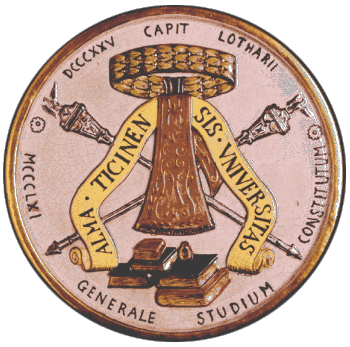
University of Pavia

Ph.D. School in Electronics, Computer Science and Electrical Engineering



Cycle XXXII

Personalized Artificial Pancreas: from identification to optimal bolus computation

Author:
Eleonora Maria Aiello

Supervisor:
Prof. Chiara Toffanin

2016-2019

Abstract

Type 1 diabetes mellitus is the more severe form of diabetes mellitus. It results from a cellular-mediated autoimmune destruction of the beta-cells of the pancreas, which are responsible for the secretion of insulin that is required for a proper blood glucose regulation. Patients with type 1 diabetes need exogenous insulin injections to keep the glucose concentration in the safe range. Artificial pancreas is an autonomous system for closed-loop blood glucose control in subjects affected by type 1 diabetes. The core of the system is the control algorithm, which receives blood glucose data from the sensor, computes the required insulin amount and transmits this information to the insulin pump. Regarding the control algorithm, one of the most promising approach revealed to be the model predictive control algorithm, which exploits a glucose-insulin model of the patient to predict near-future blood glucose values and, consequently, computes the optimal insulin dose. The performance of the control algorithm is highly influenced by the quality of the model used for prediction. Moreover, the inter-patient variability characterizing subjects with T1D increases the need of patient-tailored models. Since promising results have been obtained *in silico* using the UVA/Padova simulator, the aim of this thesis is to investigate and test the applicability of the identification techniques on free-living data collected without ad hoc clinical protocols thanks to the availability of long term trials. The individualized models show superior prediction performance with respect to the average model that was used to synthetize the controller used during the trial. This result pushed towards a detailed data analysis to improve model identification. A multiple-model predictor with different identified models on the basis of the data analysis is proposed in this thesis. The prediction capabilities are improved if compared to the performance of a predictor built using a single model identified on a daily subset. These results represent a milestone in the development for a new generation of individualized controllers for artificial pancreas. The patient-tailored models can be exploited to predict this risk of hypoglycemia in advance and therefore to alert the patient on the risk. In order to improve the safety of the artificial pancreas system, the identified models have been evaluated in terms of hypoglycaemia prevention by showing that these models are able to detect 84.53% of the hypoglycaemia events occurred during a 1-month trial on average. In this thesis, model identification has been addressed by deep learning techniques, by showing that the proposed architecture obtains state-of-the-art performance on both *in silico* and *in vivo* data, considering several prediction horizons. Finally, since the post-prandial glucose regulation remains a challenging issue for diabetes treatment, machine learning methodologies have been applied in order to improve the postprandial glucose regulation. Two algorithms are proposed to provide corrections to time and/or amount of the meal bolus. They have been tested on the *in silico* virtual population of the UVA/Padova simulator by showing the reduction of both hypoglycemia and hyperglycemia.

Contents

List of Figures

List of Tables

Abbreviations

A	**A**verage
ACC	**Acc**uracy
ANOVA	**AN**alysis **O**f **VA**riance
AP	**A**rtificial **P**ancreas
AR	**A**uto**r**egressive
AUROC	**A**rea **U**nder the curve of **ROC**
Av	**Av**erage *in silico* patient
AVG	**AV**era**G**e *in silico* model
B	**B**reakfast
BG	**B**lood **G**lucose
BM	**B**asic **M**odel
BMB	**B**asic **M**odel **B**reakfast
BMD	**B**asic **M**odel **D**inner
BMI	**B**ody **M**ass **I**ndex
BML	**B**asic **M**odel **L**unch
CA	**C**lassifier-**A**ugmented
CF	**C**orrection **F**actor
CGM	**C**ontinuous **G**lucose **M**onitor
CHO	**C**arbo**h**ydrates
CL	**C**losed-**L**oop
COD	**C**oefficient **O**f **D**etermination
CR	**C**arbo **R**atio
CSII	**C**ontinuous **S**ubcutaneous **I**nsulin **I**nfusion
CT	**C**onventional **T**herapy
D	**D**inner

DGF	**D**eep **G**lucose **F**orecasting
DKA	**D**iabetic **K**eto**A**cidosis
DM	**D**aily **M**odel
DMP	**D**aily **M**odel **P**redictor
DP	**D**ay **P**eriod
DW	**D**etection **W**indow
FC	**F**ully **C**onnected
FDA	**F**ood and **D**rug **A**dministration
FHOCP	**F**inite **H**orizon **O**ptimal **C**ontrol **P**roblem
FN	**F**alse **N**egative
FNR	**F**alse **N**egative **R**ate
FOR	**F**alse **O**mission **R**ate
FP	**F**alse **P**ositive
FPG	**F**asting **P**lasma **G**lucose
FPR	**F**alse **P**ositive **R**ate
FT	**F**ine-**T**uning
HbA1c	**H**emoglo**b**in **A1c**
HE	**H**ypoglycemia **E**vent
HPSA	**H**ypoglycemia **P**rediction **S**tate of **A**rt
IG	**I**nterstitial **G**lucose
IIG	**I**nitial **I**nterstitial **G**lucose
IIGS	**I**nitial **I**nterstitial **G**lucose **S**lope
IFG	**I**mpaired **F**asting **G**lucose
IGT	**I**mpaired **G**lucose **T**olerance
IHPA	**I**ndividualised **H**ypoglycemia **P**redictive **A**lert
IR	**I**mpulse **R**esponse
IOB	**I**nsulin **O**n **B**oard
KNN	**K**-**N**earest **N**eighbors
L	**L**unch
LSTM	**L**ong-**S**hort **T**erm **M**emory
LMPC	**L**inear **M**odel **P**redictive **C**ontrol
MCM	**M**eal **C**ontrol **M**odule
MCS	**M**ultiple **C**lassifier **S**cheme

MLR	**M**ultiple **L**inear **R**egression
MMP	**M**ulti-**M**odel **P**redictor
MPC	**M**odel **P**redictive **C**ontrol
MSE	**M**ean **S**quare **E**rror
MSG	**M**ean **S**quare of **G**roup
N	**N**ight
NEHE	**N**ot **E**valuable **H**ypoglycemia **E**vent
NPV	**N**egative **P**redictive **V**alue
O	**O**verall
OGT	**O**ral **G**lucose **T**olerance
OL	**O**pen-**L**oop
PH	**P**rediction **H**orizon
PP	**P**ost**p**randial
PPV	**P**ositive **P**redictive **V**alue
PW	**P**rediction **W**indow
RCM	**R**ange **C**ontrol **M**odule
RMSE	**R**oot **M**ean **S**quared **E**rror
ROC	**R**eceiver **O**perator **C**haracteristic
SD	**S**tandard **D**eviation
SMBG	**S**elf **M**onitoring **B**lood **G**lucose
SSE	**S**um of **S**quare **E**rror
SSG	**S**um of **S**quare of **G**roup
SST	**S**um of **S**quare **T**otal
T1D	**T**ype **1** **D**iabetes
T2D	**T**ype **2** **D**iabetes
$\mathbf{T_a}$	**T**ime **a**bove target
$\mathbf{T_{4_{hyper}}}$	**T**ime **a**bove target in severe **hyper**glycemia
$\mathbf{T_b}$	**T**ime **b**elow target
TDI	**T**otal **D**aily **I**nsulin
$\mathbf{T_{b_{hypo}}}$	**T**ime **b**elow target in severe **hypo**glycemia
TN	**T**rue **N**egative
TNR	**T**rue **N**egative **R**ate
TP	**T**rue **P**ositive

TPR	**T**rue **P**ositive **R**ate
$\mathbf{T_r}$	**T**ime in **r**ange
$\mathbf{T_{tr}}$	**T**ime in **t**ight **r**ange

Chapter 1

Introduction

Diabetes mellitus is a metabolic disease characterized by high blood glucose concentration, known as hyperglycemia. The National Institute of Health (NIH) affirms that the prevalence of diagnosed diabetes increased by 382% from 1988 to 2014. It has been measured in 2015 that in America [1] there were 30.3 million affected people that corresponds to 9.4% of the entire population, where 1.25 million of children and adults have the more severe form of diabetes mellitus. Its incidence has been increasing about 1.5 million per year, so much so, 1 in 3 American adults will have diabetes in 2050 if current trend continues. On the other hand, in the European Region [2], 60 million people with diabetes have been calculated, or about 10.3% of men and 9.6% of women aged 25 years and over. Therefore, most people are unable to identify the main signs of the disease, including weight loss, slow healing of cuts and bruises, blurred vision, being thirsty, according to recent researches. In America, only 23.1 million of 30.3 were diagnosed, but 7.2 million were undiagnosed, while in Europe around a quarter of these cases are undiagnosed. The risk factors include diet, physical inactivity, overweight and obesity, which have been estimated to account for about 65–80% of new cases. Specifically, the risk depends strongly on the age and the duration of obesity, and weight gain during adult life. If this diseas is not managed correctly, diabetes can damage the heart, blood vessels, kidneys, eyes and nerves. Half of patients with diabetes die of cardiovascular disease, and 10-20% of people die of kidney failure. Diabetes still remains the 7th leading cause of death in the United States in 2015 according to the statistics about diabetes provided by the American Diabetes Association (ADA) in 2018. Worldwide, diabetes leads to death about 3.4 million people annually. Almost 80% of these deaths occur in low- and middle-income countries, and almost half are people aged under 70 years. The numbers associated with diabetes make a strong case for devoting more resources to finding a suitable solution.

The most common form of diabetes mellitus is the Type 2 Diabetes (T2D), which is less severe

than Type 1 Diabetes (T1D) mellitus. Before developing T2D, people almost always experience the so-called prediabetes phase, which is characterized by blood glucose levels higher than normal but not yet high enough to be diagnosed as diabetes. T1D is the more severe form of diabetes mellitus. T1D results from a cellular-mediated autoimmune destruction of the beta-cells of the pancreas, which have the task of secreting the insulin that is required for a proper blood glucose regulation. Although T2D is less severe than T1D, T2D may evolve into T1D. Since hyperglycemia can cause long-term complications including damage to blood vessels, eyes, kidneys, and nerves, treatment of diabetes involves lowering blood glucose to minimize diabetes complications. T1D patients need exogenous insulin injections to keep the glucose concentration in the euglycemic range taking care to avoid hypoglycemia, a condition that could be caused by excessive insulin administration.

T1D accounts for only 5–10% of the diabetes affected population. T1D is also known as *insulin-dependent* or *juvenile* diabetes because it is usually diagnosed in children and young adults. Between 2001 and 2009, an increase of 21% in the prevalence of T1D has been observed in people under age 20 [3], and T1D is associated with an estimated loss of life-expectancy of up to 13 years. The estimated annual number of newly diagnosed cases in the United States counted 17,900 children and adolescents younger than age 20 years. The T1D diffusion explains why it has been investigated in research for the treatment of type 1 diabetes.

Artificial Pancreas (AP) is an autonomous system for maintaining the glucose concentration within a safe range. AP helps T1D subjects to reduce T1D related risks. This system monitors glucose levels through a subcutaneous glucose sensor or Continuous Glucose Monitor (CGM), and automatically provides the proper amount of insulin, which is infused through a subcutaneous insulin pump. The core of the system is the control algorithm, which receives glucose data from the sensor together with optional information provided by the patient, computes the required insulin amount and transmits this information to the insulin pump. The research is moving towards a more wearable and portable solution for the AP, integrating the control strategy on a reconfigured smartphone or directly on the insulin pump [4]. For example, every sample time, the advanced algorithm running on the smartphone analyzes the data and, if needed, adjusts the insulin doses. The development of the control algorithm has to face problems such as the noisy measurements of CGM, the physiological delays due to diffusion process from plasma to subcutaneous tissues, and the inter- and intra-variability among T1D individuals. A good regulation of blood glucose concentration, i.e. glycemia, is the main aim of the technological improvement of AP. An improved glycemia regulation can avoid conditons such as hyperglycemia and hypoglycemia, which may be very dangerous for the patients. Indeed, hyperglycemia can cause long-term complications for the subject's health, whereas hypoglycemia can lead to immediate and severe complications.

The Juvenile Diabetes Research Foundation (JDRF) has established in 2006 an international consortium for the development of an AP, where several research groups from University of Virginia, University of California in Santa Barbara, Montpellier University Hospital in France, and the Universities of Padova and Pavia in Italy, have been involved. Starting from 2008, several clinical trials have been performed in a hospital setting for a subject, then the goal was to move the AP to free-living conditions for long periods at home. In 2010, the AP@home European project was founded involving several universities and medical centres around Europe and the University of Pavia was kept as reference point for the control algorithm research. In 2014, an AP system has been used in the first randomized crossover outpatient clinical trial for eight weeks where the loop was closed during evening and night periods. Finally, in 2015 the closed-loop was continually used 24 hours per day for one month in real-life conditions. As for both *in silico* tests, *in vivo* experiments have shown the need for an individualization of both the control algorithms and the safety systems. Simultaneously, the availability of long term trials allows the collection of datasets rich of information potentially useful for the individualization purpose. The aim of this thesis is to improve the current control strategy by identifying individual models from experimental data. Moreover, the other focus of this thesis is the improvement of post-prandial glucose regulation, which remains a challenging issue for diabetes treatment. In fact, an incorrect administration of a meal bolus can result in dangerous complications for the patient.

1.1 Author's Contributions

Both *in silico* and *in vivo* experiments have shown the need for an individualization of the control algorithm, which is a Model Predictive Control (MPC) algorithm, this aim can be achieved by identifying a specific model to be included in the MPC.
An already existing identification approach based on the impulse response [5] has been extended in this thesis to the closed-loop in free-living conditions thanks to the availability of experimental data. This set-up is particularly challenging because the identification of reliable models on real-data is more difficult than on simulated data. Firstly, a single patient has been considered and the improved prediction capabilities of the patient-tailored model are presented with respect to the average model, computed by averaging the model parameters of the virtual adult population of the UVA/Padova simulator and currently used in the MPC controller during the recent trials. Given the good results obtained on a single patient, the technique has been extended to the 7 patients studied at Amsterdam clinical centre. The aim is to show the ability of the impulse-response technique to adapt to different subjects, showing its effectiveness in front

of inter-subject variability.

The patient-tailored models of the 7 patients are used to create a system to detect in advance hypoglycemia phenomena thanks to the good prediction capabilities of individualized models. The performance of the proposed system is compared to the one achieved by the algorithm for hypoglycemia prevention used in several clinical trials. Specifically, the individualized hypoglycemia predictive alarms are evaluated on 7 T1D patients with the aim of predicting in advance the unavoidable hypoglycemia events. The performance metrics of the alarm system are evaluated on the entire trial. This methodology has al- lowed to detect on average about 85% of the hypoglycemia events occurred during the trial in time to allow a rescue action with a negligible number of false alarms. Although promising results have been obtained in the identification of patient-tailored mod- els from clinical data, a in-depth analysis of the data has been required to extract hidden information from the dataset useful to improve model identification. An ANOVA analysis has been performed on the real-life data, and it has shown that there is a significant dependence between different day periods and the glucose profile. The results of the ANOVA tests have been exploited to build a multi-model able to captures different postprandial glucose dynamics along the day using different models in each sub-period.

In this thesis, a deep learning architecture to be employed in a MPC-based system has been developed: its main purpose is the prediction of the blood glucose associated to a list of possible future insulin actions in order to decide the optimal therapy for the patient. The solution entails multiple models trained on the *in silico* adult patients of the UVA/Padova simulator. Each model is used to predict a glucose profile for a fixed prediction horizon and the individual predictions are then aggregated to obtain a profile of future glucose levels. In order to improve the predictor performance, the model trained on *in silico* data is fine-tuned on data of the specific patient improving the predictive performance.

Considered that the postprandial glucose control is typically one of the most problematic aspects of glucose regulation, this thesis is also concerned with improving the conventional therapy. Since machine learning represents a new methodology recently explored in this type of applications, two new approach based on machine-learning methodologies have been proposed in order to improve meal compensation during postprandial periods. The first methods exploits the K-Nearest Neighbors classification algorithm to predict postprandial glucose profile due to the nominal therapy and to suggest a correction to time and/or amount of the meal bolus. The classifier is used to forecast the glucose response to carbohydrate intake. Then, the predictions are exploited to correct the nominal bolus computed via conventional therapy, minimizing the occurrence of hypo- and hyperglycemia. A unique classifier valid for the entire adult virtual population of the UVA/Padova simulator has been identified. Since an average model could ideally limit the performance, a personalized algorithm is developed. It is an individualized

approach able to correct the meal bolus computed with the conventional therapy in order to handle the inter-subject variability characterizing patients that may affect postprandial glucose regulation. In this case the postprandial glucose regulation has been managed in a decisional framework, where the decision variable is the correction of the insulin bolus by exploiting a multiple linear regression model able to describe the relation between glucose concentration and injected insulin. Both these approaches are compared with conventional therapy glucose profile.

This thesis is organized as follows:

- Chapter 2: a brief introduction on diabetes mellitus and an overview on diabetes diagnosis, management, recently developed technological devices, and the description of the conventional therapy are given
- Chapter 3: the 2018 version of the UVA/Padova simulator with the multi-compartmental model used in this thesis is presented
- Chapter 4: the presentation of two new approaches based on machine-learning methodologies to improve postprandial glucose regulation in subject treated via conventional therapy are shown
- Chapter 5: a complete description of artificial pancreas system and the current control strategy is given
- Chapter 6: the description of the impulse-response technique identification technique to be used on free-living data is presented and the results obtained on 7 T1D patients are shown.
- Chapter 7: the application of the *in vivo* identification technique for the development of individualized hypoglycemia safety alerts is proposed and validated on a population cohort
- Chapter 8: a detailed real-life data analysis is proposed and a data-driven identification methodology based on real-data analysis is presented
- Chapter 9: the description of a deep learning architecture able to forecast the blood glucose level of T1D patients is proposed
- Chapter 10: the conclusion is drawn and possible future developments are proposed.

Chapter 2

Diabetes Mellitus

Diabetes mellitus is a chronic metabolic disease associated with high Blood Glucose (BG) concentration (BG $>$ 180 mg/dl), known as hyperglycemia. Glucose is a major source of energy for the human body and it is essential that BG levels are maintained within a safe range $[70 - 180]$ mg/dl. In a healthy system, in order to regulate the supply of glucose to the cells, the body has to maintain the BG concentration, also called glycemia, within the safe range. Maintaining internal conditions in the body is called homeostasis, which typically occurs through the use of feedback loops that control the body's internal conditions. The control of BG is a negative feedback that allows the human body to self-stabilize and the hormones secreted from pancreatic islet cells play central roles in the whole-body glucose homeostasis. Specifically, a small portion (about 5%) of the pancreas consists of endocrine cells, which are clustered in groups (alpha, beta, gamma, delta, epsilon) known as pancreatic islets or, more specifically, islets of Langerhans. However, only the alpha and beta cells are involved in the BG regulation through hormones production.

Since glucose comes mainly from foods containing carbohydrates, a meal intake produces a glucose rise. Increased BG levels stimulate beta cells in the pancreas to secrete a hormone called insulin, which stimulates the absorption of the glucose from blood into the cells and the formation of glycogen in order to reduce BG level. Once glucose levels fall below a threshold (BG $<$ 70 mg/dl), known as hypoglycemia, there is no longer a stimulus for insulin release, and the beta cells stop releasing insulin. Hypoglycemia suppresses insulin secretion from beta cells and stimulates glucagon secretion from alpha cells of the pancreas, normalizing BG levels. Glucagon works in the opposite way of insulin. It acts by stimulating liver and muscles cells to break down stored glycogen into glucose and by allowing the formation of new glucose molecules to back up to a normal glucose concentration, known as glycogenolysis and gluconeogenesis, respectively.

Since the body cells uses glucose as fuel for energy, glucose absorption is fundamental to make

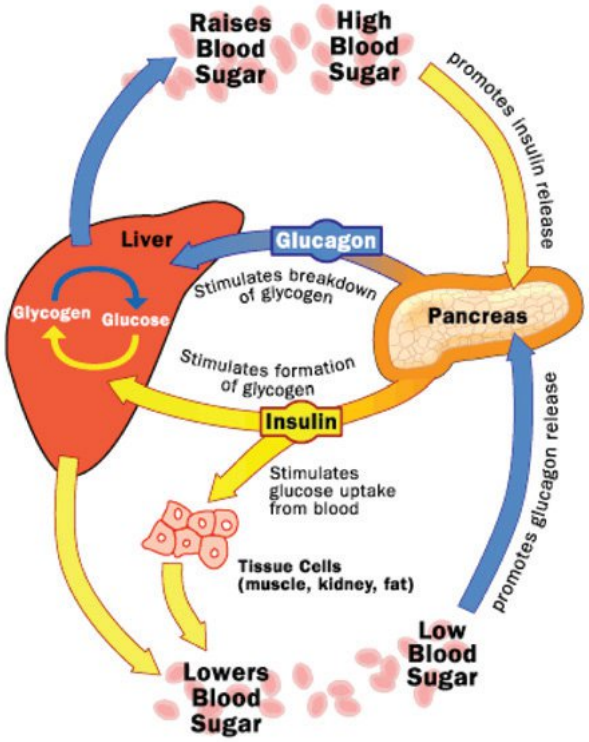


FIGURE 2.1: Scheme of glucose metabolism [6].

the organs to operate normally. In particular, the brain cells entirely depends on glucose, so maintaining a constant BG source is crucial to ensure the normal brain functions. In healthy subjects, insulin and glucagon are continually produced to maintain the glucose concentration within safe levels. This process is called glucose metabolism and it is shown in Figure 2.1. A failure of the negative feedback mechanism can result in high BG levels, which have a variety of negative health effects including damages to blood vessels, eyes, kidneys, and nerves.

If the pancreas cannot secretes insulin, consequently body cells cannot absorb glucose from the bloodstream. Moreover, the body interprets this situation as lack of glucose and stimulates the pancreatic alpha cells to secrete glucagon to release glucose by performing glycogenesis. This action leads to an increase of BG concentration, which becomes even higher in case of meal intake. High BG levels causes polyuria, that is an increased quantity of urination. The higher the BG, the more the amount of glucose filtered by kidneys is; so, if the filtered glucose exceeds the quantity that kidneys can reabsorb, glucose remains in the small tubes of kidneys by causing osmotic diuresis that is an increase of the urination rate. High BG and increased urine flow stimulate the thirst receptors of the brain by causing constant thirst, called polydipsia. Since the cells cannot receive glucose due to the absence of insulin, the brain receptors interpret this fact as a lack of glucose and, since glucose comes mostly from the ingested carbohydrates, the lack of glucose produce constant hungry, also called polyphagia. Polyuria, polydipsia and

polyphagia are diabetes symptoms, regardless of the type of diabetes.
Since the lack of insulin implies that glucose cannot be absorbed from the bloodstream into the body cells to be used as energy, the body starts breaking down fats for energy, which produces ketones. When ketones build up in the blood, they make it more acidic by causing a serious condition known as Diabetic KetoAcidosis (DKA). The most common symptoms associated with DKA are the consistent increase in polydipsia and polyuria, but also generalized weakness, nausea, vomiting, rapid weight loss, and altered consciousness.

2.1 Types of diabetes mellitus

There are three main types of diabetes:

- type 1 diabetes mellitus
- type 2 diabetes mellitus
- gestational diabetes.

2.1.1 Type 1 diabetes mellitus

T1D is probably caused by the destruction of the pancreas beta cells due to an autoimmune reaction. Since beta cells are responsible for insulin secretion, this type of diabetes is characterized by an absolute insulin deficiency and this the reason why it is also known as insulin-dependent diabetes. T1D patients need exogenous insulin injections to keep the glucose concentration in the euglycemic range (80-140 mg/dl). The beta-cell destruction is an autoimmune process and it can be caused by exposure to certain viruses, genetics or more likely a combination of environmental factors. When the beta cells die, the pancreas stops generating insulin, so T1D patients are characterized by an absolute insulin deficiency. If the body no longer produces enough insulin, the regulation of the BG levels is compromised and this can lead not only to polyuria, polydipsia and polyphagia, but also to unexplained weight loss, tiredness or fatigue, changes in vision, numbness or tingling in hand and feet, slow-healing wounds or sores, and abnormally high frequency of infection. T1D must be managed with a insulin therapy, which defines the daily exogenous insulin injections, individualized on the specific patient. If the injected insulin is not enough according to the body needs, high levels of ketones in the blood can result by leading to the DKA development, which is a common problem in patients with insulin-dependent diabetes. On the other hand, if there is an overestimation of needed insulin, the patient can

experience hypoglycemia that may lead to severe short-term complications. Thus, in order to improve the insulin management and reduce risks related with insulin dependence, the insulin therapy needs continuous adjustments. T1D is also known as juvenile diabetes because it usually appears during childhood or adolescence.

2.1.2 Type 2 diabetes mellitus

T2D occurs when the insulin resistance of the body becomes unexpectedly high, and the insulin produced by the pancreas is no longer effective on maintaining a stable glycemia. Unlike T1D, T2D is characterized by uncommon high insulin resistance, which is defined clinically as the inability of insulin to increase glucose absorption and utilization [7]. Thus, people with insulin resistance have cells that have trouble absorbing glucose, which causes a buildup of sugar in the blood. Consequently, oral medications devoted to the insulin resistance reduction are usually needed in subjects affected by T2D. Without enough insulin, excess glucose builds up in the bloodstream by causing high BG levels. T2D mellitus is also called non-insulin-dependent diabetes and it occurs in 90% to 95% of diabetics and usually occurs in adults over the age of 40, most often between the ages of 50 and 60 [8]. This is the reason why it is also known as adult-onset diabetes. The insulin resistance has several causes such as genetics, excess weight and physical inactivity, and it is a condition, that can precedes the development of T2D.
If BG levels are higher than normal, but not high enough to be considered T2D, the so-called prediabetes is diagnosed. The body reacts to insulin resistance by increasing the insulin production. Over time, the beta cells in the pancreas fail to keep up with the increased need for insulin and T2D becomes type 1. The degeneration from T2D into type 1 is highly frequent, making the treatment of T1D even more interesting to be investigated.

2.1.3 Gestational diabetes

Gestational diabetes can occur during pregnancy and every year 2% to 10% of pregnancies are affected by this disease. It is similar to T2D because it is characterized by insulin resistance, but it usually disappears after the birth. However, there are cases in which gestational diabetes can degenerate to type 2. Gestational diabetes is dangerous not only for the mother, but above all for the baby. In case of gestational diabetes, the mother's pancreas works overtime to produce insulin, but the BG concentration remains high. Although insulin does not cross the placenta, glucose and other nutrients do. So extra BG goes through the placenta, giving the baby high BG levels. This causes the baby's pancreas to make extra insulin to remove BG from bloodstream.

Since the baby is getting more energy than it needs to grow and develop, the extra energy is stored as fat. Babies with excess of insulin become children who are at risk for obesity and adults who are at risk for T2D.

2.1.4 The long-term complications

A number of serious health problems can affect people with diabetes if BG concentration is not kept under control. As a matter of fact high BG levels can lead to serious diseases affecting the heart and blood vessels, eyes, kidneys, nerves, and teeth. Moreover, people with diabetes also have a higher risk of developing infections.
The most common complications are related with cardiovascular diseases, which is the most frequent cause of death in people with diabetes. They have chances of having a stroke or heart attack 1.5 times higher than in people who don't have diabetes [8]. A frequent complication is a kidney disease called nephropathy: indeed high levels of blood sugar make the kidneys filter too much blood, overburdening the filters. After many years, the kidneys could start to not operate correctly, and useful proteins could be lost in the urine. In half of subjects affected by diabetes, diabetic neuropathy, which is a peripheral nerves malfunction, may also occur as a complication [8]. In this case the symptoms are tingling, pain, numbness, or weakness in the feet and hands. The degeneration of neuropathy can lead to foot ulcers and eventual limb amputation. Finally, long-term accumulated damage to the small blood vessels in the eye leads to diabetic retinopathy, an important cause of blindness.

2.2 Diagnosis of Diabetes

As stated in the report of the ADA [8], the diagnosis of diabetes is generally stated through the occurrence of chronic hyperglycemia, which is associated with serious long-term complications. The symptoms described in the previous section can be prevented by keeping blood sugar in the target range, also known as euglycemic range, which spans from 80 to 140 mg/dl. There are three main methods to diagnose diabetes:

- Fasting Plasma Glucose (FPG) test
- Oral Glucose Tolerance (OGT) test
- A1C test.

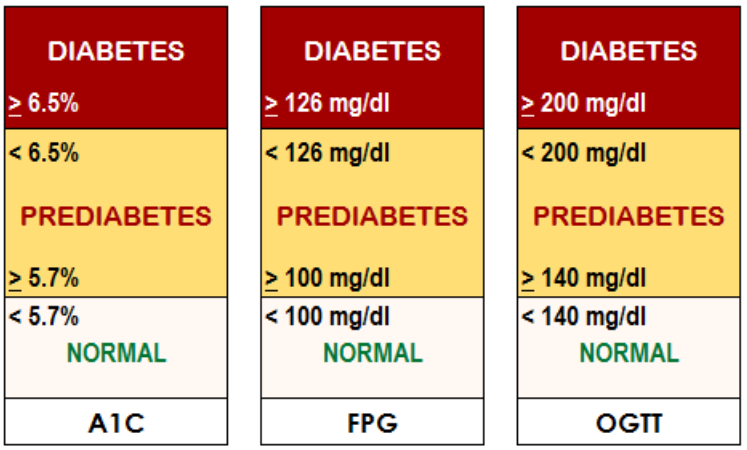


FIGURE 2.2: Test to diagnose diabetes mellitus.

The FPG test is a simple blood test taken after 8-10 hours of fasting and it diagnoses diabetes if fasting BG is greater than or equal to 126 mg/dl. In fact, due to the lack of insulin, the body is unable of absorbing the assumed carbohydrates and sugars into the skeletal muscles and fat tissues, letting the blood sugar level to increase. The OGT test checks patient's BG levels before and 2 hours after drink a special sweet drink. Diabetes is diagnosed if the two-hour BG is greater than or equal to 200 mg/dl. The A1C test, also called Hemoglobin A1c (HbA1c) measures person's average levels of BG over the past 3 months. This test is based on the attachment of glucose to hemoglobin, the protein in red blood cells that carries oxygen. In the body, red blood cells are constantly forming and dying, but typically they live for about 3 months. Diabetes is diagnosed if the average BG are higher than 140 mg/dl which is equivalent to 6.5% of HbA1c. Some of the advantages of the A1C test with respect to the other mentioned tests are that it captures better chronic hyperglycemia, fasting is not needed for A1C assessment, and no acute perturbations (e.g., stress, diet, exercise) affect A1C [9]. The target ranges adopted for the mentioned tests are shown in Figure 2.2. The National Glycohemoglobin Standardization Program certifies that manufacturers of A1C tests provide tests that are consistent with those used in a major diabetes study.

A1C is used as a key indicator of long-term BG control and the 2015 ADA Standards of Medical Care in Diabetes proposes the following optimal targets, but each target must be individualized to the needs of each patient and their disease factors [10]:

- 6.5 percent or less for people who never experiences hypoglycemia episodes
- 7 percent for many adults with diabetes
- 7.5 percent for all children with diabetes

- 8 percent or less for people with a history of severe hypoglycemia.

Thus, people will have different A1C targets according to their diabetes history.
The A1C test helps physician adjusting treatments to reduce the risk of long-term diabetes complications. Moreover, the observance of the recommended glycemic targets has been proved that reduces the occurrence of diabetes complications [11].
Prediabetes condition is defined as Impaired Glucose Tolerance (IGT) or Impaired Fasting Glucose (IFG), depending on which test it was detected. IFG is a type of prediabetes, whose feature is the high blood sugar level during fasting; it is defined by FPG concentration from 110 mg/dl to 125 mg/dl as specified by the World Health Organization criteria.[12]. Whereas, IGT is a pre-diabetic state of hyperglycemia related with insulin resistence and is defined by an elevated 2-h plasma glucose concentration from 140 mg/dl to 199 mg/dl after a 75-g glucose load on the OGT test in the presence of an FPG concentration < 126 mg/dl [12].

2.3 Insulin Therapy

The goal of diabetes treatments is to keep the BG level within euglycemic range and prevent diabetes complications. In order to regulate BG concentration, patients with T1D need exogenous insulin injections since the endogenous insulin production is deficient due to dysfunction of the beta cells [13, 14]. Since several factors can influence the choice of an insulin therpay, such as the type of diabetes, the BG levels, how much BG fluctuates throughout the day and the lifestyle, it is suited to the patient by the physician. In order to define an appropriate insulin treatment, the physician has to take into account several factors related with the patient's daily habits and physiological characteristics, such as the glycemic response of the individual to food intake. This insulin treatment is called Conventional Therapy (CT). It consists of the combination of the basal insulin, which is a piecewise constant amount of insulin in charge to maintain stable the BG levels during fasting periods, and the insulin bolus, which is an impulse-like amount specifically administered at meal times to compensate the glucose rise due to a meal [15–17]. Defining an insulin therapy means to choose the type and amount of insulin and the insulin delivery options. Insulin can be categorized in four main types that are defined on the basis of the onset time, the peak time, and duration. The first is the time needed to reach the bloodstream after the injection or infusion, the peak time is the time needed to reach its maximum effect, whereas duration represents the total time in which the insulin continues to have an effect. The four categories are:

- Rapid-acting insulin

Type	Onset	Peaktime	Duration
Rapid-acting insulin	15 minutes	30-90 minutes	3-5 hours
Short-acting insulin	30-60 minutes	2-4 hour	5-8 hours
Intermidiate-acting insulin	1-4 hours	4-12 hours	12-16 hours
Long-acting insulin	1-4 hours	no peak	up to 24 hours

TABLE 2.1: Types of insulin.

- Short-acting insulin
- Intermidiate-acting insulin
- Long-acting insulin.

Rapid-acting insulin usually reaches the bloodstream in about 15 minutes, reaches its peak between 30 to 90 minutes after the administration, and can last 3 to 5 hours. Rapid-acting insulin is taken just before or after meals, to compensate the rise of BG. Short-acting insulin can reach the bloodstream in about 30 minutes to one hour, and peaks after two to four hours. Its effects tend to last about five to eight hours. This type of insulin is also known as regular-acting insulin. Intermediate-acting insulin reaches the bloodstream in about one to four hours, and has a peak time between four and 12 hours, depending on the brand, and controls BG levels for about 12 hours or longer. Long-acting insulin has an onset within one to four hours, tends to lower glucose levels fairly evenly over a 24-hour period and it has minimal peak. Long-acting insulin covers insulin needs for about one full day and it is often combined, when needed, with rapid- or short-acting insulin to compensate meals, which highly affected BG concentration within few minutes. Table 2.1 summarizes the parameter values for every type of insulin. Diabetes treatment has to deal with the insulin daily intake. Approximately 40-50% of the total daily exogenous injected insulin has to replace insulin overnight, during fasting and between meals. This is called basal insulin replacement, whose amount is usually constant from day to day. The other 50-60% is for carbohydrate coverage during mealtime and high blood sugar correction. This is called the bolus insulin replacement. At mealtime, in order to compute the amount of insulin to be injected, the patient has to estimate the carbohydrates contained in the meal. The computation of the amount of insulin bolus involves some clinical parameter identified by the physician. If the insulin bolus results to be underestimated, the patient can inject an additional insulin bolus, called correction bolus. Rapid-acting insulin is used for insulin boluses because of the need of a fast meal compensation, whereas long-acting insulin is used to replace the basal insulin through few daily injections. Rapid-acting insulin can also used for the replacement of the basal insulin: in this case it has to be continuously infused through insulin pumps in the form of micro-boluses.

The Post-Prandial (PP) glucose regulation remains a critical issue for diabetes treatment. An incorrect administration of a meal bolus can result in dangerous complications for the patient. In fact, hyperglycemia may occur due to an underestimated bolus, while an overestimated bolus could lead to hypoglycemia. The induced hypoglycemia due to insulin overestimation is a dangerous condition for T1D subjects. Normally, a person will feel warning symptoms when the blood sugar goes low, but people with long-standing T1D may be in condition of hypoglycemia in absence of symptoms, a physical state known as hypoglycemic unawareness. Hypoglycemia unawareness is not rare, occurring in 17% of those with T1D. In such cases, hypoglycemia is typically diagnosed through a blood sugar test. The risk of hypoglycemia unawareness is far lower in people with T2D because hypoglycemia occurs less frequently. In case of diabetic hypoglycemia, the BG level can be increased through exogenous injections or infusions of glucagon, injections of glucose, or through oral administration of fast-acting glucose. Treatment of hypoglycemia is also called hypotreatment. Usually, a hypotreatment should be able to recover the patient from hypoglycemia within 5 to 10 minutes, and the symptoms should completely disappear in approximately 10 to 20 minutes.

2.3.1 The Conventional Therapy

In order to treat T1D, the patients can rely on the CT, which is adapted by the physician to the patient. The CT consists of the combination of the basal insulin i_b, defined as a stepwise constant function and responsible for maintaining stable the blood glucose (BG) levels during fasting periods, and the insulin bolus i_B, defined as an impulse-like function and specifically administered at meal times to compensate the glucose rise due to a meal. Then, the CT is defined as follows:

$$i(k) = i_b(k) + i_B(k) \tag{2.1}$$

where i_B is the insulin bolus in correspondence to the meal taken at time k, and i_b [U/h] is the basal insulin, which is defined as a piece-wise constant function with respect to k, and i_B [U/h] is the insulin bolus associated with the meal taken at time k. Its calculation is generally performed by considering the Carbohydrate-to-insulin Ratio (CR) and the Correction Factor (CF), which are clinical parameters identified by the physician from patient history/habits. CR relates how many carbohydrate grams are disposed of by one unit of insulin and it is then determined as the ratio between the amount of ingested carbohydrates and the optimal insulin bolus as follows:

$$CR = \frac{\text{ingested CHO}}{\text{optimal insulin bolus}}$$

where the optimal insulin bolus is determined according to the following criteria [18]:

- BG concentration, measured 3 hours after the meal, is between 85% and 110% of the basal glucose
- the minimum glucose concentration is above 90 mg/dl
- the maximum glucose concentration is between 40 and 80 mg/dl above the basal level.

CF or insulin sensitivity factor is the BG drop in mg/dl caused by one unit of insulin and is computed with the so-called 1700 rule [19]:

$$CF = \frac{1700}{TDI}$$

where TDI is the total daily insulin. Using the CR and CF, the dose of insulin bolus is computed as follows:

$$i_B(k) = \frac{\widehat{CHO(k)}}{CR(k)} + \frac{BG(k) - y_{CF}}{CF(k)}; \quad k = k_m \tag{2.2}$$

where $\widehat{CHO}$ [g] is the estimated amount of ingested carbohydrates, $CR(k)$ [g/U] and $CF(k)$ [mg/dl/U] are the patient time variant CR and CF, respectively, k_m is the meal time, BG [mg/dl] is the BG concentration measured just before the meal (through a glucometer), and y_{CF} [mg/dl] is a glucose target. A more complex but complete method used to calculate the insulin bolus considers also the patient Insulin On Board (IOB), as follows:

$$i_B(k) = \frac{\widehat{CHO(k)}}{CR(k)} + \frac{BG(k) - y_{CF}}{CF(k)} - IOB(k); \quad k = k_m \tag{2.3}$$

where $IOB(k)$ represents the still estimated active insulin remaining at time k from previous infusions [19].

2.4 Devices for Diabetes Management

In order to control the glycemia, T1D patients are able to perform a Self-Monitoring of BG (SMBG) using a glucometer, also called fingerstick. This device has a little needle used to sting a fingertip to obtain drops of blood for testing. The result of the test is shown in a digital display. The ADA does not establish strict recommendations for the minimum daily frequency of SMBG tests, but it is generally accepted that at least three SMBG checks per day is the minimum standard in case of T1D [20]. Over the years, the use of glucose meters has become

easier and faster by reducing also the needed amount of blood samples [21]. For its simplicity, glucose meters are now spreadly used, even though they cannot be used to gather information about the daily glucose trend and about the development of the glucose evolution after meals.

2.4.1 Subcutaneous insulin pumps

In order to automate the insulin injections, insulin pumps have been developed. It is a small automatic device that mimics a healthy pancreas. The insulin pump therapy includes the continuous basal insulin and the insulin boluses as described in Section 2.3.1, but the patient still needs to continuously check his/her glycemia. An external infusion pump delivers a continuous infusion of rapid- or short-acting insulin through a catheter inserted into the subcutaneous tissue of the abdominal wall. The pump is preprogrammed by the patient to deliver insulin continuously at a specified rate that is designed to meet the individual's basal insulin demands. For meal coverage, patients use the pump to deliver a specified bolus of insulin before eating. Moreover, current insulin pumps allow the patient to vary the rate of basal insulin replacement at different times of the day depending on requirements. For instance, the pump can be preprogrammed to increase the basal rate of insulin infusion in the early morning hours to hanlde the "dawn phenomenon", i.e. patients experience a physiologic early morning rise in BG levels. In addition, many insulin pumps have programming capabilities that facilitate dose selection for carbohydrate counting and insulin-correction boluses [22]. Although the use of an insulin pump can avoid individual insulin injections, it can also have some disadvantages as the risk for infection at the insertion site of the catheter. These infections usually are readily treatable with antibiotics and a change of insertion site. Occasionally, there may be subcutaneous abscess requiring further drainage. The interruption of basal insulin delivery due to pump malfunction or catheter disruption may occur and can rapidly lead to hyperglycemia or even the possibility of causing DKA or weight gain. However, most pump users agree the advantages outweigh the disadvantages. This is the reason why several companies invested in the production of insulin pumps, and, in particular, continuous subcutaneous insulin pumps. A subcutaneous insulin pump is battery powered and it is composed of three main parts: a processing module to coordinate the components of the device and to program insulin delivery; an infusion set,which includes a thin tube inserted with a needle in the subcutaneous tissues, and finally, a reservoir of insulin. A characteristic of subcutaneous insulin pumps is the quantization, which indicates the minimum amount of insulin that can be infused. Among the most popular subcutaneous pumps it is worth to mention the Accu-Check® Combo system (Figure 2.3). This pump is of particular interest because it has been used as a component of the AP described in Chapter 5 during the execution of the clinical trials described in Sections 6.2.1.

FIGURE 2.3: Example of subcutaneous insulin pump: the Accu-Check® Combo system. Processing module on the left and infusion set on the right.

2.4.2 Continuous Glucose Monitor

Following the subcutaneous insulin pumps, subcutaneous glucose sensors (or Continuous Glucose Monitor, CGM) have been developed. These devices are less invasive because they measure glucose in the interstitial tissues (i.e. subcutaneous tissues) rather than blood stream glucose. Unlike glucometers, which measure the current glucose level at a single point in time, CGM provides maximal information presenting a constant stream of data (measured every 1-5 minutes), not only indicating the current interstitial glucose level, but also trends in glucose direction and velocity of glucose change [23]. Since a CGM system provides glucose direction and rate of change, CGM can provide valuable information during the day improving the optimal treatment decisions for the diabetic patient. CGM systems help to minimize the guesswork that comes with making treatment decisions based solely on a limited number of a BG meter readings. If the sensor is inaccurate, i.e. there is a gap between CGM readings and finger stick measurements, an easy calibration procedure can be performed by providing a fingerstick BG measurement using a BG meter within 5 minutes. An example of CGM is Dexcom G4® Platinum CGM system with Share, shown in Figure 2.4: it has been used for performing the clinical trials presented in this thesis. Dexcom G4® Platinum CGM is composed of a sensor that measures glucose levels just underneath the skin, a trasmitter used to send glucose information wirelessly measured by the sensor, and of a receiver where the BG level is plotted on a display with the trend of the measured data with respect to adopted target range. This device allows to set a customized BG target range and provides an alarm system to alert if BG concentration falls outside that range. Moreover, the Dexcom G4® PLATINUM has a fixed hypo alert at 3.1mmol/l ($\sim$ 55.8 mg/dl) to guarantee the safety of the patient even when he/she is sleeping.

FIGURE 2.4: Dexcom G4® Platinum CGM system with Share.

A substantial disadvantage of subcutaneous sensors is that data are strongly affected by measurement noise. Indeed, CGM measurements are less accurate with respect to measurements provided by glucometers. Glucometers measure the glucose directly from the bloodstream, whereas CGM measures glucose in the interstitial tissues. Thus, considering that there is a physiological time lag due to the diffusion process of glucose from the bloodstream to the subcutaneous tissues, CGM measuremnts are affected by a delay. When glucose levels are rapidly changing, CGM measurements result wrong with respect to the correct current BG value by leading to erroneous and potentially dangerous patient response [23], e.g. hypoglycemia alarms may be delayed. CGM may be affected also by drift error, which can be compensated by performing few daily re-calibration using fingerstick measurements. Despite CGM provides delayed and noisy BG measurements, its main advantage resides on the high frequency of measures, which can lead to define glucose trends both during fasting and postprandial periods. The latest technological developments have substantially increased the CGM reliability, and the remaining inaccuracies can be compensated through few daily re-calibrations.

Chapter 3

The simulator

Simulation software are a helpful support for scientists and researchers. These tools provide several advantages like reducing testing time and guaranteeing repeatability and versatility. Indeed, virtual design can eliminate the need of many costly experiments and can speed up the development process. Moreover, simulations provide highly flexible techniques applicable for many research fields, i.e. it allows to run scenarios that cannot be easily achieved in real situations.

Several diabetes simulation tools have been developed to model glucose-insulin dynamics of T1D patients. These simulators were based on comprehensive mathematical population models, so their prediction capabilities were generally limited to describe population averages observed in the context of a clinical trial. However, an average-model approach is not sufficient for realistic *in silico* experimentation with control scenarios, where facing with inter-subject variability is particularly challenging. Moreover, the evolution of the AP would be greatly accelerated by employing mathematical modeling. In this prospective the UVA/Padova simulator was developed in 2008 [24]. A system capable of simulating the glucose-insulin dynamics of a particular person was required to develop an AP. The Universities of Virginia and Padova have developed the UVA/Padova simulator to design and test treatments for T1D. Indeed, the UVA/Padova simulator provides realistic computer simulation of clinical trials, together with invaluable information about the safety and the limitations of the control algorithms, guiding clinical experiments, and out-ruling ineffective control scenarios in a cost-effective manner [24]. Although computer simulation are realistic, simulation cannot be considered a complete replacement of clinical trials [25] but, they can demonstrate safety and efficacy of the therapy under real-life conditions before testing it in T1D subjects.

The UVA/Padova simulator has been equipped with a cohort of *in silico* subjects to cope with a wide prospective of the intra- and inter-subject variability of T1D subjects. In order to cover

all ages, the cohort includes 100 virtual adults, 100 virtual adolescents, and 100 virtual children. Virtual subjects were created by fitting a metabolic model to data of individuals collected during clinical trials. The Food and Drug Administration (FDA) approved the UVA/Padova simulator as equivalent substitute of pre-clinical animal experiments for T1D treatment [26]. This acceptance allowed the *in silico* synthesis of control algorithms directly testable on real patients. The utilization of a simulator for the development of an effective T1D treatment reduces the need of long and expensive pre-clinical experiments usually performed on animals.
The UVA/Padova simulator includes a large nonlinear compartmental model able to simulate the glucose-insulin dynamics of the diabetic population [27]. In the most recent version of the simulator, the circadian variability of insulin sensitivity and meal absorption parameters have been added [25, 28, 29] to simulate the inter-subject variability of this population. In order to simulate the glycemia of a specific patient, individualized model identification techniques have been developed. In this context, several metabolic models were developed through realistic closed-loop clinical protocols simulated via the UVA/Padova simulator [24, 27], as described in [30–33].
In addition to the cohort of *in silico* subjects, the simulation environment includes also a virtual sensor and a virtual pump that can reflect the characteristics of several commercial devices, also known as virtual hardware. Time lags and random noise of subcutaneous CGM devices have been also included in the simulator. In the following, the patient model used in the UVA/Padova simulator, and the sensor and pump models wills be presented in details.

3.1 UVA/Padova Metabolic Model

The UVA/Padova simulator is a large-scale maximal model. Specifically, it includes a compartmental model that describes the glucose-insulin dynamics of a generic T1D patient including 13 differential equations and 36 subject-specific parameters. The model representation is shown in Figure 3.1. In particular, carbohydrate intake and insulin injections are the model inputs, while the model output is the glucose concentration. The structure of the metabolic model included in the simulator is based on the physiological characteristics of a generic subject. The first version of the UVA/Padova simulator was released in 2006 and it has been continually refined over the years. In 2013 an updated version of the UVA/Padova T1D simulator was released, and approved by FDA, including an improved model of hypoglycemia and a new model of glucagon dynamic [27]. In 2017 the inter-subject variability of the virtual population was included by modelling through different sets of metabolic parameters of this model [29, 34]. This fact allows to extend the use of the UVA/Padova T1D simulator from a single meal to multiple days.

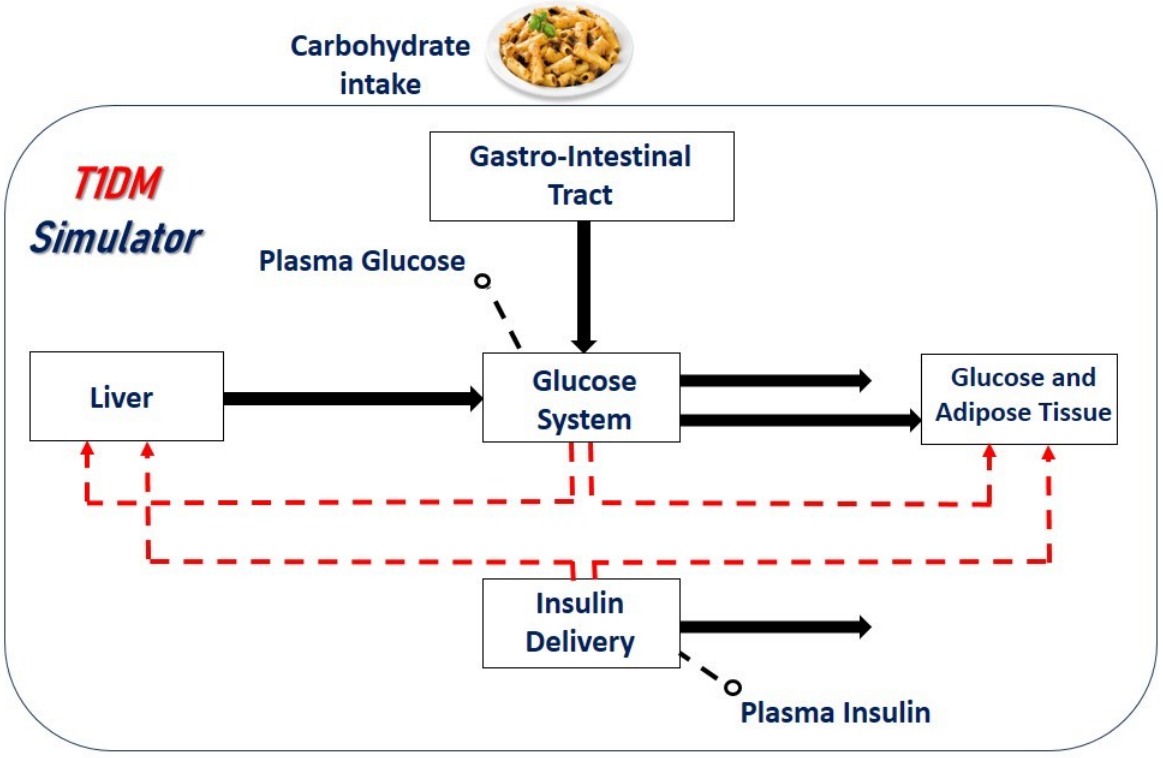


FIGURE 3.1: The representation of the metabolic model included in the UVA/Padova simulator.

This new feature guarantees more realistic testing scenarios because it allows to simulate that the glucose response to a certain perturbation should be strictly dependent from the time at which the perturbation happens. The last version of the UVA/Padova simulator is presented in this work. Given that the exogenous glucagon delivery is not used in this thesis, the entire UVA/Padova model will be presented in the following except the glucagon subsystem.

Glucose rate of appearence

The gastro-intestinal tract has been modeled to reproduce the effects of a meal on the glycemia. Three compartments, two for the stomach and one for the gut, respectively, simulate the digestion process [35]. The rate of appearance represents the rise of glycemia due to the injested carbohydrates, as follows:

$$\begin{cases} Q_{sto}(t) = Q_{sto1}(t) + Q_{sto2}(t) & Q_{sto}(0) = 0 \\ \dot{Q}_{sto1}(t) = -k_{max} \cdot Q_{sto1}(t) + Dose \cdot \delta(t) & Q_{sto1}(0) = 0 \\ \dot{Q}_{sto2}(t) = -k_{emp} \cdot (Q_{sto}(t)) \cdot Q_{sto2}(t) + k_{max} Q_{sto1}(t) & Q_{sto2}(0) = 0 \\ \dot{Q}_{gut}(t) = -k_{abs} \cdot Q_{gut}(t) + k_{emp}(Q_{sto}(t)) \cdot Q_{sto2}(t) & Q_{gut}(0) = 0 \\ Ra_{meal}(t) = \dfrac{f_c \cdot k_{abs} \cdot Q_{gut}(t)}{BW} & Ra_{meal}(0) = 0 \end{cases} \tag{3.1}$$

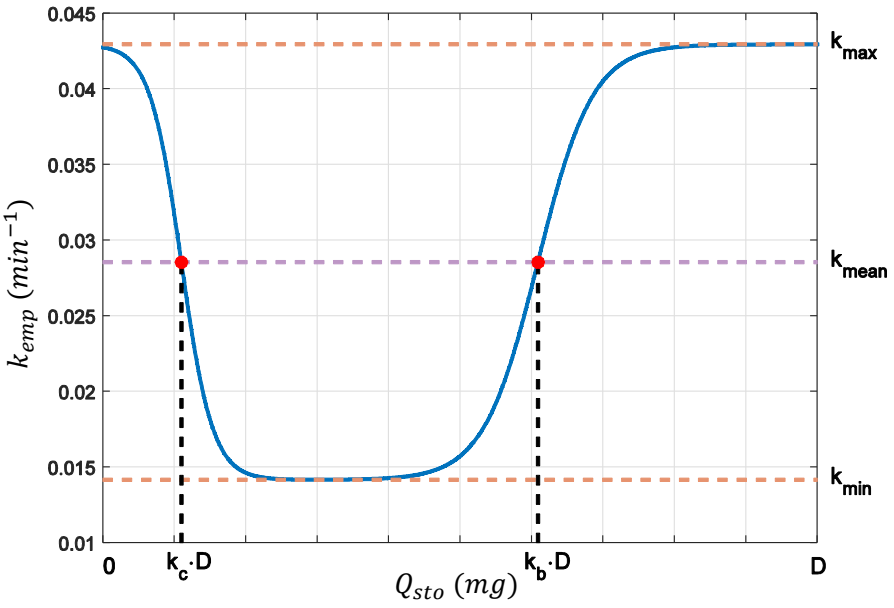


FIGURE 3.2: Gastric emptying coefficient $k_{emp}(Q_{sto}(t))$ [min^{-1}] as function of the amount of carbohydrates in the stomach $Q_{sto}(t)$ [mg].

where Q_{sto} [mg] is the total amount of carbohydrates into the stomach, Q_{gut} [mg] are the carbohydrates into the gut, k_{max} [min^{-1}] and k_{abs} [min^{-1}] are model parameters, $Dose$ [mg] is the amount of carbohydrates eaten at time t, f_c is the fraction of carbohydrates transferred into the glucose system and affects the glucose rate of appearance Ra_{meal} [mg/kg/min], and BW [Kg] is the patient body weight. k_{emp} [min^{-1}] is the time-varying nonlinear function representing the gastric emptying coefficient and is defined as follows:

$$k_{emp}(Q_{sto}(t)) = k_{min} + \left(\frac{k_{max} - k_{min}}{2}\right) \cdot \{\tanh\left[\alpha\left(Q_{sto}(t) - k_b \cdot Dose(t_i)\right)\right] - \\ - \tanh\left[\beta(Q_{sto}(t) - k_c \cdot Dose(t_i))\right] + 2\} \tag{3.2}$$

where k_{min} and k_{max} are the minimum and the maximum values of the gastric emptying coefficient, respectively, and where:

$$\alpha = \left(\frac{5}{2D(t_i)(1-k_b)}\right), \quad \beta = \left(\frac{5}{2D(t_i)k_c}\right), \quad Dose(t_i) = \int_{t_i}^{t_f} d(\tau)\, d\tau + Q_{sto}(t_i) \tag{3.3}$$

with $t_i < t$ and t_f representing the beginning and ending times of the i^{th} eaten meal, respectively. k_b and k_c are percentage values associated to the portion of non-digested meal when:

$$k_{mean} = \left(\frac{k_{max} + k_{min}}{2}\right) = k_{empt} \cdot (k_b \cdot Dose(t_i)) = k_{empt}(k_c \cdot Dose(t_i)) \tag{3.4}$$

As shown in Figure 3.2, the gastric emptying coefficient reaches its maximum value at the beginning of a meal, then decreases until it reaches its minimum values at about half digestion time, returning to its maximum value when the stomach is almost empty. The identification of the gastrointestinal tract parameters is described in [35, 36].

Glucose system

After digestion, the meal appears in the glucose system, producing the increase of glucose level in blood stream. The glucose subsystem consists of a two-compartment model of glucose kinetics.

$$\begin{cases} \dot{G}_p(t) = EGP(t) + Ra_{meal}(t) - U_{ii}(t) - E(t) - k_1 \cdot G_p(t) + k_2 \cdot G_t(t) & G_p(0) = G_{pb} \\ \dot{G}_t(t) = -U_{id}(t) + k_1 \cdot G_p(t) - k_2 \cdot G_t(t) & G_t(0) = G_{tb} \\ G(t) = \dfrac{G_p(t)}{V_G} & G(0) = G_b \end{cases} \tag{3.5}$$

where G_{pb} and G_{tb} [mg/kg] are the plasma and the tissue glucose masses, respectively, G [mg/dl] is the plasma glucose concentration, suffix b denotes basal state, EGP [mg/kg/min] is the endogenous glucose production, E [mg/kg/min] is the renal excretion, U_{ii} and U_{id} [mg/kg/min] are the insulin-independent and -dependent glucose utilizations, respectively, V_G [dl/kg] is the distribution volume of glucose, and k_1 and k_2 [min^{-1}] are rate parameters [36]. EGP is defined by the following equation [27, 37]:

$$EPG(t) = k_{p1} - k_{p2} \cdot G_p(t) - k_{p3} \cdot X^L(t) + \xi \cdot X^H(t) \tag{3.6}$$

with

$$\begin{cases} \dot{X}^L(t) = -k_i \cdot \left[X^L(t) - I'(t)\right] & X^L(0) = I_b \\ \dot{I}'(t) = -k_i \cdot \left[I'(t) - I(t)\right] & I'(0) = I_b \\ \dot{X}^H(t) = -k_H X^H(t) + k_H \cdot \max\left[(H(t) - H_b, 0\right] & X^H(0) = 0 \end{cases} \tag{3.7}$$

where X^L [pmol/L] is the delayed insulin action in the liver, X^H [ng/dl] is the delayed glucagon action on EGP, ξ is the liver responsiveness to glucagon and it is also constrained to be non-negative, H [ng/dl] is the plasma glucagon concentration, $\frac{1}{k_H}$ [min] is the delay between glucagon concentration and action, I [pmol/L] is the plasma insulin concentration, I' [pmol/L] represents a delay compartment for insulin action on glucose production, and k_i, K_{p1}, K_{p2}, K_{p3} are rate parameters [36]. EGP is also constrained to be non-negative.

The insulin-independent glucose utilization U_{ii} of Eq. 3.5 represents the insulin mainly used by the brain and takes place in the first compartment of the glucose system. This quantity is assumed to be constant. The insulin-dependent glucose utilization U_{id} takes place in the remote

compartment and depends on the glucose in the tissues by the following Michaelis-Menten relationship:

$$\begin{cases} U_{id}(t) = \dfrac{[V_{m0} + V_{mx} \cdot X(t) \cdot (1 + r_1 \cdot risk)] \cdot G(t)}{K_{m0} + G_t(t)} \\ \dot{X}(t) = -p_{2U} \cdot X(t) + p_{2U} \cdot [I(t) - i_b] \qquad X(0) = 0 \end{cases} \tag{3.8}$$

with X [pmol/L] the insulin action on glucose utilization and V_{m0}, V_{mx}, K_{m0}, p_{2U} rate parameters [36]. The $risk$ function has been introduced to model the increase of the insulin-dependent glucose utilization U_{id} when glucose decreases below a certain threshold [27]:

$$risk = \begin{cases} 0 & \text{if } G(t) \geq G_b \\ 10 \cdot [f(G(t))]^2 & \text{if } G_{th} \leq G(t) < G_b \\ 10 \cdot [f(G_{th})]^2 & \text{if } G(t) < G_{th} \end{cases}$$

with G_b the basal glucose, G_{th} the hypoglycemic threshold (set equal to 60 mg/dl),

$$f(G(t)) = \ln\left(\frac{G(t)}{G_b}\right)^{r_2}$$

and r_1, r_2 model parameters. The $risk$ function is a measure of the risk associated with a certain glucose level [38].It represents a real body behavior, that increases the insulin-dependent glucose utilization under the conditions of hypoglycemia, thus further increasing the risk of severe hypoglycemia. Glucose excretion in T1D patients increased in a proportional manner with increasing blood glucose. The renal excretion E of Eq. 3.5 is described by the following relationships:

$$E(t) = \begin{cases} k_{e1} \cdot [G_p(t) - k_{e2}] & \text{if } G_p(t) > k_{e2} \\ 0 & \text{if } G_p(t) \leq k_{e2} \end{cases}$$

where k_{e1} [min^{-1}] represents the glomerular filtration and k_{e2} [mg/kg] represents the renal threshold for glucose.

Subcutaneous glucose kinetics

A first order system has been introduced to simulate the subcutaneous glucose of the patient, that is also the model output:

$$\begin{cases} \dot{G}_s(t) = -k_{sc} \cdot G_s(t) + k_{sc} \cdot G(t) \quad G_s(0) = G_b \\ y(t) = G_s(t) \end{cases} \tag{3.9}$$

where k_{sc} [min^{-1}] is the transfer coefficient from plasma to subcutaneous glucose. The delay $\frac{1}{k_{sc}}$ models the physiological time lag due to the glucose transition from plasma to subcutaneous tissues.

Time-Varying Model of T1D Subject

The most important feature of last version concerns the model of T1D subject, which includes intra- and inter-day variability of insulin sensitivity, dawn phenomenon effect, and different distributions of Carbohydrate-to-insulin Ratio (CR) at different day time [29, 34]. Thus, the last version of the simulator allows the description of intra-subject diurnal glucose variability. Since the same treatment applied to the same patient can obtain significantly different results day by day, The UVA/Padova simulator domain of validity is extended from a single-meal to a single-day scenario.

The incorporation into the simulator is related to the result of a clinical study conducted on T1D subjects, which has revealed the existence of diurnal patterns of insulin sensitivity, i.e. different insulin sensitivity at breakfast, lunch and dinner [39]. Then, the results of this experiment have allowed to perform their estimation [18, 28]. The intra-day variability of insulin sensitivity was incorporated into the simulator by associating a certain intra-day variability pattern of model parameters V_{mx} 3.8 and k_{p3} 3.6, representing insulin action on glucose utilization by tissues and on glucose production by the liver, respectively [29, 34].

Dawn phenomenon effect consists in an increase in BG in the early morning and occurs when hormones (including cortisol, glucagon, epinephrine) are released by the body, causing an increased Endogenous Glucose Production (EGP). These hormones give rise to a brief period of insulin resistance, an increased insulin requirements have to be fulfilled from 3:00 to 7:00. The EGP variation is described as a linear increase of basal EGP (EGP_b), while, the increased insulin requirement is modelled as a decrease in insulin-dependent glucose utilization (U_{id}) [29, 34].

The CR defined in Section 2.3.1 is modelled in the last version of the UVA/Padova simulator as a time-varying parameter along the day: each *in silico* subject has been equipped with diurnal pattern of CR [18, 28]:

$$CR_j = \frac{\text{ingested CHO}_j}{\text{optimal insulin bolus}_j} \quad \text{with j=B,L,D.}$$

where CHO stands for the estimated amount of carbohydrates.

3.2 Virtual population

Simulators based on an average-model approach have limitations on capabilities to prediction. These simulators allow only prediction of population averages that would be observed during clinical trials, so they are not realistic for *in silico* experimentation. In order to simulate the glucose insulin dynamics of a specific individual, the UVA/Padova simulator is equipped with a collection of virtual patients, which spans sufficiently well the variability present in a general population of people with T1D. In order to have a complete description of the multiple system fluxes shown in Figure 3.1, in 2006, 204 nondiabetic individuals received a triple tracer meal protocol in order to estimate the virtually model-independent fluxes of the system, like the rate of appearance in plasma of ingested glucose, glucose production, glucose utilization, and insulin secretion [40]. Instantaneous values of blood glucose and insulin concentration are not sufficient to identify the involved metabolic processes. On the other hand, flux information minimizes structural uncertainties in modeling the various subsystems.

For each of the 204 subjects, the key metabolic parameters were identified and were included in a database. The adopted identification strategy is described in detail in [36]. Then, model parameters and their joint probability distribution were computed by exploiting this data collection. Since the parameter distributions were mostly log-normal, the average vector and the covariance matrix of the log-transformed parameters vector uniquely define this probability distribution. Then, *in silico* individuals to be included in the UVA/Padova simulator have been generated by using joint probability distribution. 100 virtual *in silico* subjects have been generated by randomly extracting different realizations of the model parameters from the identified joint parameter distributions. In order to represent T1D *in silico* subjects, the identified model parameters of the 100 nondiabetic *in silico* subjects were modified. In case of T1D patient, there is no endogenous insulin production, but the insulin is introduced by subcutaneous exogenous insulin infusions. Thus, the insulin secretion subsystem has substituted with an exogenous insulin delivery subsystem, where the insulin rate of appearance in the bloodstream has been introduced.

Since glycemia is monitored by using a CGM sensor, it is introduced the subcutaneous glucose kinetics, described in Eq. 3.9, to simulate the interstitial glucose measurements.

Since even single-tracer studies in type 1 diabetes are scarce, the description of the inter-subject variability was difficult. In order to obtain parameter joint distributions in T1D from those in the healthy state, the inter-subject variability was assumed the same of healthy subjects, that is having same covariance matrix. This assumption is based on the hypotesis that each subject is assumed in good control. Furthermore, the average vector of the distribution has been modified to reflect the clinically relevant differences with respect to the non-diabetic individuals, i.e. the

average basal glucose was set equal to 120 mg/dl [27].
Thus, since the description of the joint probability distribution reproduces the inter-subject variability, the simulator allows the possibility of defining and testing an individualized insulin treatment for T1D patients.
An additional *in silico* subject, known as the "average" patient, has been obtained averaging the parameters associated to the 100 *in silico* subjects. Thus, the average patient represents a patient with the average dynamics of the population.

3.3 The simulation environment

In order to create a comprehensive environment for algorithm testing, the metabolic model is included in a software environment developed in Matlab®. This software allows to perform simulations of T1D virtual patients and test several subcutaneous insulin treatments. The software allows for

- defining a testing scenario, i.e., a schedule of meals with corresponding CHO amounts, sessions of physical activity and times of simulated sensor re-calibrations
- selecting subjects
- selecting virtual hardware, i.e. CGM sensor and insulin pump
- selecting a set of outcome metrics to evaluate the performance of the insulin therapy.

The insulin therapy can be either the CT or it can be driven by a controller, which defines and adjusts periodically the current insulin treatment by exploiting the measurements coming from the virtual CGM sensor and additional information provided by the patient. The schema in Figure 3.4 shows how the UVA/Padova metabolic model is related with the virtual hardware and the virtual controller. The model allows to simulate the subcutaneous glucose measurements in response to an insulin therapy. The reading of subcutaneous glucose measurements is performed by a simulated glucose sensor and the insulin injections are infused by a simulated insulin pump. The simulated controller closes the loop: it can command the virtual pump on the basis of the measurements coming from the virtual sensor and other information from the patient.
The performance of the insulin therapy is evaluated on the basis of the output of the simulator, which is the subcutaneous glucose resulting from a specific insulin treatment of a virtual patient simulated in a specific scenario. The UVA/Padova simulator allows the so-called meal announcement, i.e. the patient can announce a meal intake in advance specifying the estimated quantity

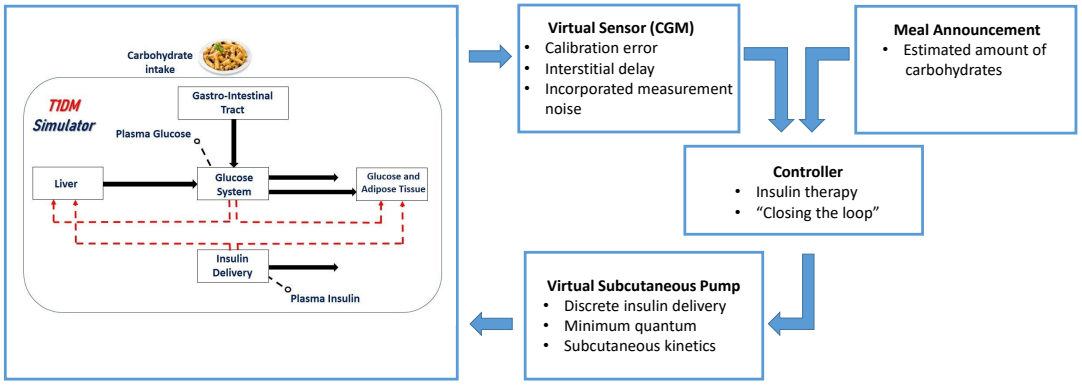


FIGURE 3.3: Schema of the UVA/Padova metabolic simulator.

of carbohydrates (CHO) included in the meal itself. The announced amount of carbohydrates can be used to compute the optimal nominal insulin bolus according to CT to compensate the induced glucose rise. The meal announcement allows to include the possibility of errors in the announcement such as a limited events of unannounced meals or meals announced with a wrong estimation of the amount in order to define a real-life scenario.

3.4 Virtual hardware

The simulation environment includes an implementation of the physical devices that are involved in the diabetes treatment: the CGM sensor and the insulin pump. In order to perform realistic simulations, the physical limitations of the devices have been implemented in the simulation software.

In silico insulin pump is used to simulate subcutaneous insulin delivery. The parameters of virtual pumps are the discrete insulin infusion corresponding to step-wise basal pump rate and insulin boluses and the minimum/maximum amount of insulin that could be injected. Moreover, the simulator allows to simulate various type of insulin pumps.

The virtual CGM sensor simulates the measurement of the subcutaneous glucose concentration in the interstitial tissues. *In silico* sensor is developed on the basis of analysis of sensor errors, and this sensor simulation model has been thought to provide worst-case scenario sensor errors [24]. Virtual sensor has been implemented as a block defined as follows:

$$CGM(t) = IG(t) + \epsilon(t) \tag{3.10}$$

where IG is the subcutaneous glucose concentration in the interstitial tissues and ϵ is a measurement noise. The error is then expressed in terms of in estimated glucose concentration in

the interstitial tissues ($\hat{IG}$) by taking into account the glucose diffusion process from blood to interstitial fluid. The relation between the interstitial glucose concentration and BG is described through a first-order system with time constant τ in the Laplace domain:

$$\frac{\hat{IG}(s)}{BG(s)} = \frac{1}{\tau s + 1} \tag{3.11}$$

where τ is representative of the lag between the blood and subcutaneous glucose concentrations. A block schema representation of error is shown in Figure 3.4. The identification of sensor error model is described in [41]. The sensor error model is defined as follows:

$$\epsilon_{PV}(k) = a_1\epsilon_{PV}(k-1) + a_2\epsilon_{PV}(k-2) + v(k) \tag{3.12}$$

where the error v is a gaussian white noise with mean μ_v and variance σ_v^2, the parameters a_1 and a_2 are coefficients of the AR model, k is the discrete sampling time and the distribution of the process initial states is characterized by mean μ_{is} and covariance matrix σ_{is}^2. The identified values associated to the error model [41] ϵ_{PV} are:

- a_1=1.5458
- a_2=-0.5708
- μ_v=0.0017 mmol/L
- σ_v^2=0.0283 (mmol/L)^2
- μ_{is}=[-0.1766 -0.1566] mmol/L
- $\sigma_{is}^2 = \begin{bmatrix} 0.7759 & 0.7895 \\ 0.7895 & 0.8603 \end{bmatrix}$ (mmol/L)^2

In vivo data coming from AP@home "CAT" clinical trial performed in 2012 have been used to identify the parameters of the model. In the following years, since the error model ϵ_{PV} has resulted to overestimate with respect to the real errors, the current error model variance has been reduced by a factor of 6 to reflect the improvement of the last generation of CGM readings. Moreover, the virtual sensor has to cover additional hardware limitations, i.e. the possible measurements of interstitial glucose concentrations are limited between 30 mg/dl and 600 mg/dl.

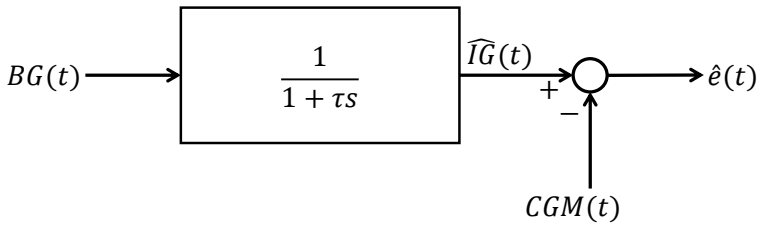


FIGURE 3.4: Block schema representation of the error $\hat{e}$.

3.5 Performance metrics

The UVA/Padova simulator allows to test and validate several open- and closed-loop therapies. The efficacy of an insulin therapy can be quantified through performance metrics commonly used during clinical trials. The consensus statement for artificial pancreas trials described in [42] identified a short set of basic, easily interpreted outcome measures to be reported in AP studies. Firstly, metrics based on average glycemia and deviations from target are considered. By taking into account the physiological glucose range of variation for patients affected by T1D, the following glycemic regions can be defined [11]:

- safe or target range (from 70 mg/dl to 180 mg/dl)
- euglycemic or tight target range (from 80 mg/dl to 140 mg/dl)
- hyperglycemia range (above 180 mg/dl)
- severe hyperglycemia range (above 250 mg/dl)
- hypoglycemia range (below 70 mg/dl)
- severe hypoglycemia range (below 60 mg/dl).

In order to evaluate the effectiveness of an insulin therapy, the calculation of percentage times spent in each of the presented ranges is computed using the generated measurements coming from virtual CGM. So, the metrics of interest include:

- average (A) and standard deviation (SD) of glucose [mg/dl]
- time in target or percentage of time spent in safe range, i.e. from 70 mg/dl to 180 mg/dl (T_r)
- tight target range or percentage of time spent in euglycemic range, i.e. 80-140 mg/dl (T_{tr})
- severe hypoglycemia range or percentage of time spent in hyperglycemia ($T_{a_{hyperThreshold}}$)

- time above target or percentage of time spent above 180 mg/dl (T_a)
- time below target or percentage of time spent below 70 mg/dl (T_b)
- severe hypoglycemia range or percentage of time spent in hypoglycemia ($T_{b_{hypoThreshold}}$) .

Additinal metrics can include the Root Mean Square Error (RMSE) respect a glucose reference of 120 [mg/dl], and the Total Daily Insulin (TDI) delivered to the patient, and variability and risk assessment. The Low Blood Glucose Index (LBGI) measures the frequency and the risk of low BG measurements, whereas the High Blood Glucose Index (HBGI) measures the frequency and the risk of high BG readings. The LBGI has been demostrated to be a good predictor of severe conditions of hypoglycemia [43, 44], while the HBGI is a relevant index for hyperglycemia condition [43]. Metrics that evaluate events and other clinically relevant characteristics are also considered.
These metrics are computed overall (O), during night (N, 0:00 - 6:00), and as an average of all the PP periods (4h after meal). Median [25^{th}; 75^{th}] percentiles for non-Gaussian distributed data and mean ($\pm$ standard deviation) otherwise are reported for the various indices. Confidence intervals on the mean or median are reported as well. The gaussianity and homoscedasticity of the data distributions are assessed by the Lilliefors test and two-sample F-test, respectively.

3.5.1 Statistical Comparison of Performance Metrics

Statistical methods have to be introduced to evaluate the significance of a comparison among different insulin therapies. In order to evaluate the significant differences, the more appropriated statistical test is selected based on the characteristics of the data distributions. If at least one distribution is non-Gaussian, the Wilcoxon rank sum test is used; if both distributions are Gaussian and homoscedastic, a two-sample t-test is performed; otherwise, if the homoscedasticity is not satisfied, the two-sample t-test with Satterthwaites approximation is used. The methods used to evaluate the statistical significance of a difference between metrics are generally based on the p-value (also denoted as p). In statistical hypothesis testing, p-value method evaluates how well the measured samples support a statistical hypothesis, also called null hypothesis. The null hypothesis states that there is no difference between metrics. If the p-value is lower than the significance level associated to the test, it suggests that the measured samples are inconsistent with the null hypothesis, so the null hypothesis can be rejected. If the null hypothesis is rejected, a result is said to be statistically significant. In case of evaluating two insulin therapies over a population, if the null hypothesis is rejected, this means the measured difference between the metrics depends on different effectiveness of the applied insulin therapies

and not on a sampling error. Sampling error can occur when the statistical characteristics of a population, like mean and standard deviation, are estimated from a subset, or sample, of that population.
Additional statistical methods like ANalysis Of VAriance (ANOVA) will be considered later for a deep analysis of clinical data.

3.5.2 ANalysis Of VAriance (ANOVA)

ANOVA provides a statistical test which is able to prove if a statistical correlation exists among the so-called indicators (quantitative variable) and the independent variables (explanatory variable), usually categorical. Each value of the explanatory variable is called group and the values of the indicators are gathered with respect to each group.
The ANOVA is used to calculate the difference between the means of the indicator of the groups, its aim is to find if the mean between groups is statistically significantly different from the mean within groups and variance is used to infer about means. For this reason the test is called analysis of the variance.
The means of the groups are significantly different if the variability between groups is larger than the variability within groups [45].
The ANOVA aims to prove which of the two following hypothesis is corrected:

- H_0: the means of all groups are equal
- H_a: the means of the groups are not all equal, that is different from: "all are unequal".

To make the results of calculation completely trustworthy, the ANOVA has some underlying assumptions:

1. Subjects are chosen via a random sample
2. The response variable is normally distributed within each group
3. The population standard deviation is the same for all group

The test of ANOVA is robust, in fact the 2^{nd} and 3^{rd} point can be relaxed in the following way:

- the data point, visualized in a normal quantile plot, fall close to a line
- after computing the standard deviation of each group, the ratio between the largest and the smallest one must be maximum two.

The ANOVA uses the F-statistics, that computes the ratio of the variability between groups and the variability within groups. If F is large, there is the rejection of the null hypothesis of equal means, it means that the specific quantitative variable is related to the interested explanatory variable. If F is small, the hypothesis of equal means is not rejected, so any type of correlation can not be drawn. Specifically, the *one-way* ANOVA is the ANOVA test used when there is just one explanatory variable.

The notation used in this section is the following one:

dim = number of groups of the explanatory variable

n_i = sample size of group i

x_{ij} = the jth value of the indicator in the group i

$\bar{x}_i$ = mean of the indicator of the i-th group = $\frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$

s_i = sample standard deviation from the i-th group = $\frac{1}{n_i-1}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2$

n = the total sample $\sum_{i=1}^{dim} n_i$

$\bar{x}$ = the mean of all values of the indicator = $\frac{1}{n}\sum_{ij} x_{ij}$.

In order to calculate the F statistic, the ANOVA uses the following metrics:

$$\text{Sum of Squares Total } (SST) = \sum_{i=1}^{dim}\sum_{j=1}^{n_i}(x_{ij}-\bar{x})^2$$

This variability has two sources:

1. variability between group means (variation around the overall mean $\bar{x}$)

$$SSG := \sum_{i=1}^{dim} n_i(\bar{x}_i-\bar{x})^2$$

2. variability within group means (variation of the observation about their group mean $\bar{x}_i$)

$$SSE := \sum_{i=1}^{dim}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2 = \sum_{i=1}^{dim}(n_i-1)s_i^2$$

So, SST is the sum of the two contributes:

$$SST = SSG + SSE$$

and finally F is equalled to:

$$F = \frac{MSG}{MSE}$$

where MSG and MSE are equalled to:

$$MSG = \frac{SSG}{dim - 1}$$
$$MSE = \frac{SSE}{n - dim}$$

The computation of these indices is done in this thesis through Matlab®, that produces the

Source	SS	df	MS	F
Model/Group	SSG	$dim - 1$	$MSG = \frac{SSG}{dim-1}$	$\frac{MSG}{MSE}$
Residual/Error	SSE	$n - dim$	$MSE = \frac{SSE}{n-dim}$	
Total	SST	$n - 1$		

TABLE 3.1: ANOVA result table.

Table 3.1.

The columns of this table are the following one:

- SS (Sum of Square)

  It lists:

  SSG: measured variation of the group means around the overall mean $\bar{x}$

  SSE: measured the variation of each observation around its group mean $\bar{x_i}$

  SST: measured variation of the data around the overall mean $\bar{x}$

- df (degrees of freedom)

  where

  $dim - 1$ is for SSG

  $n - dim$ is for SSE

  $n - 1$ is for SST

- MS (Mean Square)

  The MS is defined as: $MS = \frac{SS}{df}$

  This is like standard deviation and lists MSG and MSE defined before.

  Another formula for MSE is:

$$MSE = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_{dim} - 1)s_{dim}^2}{(n_1 - 1) + (n_2 - 1) + \cdots + (n_{dim} - 1)} \tag{3.13}$$

  MSE is also called s_p^2

- F (F statistic)

  The F statistic is defined as: $F = \frac{MSG}{MSE}$

If the null hypothesis is true, the F statistics has an F distribution with $dim - 1$ and $n - dim$ degrees of freedom in the numerator/denominator. If the alternate hypothesis is true, the F tends to be large.

The hypothesis H_0 is rejected in favor of H_a if F is sufficient large.

As in other test statistics the determination of the width of F is given by the p-value. In order to read it, it is necessary to know the degrees of freedom associated to the numerator (MSG) and the denominator (MSE). Given the ANOVA results, additional information can be extracted by them using the multiple comparison, a method able to fully understand the difference between the groups. [46]

In particular when $P(F_{dim-1,n-dim} > F_{computed}) < \alpha$, where α is the level of significance, the means of at least 2 groups are different, but no information about what groups and how many groups are involved is provided.

It is important to know which groups are different, so it is necessary to execute a significant F-statistic with *pairwise comparison* of the means and it is computed with the t-test between each pair of means:

$$t_{ij} = \frac{\bar{x}_i - \bar{x}_j}{s_p\sqrt{1/n_i + 1/n_j}}$$

where $t_{i,j}$ is the t-test between the mean of the group i and group j, n_i and n_j are the number of elements of the group i and the number of elements of group j respectively. The statistic significant of the t-test is related to the t-test table using $n - dim$ degree of freedom (the df associated with s_p).

The main advantage of this approach is that it is very easy, as well as very widely applicable. The main disadvantage is that it often is unnecessarily conservative (weak): α is smaller than it needs to be.

The procedure for performing multiple comparisons involves the Tukey's Method that test all possible pairwise differences of means to determine if at least one difference is significantly different from 0.

Chapter 4

Smart Conventional Therapy

The main purpose of a control algorithm is to achieve an insulin therapy to keep BG concentration within a predefined range by acting on insulin delivery. There are two classes of control schemes: open- and closed-loop. The Open-Loop (OL) control strategy corresponds to the application of the CT. The control of BG levels in PP period is a challenging task and represents a critical problem due to the glycemic PP peaks. PP glucose regulation is typically based both on the knowledge of an estimation of the ingested carbohydrates and the measurement of the current glucose level, in addiotion with reliable information about the patient according to the CT, which is suited to the patient by the physician. CT administers a basal insulin throughout the day to maintain stable BG levels during fasting periods and insulin boluses at mealtimes to compensate the induced glucose rise in the PP periods [15–17]. Specifically, CT is typically based on the knowledge of an estimation of the ingested carbohydrates, of the CR, of the CF, of the IOB and of a measure of the glycemia just before the meal. Despite the use of this information meal compensation is yet a key unsolved issue. The application of the CT corresponds to apply OL control strategy to control BG levels. In fact, the OL methods design an insulin therapy based only on the knowledge of consumed meals, without considering continuous BG measurements. The OL therapy is designed as a feed-forward action, which exploits the knowledge of external disturbances, namely the meals, to compensate in advance for their effects adapting the insulin bolus doses. Thus, the control of BG levels in PP period become the most critical problem when an OL therapy is adopted because an OL system has no knowledge of the output condition. An incorrect administration of a meal bolus can result in dangerous complications for the patient. In fact, hyperglycemia may occur due to an under-estimated bolus, while an over-estimated bolus could lead to hypoglycemia. Machine learning represents a new methodology recently explored in this type of applications; e.g., in [47], a K-Nearest Neighbors (KNN) methodology has been used to distinguish aerobic and anaerobic exercise metabolism to

improve the control of glucose concentration, or in [48] binary classifiers are trained to recognize the postprandial pattern for meal detection.

Two new approach based on machine-learning methodologies have been proposed in order to improve meal compensation during PP periods. The first methods exploits the KNN classification algorithm to predict PP glucose profile due to the nominal therapy and to suggest a correction to time and/or amount of the meal bolus. The second method uses an individualized regression model able to correct the meal bolus computed with the CT by exploiting measurable variables known at mealtime.

The KNN approach is used to forecast the glucose response to carbohydrate intake. Then, the predictions are exploited to correct the nominal bolus computed via CT, minimizing the occurrence of hypo- and hyperglycemia. A unique classifier valid for the entire adult virtual population of the UVA/Padova simulator has been identified and satisfactory results have been obtained. Of course, an average model could ideally limit the performance since it describes the average dynamics of the population. Therefore, even if it is not individualized to the patient, since it has been identified on the 100 adults of the UVA/Padova simulator, which is able to simulate inter-subject variability, the identified model is supposed to face the inter-subject variability. However, since we are aware that a model that describes the glucose-insulin dynamics of a specific patient can substantially improve the safety and the effectiveness of the glucose control, a personalized algorithm is developed. It is an individualized approach able to correct the meal bolus computed with the CT in order to handle the inter-subject variability characterizing T1D patients that may affect PP glucose regulation. In this case the PP glucose regulation has been managed in a decisional framework, where he decision variable is the correction of the insulin bolus by exploiting a multiple linear regression model able to describe the relation between glucose concentration and injected insulin. Both these approaches are compared with CT glucose profile in realistic scenarios in order to demonstrate that these data-based modeling methodologies are suitable to improve the problem of PP glucose regulation in CT.

4.1 Augmented CT via Meal Classification

A KNN classification algorithm [49, 50] able to predict the PP glucose profile due to the nominal CT, defined in Section 2.3.1, is presented in this section. This approach is exploited to adapt the meal bolus accordingly to the glucose prediction. So, if the classifier predicts a rise in glycemia exceeding the high safe threshold, the insulin bolus amount has to be increased to avoid hyperglycemia, while if the classifier predicts a fall in glucose concentration below the low safe threshold, the bolus amount has to be decreased to avoid hypoglycemia. Since the BG trend

in the PP period is characterized by two features, the amplitude of the excursions and the shape of the profile, a Multiple Classifier System (MCS) is implemented with a parallel architecture, where each of the two base classifiers is devoted to predict one PP feature. The outputs of the two classifiers are combined by an integration strategy to obtain a complete prediction. The description of the MCS architecture and of the base classifiers is given in the following paragraph, together with the training procedure and results in terms of classifier predictions. Then, the proposed augmented therapy based on KNN classifiers is formulated, and the testing scenario with the final results are reported in comparison with the CT.

4.1.1 Multiple Classifier Scheme

The effect of a meal on the glucose response can be characterized by two main PP variables, the amplitude of the maximum excursions and the shapes of glucose profile. The first is a measure of the glucose variability in the PP period and it is strictly related to the minimum and maximum values reached in the PP period; while the second is an information related to the times in which these values are reached. If these PP variables are predictable, the future trend of the glucose response to the meal ca be depicted and the PP glucose response can be regulated by properly modifying the insulin bolus i_B. Note that the description given by these variables is not exhaustive, but it is sufficient to define the proper bolus changes. The first aim is to identify a relation between measurable quantities, called input features, given at mealtime and the two PP variables characterizing the future meal response. An example of glucose trend after a meal is presented in Figure 4.1, where the parameters characterizing the curve are:

- G_{min} and G_{max}, the minimum and maximum values of glucose, respectively
- t_{min} and t_{max}, the time instants when G_{min} and G_{max} values are reached, respectively
- G_m, the glucose at mealtime t_m.

Since G_{min} and G_{max} define completely the excursion, while the shape is more related to t_{min} and t_{max}, the two PP variables can be considered separately by identifying two independent base classifiers, C_e and C_s, respectively. The PP variables result mutually complementary, so a MCS is proposed where each predictable variable has its own classifier algorithm. Given a single set of input features known at t_m, each base classifier will select its own subset for the classification.

The outputs of two base classifiers are combined following a parallel MCS architecture. The goal of this scheme is the improvement of the accuracy of the final prediction by exploiting

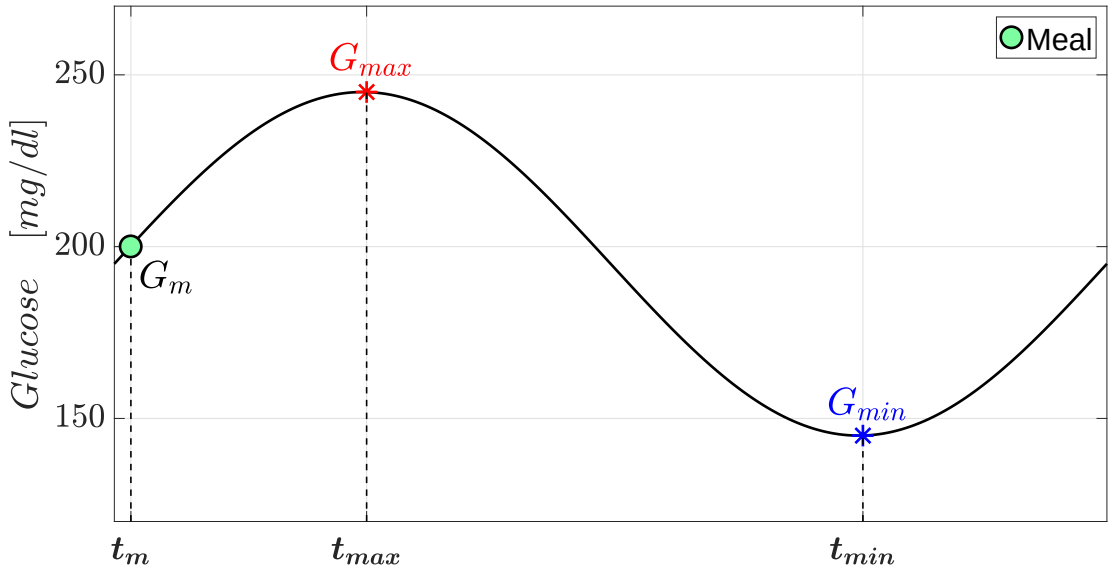


FIGURE 4.1: An example of glucose trend where G_m, G_{min}, G_{max} are the glucose value at mealtime, the minimum value of glucose, and the maximum value of glucose, respectively and t_m,t_{min}, t_{max} are mealtime, the time instants when G_{min} and G_{max} values are reached, respectively.

the classifier diversity. Indeed, since the classifiers make different misclassification errors on different test samples, diversity improves the classification accuracy. Of course, the drawback is the need to train multiple classifiers. The combination strategy of the two base classifiers is performed by integration, namely both the classifiers contribute to the final output [51].

4.1.2 Excursion Classifier (C_e) and Shape Classifier (C_s)

In order to define the first classifier, the combinations of G_{max} and G_{min} have to be associated to a finite set of classes. Hence, five classes are defined on the basis of proper glucose thresholds:

- C_e^{-2}: glucose trend characterized by hypoglycemia
- C_e^{-1}: glucose trend associated with hypoglycemia risk
- C_e^{0}: glucose trend in the desired target
- C_e^{+1}: glucose trend associated with hyperglycemia risk
- C_e^{+2}: glucose trend characterized by hyperglycemia.

$\mathbf{C_e}$	Description
$\mathbf{C_e^{-2}}$	$G_{max} < G_{high}$ and $G_{min} < G_{hypo}$
$\mathbf{C_e^{-1}}$	$G_{max} < G_{hyper}$ and $G_{hypo} < G_{min} < G_{low}$
$\mathbf{C_e^{0}}$	$G_{max} < G_{high}$ and $G_{min} < G_{low}$
$\mathbf{C_e^{+1}}$	$G_{max} > G_{high}$ and $G_{low} < G_{min} < G_{high}$
$\mathbf{C_e^{+2}}$	$G_{max} > G_{hyper}$ and $G_{min} > G_{high}$

TABLE 4.1: Excursion classification.

$\mathbf{C_s}$	Description
$\mathbf{C_s^{-2}}$	$t_u < t_o$ and t_o, $t_u > 0$
$\mathbf{C_s^{-1}}$	t_o *n.d.* and $t_u > 0$
$\mathbf{C_s^{0}}$	t_u, t_o *n.d.*
$\mathbf{C_s^{+1}}$	t_u*n.d.* and $t_o > 0$
$\mathbf{C_s^{+2}}$	$t_o < t_u$ and t_o, $t_u > 0$

TABLE 4.2: Shape classification.

Denoting with G_{high}, G_{low} the upper and lower bounds of the desired glucose range, and with G_{hyper}, G_{hypo} the glucose limits that defines hyperglycemia and hypoglycemia events, the excursion classification is described in Figure 4.2(a) and summarized in Table 4.1.

The second classifier aims to distinguish the meal responses on the base of glucose shape and in particular on the presence of significant undershoot and/or overshoot. Denoting a significant undershoot as

$$G_{min} < G_m - \Delta G_{th} \tag{4.1}$$

and a significant overshoot as

$$G_{max} > G_m + \Delta G_{th} \tag{4.2}$$

with ΔG_{th} a glucose threshold to be tuned, the shape classifier maps the input features to five categories of possible PP glucose shapes defined as follows:

- C_s^{-2}: meal response characterized by a significant undershoot at time t_u followed by a significant overshoot at time t_o
- C_s^{-1}: meal response characterized by a significant overshoot at time t_o
- C_s^{0}: meal response without significant undershoot and overshoot
- C_s^{+1}: meal response characterized by a significant undershoot at time t_u
- C_s^{+2}: meal response characterized by a significant overshoot at time t_o followed by a significant undershoot at time t_u.

If undershoot (overshoot) are not present, t_u (t_o) are not defined (n.d.). The shape classes are described in Figure 4.2(b) and summarized in Table 4.2. In order to make the classifiers more conservative in detecting events of hyper- and hypoglycemia, the threshold ΔG_{th} can be defined

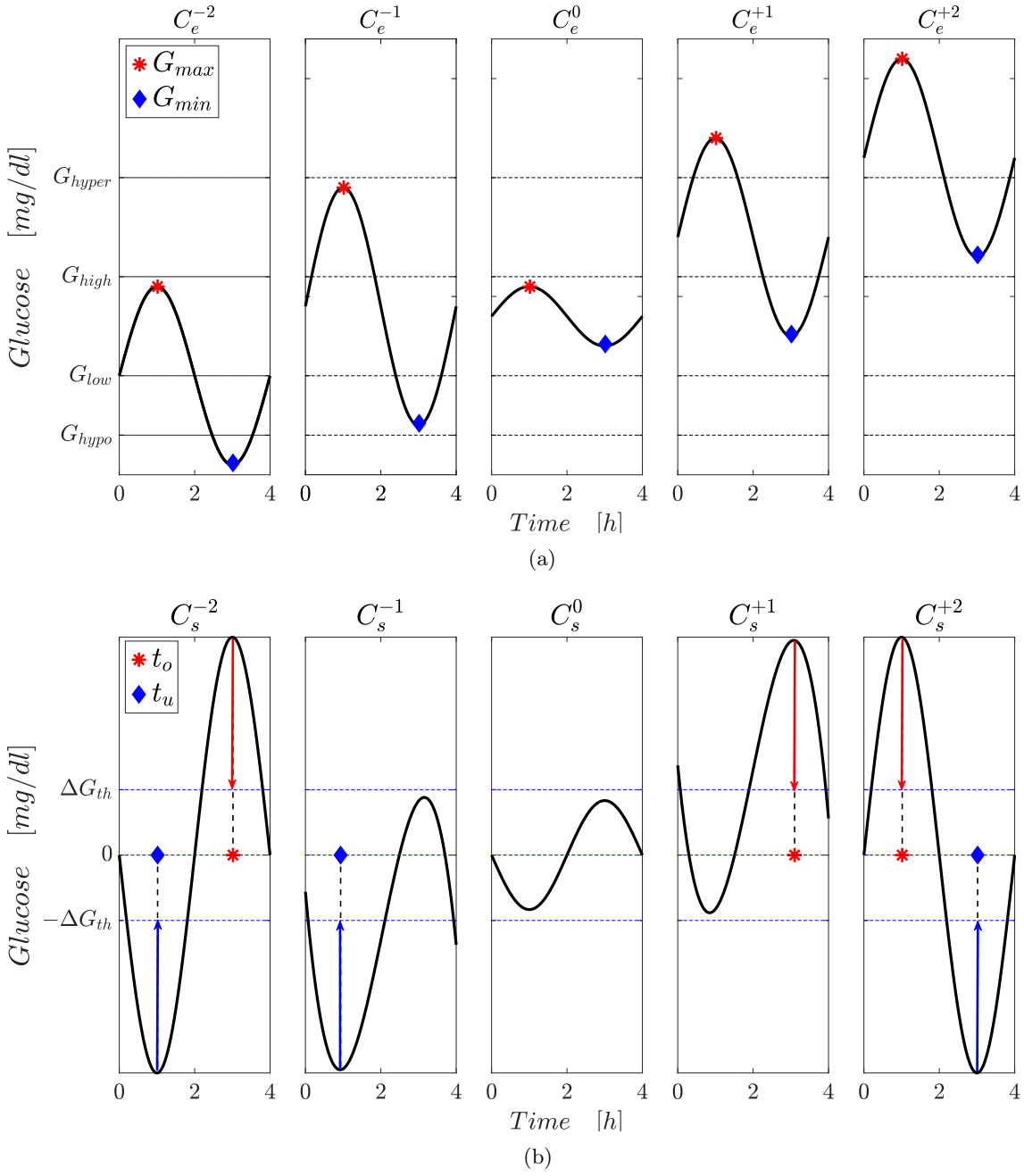


FIGURE 4.2: (a) The five classes of the *Excursion Classifier* C_e. The red stars represent G_{max}, and the blue diamonds represent G_{min}. (b)The five classes of the *Shape Classifier* C_s. The red stars represent t_o, and the blue diamonds represent t_u. The red lines represent a significant overshoot, and the blue lines represents a significant undershoot.

in a different way for overshoot and undershoot.

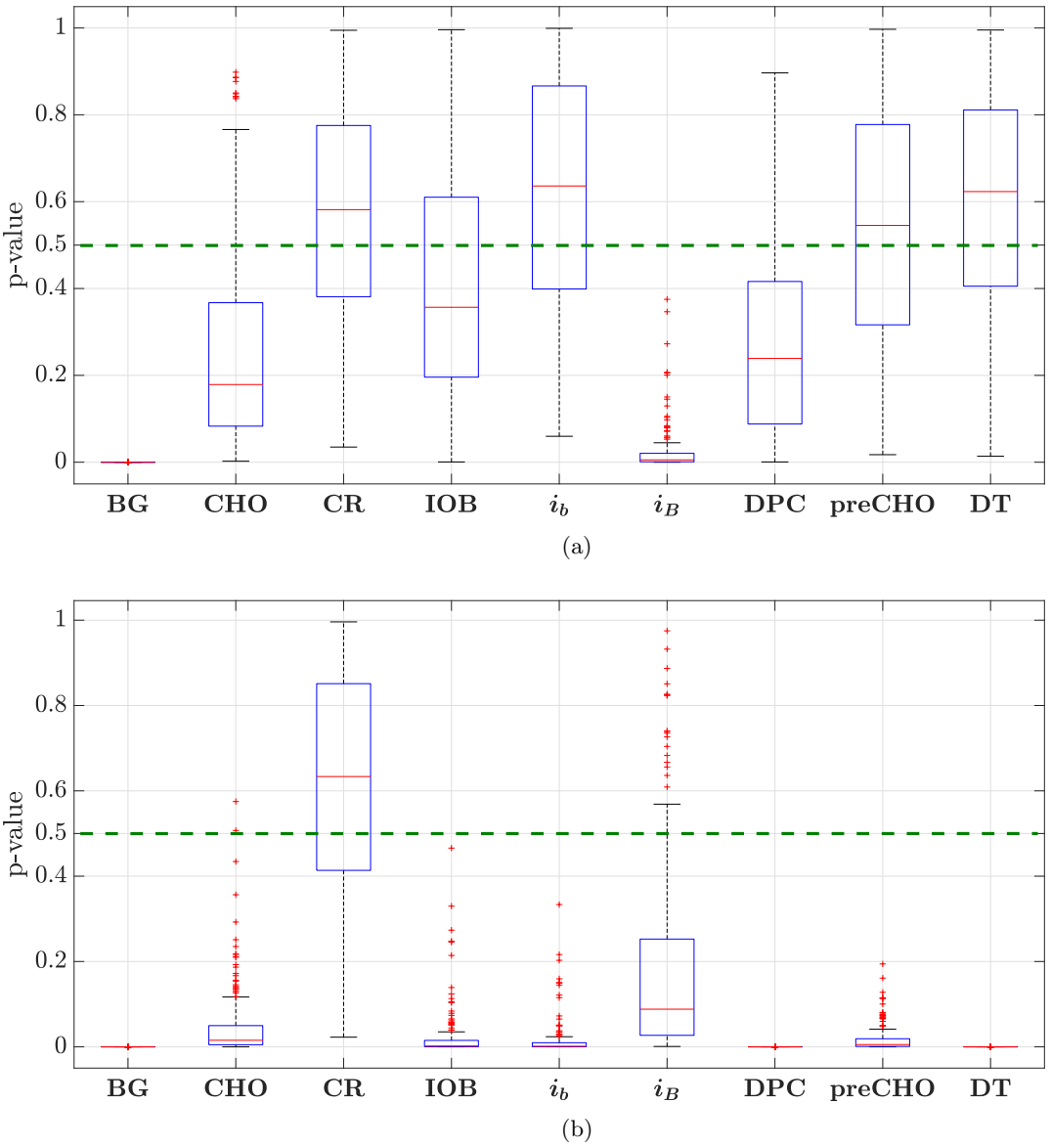


FIGURE 4.3: (a) Boxplots associated to ANOVA test of C_e. (b) Boxplots associated to ANOVA test of C_s. The green dashed lines represent the p-value theshold set to 0.05.

4.1.3 KNN Classifier

The two developed classifiers use a KNN algorithm to perform the classification. KNN algorithms are supervised non-parametric learning algorithms that learn the relationship between input and output observations and whose strenght is making very mild structural assumptions. A new input instance x^* is classified by assigning the output class of the K most similar neighbors, where similarity is defined according to a distance metric [49]. Specifically, given a set of measurements (x_i, y_i), $i = 1, \cdots, N$, known as the training data, the KNN fit is defined as

follows:

$$\hat{Y}(x^*) = \frac{1}{N_n} \sum_{x_i \in N_{N_n}(x)} y_i \tag{4.3}$$

where $N_{N_n}(x)$ is the neighborhood of x defined by the N_n closest points x_i in the training sample. In other words, the algorithm find the k observations with x_i closest to x in input space, and average their responses (y_i). KNN fits have a single parameter, the number of neighbors k: this parameter specifies how many closest points the algorithm has to use to classify the output of a new point. KNN algorithm is highly efficient for pattern recognition [50, 52] and well fits the purpose of glucose patterns classification. In order to train the classifiers, a single space set of possible input features has been defined for both the classifiers. These inputs features are selected from a set of parameters known at t_m consisting of BG, carbohydrate content (CHO), CR, IOB, i_b, i_B, amount of carbohydrate intake calculated in the interval $[t_m - 6h, t_m]$ (preCHO), Day Time (DT) with respect to midnight and the Day Period Classification (DPC). The DPC is a categorical variable defined as: morning [5:00 - 12:00], afternoon [12:00 - 19:00] and evening [19:00 - 5:00]. For each classifier, the selection of the input features is performed via an ANOVA test to determine the parameters subset statistically correlated with the considered classifier, where a statistical correlation is considered significant at the level 0.05. Other factors, like stress, physical activity, etc., not usually available at mealtime have been excluded from the analysis. Interaction effects of the input features in the ANOVA have not been considered: combined effects of inputs may loss the physical meaning in the considered application. The results of the two ANOVA tests are shown in Figure 4.3(a) and 4.3(b). Figure 4.3(a) shows the boxplots related to C_e: the input features statistically correlated with the classes of C_e are BG and i_B. The boxplot associated to BG is entirely below the adopted significance threshold (p-value ¡ 0.05); i_B presents a few outliers exceeding the threshold, but the performance are significantly better than the other parameters. Figure 4.3(b) shows the boxplots related to C_s: BG, DCP, and DT are chosen for C_s because the boxplots are entirely below the threshold, set equal to 0.05.

In addition to the input features, two other elements characterize the classifier: the dataset used in the training procedure, called training set, and the number of neighbors, N_n. Since the drawback of the KNN method is the need for storing the whole training set, which may be large, the Condensed Nearest Neighbor (CNN) rule [53] is used in this work to optimize the training set by selecting a subset of the samples contained in the original set. CNN is an iterative procedure which allows to reduce the training set by minimizing data redundancy. However, the training data can be affected by a sample error which can cause a decrease in performance; also N_n affects the prediction performance of the classifier. So, the training procedure is repeated an exhaustive number of times N_r with different training sets and different N_n. The selection

of the classifier aims to maximize the probability of correctly predicting hypoglycemia. Then, each classifier prediction is associated to a posteriori probability of success, called score, and the standard KNN algorithm selects the one with the maximum score. Considering that a misclassification of the classes $C_x^{-2}, C_x^{+2}, x \in \{e, s\}$ represents a risk of hyper or hypoglycemia, respectively, in this work the classifier C_x is made more conservative by exploiting this score. Let's consider the critical case of a C_e^{-2} prediction, the classifier C_e maintains this prediction only if the difference between the score of C_e^{-2} and C_e^{-1} is above a safety threshold ϵ, otherwise the C_e prediction is changed to C_e^{-1} in order to apply a more conservative action. The same approach is used with $C_e^{+2}, C_s^{-2}, C_s^{+2}$ prediction, while the other predictions remain unchanged. In addition to the training set, a validation set is needed to assess the classifier performance. In order to avoid the overfitting problem, training and validation sets are constrained to be independent, but with the same probability distribution.

4.1.4 MCS training

The training and validation sets are generated by using the 100 *in silico* adult patients of the UVA/Padova simulator with a 4-day scenario. These classifiers are trained on a large number of meals belonging to different patients, so a unique KNN-model is proposed for all the T1D population. This scenario includes 18 meals: 4 breakfasts (between 7:00 and 8:30), 6 snacks (2 in the morning, 3 in the afternoon and 1 in the evening), 4 lunches (between 12:30 and 14:00) and 4 dinners (between 20:00 and 20:30). The total amount of 1800 meal data over the 100 patients are split in two parts: half pertaining to the training set and half to the validation set. In order to avoid overfitting or sample error, the number of repetition of the procedure N_r is set to 200.
The algorithm parameters are set to $\Delta G = 30$ mg/dl, $G_{high} = 150$ mg/dl, $G_{low} = 100$ mg/dl, $G_{hyper} = 200$ mg/dl, $G_{hypo} = 70$ mg/dl, $\epsilon = 0.20$.

4.1.5 MCS validation

The validation of the classifiers has been performed in term of area under the curve that describe the relation between false positive and true positive classification outputs. Denoting with C_x^n a generic class of a generic classifier C_x, some preliminary definitions need to be introduced below:

- True Positive (TP): sample belonging to class C_x^n classified correctly
- True Negative (TN): sample not belonging to class C_x^n classified correctly

- False Positive (FP): sample not belonging to class C_x^n but classified as belonged to C_x^n
- False Negative (FN): sample belonging to class C_x^n but classified as not belonged to C_x^n
- True Positive Rate (TPR) or sensitivity and False Positive Rate (FPR) or specificity measure the proportion of positives that are correctly classified and positives that are wrongly classified, respectively and they are defined as follows:

$$TPR = \frac{TP}{TP + FN}, \qquad FPR = \frac{FP}{TN + FP}.$$

Defining the Receiver Operator Characteristic (ROC) as the curve describing the relation between TPR and FPR, the perfomance of the classifiers are evaluated through the Area Under the curve of ROC (AUROC) described in [54]. In other words, the ROC curve is a commonly used summary for assessing the trade off between sensitivity and $1-$ specificity of the trained classifiers. The classifier allows both the minimization of FPR and the maximization of TPR by increasing the sensitivity and specificity, respectively. Classifiers have a good performance with AUROC value more than 0.5, while AUROC equal to 0.5 means that the classifier behaves exactly as a random variable.
The AUROC values of the selected C_e and C_s are reported in Figures 4.4(a) and 4.4(b), respectively. This index presents satisfactory performance with values between 0.60 and 0.90 for all the classes of both classifiers apart for C_s^{-2}. The critical AUROC value associated to this class can be explained by the poor representation of C_s^{-2}. This limitation is acceptable, considering that the event occurs occasionally during clinical trials and that the simulated dataset distribution has been choosen to be homogeneous with the real data.

4.1.6 Classifier augmented Open-Loop

The availability of a classifier able to predict the PP glucose dynamics at the mealtime allows to modify in advance the OL therapy in order to avoid both hypoglycemia and hyperglycemia. In general, the purpose of the MCS is to classify the PP glucose dynamic: this classification allows to distinguish if the meal is be well compensated or under/over-insulinized. Thus, the CT can be improved by exploiting the prediction of MCS at mealtime k_m. Given an estimation of postprandial glucose profile at mealtime, a condition of hyperglycemia can be reduced by increasing the meal bolus, while a condition of hypoglycemia can be prevented, or at least mitigated, by decreasing it. If both conditions occur sequentially, the insulin bolus has to be properly adapted on the basis of the time and the severity of each condition. If one single phenomenon is predominant, the bolus is modified only in terms of insulin amount, otherwise

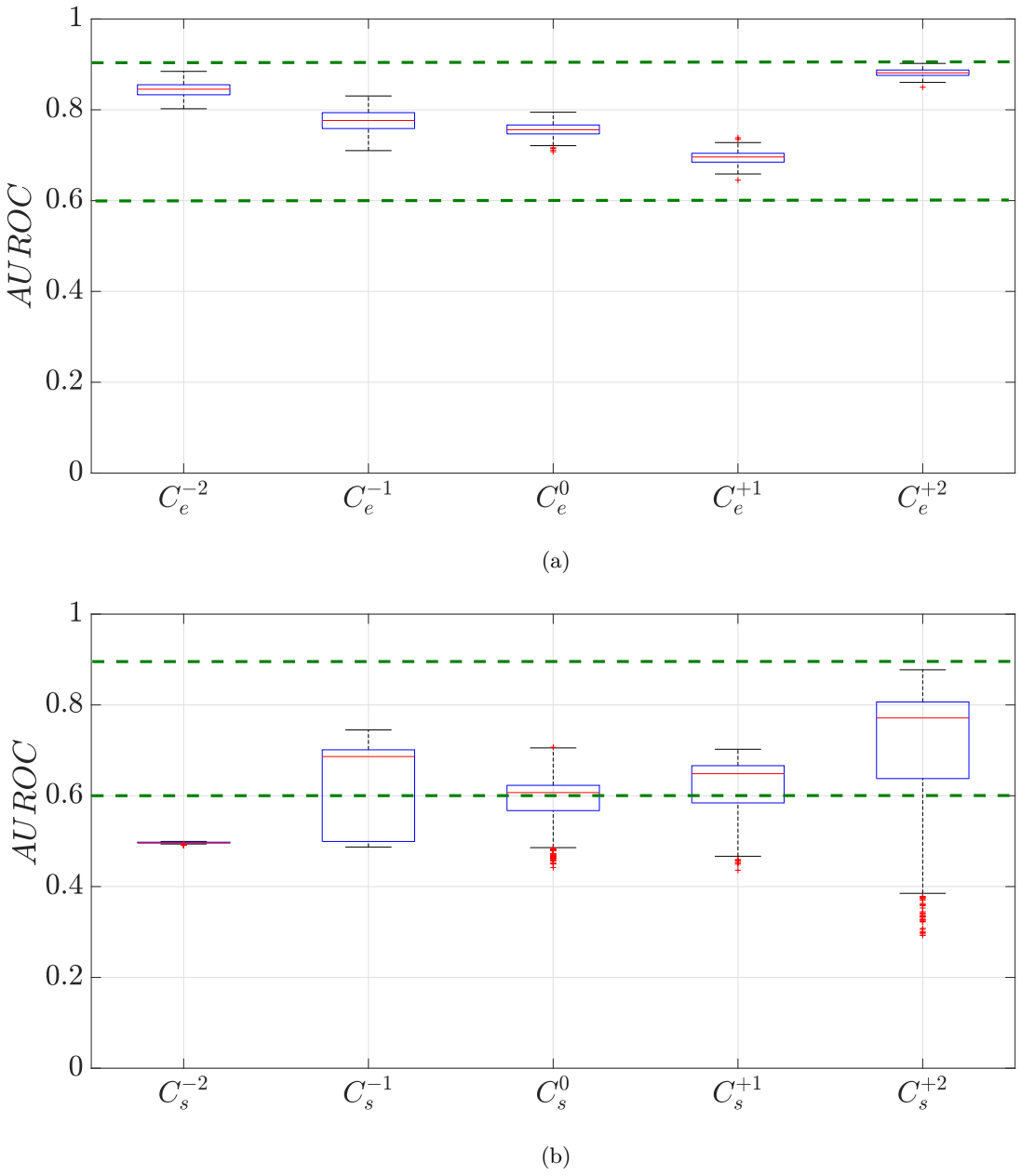


FIGURE 4.4: (a) Boxplots associated to AUROC test of C_e. (b) Boxplots associated to AUROC test of C_s. The green dashed lines represent the boundaries between 0.6 and 0.9.

both the time and the amount of the bolus have to be adapted. Following these criteria, a Classifier-Augmented (CA) OL therapy is proposed that exploits the information of both classifiers, trained offline, to modify the meal bolus, i_B. In particular, the classifier C_s defines the variation of the time, while C_e defines the range of the insulin variations. Then, the final adaptation of the bolus is the result of the integration of C_s and C_e.

Defining C_{sm} and C_{em} the classifier prediction of C_s and C_e at mealtime t_m, respectively, the

$\mathbf{C_e}$ \ $\mathbf{C_s}$	$\mathbf{C_s^{-2}}$	$\mathbf{C_s^{-1}}$	$\mathbf{C_s^{0}}$	$\mathbf{C_s^{+1}}$	$\mathbf{C_s^{+2}}$
$\mathbf{C_e^{-2}}$	α=0.2, β=0.3	α=0.2, β=0,	α=0.4, β=0,	α=0.8, β=0,	α=0.3, β=0.2
$\mathbf{C_e^{-1}}$	α=0.5, β=0.5	α=0.6, β=0,	α=0.8, β=0,	α=0.7, β=0,	α=0.5, β=0.5
$\mathbf{C_e^{0}}$	α=0.8, β=0	α=0.8, β=0,	α=1, β=0,	α=1.2, β=0,	α=1.2, β=0
$\mathbf{C_e^{+1}}$	α=0.5, β=0.8	α=1, β=0,	α=1.2, β=0,	α=1.4, β=0,	α=1.2, β=0.2
$\mathbf{C_e^{+2}}$	α=0.7, β=1	α=1.4, β=0,	α=1.4, β=0,	α=1.7, β=0,	α=1.4, β=0.6

TABLE 4.3: α and β values for patients of Category A.

insulin bolus adapted by the CA strategy, i_B^{CA}, can be defined as follows:

$$i_B{}^{CA}(t_m) = \alpha\, i_B(t_m) + \beta\, i_B(t_m - \tau) \tag{4.4}$$

where the values of α, β used for each bolus depend on C_{sm} and C_{em}. In particular, a first set of values for each possible combination of these classifiers has been defined as reported in Table 4.3 (Category A). A second more conservative category (Category B) has been defined in order to manage patients very sensitive to drastic change of their therapy, see Table 4.4. Finally, a third category (Category C) has been defined by setting $\alpha = 1$ and $\beta = 0$, in order to manage critical situations. The values of α and β have been tuned on a subset of meals taken from different patients able to represent the 25 possible combinations of classes reported in tables 4.3 and 4.4. The assignment of each patient to one category is performed by a trial and error procedure. In particular, each patient undergoes to a 16-day scenario composed of 4 periods of 4 days: the first in OL, the second with Category A and the fourth with Category B, divided by a washout period in OL. The introduction of hypotreatments in the protocol guarantees patient safety. Patients are assigned to Category C if both Categories A and B increase significantly the number of hypotreatments or no improvements of T_a (time above 180 mg/dl) and T_b (time below 70 mg/dl) are achieved. The assignment to Category A and B is performed minimizing both T_a and T_b or alternatively only one of them, giving priority to T_b. If the improvements between the two categories are similar, the category A is preferred. The parameter $\tau = 20$ minutes is set constant for all the patients. Since the classifier training is performed offline, the computational time required for the online implementation of the algorithm is negligible.

$\mathbf{C_e}$ \ $\mathbf{C_s}$	$\mathbf{C_s^{-2}}$	$\mathbf{C_s^{-1}}$	$\mathbf{C_s^{0}}$	$\mathbf{C_s^{+1}}$	$\mathbf{C_s^{+2}}$
$\mathbf{C_e^{-2}}$	α=0.3, β=0.2	α=0.3, β=0,	α=0.5, β=0,	α=0.8, β=0,	α=0.4, β=0.1
$\mathbf{C_e^{-1}}$	α=0.6, β=0.4	α=0.8, β=0,	α=1, β=0,	α=0.9, β=0,	α=0.8, β=0.2
$\mathbf{C_e^{0}}$	α=0.8, β=0	α=0.9, β=0,	α=1, β=0,	α=1.1, β=0,	α=1.2, β=0
$\mathbf{C_e^{+1}}$	α=0.8, β=0.4	α=1, β=0,	α=1, β=0,	α=1.2, β=0,	α=1.1, β=0.2
$\mathbf{C_e^{+2}}$	α=0.9, β=0.8	α=1.2, β=0,	α=1.2, β=0,	α=1.4, β=0,	α=1.2, β=0.2

TABLE 4.4: α and β values for patients of Category B.

4.1.7 Simulation settings

The proposed CA algorithm is tested on the 100 *in silico* adult patients of the UVA/Padova simulator with a 4-day scenario. The testing scenario starts at 0:00 am and involves 16 meals: 4 breakfasts (between 7:00 and 8:00), 4 snacks (1 in the morning, 2 in the afternoon and 1 in the evening), 4 lunches (between 12:30 and 13:00) and 4 dinners (between 19:00 and 20:30). Hypotreatments (ht) of 15g are administrated to the patient in case of hypoglycemia (BG $<$ 65 mg/dl). In order to evaluate the effectiveness of an insulin therapy, the performance metrics introduced in Section 3.5 have been evaluated. These metrics follow the consensus statement for artificial pancreas trials described in [42].

4.1.8 Results

The performance metrics obtained with the OL and CA strategies evaluated on the entire population on the testing scenario are reported in Table 4.5. Interesting improvements are obtained by the new technique for all the 100 patients in the PP period, period mainly affected by the meal bolus changes. The CA is able to improve the time in target by 2.1% and in tight target by 11.6%, to lower the average glucose by 5.4% and to reduce the time spent above the target by 36.7% with respect the OL therapy. The increase of the times spent below the target remains limited: T_b initially 2.9% reaches 3.2% while T_{b60} from 1.5% increases to 1.8%. The RMSE is decreased by 5.4% with and increment of the TDI of 2.9 U. All these results are statistically significant. The same observations can be drawn overall, where the CA is able to improve the times in target (5% and 10.3% for the tight target), to lower the average glucose by

		O	N	PP
A [mg/dl]	OL	**144.57 [134.86, 155.84] [142.38,148.63]**	**148.59 [137.59, 168.59] [148.49,157.77]**	**145.73 [134.84, 154.13] [142.20,148.78]**
	CA	**139.05**^{a} **[132.11, 148.43] [137.98,143.29]**	**144.76**^{a} **[133.95, 161.91] [142.75,151.62]**	**137.82**^{a} **[131.45, 148.92] [137.01,142.72]**
SD [mg/dl]	OL	**34.46 (± 12.54) [31.97,36.95]**	**19.06 [13.59, 26.83] [18.25,22.20]**	**35.08 [24.68, 41.42] [31.52,36.71]**
	CA	**35.80**^{a} **(± 12.64) [33.29,38.30]**	**22.58**^{a} **[15.75, 30.31] [20.87,25.00]**	**35.07**^{c} **[25.80, 42.70] [32.26,37.32]**
T_t [%]	OL	**80.61 [65.47, 92.74] [75.41,83.55]**	**89.00 [64.38, 100.00] [77.51,88.30]**	**80.80 [65.82, 92.57] [74.26,82.61]**
	CA	**84.64**^{a} **[68.61, 94.57] [77.78,85.46]**	**93.11**^{a} **[74.62, 100.00] [82.03,92.72]**	**82.47**^{a} **[68.86, 94.84] [77.18,85.02]**
T_{tt} [%]	OL	**38.84 (± 18.43) [35.19,42.50]**	**31.58 [11.94, 53.13] [27.54,39.28]**	**38.75 (± 18.33) [35.12,42.39]**
	CA	**42.83**^{a} **(± 17.09) [39.44,46.22]**	**41.78**^{a} **[21.27, 57.76] [35.13,45.79]**	**43.25**^{a} **(± 17.61) [39.75,46.74]**
T_a [%]	OL	**13.31 [4.30, 28.94] [12.91,20.48]**	**9.16 [0.00, 35.62] [11.63,22.25]**	**14.39 [4.86, 27.90] [13.51,21.17]**
	CA	**10.53**^{a} **[2.77, 23.54] [10.51,16.77]**	**4.67**^{a} **[0.00, 24.79] [5.54,17.51]**	**9.11**^{a} **[3.66, 23.69] [10.81,17.38]**
T_b [%]	OL	**2.92 [0.00, 5.53] [2.57,3.98]**	**0.00 [0.00, 0.00] [0.00,0.00]**	**2.87 [0.00, 5.57] [2.22,4.08]**
	CA	**3.64**^{a} **[1.41, 6.72] [3.33,4.84]**	**0.00**^{a} **[0.00, 0.00] [0.00,0.00]**	**3.17**^{a} **[0.87, 6.99] [3.07,4.83]**
T_{b60} [%]	OL	**1.81 [0.00, 3.45] [1.54,2.66]**	**0.00 [0.00, 0.00] [0.00,0.00]**	**1.48 [0.00, 3.82] [1.50,2.69]**
	CA	**2.37**^{a} **[0.00, 4.77] [2.13,3.43]**	**0.00**^{b} **[0.00, 0.00] [0.00,0.00]**	**1.82**^{a} **[0.00, 5.24] [2.08,3.52]**
RMSE	OL	**42.64 [31.94, 56.21] [40.52,47.44]**	**36.15 [27.01, 56.05] [37.25,45.62]**	**42.50 [34.09, 54.17] [40.94,47.79]**
	CA	**40.37**^{a} **[30.71, 53.89] [38.81,45.55]**	**34.79**^{a} **[26.08, 50.66] [34.03,41.96]**	**40.21**^{a} **[30.63, 52.73] [38.78,45.47]**
TDI [U]	OL	**53.76 [44.45, 68.83] [52.21,59.03]**	**8.65 [7.21, 10.30] [8.36,9.35]**	**38.74 [31.49, 48.45] [37.58,42.49]**
	CA	**55.87**^{a} **[45.40, 69.57] [54.86,62.16]**	**8.75**^{a} **[7.64, 10.76] [8.72,9.74]**	**39.86**^{a} **[32.46, 50.74] [39.45,44.74]**

TABLE 4.5: OL strategy vs new CA approach on 100 in silico patients. O, N, and PP stand for Overall, Night, and Postprandial period, respectively. Confidence intervals are reported. The significant results are highlighted in bold. p-value (p) significance levels are: $a := p < 0.001$, $b := p < 0.01$, $c := p < 0.05$

3.8% and to reduce the time spent above the target by 20.9%. The increase of the time spent below the target remains limited. The RMSE is decrease by 5.3% with an increment of TDI of 3.9 U. All these results are statistically significant and have been published in [55].

4.2 Individualized CT

In order to improve PP glucose regulation, the insulin bolus computed by CT can required a variation due to the amount. Unlike the approach presented in Section 4.1, where the correction of the insulin bolus is based on the prediction of the PP glucose levels, in this section the goal is to identify the parameters that can directly predict the effect of the insulin bolus on glycemia. In this section a data-driven modeling approach able to predict the optimal percentage to be applied to the insulin bolus is proposed. *In silico* data are used for model identification with the UVA/Padova simulator [34]. Since this version of the UVA/Padova simulator includes intra- and inter-day variability of insulin sensitivity, dawn phenomenon effect, and different distributions of CR at different day time [18], the proposed approach aims to identify specific models for different day periods on the basis of the intra-day variability of insulin sensitivity. The regression model capabilities are evaluated on three robustness testing scenarios: a nominal scenario and two perturbed scenarios. The first perturbed scenario includes random variations of meal amounts to demonstrate the robustness of the individualized model against under- or over-estimation errors of carbohydrate intakes. The second more challenging scenario is designed by adding to the first scenario a random ±15% variation of the nominal insulin sensitivity to test the proposed approach against disturbances, such as physical activity or stress which both affect insulin sensitivity.

4.2.1 Modified Conventional Therapy

This approach proposes an individualized correction of the nominal insulin bolus, which is adapted as follows:

$$i_B^{\alpha} = i_B \cdot \alpha \tag{4.5}$$

where α is the insulin percent variation at mealtime and it is the output of a Multiple Linear Regression (MLR) model [56] in the following form:

$$\begin{aligned} \alpha &= f(X_R, k_m) = \theta^I(k_m) + \theta^{II}(k_m)X_R + \theta^{III}(k_m)X_R^2 \\ s.t.\ \alpha &\in I_\alpha = \left[\alpha^{MIN}, \alpha^{MAX}\right] \end{aligned} \tag{4.6}$$

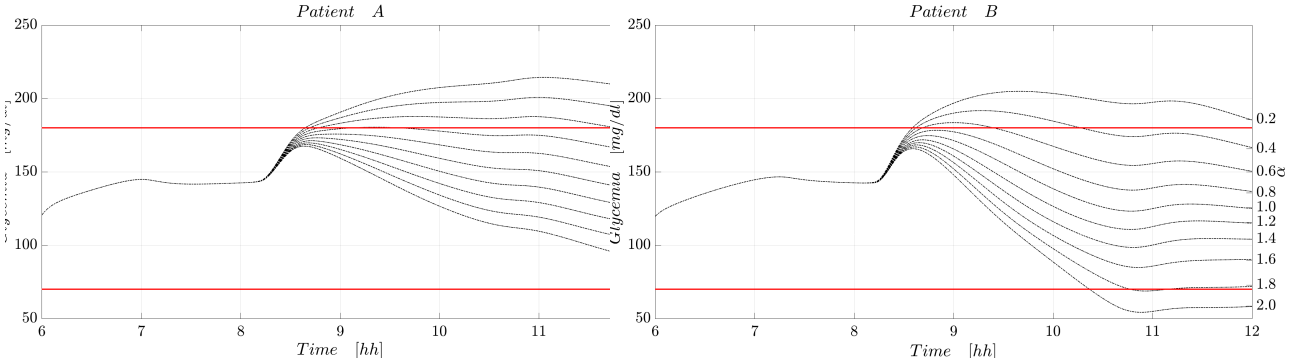


FIGURE 4.5: PP glucose profiles of two patients belonging to the adult virtual population of the UVA/Padova simulator. The profiles are obtained by correcting i_B with $\alpha = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ in order to compensate a 40 g meal taken at 8.00.

where X_R, $X_R^2 \in \mathbb{R}^N$ represent the vector of the regressors and their values squared, respectively, θ^I, θ^{II}, $\theta^{III} \in \mathbb{R}^N$ are the model parameters at mealtime k_m, I_α is the set of admissible values for α, and α^{MIN}, α^{MAX} are the lower and upper bounds of I_α, respectively. The constraint on the feasible values of insulin percent variation is imposed for safety reasons. Hence, the modified insulin bolus is bounded within the interval $I_B^\alpha = \left[i_B^{\alpha^{MIN}}, i_B^{\alpha^{MAX}}\right]$, where $i_B^{\alpha^{MIN}}$, $i_B^{\alpha^{MAX}}$ are the lower and upper bounds of i_B^α at mealtime, respectively.

The aim of MLR model is to capture the interaction between glucose and insulin for a given meal amount. Thus, the regressors are glucose-related variables and the structure of the model has to approximate the interaction between glucose and insulin. As shown in [57], a linear approximation of glucose-insulin relation does not capture the relationship between the prandial glucose profile and optimal insulin bolus. Hence, a quadratic (squared) term X_R^2 is included in the model to add flexibility to the relationship between the regressors and the dependent variable. Thus, the goal here is to choose a set of suitable regressors X_R and define a procedure to identify the parameters θ^I, θ^{II}, θ^{III}.

4.2.2 Model Customization

To improve PP glucose control, the development of an individualized model which fits the individual characteristics of each subject is necessary. Given the same initial conditions, the corrections of the bolus dose may affect differently PP glucose profile of different patients, as shown in Figure 4.5. Thus, the parameters of the model defined in Eq. 4.6 are patient-dependent and we utilize the model below to predict the proper dose percent variation for a specific Patient i:

$$\alpha_i = f_i(X_R, k_m) = \theta_i^I(k_m) + \theta_i^{II}(k_m)X_R + \theta_i^{III}(k_m)X_R^2. \tag{4.7}$$

Since PP glucose excusions have different patterns at breakfast, lunch and dinner, the parameters θ^I, θ^{II}, and θ^{III} are defined as a piecewise function with respect to day period. The definition of the day periods is based on CR is a time variant and is influenced by insulin sensitivity. Each patient of the virtual population of the UVA/Padova simulator is equipped with its own CR daily pattern optimized on the basis of its own time-varying insulin sensitivity [18]. Thus, the definition of the day period is provided by the time intervals where CR remains constant and this information is already individualized. Hence, the model is of the following structure:

$$\alpha_i = f_i(X_R, T_t) = \theta_i^I(T_t) + \theta_i^{II}(T_t)X_R + \theta_i^{III}(T_t)X_R^2, \qquad t = 1, ..., T \tag{4.8}$$

where T_t is the t^{th} interval and T is the number of day periods related to Patient i. For instance, if the CR has three different day periods (usually one for breakfast, one for lunch, and one for dinner), three models will be identified with three different set of parameters θ^I, θ^{II}, and θ^{III}, one set for each period.

4.3 Model Identification

The aim is to identify an individualized model, which is able to predict the optimal percent variation of i_B at mealtime k_m to improve PP glucose profile by exploiting measurable variables known at k_m. The main idea is to compute the optimal value of α for different values of these measurable variables in order to span all the possible space of combinations. Then, the model will be identified in order to better fit the space of optimal α values. So, the identification of the model of Eq. 4.8, is performed in two steps.

The first step is to estimate the optimal percent variation α^O of the nominal bolus in different conditions. Since the computation of $i_B(k_m)$ is mainly affected by the contribution of $\widehat{CHO}(k_m)$, $BG(k_m)$, and $IOB(k_m)$, as described in Eq. 2.3, we investigated how these variables may affect α^O. In other words, different initialization of $\widehat{CHO}$, BG, IOB, and k_m are considered to define a set of optimal insulin variations (I_α^O).

In the second step, once several α^O are collected for different conditions, the parameters of the model predictors are identified by least squares.

These two steps are performed for each day period T_t and different models for different day periods are identified. Since the data significantly affect the model performance, an experimental design is set up to collect data to identify MLR [58]. The experiments are designed to produce a sufficient excitation on the overall system and are defined by analyzing real data of T1D patients collected during clinical trials [59]. Each experiment is characterized by a tuple whose elements

are the following measurements: $\widehat{CHO}$, BG, IOB, k_m, Initial Interstitial Glucose (IIG), and Initial Interstitial Glucose Slope ($IIGS$). IIG and $IIGS$ are the glucose measurement at k_m, and the pre-prandial glucose variation in the time interval $[k_m - T_{GS}, k_m]$, where T_{GS} [min] represents the length of the pre-prandial time interval, respectively. Denoting with N_{I_α} the number of experiments, a matrix X of N_{I_α} lines is created:

$$X = \left[\begin{array}{ccccc} \widehat{CHO}_0 & BG_0 & IOB_0 & K_{m0} & IIG_0 \quad IIGS_0 \end{array} \right] \tag{4.9}$$

where $\widehat{CHO}_0$, BG_0, IOB_0, K_{m0}, IIG_0, $IIGS_0 \in \mathbb{R}^{N_{I_\alpha}}$ are the vectors of the collected measurements for all the experiments. So, each line of X, x_0, is obtained by performing a single experiment which captures the system response at different operating conditions.

4.3.1 Step 1: Definition of Optimal Insulin Variation

Since different models are identified for each day period T_t, the model identification procedure is presented for a generic time interval T_t and it will be expanded in the following sections to all the intervals. The aim of step 1 is to compute the values of α^O for every line of X. Since only a subset of information is used for this scope, X_α is introduced as follows:

$$X_\alpha = \left[\begin{array}{ccc} \widehat{CHO}_0 & BG_0 & IOB_0 \end{array} \right] \tag{4.10}$$

and let x_α denote a generic line of X_α:

$$x_\alpha = \left[\begin{array}{ccc} \widehat{CHO}_0 & bg_0 & iob_0 \end{array} \right] \tag{4.11}$$

where $\widehat{CHO}_0$, bg_0, $iob_0 \in \mathbb{R}$ indicate the measurements collected in a single experiment performed in the considered interval. Given a specific experiment, consider the following metrics of glucose variability: $T^{pp}_{Th}(x_\alpha, \alpha)$ denotes the percentage of PP time spent below the threshold Th [mg/dl] and $A^{pp}(x_\alpha, \alpha)$ [mg/dl] represents PP glucose average, respectively. Thus, the optimal bolus variation is the solution of the following optimization problem:

$$\alpha^O(x_\alpha) = \arg\min_{\alpha \in I_\alpha} J(\alpha(\cdot), x_\alpha) \tag{4.12}$$

where the cost function is defined as follows:

$$J(\alpha(\cdot), x_\alpha) = \epsilon_1 T^{pp}_{Th}(x_\alpha, \alpha) + \epsilon_2 \frac{A^{pp}(x_\alpha, \alpha) - A^{pp}(x_\alpha, 1)}{100} + \epsilon_3 \tag{4.13}$$

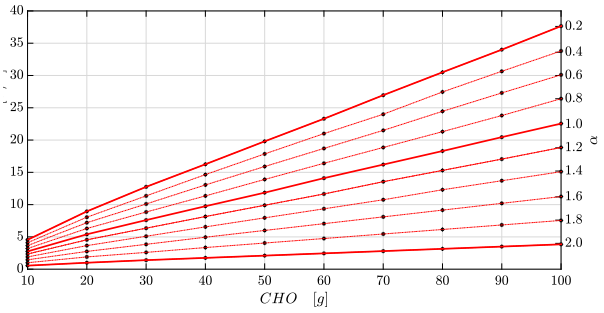


FIGURE 4.6: The black stars represent I_B^{α} with $\alpha = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$, $\widehat{CHO}_0 = [10, 100]$, $bg_0 = 160\ mg/dl$, $iob_0 = 0$, and $k_{m0} = 6.00$. The red lines show the evolution of each I_B^{α} by varying $\widehat{CHO}_0$ of Patient A of the adult virtual population of the UVA/Padova simulator. The thick red lines highlight for $\alpha = \alpha^{MIN}$ set equal to 0.2, $\alpha = \alpha^{MAX}$ set equal to 2, and $\alpha = 1.0$, i.e., the CT.

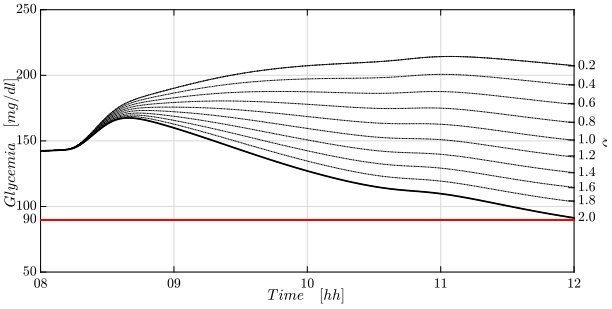


FIGURE 4.7: The black lines are PP glucose profiles of Patient A of the adult virtual population of the UVA/Padova simulator. The profiles are obtained by correcting i_B with $\alpha = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ in the experiment characterized by $\widehat{CHO}_0 = 30\ g$, $bg_0 = 160\ mg/dl$, $iob_0 = 0$ and $k_{m0} = 6.00$. The thick black line represents the optimal variation and corresponds to $\alpha = 2.0$. The red line is the threshold $Th = 90$ mg/dl.

where the weights ϵ_1 and ϵ_2 balance the trade-off between avoiding PP hypoglycemia events and decreasing the averaged PP glucose concentration. The parameter ϵ_3 represents a safety threshold to guarantee a more conservative bolus correction. This parameter can have different values on different intervals and is determined on the basis of the daily pattern of insulin sensitivity, which is lower in the morning than afternoon and evening [18]. This procedure is repeated for all entries of X_α and a set of optimal insulin variation I_α^O is obtained. Figure 4.6 shows a subset of i_B^{α} by considering α values belonging to I_α for a group of experiments characterized by different $\widehat{CHO}_0$ with fixed $bg_0 = 160$ mg/dl, and $iob_0 = 0$. Figure 4.7 shows PP glucose profile of a single experiment obtained by correcting the nominal bolus with different α.

4.3.2 Step 2: Identification of Model Parameters

The first step provides a set of optimal insulin variations with respect to several different conditons. I_α^O is split in two subsets: half is used for identification purpose, i.e., the identification subset, and half is used to perform the model validation, i.e., the testing subset. In order to identify the model of Eq. 4.6, the selection of regressors is addressed. The predictor variables have to fulfill the following requirements: they have to be measurable at k_m and glucose-related. The proposed regressors consist of IIG and $IIGS$. The choice of the number of regressors is based on the principle of parsimony, while their selection is based on physical considerations.

Specifically, the regressor IIG is the measurement of glucose concentration at mealtime and represents the actual information of the current level of glucose concentration. The regressor $IIGS$ represents the change of glucose concentration in the T_{GS} minutes before the meal. The combination of these measurements depicts the glucose trend with respect to a value at a mealtime. Hence, X_R and X_R^2 are defined as follows:

$$X_R = \begin{bmatrix} IIG & IIGS \end{bmatrix} \quad X_R^2 = \begin{bmatrix} IIG^2 & IIGS^2 \end{bmatrix} \tag{4.14}$$

and the values of the regressors for the collected experiments are the lines of last two columns of X that belong to the identification subset and are collected in the following matrix X_{R0}:

$$X_{R0} = \begin{bmatrix} IIG_0 & IIGS_0 \end{bmatrix}. \tag{4.15}$$

Thus, the optimal θ^I, θ^{II}, θ^{III} are identified by LS as follows:

$$\theta^{I_O}, \theta^{II_O}, \theta^{III_O} = \arg\min_{\theta^I, \theta^{II}, \theta^{III}} J^{LS}(\theta^I, \theta^{II}, \theta^{III}, I_\alpha^O, X_{R0}). \tag{4.16}$$

where the cost function is defined as follows:

$$\begin{aligned} &J^{LS}(\theta^I, \theta^{II}, \theta^{III}, I_\alpha^O, X_{R0}) = \sum_{j=1}^{N_{I_\alpha}/2} (||\alpha_j^O - f(X_{R0_j})||^2) = \\ &\sum_{j=1}^{N_{I_\alpha}/2} ||\alpha_j^O - \left(\theta^I + \theta^{II} X_{R0_j} + \theta^{III} X_{R0_j}^2\right)||^2 = \\ &\sum_{j=1}^{N_{I_\alpha}/2} ||\alpha_j^O - \left(\theta^I + \theta^{II_1} IIG_{0_j} + \theta^{II_2} IIGS_{0_j} + \theta^{III_1} IIG_{0_j}^2 + \theta^{III_2} IIGS_{0_j}^2\right)||^2 \end{aligned} \tag{4.17}$$

where

$$\theta^{II} = \begin{bmatrix} \theta^{II_1} \\ \theta^{II_2} \end{bmatrix} \quad \theta^{III} = \begin{bmatrix} \theta^{III_1} \\ \theta^{III_2} \end{bmatrix} \tag{4.18}$$

and X_{R0_j} is the line of X_R and IIG_{0_j} and $IIGS_{0_j} \in \mathbb{R}$ are its elements.

4.3.3 Model Summary

In order to improve PP glucose control, a MLR model able to individualize CT is proposed. MLR model represents the relation between the pre-prandial glucose measurements and the optimal meal bolus variation of a specific patient and is used to predict the correction of the nominal meal bolus. The correction of the nominal meal bolus can be formulated with the

following piecewise model:

$$i_B^{\hat{\alpha}^O} = i_B \cdot \hat{\alpha}^O \qquad s.t.\ \hat{\alpha}^O = \theta^{I_O}(T_t) + \theta^{II_O}(T_t)X_{R0} + \theta^{III_O}(T_t)X_{R0}^2 \qquad t = 1, ..., T$$

where $\hat{\alpha}^O$ is the estimated optimal bolus variation at mealtime, $i_B^{\hat{\alpha}^O}$ is the estimated optimal insulin bolus corrected by $\hat{\alpha}^O$ at mealtime, and X_{R0} is the matrix of model predictors. In order to individualize the CT, the MLR model parameters $\theta^{I_O}, \theta^{II_O}, \theta^{III_O}$ are patient-tailored and are identified by performing the following least square estimation for each day period:

$$\theta^{I_O}, \theta^{II_O}, \theta^{III_O} = \arg\min_{\theta^I, \theta^{II}, \theta^{III}} J^{LS}(\theta^I, \theta^{II}, \theta^{III}, I_\alpha^0, X_{R0}).$$

In order to identify and test proposed approach, the real optimal correction α^O have to be defined. In order to collect a set of optimal α^O for each possible condition set present in X, identification experiments are conducted. For a given test x_0, an optimization problem is formulated and each sample of I_α^O is achieved by optimizing the following cost function:

$$\alpha^O(x_\alpha) = \arg\min_{\alpha \in I_\alpha} J(\alpha(\cdot), x_\alpha)$$

where $x_\alpha \in \mathbb{R}$ is the line of X_α.

4.3.4 Simulation Settings

The results of a case study on patient (Patient A) of the adult virtual population of the UVA/-Padova simulator [18] are presented. The goal is to demonstrate that PP glucose regulation can be significantly improved by the proposed approach, which is suitable for individualization. The testing scenario lasts 4 days and its meal schedule is reported in Table 4.6. It is designed to reproduce the critical situations observed during clinical trials [59]. Three testing scenarios are defined to evaluate the performance of the identified model: Scenario I, the nominal scenario, i.e., without uncertainties; Scenario II, a perturbed scenario where a random ±25% variation of meal amounts is considered; Scenario III, a perturbed scenario with both a random ±25% variation of amount of carbohydrate intake and a random ±15% variation of the nominal insulin sensitivity. The perturbed scenarios represent a critical condition because the model is not aware of these variations. Scenario II simulates that the estimation of carbohydrate intakes is not reliable and a wrong meal compensation is happening, while Scenario III aims to reproduce as much as possible the uncertainties of a real scenario due to unknown factors. Hypotreatments (ht) of 15 g are administrated to the patient in case glucose concentration falls below 65 mg/dl.

The day periods for Patient A are the following:

$$\begin{cases} T_1 = [t_1, t_2] \\ T_2 = [t_2, t_3] \\ T_3 = [t_3, t_1] \end{cases} \tag{4.19}$$

where t_1 = 4:00, t_2 = 11:00, and t_3 = 17:00 on the basis of the optimized CR pattern of the considered patient. Once a model for each day period is identified via *in silico* trials, the insulin percent variation model is designed in the following form for Patient A:

$$\hat{\alpha}_A^O = \theta_A^{I_O}(T_t) + \theta_A^{II_O}(T_t)X_R + \theta_A^{III_O}(T_t)X_R^2,\ t = 1, 2, 3 \tag{4.20}$$

where T_t is the t^{th} time interval where CR is constant.
The lower and upper bound for the admissible region of α are set as follows: $\alpha^{MIN} = 0.1$, and $\alpha^{MAX} = 2$. The threshold Th is fixed to 90 mg/dl, the threshold T_{GS} is set to 30 minutes and the parameters ϵ_1, ϵ_2, and ϵ_3 are set as follows: $\epsilon_1 = 100$, $\epsilon_2 = 10$, $\epsilon_3(T_1) = 0.2$, $\epsilon_3(T_2) = \epsilon_3(T_3) = 0.3$. The values of the parameter ϵ_3 are based on physical considerations: they reflect the daily pattern of insulin sensitivity, which is lower in the morning than the rest of day [18]. In order to evaluate the fitting performance of the model, the lines of X which belong to the testing subset are used to compute the optimized correction $\hat{\alpha}_A^O$ and evaluate the quality of the MLR prediction with respect to the optimal values α_A^O via the computation of the Root Mean Squared Error (RMSE), i.e., the square root of the variance of the residuals, defined as follows:

$$RMSE = \sqrt{\frac{SSE}{\nu}} \tag{4.21}$$

where $\nu = \frac{N_{I_\alpha}}{2} - M$ represents the residual degrees of freedom and $M = 5$ is the number of fitted parameters, and SSE is the sum squared error, defined as follows:

$$SSE = \sum_{j=1}^{N_{I_\alpha}/2} (||\alpha_{A_j}^O - \hat{\alpha}_{A_j}^O)||^2). \tag{4.22}$$

An RMSE value closer to 0 indicates a good fit of the identified model. Moreover, the prediction performance of the model are evaluated on 100 testing samples for each day period by simulating the scenario reported in Table 4.6 and some performance metrics are evaluated. The adopted metrics, which have been presented in Section 3.5 and are in line with the consensus AP metrics [60], are computed overall (O) and take all PP periods into account (4h) as reported in Table 4.8. The evaluation of the night period is discharged because the modification of the therapy is strictly related to the PP period.

Table 4.6: 4 Day Testing scenario.

	Time	CHO [g]	Description
Day 1	06:00	35.00	Breakfast
	08:45	20.00	Snack
	12:40	30.00	Lunch
	18:30	25.00	Snack
	20:45	80.00	Dinner
Day 2	07:00	15.00	Breakfast
	08:25	30.00	Snack
	12:45	63.00	Lunch
	19:10	50.00	Dinner
Day 3	07:30	45.00	Breakfast
	13:05	57.00	Lunch
	20:45	35.00	Dinner
Day 4	06:20	35.00	Breakfast
	08:00	10.00	Snack
	12:30	15.00	Lunch
	14:00	48.00	Snack
	19:45	48.00	Dinner

4.3.5 Discussion

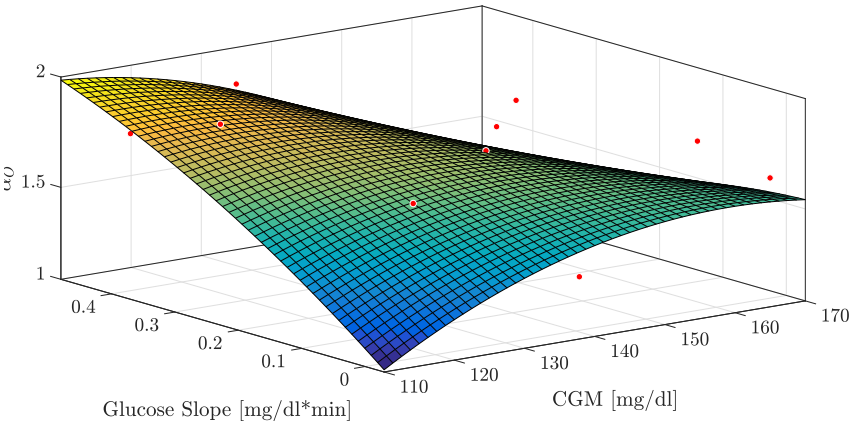
RMSE values obtained by evaluating the model by using the testing subset of the collected experiments are reported in Table 4.7, where each RMSE is associated with the corresponding identified regression model. Since RMSE indicates how close the observed data points are to the predicted values of the model, we can conclude that the models are identified with good accuracy. Moreover, although the RMSE results are sufficiently small on a one-hundred dataset, over-fitting does not occur in the regression model as shown in Figures 4.8(a), 4.8(b), and 4.8(c), where the fitted surfaces are compared with their own testing experiments. The performance metrics obtained by testing CT and the improved therapy with MLR on the three testing scenarios are reported in Table 4.8. Significant improvements are obtained with new technique in both the PP period and the overall periods. The improved MLR therapy avoids the hypoglycemic events and demonstrates that the glucose concentration never drops below 70 mg/dl. In PP period, which is mainly affected by the bolus variation, the correction of the meal bolus computed by MLR is able to decrease the time spent above the target by 80% and 73% in Scenario I and Scenario II, respectively. The time in target and in tight target are also improved by 13% and 10% in Scenario I and by 30% and 48% in the Scenario II, respectively. Finally, the average glucose is allowed to decrease by 4% in both Scenario I and Scenario II in PP periods.

TABLE 4.7: RMSE testing results.

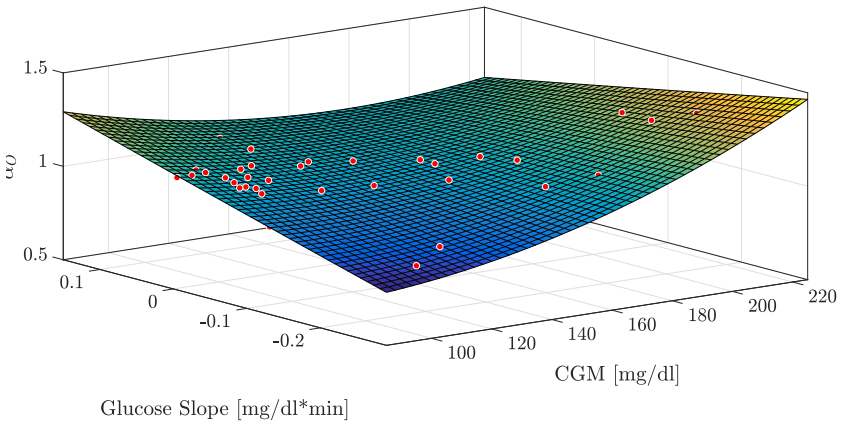
Day period T_1	Day period T_2	Day period T_3
0.22	0.19	0.07

The corrections of i_B predicted by the identified model allow BG concentration to remain in the euglycemic range for most of the time and avoid hypoglycemia events, as shown in Figure 4.9(a). In Scenario III, even more challenging than the previous scenarios, the improved MLR threapy increases the time spent in range, doubles the time spent in tight target and reduces hyperglycemia by avoiding also the hypoglycemia. The absence of hypoglycemia in Scenario III results in an increase of the average glycaemia by 2% in the PP period, but it represents a good compromise for the overall glucose regulation, as shown in Figure 4.9(b).

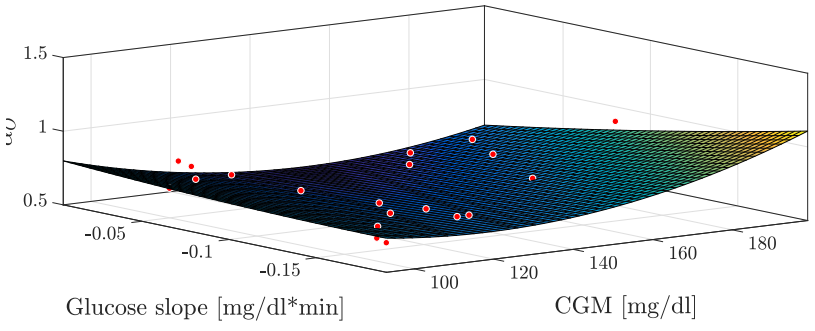
These promising results are obtained by testing the MLR model via *in silico* trials on a single subject; the application of this methodology to a real T1D patient is limited to the fact that it requires potentially dangerous tests for the patient. In order to apply this technique on real patients, the identification of an individualized model able to predict patient glucose dynamics is required [61, 62]. *In vivo* data can be exploited to identify an offline patient-tailored model, which has the capability to simulate the patient glucose dynamics also in critical conditions needed for this approach. Then, the identified model can be used to identify the parameters of MLR model. Given the satisfactory results, this approach has been published in [63].



(a)



(b)



(c)

FIGURE 4.8: (a-b-c) The regression surface identified from the dataset belonging to day periods (a) $T_1 = [t_1, t_2]$, (b) $T_2 = [t_2, t_3]$, (c) $T_3 = [t_3, t_1]$ of Patient A of the adult population of the UVA/Padova simulator. The red points represent the testing samples.

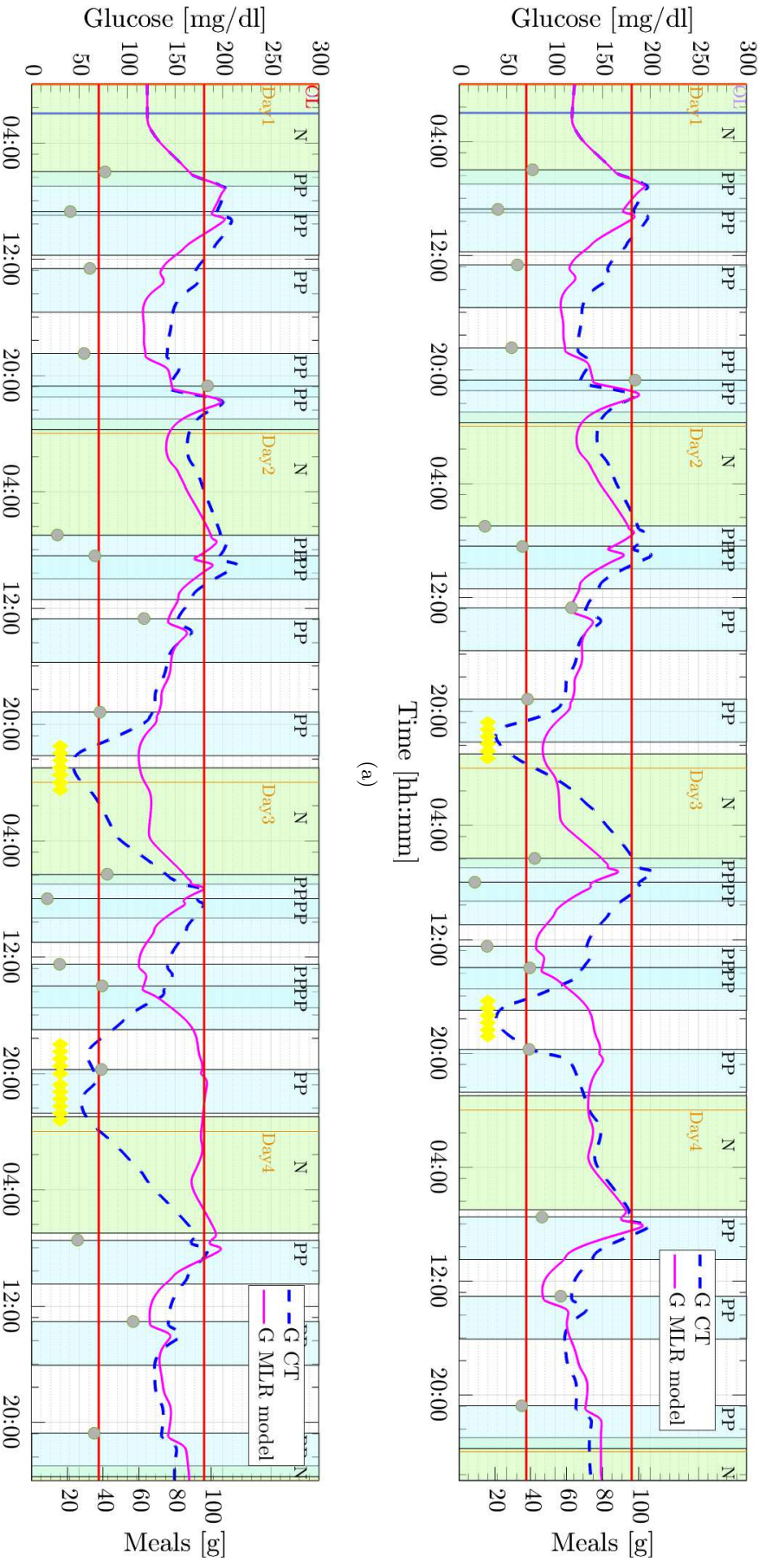


FIGURE 4.9: (a-b) Comparison of glucose trends of Patient *A* of the adult virtual population of UVA/Padova simulator on Scenario II (a) and Scenario III (b). The blue line is the glucose profile with CT, and the magenta line is the glucose profile with the MLR data-driven model. The grey circles denote the meals and the yellow diamonds are the ht events. PP denotes the PP period and night periods are highlighted in yellow.

TABLE 4.8: Results of CT vs MLR improved therapy on all testing scenarios in Patient A of the adult population of the UVA/Padova simulator.

		Scenario I		**Scenario II**		**Scenario III**	
		O	PP	O	PP	O	PP
A [mg/dl]	CT	134.31	138.52	136.92	141.94	139.28	150.88
	MLR	132.43	133.00	130.60	132.67	150.20	153.76
T_r [%]	CT	89.20	86.99	81.55	72.08	74.93	68.46
	MLR	98.20	98.01	96.75	94.07	86.72	81.29
T_{tr} [%]	CT	55.13	53.04	47.71	43.06	27.95	15.96
	MLR	62.00	61.03	68.19	64.75	37.94	31.76
T_a [%]	CT	5.80	9.61	11.99	21.27	14.41	22.55
	MLR	1.80	1.99	3.25	5.93	13.28	18.71
T_b [%]	CT	5.00	3.41	6.46	6.64	10.66	8.99
	MLR	0.00	0.00	0.00	0.00	0.00	0.00
T_{b60} [%]	CT	3.38	2.40	5.27	5.76	6.18	5.55
	MLR	0.00	0.00	0.00	0.00	0.00	0.00
$\#ht$	CT	9.00	5.00	12.00	10.00	18.00	11.00
	MLR	0.00	0.00	0.00	0.00	0.00	0.00

Chapter 5

Artificial Pancreas

The main aim of the AP is the automatic regulation of the BG concentration for people affected by type 1 diabetes through exogenous insulin administrations. The AP system is composed of three main parts: a glucose sensor, an insulin infusion device, and a control algorithm. The concept of AP was proposed for the first time in 1959 by Professor Perry McCullagh, an endocrinologist at The Cleveland Clinic. The first AP was developed in 1964 by Dr. Arnold Kadish, a Californian internist. The device had the size and shape of a large backpack and was not compatible for free-living conditions [64]. In 1974, Albisser and colleagues reported the use of an extracorporeal AP system to maintain BG concentration, i.e. glycemia, in the normal range during consumption of meals [65]. In the same year, Pfeiffer and colleagues also reported use of a computerized glucose controlled insulin infusion artificial beta cell system [66]. In 1977, Miles Laboratories has developed a Glucose Controlled Insulin Infusion System (GCIIS) designated by the Trademark (BIOSTATOR) [67]. Biostator was the first commercial AP for inpatient control. It consists of a rapid online glucose analyzer, a computer/controller for the calculation and control of insulin infusion, and a multichannel infusion system. Biostator needed venous access and was highly invasive and non-portable, forcing the patients to be hospitalized. In order to allow to use the AP in the long term, the subsequent improvements concerned with the miniaturization of the AP. Over the years, the improvements went in the direction of developing a non-invasive, safe, and portable system, by relying on the latest technological developments. Continuous Subcutaneous Insulin Infusion (CSII) pump therapy was introduced to treat patients with type 1 diabetes in the late 1970's [68]. Then, minimally invasive subcutaneous glucose sensing was commercially introduced in 1999 by the MiniMed Continuous Glucose Monitoring (CGM) system. When subcutaneous insulin pump technology was combined with a continuous BG monitoring system, subcutaneous AP were developed. The current AP components are a subcutaneous insulin pump, a CGM sensor, and a control algorithm. In the first AP system,

the control algorithm was executed on a computer, which received and interpreted the electrical signals generated by the glucose analyzer, and in turn instructed insulin pump to delivery insulin. The performances of the first algorithms were limited by the computational power of available computers. The most important step toward a non-invasive and portable AP was the replacement of the laptop and all the wires with a portable device. Currently, the used portable device is a modified smartphone where the control algorithm resides.
Despite important developments in sensor and pump technologies, the AP must cope with the delays and inaccuracies in both glucose sensing and insulin delivery. The core of the AP is the control algorithm used to calculate insulin delivery in order to perform an accurate and robust Closed-Loop (CL) control. The controller design must take into account that the physiological dynamics have a relatively slow response, so a delay has to be considered before performing the next control action. However, a slow response is not able to attenuate PP glucose peaks. Hence, the design of the control algorithm has to face a trade-off between slow-pace regulation that includes both soft control actions applicable to quasi-steady state (e.g., overnight when the glucose levels has to be maintained constant usually without meals), and aggressive responses in the PP periods [69].
Moreover, since physical devices, like glucose sensor and insulin pumps, are involved, the control theory has been implemented in discrete time. The adopted sample times can be the sample times of insulin pumps or CGM sensors, which results to be smaller with respect to the involved biological dynamics.

5.1 Control algorithms

The main purpose of the control algorithm is to achieve an insulin therapy to keep BG concentration within a predefined range by acting on insulin delivery. In the control scheme, the controlled variable is the glycemia, the manipulated variable is the infused insulin, and the measured output is the subcutaneous glucose. The subcutaneous glucose is provided by the CGM sensor, while the insulin is delivered by the subcutaneous pump. Recently, alternative routes of insulin delivery have been explored such as intraperitoneal [70, 71] and implantable pumps [72]. The system is subjected to various disturbances, of which the most significant are represented by the glucose variations induced by the meals intake. It is important to note that this disturbance may be announced, approximately known, or even predictable [73]. There are two classes of control schemes: open- and closed-loop. The main difference is that OL methods design an insulin therapy based only on the knowledge of consumed meals, and eventually some clinical parameters, without considering measurements coming form CGM sampled every 1-5

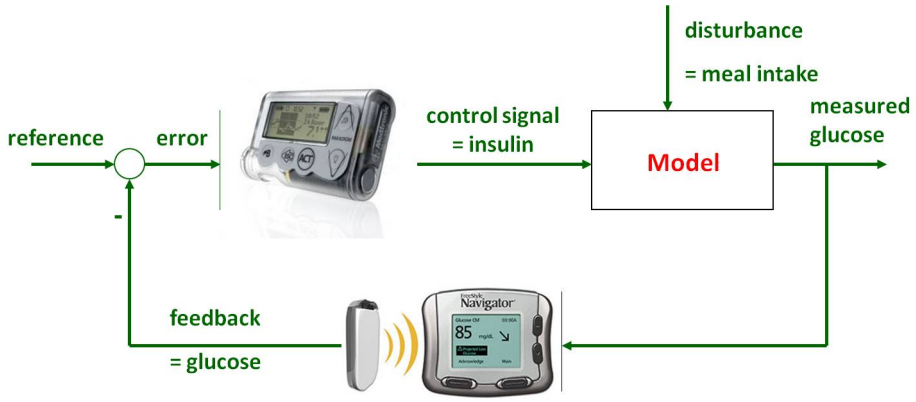


FIGURE 5.1: Conceptual block diagram of CL glucose control.

minutes, whereas CL control adapts and chooses the control action by exploiting the subcutaneous measurements.

The OL control strategy corresponds to the application of the CT, previously described in Section 2.3.1.

Closed-loop glucose control

A CL control scheme manages the insulin delivery by exploiting the knowledge of the subcutaneous glucose concentration measurements. In order to maintain the glycemia within the euglycemic range, this control approach ideally allows to correct the insulin injections when glucose level drops too low or increases too much. Furthermore, no explicit knowledge of external disturbances should be required because the effects of disturbances are reflected on CGM measurements. However, the aforementioned control scheme needs to take into account different critical aspects such as the presence of time-variable dynamics and time delays. The action of insulin on plasma glucose is subject to significant physiological delays that affect glycemia due the absorption of insulin from the subcutaneous level to the blood glucose levels. Since the glucose sensor measures the interstitial glucose, the diffusion process from plasma to interstitial tissues, i.e. subcutaneous tissues, needs to be taken into account. As a result, the information about the needed amount of insulin may arrive too late to prevent hyper- or hypoglycemic episodes. Moreover, the saturation limits of the subcutaneous pump for delivery and lack of a reliable individual model of the patient impose intrinsic limits to the time constant. The solution of increasing the responsiveness of the CL system may lead to unstable behaviors. On the other hand, the glycemic PP peaks cannot be attenuate with a conservative control strategy. PP regulation can only be managed by delivering the necessary insulin in a brief time window. The problem is to find a trade-off between PP regulation and fasting regulation. In order to obtain a faster glucose regulation in presence of meals, a feed-forward action is introduced.

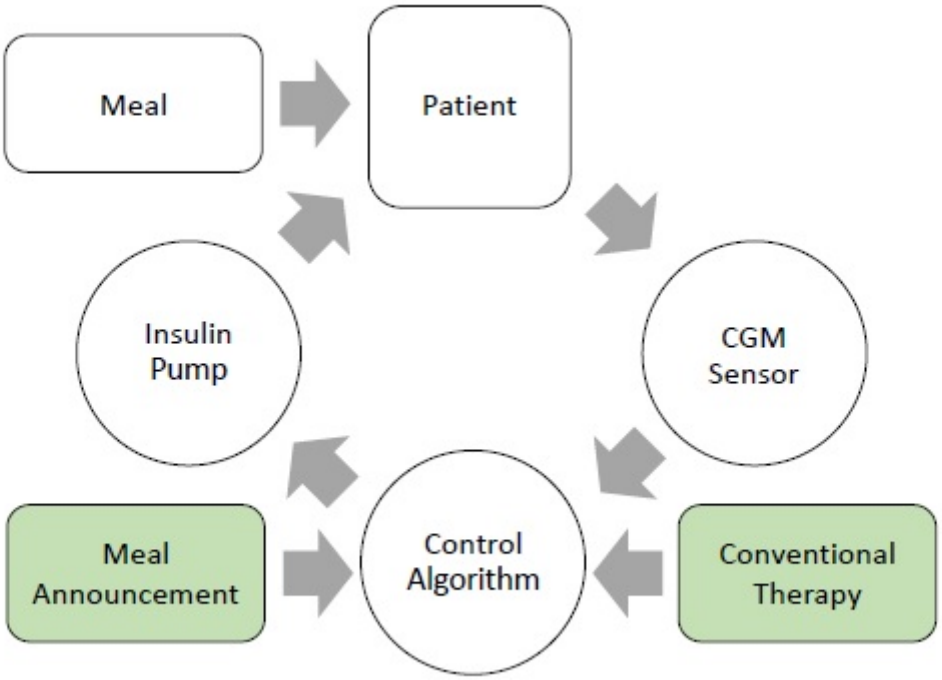


FIGURE 5.2: Conceptual AP representation.

This is the concept of meal announcement, where the patient is involved in the control loop by announcing to the controller a meal intake. The patient announces the meal and specifies the estimated quantity of carbohydrates (CHO) included in the meal itself. This information alerts the controller that prompt insulin infusions will be needed to compensate the induced glucose rise. The patient is involved in the control loop and this means that the proposed concept of AP system is not fully automated but the control strategy is partially driven by the patient actions. Meal announcement is considered an additional knowledge, which can improve the critical PP glycemia control. If the patient forgets to announce the meal, the AP system should remain able to operate safely. On the other hand, a flow of basal insulin, which typically is constant in portions, is delivered throughout the day. The controller aims to provide the corrective insulin value with respect to the basal insulin value to manage the time-variable dynamics. Thus, the conventional therapy is used to further improve the control strategy, and potentially improving the glucose control performance since it contains reliable information about the patient that are continually adjusted by the physician. Finally, the solution for a good glucose regulation is a control scheme that combines feed-forward and feedback actions [73]. The control algorithm block is shown in Figure 5.2. The required inputs are not only CGM sensor measurements, but also the information provided by the conventional therapy and meal announcements. The output is the optimal insulin administration.

5.2 Model Predictive Control

In the last years, one of the most promising approach revealed to be the Model Predictive Control (MPC) algorithm that achieved successful results in the AP context, both *in silico* and *in vivo*. Several MPC algorithms have been tested in clinical trials with satisfactorily results [59, 74–79]. In these trials the patients followed a therapy defined via a fully automatic CL control. The MPC approach exploits a glucose-insulin model of the patient to predict near-future BG values and, consequently, computes the optimal insulin dose.
Specifically, the MPC has three main components that are the model, the cost function, and the constraints. The model is useful to compute the prediction of both the future states and outputs of the system as a function of the current state, future inputs, and future values of the estimated disturbances. In the context of AP, the input is the suggested injected insulin, the output is the estimated glucose concentration and the disturbances are represented by meal intakes. At each sampling time k, the future sequence of manipulated variables is computed by optimizing a cost function $J(x(k)), u(\cdot), k)$, that is:

$$u_0(k) = \underset{u(\cdot)}{\operatorname{argmin}}\, J(x(k)), u(\cdot), k). \tag{5.1}$$

Constraints on the inputs can be added in this optimization problem. Then, according to the receding horizon principle, only the first control action is applied. The receding horizon principle states that at each time instant, the optimization problem is solved. The $u_0(k)$ is computed over the prediction horizon and only $u_0(0)$ is applied to the system as the current input. Then, the same procedure is repeated by translating the prediction horizon. The MPC approach allows to convert the controller design problem into a Finite Horizon Optimal Control Problem (FHOCP). This control approach is driven by a metabolic model that must be able to predict the glucose-insulin dynamics of the patient under control. Thus, a model with good prediction capabilities is expected to improve the glucose regulation performance.
Figure 5.3 shows the developed control scheme. In the scheme, the required insulin u^p is defined according to the conventional therapy. The signal u^p is the contribution of both the basal insulin i_b and the suggested insulin u^{MPC} computed by the MPC controller. In order to compute u^{MPC}, two different situations have to be taken into account. During fasting periods, u^{MPC} depends only on the subcutaneous glucose error $e = y_{sp} - y$, where y_{sp} is the glucose set-point and y is the CGM measurement. If there is a meal announcement, the estimated amount of carbohydrates $d = \hat{m}$ associated to the meal m is used to compute the optimal nominal insulin bolus i_B^o according to the open loop therapy, which corresponds to conventional therapy. Then, the MPC controller uses the nominal insulin bolus i_B^o, the estimated CHO $\hat{m}$, and the glucose error

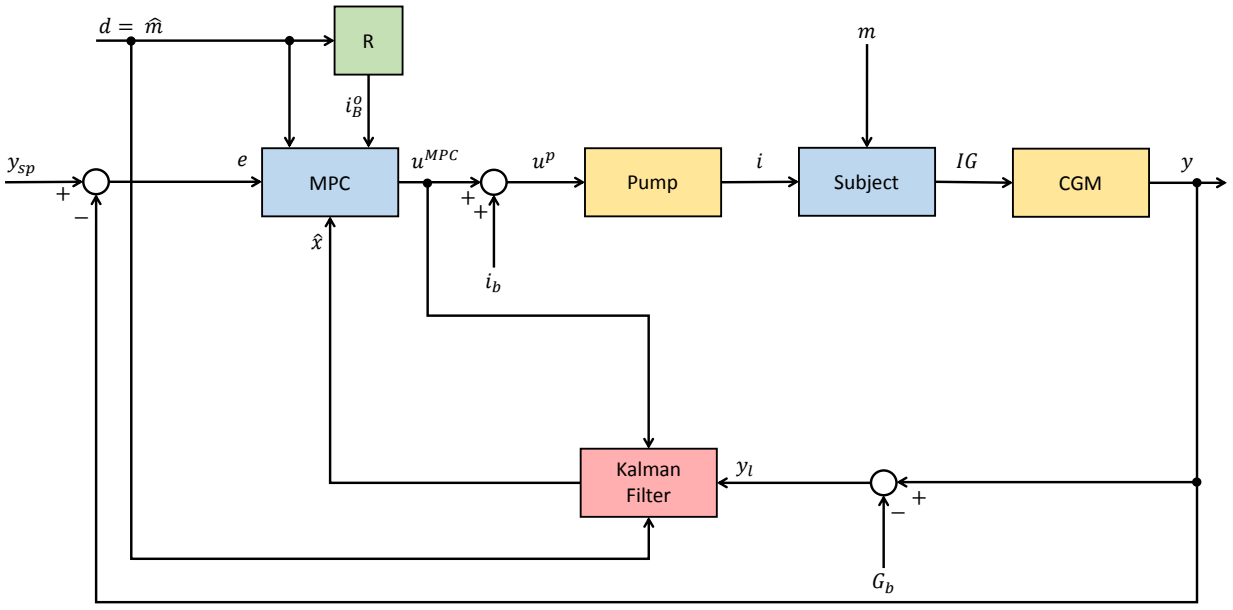


FIGURE 5.3: CL scheme for BG regulation. The blue blocks represent the MPC and the patient, yellow blocks represent the hardware, and the green block represents the OL therapy used to compute the nominal insulin boluses, which is used in the meal announcement.The red block is the Kalman filter, which provides an estimation of the non measurable state given the noisy CGM measurements and the controller insulin suggestions.

to drive u^{MPC}. Thus, the u^{MPC} can be considered as an insulin variation with respect to basal insulin i_b. Then, the pump receives the u^p and infuses the insulin i.

5.2.1 Linear MPC

Linear MPC (LMPC) exploits a linear model to perform the predictions used to compute the control inputs. A linear model is usually represented in state-space variables as follows:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \tag{5.2}$$

where the system is a linear discrete time-invariant model, and where $x \in \mathbb{R}^n$ is assumed to be measurable, $u \in \mathbb{R}^m$ is the input vector, and $y \in \mathbb{R}^p$ is the output vector. In the linear case [80], the optimal control sequence is obtained by minimizing the following cost function:

$$J(x(k)), u(\cdot), k) = \sum_{i=0}^{N-1} (||x(k+i)||_{2_Q}) + ||u(k+i)||_{2_R})) + ||x(k+N)||_{2_S}). \tag{5.3}$$

In case of non-constrained control, one of the advantages of LMPC is that the control law can be defined in closed-form, so the solution of the problem is the following:

$$u^o(k+i) = -K(i)x(k+1), \quad i = 0, 1, ..., N-1 \tag{5.4}$$

where

$$K(i) = (R + B'P(i+1)B)^{-1}B'P(i+1)A \tag{5.5}$$

and $P(i)$ is the solution of the Riccati equation

$$P(i) = Q + A'P(i+1)A - A'P(i+1)B(R + B'P(i+1)B)^{-1}B'P(i+1)A \tag{5.6}$$

with boundary condition $P(N) = S$. Given the reciding horizon criterion, the MPC control law is state-feedback, time-invariant, and is given by

$$u^{MPC}(k) = -K(0)x(k). \tag{5.7}$$

Since the metabolic model described in the previous chapter is time-variant and highly nonlinear, the LMPC cannot be directly used. Thus, the original metabolic model needs to be transformed: it is first approximated to time-invariant and then it is linearized around a fictitious basal equilibrium point. The time-invariant approximation is performed by approximating the time-varying gastric emptying coefficient k_{empt} defined in the Eq. 3.2 to its average value k_{mean}. Then, the second step consists of performing a linearization around an equilibrium point representing the steady state of the patient during fasting periods. Hence, the insulin input $i(t)$ has been imposed equal to the insulin basal value i_b, and the meal input $d(t)$ of Eq. 3.2 equal to 0. This transformation procedure can be applied to the entire virtual population, but there is not a correspondent real patient-specific model, since the virtual population has been generated by randomly extracting different realizations of the parameters vector from a joint parameters distribution obtained from real 204 patients [27]. An average nonlinear time-variant model has been defined by averaging the parameters of the entire virtual population. So, it represents the best trade-off for the controller synthesis. The resulting model is subsequently approximated to time-invariant and linearized to be included in the LMPC. The average linear time-invariant model is then discretized with a sampling time of 15 minutes (that is the sampling time chosen for the control action), obtaining the following discrete-time state-space model:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + Md(k) \\ y(k) = Cx(k) \end{cases} \tag{5.8}$$

where x and y are the differential states and output with respect to their steady-state values, respectively, $u = u^{MPC}$ represents the differential infused insulin by the controller with respect to the basal insulin i_b, and d represents the amount of CHO associated to the meal announced to the controller. Therefore, a discrete-time LMPC is derived from a unique state space linearized approximation of the nonlinear time-variant model [81].
The state of the model in general is not measurable, thus a Kalman filter has been incorporated [5]. On one hand the controller task is to achieve the basal equilibrium at the end of the prediction horizon. The linearized model is then modified as follows:

$$\begin{cases} x_l(k+1) = Ax_l(k) + Bu(k) + Md(k) + \epsilon_x^{KF}(k) \\ y_l(k) = Cx_l(k) + \epsilon_y^{KF}(k) \end{cases} \tag{5.9}$$

where $\epsilon^{KF} = [\epsilon_x^{KF} \epsilon_y^{KF}]'$ is a multivariate zero-mean WGN with covariance matrix given by

$$V = \begin{bmatrix} Q^{KF} & 0 \\ 0 & R^{KF} \end{bmatrix} \tag{5.10}$$

with $Q^{KF} > 0$ and $R_{KF} > 0$. With the assumption of a stabilizable and detectable system, LMPC control law is obtained minimizing the following quadratic cost function:

$$J(\hat{x}(k|k), u(\cdot), k) = \sum_{i=0}^{N-1} (q(y(k+i) - y_{sp}(k+i))^2 + (u(k+i) - u_0(k+i))^2 + ||x(k+N)||_{2_P}) \tag{5.11}$$

with

$$\begin{aligned} &x(k) = \hat{x}(k|k) \\ &x(k+i+1) = Ax(k+i) + Bu(k+i) + Md(k+i) \\ &y(k+i+1) = Cx(k+i+1) \\ &u^0(k+i) = u_{OL}(k+i) - i_b(k+i) \end{aligned} \tag{5.12}$$

where the state $x(k)$ is substituted with its estimation $\hat{x}(k|k)$, $q > 0$ is the weight associated to the output, y is the measured noisy subcutaneous glucose, y_{SP} is the glucose set-point, u_{OL} is the insulin that would be injected by the OL therapy, and P is the unique non-negative solution of the discrete-time Riccati equation:

$$P = qC'C + A'PA - A'PB(1 + B'PB)B'PA. \tag{5.13}$$

Finally, the steady-state Kalman filter is described by the following equations:

$$\begin{aligned}\hat{x}(k+1|k) &= A_{kf}\hat{x}(k|k) + B_{KF}u(k) + M_{KF}d(k)\\ \hat{x}(k|k) &= \hat{x}(k|k-1) + L[y(k) - C_{KF}\hat{x}(k|k-1)]\end{aligned} \tag{5.14}$$

with

$$L = P_{KF}C'_{KF}[C_{KF}p_{KF}C'_{KF} + R_{KF}]^{-1} \tag{5.15}$$

where P_{KF} is the unique positive-definite solution of the Riccati algebraic equation:

$$P_{KF} = Q_{KF} + A'_{KF}P_{FK}A_{KF} - A_{KF}P_{FK}C'_{KF}(C_{KF}p_{KF}C'_{KF} + R_{KF})^{-1}C_{KF}P_{FK}A'_{KF} \tag{5.16}$$

The Kalman filter has sampling time equal to 5 minutes, like the sampling time of the CGM. Therefore, since each control action is performed every 15 minutes, three Kalman filter predictions are computed within each control action, leading to three state estimations of which the last one is passed to the LMPC as initialization state $\hat{x}(k|k)$ in the cost function (5.11).

5.2.2 Closed-Form Implementation

In order to avoid online optimization or the computational and memory burden of an explicit MPC for constraints systems, the controller does not include explicitly the constraints [5]. Thus, it is possible to calculate the closed-form of the LMPC control law by relying on the Lagrange formula. The model output and states predictions within the horizon N can be obtained through the following formula:

$$Y(k) = \mathcal{A}_c x(k) + \mathcal{B}_c U(k) + \mathcal{M}_c D(k) \tag{5.17}$$

where $x(k)$ is the linear model state at time k and the other quantities are defined as

$$Y(k) = \begin{bmatrix} y(k+1)\\ y(k+2)\\ \vdots\\ y(k+N-1)\\ x(k+N)\end{bmatrix}, \quad U(k) = \begin{bmatrix} u(k)\\ u(k+1)\\ \vdots\\ u(k+N-1)\end{bmatrix}$$

$$D(k) = \begin{bmatrix} d(k) & d(k+1) & \cdots & d(k+N-1)\end{bmatrix}^T \tag{5.18}$$

$$\mathcal{A}_c = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^{N-1} \\ A^N \end{bmatrix}, \quad \mathcal{B}_c = \begin{bmatrix} CB & 0 & \cdots & \cdots & 0 \\ CAB & CB & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ CA^{N-2}B & \cdots & \cdots & CB & 0 \\ A^{N-1}B & \cdots & \cdots & AB & B \end{bmatrix}$$

$$\mathcal{M}_c = \begin{bmatrix} CM & 0 & \cdots & \cdots & 0 \\ CAM & CM & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{N-2}M & CA^{N-3}M & \cdots & CM & 0 \\ A^{N-1}M & A^{N-2}M & \cdots & AM & M \end{bmatrix}$$

Note that, according to the cost function (5.11), the last element of $Y(k)$ represents the model state prediction $x(k+N)$ at the horizon, and the other matrices are properly defined. The predicted trajectory $Y(k)$ depends on the applied input trajectory $U(k)$, that has to be optimized through the minimization of the cost (5.11). By defining the weight matrix $\mathcal{Q}$ as

$$\mathcal{Q} = \begin{bmatrix} q & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & q & \vdots \\ 0 & \cdots & 0 & P \end{bmatrix} \in \mathbb{R}^{(N-1+n)\times(N-1+n)}$$

with $n = 13$ representing the number of the states associated to the controller model and with $q > 0$ representing the output weight defined in Eq. 5.11, the controller cost function can be rewritten as follows:

$$\begin{aligned} J(\hat{x}(k|k), u(\cdot), k) = &\left(\mathcal{A}_c\hat{x}(k|k) + \mathcal{B}_c U(k) + \mathcal{M}_c D(k) - Y_{sp}(k)\right)^T \mathcal{Q}\left(\mathcal{A}_c\hat{x}(k|k)\right. \\ &\left.+\mathcal{B}_c U(k) + \mathcal{M}_c D(k) - Y_{sp}(k)\right) + \left(U(k) - U^0(k)\right)^T \left(U(k) - U^0(k)\right) \end{aligned} \tag{5.19}$$

where the reference vectors Y_{sp} and U^0 are defined as

$$\begin{aligned} Y_{sp}(k) &= \begin{bmatrix} y_{sp}(k+1) & y_{sp}(k+2) & \cdots & y_{sp}(k+N-1) & 0 & \cdots & 0 \end{bmatrix}^T \\ U_0(k) &= \begin{bmatrix} u_0(k) & u_0(k+1) & \cdots & u_0(k+N-1) \end{bmatrix}^T \end{aligned}$$

By zeroing the gradient with respect to U, the vector U^o containing the optimal input trajectory is achieved in the following closed-form:

$$U^o(k) = \left(\mathcal{B}_c^T\mathcal{Q}\mathcal{B}_c + I\right)^{-1}\left(-\mathcal{B}_c^T\mathcal{Q}\mathcal{A}_c\hat{x}(k|k) - \mathcal{B}_c^T\mathcal{Q}\mathcal{M}_c D(k) + \mathcal{B}_c^T\mathcal{Q}Y_{sp}(k) + U^0(k)\right)$$

where I is an identity matrix with proper dimensions. According to the *receding horizon* criterion, the time-invariant LMPC control law is given by:

$$u^{MPC}(k) = [1\ 0\ \cdots\ 0]\left(-K_x\hat{x}(k|k) - K_d D(k) + K_{Y_{sp}} Y_{sp}(k) + K_{U^0} U^0(k)\right) \tag{5.20}$$

where the gain matrices K_x, K_d, $K_{Y_{sp}}$, and K_{U^0} are defined as

$$\begin{aligned}
K_x &= \left(\mathcal{B}_c^T \mathcal{Q} \mathcal{B}_c + I\right)^{-1} \mathcal{B}_c^T \mathcal{Q} \mathcal{A}_c \\
K_d &= \left(\mathcal{B}_c^T \mathcal{Q} \mathcal{B}_c + I\right)^{-1} \mathcal{B}_c^T \mathcal{Q} \mathcal{M}_c \\
K_{Y_{sp}} &= \left(\mathcal{B}_c^T \mathcal{Q} \mathcal{B}_c + I\right)^{-1} \mathcal{B}_c^T \mathcal{Q} \\
K_{U^0} &= \left(\mathcal{B}_c^T \mathcal{Q} \mathcal{B}_c + I\right)^{-1}
\end{aligned}$$

5.3 Modular Control Architecture

In addition to MPC control strategy, a structured modular approach has proved to be a winning approach toward an automated CL glucose control. The modular AP architecture has been implemented as the combination of separate interacting components (modules) responsible for different tasks, and it involves decomposition of control, estimation, and signal management functions into multiple timescales. The main advantage of the modular AP architecture is the possibility for sequential development and clinical testing of the CL system [82].
The current architecture is described in detail in [83] and it is shown in Figure 5.4. The architecture is divided in the following four main layers and includes several modules:

- the offline layer
- the real time layer
- the continous time layer
- the hardware layer.

The offline layer is designed to perform an adaptation of the control algorithm by exploiting the available information on a specific patient. In order to initialize and individualize the control algorithm, conventional therapy and trials historical data are the currently needed information to be collected. More precisely, the conventional therapy, which includes information about the patient basal insulin, CR and CF parameters, is exploited to properly initialize the controller, while the clinical data are used to individualize the metabolic model included in the MPC. Hence, the introduction of this layer has been thought to face the patients inter-variability, but

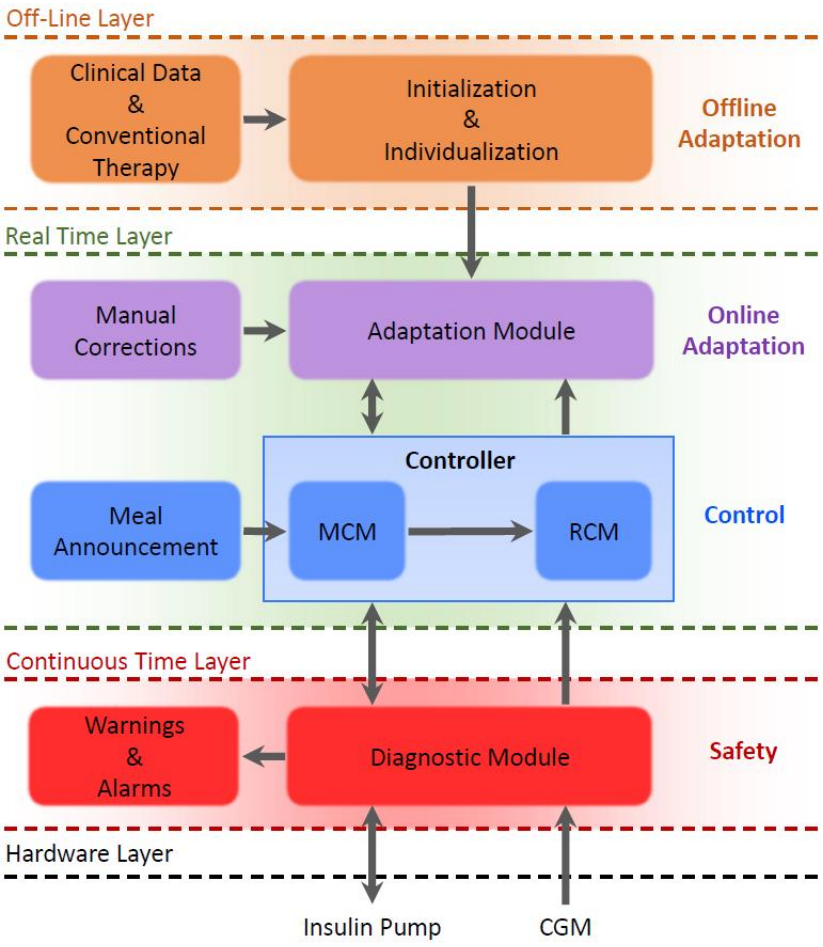


FIGURE 5.4: Representation of the AP modular architecture.

this adaptation strategy has to be computed offline because it is computationally demanding. In the real time layer, online received information can modify the offline performed controller individualization through a run-to-run strategy: it has been implemented in the adaptation module. Online information includes also manual corrections, which become important sources of knowledge to improve the current control strategy. Moreover, the system is in charge of storing the manual corrections because they can be used to improve individualization techniques.

The real time layer includes also the controller module, which is composed of two sub-modules, the Meal Control Module (MCM) and the Range Control Module (RCM). The aim of the MCM is to perform meal compensation task by exploiting meal announcements and information provided by the patient about the meal quantity estimation. Even if the meal is not announced, the AP system is implemented to recover from PP glucose peaks. The RCM has to maintain the BG level within a safe range according to the MPC approach, employing the information provided by the MCM and by the online adaptation module.

Finally, in order to prevent from possible complications, the continuous time layer includes the diagnostic module to guarantee the safety of the patient. The diagnostic module continuously

monitors the insulin pump and CGM connections statuses and detects possible hardware failures. It contains an internal model used to predict the possible risk of future hypo- or hyperglycemia based on the available CGM measurements and on the suggested insulin by RCM. When the diagnostic module has detected possible dangerous situations for the patient's health, warnings or alarms can inform the patient about the undesired current condition. The alarm systems can be divided in two categories on the base of the method used for the alarm generation: low-threshold detection and prediction. The first notifies the crossing of a critical BG level [84], while the second tries to foresee the hypoglycemia risk to allow the user to act in advance in order to avoid this event [85–88]. The latter typically requires the use of patient models to perform glucose trend predictions. According to the model prediction, a signal is provided to the patient indicating one of three possible levels of hypoglycemic risk. This signal is given in the form of a traffic light: green light codes no risk of hypoglycemia, yellow light indicates that a risk of hypoglycemia is present, and red light means imminent hypoglycemia. The color of traffic light implies an automated attenuation of the insulin pump delivery rate based on risk in a CL setting [85].

Chapter 6

Model Identification

The inter-subject variability characterizing subjects with T1D needs definition of an individualized strategy, which can substantially improve the safety and the effectiveness of the glucose control. From a control perspective, since MPC performance is highly influenced by the quality of the model used for prediction, as described in Section 5.2, an individualized control law achieved by identifying a specific model and including it in the controller could highly improve the performance of the control. A safety improvement is likewise achieved by including this model in an alarm system to predict dangerous situation such as hypoglycemia. Hence, patient-tailored models have a dual purposes: they can be exploited by both the controller and the alarm system. In order to implement monitoring and alarm systems, and to design control law, reliable patient models for the evaluation of future glucose levels are necessary. Currently, the average linear time-invariant metabolic model, which is computed by linearizing the model of the average adult virtual patient of the UVA/Padova simulator, is used in both the MPC and the diagnostic module. Of course, the average linear model could limit the performance, so the research moved to the identification of patient-tailored models. Recently, promising results have been obtained by identifying individualized models using the UVA/Padova simulator, in particular in [5] and [30] different individualization techniques have been studied and compared to the "average" model of the UVA/Padova adult population showing significant improvements in term of prediction ability. In these works, *in silico* data were collected during CL simulations of clinical protocols designed to produce a sufficient input-output excitation without compromising the patient safety. Indeed, sufficiently noisy data are needed to convey enough information about dynamics of the signals. Consequently, identification data coming from CL simulations are preferred with respect to identification data coming from the simulation of the conventional therapy because CL data are usually characterized by more variable inputs.

The identification technique proposed in [5] has been extended to identify a patient-tailored

model from CL data collected in free-living conditions [59]. This set-up is particularly challenging because the identification of reliable models on real-data is more difficult than on simulated data. The glucose-insulin model identification problem is subject to the issue of a poor input excitation, due to the synchronization and proportion of meals and pre-meal insulin boluses. The inputs of the model are the injected insulin and the meals, which have opposite effects on the glucose levels. Indeed, meals and insulin infusions increase and reduce the BG concentration, respectively. In order to compensate the meal intake, an insulin bolus has to be injected in correspondence to the meal. Thus, insulin boluses and meals can be considered simultaneous inputs, and their effects of BG are superimposed. The problem is that these two different inputs cannot be temporarily separeted without compromising the patient safety. This aspect can be manage *in silico* but it can not be avoided in real-life for patient safety reasons and it makes the identification of the effect of each input signal difficult. This techinque will be described in details in Section 6.2.1.

In [30], the successful identification of a good model has been shown to be strictly related with the identification scenario. The identification technique is based on the input-output data, so the selection of the identification data plays a key role. In order to identify models from clinical data, it would have been advisable to reproduce *in vivo* the same protocols used in simulation. Unfortunately, the clinical trials have been not conducted for collecting data for identification purpose. Therefore, the availability of experimental data on long outpatients trials motivated the identification techniques applicable to free-living patient. Since clinical trials have been conducted in free-living conditions, many factors can alter the quality of data. There may be present errors related with meal announcement, i.e. patiens may under/overestimate the amount of announced CHO, which may lead to a wrong computation of the insulin bolus according to Eq. 2.3. The patient can also forget to announce the meal or announces the meal twice. This error is reflected on unexpected chances of glucose level. In order to guarantee the causal relationship between glucose and CHO, this wrong information has to be managed in the identification data. The data portions that contain these input errors have to be discharged because they compromise the identification of glucose response to meal intake. A malfunction in the AP system is another factor that is never present in the simulation data. In order to use data collected during a proper functioning of AP system, the data portions for identification have to belong to time intervals when CL is guaranteed.

The individualization techniques belong to the offline layer of Figure 5.4, where the patient's available information are used both to initialize the controller and to individualize the metabolic model included in the MPC. The identification of an individualized glucose-insulin model is usually based on identification techniques that are computationally demanding, and this is the reason why the individualization module is implemented in the offline layer of the previously

presented modular architecture.

6.1 Model Identification: Impulse-Response (IR) technique

The measurable inputs of the patient model are the injected insulin in pmol/min/Kg, $i(k)$, and the carbohydrates content in mg, $m(k)$. The model output is the glucose concentration measured by the CGM sensor, $CGM(k)$. All these signals are collected every T_s minutes, with $T_s = 5$ minutes. Denoting with $I(z)$, $M(z)$ and $CGM(z)$ the Z-transforms of inputs and output, the model has the following structure:

$$CGM(z) = G_i(z)I(z) + G_m(z)M(z) + E(z) \tag{6.1}$$

where $G_i(z)$ and $G_m(z)$ are transfer functions to be estimated from the data and $E(z)$ is the Z-transform of the residual error $e(k)$. Besides insulin and meal, a number of other unmeasurable factors affect blood glucose concentration, first and foremost physical exercise, but also stress, illness, menstrual cycle, etc. The effect of these unmeasured factors and other unmodeled dynamics are partially accounted for by assuming $e(k)$ to be a colored noise, i.e. assuming that $e(k)$ is correlated with the past errors $e(k-1), e(k-2), \ldots$. Also the spectral characterization of the error has to be estimated from the data.

6.1.1 Continuous-time Impulse Response Model

In order to successfully identify a black-box model, it is necessary to have sufficiently exciting input data and to properly define the order of the system. Impulse signals are the most exciting inputs and are naturally used in continuous time. Hence, following the procedure described in [5], we first identify a continuous-time model to describe the deterministic part of the system:

$$CGM(s) = G_i(s)I(s) + G_m(s)M(s)$$

where $G_i(s)$ and $G_m(s)$ are transfer functions to be estimated from the data, $I(s)$, $M(s)$ and $CGM(s)$ are the Laplace transforms of inputs, $i(t)$ and $m(t)$, and output, $CGM(t)$. Due to the impossibility of performing extensive and potentially dangerous experiments on human subjects, the identification technique is divided in two steps: the first one is entirely developed on the "average" *in silico* patient (Av) of the UVA/Padova simulator [27] with highly exciting input data that could not be applied on human subjects. For example, it is important to have insulin boluses (impulse-like amount) without meal intake as well as uncontrolled meals (meals without

insulin boluses). The outputs of this first step are the two transfer functions $G_i(s)$ and $G_m(s)$ that describe the dynamics of the "average" patient. Starting from the Av model obtained in step 1, the goal of the second step is to identify a specific patient model using patient input-output real data.

6.1.1.1 Step 1: linear average model

The transfer functions for the first step are identified starting from single impulse response experiments performed on the nonlinear model of the Av patient of the UVA/Padova simulator. The simulations lasted 1 day and involved a single meal of 10 g without any insulin boluses to identify G_m and a single insulin bolus without meal administration to identify G_i. The insulin bolus was calibrated to obtain a glucose slope of 10 mg/dl using the CF.
The transfer functions are defined as

$$\begin{aligned} G_i(s) &= \frac{\mu_i}{(1+sT_{i1})(1+sT_{i2})(1+sT_{i3})(1+sT_{i4})} \\ G_m(s) &= \frac{\mu_m}{(1+sT_{m1})(1+sT_{m2})(1+sT_{m3})} \end{aligned} \tag{6.2}$$

where the orders have been selected by a trial and error procedure and the vector of the parameters to be identified is

$$\theta = \begin{bmatrix} \theta_i \\ \theta_m \end{bmatrix}, \theta_i = \begin{bmatrix} \mu_i \\ T_{i1} \\ T_{i2} \\ T_{i3} \\ T_{i4} \end{bmatrix}, \theta_m = \begin{bmatrix} \mu_m \\ T_{m1} \\ T_{m2} \\ T_{m3} \end{bmatrix}$$

The optimal parameters vector, denoted by θ^{Av}, is obtained by two independent constrained optimization problems, one to compute the optimal value of θ_i, denoted by θ_i^{Av}, and the other to compute the optimal value of θ_m, denoted by θ_m^{Av}. Both aim to minimize the Sum of Squares Residuals (SSR) computed as differences between the observed CGM data (CGM) and the CGM estimation ($\widehat{CGM}$) obtained by running a simulation using the model:

$$SSR = ||CGM(t) - \widehat{CGM}(t)||_2 \tag{6.3}$$

with $||x(k)||_2$ the l_2 norm of the signal $x(k)$, namely $\sqrt{\Sigma_{k=1}^{N} x(k)^2}$, where N is the number of data collected every minute. In order to ensure the stability of the system, the minimization problems are performed with the constraint of non-negativity of the time constants performed by adding a barrier function to the SSR.

Since the residuals are described by a nonlinear function of the model parameters, an iterative nonlinear least squares algorithm is used.

The initialization of this procedure involves the estimation of the gain and of the slow time constant of each transfer function, while the other time constants are initialized to 1. The gains are set equal to the Area Under the Curves (AUC) of the impulse response data, while the slow time constant (T_1) is estimated from the final part of these data ($t \in [t_{start}, t_{end}]$), with the assumption that the effects of the other faster time constants are negligible in this data window. In particular, under this assumption, the selected data can be approximated to the impulse response of a first order dynamic system described by

$$CGM(t) = \frac{\mu}{T_1} e^{-\frac{t}{T_1}}, \quad t \in [t_{start}, t_{end}]$$

So that

$$ln\left(CGM(t)\right) = ln\left(\frac{\mu}{T_1}\right) - \frac{t}{T_1}, \quad t \in [t_{start}, t_{end}]$$

Considering the CGM data in the interval $[t_{start}, t_{end}]$ and defining the following vectors

$$\vartheta = \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} = \begin{bmatrix} ln\left(\frac{\mu}{T_1}\right) \\ -\frac{1}{T_1} \end{bmatrix}, Y = \begin{bmatrix} ln\left(CGM(t_{start})\right) \\ \vdots \\ ln\left(CGM(t_{end})\right) \end{bmatrix},$$

$$\phi = \begin{bmatrix} 1 & t_{start} \\ \vdots & \vdots \\ 1 & t_{end} \end{bmatrix}$$

we can describe the vector Y by the model

$$Y = \phi\vartheta + V$$

where the vector V represents the measurements error of the model. Then, the time constant T_1 can be estimated through the linear Minimal Mean Squared Error (MMSE) estimator:

$$\vartheta^{MMSE} = \arg\min_{\vartheta} \quad (Y - \phi\vartheta)^T (Y - \phi\vartheta)$$

Since the identifiability constraint is respected, the global minimum results to be

$$\vartheta^{MMSE} = (\phi^T \phi^{-1}) \phi^T Y$$

and the slow time constant can be obtained as

$$T_1 = -\frac{1}{\vartheta_2^{MMSE}}$$

6.1.1.2 Step 2: linear individualized model

The goal of the second step is to identify an individualized model starting from individual data collected without ad-hoc experiments. An insulin absorption delay (τ_I) has been added to the transfer function G_i in order to consider a physiological delay that can be relevant in some patients. Different values of τ_I have been considered; for each one a set of different optimal individual parameters has to be computed by solving a unique constrained minimization problem for both θ_i and θ_m using the individual data. In fact, differently from the first step it is not possible to separate the effects of insulin and meal. In view of the non-convexity of the optimization problem, the problem of local minima must be addressed by a proper initialization. The initialization can affect the final result. Two choices are adopted in this thesis both leaded by the idea that converging to a local minimum compatible with clinical experience is better than to find a global minimum far from them. Two different initializations were explored: the first one (θ^{Av}) already described in [5], uses all the parameters estimated for the Av patient in step 1, while the second one (θ^{Cp}) complements the time constants of the Av patient with some available individual clinical information to initialize the gains μ_i and μ_m. In particular, μ_i is initialized to the CF and the meal gain μ_m is initialized to $\frac{CF}{CR}$. Given the computational load, the optimization problem has been divided in two parts. The first one has the goal to optimize the parameters θ_i by keeping fixed θ_m to the one chosen as initialization value:

$$\begin{aligned} \theta_{i1}^* = \quad & \arg\min_{\theta_i} \quad SSR \\ & \text{subject to} \quad T_{ij} > 0, j = 1, \ldots, 4 \end{aligned}$$

with $\theta_i^{init} \in \{\theta_i^{Av}, \theta_i^{Cp}\}$, where

$$\theta^{Av} = \begin{bmatrix} \theta_i^{Av} \\ \theta_m^{Av} \end{bmatrix}, \theta^{Cp} = \begin{bmatrix} \theta_i^{Cp} \\ \theta_m^{Cp} \end{bmatrix}$$

This choice is due to the fact that insulin parameters are more variable than the meal ones. In the second part, the insulin parameters are initialized to θ_{i1}^* and the entire θ, meaning both θ_i

and θ_m, is estimated to obtain its most accurate estimation:

$$\begin{aligned} \theta^* = \quad & \arg\min_{\theta} \quad SSR \\ & \text{subject to} \quad T_{ij} > 0, j = 1, \ldots, 4 \\ & \qquad\qquad\quad T_{ml} > 0, l = 1, \ldots, 3 \end{aligned}$$

with $\theta_i^{init2} = \theta_{i1}^*$, $\theta_m^{init2} \in \{\theta_m^{Av}, \theta_m^{Cp}\}$ and $\theta^* = \begin{bmatrix} \theta_{i2}^* & \theta_{m2}^* \end{bmatrix}'$. At the end of this procedure the transfer functions (6.2) are completely individualized.

6.1.2 Discrete-time Model

The continuous-time model identified in the previous section is discretized via zero-order hold method obtaining $G_i(z)$ and $G_m(z)$. Then, we identify the stochastic part of the model (8.1) by describing the residual error $e(k)$ as an AR process of order n:

$$e(k) = a_1 e(k-1) + \cdots + a_n e(k-n) + \epsilon(t)$$

with $\epsilon(t)$ a zero-mean white noise with variance λ. The parameters $a_1, \ldots, a_n$ and λ are estimated from the data by minimizing the 1 steps ahead prediction. The complexity of the AR model is fixed a priori and chosen by trial and error to $n = 5$.

6.2 Data

6.2.1 Experimental set-up

The dataset used was collected during experiments involved in the "AP@home" project [89] in 2015. The clinical trials took place in three clinical centres of Padova (Italy), Montpellier (France) and Amsterdam (Netherlands) [59]. 18 patients have been enrolled in a 1-month trial aimed to test the day-and-night use of a CL controller implemented in an AP system in free-living conditions. The baseline characteristics of these subjects are presented in Table 6.1. The considered clinical trial has been conducted through a fully automatic closed-loop control [59]. The patients worn the AP prototype consisting of an suitably modified android smartphone (the DiAs platform, [90]), communicating wirelessly with the G4 Platinum CGM system, Dexcom Inc. and the AccuCheck Spirit Combo insulin pump, Roche Diagnostic. The computational unit run the MPC controller described in [41]. Since data were collected in real-life condition,

Variable	*Study population (n=18)*		
	AMS (n=7)	MPL (n=4)	PAD (n=7)
Age [years] (SD)	37.71 (12.21)	51.50 (5.80)	46.85 (9.42)
Sex Male (n)(%)	3 (42.86)	2 (50)	2 (28.57)
Sex Female (n)(%)	4 (57.14)	2 (50)	5 (71.43)
Body mass index (BMI) [kg/m^2] (SD)	25.12 (4.47)	25.63 (4.16)	24.06 (2.94)
HbA_{1c} [%] (SD)	7.60 (0.73)	7.44 (0.31)	7.54 (0.42)
Diabetes duration [years] (SD)	21.71 (12.51)	33.75 (10.43)	29 (12.63)
Insulin delivery mode, CSII [n] (%)	7 (100)	4 (100)	7 (100)
Duration of CSII use [years] (SD)	6.42 (3.50)	9.25 (3.77)	10.57 (7.02)
Total daily insulin dose [U/kg] (SD)	0.59 (0.04)	0.52 (0.10)	0.46 (0.11)

TABLE 6.1: Baseline characteristics of patients who participated in the extension. For categorical variables, n (%) is presented. For continuous variables, mean (SD) is presented. CSII, Continuous Subcutaneous Insulin Infusion.

patients have received appropriate training for the safe use of the CL insulin delivery system. Carbohydrates ingested at meal time or for snack and those used to treat hypoglycemic episodes were manually entered by the patient into the system. Capillary blood glucose measurements, obtained by pricking patient's finger (Self-Monitoring Blood Glucose measurements, SMBG) were performed for CGM calibration, at meal and to confirm hypo- or hyperglycemia detected by the CGM sensor.

It should be noted that this prototype was not specifically designed to collect data for model identification, posing a number of technical issues regarding device synchronization, completeness of stored data and reliability of patient's provided information. Furthermore, a few malfunctions occurred during the trial hampering the reliability of the associated data portions, as shown in Figure 6.1. Hence, a careful data selection phase has been performed before the model identification. Moreover, since a clinical trial under free-living condition means that the patient has a normal life without any type of restriction, the original data need to undergo preprocessing before the identification.

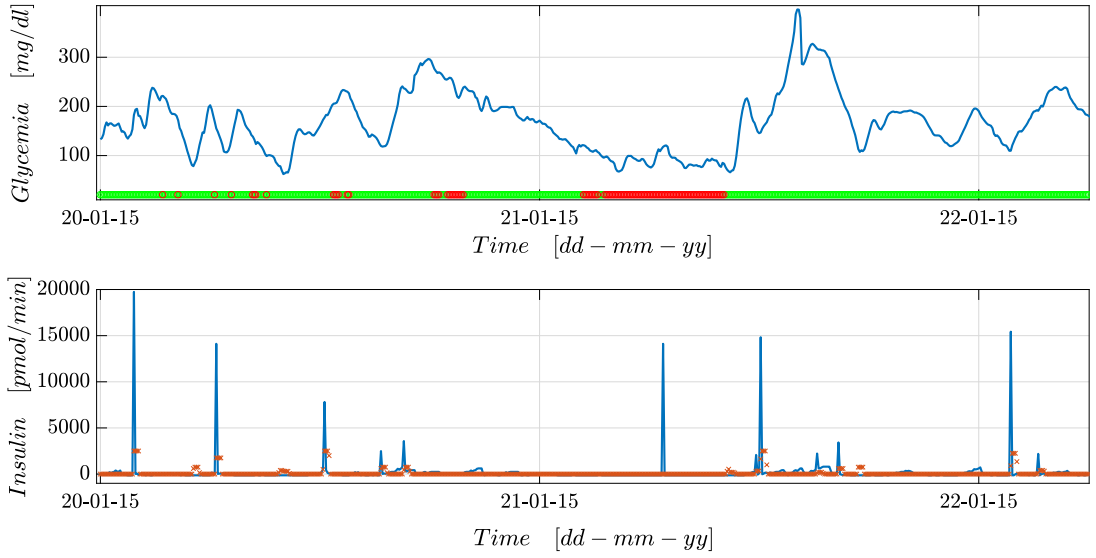


FIGURE 6.1: A portion of the clinical trial. In the first subplot, the blue line represents the CGM data, the green circles corresponds to the times in which the system was in CL mode, whereas the red circles emphasize some system malfunctions. In the second subplot, the orange crosses represent the meal fluxes and the blue signal is the infused insulin.

6.2.2 Data Preprocessing

Current CGM sensors have to be calibrated two times/day by using SMBG measurements to produce reliable glucose readings. Imperfect calibration induces a systematic distortion in CGM measurements as illustrated in Figure 6.2 and modeled in [91]. Denoting with $CGM_{\text{b.p.}}(t)$ the CGM before preprocessing, a simplified version of the model in [91] is:

$$CGM_{\text{b.p.}}(t) = \alpha g(t) + \beta + \gamma t + e_{CGM}(t)$$

where $g(t)$ is the true glucose concentration, $e_{CGM}(t)$ is a colored noise and α, β, γ are the "decalibration" model parameters. The interstitial glucose concentration $g(t)$ can be related to the blood glucose concentration through a first order continuous-time dynamical system with unit gain and time constant τ. The parameters $\alpha, \beta, \gamma, \tau$ abruptly change every time a calibration is performed.

Apparently, this distortion can affect the estimate of model coefficients and introduce spurious jumps and additional dynamics. To mitigate these artifacts, we employed a preprocessing algorithm known as "retrofitting" [92], that retrospectively corrects this calibration-induced distortion by leveraging on the additional SMBG measurements collected during the trial. Here we only illustrate the effect of the algorithm in Figure 6.2, while we refer the interested reader to the original paper for more details.

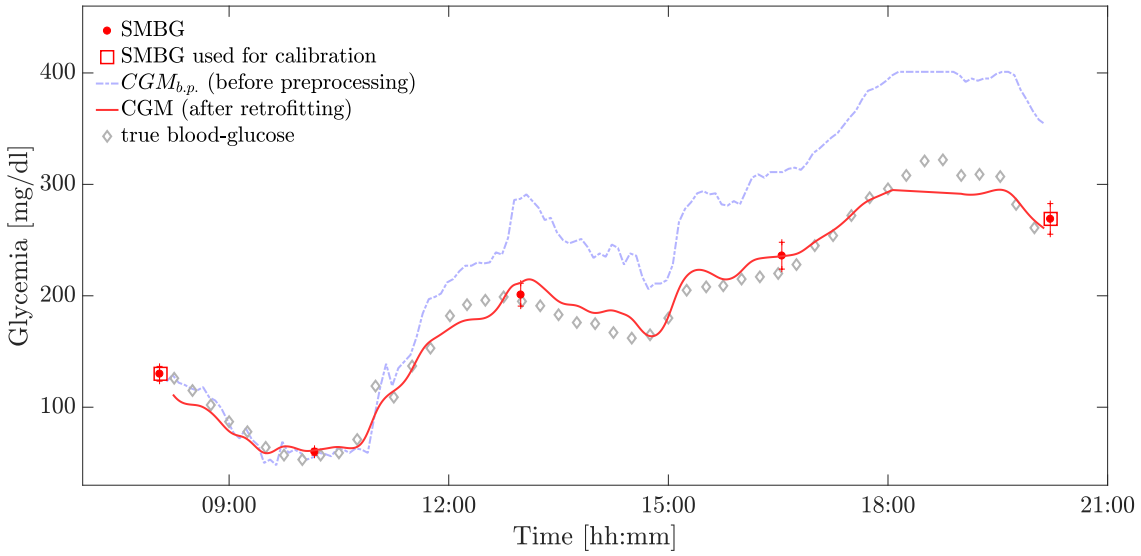


FIGURE 6.2: Illustration of CGM data pre-processing performed by the retrofitting algorithm, taken from [93]. Due errors and uncertainty in the calibration process the CGM measurement (dashed blue line) overestimates the true blood glucose (gray diamond, not available in our dataset). The retrofitting algorithm, leveraging on a the additional SMBG measurements collected during the trial (red dots), compensates for calibration error and the output of the method (red solid line) is closer to the true glucose concentration.

From now on, $CGM(t)$ always refers to the CGM trace *after* the preprocessing by retrofitting. It should be noted that the retrofitting algorithm improves the accuracy of the CGM, but it does not solve the issue of data reliability previously mentioned (e.g. due to AP malfunctioning or human errors on patient-provided data). Thus, the extraction of a "clean" data portion remains key also after preprocessing.

6.3 Performance metrics

To assess the efficacy of an identified model we compare the model predictions of future CGM ($\widehat{CGM}$) against its actual values ($CGM(t)$). Since the identified model is meant to be used in our MPC controller [41] with a specific Prediction Horizon (PH) the prediction capabilities over finite PHs have to be evaluated. In particular, the accuracy of the model predictions is assessed by considering various PHs. The considered PHs are expressed in terms of steps, where each step corresponds to the sampling time, $T_s = 5$ minutes. In this work, the authors consider PHs from 1 step, i.e. 5 minutes, to 12 steps, i.e. 60 minutes, where one hour can be considered a sufficiently long PH with respect to glucose dynamics and coincides with the PH of the MPC presented in Section 5.2. Moreover, simulation capabilities of the identified models are evaluated by testing the prediction capabilities over an infinite PH. Hence, the vector of all possible PHs, PH can be defined as follows $PH = [T_s, 2T_s \ldots, 12T_s, +\infty] = [5, 10, \ldots, 60, +\infty]$ minutes.

In details, let us denote with $\widehat{CGM}(t|t-PH)$ the PH-steps ahead prediction of a model, i.e. the prediction obtained by exploiting past glucose values up time $t-PH$, $CGM(t-PH)$, $CGM(t-PH-1)$, ... and inputs up to time t, $i(t)$, $i(t-1)$, ..., $m(t)$, $m(t-1)$, ... Furthermore, let us denote with $PH = +\infty$ the glucose simulation, i.e. the output of the model when fed with the inputs $i(t)$, $i(t-1)$, ..., $m(t)$, $m(t-1)$, ... without taking advantage of any of the measure outputs.

The signal $\widehat{CGM}(t|t-PH)$ and CGM are compared using the metrics listed below. The starting point is Root Mean Square Error (RMSE) defined as:

$$RMSE(PH) = \frac{1}{N}||CGM(t) - \widehat{CGM}(t|t-PH)||_2$$

where we denote with $||x(t)||_2$ the l_2 norm of the signal $x(t)$, namely $\sqrt{\Sigma_{t=1}^{N} x(t)^2}$, N being the length of the dataset.

RMSE assesses the variance of the prediction error: the larger it is, the poorer is the prediction. Instead of presenting this absolute quantity, we report two normalized versions commonly used in system identification [58, 94].

Metric 1: Index of fitting (FIT).

Defined as

$$FIT(PH) = 100\left(1 - \frac{||\widehat{CGM}(k|k-PH) - CGM(k)||_2}{||CGM(k) - \overline{CGM}||_2}\right)$$

where $\overline{CGM}(t)$ is the sample mean of the glucose signal. FIT is equal to 100% if and only if $\widehat{CGM}(k) = CGM(k)\ \forall k = 1, \ldots, N$ (perfect prediction), and smaller than 100% otherwise. Note that FIT can become negative.

Metric 2: Coefficient of Determination (COD).

Defined as

$$COD(PH) = 100\left(1 - \frac{||\widehat{CGM}(k|k-PH) - CGM(k)||_2^2}{||CGM(k) - \overline{CGM}||_2^2}\right)$$

Similarly to FIT, COD is equal to 100% for perfect predictions, smaller than 100% and possibly negative otherwise.

Metric 3: Pearson's correlation coefficient ρ.

Defined as

$$\rho(PH) = \frac{\sum_{t=PH}^{t_{max}} (CGM(t) - \overline{CGM})(\widehat{CGM}(t|t-PH) - \overline{\widehat{CGM}}(t|t-PH))}{||CGM(t) - \overline{CGM}||_2 \cdot ||\widehat{CGM}(t|t-PH) - \overline{\widehat{CGM}}(t|t-PH)||_2}$$

with $\overline{\widehat{CGM}}(t|t\text{–}PH)$ being the sample mean of the predicted CGM.

All the metrics mentioned above are function of the prediction horizon PH. Furthermore, the average value of each metric ($\overline{FIT}$, $\overline{COD}$, $\overline{\rho}$) over the considered PH was used as the primary outcome to evaluate the models. In addition to the performance metrics for individuals, indeces of performance for a patient cohort are also needed. Hence, for every PH, a secondary outcome is computed as the average value of each metric ($\overline{\overline{FIT}}$, $\overline{\overline{COD}}$, $\overline{\overline{\rho}}$) for every PH over the considered population. For each index, the mean ($\pm$ Standard Deviation (SD)) for normally distributed data or median [25^{th} – 75^{th} percentiles] for the non Gaussian case have been computed over the patient cohort as cumulative index of the entire population.

6.4 Identification Results

In order to verify the glucose prediction capabilities of the presented identification approach, an identification process has been performed on a single *in vivo* patient. For this first case-study, trials data belonging to a patient affiliated to the Amsterdam clinical centre have been used [59]. The aim is to show the prediction capabilities of the patient-tailored model identified by the technique described in Section 6.1.1 with respect to the average model, currently used in the controller.

Before proceeding with the achieved results, an important remark concerns the identification dataset. In [30], the authors have shown how much fundamental is the role of the identification scenario for the identification of a reliable model. Unfortunately, the clinical trials have not been conducted for collecting data for identification purpose purposes and, for this reason, a preliminary phase consisted in finding suitable scenarios that allow to capture the correct dynamics of the patient. The choice of the identification scenario consists of a total of 49 hour of data, selected randomly among the available ones, and containing two (not necessarily consecutive) data sub-portions, used to estimate the parameters of the input-output relation (30 hours) and the stochastic part (19 hours). Once the identification scenario is defined, ten models of the patient have been identified using all the possible combinations of the initializations, $\theta^{init} \in \{\theta^{Av}, \theta^{Cp}\}$, and delays, $\tau_I \in \{0, 15, \ldots, 60\}$ as reported in Table 6.2. The identified models are tested on two testing protocols to assess model prediction. Firstly, a short 1-day scenario has been considered. This dataset is completely disjoint from the training-set and manually selected by visual inspection among those free from technical issues of the AP system including infusion set failures and possible errors on patient provided information. Then, the entire 1-month dataset containing all data from the trial for the selected patient is considered

without any manual ad-hoc exclusion. The obtianed results, which have been published in [61], are particularly relevant due to the fact that the entire 1-month dataset is really challenging because it includes all the problems experienced during the whole trial

Model	Parameters		
	μ_i^0	μ_m^0	τ
M1	μ_i^{Av}	μ_m^{Av}	0
M2	μ_i^{Av}	μ_m^{Av}	15
M3	μ_i^{Av}	μ_m^{Av}	30
M4	μ_i^{Av}	μ_m^{Av}	45
M5	μ_i^{Av}	μ_m^{Av}	60
M6	μ_i^{Cp}	μ_m^{Cp}	0
M7	μ_i^{Cp}	μ_m^{Cp}	15
M8	μ_i^{Cp}	μ_m^{Cp}	30
M9	μ_i^{Cp}	μ_m^{Cp}	45
M10	μ_i^{Cp}	μ_m^{Cp}	60

TABLE 6.2: List of all considered models.

The performance metrics of the identified models on both testing datasets are reported in Tables 6.3 and 6.4, respectively. In both Tables the average value of the different metrics are computed considering the prediction horizons (PH) varying from 1 to 12 (i.e. from 5 to 60 minutes). Moreover, the performance of the "average" *in silico* model (Av) is also reported for comparison with the patient-tailored models.

It is evident that the technique is rather independent from algorithm initialization and insulin delay estimation. The best patient-tailored model is M7 for both test-sets, it uses the clinical patient information as initial condition and assumes the insulin delay $\tau_I = 15$ minutes. M7 is able to improve the performance metrics of the Av model of almost three times in terms of $\overline{FIT}$ and $\overline{COD}$ (247% $\overline{FIT}$, 245% $\overline{COD}$) and of 22% in terms of $\overline{\rho}$. Figure 6.3 reports the performance metrics (FIT (a), COD (b) and ρ (c)) as a function of PH for all the considered models with the daily testing dataset. By increasing the PH, the FIT, COD and ρ are expected to decrease. The figure clearly shows that the performance of the individualized models decreases significantly less than the Av model. If the entire trial is considered, the best patient-tailored model remains M7. The improvements of M7 with respect to the Av model increase with respect to 49-hour scenario; they are of 275% in terms of $\overline{FIT}$, 335% in terms of $\overline{COD}$ and 27% in terms of $\overline{\rho}$. Figure 6.4 reports the performance metrics (FIT (a), COD (b) and ρ (c)) as a function of PH

Model	$\overline{FIT}$ [%]	$\overline{COD}$ [%]	$\overline{\rho}$
M1	65.8 (± 22.6)	83.6 (± 16.2)	0.92 (± 0.08)
M2	65.9 (± 22.5)	83.8 (± 16.1)	0.92 (± 0.08)
M3	66.9 (± 21.7)	84.7 (± 15.1)	0.93 (± 0.07)
M4	65.4 (± 22.6)	83.4 (± 16.4)	0.92 (± 0.08)
M5	67.6 (± 21.0)	85.5 (± 14.3)	0.93 (± 0.07)
M6	65.6 (± 22.8)	83.4 (± 16.5)	0.92 (± 0.08)
M7	**68.6 (± 20.3)**	**86.4 (± 13.2)**	**0.93 (± 0.07)**
M8	65.6 (± 22.7)	83.5 (± 16.4)	0.92 (± 0.08)
M9	67.1 (± 21.4)	85.0 (± 14.7)	0.93 (± 0.07)
M10	67.0 (± 21.6)	84.9 (± 14.9)	0.93 (± 0.07)
Av	**19.8 (± 34.0)**	**25.1 (± 52.4)**	**0.76 (± 0.15)**

TABLE 6.3: Performance metrics of the identified models for patient 1 on a daily scenario.

Model	$\overline{FIT}$ [%]	$\overline{COD}$ [%]	$\overline{\rho}$
M1	57.7 (± 26.6)	75.6 (± 22.6)	0.88 (± 0.11)
M2	57.1 (± 27.1)	74.8 (± 23.4)	0.88 (± 0.11)
M3	58.0 (± 26.3)	76.0 (± 22.1)	0.89 (± 0.11)
M4	55.9 (± 27.6)	73.5 (± 24.4)	0.87 (± 0.12)
M5	58.3 (± 26.0)	76.4 (± 21.8)	0.89 (± 0.10)
M6	57.4 (± 26.9)	75.3 (± 23.0)	0.88 (± 0.11)
M7	**59.2 (± 25.4)**	**77.4 (± 20.7)**	**0.89 (± 0.10)**
M8	57.0 (± 27.1)	74.8 (± 23.4)	0.88 (± 0.11)
M9	57.9 (± 26.4)	75.9 (± 22.3)	0.88 (± 0.11)
M10	58.4 (± 26.0)	76.5 (± 21.7)	0.89 (± 0.11)
Av	**15.8 (± 35.0)**	**17.8 (± 56.4)**	**0.71 (± 0.20)**

TABLE 6.4: Performance metrics of the identified models for patient 1 on the entire trial.

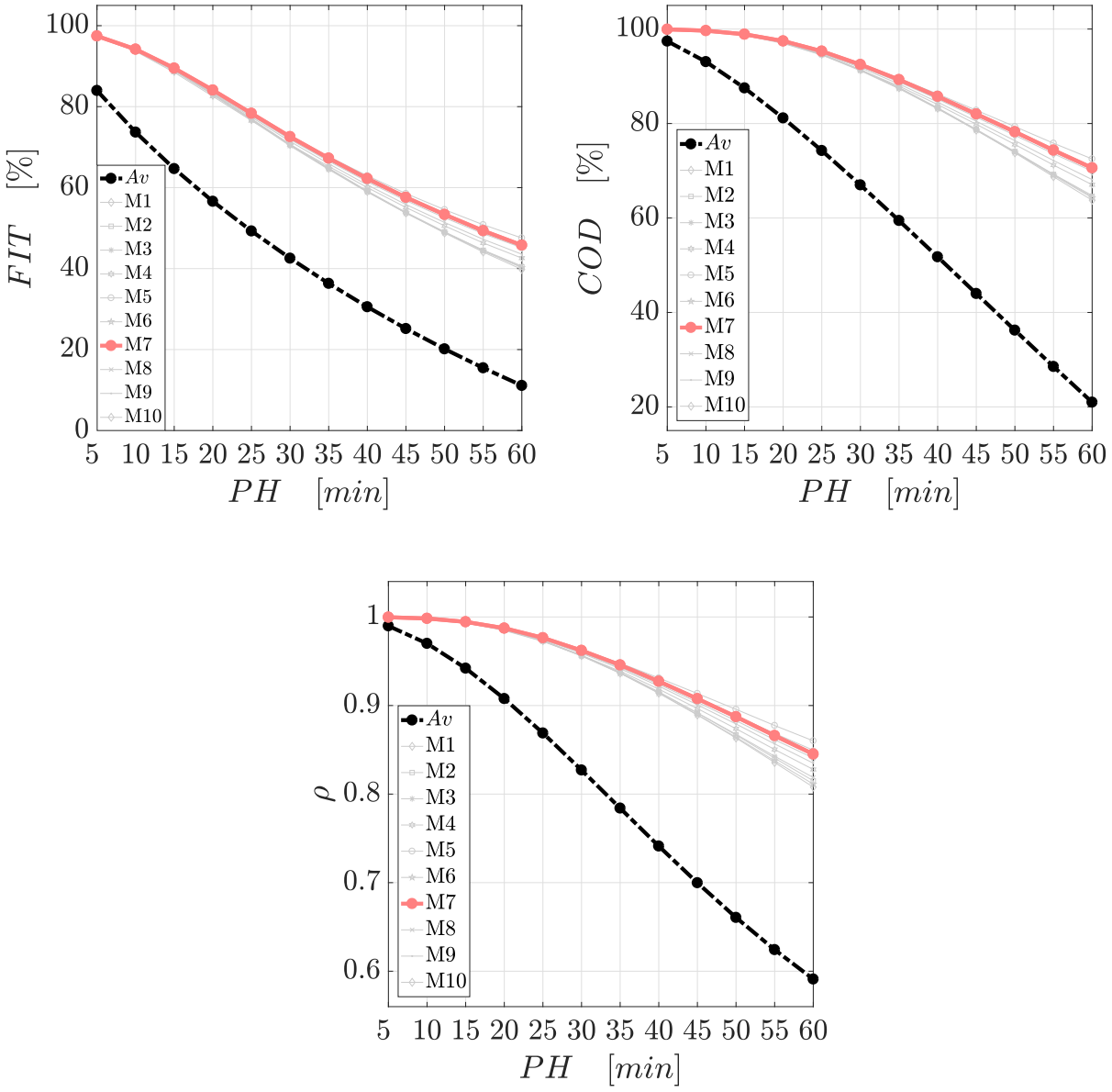


FIGURE 6.3: A comparison of the prediction performance of the proposed technique compared to the "average" patient of the *in silico* population in term of FIT (a), COD (b) and ρ (c) on the 49-hour scenario.

for all the considered models with the second testing scenario. The results obtained with the first scenario are confirmed also by considering the entire trial; this result is very important considering the large impact of the many confounding factors and technical issues that affect the entire trial testing dataset. In general, both initializations show a good performance with a preference for the clinical one that includes additional information about the patient; the introduction of an a priori insulin delay τ_I seems to improve the prediction ability of the model, particularly with a value of 15 minutes. The improvement of the prediction capability of the individualized model with respect to "average" model used to synthetize this MPC algorithm

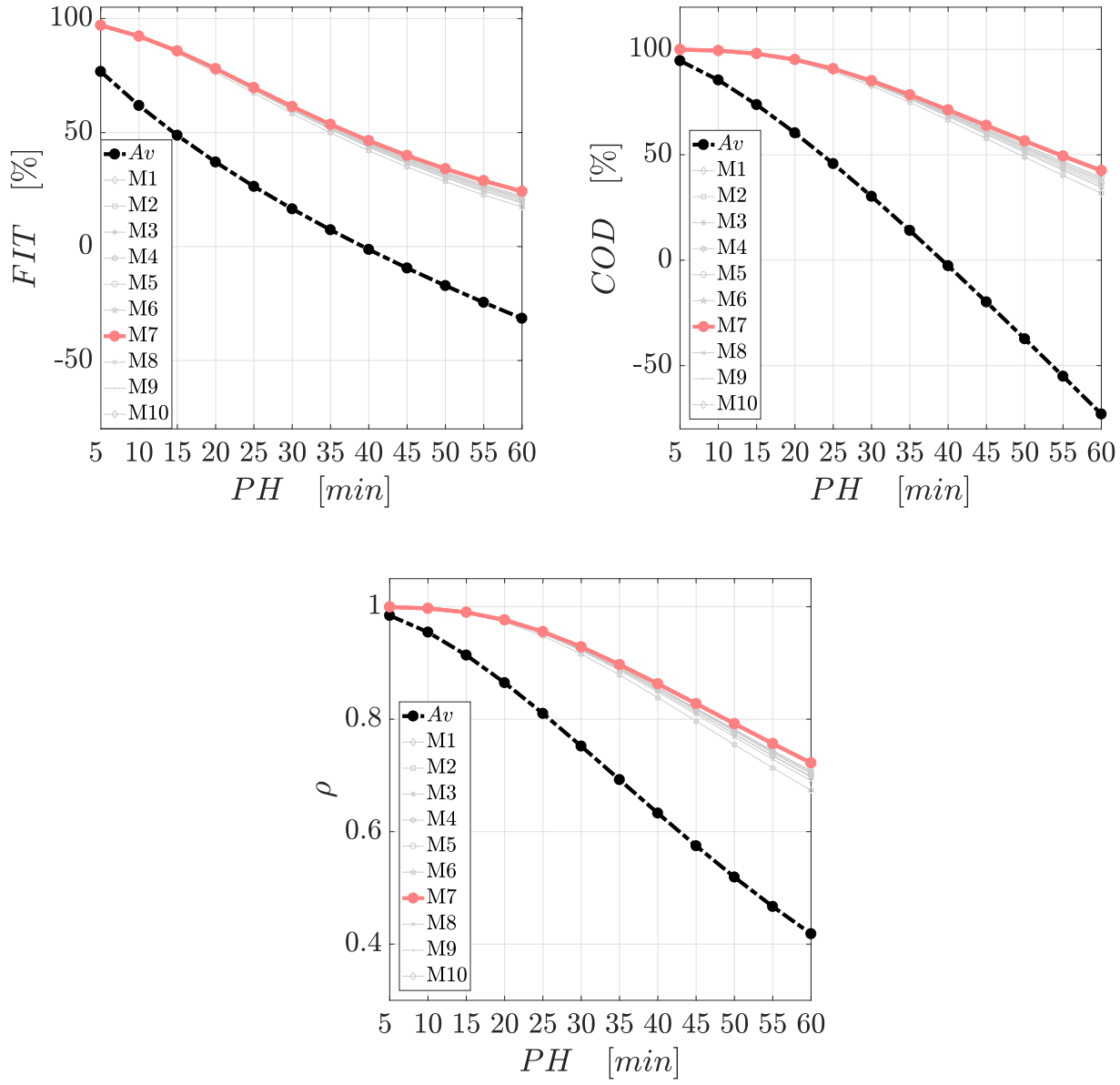


FIGURE 6.4: A comparison of the prediction performance of the proposed technique compared to the "average" patient of the *in silico* population in term of FIT (a), COD (b) and ρ (c) on the entire trial.

paves the way for a new generation of individualized glucose control strategies for AP.

6.5 Individualized models of population

Given the good results obtained on a single patient, the technique has been extended to all the 7 T1D patients studied at Amsterdam clinical centre. The aim is to show the ability of the IR

technique to adapt to different subjects, showing its effectiveness in front of inter-subject variability that is the main reason for the need of patient-tailored models. Patients inter-variability is caused by different biological characteristics like medical history, gender, age, weight, height and metabolism and implies that different patients have different insulin responses. Therefore, the identification of reliable models on real data is particularly difficult. The available patient clinical data plays a key role to identify individualized models and to deal with the patients inter-variability. Since clinical data have been collected in free-living conditions, an additional difficulty is to take into account that each patient can experience different problems during the trial and different factors of uncertainty can affect their data. Some examples are the physical exercise or differences in daily activities, human errors in patient-provided information or technical issues affecting the AP prototype adopted during the trial.

Among the possible initializations for the optimization, for the population study the initialization that uses some available clinical parameters to adapt the two gains μ_i and μ_m has been considered in order to define the initialization vector since it showed the best performance in the case-study presented in Section 6.4. Regarding the insulin absorption delay (τ_I), the parameter has been set constant to 15 minutes on the basis of the results obtained in Section 6.4. For each patient of the Amsterdam clinical center an individualized model has been identified by using an identification protocol that consists of 49 hours of data picked up randomly among the available ones. A unique testing scenario is considered, that is the entire trial. The set intervals used for the identification have not been considered to compute the performance metrics of these models, for this reason the performance of Patient #1 are slightly different with respect to the ones showed in Table 6.4.

Patient	$\overline{FIT}$	$\overline{COD}$	$\overline{\rho}$
Patient #1	58.72	76.95	0.89
Patient #2	60.56	79.55	0.9
Patient #3	74.42	91.35	0.96
Patient #4	67.92	86.23	0.93
Patient #5	70.77	88.18	0.94
Patient #6	45.68	58.49	0.83
Patient #7	68.12	86.37	0.93
Population mean (±SD)	63.74 (± 9.67)	81.02 (± 11.11)	0.91 (± 0.04)

TABLE 6.5: Predictive performance of patient-tailored models identified for the 7 patients. All indices are normally distributed, so mean ± SD are reported for the entire population.

The results reported in Table 6.5 includes the average values of the different metrics computed by considering the prediction horizons varying from 1 to 12 (i.e. from 5 to 60 minutes). From the results it is possible to note that the $\overline{FIT}$ computed on the entire population remains above 60% (63.74%) with a $\overline{COD}$ of 81.02%. The results are in line with those in Section 6.4. Patient #1 is the test-case used in Section 6.4 to identify a patient-tailored model. It is worth to note that the training set interval was so limited with respect to the entire trial that the performance is not significantly affected by excluding it from the testing dataset: $\overline{FIT}$ passes from 59.2% to 58.72%, $\overline{COD}$ from 77.4% to 76.65, and $\overline{\rho}$ remains constant. One patient (Patient #6) performed worse than the others. However, considering the large changes of the patients' habits and the time-varying nature of the system under study, all the results seem acceptable. The

PH	$\overline{\overline{FIT}}$	$\overline{\overline{COD}}$	$\overline{\overline{\rho}}$
1	96.83 (± 0.84)	99.91 [99.86, 99.94]	1 [1,1]
2	92.47 (± 1.69)	99.41 (± 0.27)	1 [1,1]
3	86.84 (±2.93)	98.19 (±0.8)	0.99 [0.99, 0.99]
4	80.34 (±4.45)	95.97 (±1.81)	0.98 (±0.01)
5	73.45 (±6.21)	92.62 (±3.45)	0.96 (±0.02)
6	66.53 (±8.14)	88.23 (±5.79)	0.94 (±0.03)
7	59.79 (±10.21)	82.94 (±8.89)	0.92 (±0.04)
8	53.35 (±12.36)	76.93 (±12.71)	0.89 (±0.06)
9	47.24 (±14.52)	70.36 (±17.18)	0.86 (±0.07)
10	41.43 (±16.63)	63.32 (±22.19)	0.83 (±0.08)
11	35.92 (±18.65)	55.96 (±27.6)	0.8 (±0.1)
12	30.72 (±20.56)	48.38 (±33.29)	0.76 (±0.11)

TABLE 6.6: Predictive performance of patient-tailored models identified for the entire population across the PH. The indices normally distributed are reported in terms of mean ± SD and the others as median [25^{th}, 75^{th} percentiles].

results reported in Table 7.3 show the population values of $\overline{\overline{FIT}}$, $\overline{\overline{COD}}$, and $\overline{\overline{\rho}}$, which have been computed by considering each single prediction horizons varying over the patient cohort from 1 step, i.e. 5 minutes, to 12 steps, i.e. 60 minutes, where one hour can be considered a sufficiently long PH with respect to glucose dynamic. The performance decreases with the increase of the PH as expected.

The IR technique has resulted to be acceptable and suitable to identify patient-tailored models from real-life data. This technique is rather simple, and at the same time flexible to be extended

to be used on free-living data even though a very tough training dataset. The identified models showed satisfactory results in prediction capabilities as shown in Table 6.5. There is only one patient (Patient #6) whose prediction results are worse than the average results obtained on the entire population. This can be addressed to the complex and tricky dataset, which requires a detailed data analysis to improve model identification.

Chapter 7

Individualized models for alarm system

One of the main concerns of glucose control is the avoidance of the hypoglycemia events. In this perspective, in the last year several alarm systems have been studied to predict dangerous situation such as hypoglycemia and to allow the patient to prevent them. The alarm systems can be divided in two categories on the base of the method used for the alarm generation: low-threshold detection and prediction. Hypoglycemia alarms based on low-threshold detection alert the user of the risk of hypoglycemia when the glucose concentration falls below a certain threshold [84]. Predictive alarms assess the risk of hypoglycemia on the basis of the estimated evolution of glucose concentration. If the future glucose level can be predicted, the hypoglycemic event can be avoided by taking a rescue action, i.e. suspending the insulin infusion [85–88]. The core of an hypoglycemia predictive alarm system is the model which provides glucose predictions and allows to forewarn the patient in case of potentially dangerous events ahead in time. Different glucose-insulin models can be used in the alarm systems: for example, minimal, maximal or black-box linear models. The first class is accepted both as clinical tool and as an approach to understand the composite effects of insulin on glucose tolerance [95, 96]; the second is able to better represent the significant inter-patient variability that characterizes T1D population, e.g. the one included in the UVA/Padova simulator [24, 34]. The third class includes models typically identified from real-life data. Although the identification of reliable models on real data is more difficult, an individualized model that describes the glucose-insulin dynamics of the specific patient has better performance with respect to a population average model, as shown in Section 6.4. In this chapter a method to use an individualized models of the patient to develope and validate Individualised hypoglycemia Predictive Alerts (IHPAs) on a rather long period (1 month) is presented. Once again, firstly the performance of the proposed system is compared

for a single case-study to the Hypoglycemia Prediction State of Art (HPSA). It is set to the algorithm for hypoglycemia prevention [85] used in [59, 75]. Then, the technique is validated for all the 7 patients studied at the University of Amsterdam Medical Centre and the performance achieved are reported for each single patient and as mean of the entire cohort.

7.1 Individualised Hypoglycemia Predictive Alert (IHPA)

At time k^*, IHPA algorithm exploits the model to predict the future glycemia during a Prediction Window (PW),i.e. from k^*+1 to k^*+PW (i.e. $PH = PW$) assuming that the all future insulin administration will be stopped, i.e. setting

$$i(k^*) = i(k^* + 1) = \cdots = i(k^* + PW) = 0$$

If for at least N_{samp} samples, the predicted glucose values are below the threshold G_{hypo}^{IHPA} in the PW, an hypo alarm is issued. In fact, according to the model, the impending hypoglycemia can not be avoided only by suspending insulin and, hence, the subject with T1D has to take some rescue carbohydrates. This alarm system is designed to be used in conjunction with an AP, where the insulin suspension is automatically performed by the system to optimize the glucose profile. So, the main purpose of this alarm is to notify when the AP can not avoid the hypoglycemia by itself by only stopping the insulin delivery. If data are missing in the previous hour for more than 20 minutes, insufficient data are available for a reliable prediction and the IHPA is preventively shut off. The parameters that characterize the algorithm have been set to $PW = 8$ (i.e. 40 minutes), $G_{hypo}^{\text{IHPA}} = 70$ mg/dl and $N_{samp} = 2$ (i.e. 10 minutes). The prediction capabilities of the patient-tailored models reported in Table 6.5 bodes well to build a reliable alarm system. In Section 6.4, the individualized model of Patient #1 has showed a significant gap in predictive performance computed on real-life data with respect to the average model. Moreover, Table 6.6 shows that the individualized models of the entire population have satisfactory performance in prediction, in particular considering PHs shorter than 40 minutes. An average value of FIT and COD of 76.20 % and 91.77 % have been achieved by averaging the FIT and COD values over the PHs from 1 (5 minutes) to 8 (40 minutes) on the all 7 patients. All this, together with the fact that a FIT value $> 50\%$ and a COD value > 70 % have been obtained in the 40-minute ahead prediction is expected to be good in improving the reliability of a safety system.

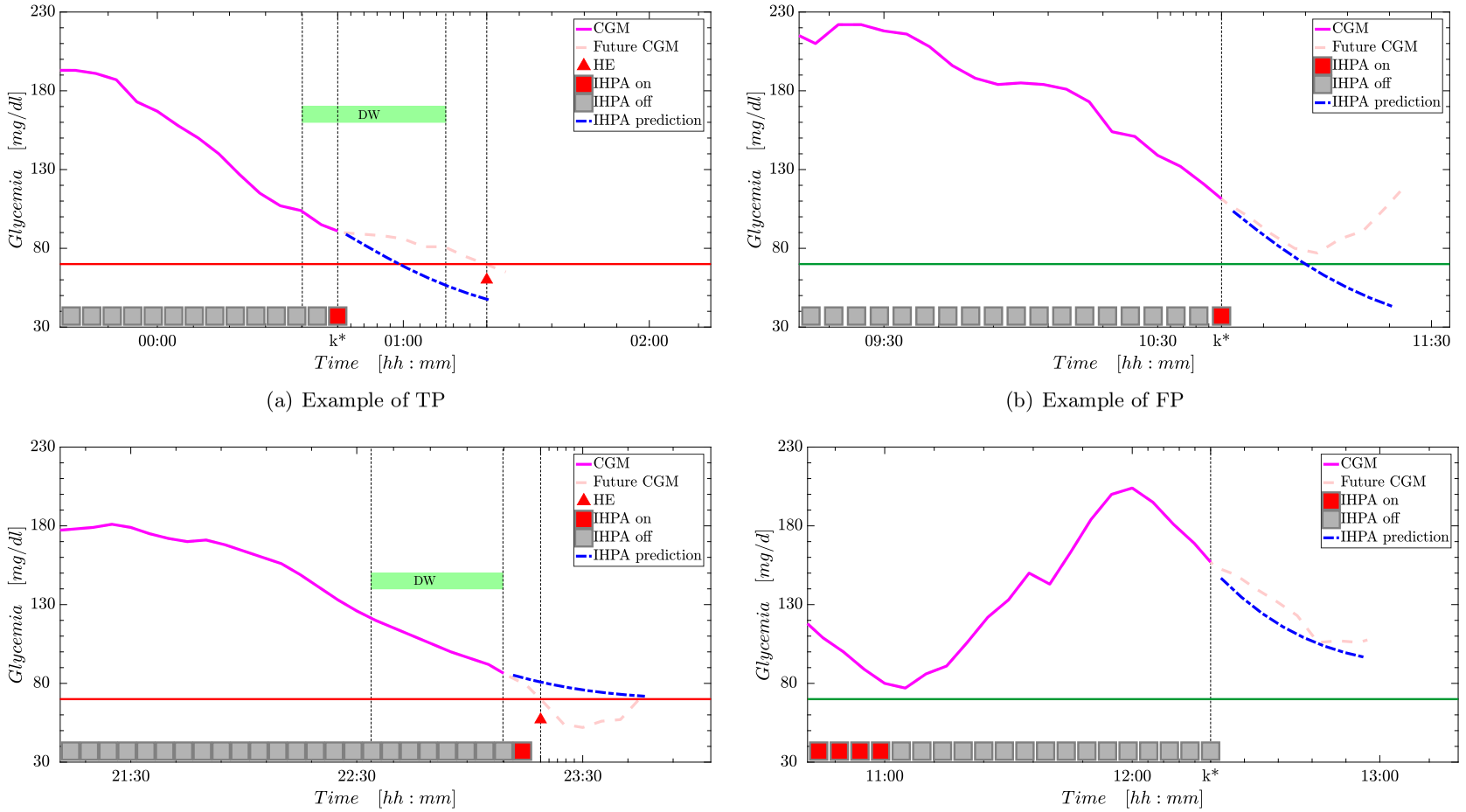
7.2 Performance Metrics for Hypoglycemia Alarms Evaluation

Before defining the performance metrics, the following definitions are introduced:

- *Hypoglycemia Event* (HE): it starts when the glycemia of the patient falls below the threshold G_{hypo} and it ends when the glycaemia remains above the threshold for more than t_{rec} minutes.
- *True Positive* (TP): when an hypoglycemia event occurred at time k_h and an alarm is activated by the algorithm in the Detection Window (DW) $[k_h - DW_s, k_h - DW_e]$, as shown in Figure 7.1(a). Note that the hypoglycemia event has to be notified at least DW_e minutes before its occurrence and that the alarm is considered correct if clearly related to the event (not too far in the past too far by DW_s).
- *False Positive* (FP): when an alarm at k^* is activated and no hypoglycemia occurred in the window $[k^*, k^* + DW_s]$, as shown in Figure 7.1(b). Given that the alarm has to inform of an upcoming hypoglycemia not avoidable without the administration of carbohydrate by the patient, an episode is not counted as a FP if a meal or an hypotreatment are administrated in the future $[k^*, k^* + DW_s]$. Note also that "late" alarms, i.e. alarms occurring after DW_e are not associated to a FP event even if they do not count as TP.
- *False Negative* (FN): when an hypoglycemia event occurs and no alarm is issued in the DW interval, as showed in Figure 7.1(c). The DW does not contain any alarm, in fact, the alarm system turns on the alarm too late, just after the DW end resulting in a FN.
- *True Negative* (TN): when no alarm is issue and no hypoglycemia occurred in the window $[k^*, k^* + (DW_s - DW_e)]$, as shown in Figure 7.1(d).
- *Not Evaluable Hypoglycemia Event* (NEHE): an hypoglycemia event occurred just after a period characterized by too many missing data and/or system failures to consider the data reliable. In particular, this happens if in the last DW samples there are less than N_A values. Note that if system failures occurred during the trial but CGM, insulin and meal data are available, IHPA runs correctly while the alarms of HPSA are not available.

The parameters related to these definitions have been set $t_{\text{rec}} = 20$ minutes, $DW_s = 9$ (i.e. 45 minutes), $DW_e = 2$ (i.e. 10 minutes) and $N_A = 3$.

The performance of the method is evaluated through the following metrics:



(a) Example of TP

(b) Example of FP

(c) Example of FN

(d) Example of TN

FIGURE 7.1: Example of TP, FP, FN and TN of prediction by IHPA. In each panel, the CGM measured before the considered time instant k^* (magenta), the future real CGM measurements (dashed pink) and the glucose prediction obtained via the model presented in 6.4 used in the IHPA (dashed blue) are shown. At each sample time, the IHPA alarm state is reported (gray/red square); if present, hypoglycemia events with their detection windows are shown.

- True Positive Rate (TPR), or sensitivity, and True Negative Rate (TNR), or specificity, to measure the proportion of positives and negatives that are correctly identified, respectively; these indices can be computed as follows:

$$TPR = \frac{TP}{TP + FN}, \quad TNR = \frac{TN}{TN + FP}$$

- Positive Predictive Value (PPV), or precision, and Negative Predictive Value (NPV) to measure the proportion of positives and negatives that are correctly identified over all the positive or negative predictions;

$$PPV = \frac{TP}{TP + FP}, \quad NPV = \frac{TN}{TN + FN}$$

- False Positive Rate (FPR) and False Negative Rate (FNR) to measure the proportion of positives and negatives that are wrongly identified, respectively;

$$FPR = \frac{FP}{FP + TN}, \quad FNR = \frac{FN}{TP + FN}$$

- Accuracy (ACC) to measure the proportion of classifications, both positives and negatives, that are correctly identified:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

- False Omission Rate (FOR) to measure the proportion of negatives wrongly predicted over all the negatives:

$$FOR = \frac{FN}{TN + FN}$$

- F1 score (F1) to measure the harmonic average of the precision and sensitivity, F1 score reaches its best value at 1 (perfect precision and sensitivity) and worst at 0:

$$F1 = \frac{2 * TPR * PPV}{TPR + PPV}$$

It should be noted that, from a classification point of view, the considered dataset is strongly unbalanced since the percent of time spent in hypoglycemia by the patient under study is only 3.07% of the total. Moreover, in the studied patients population, time spent in hypoglycemia was 1.9% ± 1.1% (mean ± SD) [59]. This poses well known problems when interpreting of the metrics reported above. In fact, in this condition TN is very large with any reasonable alarm system, and it saturates the metrics influenced by this quantity, e.g. TNR will be close to 100%,

Parameter	HPSA	IHPA
TP	26	34
FN	6	2
$NEHE$	4	0
FP	23	10
TN	6856	7050
Total	6915	7096

TABLE 7.1: Comparison between HPSA and IHPA on a single patient (Patient #1).

making these metrics less informative than in the case of balanced datasets. For instance, in an unbalanced dataset any degenerate alarm/classification algorithm that detects always the most common class (not hypoglycemia in our case) scores a very high accuracy. In view of this, the results are focused on sensitivity and precision, not depending on TN, even if all the performance metrics reported before are computed for completeness.

7.3 Evaluation of IHPA algorithm

A preliminary study and evaluation of the algorithm has been conducted on a single patient (Patient #1). The performance metrics of the alarm sytem are evaluated on the entire trial where 36 hypoglycemia events were found. The number of TP, FN, FP and TN detected for each algorithm are reported in Table 7.1. HPSA algorithm was not evaluated for 4 events because the system was off when the hypoglycemia occurred during the trial. In 3 of these cases, IHPA is able to detect correctly the event. IHPA is able to increase the number of TP of 30.77% respect to HPSA, decreasing at the same time the FN of 66.67% and the FP of 56.52%. The prediction performance achieved by HPSA and IHPA have been computed as described in Section 7.2 and reported in Table 7.2. The new methodology is able to improve all the considered metrics: the sensitivity (TPR) of the predictor by 16.23% and the precision (PPV) by 45.63%; the FPR by 58.96%, the FNR by 70.37% and the F1 by 32.40%. Regarding the performance metrics affected by the unbalanced nature of the dataset, the specificity (TNR) decreases by 3.01%, the accuracy (ACC) improves by 0.26%, the NPV by 0.06%, and FOR by 68.58%. The improvements are limited where HPSA had already good performance, while the new methodology is able to optimize the other characteristics. For example, the sensitivity of the alarm system IHPA ensures to predict an unavoidable hypoglycemia in 94.44% of the cases, giving the possibility to the patient to implement recovery treatments on time. It is worthy to note that IHPA system

Performance index	**HPSA**	**IHPA**
TPR	81.25 %	94.44 %
PPV	53.06%	77.27 %
FPR	18.75 %	5.56 %
FNR	0.33 %	0.14 %
$F1$	0.64	0.85
TNR	95.92 %	93.03 %
NPV	99.91 %	99.97 %
ACC	99.58 %	99.84 %
FOR	0.09 %	0.03 %

TABLE 7.2: Performance index comparison between HPSA and IHPA on a single patient (Patient #1).

is designed to detect unavoidable hypoglycemia events, so part of the undetected 5.56% could be relative to avoidable hypoglycemia events for which the pump had not been shut off during the trial. Moreover, this was a retrospective study, not designed ad hoc for this purpose, where the hypotreatment was driven by HPSA alarm. For this reason, FPs of HPSA could be hidden by not necessary hypotreatment delivered only because of the presence of the alarm. In spite of this favorable limitation for HPSA, IHPA is able to improve its performance. This alarm system has shown significant improvements in hypoglycaemia detection in comparison with the safety system used in the trial: the case-study conducted on Patient #1 has been published in [62].

7.4 Results on the patient cohort

The good results on a single patient obtained by considering the entire month, where the patient habits have widely changed, have represented a starting point for the validation of the IHPA on the entire cohort of the patient of the Amsterdam Medical Centre. The patients experienced about 20 HE on average. The performance indices of IHPAs on each patient and on the entire the population are reported in Table 6.6. In case of an unbalanced dataset, the specificity could provide false assumptions on the algorithm. Hence it is better to rely on sensitivity and precision. The results include the sensitivity (TPR) of the predictor of 84.67%: since sensitivity measures the accuracy of positive cases, an high value of TPR means that this algorithm is

able to forecast correctly the hypoglycemic events with great accuracy. Moreover, precision (PPV) of 41.41% has been obtained. Precision on the other hand is a measure of a model's exactness, thus higher precision value for this algorithm is an indication of a good prediction capabilities. Simultaneously, the IHPA has shown FPR of 0.328% and FNR of 15.33% and F1 reaches 52.82% on average for all 7 patients. Regarding the performance metrics affected by the unbalanced nature of the dataset, the specificity (TNR), accuracy (ACC), NPV (all $> 96\%$) and FOR ($< 0.042\%$) indicate alarms activated mostly when needed. The IPHAs based on patient-tailored models showed a sensitivity (TPR) of 84.67% with FPR of 0.328% on the entire cohort of patients: on average about 85% of the hypoglycemia events occurred during the trial have been detected in time to allow a rescue action with a negligible ($< 0.4\%$) number of false alarms. Since the results obtained on the entire population have confirmed the validity and the reliabilty of the proposed hypoglycemia detection algorithm, this approach has been published in [97].

Metric	Patient #1	Patient #2	Patient #3	Patient #4	Patient #5	Patient #6	Patient #7	Population
HE	34	13	29	14	20	16	11	19.57 (±8.73)
TPR	97.06	84.62	72.41	78.57	85.00	75.00	100.00	84.67 (±10.56)
PPV	76.74	31.43	48.84	52.38	44.74	9.52	26.19	41.41 (±21.57)
FPR	0.143	0.328	0.328	0.130	0.264	1.291	0.404	0.328 [0.173 0.385]
FNR	2.94	15.38	27.59	21.43	15.00	25.00	0.00	15.33 (±10.56)
F1	85.71	45.83	58.33	62.86	58.62	16.90	41.51	52.82 (±21.27)
TNR	93.12	97.71	94.76	98.42	96.82	96.15	97.70	96.38(±1.88)
NPV	99.99	99.97	99.88	99.96	99.96	99.95	100.00	99.96 (± 0.04)
ACC	99.84	99.65	99.56	99.83	99.70	98.67	99.60	99.65 [99.57 99.80]
FOR	0.014	0.027	0.119	0.0389	0.038	0.046	0.000	0.041(± 0.038)

TABLE 7.3: IPHA performance indices computed on the selected cohort of 7 patients. The performance computed on the entire population is reported in terms of mean ± SD if the data are normally distributed and in terms of median [25^{th} – 75^{th} percentiles] otherwise.

Chapter 8

Individualized model for controller

Currently, the model included in the MPC is the average linear time-invariant metabolic model, which is computed by linearizing the model of the average adult virtual patient of the UVA/-Padova simulator. The performance of the controller can probably be improved by synthesizing the controller on an individualized model. Since promising results have been obtained in the identification of patient-tailored models from clinical data as reported in Chapter 6, the application of these techniques has been extended to adapt the patient-tailored model to the patient changes over time. In order to capture the variations in patient's dynamics, an in-depth analysis of the data can be used to extract hidden information from the dataset useful to improve model identification and understand better the time changes. An ANOVA analysis has been performed on the real-life data, and it has shown that there is a significant dependence between different day periods and the glucose profile, i.e. different glucose dynamics are present during the day. The results of the ANOVA tests have been exploited to build a multi-model able to captures different PP glucose dynamics along the day using different models in each sub-period. In particular, a data-driven Multiple-Model Predictor (MMP) based on real-data analysis is proposed and analyzed in this chapter in terms of predictive performance. With data-driven MMP we indicate a predictor that exploits different identified models on the basis of the a priori knowledge acquired through the analysis of real patient. The MMP uses three Basic Models (BMs) defined on the basis of the data analysis: each one is specific of a Day Period (DP). The DPs considered in this work are breakfast (B), lunch (L) and dinner (D). The use of the DPs is strictly dependent from the results obtained via the statistical analysis that correlates them to the PP glucose dynamics. This correlation could be not universally verified since it is a characteristic typical of the dataset. However, this approach allows to exploit correlation to other meal parameters to build the MMP. The statistical analysis is here performed on patient data collected during clinical trials described in Section 6.2.1. The main focus of the analysis

is the PP glucose control, which is typically one of the most problematic aspects of glucose regulation. From the collected dataset a correlation between PP glucose profiles and different daytimes was observed. This feature is confmed by the presence of the intra-day variability, a phenomenon well-known in the literature [18], which results in different responses to meals and insulin during the day. The presented data analysis is focused on the PP glucose dynamics with the aim of proving the existence of a statistical correlation between them and the DPs they belong to.

8.1 Data Analysis

The availability of long term trials in CL [59, 75, 98] allows the collection of datasets rich of information potentially useful to improve model identification. A deep analysis of this large amount of data can lead to understand the correlation between some known parameters and the meal response, if it exists. In Chapter 4 this kind of analysis was carried out on *in silico* data. The relation between the two main PP variables that describe the meal response, the amplitude of the excursion and the shape of the glucose profile, and several parameters available at meal time such as BG, Carbohydrate intake (CHO), DP, etc. was considered. An interesting result obtained *in silico* was the correlation between the DP and the shape of the meal response, also justified by the well-known intra-day variability of PP glucose dynamics. In this work the influence of this feature on the characteristics of the meal glucose response is studied for all the patients belonging to the three centres involved in the trial presented in [59].
In order to study this correlation, the meal glucose response characteristics have to be formally defined. A meal intake induces a rise in BG concentration, while an insulin bolus results in a decrease in glucose level. The rise in BG concentration due to the absorption of carbohydrate may be immediate or delayed with respect to the mealtime depending on the meal composition and other factors [99]. On the other hand, it is well-known from physiology that the insulin response is also affected by absorption delays. Thus, the magnitude variation in glucose concentration and the time interval amplitudes mainly depict the PP glucose response as shown in Figure 8.1. The magnitude variations with respect to the glycemia value at mealtime are positive or negative if they are due to meal or insulin response, respectively. The time interval amplitudes are greater or equal to zero depending on the absorption delays.

The parameters characterizing the PP curve considered in this thesis are shown in Figure 8.1 where:

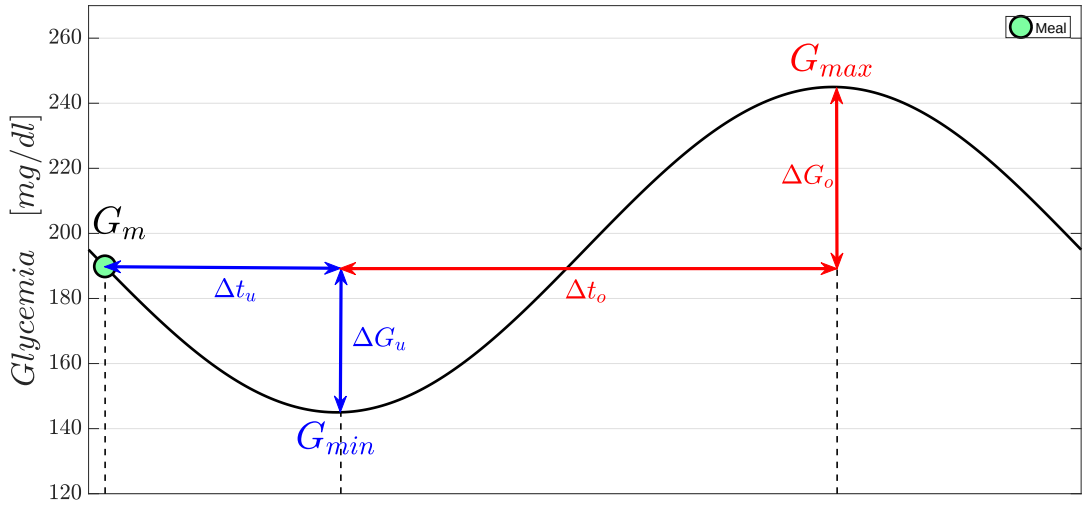


FIGURE 8.1: An example of glucose trend with the indicators $\Delta G_u, \Delta t_u$ (in blue) and $\Delta G_o, \Delta t_o$ (in red).

Indicators	**Description**
ΔG_u	$G_m - G_{min}$
ΔG_o	$G_{max} - G_m$
Δt_u	$t_{min} - t_m$
Δt_o	$t_{max} - t_{min}$

TABLE 8.1: The fours indicators of the glucose trend with relative definitions.

- G_{min} is the minimum glucose below G_m reached before the glucose starts rising due to the meal intake
- G_{max} is the maximum value of PP glucose concentration
- t_{min} and t_{max} are the time instants when G_{min} and G_{max} values are reached, respectively

with G_m, the glucose at meal time, t_m. If the glucose starts rising immediately after the meal intake, G_{min} and t_{min} are set equal to G_m and t_m, respectively. Denote with ΔG_u the amplitude of the maximum negative BG variation between G_m and G_{min} and ΔG_o the amplitude of the positive BG variation between G_m and G_{max}. Furthermore, define Δt_u as the time interval between t_{min} and t_m and Δt_o as the time interval between t_{max} and t_{min}. All these parameters are computed on a PP period of 7 hours [100]. Table 8.1 summarizes the definitions of the fours indicators of the glucose trend. In order to study the relation between DP and these characteristics, the DPs have to be defined. The idea is to take into account the intra-day variability

of PP glucose dynamics to enhance the predictive capabilities of the patient model. In this prospective, the DPs have been defined on the base of well-known clinical parameters individualized for each subject. In fact, information about the intra-day variability for each diabetic patient can be extracted by the open-loop CT. The computation of the insulin bolus involves CR that is time-varying along the day. CR is influenced by the patient insulin sensitivity strictly connected to the intra-day patient variability, as well as the basal insulin. Thus, the information needed to define the day periods can be extracted by the time intervals used to set the CR or the ones used for the basal; since each patient has his/her own daily pattern, this information is already individualized. Three different ways to define the DPs have been explored: one based on the basal intervals (DP_b), another based on the CR intervals (DP_C) and, as standard, one derived from the common three daily meals (DP_S, with t_B=5:00, t_L=12:00 and t_D=19:00). This decision is due to the awareness that these parameters contain specific information about the intra-day variability, but which one is the most accurate is not known a priori. Each meal is assigned to a specific DP $[t_{start}, t_{end}]$ on the base of t_m, i.e. if $t_m \in [t_{start}, t_{end} - 30]$. If t_m is close to a switch time (less than 30 minutes), the meal is assigned to the next DP.

Once the indicators ΔG_o, ΔG_u, Δt_o, Δt_u and the input feature DP_b, DP_C, DP_S have been defined, an ANOVA test is performed for every patient of the dataset to determine if a statistical correlation exists, where a statistical correlation is significant at the level 0.05. This test allows also to chose the proper DP definition for each patient. We define that PP glucose profile is statistically correlated with DP if at least 3 indicators are dependent from DP.

Since a clinical trial under free-living condition means that the patient has a normal life without any type of restriction. The original data need to undergo preprocessing before the analysis. For example, PP glucose dynamics belonging to different meals closed to each other are usually overlapped. In order to cope with this problem, the concept of *cumulative meal* is introduced. Since the glucose dynamics are slow, two meals consumed in a small time interval (<40 minutes) do not produce distinct PP glycemic response (peaks 15 minutes apart). So, the indicators of each single meal defined in Section 8.1, can not be correctly computed. A cumulative meal, having the carbohydrate amount equal to the sum of the two individual meals, is considered as a single meal consumed at time of the first meal and associated to the overall PP glucose profile. The drawback of this approach is related to the combination of meals belonging to different groups of indicators.

Since the datasets of some patients were characterized by many inaccuracies, which alter the reliability of the data, thirteen patients of the total cohort [59] are considered. The selection of these patients is based on the accurate graphical analysis of the individual glucose profile of each patient, also used to deduct the eating habits. In particular, the preprocessing step reduced the total amount of meals available from the trial by 11%. For each patient 116.38 ($\pm$38.46) meals

Patients	*Indicators*			
	$\mathbf{\Delta t_u}$	$\mathbf{\Delta t_o}$	$\mathbf{\Delta G_u}$	$\mathbf{\Delta G_o}$
Patient #1		$\mathbf{S}, C, B$		
Patient #2	S	$\mathbf{S}, B$	$\mathbf{S}, B$	$\mathbf{S}, B$
Patient #3	S	$\mathbf{S}, B$	$\mathbf{S}, C, B$	$\mathbf{S}, B$
Patient #4				
Patient #5		B		
Patient #6	B		B	$\mathbf{B}, C$
Patient #7	$\mathbf{S}, C, B$	$\mathbf{S}, C, B$	$\mathbf{S}, C, B$	
Patient #8	S	S		$\mathbf{S}, B$
Patient #9	$\mathbf{S}, C, B$	S	S	
Patient #10	$\mathbf{B}, C$	$\mathbf{B}, S$	$\mathbf{B}, C$	$\mathbf{B}, S, C$
Patient #11		S	S	
Patient #12		C		
Patient #13	S	S	S	S

TABLE 8.2: Results obtained via ANOVA test. The presence of a letter indicates a statistical correlation between DP and the specific indicator. The letters indicate the DP definitions ($B = DP_b$, $C = DP_{CR}$, $S = DP_S$) that satisfy this correlation. The strongest correlation is highlighted in bold.

were included in the dataset (total considered meals, M=2095). The patients with a dataset size below mean-SD (78 meals) were excluded from the statistical correlation analysis in the following section. Five patients did not respect this criteria, so they were totally discharged. The remaining patients did not show imbalance between the considered DPs.

8.1.1 Statistical correlation analysis

The aim of this section is to establish the existence of a statistical correlation between PP glucose profile, depicted by four indicators ($\Delta G_u, \Delta G_o, \Delta t_o, \Delta t_u$) presented in Section 8.1, and DPs. This goal is achieved through an ANOVA analysis performed independently for each of the 13 considered patients: Table 8.2 shows the results of ANOVA test, where the presence of a letter indicates a statistical correlation between DP and the specific indicator. In particular, for a given patient, it indicates that the correspondent indicator of PP glucose dynamics in a particular DP is different with respect to the same indicator in at least another DP. The letter in Table 8.2 specifies which DP definition satisfies the dependence; in particular, the primary

correlation is reported in each cell in bold with the other secondary ones in normal style. Eight patients out of 13 have at least 3 indicators significantly dependent from DP; in particular, Patients #2 and #4 have all the indicators dependent from DP_S, that means the meals belonging to different DPs have PP glucose profiles completely different in terms of all the selected characteristics (ΔG_o, ΔG_u, Δt_o, Δt_u). Also Patient #6 shows 3 indicators dependent from DP_B. Patient #1 is the test-case used in Section 6.5 to identify a patient-tailored model. The ANOVA test confirms that for this patients the PP glucose profile is not dependent from the DPs, so the use of the Multiple-Model (MM) approach is not expected to improve the prediction capability for this subject. Patient #13 has all the indicators dependent from DP_S, while Patient #10 from DP_b. For Patient #7, all the DP definitions resulted to be significant correlated to 3 indicators: ΔG_u, Δt_o, Δt_u. Since, both his/her CR and basal contains 5 intervals, DP_S as been selected as primary definition for this patient. The other patients do not show relations so significant to any DPs. The detailed results of ANOVA for Patient #7 are reported in Figure 8.2. The DPs selected for this patient are: breakfast DP (B, in red) [5:00-12:00], lunch DP (L, in blue) [12:00-19:00] and dinner DP (D, in green) [19:00-5:00]. In Figure 8.2, for each DP the mean of the specific indicator with the range of variation in terms of SD is reported. For example, the Δt_u has a mean of 14 minutes for the breakfasts, of 40 minutes for the lunches and of 78 minutes for the dinners. On the other hand, the ΔG_u has a mean of 3 mg/dl for the breakfasts, of 19 mg/dl for the lunches and of 40 mg/dl for the dinners. These results denote a PP glucose response characterized by a small undershoot limited in time for the breakfasts, while a large and long undershoot for the dinners. The lunches have average characteristics of the other two that sometimes overlap one or the other category. Figure 8.3 represents an example of a day profile where these characteristics are clearly shown: the first two meals that belong to the breakfast DP are characterized by a $\Delta t_u = 0$, while the last meal belonging to the dinner DP has significant Δt_u.

8.2 Multiple-Model Predictor

The identification of a patient-tailored model is one of the most challenging aspects of glucose regulation. Given the high inter-patient variability that characterizes the system, the availability of a specific model for each patient represents an improvement in terms of glucose control and hypoglycemia prevention. However, considering the intra-day variability that characterizes the diabetic patients, a linear time-invariant model is intrinsically limited. In particular, considering the high correlation between PP glucose response and DP demonstrated in Section 8.1.1, the representation of different dynamics with a time-invariant model limits the model prediction

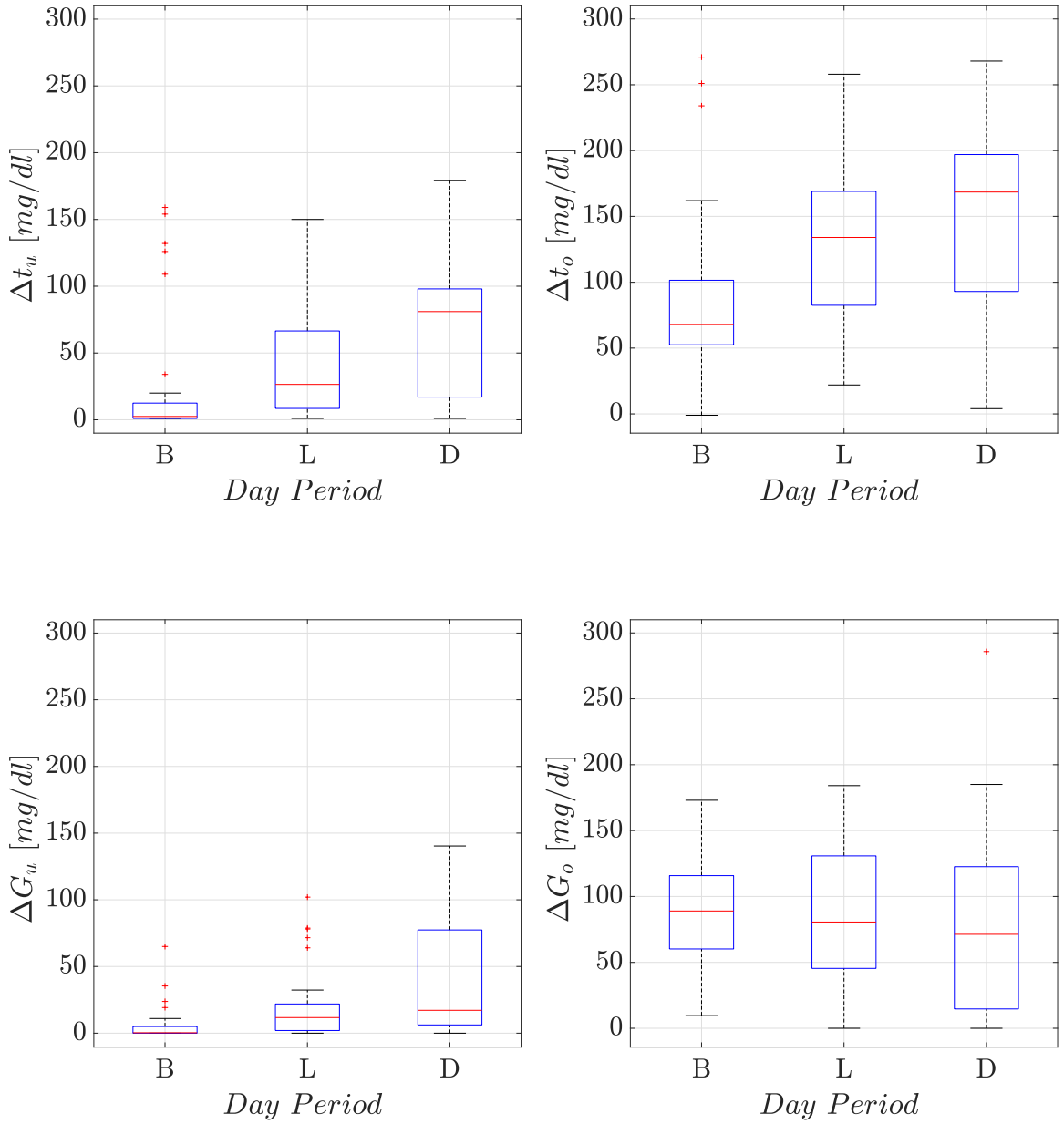


FIGURE 8.2: Boxplots associated to ANOVA test of each indicator in the three selected DPs: breakfast (B), lunch (L) and dinner (D).

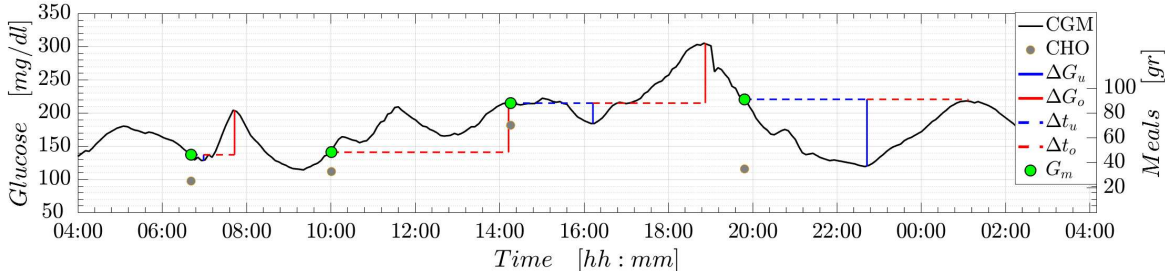


FIGURE 8.3: Glucose trend of Patient #7.

performance. In order to keep the model simple enough to guarantee the implementation of a linear MPC based on this model on a smartphone, a Multiple-Model Predictor (MMP) is proposed to cope with the intra-day variability of the system. For each DP defined as reported in Section 8.1 a linear time-invariant Basic Model (BM) is identified via the IR identification approach presented in Section 6.1.1. The ANOVA tests have shown that Patient #7 has a very large values of Δt_o as shown Figure 8.2. The Δt_u has a mean of 81 minutes for the breakfasts, of 127 minutes for the lunches and of 148 minutes for the dinners. A large Δt_o implies that there is a significant delay in the rise of glucose levels after the meal. In order to take into account this delay, the delay of meal absorption (τ_M) has been introduced in the IR technique. In each DP, the prediction is performed with the specific model identified for that DP. The BM models will be referred in the following as BM Breakfast (BMB), BM Lunch (BML) and BM Dinner (BMD).

Each BM has two measurable inputs, the injected insulin, $i(k)$, and the carbohydrates content, $m(k)$ and one output, the glucose concentration measured by the CGM sensor, $CGM(k)$, collected every T_s minutes. The model has the following structure:

$$CGM^p(z) = G_i^p(z)I^p(z) + G_m^p(z)M^p(z) + E(z) \tag{8.1}$$

where $I^p(z)$, $M^p(z)$ and $CGM^p(z)$ denote the Z-transforms of inputs and output in a particular DP, p, $G_i^p(z)$ and $G_m^p(z)$ are transfer functions to be estimated from the data collected in p and $E(z)$ is the Z-transform of the residual error $e(k)$. Once a model for each DP is identified, the predictor is designed in the following form:

$$\begin{cases} x_B(k+1) = A_B x_B(k) + B_B u(k) + K_B e(k) \\ x_L(k+1) = A_L x_L(k) + B_L u(k) + K_L e(k) \\ x_D(k+1) = A_D x_D(k) + B_D u(k) + K_D e(k) \\ y(k) = \alpha C_B x_B(k) + \beta C_L x_L(k) + \gamma C_D x_D(k) + e(k) \end{cases} \tag{8.2}$$

where $x_B, x_L, x_D \in \Re^n$, with $n = 12$, are the space vectors of the BMB, BML and BMD respectively, and $\{A_B, B_B, C_B, K_B\}$, $\{A_L, B_L, C_L, K_L\}$ and $\{A_D, B_D, C_D, K_D\}$ are their minimal state-space realization, and $u(k) = [i(k), m(k)]'$ is the input vector. The parameters $\alpha, \beta, \gamma \in \{0, 1\}$ determine the switch between the three models that remain always active: $\alpha = 1$ if $k \in B$, $\beta = 1$ if $k \in L$, $\gamma = 1$ if $k \in D$ and α, β, γ are set to 0 otherwise.

In order to evaluate the improvement of the new approach, a model identified on a daily subset, called Daily Model (DM), using the standard procedure described in Section 6.1.1 is considered. The performance of MMP is here compared to the one achieved by DM Predictor (DMP), a predictor based on DM for a test-case study. The selected patient is Patient #7 since from the

data analysis he resulted particularly suitable for the MMP approach as reported in Section 8.1.1.

For each model, two datasets were selected: an identification dataset of variable length depending on the type of model (DM, BMB, BML or BMD) and a testing dataset, including the entire 1-month dataset from the trial. The identification dataset consists of three different disjoint subsets of data: two are used to compute an average SSR as shown in Eq. 6.3, in order to avoid overfitting, while the third is used to estimate the stochastic part of the model. The first two subsets are specific of the type of model to identify and in the following they will be referred as training-set A and training-set B. The third subset is common to all types of models and includes 24 hours of data; it will be referred as test-set.

For DM two 24-hour subsets have been selected as training-set A and B. For each BM model, a different sub-portion of these subsets have been selected: for BMB two segments of 5 hours starting from 5:00; for BML two portions of 6 hours starting from 12:00 and for BMD two segments of 7 hours starting from 19:00.

8.3 Identification results

All type of models (DM and BMs) using all the possible combination of the initialisations, $\theta^{init} \in \{\theta^{Av}, \theta^{Cp}\}$, and delays, $\tau_I \in \{0, 15, \ldots, 60\}$ and $\tau_M \in \{0, 15, \ldots, 75\}$ have been identified. DMP with the best performance uses a DM with $\tau_I = 45$ minutes, $\tau_M = 60$ minutes and the clinical initialization of the gains, while MMP uses a BMB with $\tau_I = 60$ minutes, $\tau_M = 15$ minutes and the average initialization of the gains, a BML with $\tau_I = 45$ minutes, $\tau_M = 45$ minutes identified on the dinner subset with the clinical initialization of the gains, and a BMD with $\tau_I = 15$ minutes, $\tau_M = 75$ minutes and the average initialization of the gains. It is important to note that for MMP, the best BMs resulted to have τ_M equal to the average indicator Δt_u for the specific DP. In fact, as reported in Section 8.1.1, the Δt_u has a mean of 14 minutes for the breakfasts, of 40 minutes for the lunches and of 78 minutes for the dinners. So, the models with τ_M set to 15 minutes for BMB, to 45 minutes for BML and to 75 minutes for BMD obtained the best results. From the ANOVA analysis, the lunches resulted not to be clearly distinguishable from the other two meals. This is the only case where the model with the best performance in a DP resulted to be the one identified using data belonging to another DP.

In Table 8.3 the metrics of these DMP and MMP achieved on the test-set are reported. MMP obtained a limited improvement of the prediction performance in terms of $\overline{FIT}$ (3.87%), in

Predictor	$\overline{FIT}$ [%]	$\overline{COD}$ [%]	$\overline{\rho}$
DMP	61.44	80.78	0.89
MMP	63.82	83.15	0.91

TABLE 8.3: Prediction metrices.

Predictor	$\overline{FIT}$ [%]	$\overline{COD}$ [%]	$\overline{\rho}$
DMP Breakfast	42.81	58.27	0.75
MMP Breakfast	50.82^b	69.67^b	0.84^a
DMP Lunch	66.5	85.32	0.92
MMP Lunch	67.32	86.06	0.93
DMP Dinner	63.1	82.19	0.9
MMP Dinner	63.19	82.19	0.91

TABLE 8.4: Prediction metrics in each DP. p-value (p) significance levels are: $a := p < 0.001$, $b := p < 0.01$, $c := p < 0.05$.

terms of $\overline{COD}$ (2.93%) and in terms of $\overline{\rho}$ (2.25%). In order to better understand the nature of these limited improvements, a individual analysis has been conducted for each single DP. In Table 8.4 the performance achieved on the day period B, L and D is reported. The meals that presented problematic characteristics as discussed in Section 8.1 have not been considered in this analysis; in particular, 6 breakfasts, 2 lunches and 7 dinners have been excluded. MMP presents similar performance for lunches and dinners (differences not statistically significant), but it is able to improve the prediction performance during the breakfast by 18.72% in terms of $\overline{FIT}$, by 19.56% in terms of $\overline{COD}$ and by 11.39% in terms of $\overline{\rho}$. All these results are statistically significant and confms the existence of different dynamics characterizing different categories of meals.

It is worthy to note that the parameters α, β, γ cause a sudden switch between a model to another. In Figure 8.4 the predicted glucose profiles obtained by MMP and DMP with PH=12 (60 minutes) are compared with the real glucose. The profile is related to a single day, where the different dynamics characterizing B with respect to L and D are well represented. In particular, the DMP is not able to predict the fast dynamic of B (DMP 38.81% vs MMP 65.51% in terms of FIT) since L and D requed slower dynamics. This aspect is not highlighted in Table 8.3 due to reduced number of breakfasts with respect to the total of lunches and dinners. The sudden switch between the models is not particularly critical as shown in Figure 8.4. However, the design of MPC for switching dynamics is not trivial, even if with limited switching time

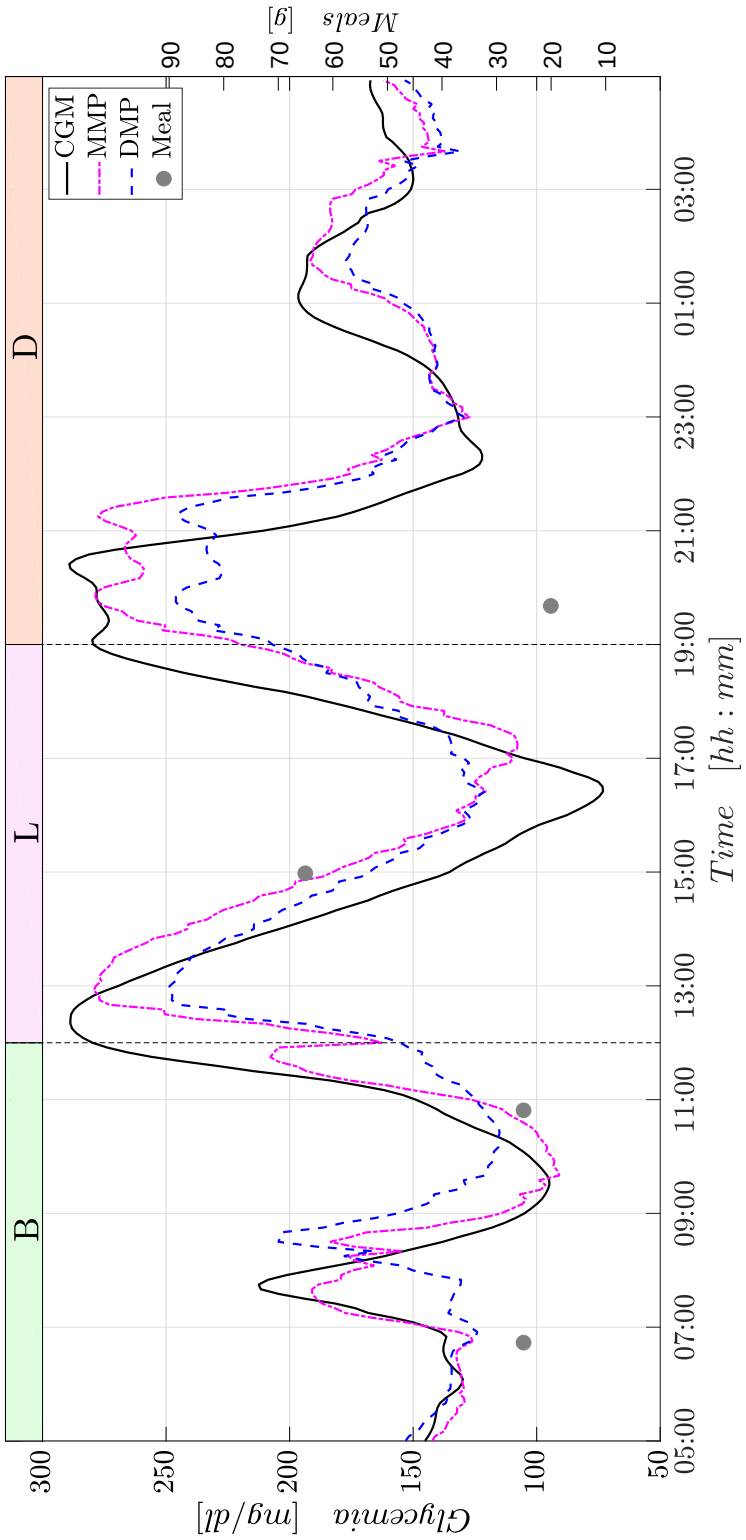


FIGURE 8.4: Glucose profile of a 24h study-case: the real glucose measured in the trial is presented as reference, the predictions obtained by MMP and DMP with DP=60 minutes and the DPs (top) are reported.

as proposed here (three per day) the critical points are restrained. However, even if only one patient is considered, since the strength of the results obtained is enforced by the large changes of the patient habits and the time-varying nature of the system under study, the MMP approach has been published in [101].

Chapter 9

Deep Learning for glycemia forecasting

The MPC approach described in Section 5.2 exploits a glucose-insulin model to forecast the BG values in order to compute the optimal insulin therapy. For this reason, the predictive performance of the model plays a key role in the overall control performance as already reported in the previous chapters. Classical mathematical model used in the AP research field are not able to fully describe the nonlinear glucose-insulin dynamics. In order to overcome this limitation, the complexity of the model has to be increased and new effective identification techniques are required. Data-driven approaches based on deep learning architecture have received an increasing attention in the last few years mainly because of the remarkable performance obtained in several research fields [102, 103]. Depending on the task at hand (i.e., classification, regression, prediction, etc.) the aim of these approaches is to learn a model directly from the data. Since a huge amount of data collected during long-period trials is available, new data-driven approaches have been explored in this research field. Among these approaches, recurrent neural networks represent a family of deep learning architectures which have been explicitly designed to model the evolution over time of a phenomenon. In particular, given an input composed of a sequence of observations from a signal, such as the BG level in our scenario, these models try to predict its future value or values. Thus, the development of a new forecasting model able to predict the future BG of a patient subject to several possible insulin treatments is explored in order to define his/her optimal future insulin therapy. In particular, the final purpose of this research is to have a model to be included in the MPC described in Section 5.2. In this perspective, a deep learning architecture able to forecast the BG level of T1D patients is studied. The model architecture is composed of two models, one observing the CGM measurements, insulin injections and carbohydrate intakes up to a given time and a second model that receives as input

the future insulin that will be administered to the patient and the future carbohydrates that he/she will assume. Both models are composed of stacked Long-Short Term Memory (LSTM) networks [104]. The output of the two models is combined and given as input to a Fully Connected (FC) layer which is used to predict the future values of the Interstitial Glucose (IG), considering a fixed prediction horizon. Training is performed in a supervised fashion on a subset of identities, separated from those that will be considered as test, in order to obtain a model which is able to generalize to new unseen data.

9.1 Solutions based on Neural Networks

The first solution exploiting neural networks for modeling the BG metabolism of a T1D patient was proposed in [105]. In particular, the authors tried to predict the glucose level of a diabetic patient by training a Recurrent Neural Network (RNN) architecture which receives insulin levels, meals and level of exercise as inputs, alongside current and previous estimates of BG. However, the data used for both train and test was acquired from a single patient and this may result in a lack of generalization for the final model. Recently, a few solutions exploiting deep learning techniques for glucose level prediction in diabetic patients have been proposed [106–110]. Similarly to [105], Allam et. al. [106] proposed the use of CGM signals to train a RNN for predicting future values of the glucose concentration, considering several PHs. Again, the data used for both train and test were selected from the same population, which may result in a model that hardly generalize to new unseen data. LSTM networks achieve state-of-the-art performance in modeling several time-dependent phenomena. For this reason, the authors of [107] proposed to exploit LSTM in a model which takes CGM values, insulin dosages and carbohydrate intake as inputs and tries to predict the glucose level at PHs of 30 and 60 minutes. Data incoming from four patients acquired using different CGM devices has been used in both train and test. Unfortunately, the training needed to be repeated multiple times, mainly because of initialization issues which left the optimization stuck in a bad local optima. An LSTM-based architecture has also been exploited in [108]. In this case, the model is trained on the measurements provided by CGM systems, and used to predict a singular value after a pre-defined prediction horizon. The output is modeled as a univariate Gaussian distribution, so as to be able to follow the uncertainty of the prediction. The LSTM dimension was set to 128 and it was trained on the Ohio T1DM Dataset [111], considering the first 80% of the glucose level as training data for each patient, and validating on the last 20%. A more complex architecture was designed by Sun et. al. [109]. In particular, they proposed to use a sequential model with one LSTM layer, one bidirectional LSTM layer and several fully connected layers to predict BG levels for different

PHs. The model was trained on the CGM measurements of both *in silico* and *in vivo* data coming from 20 real patients. Convolutional RNNs have also been exploited to predict the BG level [110]. The concatenated time series of glucose level, carbohydrate and insulin has been firstly preprocessed by a deep convolutional networks, so that the recurrent LSTM layers accepts these features instead of the CGM measurements directly. The model has been trained on *in silico* data consisting on a small sample of 10 adult T1D subjects simulated using the UVA/Padova simulator.

9.2 Long-Short Term Memory Architecture

Glucose concentration depends mainly on the injected insulin and carbohydrate intake, which have opposite effects on glucose levels, and its future evolution is also influenced by its current value and trend. All these variables can be easily measured without an invasive data collection. For these reasons, the inputs of the proposed model will be composed by these three measurable signals sampled at a given rate T_s: the injected insulin ($i(t)$) recorded by subcutaneous insulin pump, the carbohydrate amount ($m(t)$) inserted manually by the patient, and the glucose concentration ($CGM(t)$) measured by the CGM sensor. The output of the model is the interstitial fluid glucose concentration ($ig(t)$). The signal $CGM(t)$ is the interstitial (i.e. subcutaneous) glucose concentration measured by a CGM device and affected by measurement noise, while $ig(t)$ is the real interstitial glucose. The available measures are related to the CGM, while the variable of interest is the IG. All these signals have different ranges of values according to the units adopted by the UVA/Padova simulator: injected insulin doses [pmol/min] and carbohydrate [mg] amounts are about 100 times larger than the glucose measurements [mg/dl] in this dataset. In order to eliminate the units of measurements for data and to guarantee that all features contribute equally in the training process, a data preprocessing step is introduced. In particular, each signal is independently rescaled using the minimum and maximum values and then subdivided in intervals of fixed size, depending on the PH in analysis. These sub-samples constitute the training and testing data for our model, as detailed in Section 9.3. Denoting the current time with t_0 and given PH $\in PH$, where $PH = [5, 10, \cdots, 60]$ is the set of the considered PHs as described in Section 6.3, let's define the following signals:

$$
\begin{aligned}
\overleftarrow{\mathbf{CGM}}(t_0, \mathrm{PH}) &= \Big[CGM(t_0 - \mathrm{PH}), CGM(t_0 - \mathrm{PH} + 1), \cdots, CGM(t_0 - 1)\Big], \\
\overleftarrow{\mathbf{i}}\,(t_0, \mathrm{PH}) &= \Big[i(t_0 - \mathrm{PH}), i(t_0 - \mathrm{PH} + 1), \cdots, i(t_0 - 1)\Big], \\
\overleftarrow{\mathbf{m}}(t_0, \mathrm{PH}) &= \Big[m(t_0 - \mathrm{PH}), m(t_0 - \mathrm{PH} + 1), \cdots, m(t_0 - 1)\Big], \\
\overrightarrow{\mathbf{i}}\,(t_0, \mathrm{PH}) &= \Big[i(t_0), i(t_0 + 1), \cdots, i(t_0 + \mathrm{PH} - 1)\Big], \\
\overrightarrow{\mathbf{m}}(t_0, \mathrm{PH}) &= \Big[m(t_0), m(t_0 + 1), \cdots, m(t_0 + \mathrm{PH} - 1)\Big], \\
\hat{\mathbf{ig}}(t_0, \mathrm{PH}) &= \Big[\hat{ig}(t_0), \hat{ig}(t_0 + 1), \cdots, \hat{ig}(t_0 + \mathrm{PH} - 1)\Big]
\end{aligned}
\tag{9.1}
$$

where $\overleftarrow{\mathbf{CGM}}(t_0, \mathrm{PH})$, $\overleftarrow{\mathbf{i}}\,(t_0, \mathrm{PH})$, and $\overleftarrow{\mathbf{m}}(t_0, \mathrm{PH})$ are the CGM data, the delivered insulin and the ingested carbohydrates in the past PH minutes, respectively, while $\overrightarrow{\mathbf{i}}\,(t_0, \mathrm{PH})$, $\overrightarrow{\mathbf{m}}(t_0, \mathrm{PH})$ are the suggested amount of insulin and the meal information in the future PH minutes. For each PH a single model is identified as described in Section 9.2.1. The aim of each model is to depict the relation between ig in the future PH minutes $\Big(\hat{ig}(t_0), \hat{ig}(t_0 + 1), \cdots, \hat{ig}(t_0 + \mathrm{PH} - 1)\Big)$, collected in the vector $\hat{\mathbf{ig}}$, and the above mentioned signals. In particular, each single $\hat{ig}$ value can be described as:

$$
\hat{ig}(t_0 + k, \mathrm{PH}) = g\left(\overleftarrow{\mathbf{CGM}}(t_0, \mathrm{PH}), \overleftarrow{\mathbf{i}}\,(t_0, \mathrm{PH}), \overleftarrow{\mathbf{m}}(t_0, \mathrm{PH}), \overrightarrow{\mathbf{i}}\,(t_0, \mathrm{PH}), \overrightarrow{\mathbf{m}}(t_0, \mathrm{PH})\right) \tag{9.2}
$$

with $k = 0, 1, 2, \cdots, \mathrm{PH} - 1$. In the perspective of employing our solution in an MPC and in order to be able to accurately predict a glucose trend, an ensemble of these models can be trained, independently for each PH, and the predictions from these models can be combined to obtain a trend of future glucose concentration as follows:

$$
\begin{aligned}
\hat{\mathbf{ig}}\,(t_0) = [&\hat{ig}(t_0, 5), \hat{ig}(t_0 + 1, 5), \cdots, \hat{ig}(t_0 + 4, 5), \\
&\hat{ig}(t_0 + 5, 10), \hat{ig}(t_0 + 6, 10), \cdots, \hat{ig}(t_0 + 9, 10), \\
&\vdots \\
&\hat{ig}(t_0 + 54, 60), \hat{ig}(t_0 + 55, 60), \cdots, \hat{ig}(t_0 + 59, 60)].
\end{aligned}
\tag{9.3}
$$

9.2.1 Proposed Architecture

A simple model based on stacked Long Short-Term Memory (LSTM) cells [104, 112] has been chosen. LSTMs are a special kind of RNNs, which are able to learn how to filter (e.g. forget) part of their hidden state during the inference process in order to model long-term temporal dependencies.

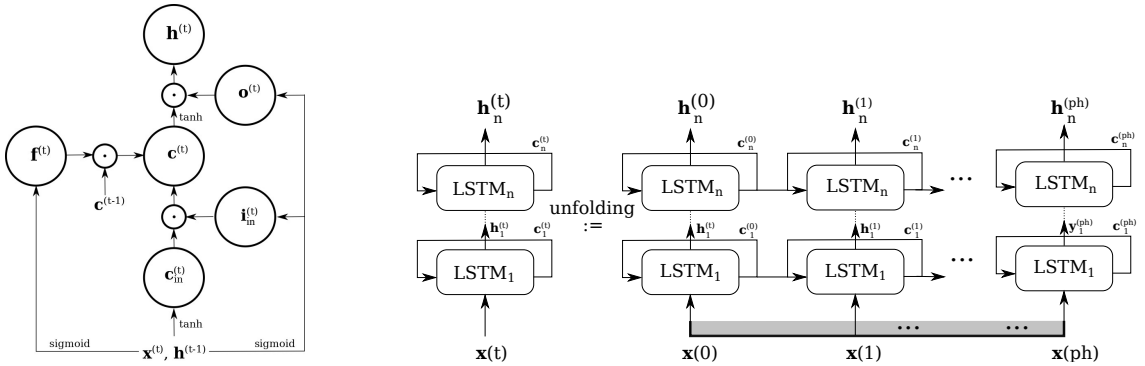


FIGURE 9.1: (a) The basic structure of an LSTM cell. For each arrow pointing to a circle, an addition is performed. Dots represent vector/matrix multiplications. (b) Temporal unfolding and data flow on n stacked LSTM cells.

Formally, a single LSTM cell with input $\mathbf{x}(t)$, output $\mathbf{h}(t)$ and an internal cell state $\mathbf{c}(t)$ is described by the following equations, also represented in graphical form in Figure 9.1(a):

$$\begin{aligned}
\mathbf{c}_{\text{in}}(t) &= \tanh(\mathbf{W}_{xc}\mathbf{x}(t) + \mathbf{W}_{hc}\mathbf{h}(t-1) + \mathbf{b}_c) \\
\mathbf{i}_{in}(t) &= \text{sigmoid}(\mathbf{W}_{xi}\mathbf{x}(t) + \mathbf{W}_{hi}\mathbf{h}(t-1) + \mathbf{b}_i) \\
\mathbf{f}(t) &= \text{sigmoid}(\mathbf{W}_{xf}\mathbf{x}(t) + \mathbf{W}_{hf}\mathbf{h}(t-1) + \mathbf{b}_f) \\
\mathbf{o}(t) &= \text{sigmoid}(\mathbf{W}_{xo}\mathbf{x}(t) + \mathbf{W}_{ho}\mathbf{h}(t-1) + \mathbf{b}_o) \\
\mathbf{c}(t) &= \mathbf{f}(t)\mathbf{c}(t-1) + \mathbf{i}_{in}(t)\mathbf{c}_{\text{in}}(t) \\
\mathbf{h}(t) &= \mathbf{o}(t)\tanh\mathbf{c}(t)
\end{aligned} \tag{9.4}$$

where each weight matrix $\mathbf{W}_x$, $\mathbf{W}_h \in \mathbb{R}^{d\times d}$ and $\mathbf{b}$, $\mathbf{x}(t)$, $\mathbf{h}(t)$, $\mathbf{c}_{in}(t)$, $\mathbf{i}_{in}(t)$, $\mathbf{f}(t)$, $\mathbf{o}(t)$, $\mathbf{c}(t) \in \mathbb{R}^d$ while d represent the LSTM *dimension*, an hyperparameter defined upfront by design and constant among all cells. Respectively, $\mathbf{i}_{in}(t), \mathbf{f}(t), \mathbf{o}(t)$ are called the *input*, *forget* and *output* gates, while $\mathbf{c}_{\text{in}}(t)$ contains a vector of new candidate values for the cell state. During temporal unfolding, both $\mathbf{h}(t)$ and $\mathbf{c}(t)$ are passed to the temporal replica of the next cell in the fold. Models made of multiple, stacked LSTM cells can be easily conceived, by making the output of a given cell the input of the next one in the stack. The process of training through unfolding n-stacked LSTM cells is illustrated in Figure 9.1(b). Multiple models have been trained, one for each PH $\in PH$. Depending on PH, the whole signal is sampled as described in Eq. 9.1 and each sub-sample is split into two arrays $\overleftarrow{\mathbf{X}}$ and $\overrightarrow{\mathbf{X}}$, the former representing past information given to the model, the latter representing the suggested therapy and meals for the future:

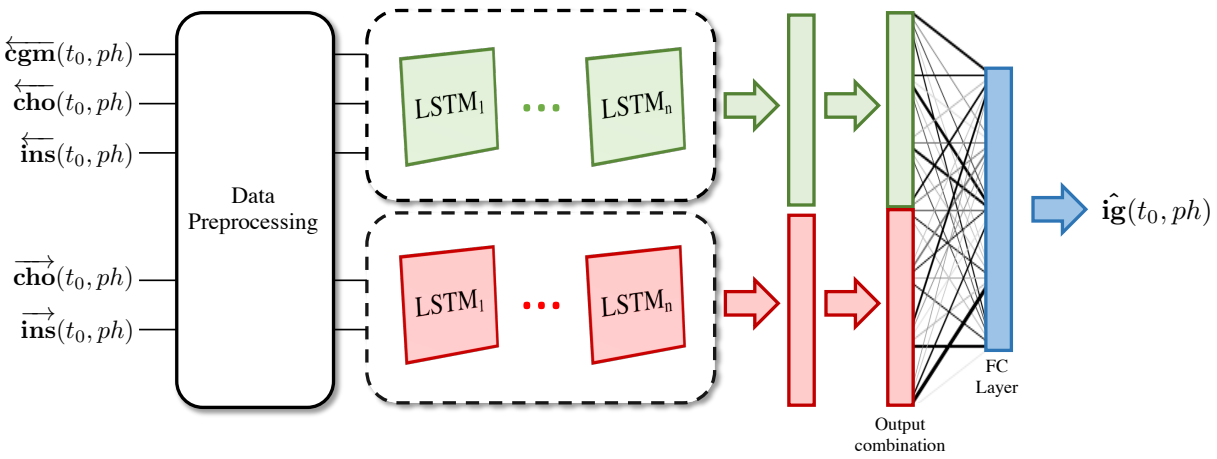


FIGURE 9.2: Deep Glucose Forecasting (DGF) architecture. The input is split into two sets: past observations and estimated future inputs. Both branches are processed by n-stacked LSTM cells with dimension d. The output of the branches is concatenated into a final Fully Connected (FC) layer.

$$\overleftarrow{\mathbf{X}}(t_0, \text{PH}) = \begin{bmatrix} \overleftarrow{\mathbf{CGM}}(t_0, \text{PH}) \\ \overleftarrow{\mathbf{i}}(t_0, \text{PH}) \\ \overleftarrow{\mathbf{m}}(t_0, \text{PH}) \end{bmatrix}, \quad \overrightarrow{\mathbf{X}}(t_0, \text{PH}) = \begin{bmatrix} \overrightarrow{\mathbf{i}}(t_0, \text{PH}) \\ \overrightarrow{\mathbf{m}}(t_0, \text{PH}) \end{bmatrix} \tag{9.5}$$

$\overleftarrow{\mathbf{X}}$ and $\overrightarrow{\mathbf{X}}$ are separately processed through two identical branches of the architecture, each being a stack of n LSTM cells. The output of both branches is then concatenated and processed through a final fully connected layer that produces the intended output. Since the main goal of this work is to predict the future BG of a patient subject to different insulin therapy in order to define the optimal treatment, the second branch containing the suggested future therapy cannot be excluded. As the model aims to forecast the IG signals, the supervised architecture assumes to have access to the IG signal during training in order to use them as ground truth. More formally, leaving out the flowing of the internal cell states, the model is described by:

$$\begin{aligned}
\overleftarrow{\mathbf{h}}_1(t_0, \text{PH}) &= \text{LSTM}_1(\overleftarrow{\mathbf{X}}(t_0, \text{PH})) & \overrightarrow{\mathbf{h}}_1(t_0, \text{PH}) &= \text{LSTM}_1(\overrightarrow{\mathbf{X}}(t_0, \text{PH})) \\
\overleftarrow{\mathbf{h}}_2(t_0, \text{PH}) &= \text{LSTM}_2(\overleftarrow{\mathbf{h}}_1(t_0, \text{PH})) & \overrightarrow{\mathbf{h}}_2(t_0, \text{PH}) &= \text{LSTM}_2(\overrightarrow{\mathbf{h}}_1(t_0, \text{PH})) \\
&\vdots & &\vdots \\
\overleftarrow{\mathbf{h}}_n(t_0, \text{PH}) &= \text{LSTM}_n(\overleftarrow{\mathbf{h}}_{n-1}(t_0, \text{PH})) & \overrightarrow{\mathbf{h}}_n(t_0, \text{PH}) &= \text{LSTM}_n(\overrightarrow{\mathbf{h}}_{n-1}(t_0, \text{PH}))
\end{aligned} \tag{9.6}$$

$$\hat{\mathbf{ig}}(t_0, \text{PH}) = \mathbf{W}_{FC} \left[\overleftarrow{\mathbf{h}}_n(t_0, \text{PH}) \;\; \overrightarrow{\mathbf{h}}_n(t_0, \text{PH})\right] + \mathbf{b}_{FC}$$

where $\mathbf{W}_{FC} \in \mathbb{R}^{d \times ph}$, $\mathbf{b}_{FC} \in \mathbb{R}^{ph}$ and LSTM_n represents the n-th LSTM layer in the stack and it is described by Eq. 9.4.

The training process uses a Mean Squared Error loss function (MSE) with a default Adam optimizer (learning rate 10^{-3}, batch size of 200, 180 epochs), so that for each sample:

$$\text{MSE} := \frac{1}{\text{PH}} \sum_{t=t_0}^{t_0+\text{PH}} (\hat{\mathbf{ig}}(t, \text{PH}) - \mathbf{ig}(t, \text{PH}))^2. \tag{9.7}$$

The complete solution is shown in Figure 9.2 and from now on the reader will refer to it as therapy-driven Deep Glucose Forecasting (DGF).

9.3 Dataset

The *in silico* dataset has been generated using the UVA/Padova simulator [24, 27] in its most recent version where the circadian variability of insulin sensitivity and meal absorption parameters have been [25, 28, 29], as described in Chapter 3. Two different scenarios are designed in order to to simulate the realistic intra-subject change in eating habits in terms of timing and meal size variations. Different food habits imply different insulin therapies, which in turn impact differently on glucose levels. Hence, *in silico* data collected by running two different scenarios allow to a richer and more realistic data set. Table 9.1 shows Scenario 1, the training scenario, which is a 4-day protocol simulated in closed-loop using the MPC described in Section 5.2.2 to define the optimal insulin therapy. The first three days are used for model training, while the remaining day is used for validation purposes. The training scenario involves three meals per day with additional snacks in each day. Moreover, in order to define a real-life scenario, possible errors in the meal announcement are included, i.e. a limited events of unannounced meals or meals announced with a wrong estimation of the amount. Scenario 2 lasts three days and it is reported in Table 9.2. The meal amounts and times of this protocol are designed to reproduce a real-life scenario. Hypotreatments of 15 g are administrated to the patient in case glucose concentration falls below 65 mg/dl in both scenarios. Scenario 1 is used to perform model training and validation, while Scenario 2 is exploited to assess the prediction capabilities of the proposed model. Specifically, Scenario 2 is defined to reproduce eating patterns different from those present in Scenario 1. Moreover, an *in vivo* dataset is considered and it is composed of clinical data of a single T1D patient of the Padova clinical centre collected during experiments involved in the "AP@home" project [17]. This dataset is challenging because the clinical trial has been conducted in free-living conditions and it is the one used to identify the classic mathematical model described in Chapter 8 [101], which represents the state-of-art reached so far via the classic identification techniques developed in this thesis. Testing on a dataset not

	Time	CHO [g]	Insulin Bolus
Day 1	08:00	60	Bolus on time
	13:00	60	Bolus at 14:00
	17:00	30	No bolus
	20:00	80	Bolus on time
Day 2	08:00	50	Bolus on time
	10:00	15	No bolus
	13:00	35	Bolus on time
	19:00	80	Bolus on time
	22:00	20	Bolus on time
Day 3	06:30	40	Bolus on time
	09:30	20	No bolus
	12:30	45	Bolus at 12:00 for 50 g
	17:00	20	Bolus on time
	20:00	70	Bolus at 20:30
	23:00	20	No bolus
Day 4	08:00	35	Bolus on time
	11:30	20	Bolus on time
	13:30	60	Bolus at 13:30 for 30 g
	16:30	20	No bolus
	20:30	90	Bolus on time

TABLE 9.1: Training Scenario.

belonging to the training set and acquired following a real-life scenario allows the evaluation of the robustness of this approach to new unseen data and subjects.

9.4 Parameter Settings and Evaluation Protocol

The accuracy of the model predictions is assessed by considering various PHs expressed in terms of minutes. For this application, PHs from 5 minutes to 60 minutes are considered. Since the proposed model is aimed to be included in the MPC controller, that is characterized by a sample time of $T_s = 5$ minutes for the predictions, the vector of the considered PHs is

	Time	CHO [g]	Insulin Bolus
Day 1	08:00	50	Bolus on time
	13:00	50	Bolus on time
	19:00	70	Bolus on time
	23:00	20	Bolus on time
Day 2	06:30	50	Bolus on time
	09:30	15	No bolus
	13:00	60	Bolus at 12:00 for 50 g
	17:00	25	Bolus on time
	20:00	90	Bolus at 20:15 for 70 g
	23:00	15	No bolus
Day 3	08:30	50	Bolus on time
	11:30	20	Bolus on time
	14:00	60	Bolus at 13:00 for 30 g
	17:00	20	No bolus
	20:30	100	Bolus on time

TABLE 9.2: Testing Scenario.

$PH = [5, 10, \cdots, 60]$ and the predicted signals are sampled every T_s. For a given patient p and a specific PH $\in PH$, denote with $\widehat{\mathbf{ig}}(\cdot, \mathrm{PH})$ the PH-steps ahead prediction on the entire testing scenario, $\mathbf{ig}(\cdot)$ the considered reference, and $\overline{\mathrm{ig}}$ its average value. The predictions of the model are evaluated in terms of COD, FIT, and RMSE. These metrics are the standards used to evaluate performance in system identification [94, 113] described in Section 6.3 and they allow a fair comparison of the proposed model with respect to previously published solutions [101]. These metrics are defined as follows:

$$\begin{aligned} COD_p(\mathrm{PH}) &= 100 * \left(1 - \frac{||\widehat{\mathbf{ig}}(j, \mathrm{PH}) - \mathbf{ig}(j)||^2}{||\,\mathbf{ig}(j) - \overline{\mathrm{ig}}||^2}\right) \\ FIT_p(\mathrm{PH}) &= 100 * \left(1 - \frac{||\widehat{\mathbf{ig}}(j, \mathrm{PH}) - \mathbf{ig}(j)||}{||\,\mathbf{ig}(j) - \overline{\mathrm{ig}}||}\right) \\ RMSE_p(\mathrm{PH}) &= \frac{1}{N_{\mathrm{sample}}} ||\,\mathbf{ig}(j) - \widehat{\mathbf{ig}}(j, \mathrm{PH})|| \end{aligned} \tag{9.8}$$

where $j = \mathrm{PH}, \mathrm{PH} + T_s, \cdots, N_{\mathrm{sample}} \cdot T_s$ is used to index the considered samples, and N_{sample} is the number of samples of the signal when sampled every T_s minutes. The average value of

each metric ($\overline{COD}$,$\overline{FIT}$,$\overline{RMSE}$) for all PH is used as main outcome to evaluate the overall performance as follows:

$$\begin{aligned}
\overline{COD} &= \frac{1}{N_{\mathrm{PH}}} \sum_{i=1}^{N_{\mathrm{PH}}} \left(\frac{1}{N_p} \sum_{p=1}^{N_p} COD_p(\mathrm{PH}(i)) \right) \\
\overline{FIT} &= \frac{1}{N_{\mathrm{PH}}} \sum_{i=1}^{N_{\mathrm{PH}}} \left(\frac{1}{N_p} \sum_{p=1}^{N_p} FIT_p(\mathrm{PH}(i)) \right) \\
\overline{RMSE} &= \frac{1}{N_{\mathrm{PH}}} \sum_{i=1}^{N_{\mathrm{PH}}} \left(\frac{1}{N_p} \sum_{p=1}^{N_p} RMSE_p(\mathrm{PH}(i)) \right)
\end{aligned} \tag{9.9}$$

where N_{ph} is the total number of PH $\in PH$ (i.e. $N_{\mathrm{PH}} = 12$) and N_p is the total number of patients involved in each testing experiment.

9.5 Results

In order to maximize the generalization capability of the proposed algorithm, a training and a testing data sets–large enough to describe the dynamics under study–are required. Since these data sets are supposed to be fairly different, two disjoint scenarios have been adopted to simulate the realistic intra-subject change in eating habits in terms of timing and meal size variations. Different food habits imply different insulin therapies and a different glucose control. The model has been trained on different patients with eating habits drastically different with respect to those observed in the testing scenario by running two different scenarios, which do not contain the same set of meals. To do so, the population of 100 adults is split in two parts ($N_p = 50$): the model is trained on the first N_p patients in Scenario 1, and the tests are conducted on the second half of the patients in Scenario 2. The same is performed but considering the other half of patients in each scenario. The final results are obtained by averaging the performance from the two different train and test groups. This data separation has been chosen in order to test the capability of the model to represent subjects not belonging to the training set but also to check the model robustness against a variation in meal sizes and correlation of meal sizes between a day. The same experiments have also been performed by testing the two trained models described above on the real patient (i.e. *in vivo*) and taking the average of the two results. This experiment has been performed in order to assess the generalization capability of the model in a real-life scenario.

Firstly, the study focuses on the choice of both the size of the hidden units in each layer ($d \in \{16, 32, 64\}$) and the number of LSTM layers ($n \in \{1,2,3\}$). The choice of powers of 2 in

		$\overline{COD}$ [%]			$\overline{FIT}$ [%]			$\overline{RMSE}$		
	d \ n	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
SILICO	**16**	79.48	79.27	79.12	55.93	55.59	55.40	12.57	12.65	12.72
	32	81.31	81.14	79.60	57.94	57.12	56.27	11.93	11.79	12.41
	64	82.10	81.93	81.32	58.84	58.66	57.98	11.68	11.72	11.93
VIVO	**16**	67.67	71.92	71.61	45.03	48.17	47.92	29.09	27.43	27.56
	32	67.20	72.57	72.05	44.96	48.93	48.65	29.13	27.03	27.18
	64	62.06	71.70	72.50	41.41	48.44	49.17	31.01	27.29	26.91

TABLE 9.3: Performance metrics on each combination of $d - n$.

the number of hidden units follows a standardized practice. The performance of the scenarios characterized by either $n >= 3$ and $d >= 128$ have also been studied. In both scenarios the performance dropped sharply and the training time increase significantly, so the results are not reported here. From Table 9.3 it is possible to observe that increasing the number of hidden units for each LSTM entails a slight improvement of the performance indices while the number of stacked LSTM does not significantly affect the final performance. Generally speaking models with more parameters are able to improve prediction performance only up to a point, that is when the amount and variability of available data is sufficient to train the model. The result presented in this section suggest that given the available data, the best configuration is the one with a single LSTM and $d = 64$. However, the performance of the single-LSTM implementation drops sharply (by more than 5%) on the real patient. As the only significant difference is in the number of parameters, this behavior is associated to over fitting on the training set. This behavior on the real patients was not observed on models with multiple LSTMs. For these reasons all subsequent experiments have been performed with the configuration $n = 2, d = 64$.

Secondly, an analysis has been performed regarding how the different features considered as input for the network influence the final performance. For this reason, past insulin ($\overleftarrow{\mathbf{i}}$) and carbohydrates ($\overleftarrow{\mathbf{m}}$) have been removed from the input stream. Denoting with $\overleftarrow{\mathbf{X}}^*$ the modified input array, the vector representing past information given to the model is defined as follows:

$$\overleftarrow{\mathbf{X}}^*(t_0, \text{PH}) = \left[\overleftarrow{\mathbf{CGM}}(t_0, \text{PH}) \right]. \tag{9.10}$$

Figure 9.3 shows the comparison of the prediction performance of the models with

- $\left[\overleftarrow{\mathbf{X}^*}(t_0, \text{PH}) \;\; \overrightarrow{\mathbf{X}}(t_0, \text{PH})\right]$
- $\left[\overleftarrow{\mathbf{X}}(t_0, \text{PH}) \;\; \overrightarrow{\mathbf{X}}(t_0, \text{PH})\right]$

as inputs, respectively. It denotes that the introduction of the information regarding the past insulin therapy and ingested meals guarantees an improvement in the prediction performance of the model. In particular, considering the results obtained by using *in silico* data, there is a slight improvement in performance. Indeed, the virtual patients belonging to the training and testing groups are subsets of the same population.

Since the LSTM is able to learn the behaviour of the population, the additional information about past history does not provide a significant improvement. On the other hand, Figure 9.3 shows a significant gap in performance computed on real-life data. Indeed, if the LSTM is trained on *in silico* data, the lack of past therapy information lowers the performance on *in vivo* testing dataset. Hence, the past evidence $\overleftarrow{\mathbf{i}}$ and $\overleftarrow{\mathbf{m}}$ help mitigating the differences in the data distribution. The model obtained considering these additional information is able to generalize to new unseen data and improve the overall glucose control.

9.6 Discussion

The proposed solution is a population average model identified on the 100 adults of the UVA/-Padova simulator. An average model could ideally limit the performance since it describes the average dynamics of the population, so it is necessary to test both its prediction and generalization capabilities on a dataset different from the one used in training. Firstly, the model has been re-trained considering the entire adult virtual population as a training dataset in order to maximize the available information provided to the training procedure. Then, the proposed algorithm is tested on a 1-month dataset containing all the data collected during the clinical trial [59] for a single patient. This dataset is challenging because it includes eating patterns not present in the training dataset, but also it includes all the problems experienced during the clinical trial.

The performance obtained by the DFG model and the linearized average model (AVG) of the UVA/Padova simulator are reported in Table 9.4. By considering the first two rows of Table 9.4, the DFG model shows superior prediction performance against the AVG. Moreover,

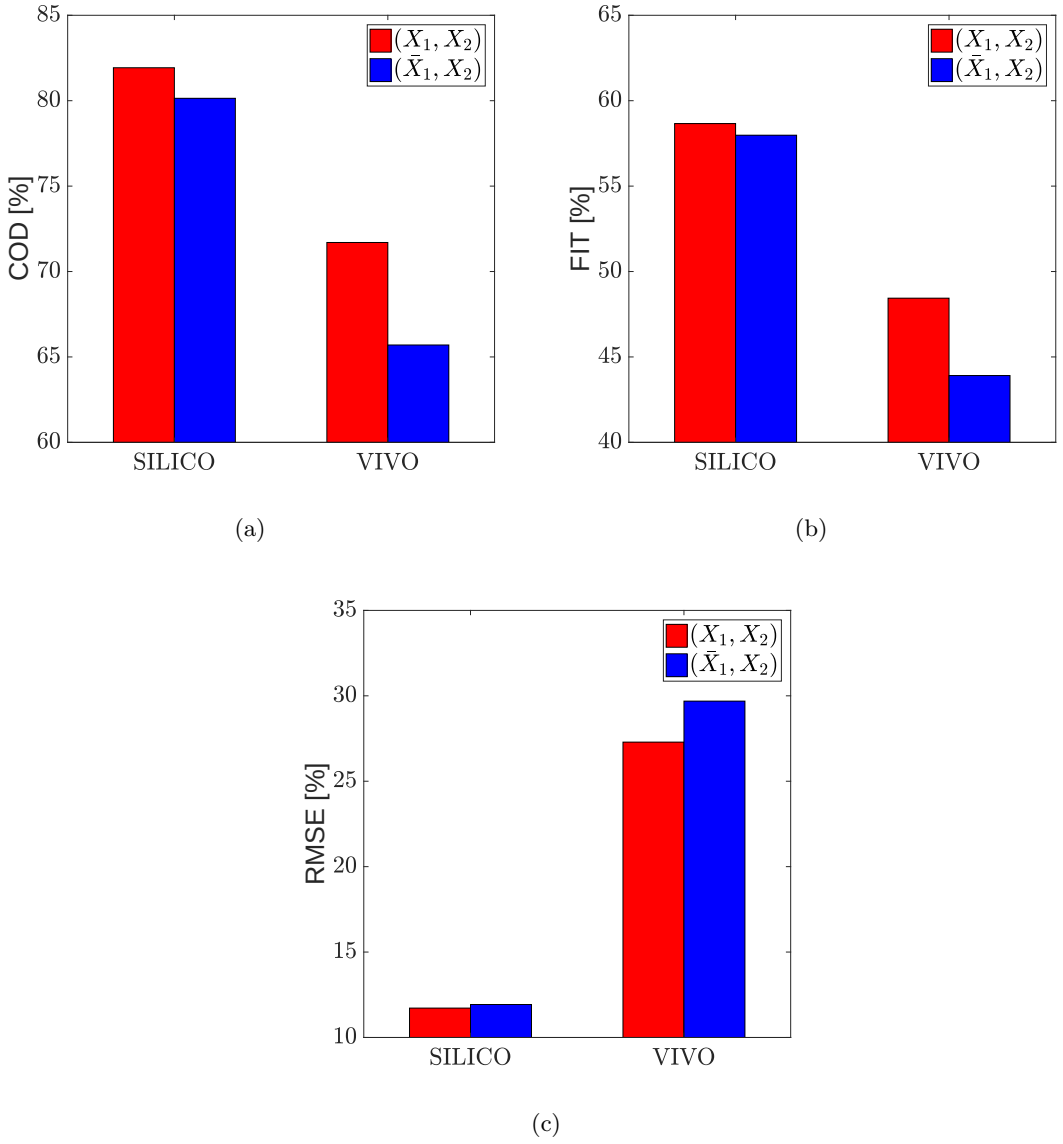


FIGURE 9.3: Comparison of the prediction performance on both *in silico* and *in vivo* testing data in term of COD (a), FIT (b) and RMSE (c) on Scenario 2 and real-life data.

DGF approach proves to be able to generalize over different datasets by achieving interesting improvements with respect to AVG, despite being an average model.

9.6.1 Fine tuning

The drawback of an average model is that it cannot fully describe the variety of individual glucose response of the entire population. The definition of an individualized insulin therapy by exploiting a patient-tailored model can substantially improve the effectiveness of the glucose control as showed in Chapters 6 and 7. Hence, in order to improve the DGF model performance, the LSTM trained on *in silico* data is fine-tuned on data of the specific patient.

Predictor	**COD** [%]	**FIT** [%]	**RMSE**
AVG	15.23	11.48	46.82
DGF	69.85	48.44	27.29
DMP [101]	80.79	61.44	–
DGF Fine Tuned	73.23	49.69	26.63
DGF Fine-tuned + Exp. Filtering	84.05	60.14	21.09

TABLE 9.4: Prediction matrices.

In this context, Fine-Tuning (FT) slightly modifies the weights of the LSTM in order better fit the individual behaviour of the real patient. For the purpose of this work, the partial retraining entailed by any FT has been extended to the entire architecture, by using a suitably small learning rate (10^{-5}) for 10 additional epochs on the FT dataset. A filtered version of the collected CGM data (ig_R) is used as reference for the signal *ig* since this signal is not available *in vivo*. The signal ig_R is obtained by using a retrofitting algorithm, which is able to reconstruct an accurate continuous-time BG profile by exploiting BG samples from the fingerstick and CGM data from the sensor [92]. In order to provide a good amount of information to describe the dynamic of a specific patient, the FT dataset contains two days picked up randomly among the available ones.

The FT technique improves the accuracy by 6% on the 2-layer stacked LSTM, as reported in the row 4 of Table 9.4. Figure 9.4 reports the performance metrics as a function of PH for the 2-layer stacked LSTM with *in vivo* testing dataset. The higher the PH, the lower the performance. This is motivated by the fact that the further you want to predict the more difficult it becomes. The improvements of the fine-tuned models are evident for large PH values where the performance increases with respect to the original model without FT. These improvements are less obvious if their performance are compared against the performance of the individualized model (Daily Model Predictor, DMP) presented in Section 8.3 and reported in Table 9.4. The individualized model is identified from real-data on the base of the a-priori knowledge acquired through the analysis of the patient real-data, while here real-data are used to adjust the model pre-trained on *in silico* data.

9.6.2 Output filtering

The main limitation of the proposed approach is that the network is trained on noisy input measurements but a noiseless signal is required as output. Since the measurement noise that

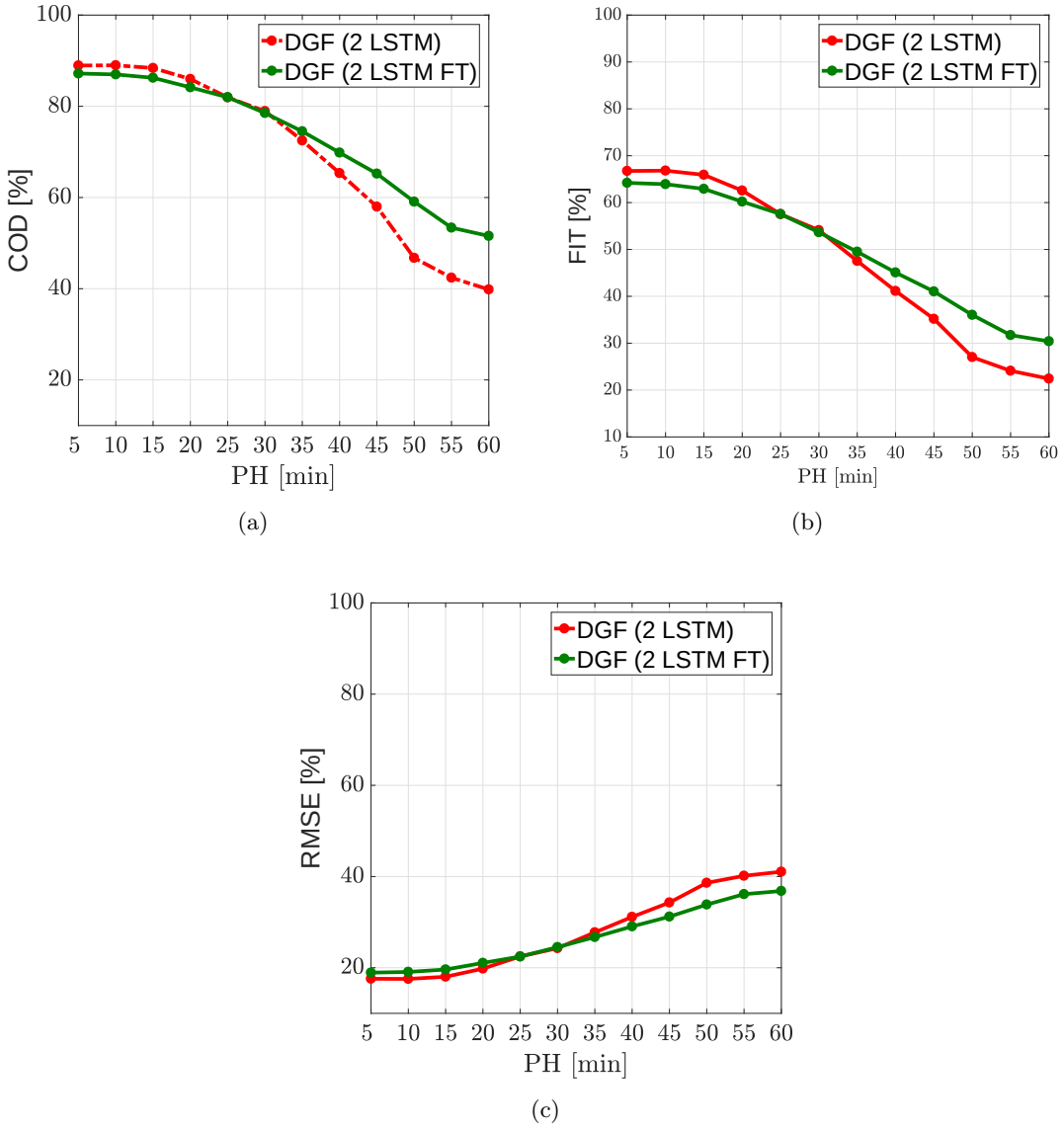


FIGURE 9.4: A comparison of the prediction performance of 2-layer stacked LSM compared to its fine-tuned version in term of COD (a), FIT (b) and RMSE (c) on the clinical dataset.

affects the CGM data is not negligible, the network would try to reduce the noise on the output in order to improve the overall quality of the prediction. However, this effect can be limited and in order to further improve the prediction smoothness an exponential filtering is applied a posteriori to the predictions. The exponential filter decreases gradually the weights on the past observations and, considering the decay of the effects of meals and insulin on the glucose, it represents the natural choice for this kind of applications. The exponential filter also allows to forget erroneous values predicted in the previous steps.

The predicted $\hat{ig}(\cdot, \text{PH})$ is filtered by the following exponential filter:

$$\hat{ig}_{\text{exp}}(t_0+k, ph) = \alpha \cdot \hat{ig}(t_0+k, ph) + (1-\alpha) \cdot \hat{ig}_{\text{exp}}(t_0+k-1, ph) \qquad k = 0, 1, \cdots, \text{PH}-1 \quad (9.11)$$

where $\alpha \in (0, 1)$ is the smoothing factor and is defined as follows:

$$\alpha = \frac{2}{w_{\text{exp}} + 1} \tag{9.12}$$

and w_{exp} is the length of the window used by the filter. It is set to 5, i.e. the minimum window observed by the model and it is kept fixed for all models. Figure 9.5 shows the noisy CGM data, the output of the retrofitting algorithm ig_R, which represents the ground truth, and the output of the exponential filter $\widehat{\text{ig}}_{\text{exp}}$. It may be noted in Figure 9.5 that there is a scaling problem in the signal $\widehat{\text{ig}}_{\text{exp}}$.

This effect can be explained by considering that $\widehat{\text{ig}}_{\text{exp}}$ is rescaled with the minimum and maximum values of **ig** computed in training while these two values for the test data cannot be known *a priori*. However, it is important to highlight that the glucose metabolism is highly affected by the food and lifestyle habits of the specific patient. This implies that the range of the injected insulin, carbohydrate amounts and glucose levels are individual characteristics of the patient. In order to cope with this problem, a larger dataset of the patient is required to capture the individual variability without compromising the testing dataset. The last row of Table 9.4 shows the results obtained by applying the exponential filter to the output of the 2-layer stacked LSTM model. It is evident that the filtering technique is able to highly improve the performance by partially removing the noise that affects the CGM data. The proposed approach concerned with therapy-driven deep glucose forecasting has been published in [114]. However, the use of a single real patient for the final validation of the model is the main limitation of this study, but the strength of the results obtained is enforced by the large changes of the patient habits and the time-varying nature of the system under study.

In view of the use of this model for the synthesis of the MPC presented in Section 5.2, it is worth to be noted that this process is a challenging task. The implementation of a nonlinear MPC is heavy from a computational point of view because the control law cannot be defined in closed-form and an optimization problem has to be solved online. Future work will be devoted to this purpose.

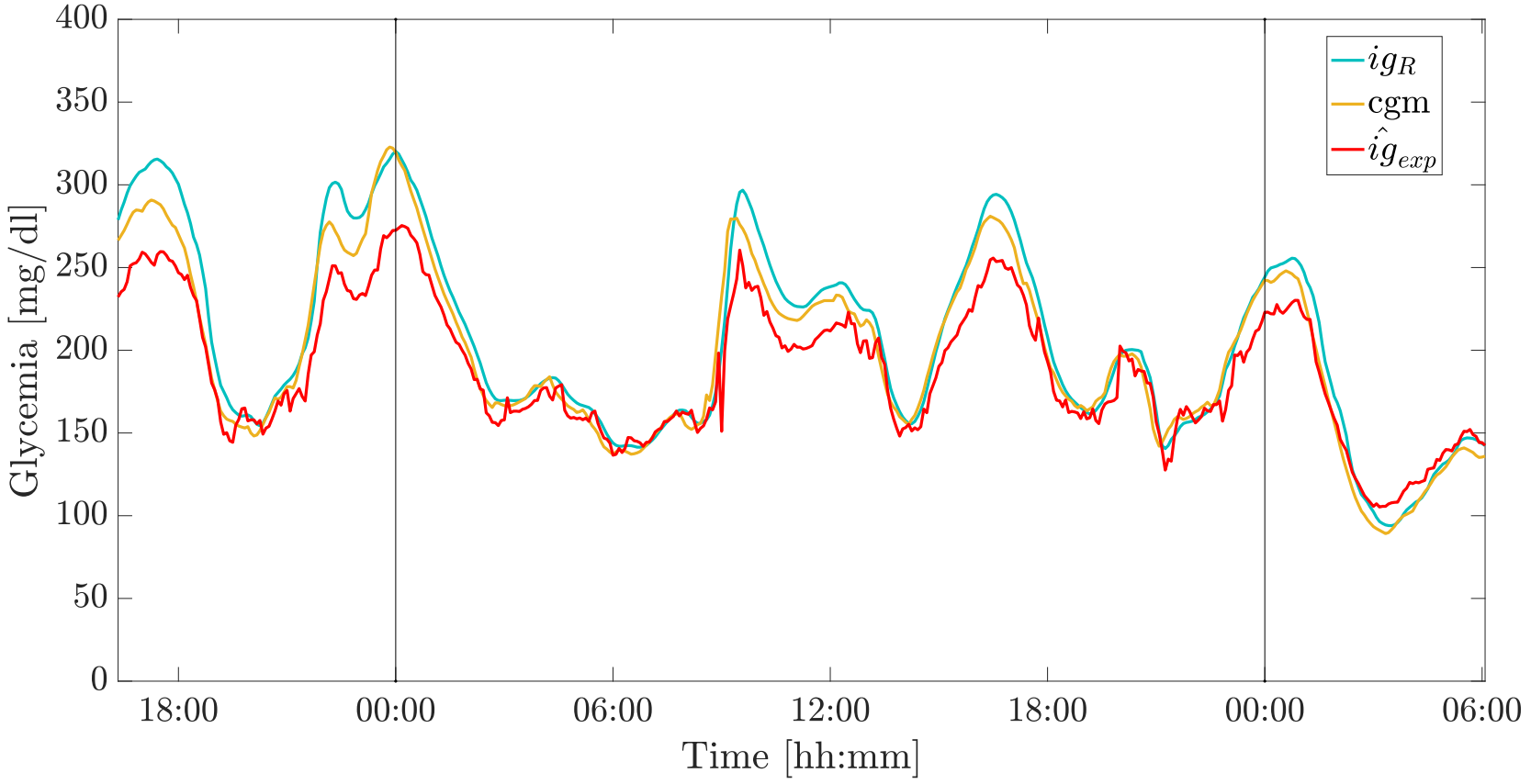


FIGURE 9.5: Glucose profile of a 24h study-case: the yellow line is the noisy CGM data, the light blue line is the output of the retrofitting algorithm ig_R presented as reference, and the red line is the output of the exponential filter $\widehat{\mathrm{ig}}_{\mathrm{exp}}$.

Chapter 10

Conclusions

The improvements in the safety and the reliability of the artificial pancreas system have been allowed by an effective control approach usable for long periods by the patients during their normal lives. The aim of the presented research is the improvement of the overall glucose control performance achieved by acting on the conventional therapy, and by identifying individualized models of T1D patients from real-life data to be exploited in the MPC strategy and in the alarm systems.

Regarding the post-prandial glucose regulation, two new approaches based on machine-learning methodologies have been proposed. A multiple KNN classification algorithm able to predict post-prandial glucose profile due to the nominal therapy is proposed in order to compute corrections to time and/or amount of the meal bolus. Satisfactory results have been obtained in terms of reduction of the average glucose and of hyperglycemia, and in terms of increment of the time in target with limited increase of hypoglycemia. In order to handle the inter-patient variability, an individualized multi-linear regression model able to enhance conventional therapy has been designed. The model identification was performed in two steps and the model validation was computed in an ideal and in perturbed scenarios. A test study demonstrated that this approach is able to successfully handle a personalized post-prandial glucose regulation but requires potentially dangerous experiments on real patient. New researches can be conducted in this prospective.

A new individualization technique has been defined to be used on free-living data collected without ad-hoc clinical protocols. This approach has been used to identify patient-tailored models for all patients belonging to the Amsterdam medical centre. The individualized models are compared with the "average" model used to synthetize the MPC controller employed in the trial. The patient-tailored models show a performance improvement in thier prediction capabilities. Moreover, the identified patient-tailored models have been used to develop an individualized

alarm system for each patient. The use of individualized models showed an high sensitivity of the system on the entire cohort of patients. Most of the hypoglycemia events (85%) occurred during the trial have been detected in time to allow a rescue action.

In order to fully exploit the available data, a complete data analysis on the entire group of patients involved in the trial has been presented. The relation between day period and the characteristics of the meal response already observed *in silico* has been detected also in several real subjects. These characteristics have been exploited to define a new predictor based on different models to predict the glucose during different day periods. This predictor is compared with a predictor based on a unique daily model identified on an entire day with satisfactory results.

Moreover, deep learning technique have been explored to derive a therapy-driven approach in order to predict a trend of future glucose concentration in T1D patients. The approach entails multiple models trained on the *in silico* adult patients of the UVA/Padova simulator. Each model is used to predict a glucose profile for a fixed prediction horizon and the individual predictions are aggregated to obtain a profile of future glucose levels. The achieved results show that the proposed approach can significantly improve predictive performance despite being an average model. In order to individualize the trained models, fine-tuning is applied to each model separately considering a small portion of the data pertaining to a specific patient. Satisfactory results have been obtained in terms of prediction capabilities.

Future works will be focused on the synthesis of individualized MPC algorithms based on these individualized models, on the investigation of new model identification techniques based on dynamical model decomposition, and on the definition of the control law for intraperitoneal insulin delivery for new protptypes of intraperitoneal pumps under development.

Bibliography

[1] Centers for Disease Control and Prevention *et al.*, "National diabetes statistics report, 2017," *Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services*, 2017.

[2] G. Roglic *et al.*, "Who global report on diabetes: A summary," *International Journal of Noncommunicable Diseases*, vol. 1, no. 1, p. 3, 2016.

[3] D. Dabelea, E. J. Mayer-Davis, S. Saydah, G. Imperatore, B. Linder, J. Divers, R. Bell, A. Badaru, J. W. Talton, T. Crume, *et al.*, "Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009," *Jama*, vol. 311, no. 17, pp. 1778–1786, 2014.

[4] E. Smalley, "Medtronic automated insulin delivery device gets fda nod," 2016.

[5] P. Soru, G. De Nicolao, C. Toffanin, C. Dalla Man, C. Cobelli, L. Magni, A. home consortium, *et al.*, "MPC based artificial pancreas: Strategies for individualization and meal compensation," *Annual Reviews in Control*, vol. 36, no. 1, pp. 118–128, 2012.

[6] U.S. Department of Health and Human Services, "Nih study shows how insulin stimulates fat cells to take in glucose." `https://www.nih.gov/news-events/`,.

[7] H. Lebovitz, "Insulin resistance: definition and consequences," *Experimental and clinical endocrinology & diabetes*, vol. 109, no. Suppl 2, pp. S135–S148, 2001.

[8] Centers for Disease Control and Prevention *et al.*, "National diabetes statistics report: estimates of diabetes and its burden in the united states, 2014," *Atlanta, GA: US Department of Health and Human Services*, vol. 2014, 2014.

[9] E. Bonora and J. Tuomilehto, "The pros and cons of diagnosing diabetes with a1c," *Diabetes care*, vol. 34, no. Supplement 2, pp. S184–S190, 2011.

[10] American Diabetes Association *et al.*, "6. glycemic targets," *Diabetes care*, vol. 38, no. Supplement 1, pp. S33–S40, 2015.

[11] The Diabetes Control and Complications Trial Research Group, "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus," *The New England Journal of Medicine*, vol. 329, no. 14, pp. 977–986, 1993.

[12] R. Kahn, "Follow-up report on the diagnosis of diabetes mellitus: the expert committee on the diagnosis and classifications of diabetes mellitus," *Diabetes care*, vol. 26, no. 11, p. 3160, 2003.

[13] D. Nathan, P. Cleary, J. Backlund, S. Genuth, J. Lachin, T. Orchard, P. Raskin, B. Zinman, D. Control, and DCCT/EDIC Study ResearchGroup, "Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes," *New England Journal of Medicine*, vol. 353, no. 25, 2005.

[14] N. White, W. Sun, P. Cleary, R. Danis, M. Davis, D. Hinsworth, L. Hubbard, J. Lachin, and D. Nathan, "Prolonged effect of intensive therapy on the risk of retinopathy complications in patients with type 1 diabetes mellitus: 10 years after the diabetes control and complications trial," *Archives of Ophthalmology*, vol. 126, no. 12, 2008.

[15] H. Zisser, L. Jovanovič, F. J. Doyle III, P. Ospina, and C. Owens, "Run-to-run control of meal-related insulin dosing," *Diabetes Technology & Therapeutics*, vol. 7, no. 1, pp. 48–57, 2005.

[16] C. C. Palerm, H. Zisser, L. Jovanovič, and F. J. Doyle III, "A run-to-run framework for prandial insulin dosing: handling real-life uncertainty," *International Journal of Robust and Nonlinear Control*, vol. 17, no. 13, pp. 1194–1213, 2007.

[17] C. C. Palerm, H. Zisser, L. Jovanovič, and F. J. Doyle III, "A run-to-run control strategy to adjust basal insulin infusion rates in type 1 diabetes," *Journal of Process Control*, vol. 18, no. 3–4, pp. 258–265, 2008.

[18] C. Toffanin, R. Visentin, M. Messori, F. Di Palma, L. Magni, and C. Cobelli, "Towards a run-to-run adaptive artificial pancreas: In silico results," *IEEE Transactions on Biomedical Engineering*, 2017.

[19] P. C. Davidson, H. R. Hebblewhite, B. W. Bode, R. D. Steed, N. S. Welch, and M. C. Greenlee et al., "Statistically based CSII parameters: correction factor and CF (1700 rule) and carbohydrate-to-insulin ratio and CIR (2.8 rule) and and basal-to-total ratio," *Diabetes Technology & Therapeutics*, vol. 5, no. 5, p. A237, 2003.

[20] S. R. Patton and M. A. Clements, "Continuous glucose monitoring versus self-monitoring of blood glucose in children with type 1 diabetes-are there pros and cons for both?," *US endocrinology*, vol. 8, no. 1, p. 27, 2012.

[21] L. Olansky and L. Kennedy, "Finger-stick glucose monitoring issues of accuracy and specificity," *Diabetes Care*, vol. 33, no. 4, pp. 948–949, 2010.

[22] R. Retnakaran and B. Zinman, "Treatment of type 1 diabetes mellitus in adults," in *Endocrinology: Adult and Pediatric (Seventh Edition)*, pp. 770 – 787, Philadelphia: W.B. Saunders, seventh edition ed., 2016.

[23] H. Kirchsteiger, J. B. Jørgensen, E. Renard, and L. del Re, *Prediction Methods for Blood Glucose Concentration*. Springer, 2016.

[24] B. P. Kovatchev, M. D. Breton, C. Dalla Man, and C. Cobelli, "In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes," *Journal of Diabetes Science and Technology*, vol. 3, no. 1, pp. 44–55, 2009.

[25] R. Visentin, C. Dalla Man, B. Kovatchev, and C. Cobelli, "The university of uva/padova type 1 diabetes simulator matches the glucose traces of a clinical trial," *Diabetes technology & therapeutics*, vol. 16, no. 7, pp. 428–434, 2014.

[26] B. P. Kovatchev, M. D. Breton, C. Dalla Man, and C. Cobelli, "In silico model and computer simulation environment approximating the human glucose/insulin utilization," *Food and Drug Administration*, vol. Master File MAF 1521, 2008.

[27] C. Dalla Man, F. Micheletto, D. Lv, M. Breton, *et al.*, "The UVA/Padova type 1 diabetes simulator: New features," *Journal of Diabetes Science and Technology*, vol. 8, no. 1, pp. 26–34, 2014.

[28] R. Visentin, C. Dalla Man, Y. C. Kudva, A. Basu, and C. Cobelli, "Circadian variability of insulin sensitivity: Physiological input for in silico artificial pancreas," *Diabetes Technology & Therapeutics*, vol. 17, no. 1, pp. 1–7, 2015.

[29] R. Visentin, C. Dalla Man, and C. Cobelli, "One-day bayesian cloning of type 1 diabetes subjects: toward a single-day uva/padova type 1 diabetes simulator," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, pp. 2416–2424, 2016.

[30] M. Messori, C. Toffanin, S. Del Favero, G. De Nicolao, *et al.*, "Model individualization for artificial pancreas," *Computer Methods and Programs in Biomedicine*, 2016.

[31] S. Del Favero, G. Pillonetto, C. Cobelli, and G. De Nicolao, "A novel nonparametric approach for the identification of the glucose-insulin system in type 1 diabetic patients," *20th IFAC World Congress*, pp. 8340–8346, 2011.

[32] A. J. Laguna, P. Rossetti, F. J. Ampudia-Blasco, J. Vehí, and J. Bondia, "Identification of intra-patient variability in the postprandial response of patients with type 1 diabetes," *Biomedical Signal Processing and Control*, vol. 12, pp. 39–46, 2014.

[33] A. J. Laguna, P. Rossetti, F. J. Ampudia-Blasco, J. Vehí, and J. Bondia, "Experimental blood glucose interval identification of patients with type 1 diabetes," *Journal of Process Control*, vol. 24, no. 1, pp. 171–181, 2014.

[34] R. Visentin, E. Campos-Náñez, M. Schiavon, D. Lv, M. Vettoretti, M. Breton, B. Kovatchev, C. Dalla Man, and C. Cobelli, "The UVA/Padova type 1 diabetes simulator goes from single meal to single day," *Journal of Diabetes Science and Technology*, vol. 12, pp. 273–281, 2018.

[35] C. Dalla Man and C. Cobelli, "A system model of oral glucose absorption: Validation on gold standard data," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2472–2478, 2006.

[36] C. Dalla Man, R. A. Rizza, and C. Cobelli, "Meal simulation model of the glucose-insulin system," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 10, pp. 1740–1749, 2007.

[37] C. Dalla Man, G. Toffolo, R. Basu, R. A. Rizza, and C. Cobelli, "A model of glucose production during a meal," in *Proceedings of the 28th IEEE EMBS Annual International Conference*, pp. 5647–5650, 2006.

[38] B. P. Kovatchev, M. Straume, D. J. Cox, and L. S. Farhy, "Risk analysis of blood glucose data: A quantitative approach to optimizing the control of insulin dependent diabetes," *Journal of Theoretical Medicine*, vol. 3, no. 1, pp. 1–10, 2000.

[39] L. Hinshaw, C. Dalla Man, D. K. Nandy, A. Saad, A. E. Bharucha, J. A. Levine, R. A. Rizza, R. Basu, R. E. Carter, C. Cobelli, Y. C. Kudva, and A. Basu, "Diurnal pattern of insulin action in type 1 diabetes: Implications for a closed-loop system," *Diabetes*, vol. 62, no. 7, pp. 2223–2229, 2013.

[40] R. Basu, C. Dalla Man, M. Campioni, A. Basu, G. Klee, G. Toffolo, C. Cobelli, and R. A. Rizza, "Effects of age and sex on postprandial glucose metabolism: Differences in glucose

turnover and insulin secretion and insulin action and and hepatic insulin extraction," *Diabetes*, vol. 55, no. 7, pp. 2001–2014, 2006.

[41] C. Toffanin, M. Messori, F. Di Palma, G. De Nicolao, C. Cobelli, and L. Magni, "Artificial pancreas: Model predictive control design from clinical experience," *Journal of Diabetes Science and Technology*, vol. 7, no. 6, pp. 1470–1483, 2013.

[42] D. M. Maahs, B. A. Buckingham, J. R. Castle, A. Cinar, E. R. Damiano, E. Dassau, J. H. DeVries, and F. J. Doyle III et al., "Outcome measures for artificial pancreas clinical trials: A consensus report," *Diabetes Care*, vol. 39, no. 7, pp. 1175–1179, 2016.

[43] B. P. Kovatchev, D. J. Cox, A. Kumar, L. Gonder-Frederick, and W. L. Clarke, "Algorithmic evaluation of metabolic control and risk of severe hypoglycemia in type 1 and type 2 diabetes using self-monitoring blood glucose data," *Diabetes Technology & Therapeutics*, vol. 5, no. 5, pp. 817–828, 2003.

[44] B. P. Kovatchev, D. J. Cox, L. A. Gonder-Frederick, D. Young-Hyman, D. Schlundt, and W. Clarke, "Assessment of risk for severe hypoglycemia among adults with iddm: validation of the low blood glucose index," *Diabetes Care*, vol. 21, no. 11, pp. 1870–1875, 1998.

[45] D. Freedman, R. Pisani, and R. Purves, "Statistics (international student edition)," *Pisani, R. Purves, 4th edn. WW Norton & Company, New York*, 2007.

[46] M. L. McHugh, "Multiple comparison analysis testing in anova," *Biochemia medica: Biochemia medica*, vol. 21, no. 3, pp. 203–209, 2011.

[47] K. Turksoy, E. Littlejohn, and A. Cinar, "Multimodule, multivariable artificial pancreas for patients with type i diabetes," *IEEE Control Systems Magazine*, vol. 38, no. 1, pp. 105–124, 2018.

[48] K. Kolle, T. Biester, S. Christiansen, A. L. Fougner, and O. Stavdahl, "Pattern recognition reveals characteristic postprandial glucose changes: Non-individualized meal detection in diabetes mellitus type 1," *IEEE Journal of Biomedical and Health Informatics*, 2019.

[49] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[50] N. Bhatia *et al.*, "Survey of nearest neighbor techniques," *arXiv preprint arXiv:1007.0085*, 2010.

[51] T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems," *Ieee Transaction on patter analysis and machine intelligence*, vol. 16, no. 1, pp. 66 – 75, 1994.

[52] A. Kataria and M. Singh, "A review of data classification using k-nearest neighbour algorithm," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 6, pp. 354–360, 2013.

[53] P. Hart, "The condensed nearest neighbors rule," *IEEE Transaction Information Theory*, vol. 14, no. 3, pp. 515–516, 1968.

[54] C. M. Florkowski, "Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: Communicating the performance of diagnostic tests," *Diabetes Care*, vol. 29, no. 1, pp. 83–87, 2008.

[55] E. M. Aiello, C. Toffanin, M. Messori, C. Cobelli, and L. Magni, "Postprandial glucose regulation via KNN meal classification in Type 1 diabetes," *IEEE Control Systems Letters*, vol. In Press, 2018.

[56] A. Gelman, J. B. Carlin, H. S. Stern, D. Dunson, A. Vehtari, and D. Rubin, "Bayesian data analysis, 3rd edition," pp. 353–380, 2013.

[57] M. Messori, M. Ellis, C. Cobelli, P. D. Christofides, and L. Magni, "Improved postprandial glucose control with a customized model predictive controller," *Procedings of the American Control Conference*, pp. 5108–5115, 2015.

[58] D. A. Finan, C. C. Palerm, F. J. Doyle III, D. E. Seborg, *et al.*, "Effect of Input Excitation on the Quality of Empirical Dynamic Models for Type 1 Diabetes," *AIChE Journal*, vol. 55, pp. 1135–1146, 2009.

[59] E. Renard, A. Farret, J. Kropff, D. Bruttomesso, M. Messori, J. Place, R. Visentin, R. Calore, C. Toffanin, F. Di Palma, *et al.*, "Day-and-night closed-loop glucose control in patients with type 1 diabetes under free-living conditions: results of a single-arm 1-month experience compared with a previously reported feasibility study of evening and night at home," *Diabetes Care*, vol. 39, no. 7, pp. 1151–1160, 2016.

[60] T. Danne, R. Nimri, T. Battelino, R. M. Bergenstal, K. L. Close, J. H. DeVries, S. Garg, L. Heinemann, I. Hirsch, S. A. Amiel, R. Beck, E. Bosi, B. Buckingham, C. Cobelli, *et al.*, "International consensus on use of continuous glucose monitoring," *Diabetes Care*, vol. 40, no. 12, pp. 1631–1640, 2017.

[61] C. Toffanin, S. Del Favero, E. Aiello, M. Messori, C. Cobelli, and L. Magni, "MPC model individualization in free-living conditions: A proof-of-concept case study," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1181–1186, 2017.

[62] C. Toffanin, S. Del Favero, E. Aiello, M. Messori, C. Cobelli, and L. Magni, "Glucose-insulin model identified in free-living conditions for hypoglycaemia prevention," *Journal of Process Control*, vol. 64, pp. 27–36, 2018.

[63] E. M. Aiello, Z. Wu, P. D. Christofides, C. Toffanin, C. Cobelli, and L. Magni, "Improving diabetes conventional therapy via machine learning modeling," in *2019 American Control Conference (ACC)*, pp. 4136–4143, IEEE, 2019.

[64] A. H. Kadish, "Automation control of blood sugar. i. a servomechanism for glucose monitoring and control," *The American journal of medical electronics*, vol. 3, pp. 82–86, 1964.

[65] A. M. Albisser, B. Leibel, T. Ewart, Z. Davidovac, C. Botz, W. Zingg, H. Schipper, and R. Gander, "Clinical control of diabetes by the artificial pancreas," *Diabetes*, vol. 23, no. 5, pp. 397–404, 1974.

[66] E. F. Pfeiffer, C. Thum, and A. H. Clemens, "The artificial beta cell – a continuous control of blood sugar by external regulation of insulin infusion (glucose controlled insulin infusion system)," *Hormone and Metabolic Research*, vol. 6, pp. 339–342, 1974.

[67] A. Clemens, "Feedback control dynamics for glucose controlled insulin infusion system.," *Medical progress through technology*, vol. 6, no. 3, pp. 91–98, 1979.

[68] J. C. Pickup, H. Keen, J. A. Parsons, and K. G. Alberti, "Continuous subcutaneous insulin infusion: an approach to achieving normoglycaemia," *The British Medical Journal*, vol. 1, no. 6107, pp. 204–207, 1978.

[69] C. Cobelli, E. Renard, and B. Kovatchev, "Artificial pancreas: past, present, future," *Diabetes*, vol. 60, pp. 2672–2682, Nov 2011.

[70] A. Fougner, K. Kölle, H. E. Frøyen, R. Ellingsen, S. M. Carlsen, and Ø. Stavdahl, "In–silico comparison of intravenous, intraperitoneal, subcutaneous and combined approaches for closed–loop glucose control," *Journal of Diabetes Science and Technology*, vol. 1, no. 3, pp. 323–330, 2007.

[71] E. Dassau, E. Renard, J. Place, A. Farret, M.-J. Pelletier, J. Lee, L. M. Huyett, A. Chakrabarty, F. J. Doyle III, and H. C. Zisser, "Intraperitoneal insulin delivery provides superior glycaemic regulation to subcutaneous insulin delivery in model predictive

control-based fully-automated artificial pancreas in patients with type 1 diabetes: a pilot study," *Diabetes, Obesity and Metabolism*, vol. 19, no. 12, pp. 1698–1705, 2017.

[72] N. A. Spaan, A. E. Teplova, E. Renard, and J. A. Spaan, "Implantable insulin pumps: an effective option with restricted dissemination," *The Lancet Diabetes & Endocrinology*, vol. 2, no. 5, pp. 358–360, 2014.

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

[74] H. Thabit, M. Tauschmann, J. M. Allen, L. Leelarathna, S. Hartnell, M. E. Wilinska, C. L. Acerini, S. Dellweg, C. Benesch, L. Heinemann, *et al.*, "Home use of an artificial beta cell in type 1 diabetes," *New England Journal of Medicine*, vol. 373, no. 22, pp. 2129–2140, 2015.

[75] J. Kropff, S. Del Favero, J. Place, C. Toffanin, R. Visentin, M. Monaro, M. Messori, F. Di Palma, G. Lanzola, A. Farret, *et al.*, "2 month evening and night closed-loop glucose control in patients with type 1 diabetes under free-living conditions: a randomised crossover trial," *The lancet Diabetes & endocrinology*, vol. 3, no. 12, pp. 939–947, 2015.

[76] S. M. Anderson, D. Raghinaru, J. E. Pinsker, F. Boscari, E. Renard, B. A. Buckingham, R. Nimri, F. J. Doyle III, S. A. Brown, P. Keith-Hynes, *et al.*, "Multinational home use of closed-loop control is safe and effective," *Diabetes Care*, vol. 39, no. 7, pp. 1143–1150, 2016.

[77] R. M. Bergenstal, S. Garg, S. A. Weinzimer, B. A. Buckingham, B. W. Bode, W. V. Tamborlane, and F. R. Kaufman, "Safety of a Hybrid Closed-Loop Insulin Delivery System in Patients With Type 1 Diabetes," *JAMA*, vol. 316, pp. 1407–1408, Oct 2016.

[78] E. Dassau, H. Zisser, R. A. Harvey, M. W. Percival, B. Grosman, W. Bevier, E. Atlas, S. Miller, R. Nimri, L. Jovanovič, and F. J. Doyle III, "Clinical evaluation of a personalized artificial pancreas," *Diabetes care*, p. DC_120948, 2012.

[79] J. E. Pinsker, A. J. Laguna Sanz, J. B. Lee, M. M. Church, C. Andre, L. E. Lindsey, F. J. Doyle III, and E. Dassau, "Evaluation of an artificial pancreas with enhanced model predictive control and a glucose prediction trust index with unannounced exercise," *Diabetes technology & therapeutics*, vol. 20, no. 7, pp. 455–464, 2018.

[80] L. Magni and R. Scattolini, *Advanced and multivariable control*. Pitagoga edirice Bologna, 2014.

[81] L. Magni, D. M. Raimondo, C. Dalla Man, G. De Nicolao, B. Kovatchev, and C. Cobelli, "Model predictive control of glucose concentration in type i diabetic patients: An in silico trial," *Biomedical Signal Processing and Control*, vol. 4, no. 4, pp. 338–346, 2009.

[82] B. P. Kovatchev, S. Patek, E. Dassau, F. J. Doyle III, L. Magni, G. De Nicolao, and C. Cobelli, "Control to range for diabetes: Functionality and modular architecture," *Journal of Diabetes Science and Technology*, vol. 3, no. 5, pp. 1058–1065, 2009.

[83] M. Messori, C. Cobelli, and L. Magni, "Artificial pancreas: from in-silico to in-vivo," *IFAC 9th International Symposium on Advanced Control of Chemical Processes*, pp. 1301–1309, 2015.

[84] E. Zijlstra, T. Heise, L. Nosek, L. Heinemann, and S. Heckermann, "Continuous glucose monitoring: quality of hypoglycaemia detection," *Diabetes, Obesity and Metabolism*, vol. 15, no. 2, pp. 130–135, 2013.

[85] C. Hughes, S. Patek, M. Breton, and B. Kovatchev, "Hypoglycemia Prevention via Pump Attenuation and Red-Yellow-Green "Traffic" Lights using CGM and Insulin Pump Data," *Journal of Diabetes Science and Technology*, vol. 4, pp. 1146–1155, 2010.

[86] K. Becker, B. Thull, H. Käsmacher-Leidinger, J. Stemmer, G. Rau, G. Kalff, and H.-J. Zimmermann, "Design and validation of an intelligent patient monitoring and alarm system based on a fuzzy logic process model," *Artificial intelligence in medicine*, vol. 11, no. 1, pp. 33–53, 1997.

[87] X. Mo, Y. Wang, and X. Wu, "Hypoglycemia prediction using extreme learning machine (elm) and regularized elm," in *Control and Decision Conference (CCDC), 2013 25th Chinese*, pp. 4405–4409, IEEE, 2013.

[88] K. Turksoy, E. S. Bayrak, L. Quinn, E. Littlejohn, D. Rollins, and A. Cinar, "Hypoglycemia early alarm systems based on multivariable models," *Industrial & engineering chemistry research*, vol. 52, no. 35, pp. 12329–12336, 2013.

[89] L. Heinemann, C. Benesch, J. H. DeVries, and A. home consortium, "Ap home: a novel european approach to bring the artificial pancreas home," 2011.

[90] P. Keith-Hynes, B. Mize, A. Robert, and J. Place, "The Diabetes Assistant: A smartphone-based system for real-time control of blood glucose," *Electronics*, vol. 3, no. 4, pp. 609–623, 2014.

[91] A. Facchinetti, S. Del Favero, G. Sparacino, and C. Cobelli, "Model of glucose sensor error components: identification and assessment for new Dexcom G4 generation devices," *Med Biol Eng Comput*, vol. 53, pp. 1259–1269, Dec 2015.

[92] S. Del Favero, A. Facchinetti, G. Sparacino, and C. Cobelli, "Improving accuracy and precision of glucose sensor profiles: retrospective fitting by constrained deconvolution," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1044–1053, 2014.

[93] S. D. Favero, A. Facchinetti, G. Sparacino, and C. Cobelli, "Retrofitting real-life dexcom g5 data," *Diabetes technology & therapeutics*, vol. 19, no. 4, pp. 237–245, 2017.

[94] M. Cescon and R. Johansson, "Glycemic trend prediction using empirical model identification," *Proc. Joint 48th IEEE CDC & CCC 2009, Shanghai,China*, pp. 3501–3506, 2009.

[95] R. N. Bergman, "Minimal model: perspective from 2005," *Hormone Research in Paediatrics*, vol. 64, no. Suppl. 3, pp. 8–15, 2005.

[96] C. Cobelli, C. Dalla Man, G. Toffolo, R. Basu, A. Vella, and R. Rizza, "The oral minimal model method," *Diabetes*, vol. 63, no. 4, pp. 1203–1213, 2014.

[97] C. Toffanin, E. Aiello, C. Cobelli, and L. Magni, "Hypoglycaemia prevention via personalised glucose-insulin models identified in free-living condition," *Journal of Diabetes Science and Technology*, 2019.

[98] M. Messori, J. Kropff, S. Del Favero, J. Place, R. Visentin, R. Calore, C. Toffanin, F. Di Palma, G. Lanzola, A. Farret, *et al.*, "Individually adaptive artificial pancreas in subjects with type 1 diabetes: a one-month proof-of-concept trial in free-living conditions," *Diabetes Technology & Therapeutics*, vol. 19, no. 10, pp. 560–571, 2017.

[99] K. J. Bell, C. E. Smart, G. M. Steil, J. C. Brand-Miller, B. King, and H. A. Wolpert, "Impact of fat, protein, and glycemic index on postprandial glucose control in type 1 diabetes: implications for intensive diabetes management in the continuous glucose monitoring era," *Diabetes Care*, vol. 38, no. 6, pp. 1008–1015, 2015.

[100] American Diabetes Association *et al.*, "Postprandial blood glucose," *Diabetes care*, vol. 24, no. 4, pp. 775–778, 2001.

[101] C. Toffanin, E. Aiello, S. Del Favero, C. Cobelli, and L. Magni, "Multiple models for artificial pancreas predictions identified from free-living condition data: A proof of concept study," *Journal of Process Control*, vol. 77, pp. 29–37, 2019.

[102] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25* (F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, 2012.

[103] O. Ronneberger, P.Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 9351 of *LNCS*, pp. 234–241, 2015.

[104] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997.

[105] V. Tresp, T. Briegel, and J. Moody, "Neural network models for the blood glucose metabolism of a diabetic," *IEEE Transactions on Neural Networks*, vol. 10, pp. 1204–1213, 1999.

[106] F. Allam, Z. Nossai, H. Gomma, I. Ibrahim, and M. Abdelsalam, "A recurrent neural network approach for predicting glucose concentration in type-1 diabetic patients," in *Engineering Applications of Neural Networks*, pp. 254–259, 2011.

[107] C. Meijner and S. Persson, "Blood glucose prediction for type 1 diabetes using machine learning," 2017.

[108] J. Martinsson, A. Schliep, B. Eliasson, C. Meijner, S. Persson, and O. Mogren, "Automatic blood glucose prediction with confidence using recurrent neural networks," in *Proceedings of the 3rd International Workshop on Knowledge Discovery in Healthcare Data co-located with the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence (IJCAI-ECAI 2018), Stockholm, Schweden, July 13, 2018.*, pp. 64–68, 2018.

[109] Q. Sun, M. V. Jankovic, L. Bally, and S. G. Mougiakakou, "Predicting blood glucose with an LSTM and bi-lstm based deep neural network," *CoRR*, vol. abs/1809.03817, 2018.

[110] K. Li, J. Daniels, P. H. Viñas, C. Liu, and P. Georgiou, "Convolutional recurrent neural networks for blood glucose prediction," *CoRR*, vol. abs/1807.03043, 2018.

[111] C. Marling and R. Bunescu, "The ohiot1dm dataset for blood glucose level prediction," in *The 3rd International Workshop on Knowledge Discovery in Healthcare Data, Stockholm, Sweden*, 2018.

[112] F. A. Gers *et al.*, *Learning to forget: Continual prediction with LSTM.* 1999.

[113] D. A. Finan, C. C. Palerm, F. J. Doyle III, D. E. Seborg, H. Zisser, W. C. Bevier, and L. Jovanovič, "Effect of input excitation on the quality of empirical dynamic models for type 1 diabetes," *AIChE journal*, vol. 55, no. 5, pp. 1135–1146, 2009.

[114] E. M. Aiello, G. Lisanti, L. Magni, M. Musci, and C. Toffanin, "Therapy-driven deep glucose forecasting," *Engineering Applications of Artificial Intelligence*, vol. 87, 2020.