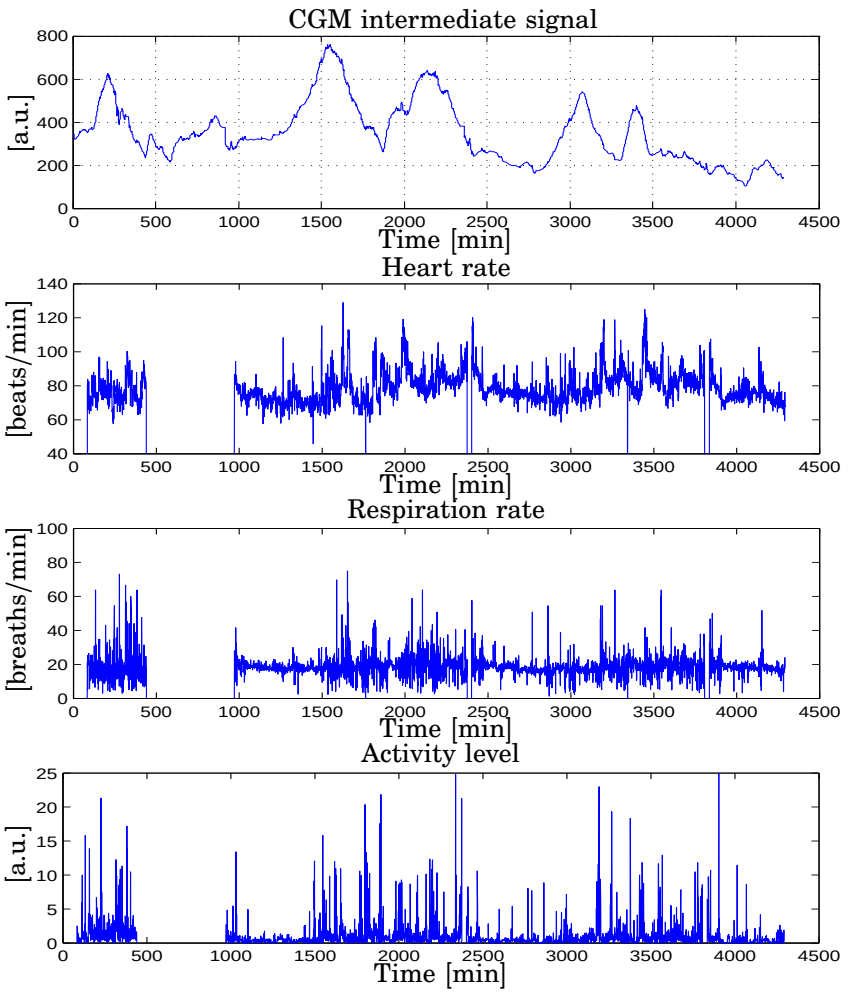

Figure A.14 Patient CHU0115 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

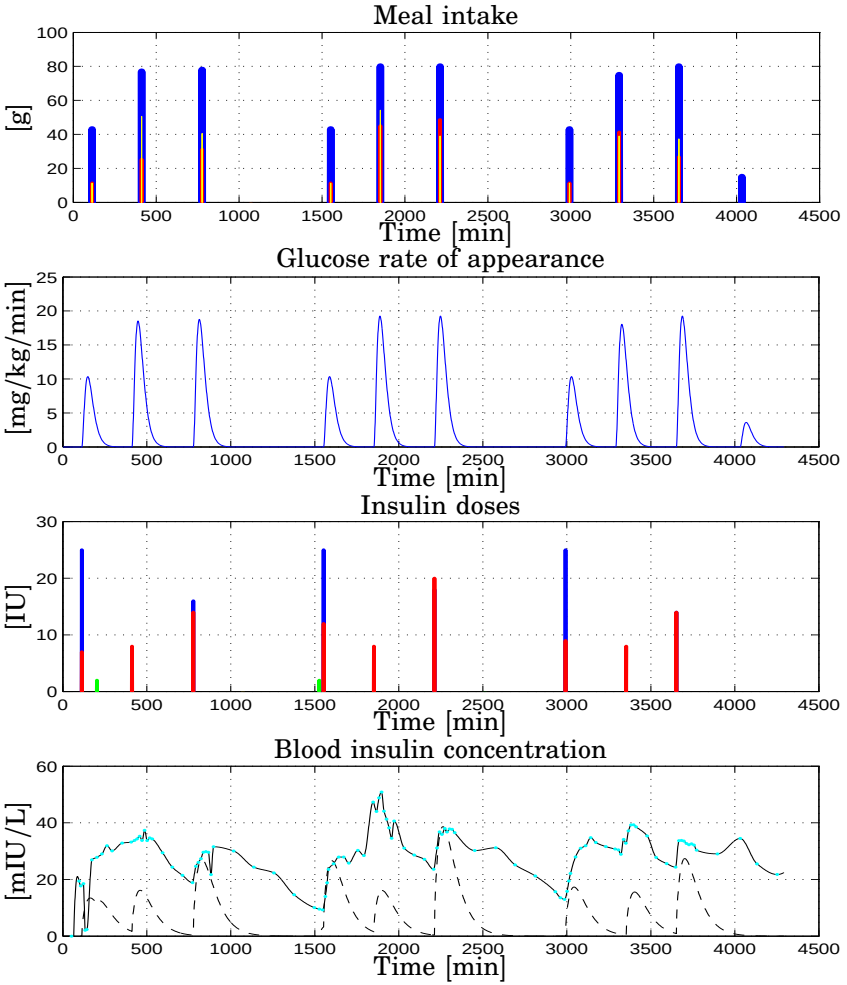

Figure A.15 Patient CHU0115 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

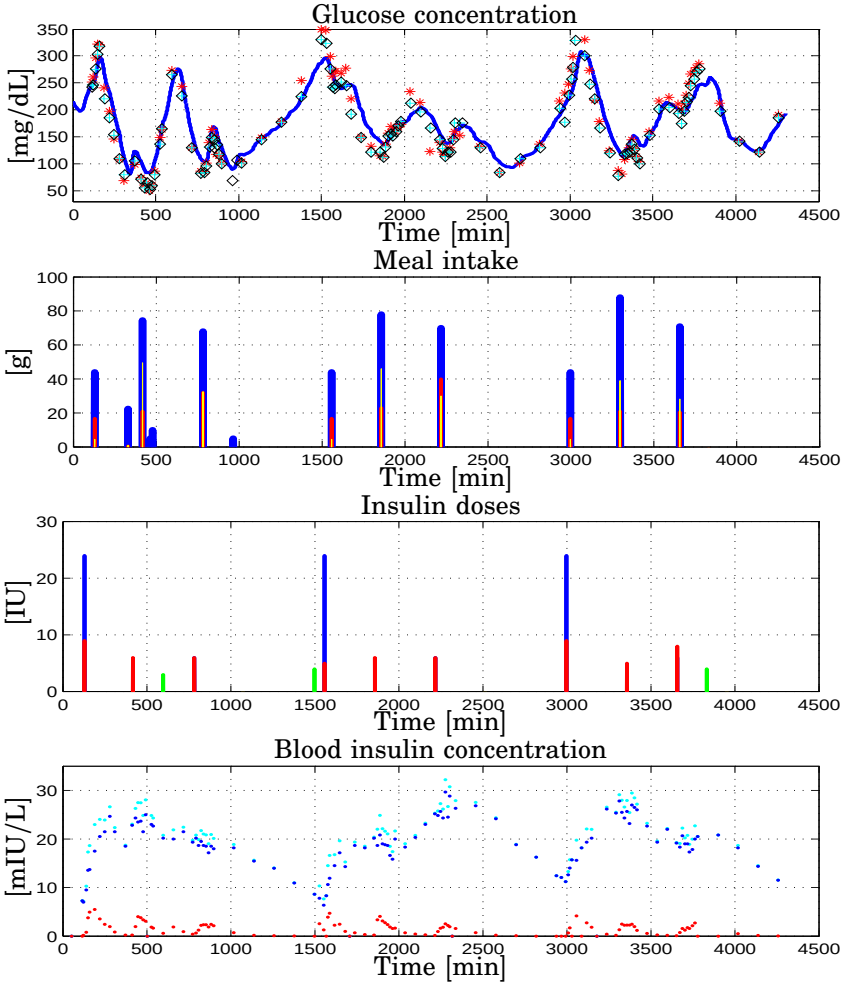

Figure A.16 Patient CHU0120 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

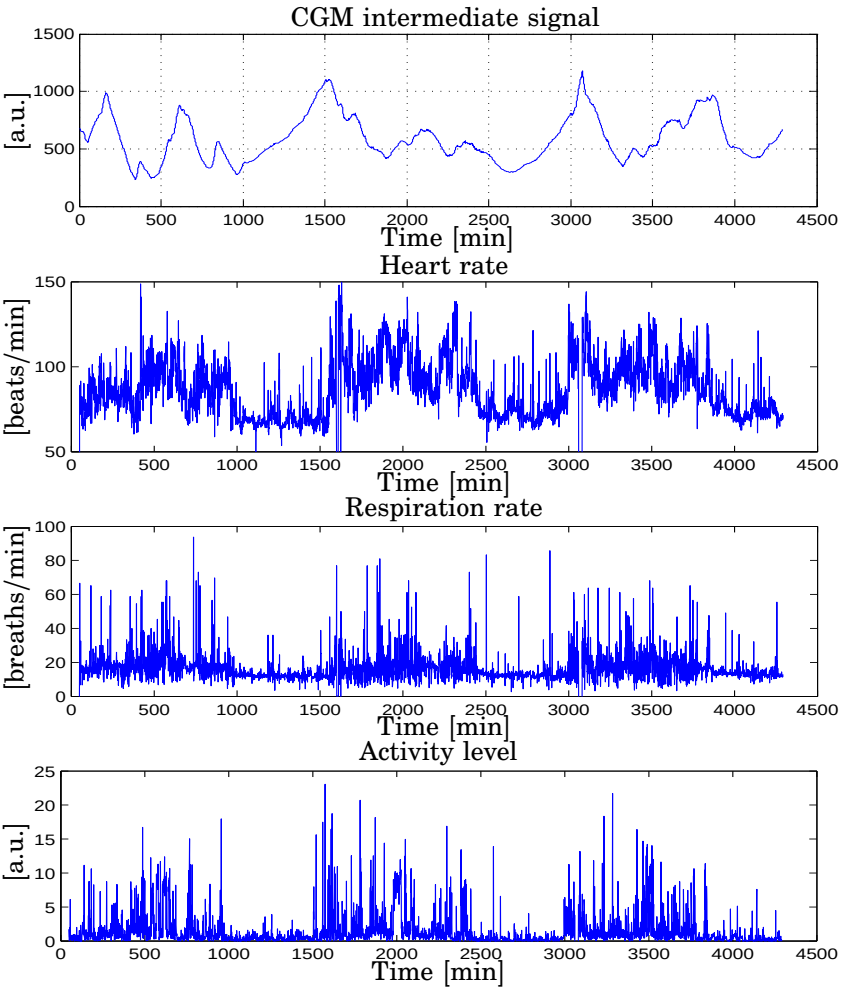

Figure A.17 Patient CHU0120 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

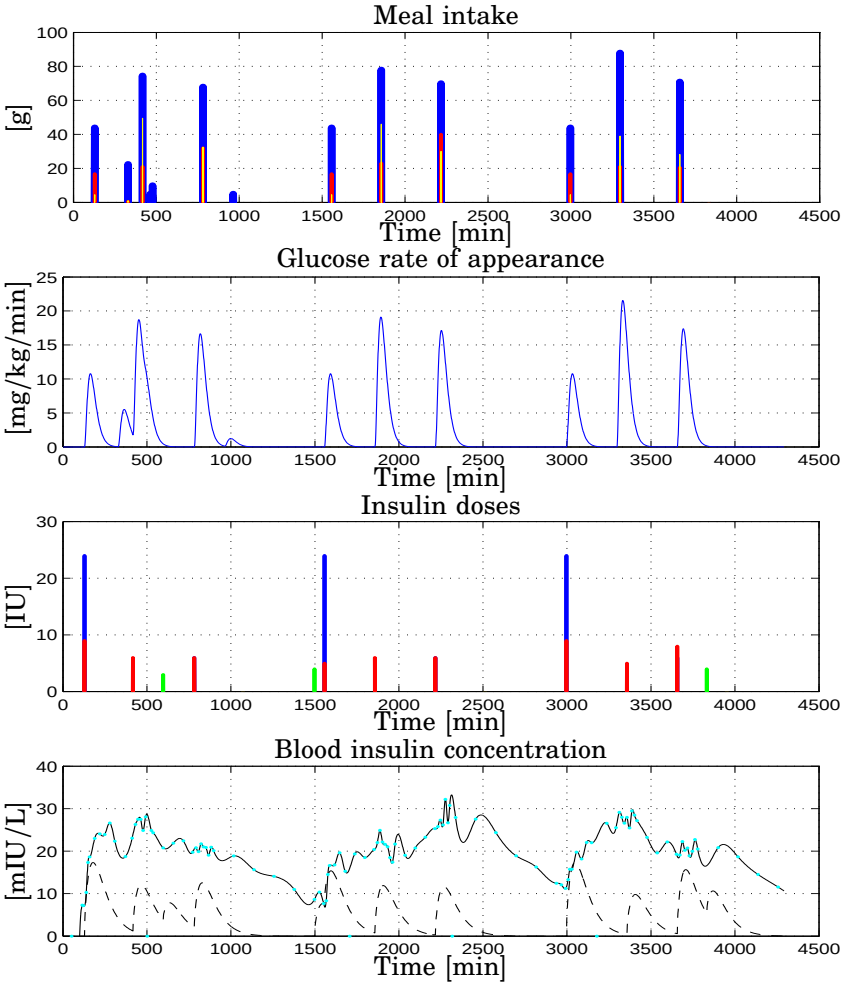

Figure A.18 Patient CHU0120 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

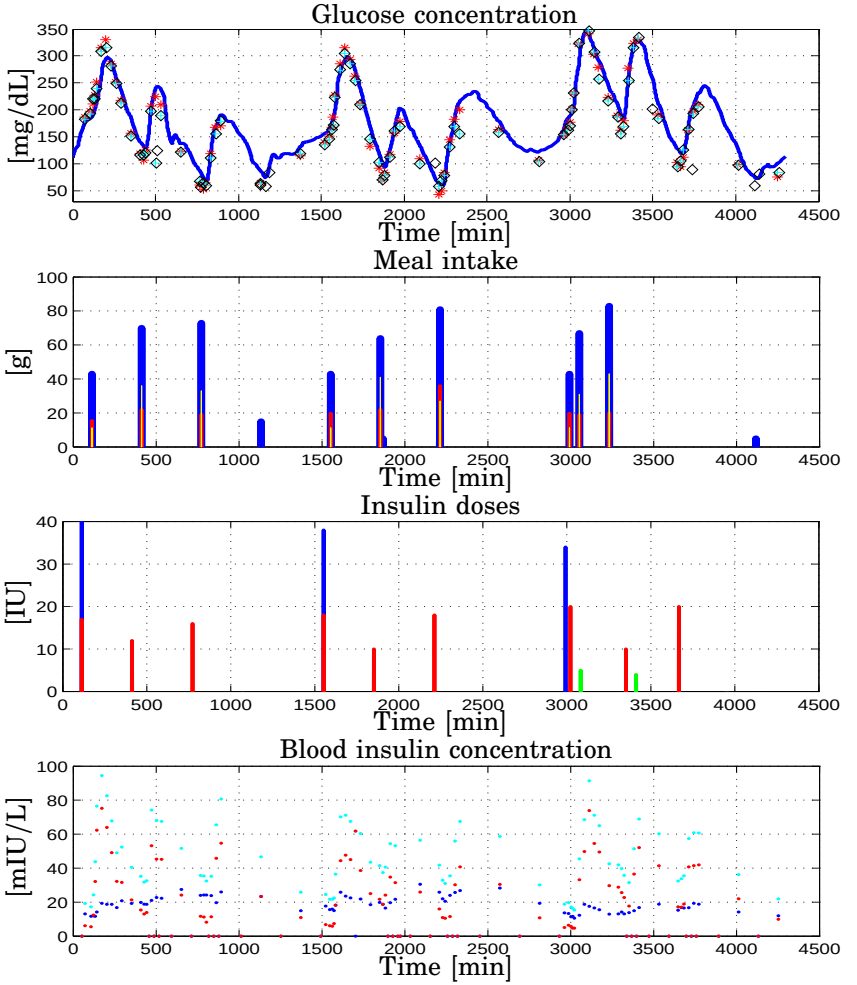

Figure A.19 Patient CHU0130 data vs. Time [min]. *Top* Glucose concentration [mg/dL]: interstitial (blue), plasma (red), finger stick (cyan and black); *Upper Center* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Blood insulin concentration [mIU/L]: basal (blue), bolus (red), total (cyan)

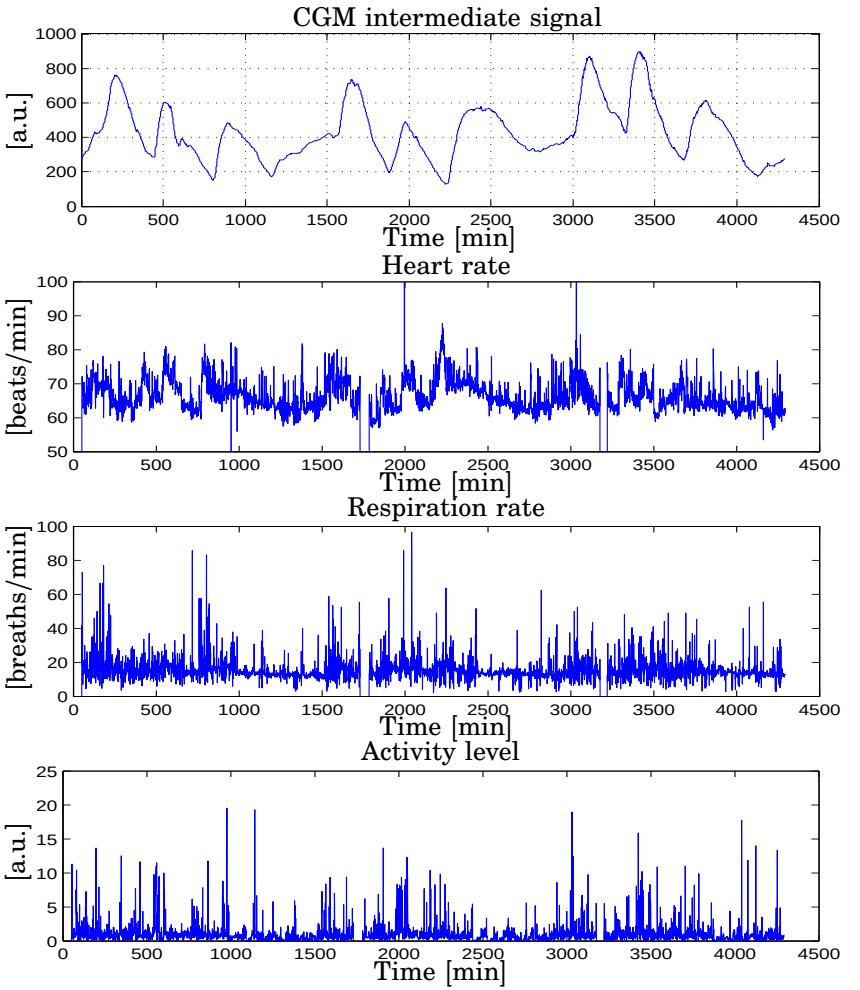

Figure A.20 Patient CHU0130 data vs. Time [min]. *Top* Intermediate current signal from CGM device [a.u.]; *Upper Center* Heart rate [breaths/min]; *Lower Center* Respiration rate [beats/min]; *Bottom* Activity level [a.u.]

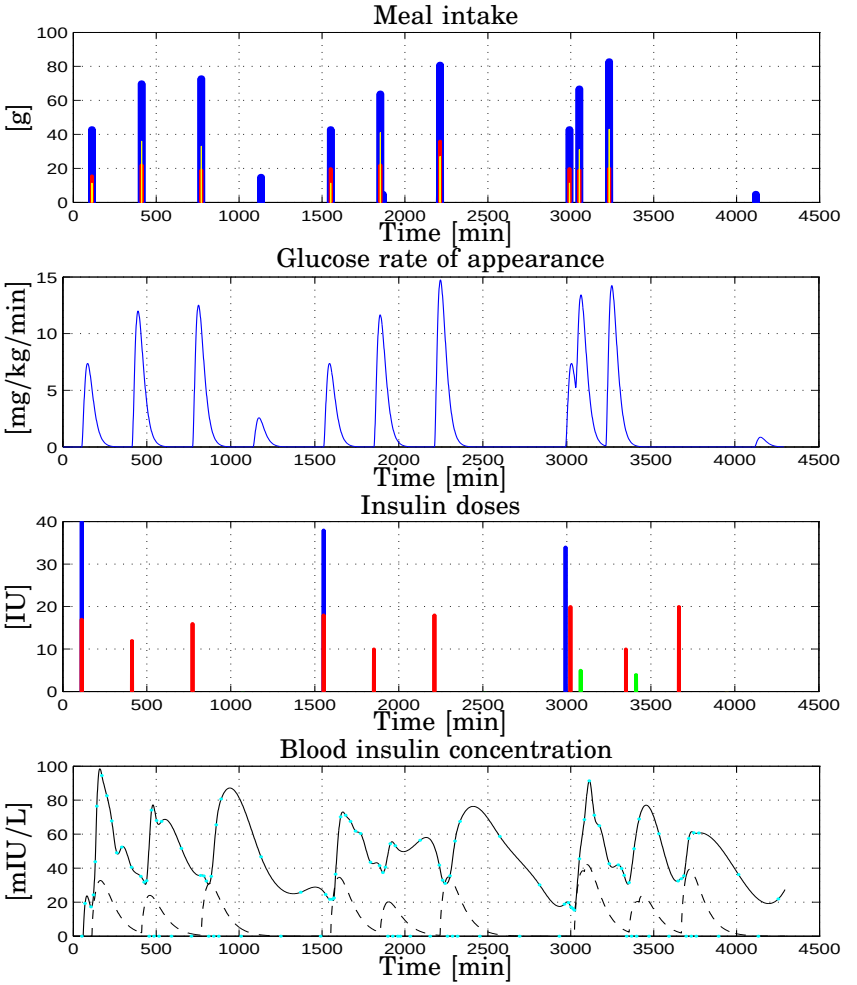

Figure A.21 Patient CHU0130 data vs. Time [min]. *Top* Meal intake [g]: carbohydrates (blue), lipids (red), proteins (yellow); *Upper Center* Glucose Rate of appearance in plasma after a meal [mg/kg/min]; *Lower Center* Insulin doses [IU]: basal (blue), bolus (red), correction (green); *Bottom* Total blood insulin concentration [mIU/L]: interpolated (solid black), blood samples (cyan), from physiological model (dashed black)

B

Power Spectrum Analysis

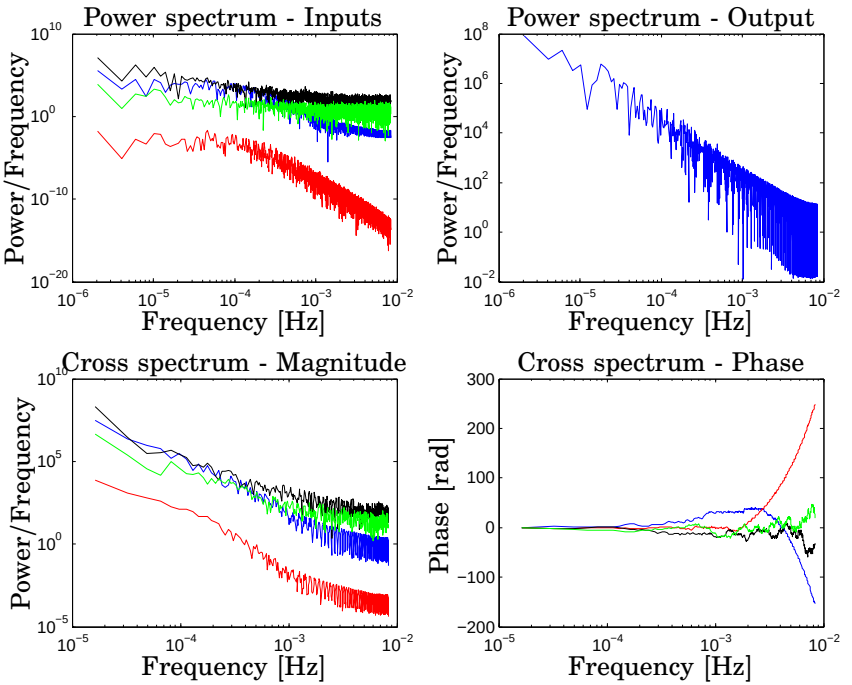

Figure B.1 Patient CHU0103 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2\ (\text{Hz})]$ (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

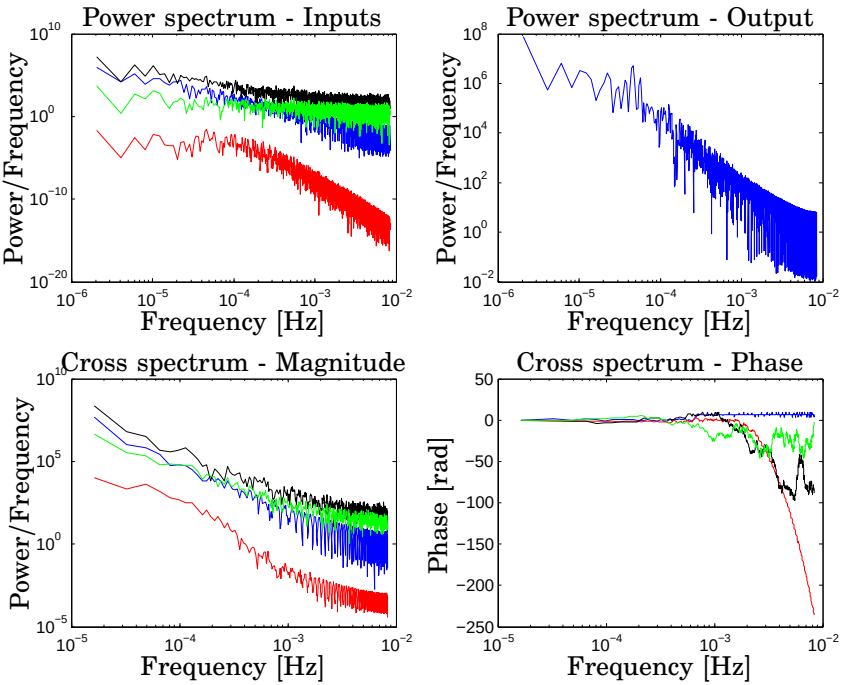

Figure B.2 Patient CHU0104 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2$ (Hz)] (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

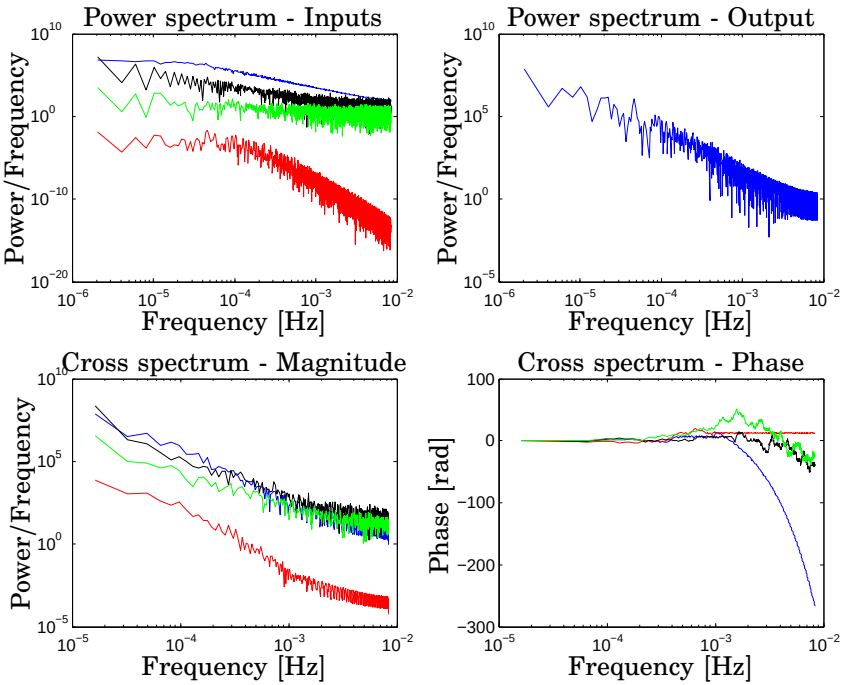

Figure B.3 Patient CHU0105 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2\ (\text{Hz})]$ (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

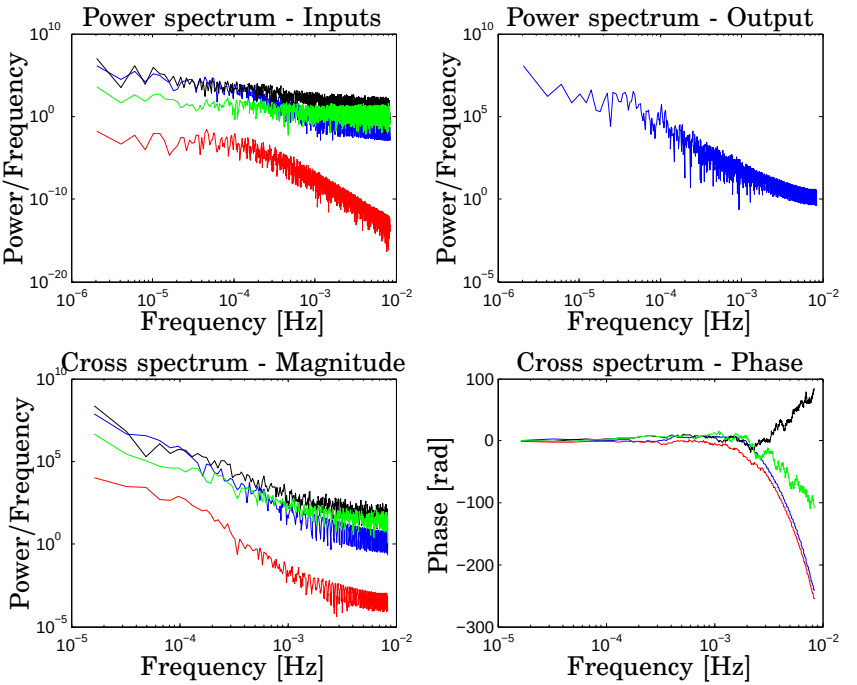

Figure B.4 Patient CHU0106 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2$ (Hz)] (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

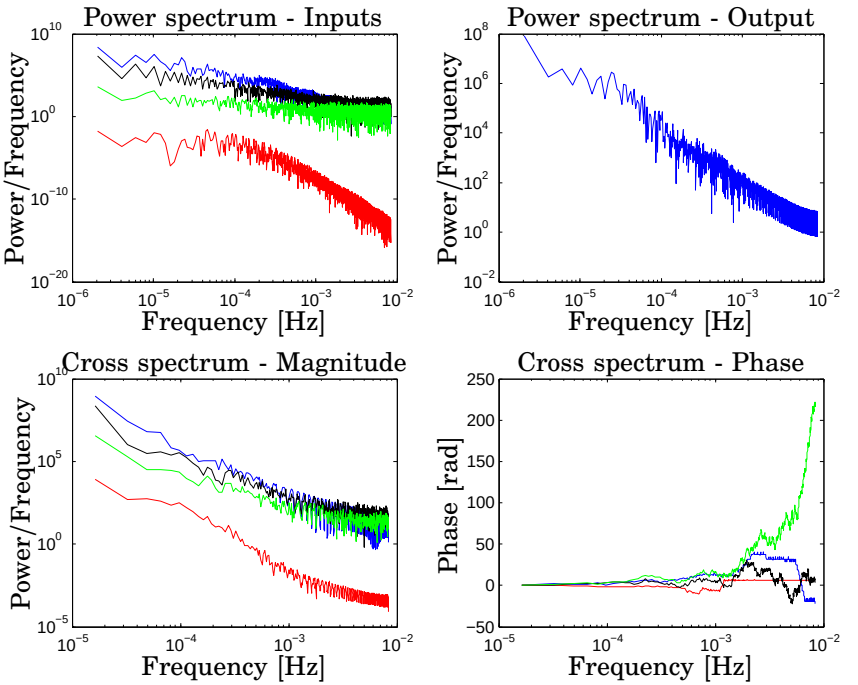

Figure B.5 Patient CHU0115 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2\ (\text{Hz})]$ (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

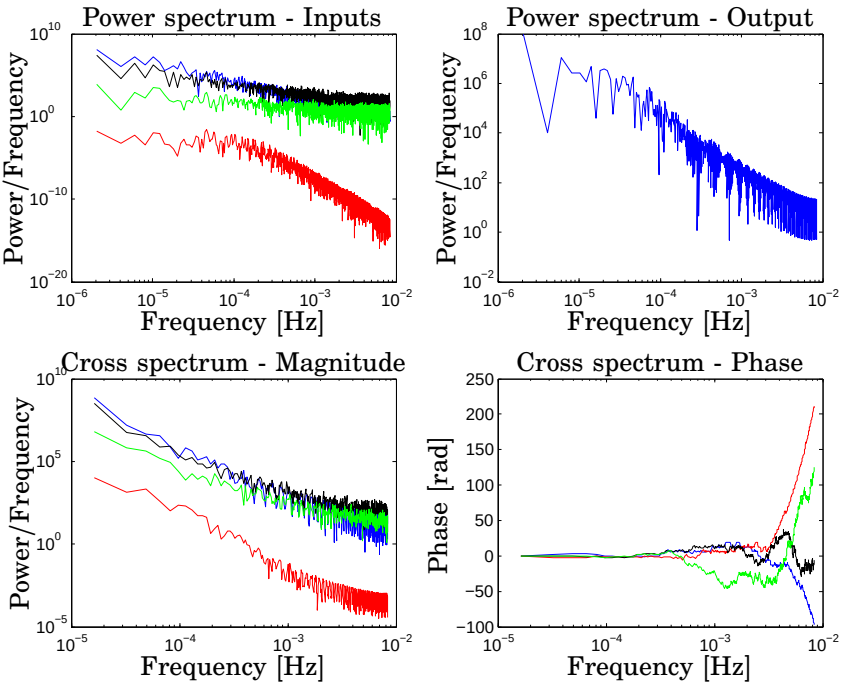

Figure B.6 Patient CHU0120 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2$ (Hz)] (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

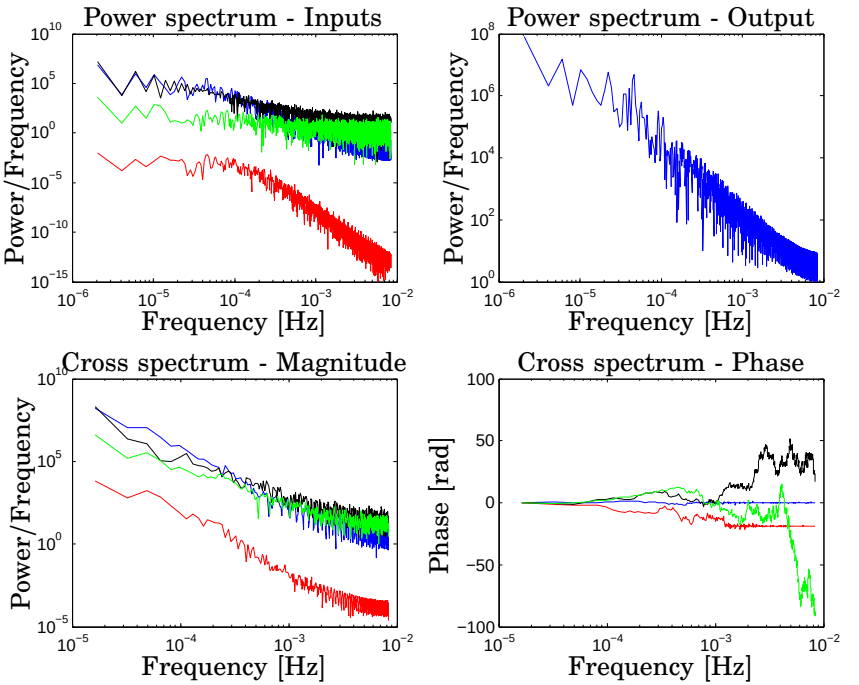

Figure B.7 Patient CHU0130 *Top Left* Magnitude of Power spectrum of inputs: total plasma insulin $[(\text{mIU/L})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance $[(\text{mg/kg/min})^2/(\text{Hz})]$ (red), heart rate $[(\text{beats/min})^2/(\text{Hz})]$ (black), activity [a.u./(Hz)] (green); *Top Right* Magnitude of Power spectrum of output: blood glucose $[(\text{mg/dL})^2/(\text{Hz})]$; *Bottom Left* Magnitude of cross spectrum: total plasma insulin, blood glucose $[(\text{mIU/L})^2(\text{mg/dL})^2/(\text{Hz})]$ (blue), plasma glucose rate of appearance, blood glucose $[(\text{mg/kg/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (red), heart rate, blood glucose $[(\text{beats/min})^2(\text{mg/dL})^2/(\text{Hz})]$ (black), activity, blood glucose $[(\text{a.u.})(\text{mg/dL})^2\ (\text{Hz})]$ (green); *Bottom Right* Phase of cross spectrum [rad]: total plasma insulin (blue), plasma glucose rate of appearance (red), heart rate (black), activity (green). All the spectra vs. Frequency [Hz]

C

Coherence spectrum analysis

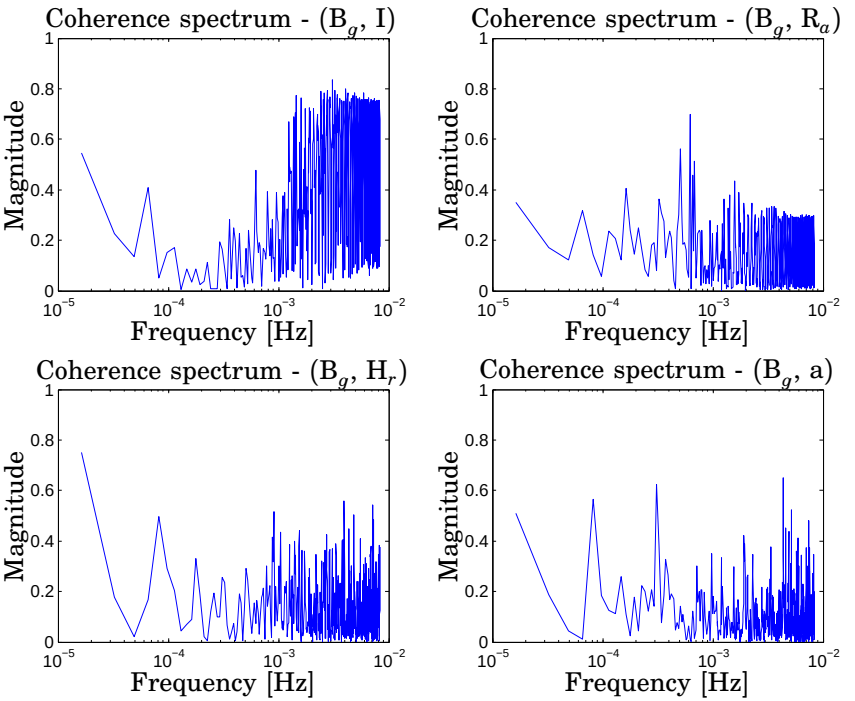

Figure C.1 Patient CHU0103. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

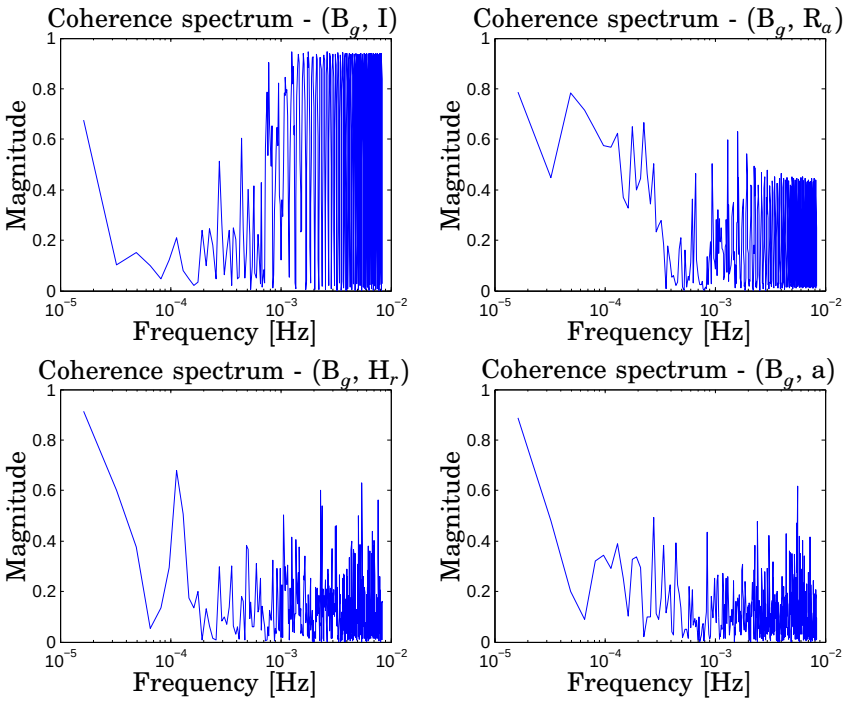

Figure C.2 Patient CHU0104. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

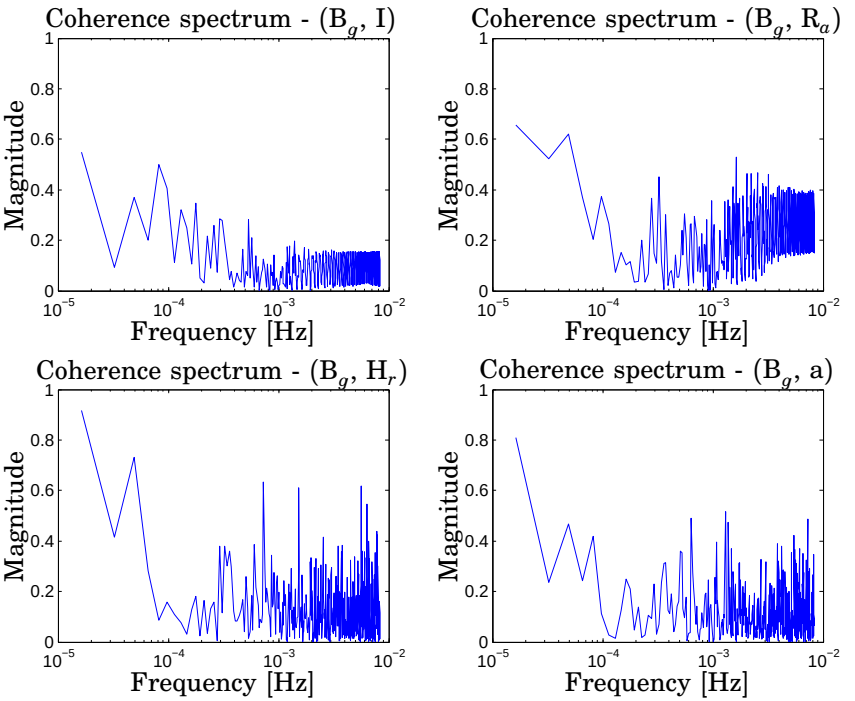

Figure C.3 Patient CHU0105. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

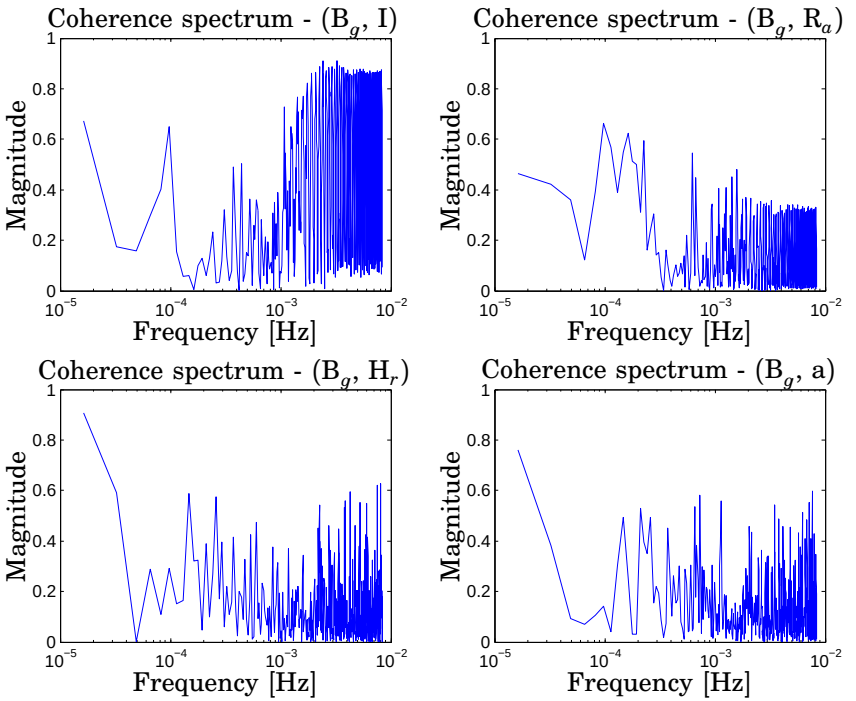

Figure C.4 Patient CHU0106. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

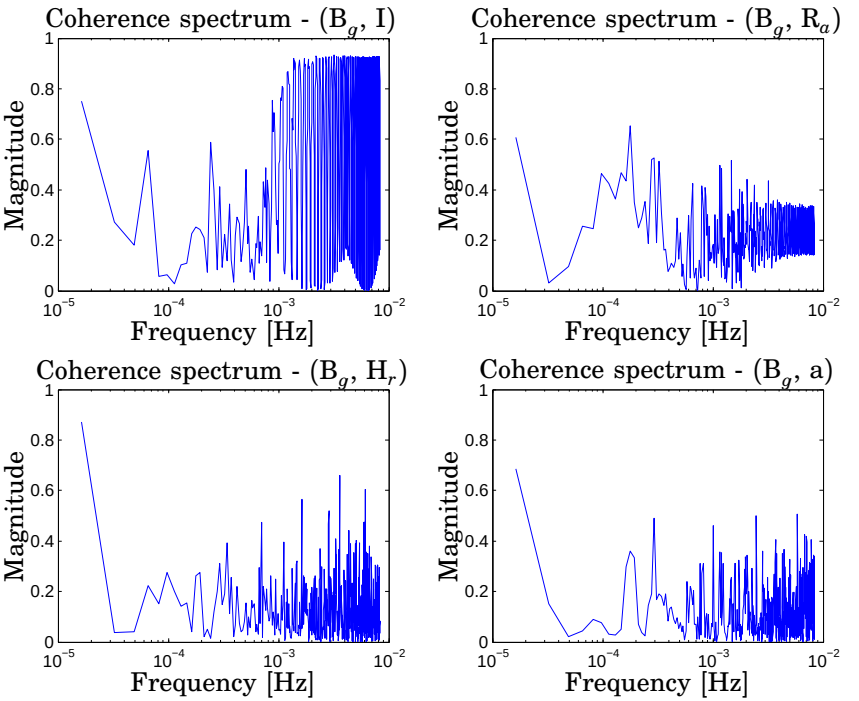

Figure C.5 Patient CHU0115. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

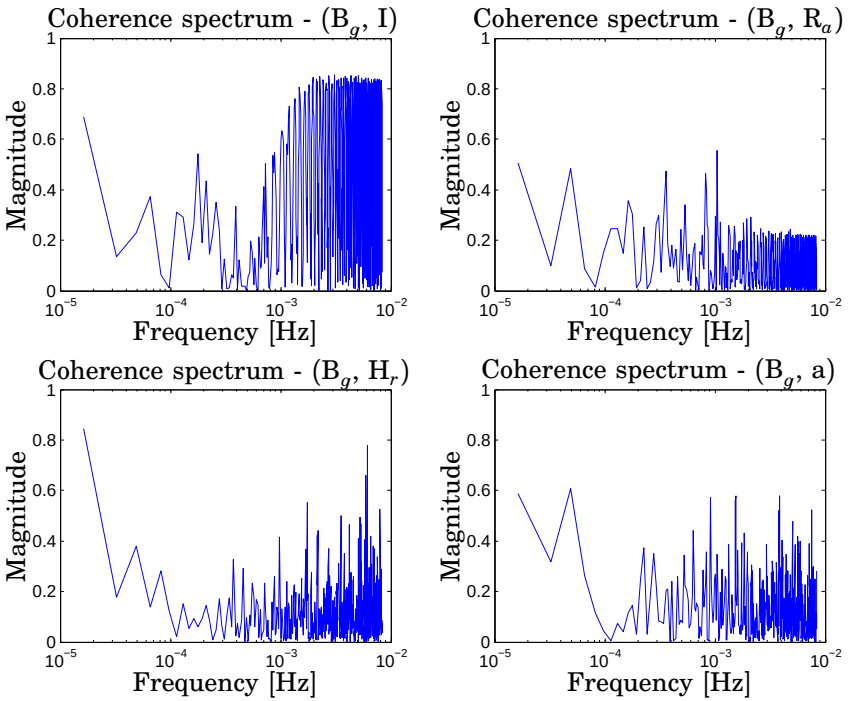

Figure C.6 Patient CHU0120. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

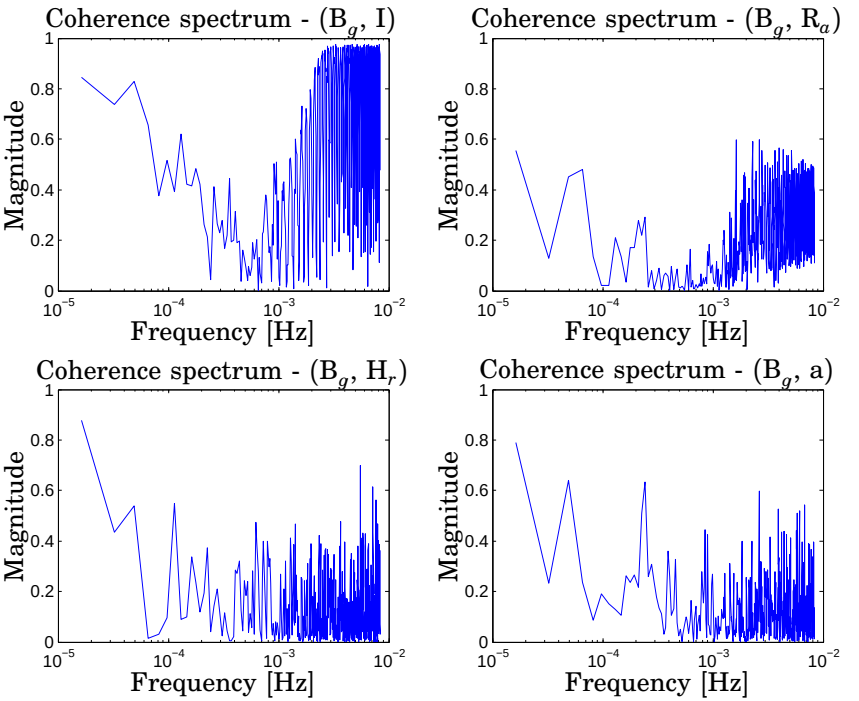

Figure C.7 Patient CHU0130. Coherence spectra between blood glucose and *Top Left* total plasma insulin; *Top Right* plasma glucose rate of appearance; *Bottom Left* Heart Rate; *Bottom Right* Activity. All the spectra vs. Frequency [Hz]

D

Models

This appendix presents one model per patient. Characteristics required to the models were:

- stability
- simplicity
- white residuals
- qualitatively correct responses to inputs

For details see Chapter 4.

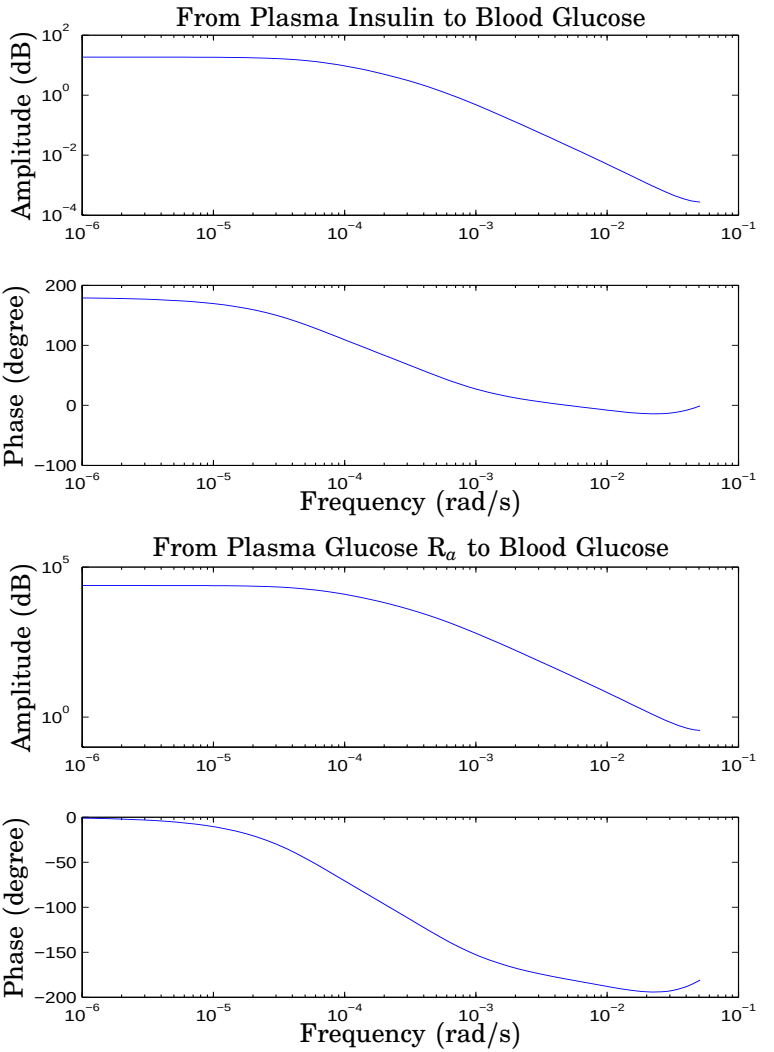

Figure D.1 Patient CHU0103. ARMAX 3rd-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

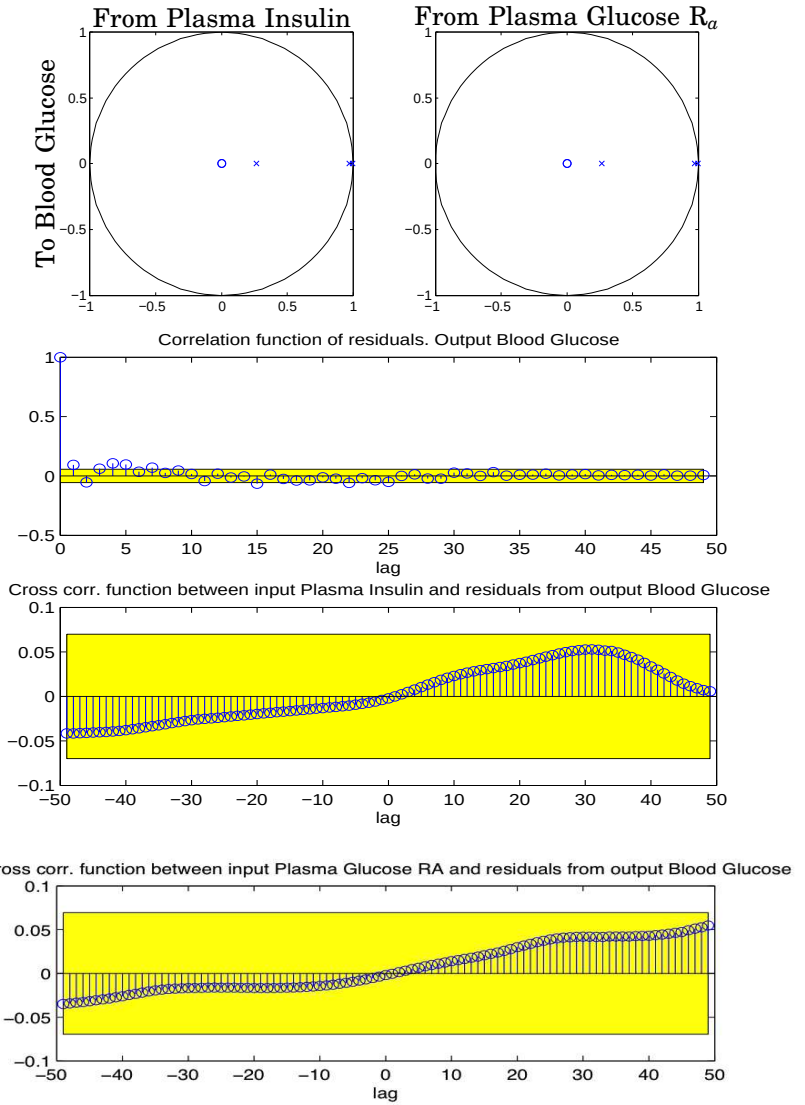

Figure D.2 Patient CHU0103. ARMAX 3rd-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

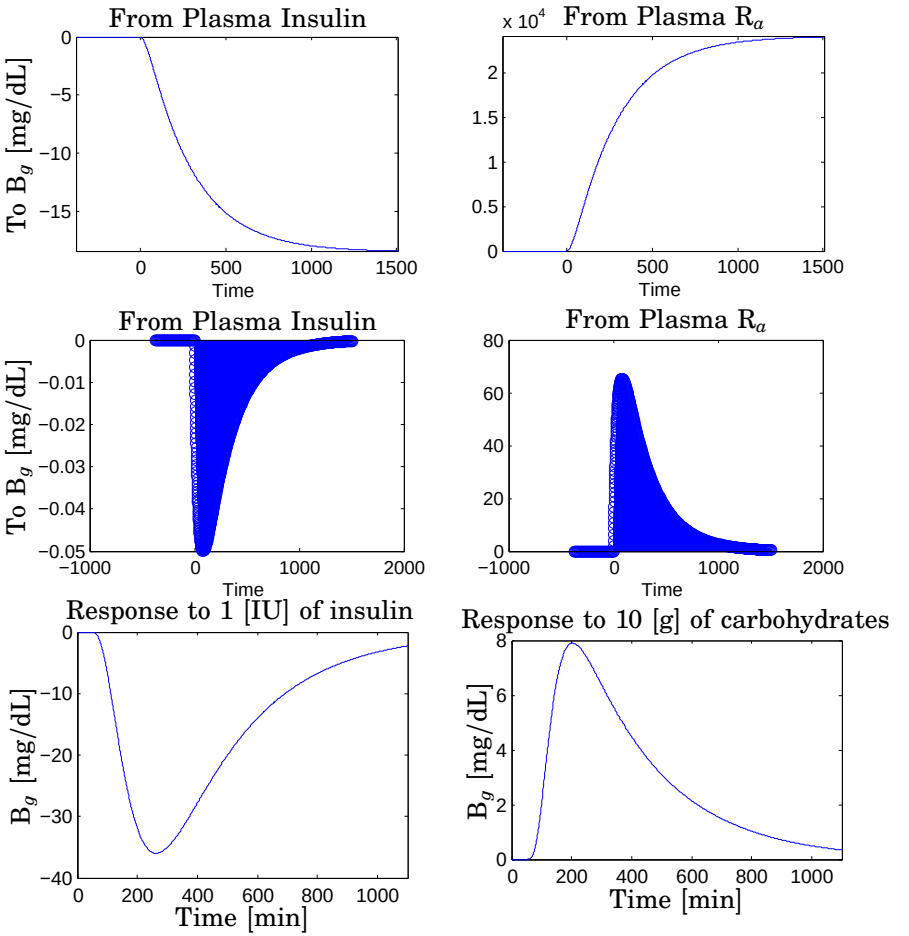

Figure D.3 Patient CHU0103. ARMAX 3rd-order model. *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

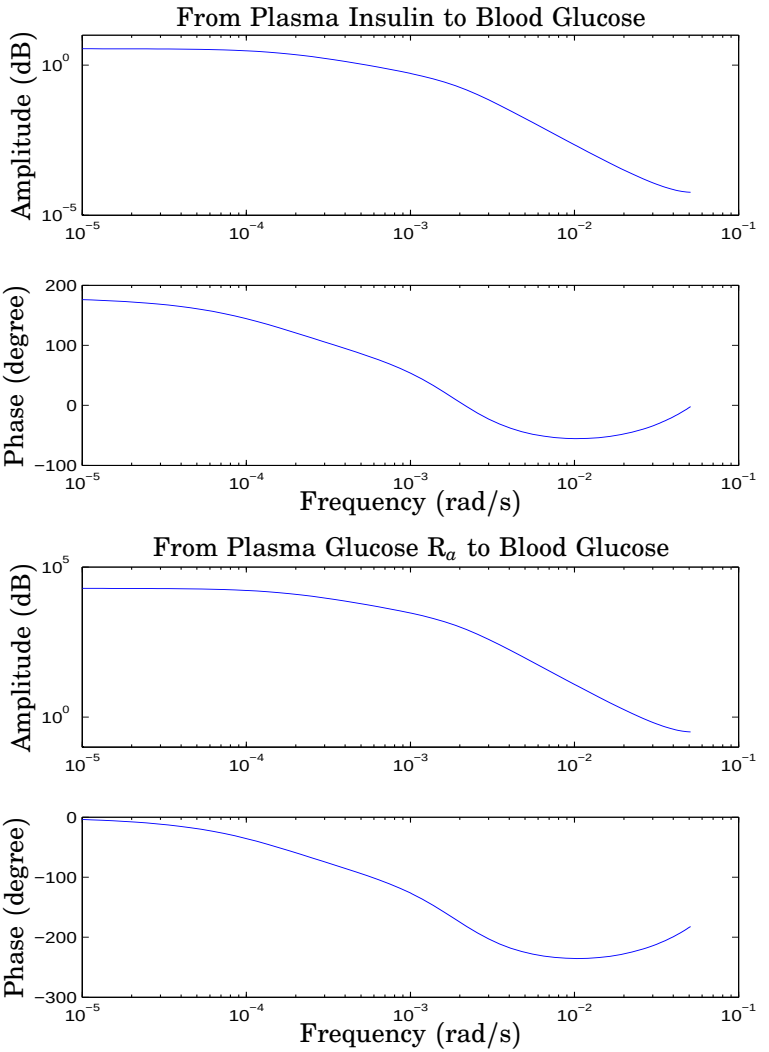

Figure D.4 Patient CHU0104. ARMAX 3rd-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

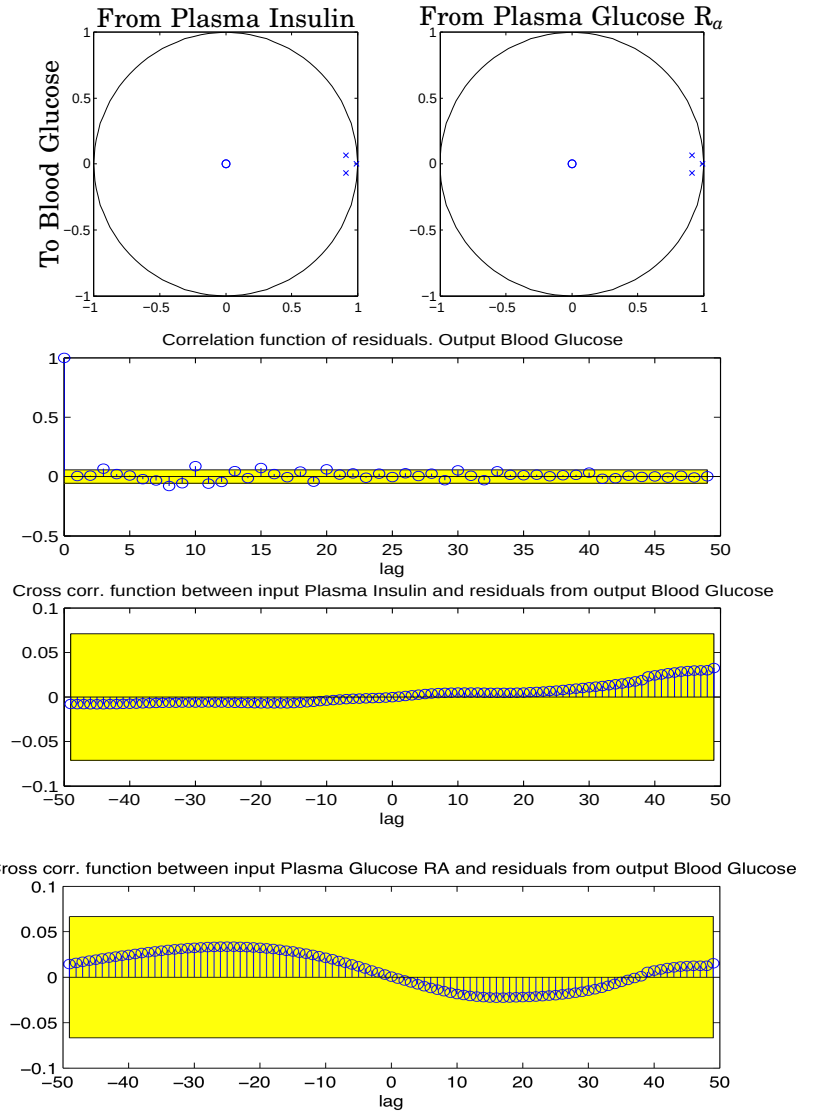

Figure D.5 Patient CHU0104. ARMAX 3rd-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

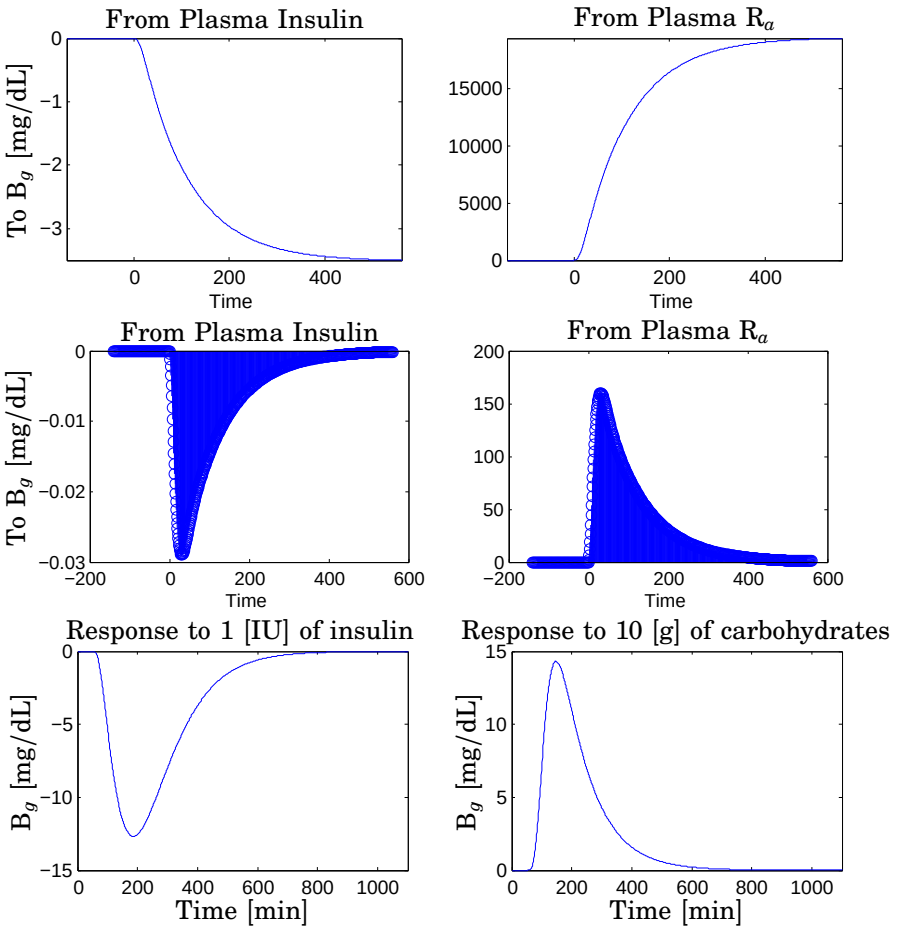

Figure D.6 Patient CHU0104. ARMAX 3rd-order model. *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

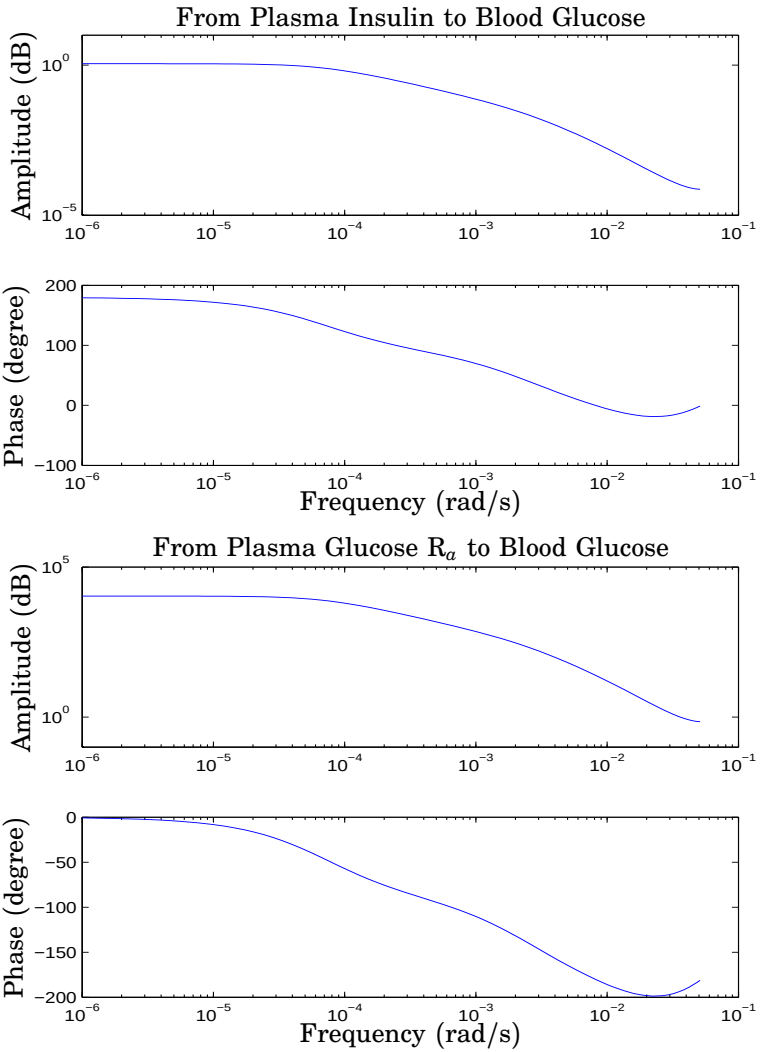

Figure D.7 Patient CHU0105. ARMAX 3rd-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

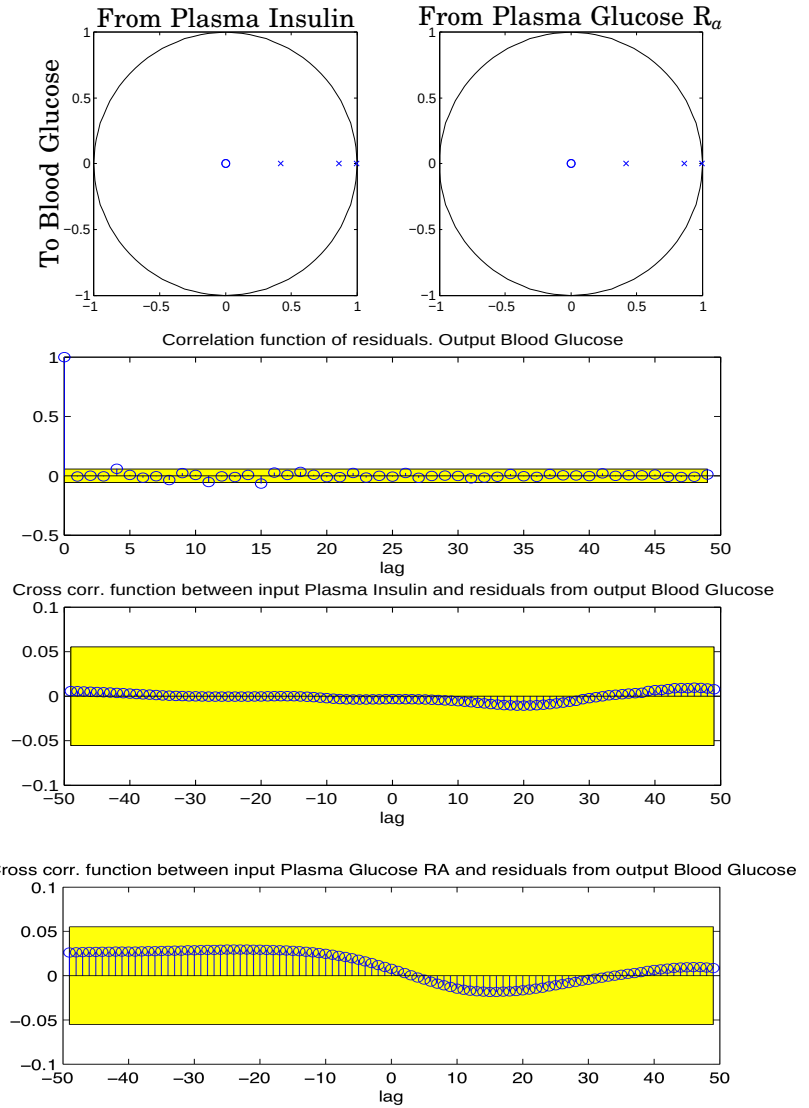

Figure D.8 Patient CHU0105. ARMAX 3rd-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

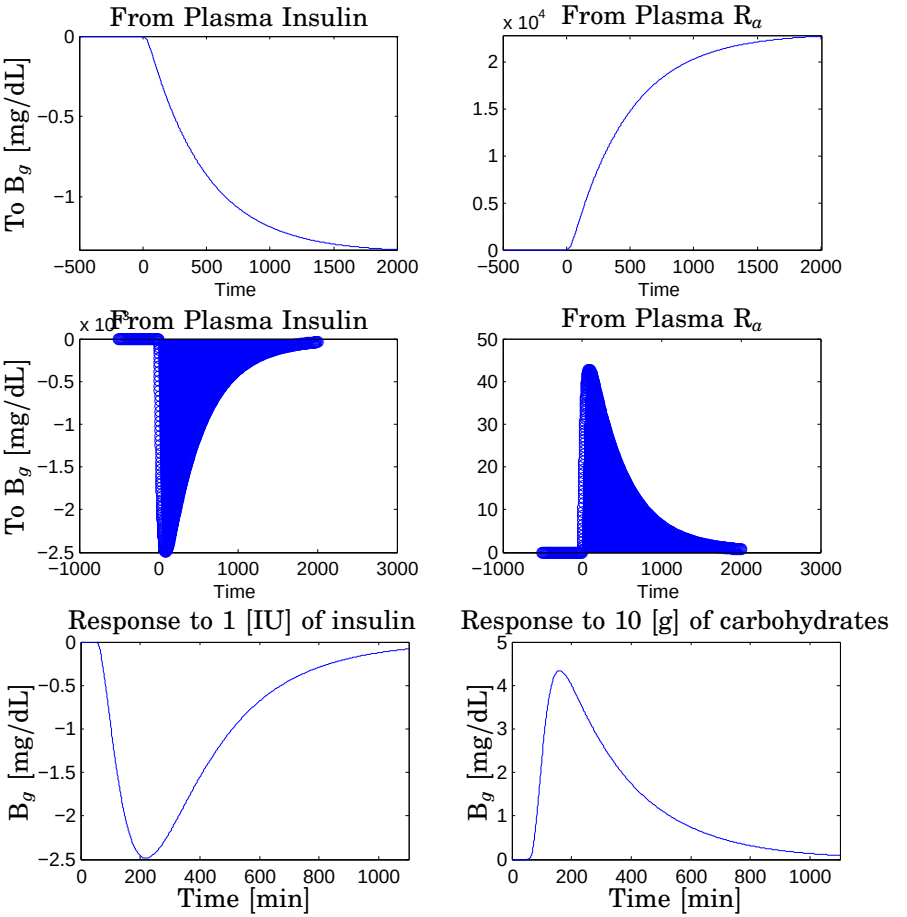

Figure D.9 Patient CHU0105. ARMAX 3rd-order model. *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

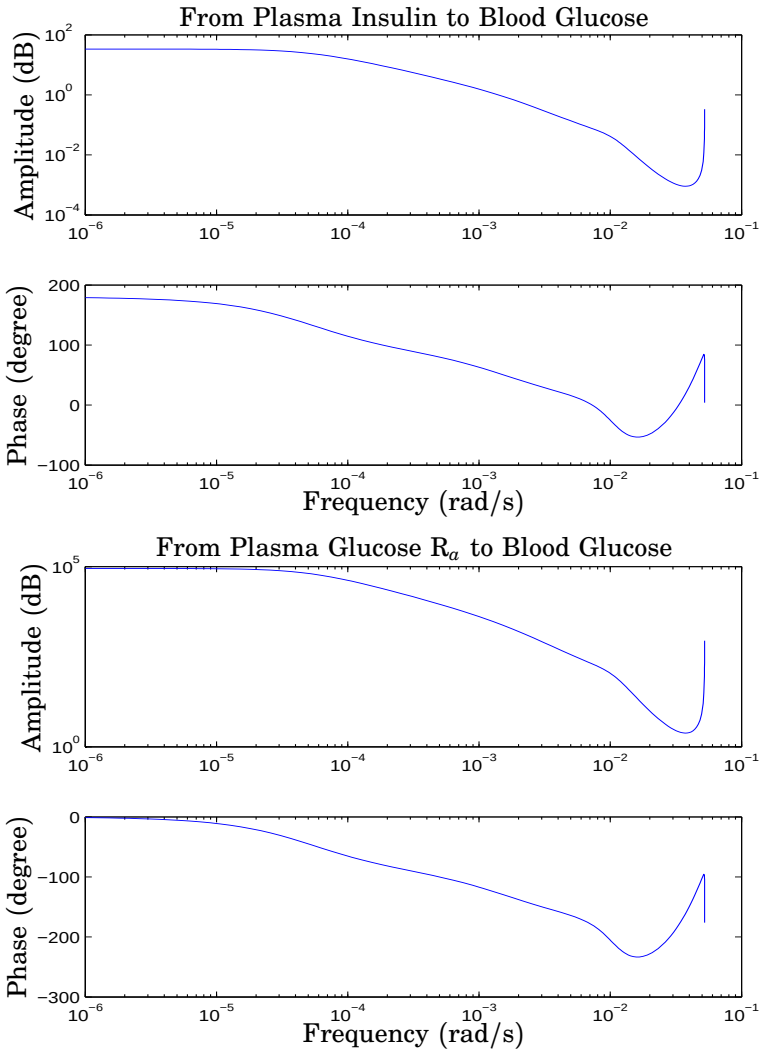

Figure D.10 Patient CHU0106. ARMAX 6th-order model. Bode diagrams of the estimated transfer functions: Top Panel From Plasma Insulin to Blood Glucose; Bottom Panel From Plasma Glucose R_a to Blood Glucose

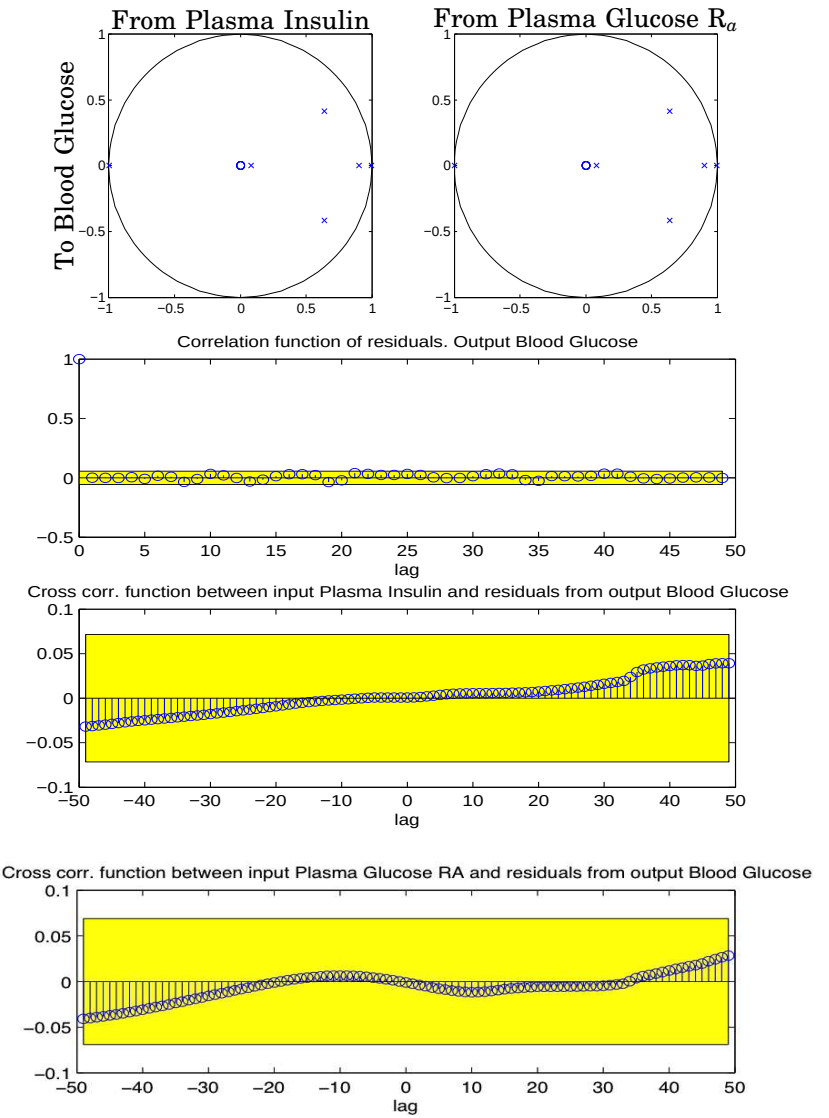

Figure D.11 Patient CHU0106. ARMAX 6th-order model *Top* Pole-zero diagram for the identified ARMAX model. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

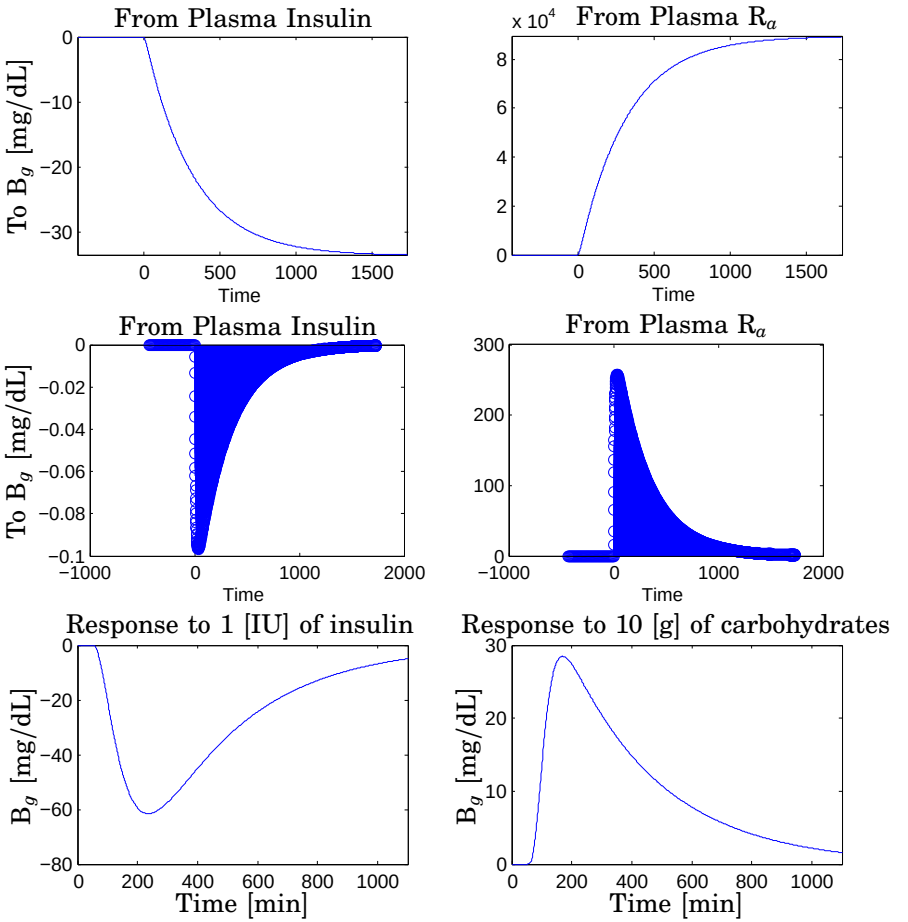

Figure D.12 Patient CHU0106. ARMAX 6th-order model *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

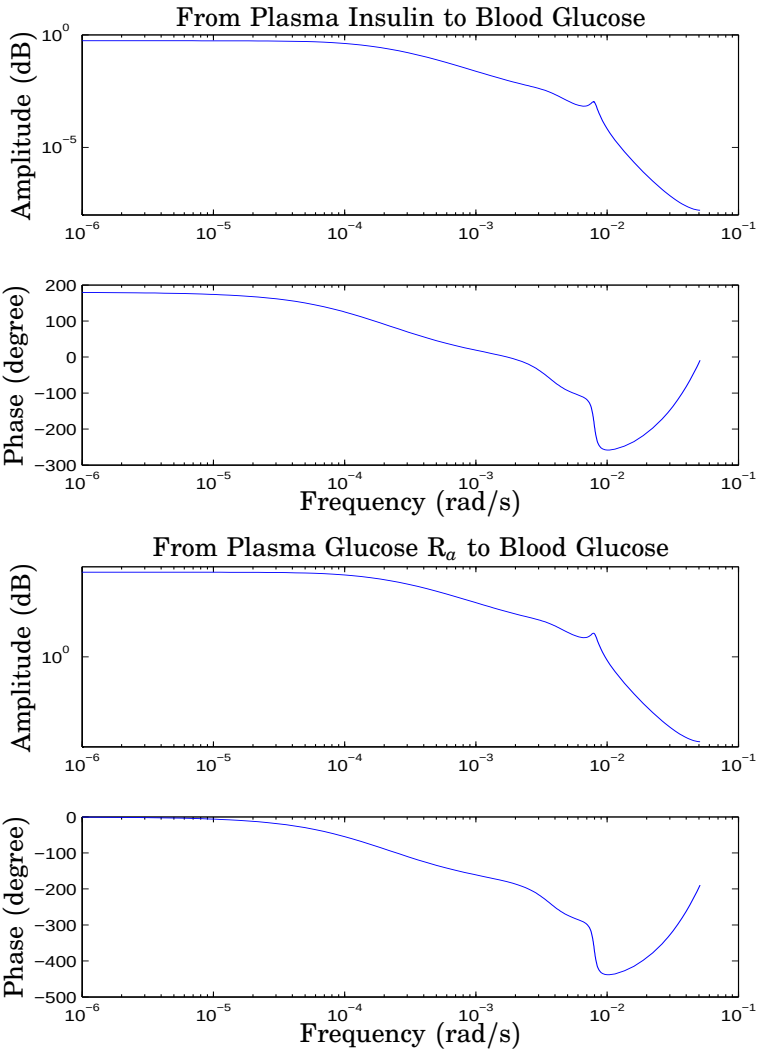

Figure D.13 Patient CHU0115. ARMAX 6th-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

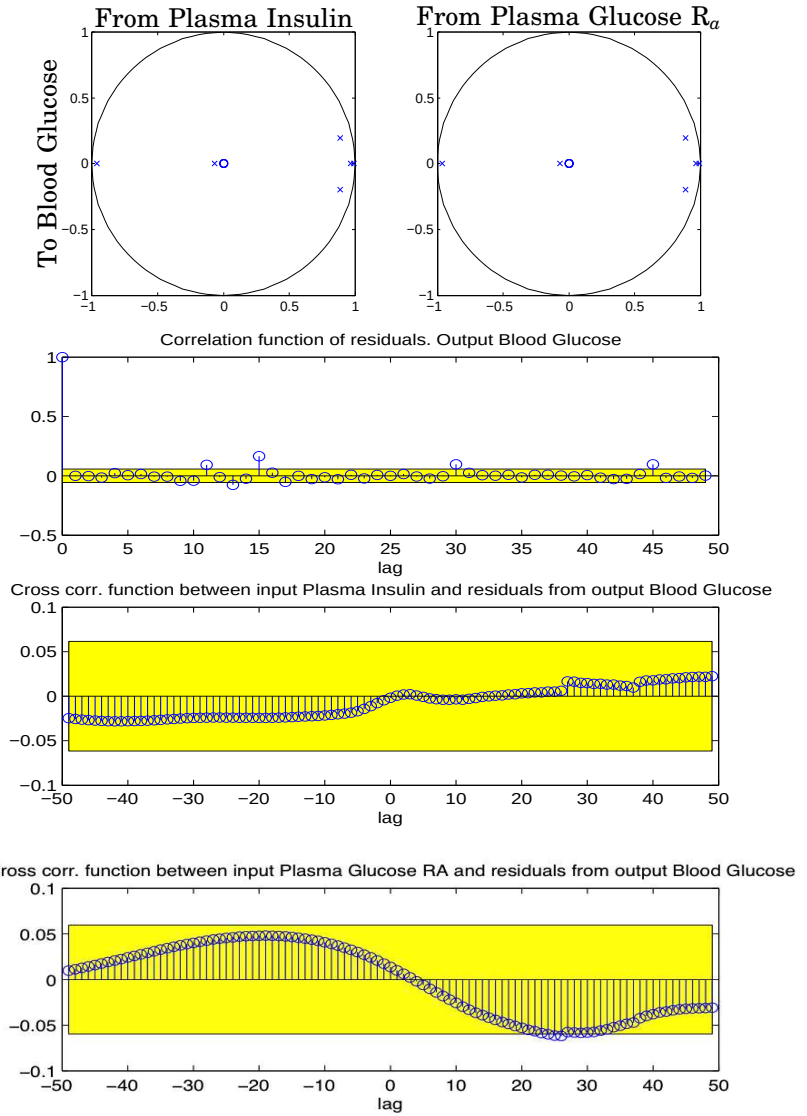

Figure D.14 Patient CHU0115. ARMAX 6th-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

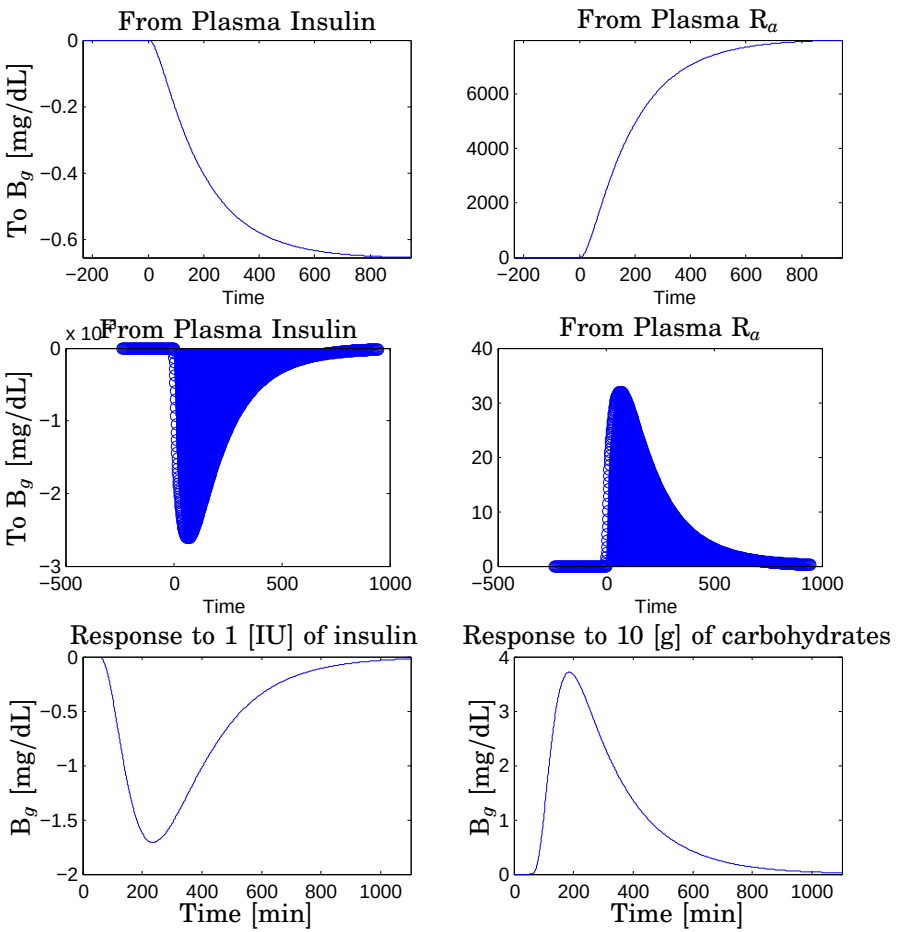

Figure D.15 Patient CHU0115. ARMAX 6th-order model. *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

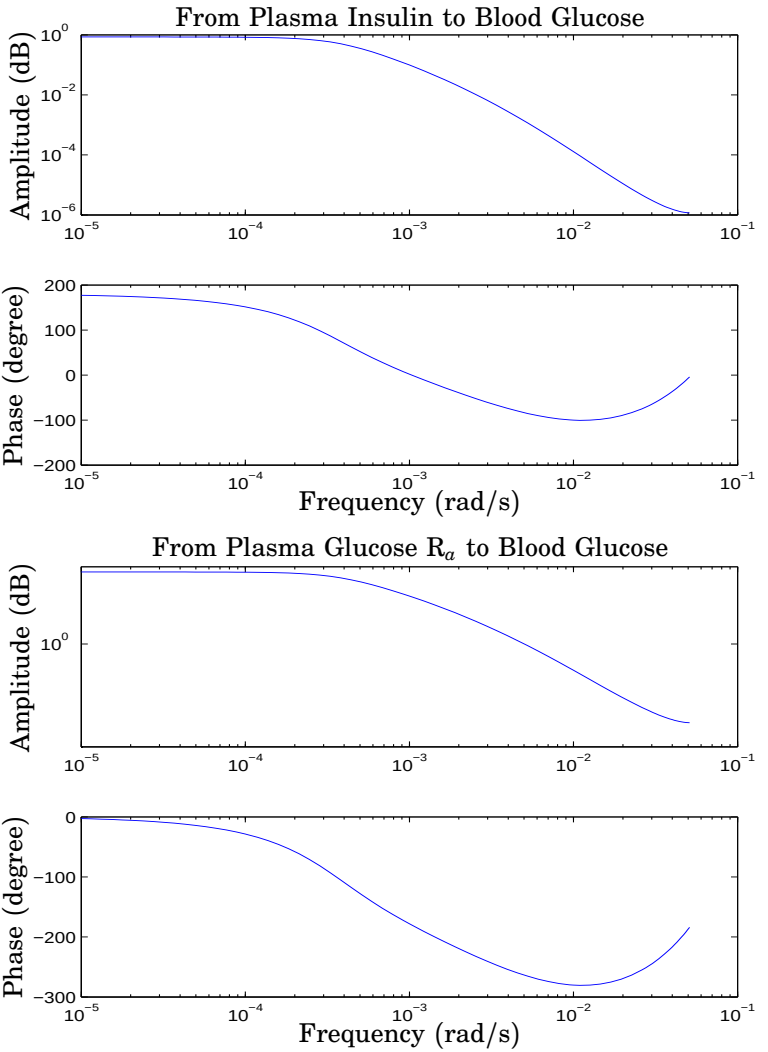

Figure D.16 Patient CHU0120. ARMAX 4th-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

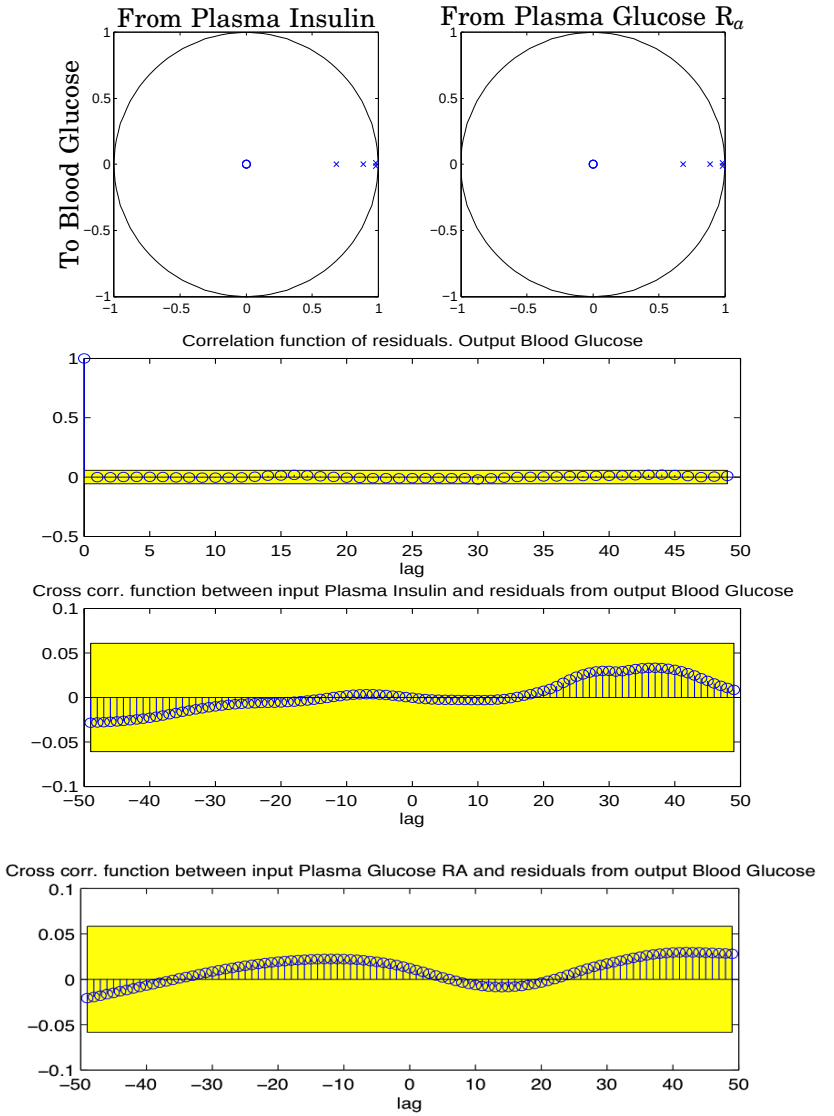

Figure D.17 Patient CHU0120. ARMAX 4th-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

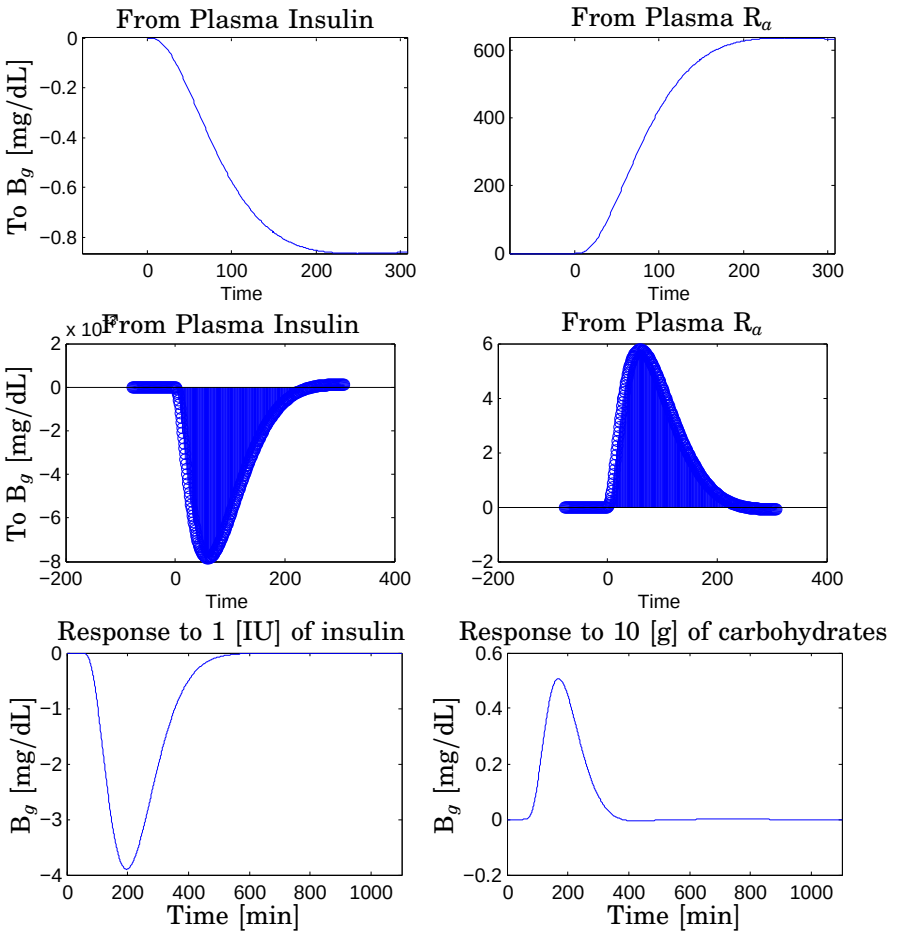

Figure D.18 Patient CHU0120. ARMAX 4th-order model. *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

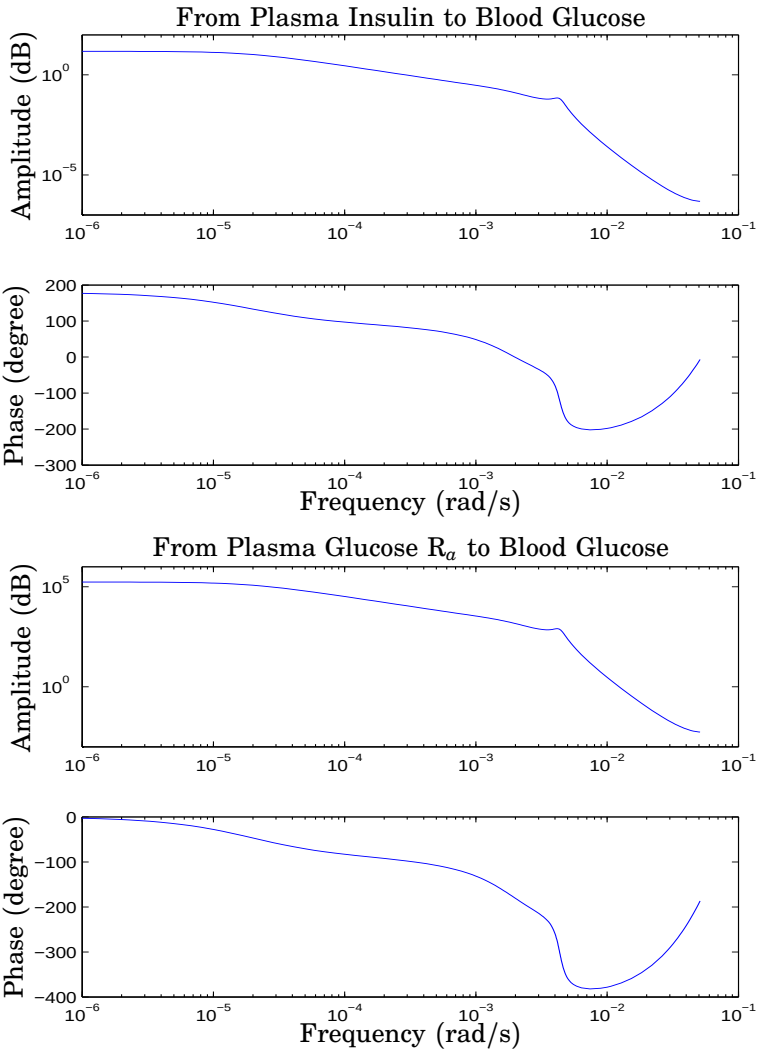

Figure D.19 Patient CHU0130. ARMAX 5th-order model. Bode diagrams of the estimated transfer functions: *Top Panel* From Plasma Insulin to Blood Glucose; *Bottom Panel* From Plasma Glucose R_a to Blood Glucose

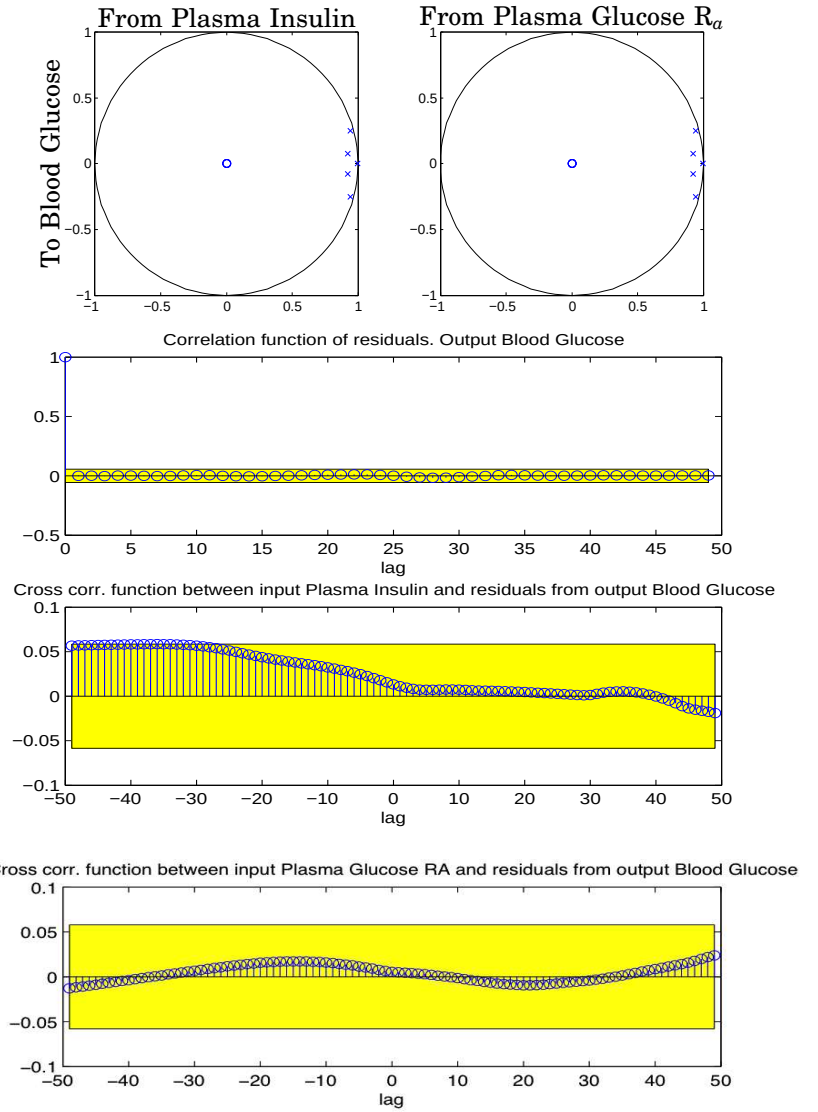

Figure D.20 Patient CHU0130. ARMAX 5th-order model. *Top* Pole-zero diagram. The symbols 'x' and 'o' denote pole and zeros, respectively; *Bottom* Residual analysis on identification data

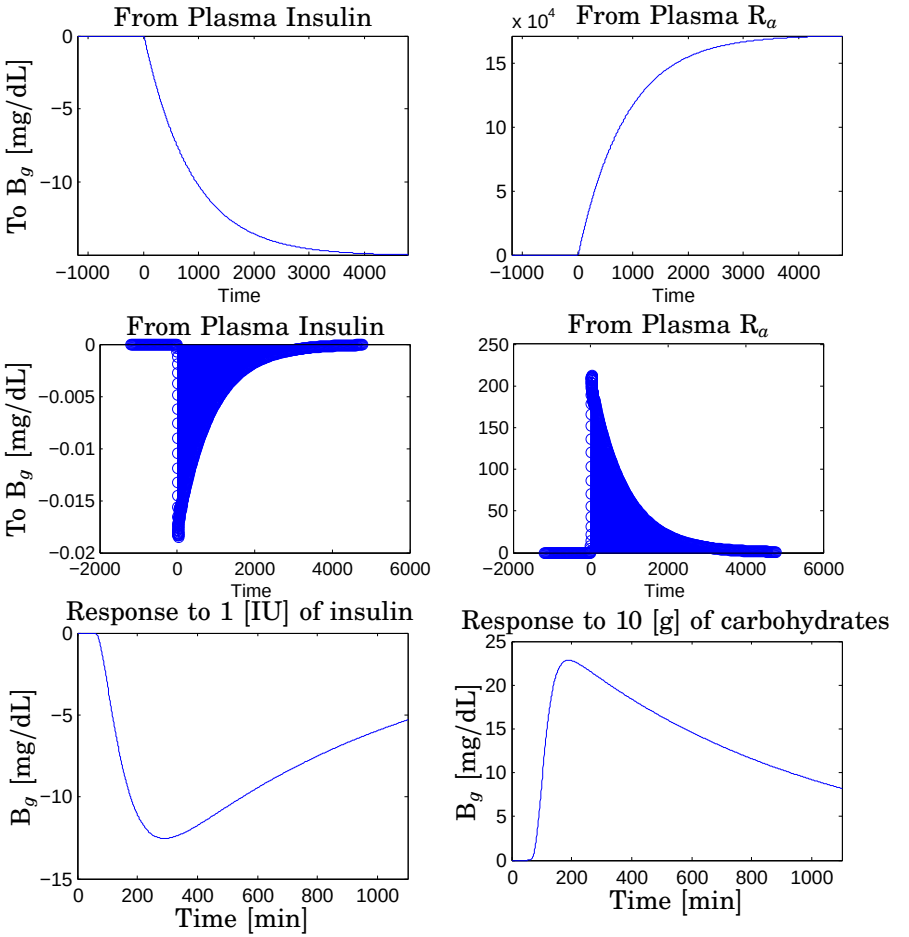

Figure D.21 Patient CHU0130. ARMAX 5th-order model. *Top* Step responses; *Center* Impulse responses; *Bottom Left* BG response to 1[IU] of fast-acting insulin; *Bottom Right* BG response to 10[g] of carbohydrates

E

Model-based predictors

This appendix shows the prediction performances of the models in Appendix D. Predictions were obtained with the Matlab® System Identification Toolbox command *predict.m*. The initial states were set to a value that minimizes the norm of the prediction error over the available data range.

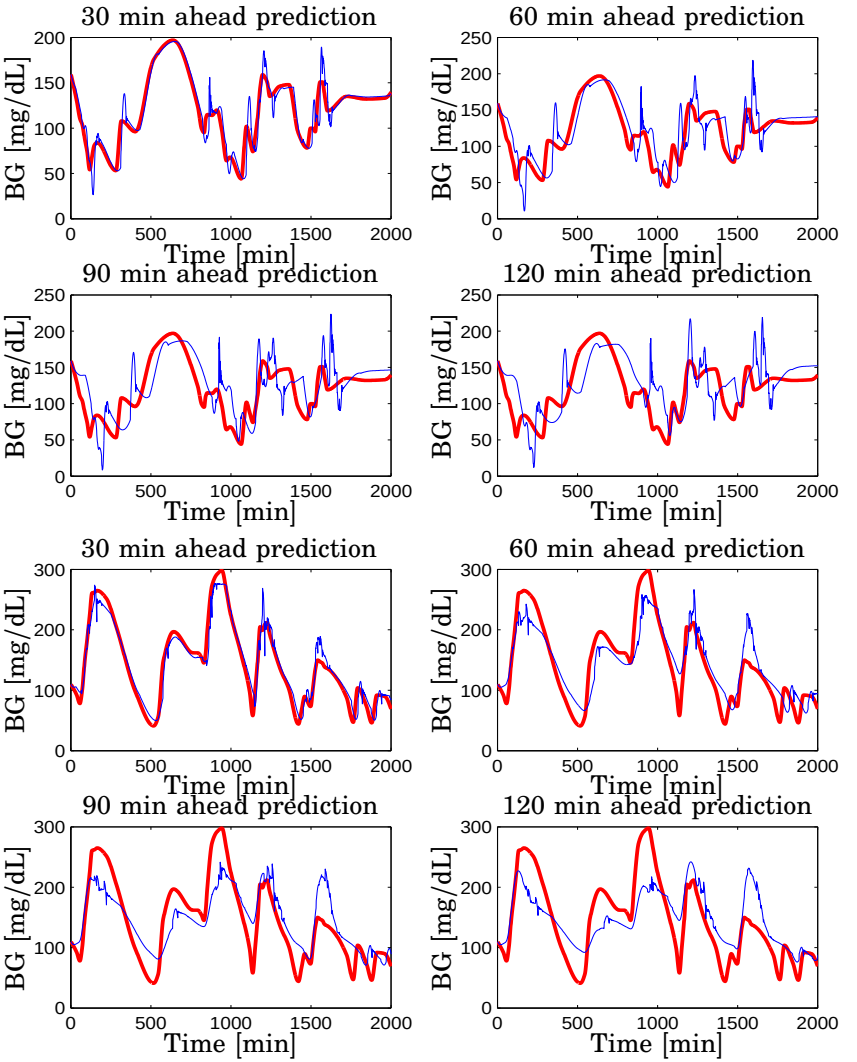

Figure E.1 *Top Panel* Patient CHU0103, ARMAX 3rd-order model; *Bottom Panel* Patient CHU0104, ARMAX 3rd-order model. Evaluation on validation data: model based predictor (thin) and measured blood glucose (thick) [mg/dL] vs. time [min]. Top left 30 minutes; Top right 60 minutes; Bottom left 90 minutes; Bottom right 120 minutes ahead prediction

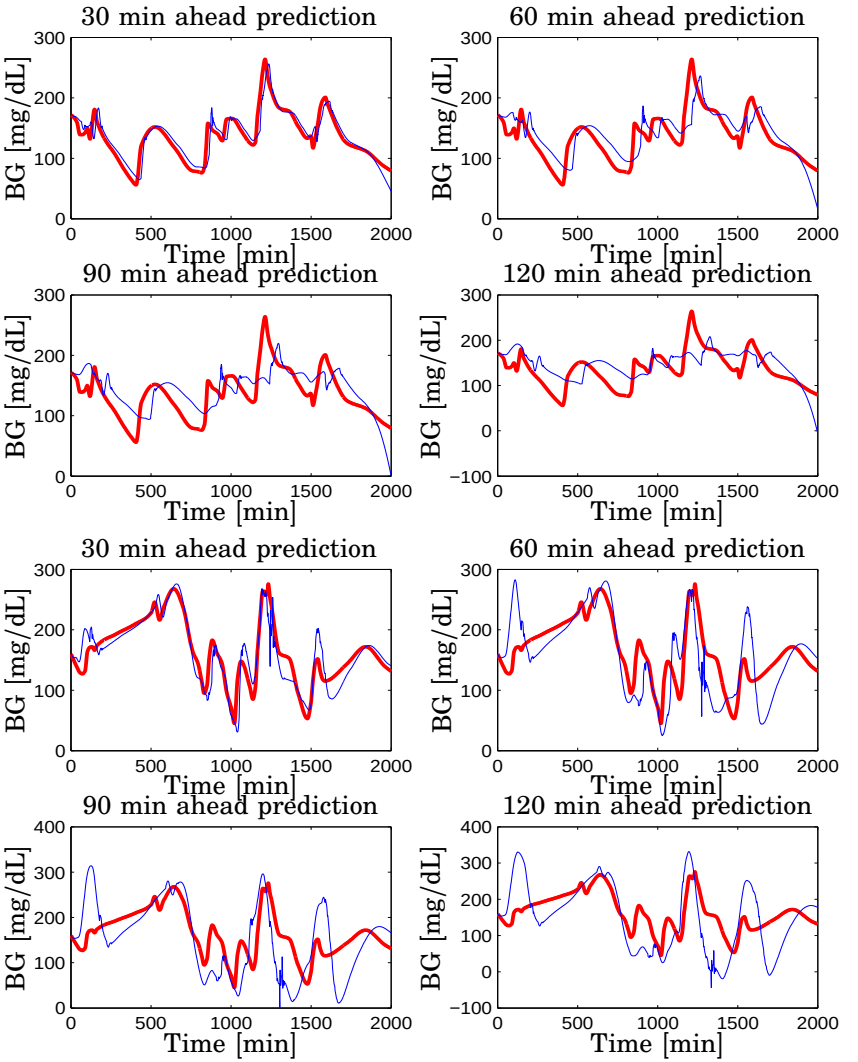

Figure E.2 *Top Panel* Patient CHU0105, ARMAX 3rd-order model; *Bottom Panel* Patient CHU0106, ARMAX 6th-order model. Evaluation on validation data: model based predictor (thin) and measured blood glucose (thick) [mg/dL] vs. time [min]. Top left 30 minutes; Top right 60 minutes; Bottom left 90 minutes; Bottom right 120 minutes ahead prediction

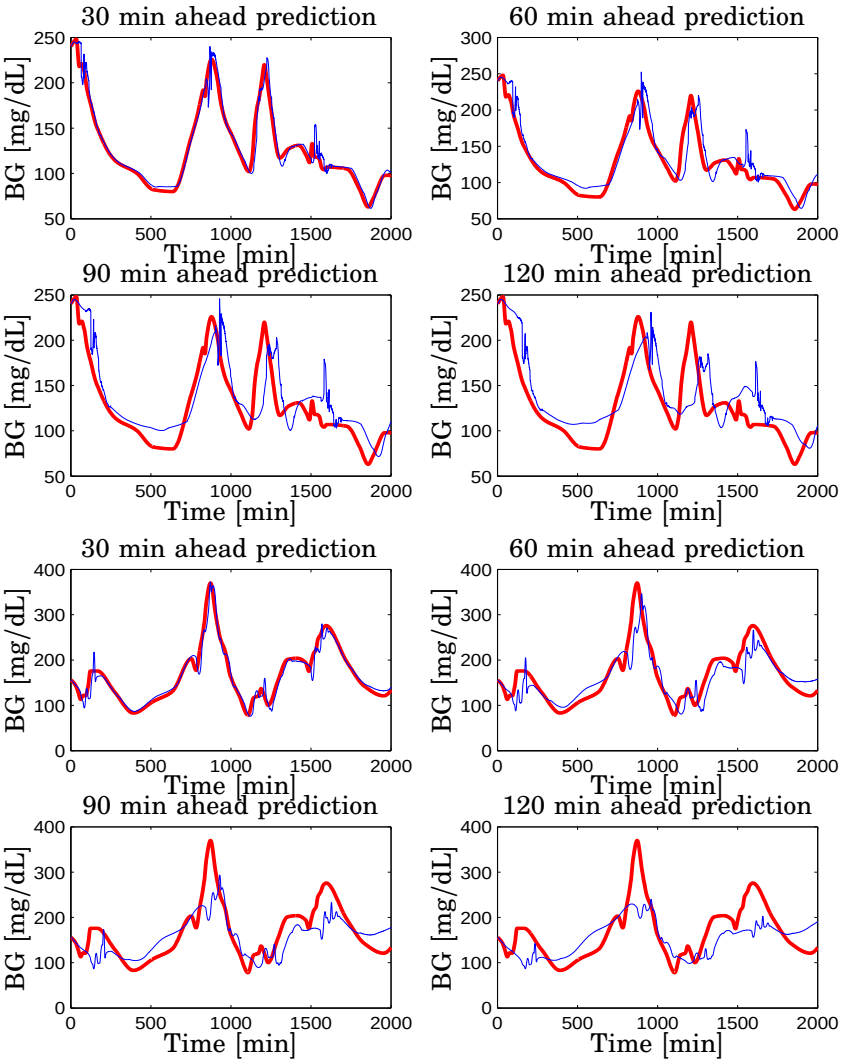

Figure E.3 *Top Panel* Patient CHU0115, ARMAX 6th-order model; *Bottom Panel* Patient CHU0120, ARMAX 4th-order model. Evaluation on validation data: model based predictor (blue) and measured blood glucose (red) [mg/dL] vs. time [min]. Top left 30 minutes; Top right 60 minutes; Bottom left 90 minutes; Bottom right 120 minutes ahead prediction

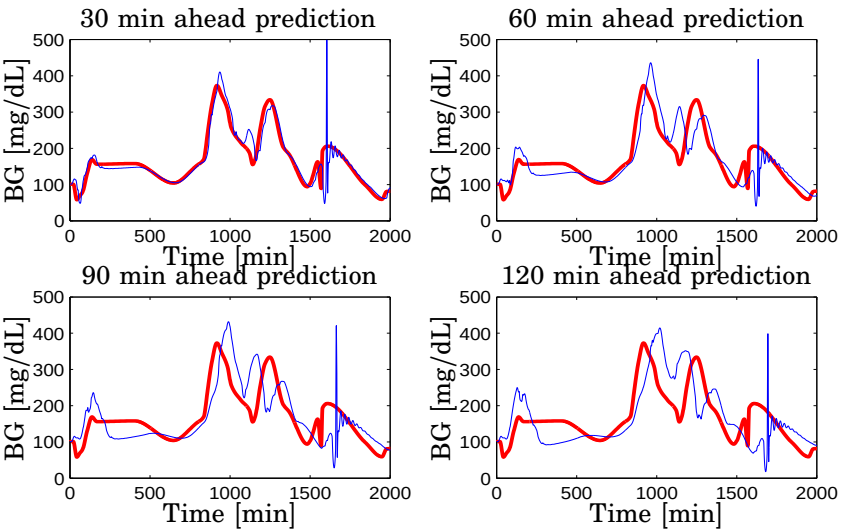

Figure E.4 Patient CHU0130, ARMAX 5th-order model. Evaluation on validation data: model based predictor (blue) and measured blood glucose (red) [mg/dL] vs. time [min]. Top left 30 minutes; Top right 60 minutes; Bottom left 90 minutes; Bottom right 120 minutes ahead prediction

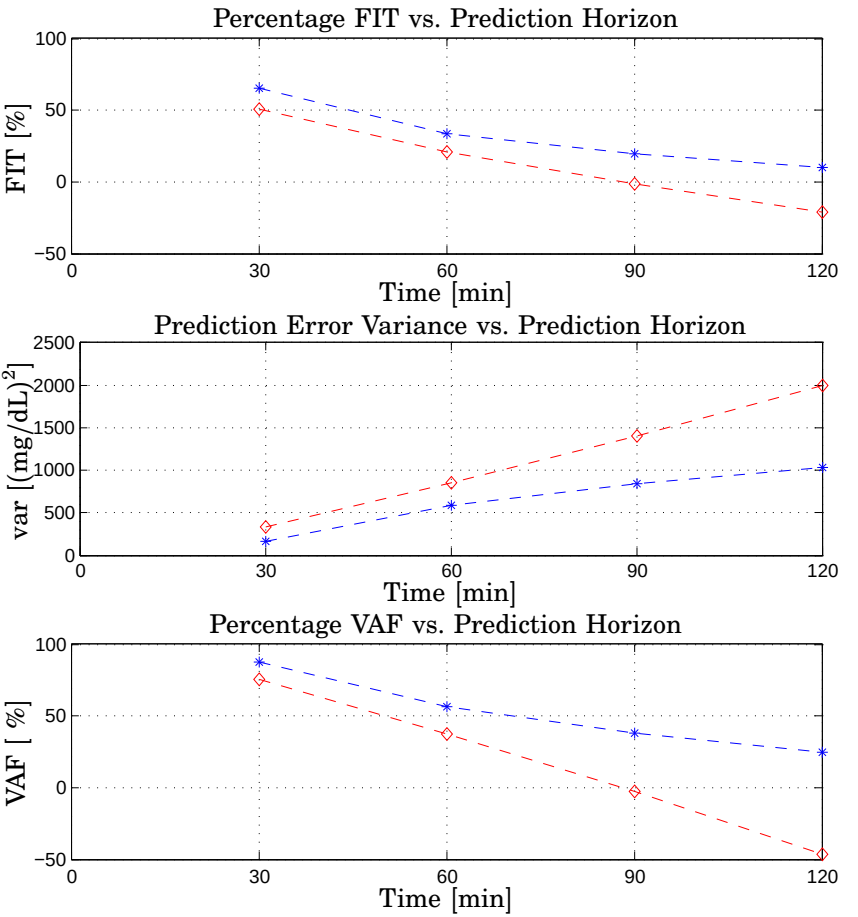

Figure E.5 Patient CHU0103. ARMAX 3rd-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

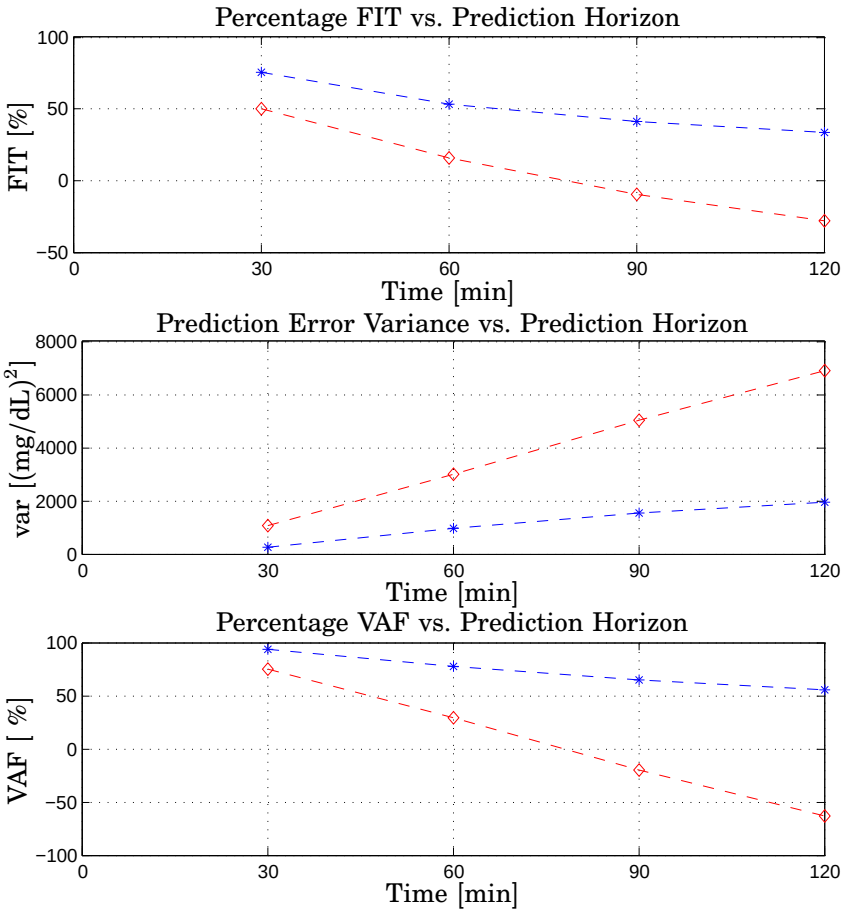

Figure E.6 Patient CHU0104. ARMAX 3rd-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance [$(mg/dL)^2$]; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

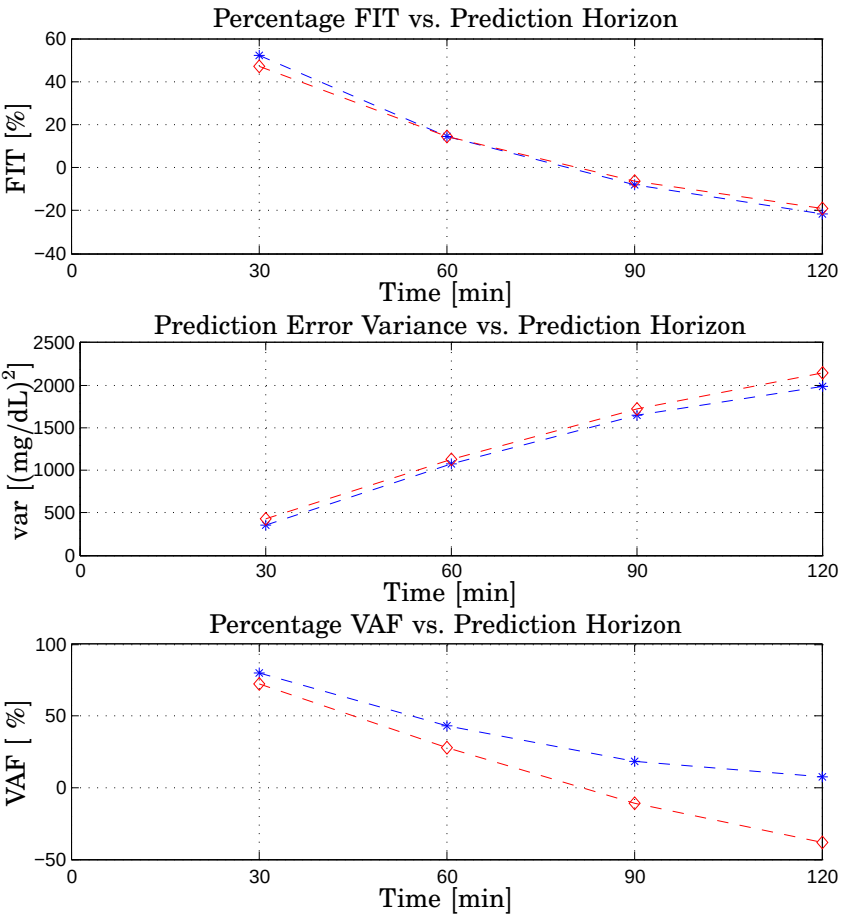

Figure E.7 Patient CHU0105. ARMAX 5th-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

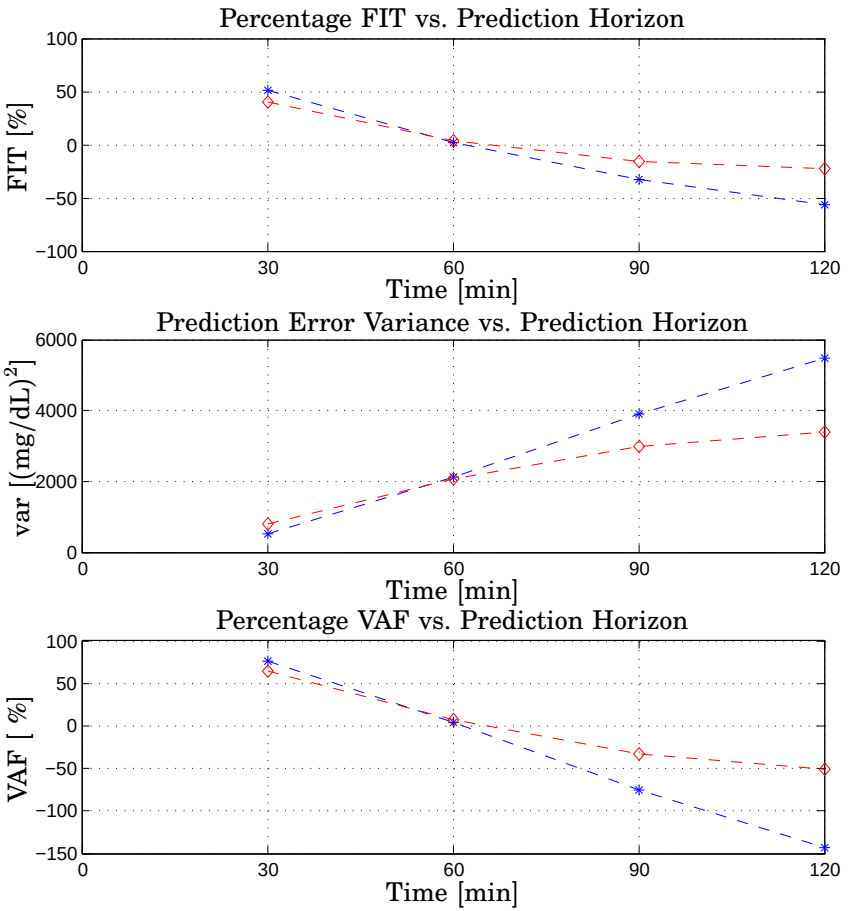

Figure E.8 Patient CHU0106. ARMAX 6th-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

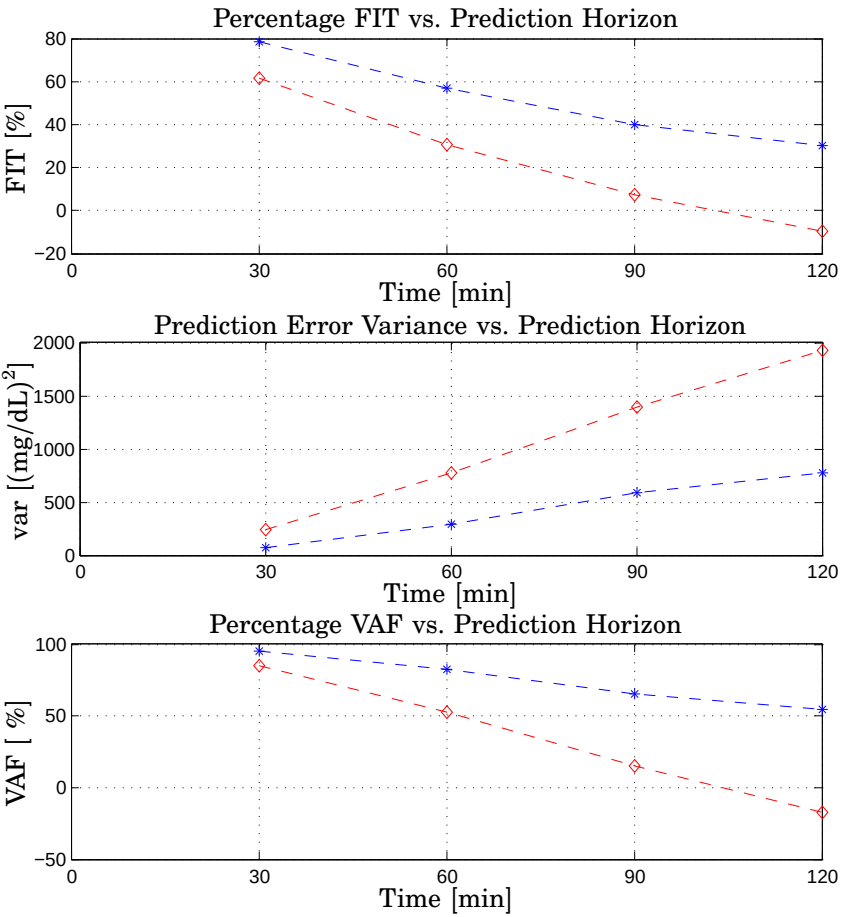

Figure E.9 Patient CHU0115. ARMAX 6th-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

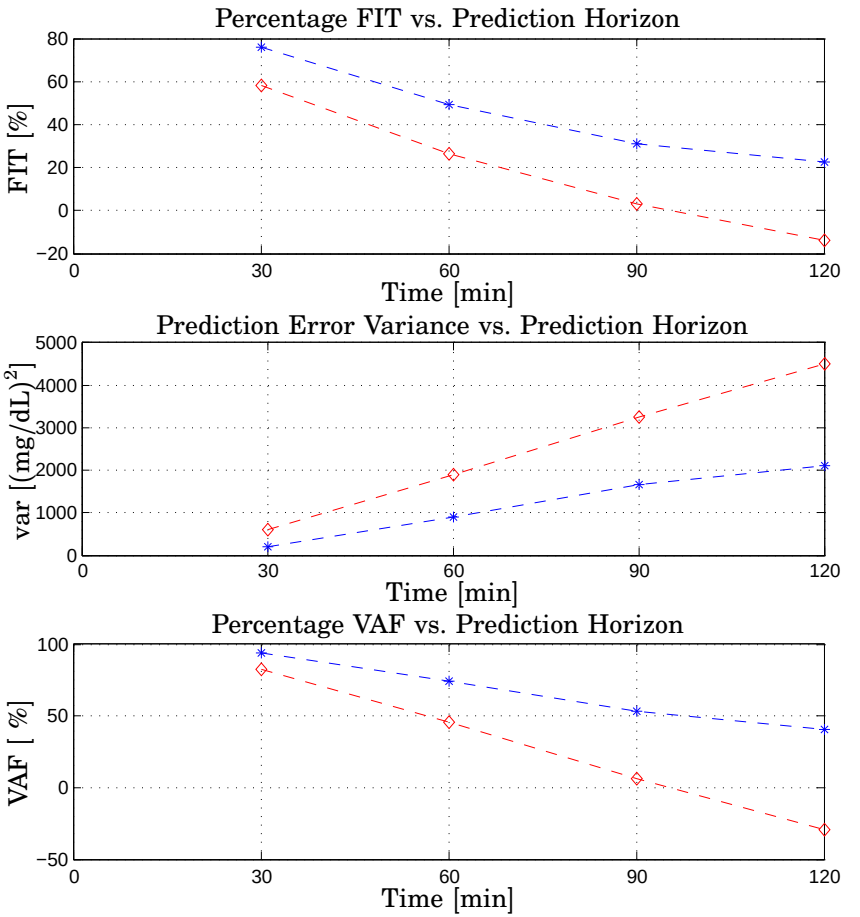

Figure E.10 Patient CHU0120. ARMAX 4th-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

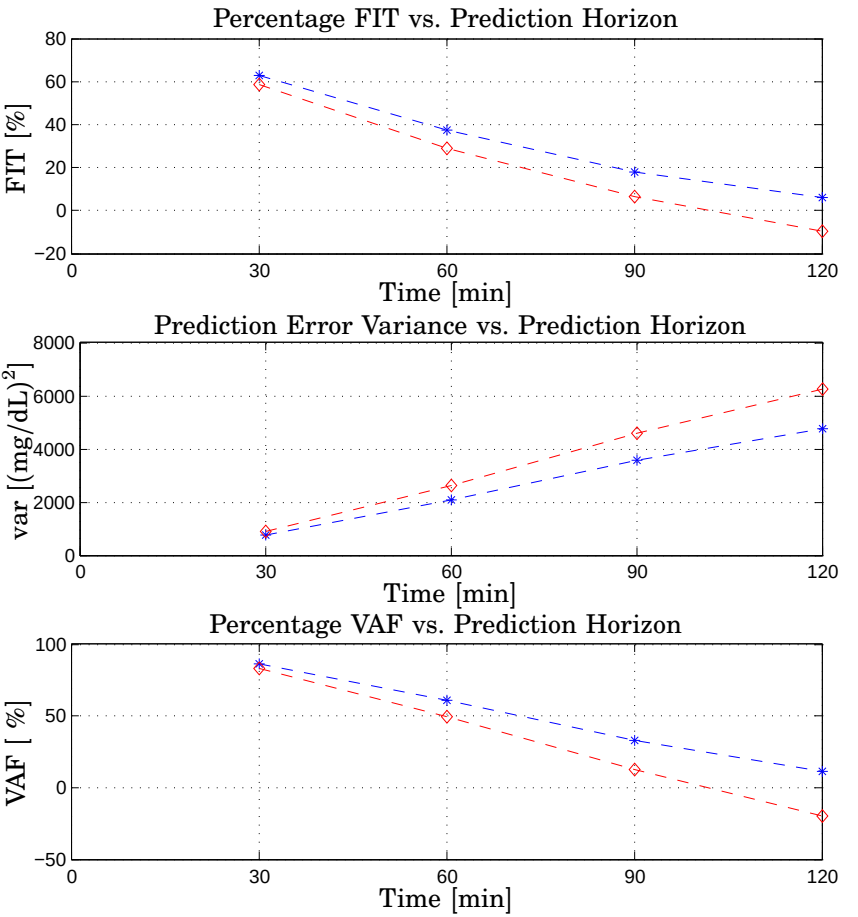

Figure E.11 Patient CHU0130. ARMAX 5th-order model. Model-based predictor (star), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

F

Linear Predictors

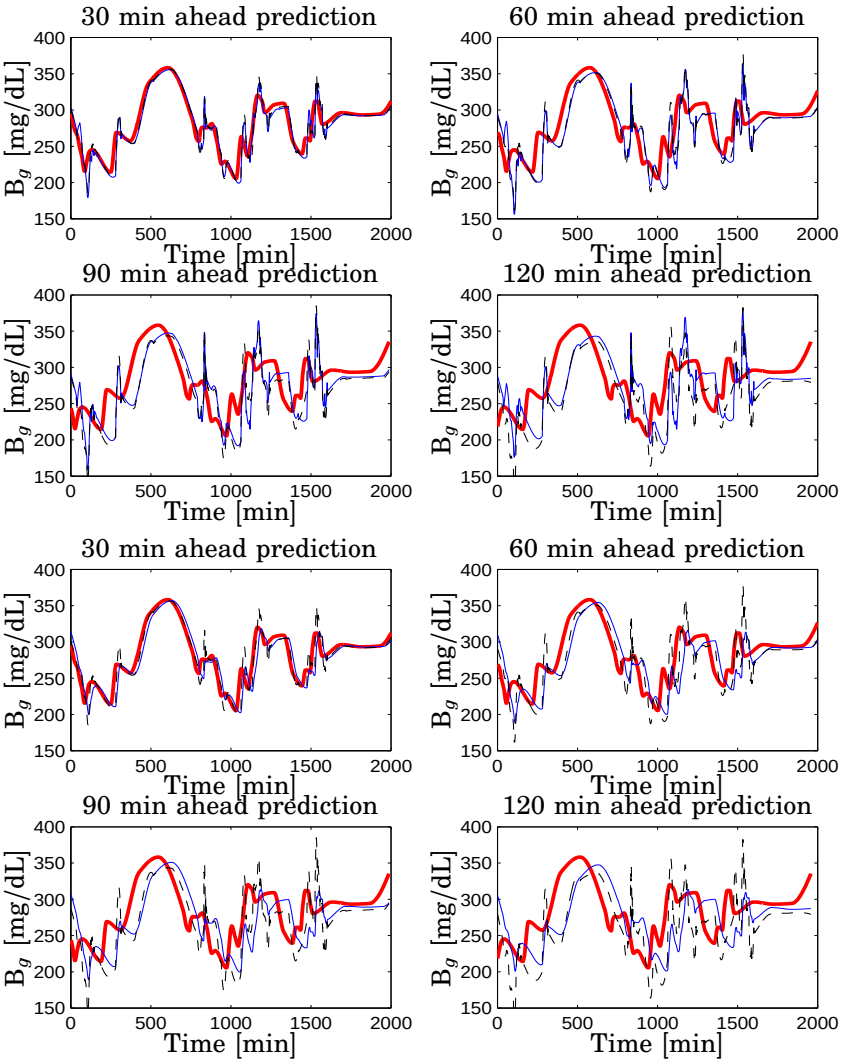

Figure F.1 Patient CHU0103. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

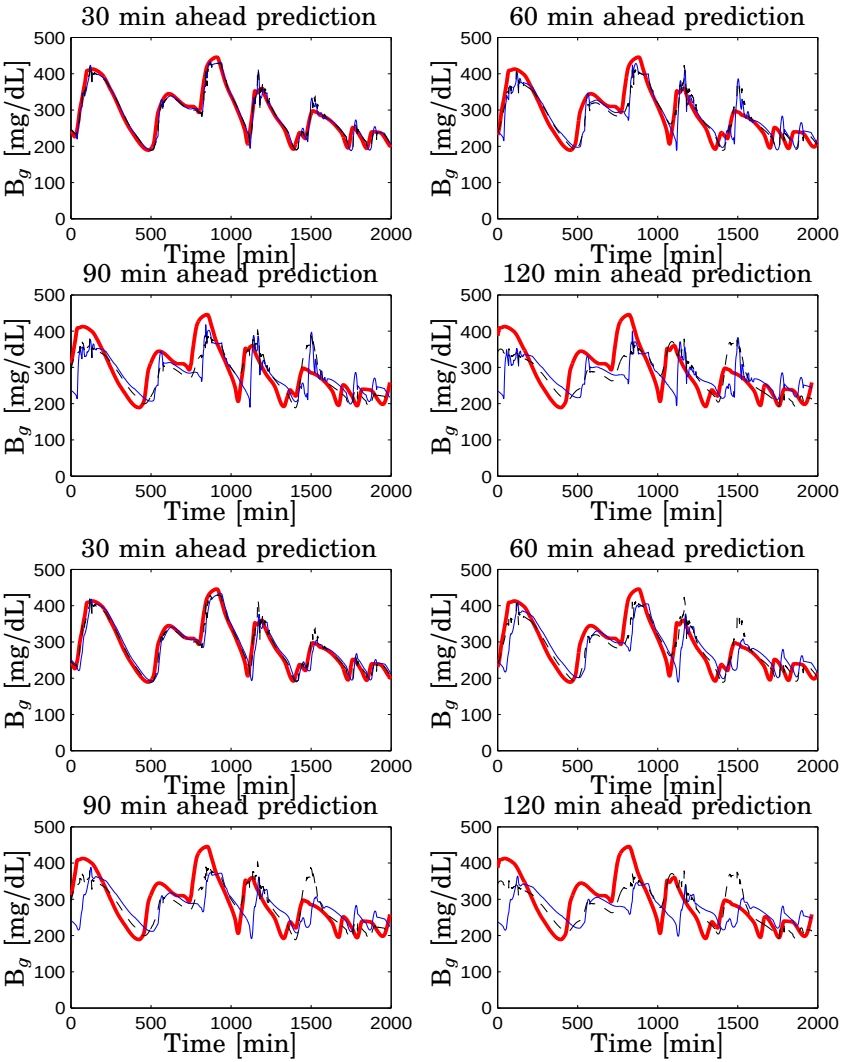

Figure F.2 Patient CHU0104. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

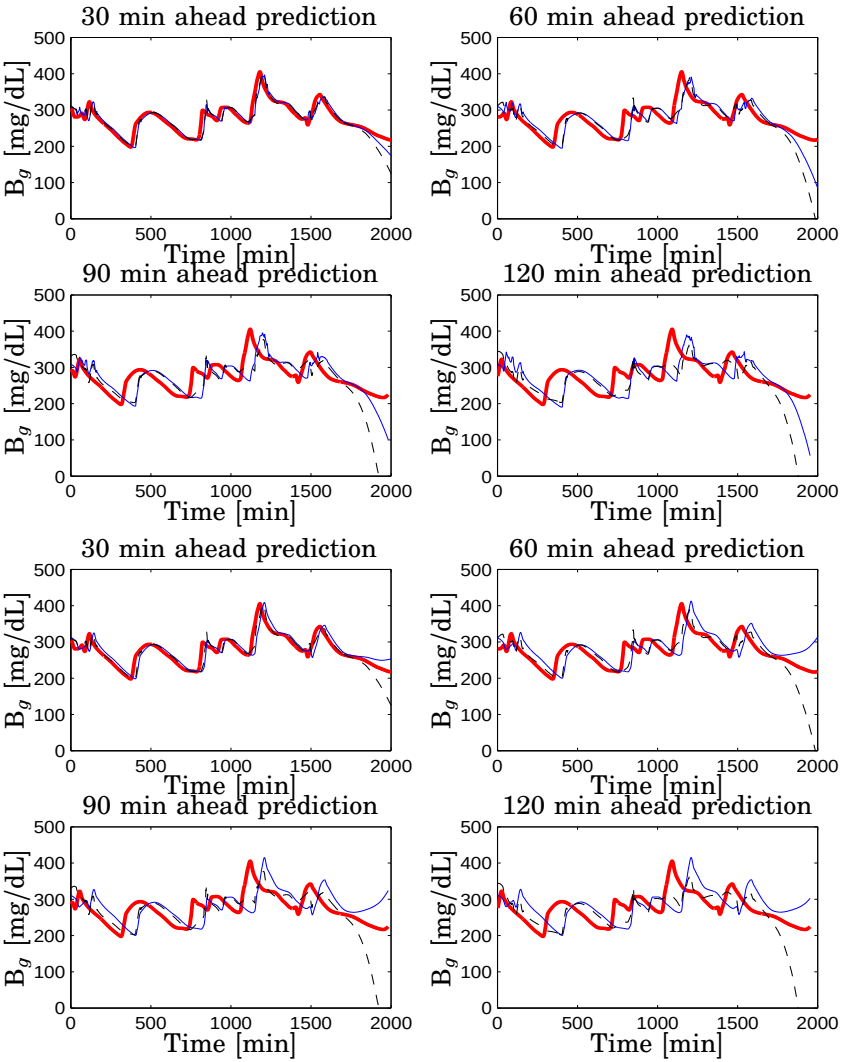

Figure F.3 Patient CHU0105. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

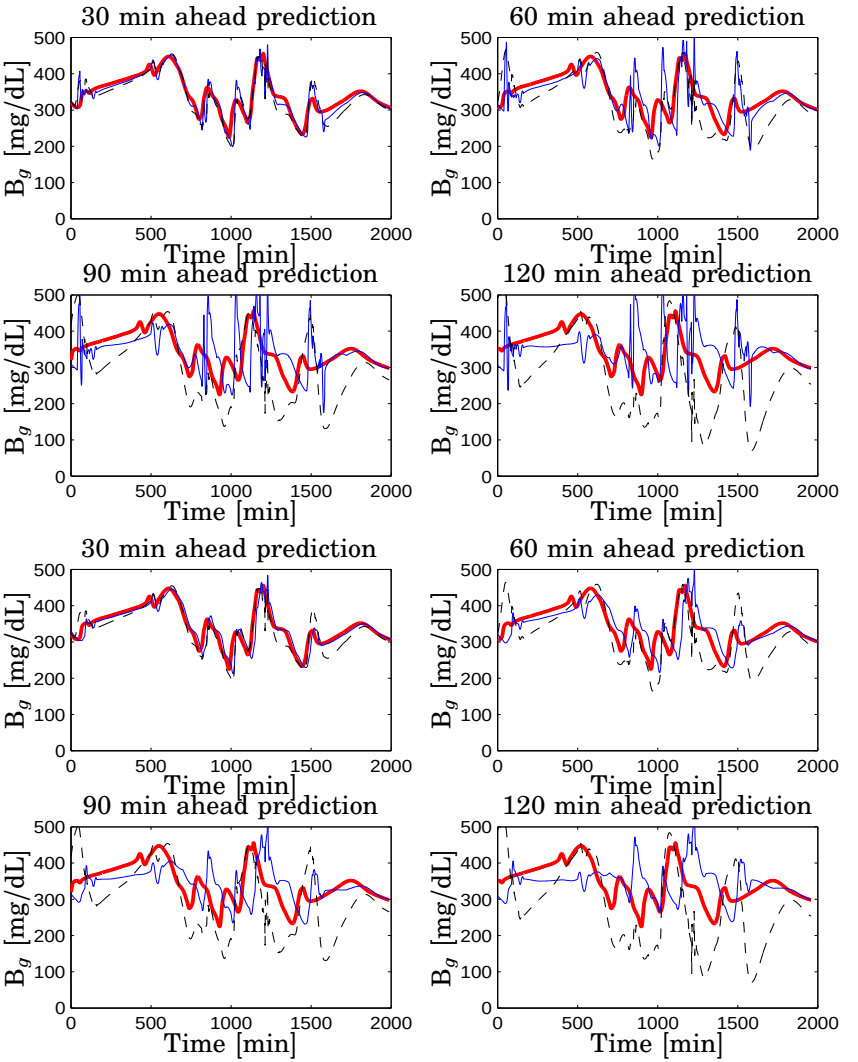

Figure F.4 Patient CHU0106. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

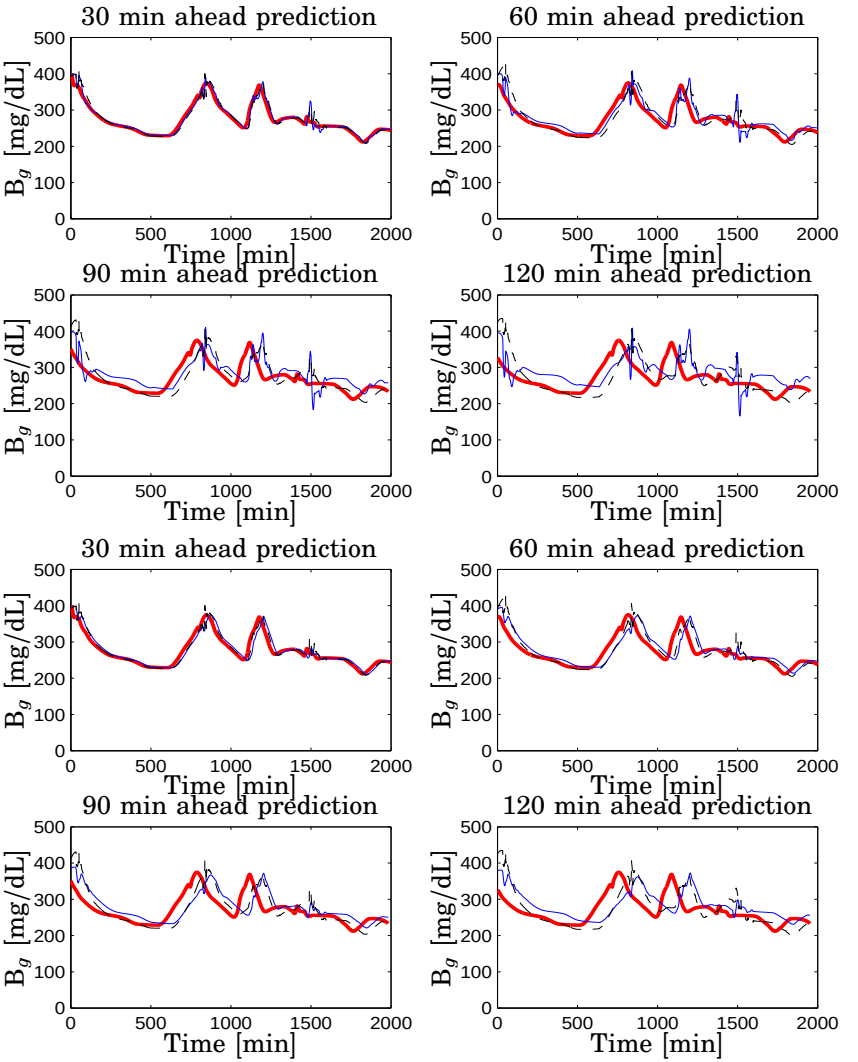

Figure F.5 Patient CHU0115. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

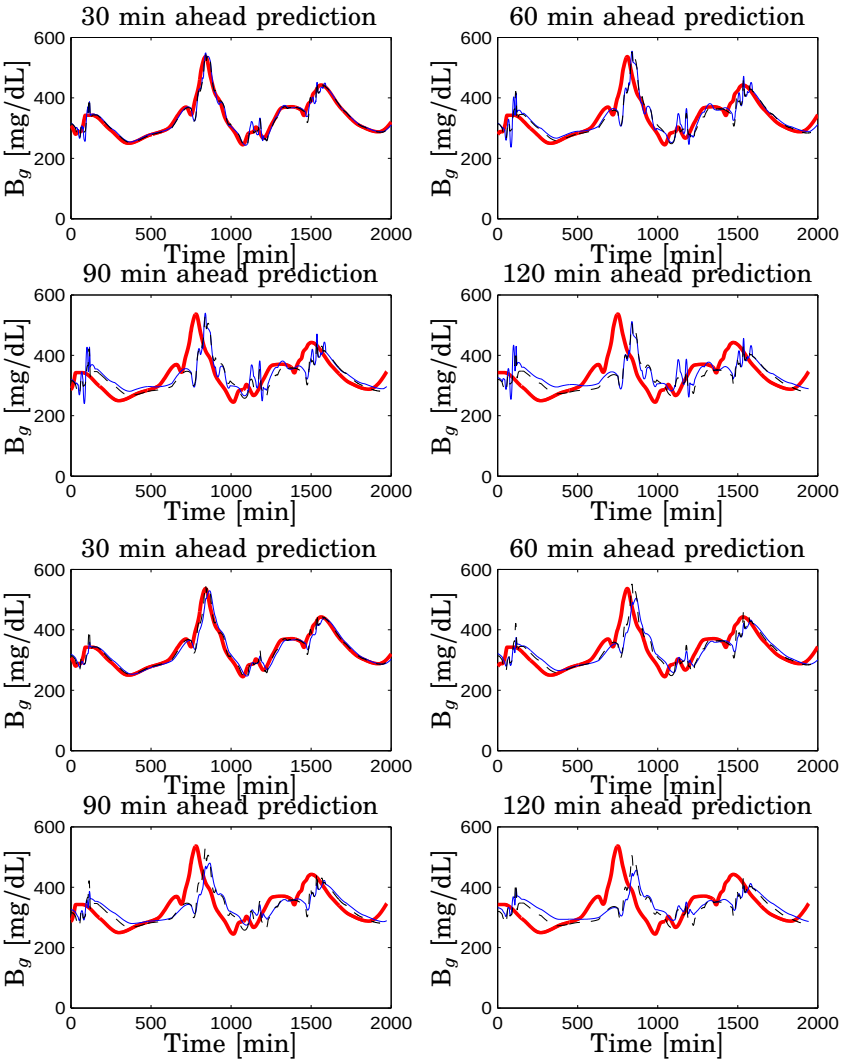

Figure F.6 Patient CHU0120. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

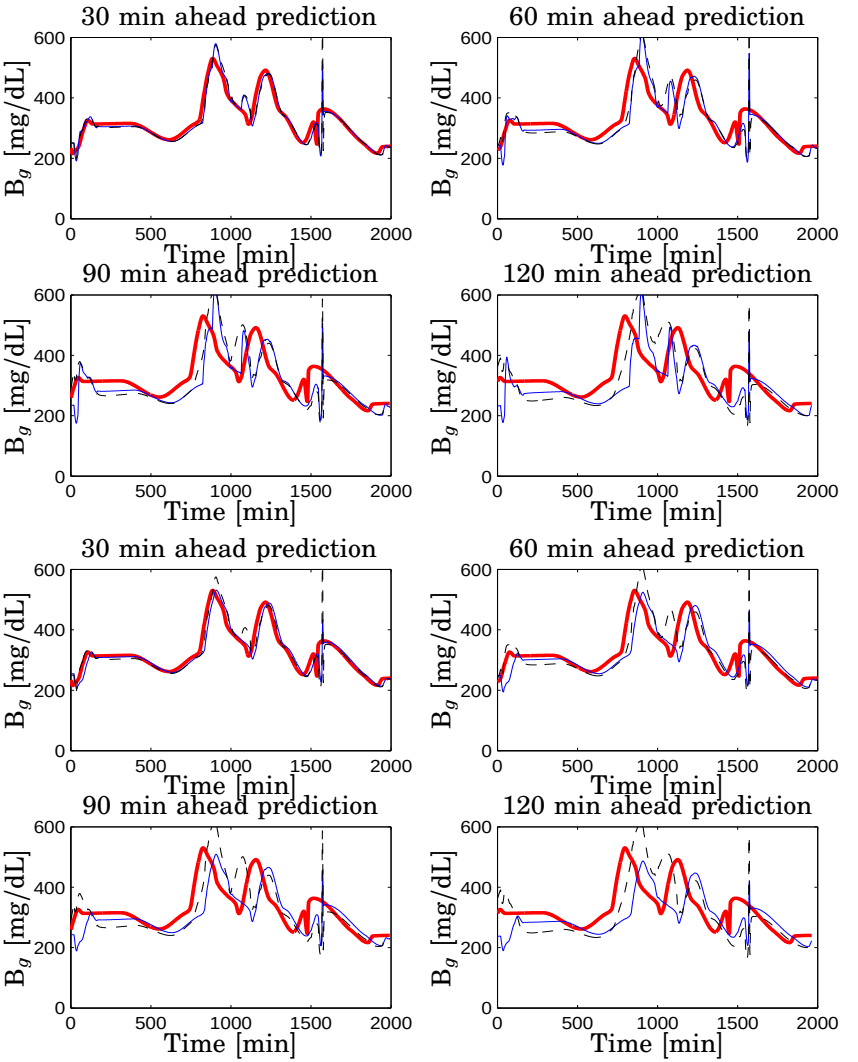

Figure F.7 Patient CHU0130. Evaluation on validation data. *Top Panel* 3rd-order ARX-based predictor (thin), 3rd-order ARMAX-based predictor (dashed) *Bottom Panel* Regularized 3rd-order ARX predictor (thin), 3rd-order ARMAX-based predictor (dashed) compared to actual blood glucose (thick) [mg/dL] vs. time [min]. Top left 30; Top right 60; Bottom left 90; Bottom right 120-minutes ahead prediction

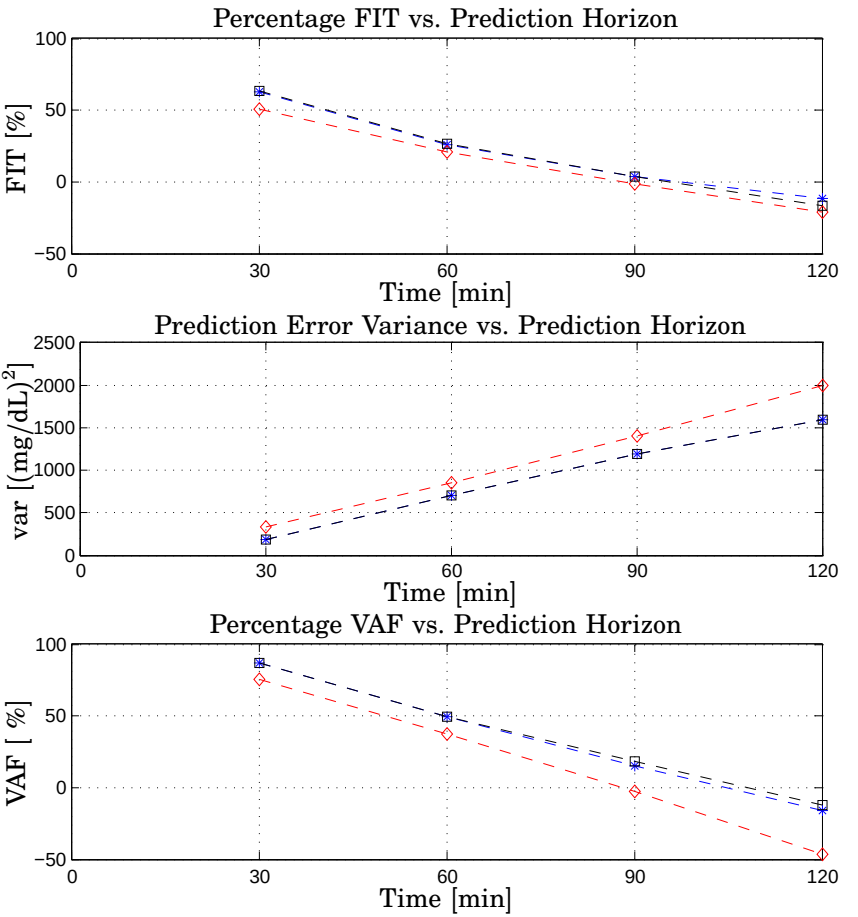

Figure F.8 Patient CHU0103. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

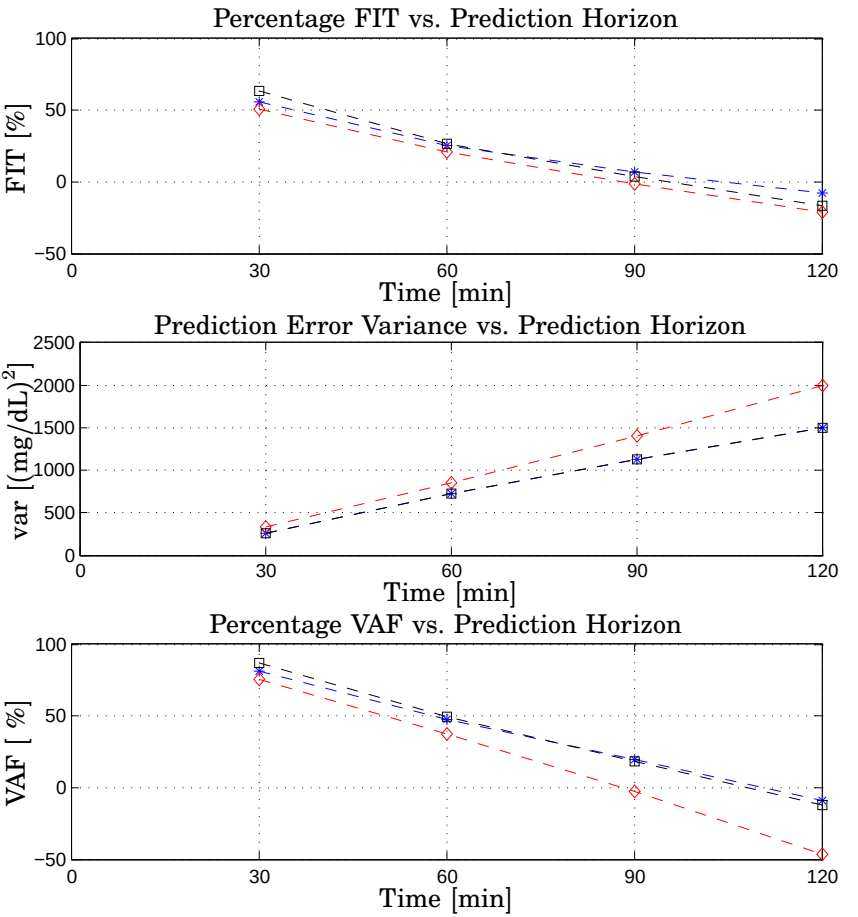

Figure F.9 Patient CHU0103. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

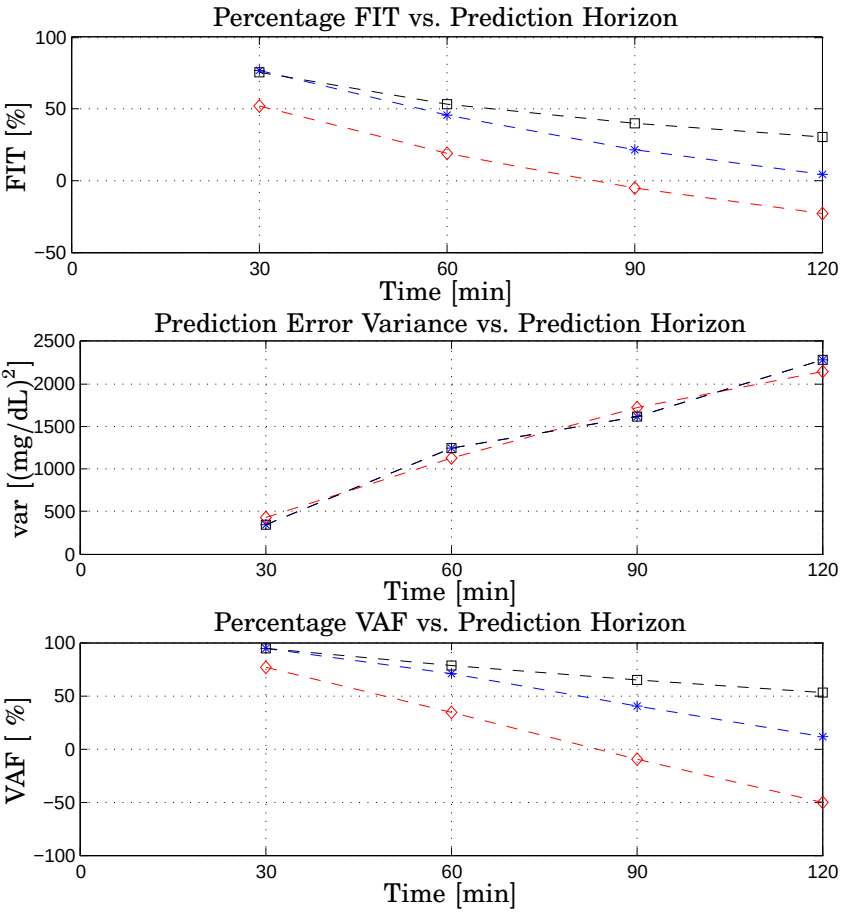

Figure F.10 Patient CHU0104. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance [$(mg/dL)^2$]; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

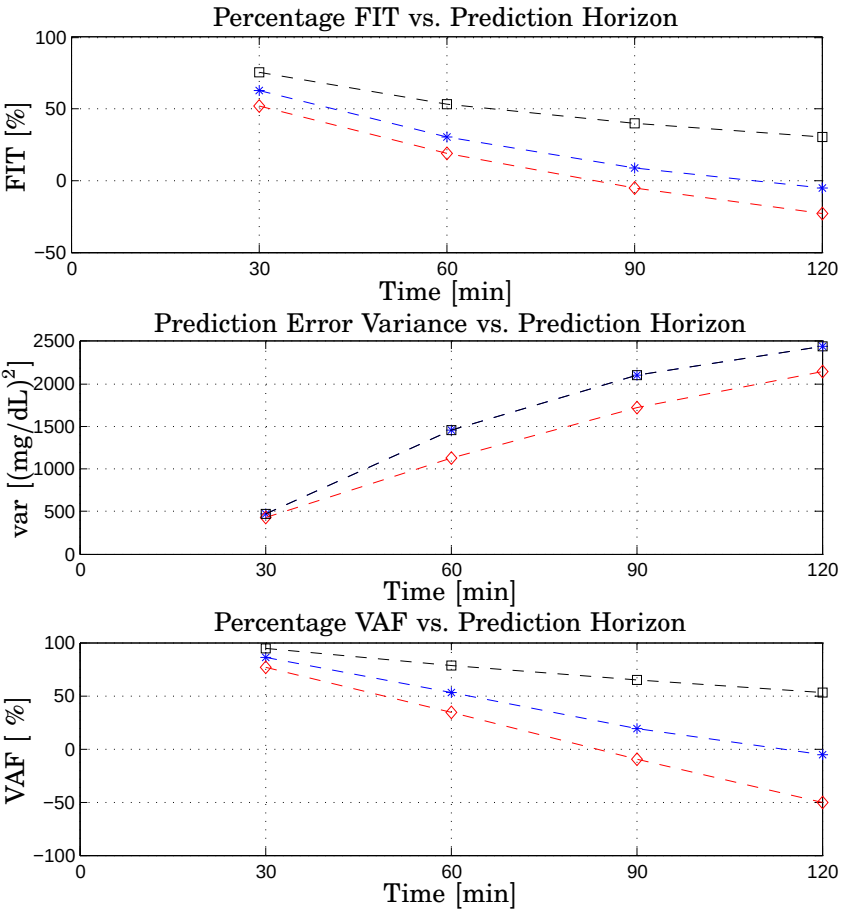

Figure F.11 Patient CHU0104. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

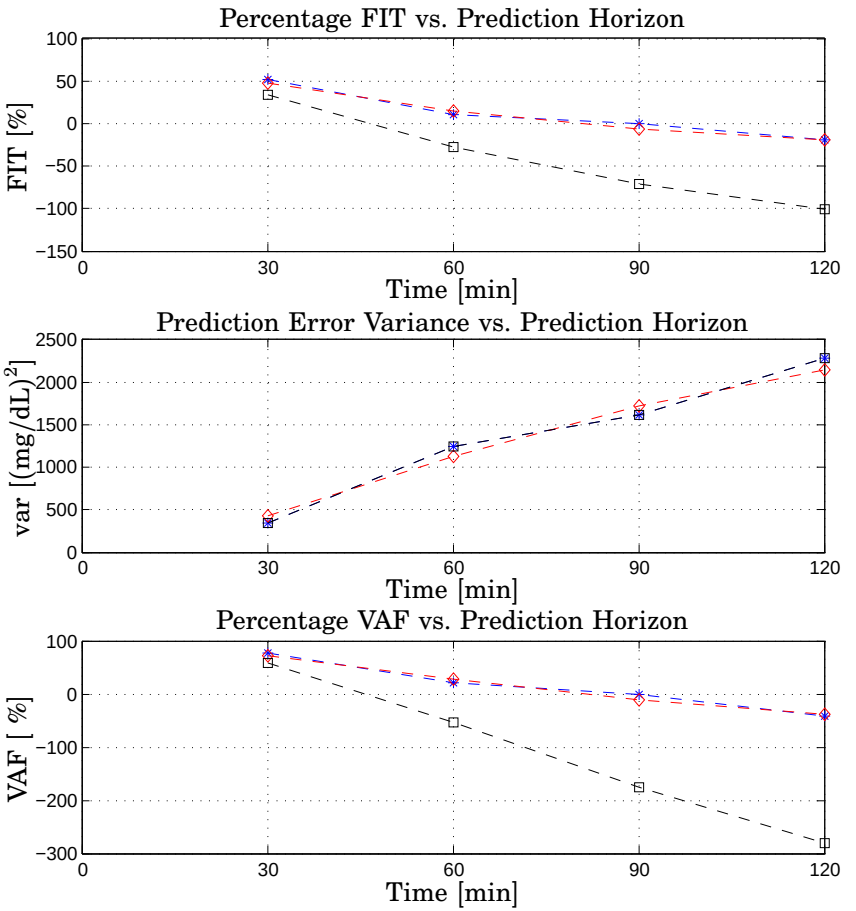

Figure F.12 Patient CHU0105. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

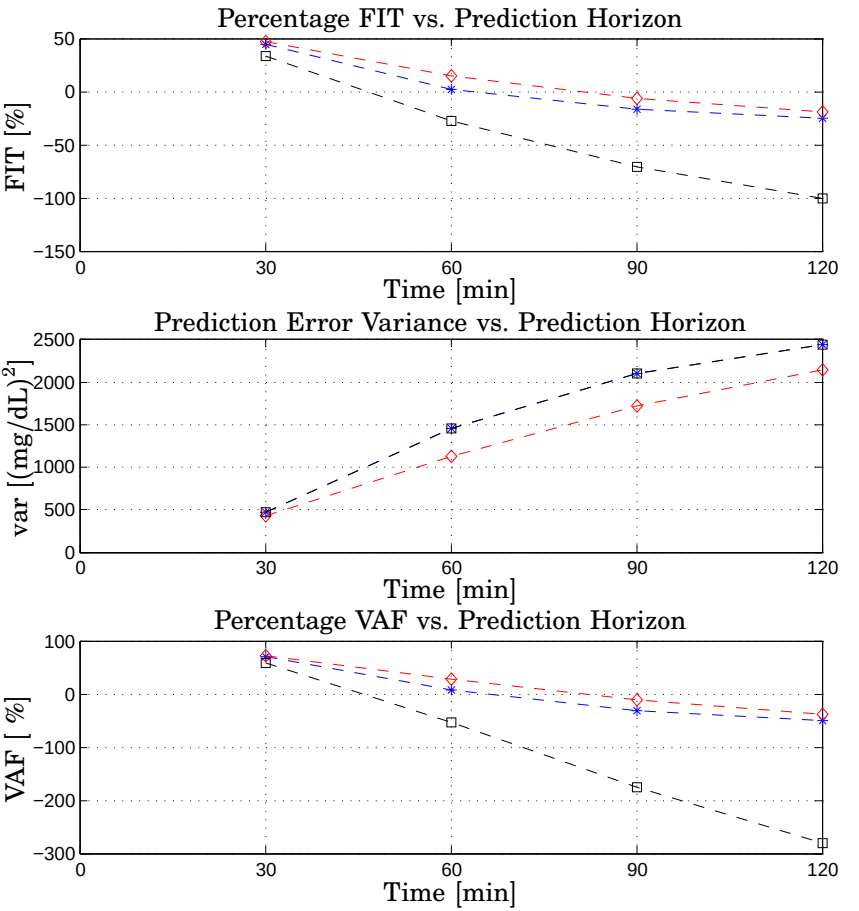

Figure F.13 Patient CHU0105. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

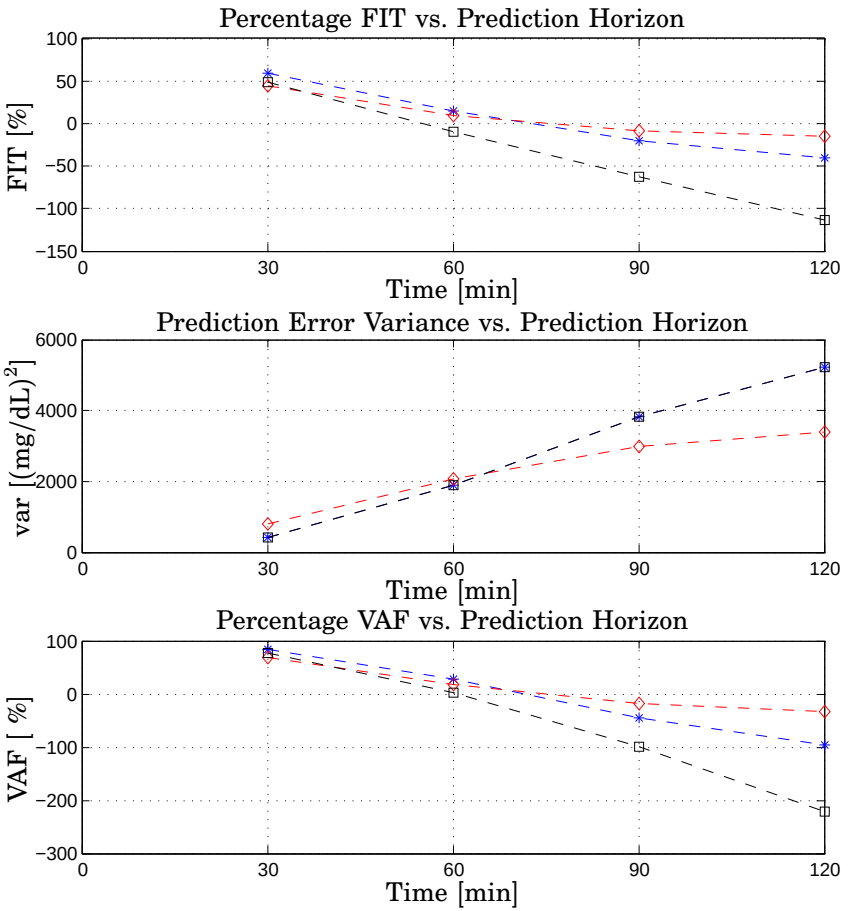

Figure F.14 Patient CHU0106. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

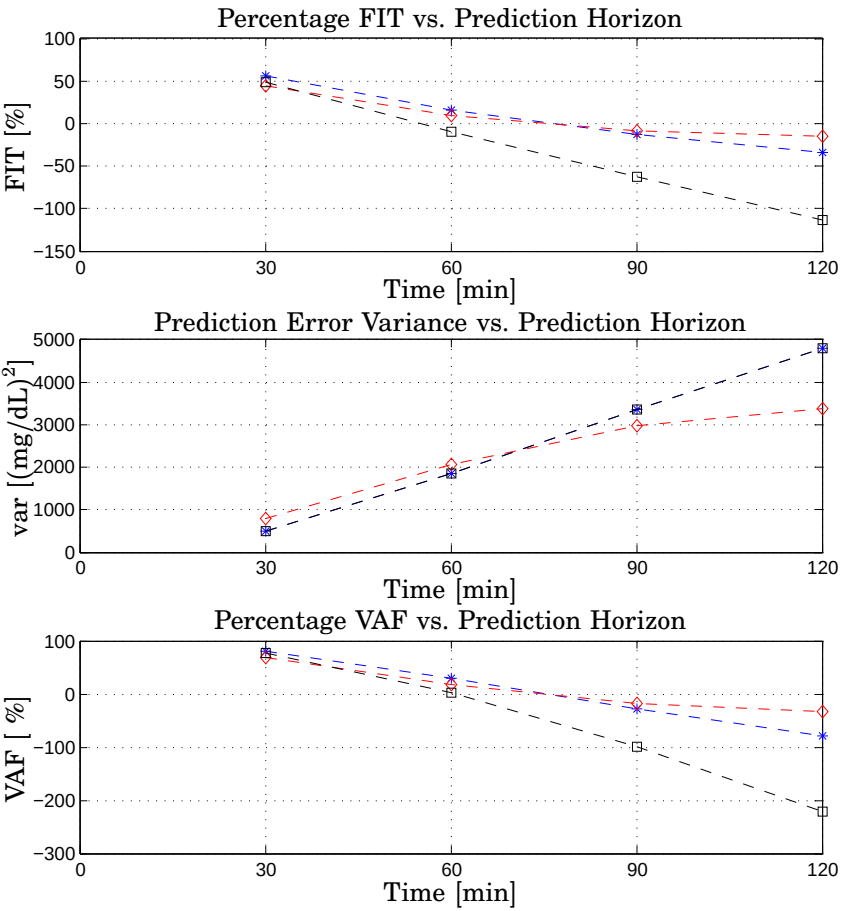

Figure F.15 Patient CHU0106. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

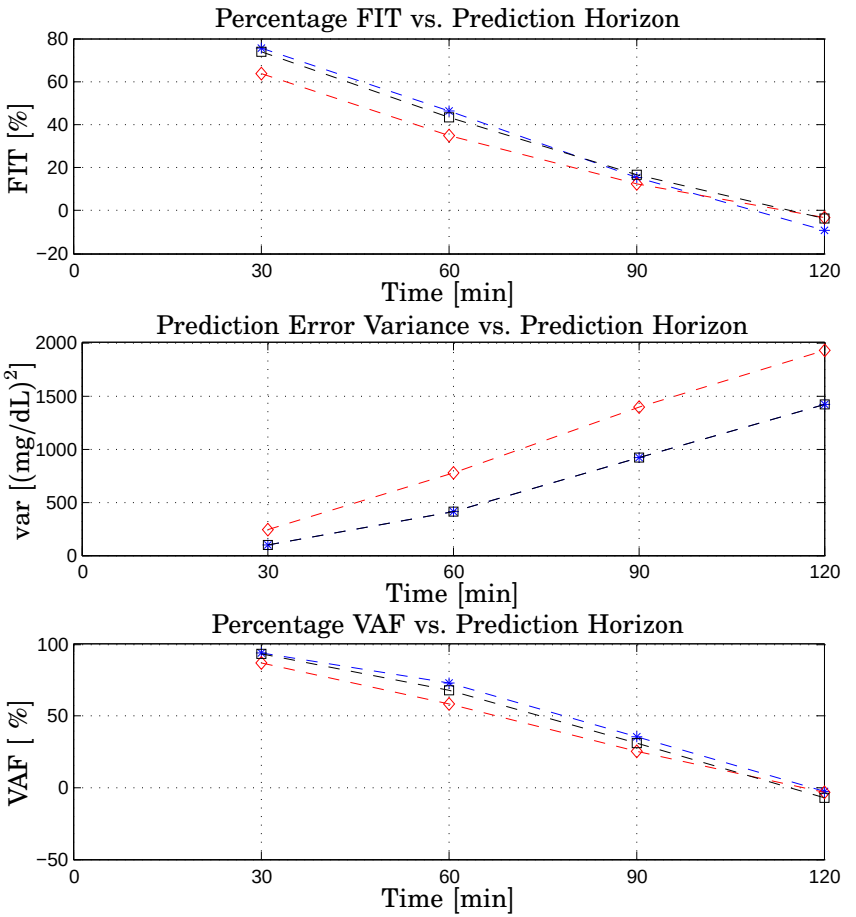

Figure F.16 Patient CHU0115. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

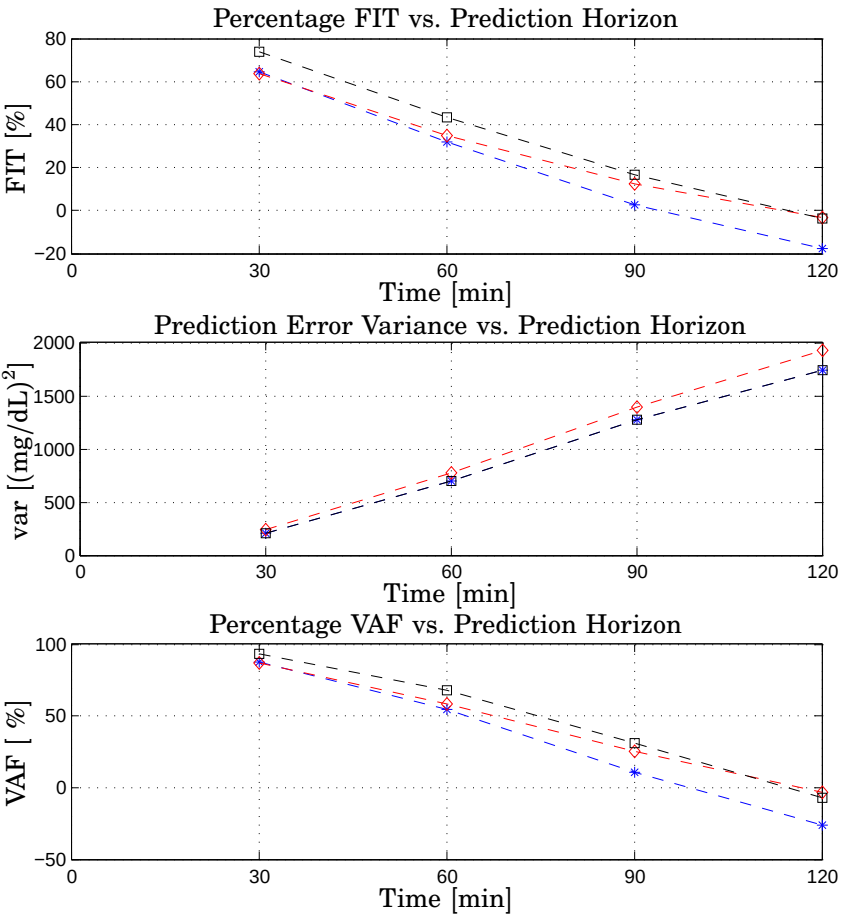

Figure F.17 Patient CHU0115. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\text{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

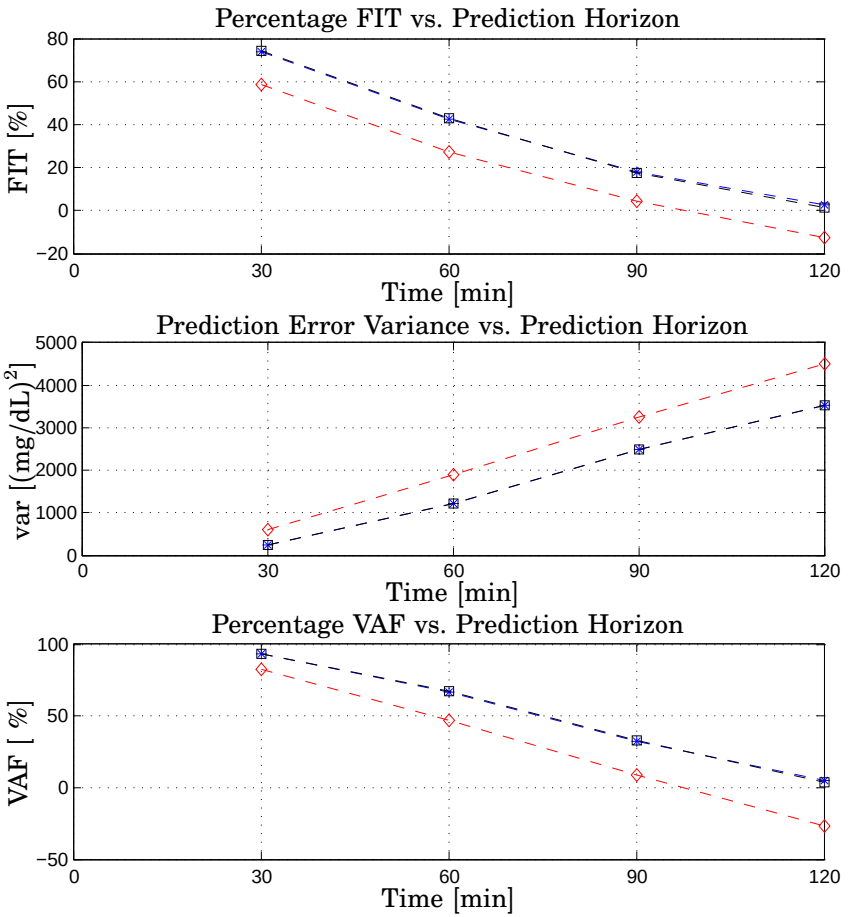

Figure F.18 Patient CHU0120. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

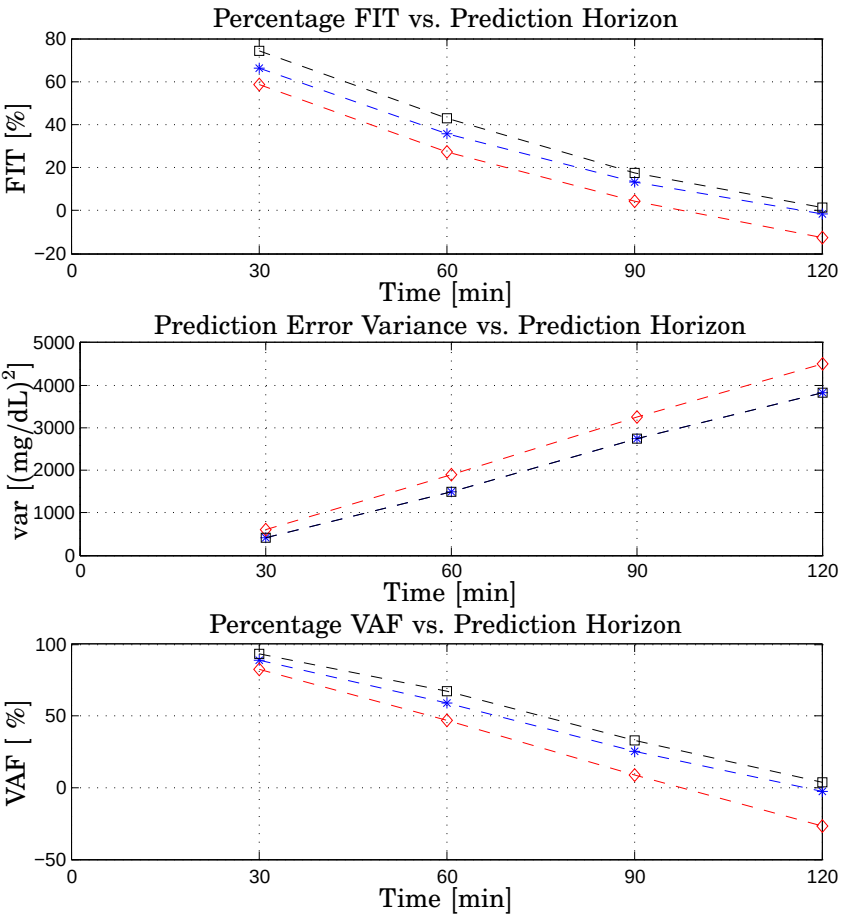

Figure F.19 Patient CHU0120. Regularized 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

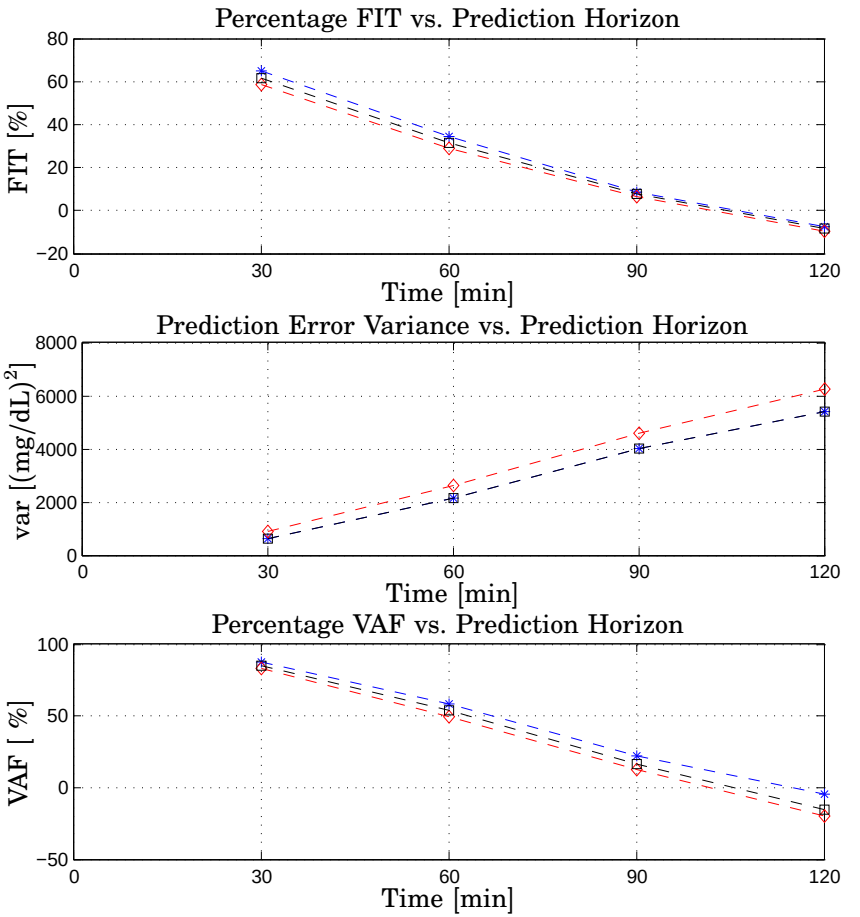

Figure F.20 Patient CHU0130. 3rd-order ARX-based predictor (star), 3rd-order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

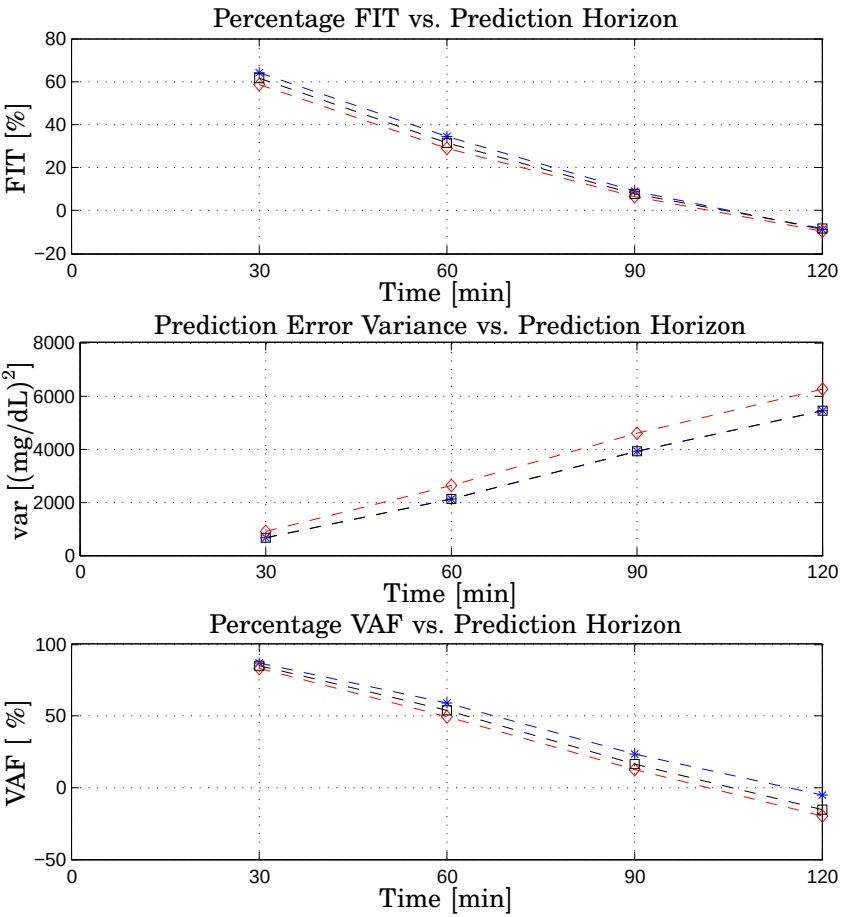

Figure F.21 Patient CHU0130. Regularized 3rd order ARX-based predictor (star), 3rd order ARMAX-based predictor (square), ZOH (diamond). *Top* Percentage FIT [%]; *Center* Prediction Error Variance $[(\mathrm{mg/dL})^2]$; *Bottom* Percentage VAF [%]. All the metrics on validation data vs. Prediction Horizon [min]

G

Bibliography

Abbott (2011): "Abbott Diabetes Care." `www.abbottdiabetescare.com`. Page retrieved May 2011.

Ackerman, J. and W. McGucking (1964): "A mathematical model of the glucose tolerance test." *Phys. Medicine Biol.*, **9**, pp. 203–213.

Anderson, B. and J. Moore (1979): *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ.

AP (2011): "The artificial pancreas project." `www.artificialpancreasproject.com/`. Page retrieved May 2011.

Åström, K.(1970): *Introduction to Stochastic Control Theory*. Academic Press, New York.

Åström, K. (1980): "Maximum likelihood and prediction error methods." *Automatica*, **16**, pp. 551–574.

Bamieh, B. and L. Giarre (2002): "Identification of linear parameter-varying models." *International Journal of Robust Nonlinear Control*, **12:9**, p. 841–853.

Barry Kenan, D., J. Mastrototaro, G. Voskanyan, and G. Steil (2009): "Delays in minimally invasive continuous glucose monitoring devices: A review of current technology." *Journal of Diabetes Science and Technology*, **3:5**, pp. 1207–1214.

Bergman, R., L. Phillips, and C. Cobelli (1981): "Physiologic evaluation of factors controlling glucose tolerance in man: Measurement of insulin sensitivity and beta-cell sensitivity from the response to

intravenous glucose." *Journal Clinical Investigation*, **68**, December, pp. 1456–1467.

Bergman, R. N., Y. Ziya Ider, C. R. Bowden, and C. Cobelli (1979): "Quantitative estimation of insulin sensitivity." *American Journal of Physiology*, **236:6**, pp. E667–E677.

Bolie, V. (1961): "Coefficients of normal blood glucose regulation." *J. Appl. Physiol.*, **16**, pp. 783–788.

Bremer, T. and D. A. Gough (1999): "Is blood glucose predictable from previous values?" *Diabetes*, **48**, March, pp. 445–451.

Breton, M., D. Shields, and B. Kovatchev (2008): "Optimum subcutaneous glucose sampling and fourier analysis of continuous glucose monitors." *Journal of Diabetes Science and Technology*, **2**, pp. 495–500.

Brouns, F., I. Bjork, K. Frayn, A. Gibbs, V. Lang, G. Slama, and T. Wolever (2005): "Glycaemic index methodology." *Nutrition Research Review*, **18**, pp. 145–171.

Cameron, F., G. Niemayer, K. Gundy-Burlet, and B. Buckingham (2008): "Statistical hypoglycemia prediction." *Journal of Diabetes Science and Technology*, **2:4**, pp. 612–621.

Cameron, F., G. Niemeyer, and B. Buckingham (2009): "Probabilistic evolving meal detection and estimation of meal total glucose appearance." *Journal of Diabetes Science and Technology*, **3:5**, pp. 1022–1030.

Castillo-Estrada, G., L. Del Re, and E. Renard (2010): "Nonlinear gain in online prediction of blood glucose profile in type 1 diabetic patients." In *Proc. 49th IEEE Conference on Decision and Control (CDC2009)*, pp. 1668–1673. Atlanta, GA, USA.

Castillo Estrada, G., H. Kirchsteiger, L. Del Re, and E. Renard (2009): "Model based validation of meal inputs in diabetes therapy." In *Proc. of System Identification Symposium (SYSID2009)*, pp. 239–244. Saint-Malo, France.

Castillo-Estrada, G., H. Kirchsteiger, L. Del Re, and E. Renard (2010): "Innovative approach for online prediction of blood glucose profile

in type 1 diabetes patients." In *Proc. American Control Conference ACC2010*, pp. 2015–2020. Baltimore, MD, USA.

Cescon, M. (2011): "Adaptive subspace-based prediction of T1DM glycemia." In *Proc. 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC2012)(Submitted)*. Orlando, FL.

Cescon, M., I. Dressler, R. Johansson, and A. Robertsson (2009): "Subspace-based identification of compliance dynamics of parallel kinematic manipulator." In *Proc. 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM2009)*, pp. 1028–1033. Singapore.

Cescon, M. and R. Johansson (2009): "Glycemic trend prediction using empirical model identification." In *Proc. 48th IEEE Conference on Decision and Control (CDC2009)*, pp. 3501–3506. Shanghai, P.R.China.

Cescon, M. and R. Johansson (2010): "Multi-step-ahead multivariate predictors: a comparative analysis." In *Proc. 49th IEEE Conference on Decision and Control (CDC2010)*, pp. 2837–2842. Atlanta, USA.

Cescon, M. and R. Johansson (2011): "On data-driven multistep subspace-based linear predictors." In *Proc. 18th IFAC World Congress*. Milano, Italy.

Cescon, M., F. Ståhl, R. Johansson, and M. Landin-Olsson (2009a): "Short-term diabetes blood glucose prediction based on blood glucose measurements." In *Proc. 2nd International Conference on Advanced Technologies and Treatments for Diabetes (ATTD2009)*. Athens, Greece.

Cescon, M., F. Ståhl, M. Landin-Olsson, and R. Johansson (2009b): "Subspace-based model identification of diabetic blood glucose dynamics." In *Proc. 15th Symposium on System Identification (SYSID2009)*, pp. 233–238. Saint-Malo, France.

Chevillon, G. (2007): "Direct multi-step estimation and forecasting." *Journal of Economic Surveys*, **21:4**, pp. 746 – 785.

Chiuso, A. (2005): "On the relation between CCA and predictor-based subspace identification." In *Proc. 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC2005)*, pp. 4976–4982. Seville.

Chiuso, A. (2007): "The role of vector autoregressive modeling in subspace identification." *Automatica*, **43:6**, pp. 1034–1048.

Cobelli, C. and E. Carson (2008): *Introduction to Modeling in Physiology and Medicine*. Elsevier, New York.

Cobelli, C., C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, and B. Kovatchev (2009): "Diabetes: Models, signals and control." IEEE Reviews in Biomedical Engineering.

Dalla Man, C., M. Camilleri, and C. Cobelli (2006): "A system model of oral glucose absorption: Validation on gold standard data." *IEEE Transactions on Biomedical Engineering*, **53:12**, pp. 2472–2477.

Dalla Man, C., A. Caumo, R. Basu, R. Rizza, G. Toffolo, and C. Cobelli (2004): "Minimal model estimation of glucose absorption and insulin sensitivity from oral test: validation with a tracer method." *American Journal of Physiology Endocrinology Metabolism*, **287**, May, pp. E637–E643.

Dalla Man, C., A. Caumo, and C. Cobelli (2002): "The oral glucose minimal model: Estimation of insulin sensitivity from a meal test." *IEEE Transactions on Biomedical Engineering*, **49:5**, pp. 419–429.

Dalla Man, C., R. R. Rizza, and C. Cobelli (2007): "Meal simulation model of the glucose-insulin system." *IEEE Transactions on Biomedical Engineering*, **54:10**, pp. 1740–1749.

Dalla Man, C., K. E. Yarasheski, A. Caumo, H. Robertson, G. Toffolo, K. S. Polonski, and C. Cobelli (2005): "Insulin sensitivity by oral glucose minimal models: validation against clamp." *American Journal of Physiology Endocrinology Metabolism*, **289**, July, pp. E954–E959.

De Gaetano, A. and O. Arino (2000): "Mathematical modelling of the intravenous glucose tolerance test." *Journal of Mathematical Biology*, **40**, pp. 136–168.

De Nicolao, G., G. Sparacino, and C. Cobelli (1997): "Nonparametric input estimation in physiological systems: Problems, methods and case studies." *Automatica*, **33:5**, pp. 851–870.

DIAdvisor (2011): "The DIAdvisor^{TM}." `www.diadvisor.eu`.

Eren-Oruklu, M., A. Cinar, L. Quinn, and D. Smith (2009): "Estimation of future glucose concentration with subject-specific recursive linear models." *Diabetes Technology and Terapeutics*, **11:4**, pp. 243–253.

Fernandez, M., M. Villasana, and D. Streja (2007): "Glucose dynamics in type I diabetes: Insights from the classic and linear minimal models." *Computers in Biology and Medicine*, **37**, pp. 611–627.

Finan, D. A., C. C. Palerm, F. J. Doyle, D. E. Seborg, H. Zisser, W. C. Bevier, and L. Jovanovic (2009): "Effect of input excitation on the quality of empirical dynamic models for type 1 diabetes." *AIChE Journal*, **55:5**, pp. 1135–1146.

Finan, D. A., C. C. Palerm, F. J. Doyle, H. Zisser, L. Jovanovic, W. C. Bevier, and D. E. Seborg (2008): "Identification of empirical dynamic models from type 1 diabetes subject data." In *Proc. of American Control Conference (ACC2008)*, pp. 2099–2104. Seattle, Washington, USA.

Finan, D. A., H. Zisser, L. Jovanovic, W. C. Bevier, and D. E. Seborg (2006): "Identification of linear dynamic models for type 1 diabetes: a simulation study." In *Proc. of IFAC International Symposium on Advanced Control of Chemical Processes (ADCHEM2006)*. Gramado, Brazil.

Finan, D. A., H. Zisser, L. Jovanovic, W. C. Bevier, and D. E. Seborg (2007): "Practical issues in the identification of empirical models from simulated type 1 diabetes data." *Diabetes Technology and Therapeutics*, **9:5**, pp. 438–450.

G. Markakis, M., G. D. Mitsis, and V. Z. Marmarelis (2008): "Computationsl study of an augmented minimal model for glycaemia control." In *Proc. of 30th IEEE EMBS Eng. Medicine and Biology Society Conference (EMBC2008)*, pp. 5445–5448. Vancouver, British Columbia, Canada.

Gani, A., A. Gribok, L. Yinghui, W. Ward, R. Vigersky, and J. Reifman (2010): "Universal glucose models for predicting subcutaneous glucose concentration in humans." *IEEE Transactions on Information Technology in Biomedicine*, **14:1**, pp. 157–165.

Gani, A., A. V. Gribok, S. Rajaraman, W. K. Ward, and J. Reifman (2009): "Predicting subcutaneous glucose concentration in humans: Data-driven glucose modeling." *IEEE Transactions on Biomedical Engineering*, **56:2**, pp. 246–254.

Golub, G. and C. Van Loan (1996): *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD.

Grossman, P. (2004): "The lifeshirt: a multi-function ambulatory system monitoring health, disease, and medical intervention in the real world." *Studies in Health Technology and Informatics*, **108**, pp. 133–141.

Guerra, S., G. Sparacino, A. Facchinetti, M. Schiavon, C. Dalla Man, and C. Cobelli (2011): "A dynamic risk measure from continuous glucose monitoring data." *Diabetes Technology and Therapeutics (Accepted for publication)*.

Hann, C., G. Chase, J. Lin, T. Lotz, C. Doran, and G. Shaw (2005): "Integral-based parameter identification for long-term dynamic verification of a glucose-insulin system model." *Computer Methods and Programs in Biomedicine*, **77**, pp. 259–270.

Hannan, E. and D. Poskitt (1988): "Unit canonical correlations between future and past." *The Annals of Statistics*, **16:2**, pp. 784–790.

Haverkamp, B. and M. Verhaegen (1997): *SMI Toolbox: State space Model Identification software for multivariable dynamical systems*. TU Delft, Delft, The Netherlands, 1.0 edition.

Hemocue (2011): "Hemocue." `www.hemocue.com`. Page retrieved May 2011.

Holst, J. (1977): *Adaptive Prediction and Recursive Estimation*. PhD thesis, Department of Automatic Control, Lund Institute of Technology, Sweden.

Hovorka, R., V. Canonico, L. Chassin, U. Haueter, M. Massi-Benedetti, M. Orsini Federici, T. Pieber, H. Schaller, L. Schaupp, T. Vering, and M. Wilinska (2004): "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes." *Physiol. Meas.*, **25**, pp. 905–920.

IDF (2011): "The International Diabetes Federation." `www.idf.org`. Page retrieved May 2011.

Johansson, R. (1993): *System Modeling and Identification*. Prentice Hall, Englewood Cliffs, NJ.

Katayama, T. and G. Picci (1999): "Realization of stochastic systems with exogenous inputs and subspace identification methods." *Automatica*, **35**, pp. 1635–1652.

Keenan, B., J. Mastrototaro, G. Voskanyan, and S. G.M. (2009): "Delays in minimally invasive continuous glucose monitoring devices: a review of current technology." *Journal of Diabetes Science and Technology*, **3:5**, pp. 1207–1214.

Kildegaard, J., J. Randloev, J. U. Poulsen, and O. Hejlesen (2007): "The impact of non-model-related variability on blood glucose prediction." *Diabetes Technology and Therapeutics*, **9:4**, pp. 385–393.

Kiriakidis, K. and R. O. Brien (2006): "Optimal estimation of blood insulin from blood glucose." In *Proc. of ASME IMECE 2006*. Chicago, Illinois, USA.

Knobbe, E. J. and B. Buckingham (2005): "The extended Kalman filter for continuous glucose monitoring." *Diabetes Technology and Therapeutics*, **7:1**, pp. 15–27.

Kollman, C., D. Wilson, T. Wysocki, A.B.P.P., W. Tamborlane, and R. Beck (2005): "Limitations of statistical measures of error in assessing the accuracy of continuous glucose sensors." *Diabetes Technology and Therapeutics*, **7:5**, pp. 665–672.

Kovatchev, B., D. Cox, L. Gonder-Frederick, and W. Clarke (1997): "Symmetrization of the blood glucose measurement scale and its applications." *Diabetes Care*, **20:11**, pp. 1655–1658.

Kovatchev, B., L. Gonder-Frederik, D. Cox, and W. Clarke (2004): "Evaluating the accuracy of continuous glucose-monitoring sensors." *Diabetes Care*, **27:8**, pp. 1922–1928.

Kovatchev, B., C. King, M. Breton, S. Anderson, and W. Clarke (2006): "Clinical assessment and mathematical modeling of the accuracy of continuous glucose sensors (CGS)." In *Proc. 28th International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2008)*, pp. 71–74. New York City, USA.

Kulku, E., J. Tamada, G. Reach, R. Potts, and M. Lesho (2003): "Phisiological differences between interstitial glucose and blood glucose measured in human subjects." *Diabetes care*, **26:8**, pp. 2405–2409.

Larimore, W. (1990): "Canonical variate analysis in identification, filtering and adaptive control." In *Proc. 29th IEEE Conference on Decision and Control (CDC1990)*, pp. 596–604. Honolulu, Hawaii.

Leal, Y., W. Garcia-Gabin, J. Bondia, E. Esteve, W. Ricart, J. Fernandez-Real, and J. Vehi (2010): "Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models." *Journal of Diabetes Science and Technology*, **4:2**, pp. 391–403.

Lehmann, E. and T. Deutch (1992): "A physiological model of glucose-insulin interaction in type 1 diabetes mellitus." *Journal of Biomedical Engineering*, **14**, pp. 235–242.

Ljung, L. (1999): *System Identification: Theory for the User*, second edition. Prentice Hall, Upper Saddle River, NJ.

Lynch, S. and W. Bequette (2002): "Model predictive control of blood glucose in type i diabetics using subcutaneous glucose measurements." In *Proc. of the American Control Conference (ACC2002)*, pp. 2714–2719. Anchorage.

M. Lovera, T. G. and M. Verhaegen (2000): "Recursive subspace identification of linear and non-linear wiener state-space models." *Automatica*, **36**, pp. 1639–1650.

Maciejovski, J. (2002): *Predictive Control with Constraints*. Prentice Hall, London.

Magni, P., G. Sparacino, R. Bellazzi, and C. Cobelli (2006): "Reduced sampling schedule for the glucose minimal model: importance of bayes estimation." *American Journal of Physiology Endocrinology Metabolism*, **290**, January, pp. E177–E184.

Makroglou, A., J. Li, and Y. Kuang (2006): "Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview." *Applied Numerical Mathematics*, **56**, pp. 559–573.

Maran, A., C. Crepaldi, A. Tiengo, G. Grassi, E. Vitali, G. Pagano, S. Bistoni, G. Calabrese, F. Santeusanio, F. Leonetti, M. Ribaudo, U. Di Mario, G. Annuzzi, S. Genovese, G. Riccardi, M. Previti, D. Cucinotta, F. Giorgino, A. Bellomo, R. Giorgino, A. Poscia, and M. Varalli (2002): "Continuous subcutaneous glucose monitoring in diabetic patients." *Diabetes Care*, **25:2**, pp. 347–352.

MathWorks (2011): "MathWorks." `www.mathworks.com/products/matlab/`. Page retrieved May 2011.

Miller, M. and P. Strange (2007): "Use of fourier models for analysis and interpretation of continuous glucose monitoring glucose profiles." *Journal of Diabetes Science and Technology*, **1**, pp. 630–638.

Morbiducci, U., A. Tura, and M. Grigioni (2005): "Genetic algorithms for parameter estimation in mathematical modeling of glucose metabolism." *Computers in Biology and Medicine*, **35**, pp. 862–874.

Mosca, E., G. Zappa, and J. Lemos (1989): "Robustness of multipredictor adaptive regulator: MUSMAR." *Automatica*, **25:4**, pp. 521–529.

Nucci, G. and C. Cobelli (2000): "Models of subcutaneous insulin kinetics. a critical review." *Computer Methods and Programs in Biomedicine*, **62**, pp. 249–257.

Palerm, C. C., M. Rodriguez-Fernandez, W. C. Bevier, H. Zisser, J. R. Banga, L. Jovanovic, and F. J. Doyle (2006): "Robust parameter estimation in a model for glucose kinetics in type 1 diabetes subjects." In *Proc. 28th International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2006)*, pp. 319–322. New York City, USA.

Paoletti, S., A. Juloski, G. Ferrari-Trecate, and R. Vidal (2007): "Identification of hybrid systems: a tutorial." *European Journal of Control*, **13**, pp. 242–260.

Pappada, S., B. Cameron, and P. Rosman (2008): "Development of a neural network for prediction of glucose concentration in type 1 diabetes patients." *Journal of Diabetes Sciences and Technology*, **2:5**, pp. 792–801.

Percival, M. W., W. C. Bevier, H. Zisser, L. Jovanovic, D. E. Seborg, and F. J. Doyle (2008): "Prediction of dynamic glycemic trends using optimal state estimation." In *Proc. of 17th IFAC World Congress*, pp. 4222–4227. Seoul, South Korea.

Perez-Gandia, C., A. Facchinetti, G. Sparacino, C. Cobelli, E. Gomez, M. Rigla, A. de Leiva, and M. Hernando (2010): "Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring." *Diabetes Technology and Therapeutics*, **12:1**, pp. 81–88.

Pillonetto, G., G. Sparacino, and C. Cobelli (2003): "Numerical non-identifiability regions of the minimal model of glucose kinetics: superiority of bayesian estimation." *Mathematical Biosciences*, **184**, pp. 53–67.

Reifman, J., S. Rajaraman, A. Gribok, and W. Ward (2007): "Predictive monitoring for improved management of glucose levels." *Journal of Diabetes Sciences and Technology*, **1:4**, pp. 478–486.

Rivera-Montalvo, A., A. Stamps, and E. Gatzke (2008): "Application of dynamic optimization-based parameter estimation to a diabetes mellitus patient model." In *Proc. of American Control Conference (ACC2008)*, pp. 1382–1387. Seattle, Washington, USA.

Roy, A. and R. S. Parker (2006): "Dynamic modeling of free fatty acid, glucose, and insulin: an extended minimal model." *Diabetes Technology and Therapeutics*, **8:6**, pp. 617–626.

Söderström, T. and P. Stoica (1989): *System Identification*. Prentice Hall, Upper Saddle River, NJ.

Sorensen, J. (1985): *A Physiologic Model of Glucose Metabolism in Man and Its Use to Design and Assess Improved Insulin Therapies for Diabetes*. PhD thesis, Massachussetts Institute of Technology, Department of Chemical Engineering.

Sparacino, G., A. Facchinetti, A. Maran, and C. Cobelli (2008): "Continuous glucose monitoring time series and hypo/ hyper glycemia prevention: requiremetns, methods, open problems." *Current Diabetes Reviews*, **4:3**, pp. 181–192.

Sparacino, G., F. Zanderigo, A. Maran, A. Facchinetti, and C. Cobelli (2007): "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series." *IEEE Transactions on Biomedical Engineering*, **54:5**, pp. 931–937.

Spurr, G., A. Prentice, P. Murgatroyd, G. Goldberg, J. Reina, and N. Christman (1988): "Energy expenditure from minute-by-minute heart-rate recording: a comparison with indirect calorimetry." *The American Journal of Clinical Nutrition*, **48:3**, pp. 552–559.

Ståhl, F., M. Cescon, R. Johansson, and E. Renard (2009): "Infinite horizon prediction of postprandial breakfast plasma glucose excursion." In *Proc. 9th Annual Diabetes Technology Meeting (DTM2009)*, p. A163. San Francisco, CA.

Ståhl, F. and R. Johansson (2008): "Short-term diabetes blood glucose prediction based on blood glucose measurements." In *Proc. of 30th IEEE EMBS Eng. Medicine and Biology Society Conference (EMBC2010)*, pp. 291–294. Vancouver, British Columbia, Canada.

Ståhl, F. and R. Johansson (2009): "Diabetes mellitus modeling and short-term prediction based on blood glucose measurements." *Mathematical Biosciences*, **217**, pp. 101–117.

Steil, G., K. Rebrin, F. Hariri, S. Jinagonda, S. Tadros, C. Darwin, and M. Saad (2005): "Interstitial fluid glucose dynamics during insulin-induced hypoglycemia." *Diabetologia*, **48**, pp. 1833–1840.

The American Diabetes Association (2010): "Standards of medical care in diabetes 2010." *Diabetes Care*, **33:Supplement 1**, pp. S11–S61.

The Diabetes Control and Complications Trial Research Group (1993): "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus." *N Eng J Med*, **329:14**, pp. 977–986.

The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Study Research Group (2005): "Intensive diabetes treatment and cardiovascular disease in patients with rype 1 diabetes." *N Eng J Med*, **353:25**, pp. 2643–2653.

Van Overschee, P. and B. De Moor (1994): "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems." *Automatica*, **30**, pp. 75–93.

Van Overschee, P. and B. De Moor (1996): *Subspace Identification for Linear Systems : Theory-Implementation-Application*. Kluwer Academic Publishers, Boston.

Verdult, V. and M. Verhaegen (2002): "Subspace identification of multivariable linear parameter-varying systems." *Automatica*, **38**, p. 805–814.

Verhaegen, M. (1994): "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data." *Automatica*, **30:1**, pp. 61–74.

Wayne Bequette, B. (2010): "Continuous glucose monitoring: Real-time algorithms for calibration, filtering and alarms." *Journal of Diabetes Science and Technology*, **4:2**, pp. 404–418.

Weiss, A. (1991): "Multi-step estimation and forecasting in dynamic models." *Journal of Econometrics*, **48**, pp. 135 – 149.

Wilinska, M. E., L. J. Chassin, H. C. Schaller, L. Schaupp, T. R. Pieber, and R. Hovorka (2005): "Insulin kinetics in type-1 diabetes: Continuous and bolus delivery of rapid acting insulin." *IEEE Transactions on Biomedical Engineering*, **52:1**, pp. 3–12.

Williams, G. and J. Pickup (1992): *Handbook of Diabetes second edition*. Blackwell Science, Oxford.

Department of Automatic Control
Lund University
Box 118
SE-221 00 Lund Sweden

Document name
LICENTIATE THESIS

Date of issue
June 2011

Document Number
ISRN LUTFD2/TFRT--3250--SE

Author(s)
Marzia Cescon

Supervisor
Rolf Johansson, Per Hagander

Sponsoring organisation
DIAdvisor (EU FP7)

Title and subtitle
Linear Modeling and Prediction in Diabetes Physiology

Abstract

Diabetes Mellitus is a chronic disease characterized by the inability of the organism to autonomously regulate the blood glucose level due to insulin deficiency or resistance, leading to serious health damages. The therapy is essentially based on insulin injections and depends strongly on patient daily decisions, being mainly based upon empirical experience and rules of thumb. The development of a prediction engine capable of personalized on-the-spot decision making concerning the most adequate choice of insulin delivery, meal intake and exercise would therefore be a valuable initiative towards an improved management of the desease.

This thesis presents work on data-driven glucose metabolism modeling and short-term, that is, up to 120 minutes, blood-glucose prediction in Type 1 Diabetes Mellitus (T1DM) subjects.

In order to address model-based control for blood glucose regulation, low-order, individualized, data-driven, stable, physiological relevant models were identified from a population of 9 T1DM patients data. Model structures include: autoregressive moving average with exogenous inputs (ARMAX) models and state-space models.

ARMAX multi-step-ahead predictors were estimated by means of least-squares estimation; next regularization of the autoregressive coefficients was introduced. ARMAX-based predictors and zero-order hold were computed to allow comparison.

Finally, preliminary results on subspace-based multi-step-ahead multivariate predictors is presented.

Key words
system identification, linear prediction, biological systems

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages
214

Recipient's notes

Security classification