Modeling and control to improve blood glucose concentration for people with diabetes

Yong Mei
Iowa State University
Modeling and control to improve blood glucose concentration for people with diabetes

by

Yong Mei

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Co-majors: Chemical Engineering and Statistics

Program of Study Committee:
Derrick K. Rollins, Co-major Professor
Huaiqing Wu, Co-major Professor
Yehua Li
Kurt Hebert
Ian Schneider

Iowa State University
Ames, Iowa
2017

Copyright © Yong Mei, 2017. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Motivation</td>
<td>4</td>
</tr>
<tr>
<td>Dissertation organization</td>
<td>6</td>
</tr>
<tr>
<td>Literature Cited</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>10</td>
</tr>
<tr>
<td>Introduction</td>
<td>10</td>
</tr>
<tr>
<td>Type 1 diabetes</td>
<td>13</td>
</tr>
<tr>
<td>Risk factors</td>
<td>14</td>
</tr>
<tr>
<td>Genetics</td>
<td>14</td>
</tr>
<tr>
<td>Environmental</td>
<td>14</td>
</tr>
<tr>
<td>Treatment</td>
<td>15</td>
</tr>
<tr>
<td>Insulin pump therapy</td>
<td>15</td>
</tr>
<tr>
<td>Pancreas transplant</td>
<td>15</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>15</td>
</tr>
<tr>
<td>Risk factors</td>
<td>16</td>
</tr>
<tr>
<td>Lifestyle</td>
<td>16</td>
</tr>
<tr>
<td>Genetics</td>
<td>16</td>
</tr>
<tr>
<td>Treatment</td>
<td>16</td>
</tr>
<tr>
<td>Change of lifestyle</td>
<td>16</td>
</tr>
<tr>
<td>Medication</td>
<td>17</td>
</tr>
<tr>
<td>Artificial pancreas (AP)</td>
<td>17</td>
</tr>
<tr>
<td>PID Controller</td>
<td>18</td>
</tr>
<tr>
<td>MPC Controller</td>
<td>19</td>
</tr>
<tr>
<td>Fuzzy Logic/MD-Logic Controller</td>
<td>21</td>
</tr>
<tr>
<td>Feedforward Controller</td>
<td>21</td>
</tr>
<tr>
<td>In silico studies</td>
<td>23</td>
</tr>
<tr>
<td>Literature Cited</td>
<td>24</td>
</tr>
<tr>
<td>CHAPTER 3: THE DEVELOPMENT OF A VIRTUAL SENSOR IN GLUCOSE MONITORING</td>
<td>29</td>
</tr>
<tr>
<td>FOR NON-INSULIN DEPENDENT PEOPLE</td>
<td></td>
</tr>
</tbody>
</table>
Concluding remarks.................................................................................................................................. 173
Literature Cited......................................................................................................................................... 175

CHAPTER 8: A FEEDFORWARD CONTROL APPROACH WHEN THE TIME DELAY OF
THE MANIPULATED VARIABLE IS GREATER THAN THE INPUT VARIABLE ............... 178

Abstract..................................................................................................................................................... 178
Introduction............................................................................................................................................... 179
Methodology ............................................................................................................................................. 180
Simulated CSTR ................................................................................................................................... 180
Modeling Methodology .......................................................................................................................... 181
FFPC Law ............................................................................................................................................. 185

Results ....................................................................................................................................................... 186
Fit of the WM........................................................................................................................................ 186
Closed-loop Control.............................................................................................................................. 186

Concluding Remarks ................................................................................................................................. 189
Acknowledgements ................................................................................................................................... 190
Literature Cited ......................................................................................................................................... 190

CHAPTER 9: DEMONSTRATION OF THE EFFECTIVENESS OF FEEDBACK PREDICTIVE
MODELING AND CONTROL METHODOLOGY IN PROCESSES WITH SIGNIFICANT TIME
DELAY AND UNMEASURED DISTURBANCES................................................................................ 193

Introduction ............................................................................................................................................... 194
Methodologies ........................................................................................................................................... 196
Simulated CSTR ................................................................................................................................... 196
Modeling Methodology .......................................................................................................................... 197
Mathematical Model ............................................................................................................................. 199
Control Algorithms ................................................................................................................................... 201

Results ....................................................................................................................................................... 202
Fit of the WM........................................................................................................................................ 202
Closed-loop Control.............................................................................................................................. 204

Concluding Remarks ................................................................................................................................. 206
Literature Cited ......................................................................................................................................... 207

CHAPTER 10: AN IN-SILICO STUDY OF FEEDFORWARD PREDICTIVE CONTROL IN
BLOOD GLUCOSE CONCENTRATION FOR PEOPLE WITH TYPE 1 DIABETES .................. 209

Abstract..................................................................................................................................................... 209
Introduction............................................................................................................................................... 209
Modeling Methodology .......................................................................................................................... 212
Control Algorithms ................................................................................................................................... 215
FFPC ..................................................................................................................................................... 215
MPC ...................................................................................................................................................... 216
MB ........................................................................................................................................................ 217
Training Performance Statistics ................................................................................................................ 218
Protocols of the In-Silico trials .................................................................................................................. 219
Results and Discussion ............................................................................................................................. 220
Concluding Remarks ................................................................................................................................. 223
Literature Cited ......................................................................................................................................... 224

CHAPTER 11: A NOVEL FEEDBACK PREDICTIVE CONTROL ALGORITHM APPLIED IN BLOOD GLUCOSE CONCENTRATION CONTROL FOR PEOPLE WITH TYPE 1 DIABETES.....227
Abstract ..................................................................................................................................................... 227
Introduction ............................................................................................................................................... 228
Methodology ............................................................................................................................................. 230
  Control Algorithms ............................................................................................................................... 230
  Modeling Methodology ........................................................................................................................ 231
  Protocols of In-silico Study .................................................................................................................. 234
  Evaluation Metrics ................................................................................................................................ 236
  Training Performance Statistics ............................................................................................................ 236
Results and Discussion ............................................................................................................................. 237
Concluding Remarks ................................................................................................................................. 241
Literature Cited ......................................................................................................................................... 242

CHAPTER 12: FUTURE WORKS .......................................................................................................... 244
ACKNOWLEDGMENTS

I would like to thank my major professors, Dr. Rollins and Dr. Wu, and my committee members, Dr. Li, Dr. Hebert, and Dr. Schneider for their guidance and support throughout the course of this research.

In addition, I would also like to thank my friends, colleagues, the department faculty and staff for making my time at Iowa State University a wonderful experience.

Most of all, I want to also offer my thanks to my wife, my parents and other family members who supported me during the hardship I have gone thorough in the past year.
ABSTRACT

Diabetes mellitus is a chronical condition that features either the lack of insulin or increased insulin resistance. It is a disorder in the human metabolic system. To combat insufficiency of insulin released by pancreas, a closed-loop control system, also known as artificial pancreas (AP) in this application, have been created to mimic the functionality of a human pancreas. An AP is used to regulate blood glucose concentration (BGC) by managing the release of insulin. Therefore, an algorithm, which can administer insulin to reduce the variation of BGC and minimize the occurrences of hyper-/ hypoglycemia episodes, is the key component of an AP. The objective of the dissertation is to develop an optimal algorithm to better control BGC for people with diabetes.

For people with Type 2 diabetes, prevention or treatment of diabetes mellitus can typically be done via a change of lifestyle and weight management. A virtual sensing system that does not require many manual inputs from patients can ease the burden for people with Type 2 diabetes. This dissertation covers the development of a monitoring system for Type 2 diabetes.

To achieve the goal of tighter control of BGC for people with Type 1 diabetes, dynamic modeling methodology for capturing the cause-and-effect relationship between manipulated variable (i.e. insulin) and controlled variable (i.e. BGC) has been developed. Theoretically, this dissertation has established that physiologically based nonlinear parameterized wiener models being superior to nonlinear autoregressive moving average with exogenous inputs (NARMAX) models in capturing dynamic relationships in processes with correlated inputs. Based on these results, wiener models have been applied in the modeling of BGC for real subjects with Type 1 diabetes under free-living conditions. With promising results shown in wiener models, an
extended physiologically based model (i.e. semi-coupled model) has been developed from wiener structure, which enables the development of a phenomenologically sound feedforward control law. The feedforward control law based on wiener models has been tested in simulated continuous-stirred-tank reactor (CSTR) that demonstrates tight control of controlled variables. Further simulation runs with a CSTR also shows feedforward predictive control (FFPC) can provide tighter control over model predictive control (MPC). Lastly, for the special application of BGC control for people with Type 1 diabetes, FFPC demonstrates tighter control than MPC under simulation environment. To account for unmeasured disturbances and inaccurate models for manipulated variable in real life scenarios, feedback predictive control (FBPC) is developed and proven to be a more effective control algorithm under both CSTR and diabetes simulation environment, which can establish the foundation for tightening BGC in real subject clinical studies.

**Key Words:** Artificial pancreas, feedforward control, model predictive control, dynamic modeling, virtual sensors, diabetes mellitus
CHAPTER 1: INTRODUCTION

Diabetes, a dysfunction in human metabolic system, has become one of the most prevalent health problems around the world. According to the report by Centers for Disease Control and Prevention (CDC) in 2014, there is a record that 9.3% of the people among population in United States are plagued by diabetes, while 9% of the world population are affected by various types of diabetes [1]. The risk of death is 50% higher for people with diabetes than people without diabetes. In addition, more than 80% of deaths related to diabetes occur in low- or middle-income countries with insufficient healthcare provided [2]. In China, up until 2015, 11.6% of population in has diabetes or pre-diabetes [3,4].

Diabetes, as a disorder in human metabolic/regulatory system, features deficiency in insulin, and is commonly divided into two major categories: Type 1 Diabetes (T1D) features the reduced or total loss of the ability to produce insulin from beta cells in pancreas, and it is usually caused by beta cells targeted and eliminated by immune system; Type 2 Diabetes (T2D) usually features cell resistance of insulin or reduced efficiency of insulin. Both major types of diabetes could result in dysfunctional behavior in insulin/blood glucose regulatory. For T1D, their blood insulin level is often less than that of a healthy person, while for T2D, the blood insulin level can be less than, equal to or more than that of a healthy person. But both types feature inability to lower blood glucose concentration (BGC) properly due to either the lack of insulin (T1D) or ineffectiveness of insulin (T2D). Consequently, for people with either T1D or T2D, their BGCs will often be higher than nominal level. The focus of this dissertation is on T1D or sometimes called insulin dependent diabetes (IDD) since the dynamic relationship between insulin and glucose is well studied and there are advanced technologies that can administer insulin (i.e. insulin pumps), while the mechanism behind T2D is more complex and the treatments for T2D
varies greatly from person to person. This dissertation will also cover monitoring systems for T2D based on input-only models.

Hyperglycemia (i.e. high BGCs) poses the major concern for people with diabetes. If insulin concentration in the human body is too low, the body will fail to provide sufficient energy via blood glucose and blood glucose could remain under-utilized. As a result, high level of blood glucose occurs. High concentration of blood glucose accumulating in blood stream will cause irreversible damage to body organs as glucose crystals can damage blood capillaries. Severe complications can follow diabetes, and the risk of death for population with diabetes is higher [1,2].

To combat deficiency of blood insulin in people with diabetes insulin injection/infusion are traditional treatments. Different insulin injection/infusion methods have been developed. Traditionally, 3 to 4 injections before meals per day is usually used to prevent BGC from rising due to meals. With the advancement of modern technology, equipment such as insulin pumps (devices that can release insulin, via a catheter inserted under the skin, at pre-programmed time with precise dosage), continuous glucose monitors (devices outputting BGC readings at a frequency up to 1 reading per min) can improve the glucose management in diabetes patients. However, the fundamental issue with glucose management for T1D, the dosage and timing of insulin release, remains a problem.

One aspect of that problem is the timing of insulin administration. In self-management of glucose, when a person with diabetes eats, it usually takes at least 15-30 mins before a rise in blood glucose occurs [5]. If that person injects rapid acting insulin after the glucose rise has been detected by a monitor, it will take at least 15 mins before insulin reaches the blood stream, and another 60mins before insulin action (i.e. lowering BGC) reaches its peak effect [6]. Summing
up all the time delays described above, it will take a long time (e.g. over 1 hour) before injected insulin can bring down the blood glucose after that meal for a person with diabetes. Thus, there will be extended period of hyperglycemia for T1D patients.

The other aspect of blood glucose management problem is the dosage used in insulin administration. When the insulin level in the blood stream is too low, the BGCs will be high and hyperglycemia ensues. But if the insulin level is too high, the BGCs will be too low (i.e. hypoglycemia) so that the brain will have to gradually shut off first peripheral organs then its own functions as BGC goes down and eventually people will die from low BGC [7].

To summarize the problem of blood glucose management, with the time delays associated with insulin administration and meals, the goal is to develop an algorithm to compute insulin dosage needed and time of the insulin injection/infusion to prevent large deviation of BGC from desired level. This can help T1D patients to achieve tight control over BGCs and improve their quality of life. For people with T2D, the key challenge is to develop an input-only monitoring system that can provide BGC information to help people improve BGC control. Ababstracted from the context of human metabolic/regulatory system, tighter blood glucose control/blood glucose management problems essentially could be viewed as a typical process control problem as seen in industry. But the complication is that the human metabolic system is far more complex than many industrial mass balance systems, as a lot of disturbances are unmeasurable or unmeasurable online (e.g. blood insulin), and each person has a unique set of dynamic glucose/insulin system that differs dynamically from one another.

There are other treatments to T1D or T2D. Most notably, various diabetes medicine can provide effective treatments for people with diabetes in certain cases according to the prescription of physicians. For example, metformin is an oral medicine can help improve BGC
control with T2D, exenatide is a glucagon-like peptide-1-agonist medication that works with T2D, and insulin glargine is a version of man-made long acting insulin substitute that can be used to treat both T1D and T2D. These treatments are not within the coverage of this dissertation since the goal of this article is to achieve tighter BGC control from the aspect of controlling devices not medication treatments.

**Motivation**

In order to help people with T1D to improve their blood glucose management, the concept of the artificial pancreas (AP) was introduced. It usually consists of an insulin pump, a control algorithm, and a continuous glucose monitor (CGM). The artificial pancreas shall be able to utilize frequent glucose measurements (up to 1 reading per min) from CGM to calculate the amount of insulin needed based on the control algorithm. Since the AP aims at glucose management without human intervention, it is also being called closed-loop control in process control terminology.

The core part of the AP is the control approach adopted, and the performance of AP depends on the control algorithm. There are two basic categories of control algorithms: feedback control (FBC) and feedforward control (FFC). FBC is a type of “reactive” control. There has to be a deviation from the set point before the control action takes place [8]. In the context of diabetes, the BGC has to deviate from its desired value before the algorithm determines the amount of insulin needed. The disadvantage of FBC is obvious- it is difficult to achieve tight glucose control and large swings of BGC after meal is inevitable. The advantage of FBC is that it’s usually easy and simple to implement the algorithm and does not require too much tuning for parameters.

The most popular type of control approach currently in this application is model predictive control (MPC). MPC is a sub-genre of FBC. It is still a “reactive” type of control
concept; however, the deviation required for control action is not based on what has happened but what is going to happen. Hence, MPC can be thought of as futuristic feedback control. The disadvantage of MPC usually lies in the models used: autoregressive moving average (ARMA) or autoregressive moving average with exogenous inputs (ARMAX) that is usually used for control algorithm [9-11]. The model structure benefits from its identifiability and easy implementation. However, the prediction horizon is limited by the models since they cannot properly represent the true mechanism behind the dynamic behaviors of blood glucose and insulin. When past the limited prediction horizon, the prediction will become neither accurate nor precise. This issue can often be alleviated by adaptive control [8]. Since diabetes patients require a certain amount of time in advance to determine the amount of insulin infusion, MPC with output dependent models (e.g. ARMAX) can be less reliable when time delays are high and the correlation structure changed.

Feedforward control (FFC) is a control approach that features “proactive” control actions. In principle, FFC that seeks to nullify the measured input effects before there is a deviation for set point or the target for the controlled variable. In the context of diabetes, FFC can determine the amount of insulin required to cancel the inputs measured such as meal contents, physical activities, and emotional stress. This is the control approach of this research. The FFC model does not depend on output measurements such as BGCs [12]. With an effective input only model implemented into the algorithm, there is a potential to greatly reduce the variation of blood glucose in T1D patients and to achieve tighter blood glucose control.

Moreover, to achieve tight glucose control, the FFC system must capture a significant amount of cause-effect relationships associated with major disturbances, must accurately achieve subject-specific modeling, and must represent a critical degree of physiological soundness, in
order to manage the timing and size of insulin administration. Feedforward predictive control (FFPC), a new type of control idea, in this dissertation has the potential of achieving this goal.

However, in real life scenario, or free-living condition where there is no restriction on the life style of patients, it is often difficult to come up with an accurate, yet simple input only model to be used in FFPC. It could be due to the inaccuracy of disturbance sensors that quantitatively measure disturbances related to physical activities, or emotional stresses. To compensate for the lack of accurate input only models for FFPC, a new type of feedback predictive control (FBPC) method is developed. It combines feedforward features from FFPC and feedback tuning features from FBC. All it requires is an accurate model that does not drift over a long period of time. This requirements on models can be satisfied via the combination of input only models and output correction terms. As a result, FBPC could see great potential in the application of BGC management in both simulation studies and clinical trials.

To evaluate different control schemes for comparison, the most effective way is to run a simulation study. With a simulation study, the robustness of the control strategy against interindividual variability can also be easily evaluated in a large-scale simulation. Also, for the safety of test subjects, simulation tests on virtual patients are required before proceeding to clinical trials for people with T1D. In this dissertation, simulation results will be given to demonstrate the superiority of FFPC and FBPC methods. Additionally, a framework of BGC monitoring system will be shown for people with T2D.

**Dissertation organization**

This dissertation covers modeling and control to improve blood glucose concentration for people with diabetes, and will be divided into the following sections:
Chapter 2 will cover the literature review on the origin, mechanism, and typical treatments on different diabetes mellitus, as well as the current progress of artificial pancreas (AP) projects, especially with the focus on results from other simulation studies.

Chapter 3 will demonstrate the feasibility of low order semi-empirical wiener structure in accurately monitoring BGC in a continuous fashion with 22 subjects with type 2 diabetes over the length of 4 weeks under free-living conditions. (Modeling and Co-authoring)

Chapter 4 will showcase the superiority of nonlinear parameterized wiener structure in dealing with correlated inputs over nonlinear autoregressive moving average with exogenous inputs (NARMAX) models in a distillation column setting. (Co-authoring)

Chapter 5 will explore the application of wiener structure in the modeling of BGC under outpatient free-living settings for people with type 1 diabetes. (Data collection and Co-authoring)

Chapter 6 will propose a novel semi-coupled structure to better represent the physiological insulin BGC metabolic structure within human bodies, and demonstrate the advantages of semi-coupled structure over wiener structure. (Data collection and Co-authoring)

Chapter 7 will introduce a powerful feedforward control law that potentially could greatly reduce the variation of BGC around its desired levels. Its effectiveness has been tested in a simulated continuous-stirred-tank reactor (CSTR). (Co-authoring)

Chapter 8 demonstrates significant reduction in standard deviation around mean for controlled variable with the implementation of feedforward predictive control (FFPC) law with the presence of significant time delays associated with manipulated variables (i.e. time delays greater than other measured disturbances) and unmeasured disturbances. (Co-authoring)
Chapter 9 demonstrates the feedback predictive control (FBPC) algorithm applied in simulated CSTR can effectively tighten control of controlled variables compared against MPC.

(Co-authoring)

Chapter 10 indicates the applications of feedforward predictive control in BGC control for people with type 1 diabetes, could provide as tight as, or tighter control of BGC than the typical model predictive control methods. (Modeling, simulated control runs and Co-authoring)

Chapter 11 demonstrates the proposed feedback predictive control method can provide significantly tighter control than feedforward predictive control and model predictive control methods. (Modeling, simulated control runs and Co-authoring)

Chapter 12 discusses future works that could be done on the modeling and control to improve BGC for people with diabetes.

Literature Cited


CHAPTER 2: LITERATURE REVIEW

Introduction

Diabetes mellitus is a chronic metabolic disorder that features the lack of insulin in regulating blood glucose concentration (BGC) within human bodies. Therefore, to understand diabetes mellitus, one must first understand the metabolic systems that have been disturbed by diabetes.

Blood glucose is a form of sugar that exists in bloodstream and can be carried into cells within human bodies to provide energy [1,2]. A major source for blood glucose is carbohydrate from meal ingestion. Through meal ingestion, carbohydrate after digestion will transform into glucose, galactose or fructose, all of which will be absorbed into the portal vein in gastrointestinal tract. From there, glucose enters glucose metabolic system and becomes blood glucose [2,3]. Glucose metabolic system is composed of two major pathways: in one pathway, blood glucose works as energy fuels. While in the other pathway, it works as the storage unit. When providing energy, blood glucose can be used by most muscle and adipose tissues as energy fuel through biochemical reactions, and the uptake of blood glucose by those tissues is facilitated by a hormone named insulin. When working as a storage unit, blood glucose could be stored in liver in the form of glycogen which is an alternate version of glucose that can help the liver to regulate BGC when there is too much glucose circulating in bloodstream. The uptake and storage of blood glucose in the liver is also facilitated by insulin. Fig. 1 illustrates this blood glucose pathways.

In addition to the major glucose pathway introduced above, there are some minor yet still important pathways for blood glucose. For example, blood glucose could emerge from other sources as well: it could be formed from glycogen stored in liver through glycogenesis, which is
a process facilitated by hormones such as epinephrine and glucagon. Another source of blood glucose is through gluconeogenesis from various precursors of glucose. Those glucose precursors can be generated via various glucose metabolic pathways: lactic acid oxidized from glucose can be reform into glucose in the liver, and glycerol from adipose tissues can be delivered to the liver to revert into glucose via gluconeogenesis as well. Moreover, besides providing energy to muscle and adipose tissues, blood glucose can also fuel the nervous system, specifically, the brain. However, the mechanism for fueling brain is quite different from others. As the brain makes use of blood glucose via diffusion of glucose from high concentration to low concentration and insulin does not facilitate this process. As a result, the level of blood glucose directly impacts the functionalities of the brain. If BGC drops below a certain threshold, the brain would be unable to utilize blood glucose as fuels. Therefore, brain functions could be reduced or shut off, and this situation could be life-threatening.

Figure 1. Blood glucose major metabolic pathway

Several key metabolic pathways in human bodies are made possible by a hormone called insulin. Insulin is secreted within human bodies. It was first discovered in 1922 [4]. Then it was revealed that insulin is produced by beta cells in pancreas and released into metabolic system.
Insulin plays a very important role in glucose metabolic pathways. High BGC stimulates the secretion of insulin. In turn, insulin facilitates the uptake and storage of blood glucose in liver and kidney to remove blood glucose from blood stream. As a result, euglycemia (i.e. normal concentration of blood glucose) can be achieved within human bodies [5]. The following Fig. 2 illustrates major insulin metabolic pathways. Insulin delivery represents exogeneous insulin intake for people requiring insulin to be infused/injected due to insufficient insulin produced by pancreas. For a healthy person, insulin is produced by beta cells in the pancreas, and decomposed primarily in liver and kidney via insulin degradation enzyme. When plasma insulin level is high, insulin will inhibit the secretion of insulin from pancreatic beta cells, when the insulin level is low, the inhibition effect will be lifted.

**Figure 2.** Plasma major insulin metabolic pathways

Other hormones involved in glucose and insulin metabolic systems such as glucagon and epinephrine usually work in conjugation with insulin in the regulation of BGC. Glucagon, secreted by alpha cells in pancreas, can regulate the process of converting glycogen into glucose, and thus increase BGC in blood stream. Epinephrine, also known as adrenalin, is another type of hormone that is released by adrenal glands and has the effect of increasing BGC among other
effects. Those hormones work together with insulin to regulate the level of blood glucose and maintain it within euglycemia.

Diabetes mellitus can be usually diagnosed through various tests. Fasting glucose test involves taking measurements of BGC for people under fasting conditions. When fasting glucose is greater than 126 mg/dl, that person can be at the risk of diabetes mellitus. A more formal method is the measurement of hemoglobin A1c (HA1c) [6], which indicates the average BGC for the past several months. The criteria for diabetes usually can be treated as HA1c values being greater than 6.5%.

With the knowledge of physiological background for the metabolic pathways for blood glucose and insulin, greater insight can be gained towards diabetes mellitus. The rest of this chapter will be divided into following sections: different types of diabetes mellitus, their risk factors, and general treatments. Then, the latest development of closed-loop control system for Type 1 diabetes (i.e. artificial pancreas) will be covered, especially the development of in silico studies.

**Type 1 diabetes**

Though symptoms may be similar, there are different categories of diabetes based on the mechanism behind them. Two major categories of diabetes mellitus plague the vast majority of those with diabetes.

Type 1 diabetes (T1D), having its root in the inability of beta cells in the pancreas to generate insulin, affects about 5% to 10% of all people with diabetes. T1D is also known as juvenile diabetes since its onset is usually diagnosed in adolescence. Its typical symptom starts with increased thirst, frequent urination, fatigue, weight loss, acetone smell in breath, etc. In addition, diabetic ketoacidosis (DKA) can often be diagnosed along with the onset of T1D. DKA
is a life-threatening complication with diabetes mellitus. Severe DKA can cause confusion and sometimes coma [7]. Besides the typical symptoms and complications, T1D causes other long-term effect such as permanent damage to organs, loss of eye sight, limb loss, etc. [8]

**Risk factors**

There are various studies on the risk factors of T1D. According to those researches, several factors contributed to the onset of T1D.

**Genetics**

Genetic factors play an important role in the development of T1D. T1D involves more than 50 genes. The heritability of T1D is estimated at around 80% to 86% [9]. Although genes involved in T1D can be dominant, recessive or status fall somewhere in between, there is still great risks of T1D from the aspect of genetics. For example, if a father has T1D, the child has 5% chance having it. And if one identical twin has T1D, the chance of the other one has it is around 50% [10].

**Environmental**

Evidence shows environmental factors influence the onset of T1D. The same example of identical twins can demonstrate that even with the same genes, the second twin does not have 100% chance of T1D [11]. There are other specific environmental factors that could cause T1D. Certain virus in human bodies can trigger an autoimmune reaction, where the immune system will destroy virus-infected cells along with beta cells that produce insulin. Plus, being in contact with some chemicals or drugs can induce the onset and development of T1D. Research also shows gluten could be a factor in the development of T1D. However, the mechanism is unclear yet [12].
Treatment

Insulin pump therapy

Currently the most popular and feasible treatment for people with T1D is insulin pump therapy. The traditional treatment features insulin bolus injection, which requires patients to inject bolus insulin based on their diet schedule and needs to bring down BGC. With the rapid development of continuous subcutaneous insulin infusion (CSII), insulin pumps have been widely adopted in insulin administering. In insulin pump therapy, patients will have long-acting basal insulin infused into body to balance BGC around their fasting glucose level, and rapid-acting insulin, also known as bolus insulin, will be used to counter act the rise in BGC due to meal ingestion. The amount of basal insulin is calculated based on personal attributes of the patients. And the amount of bolus insulin, in insulin pump therapy, is calculated by the carbohydrate/insulin ratio, whose value indicates the amount of carbohydrate countered by one unit of insulin.

Pancreas transplant

A pancreas transplant is usually performed along with a renal transplant. The recipients of pancreas will typically have their newly transplanted organ attached to a different location while the original organ untouched since removing the original organ could increase mortality rate. Around 90% of people with transplanted pancreas are insulin-independent at the one-year mark, while at the five-year mark, the percentage drops to as low as 65% [10,13].

Type 2 diabetes

Type 2 diabetes (T2D) is the most prevalent form of diabetes that affects around 90% of people with diabetes [14]. People with T2D usually can still generate insulin from their beta cells in pancreas. However, insulin resistance in organs prevent the body from utilizing insulin
effectively. The symptoms of T2D also include thirst, frequent urination, and weight loss. The long-term effects include blindness, kidney failure and limb loss, etc. People with T2D can have 10 years shorter life expectancy [15]. However, unlike T1D, T2D is preventable and treatable in some cases by proper weight management and regular exercise.

**Risk factors**

**Lifestyle**

T2D often comes along with obesity and lack of exercise. As a result, lifestyle becomes an important factor in the onset and development of T2D. Overweight, defined by body mass index (BMI) greater than 25, is a major factor associated with T2D. For instance, obesity accounts for around 30% of cases of T2D in people with Chinese ancestry [16]. The lack of regular exercise can account for 7% of T2D population. In addition, diet composition (in terms of glycemic index) can also impact the risks for T2D.

**Genetics**

Over 30 genes are involved in the onset and development of T2D. Overall, heritability of T2D is less than that of T1D. Genetics factors only account for 10% of T2D. However, if one identical twin has T2D, the risk increased for the other twin is 90% [11].

**Treatment**

**Change of lifestyle**

The most simple and readily available treatment for most people with T2D is to change the lifestyle including losing weight, regular exercise, proper diet management, etc. Sufficient exercise and healthy eating habits can help to improve insulin resistance [17]. Exercise, especially aerobic exercise can reduce HbA1c and lead to tighter BGC control. A diet with a low
glycemic index or low carbohydrate can also improve BGC control. Typically, a healthy lifestyle can help people with T2D to improve their BGC control within weeks.

**Medication**

Medication is also an effective treatment for people with T2D. Oral medicine such as metformin can help people with T2D to better control BGC. Injection of insulin is usually not required during the early phases of T2D. However, if lifestyle change and medication do not work as intended, long-acting insulin or insulin substitute may become a choice to better control BGC for some cases.

**Artificial pancreas (AP)**

Recently, development has been made to show the AP/automatic closed-loop control can improve the percentage of time that BGC stays within desired glycemic range (70-180 mg/dl) [18]. Various strategies were applied in glucose management in order to maintain euglycemia (i.e. normal level of BGC).

While clinical trials are an essential step toward final products for patients, there are a lot of groups working on simulation studies. With the highly dynamic, highly nonlinear nature of human metabolic system, and the between subject variation of the system, computer simulation with complex compartment models, such as University of Padova/University of Virginia models [19] and a simulator based on Cambridge model [20], can provide insight into the dynamic relationship between glucose and insulin, but they cannot provide much useful information towards real-life scenarios or free living conditions (i.e. people live their life without any constraints) since there are a lot of unknown disturbances to account for. However, simulation studies can provide researchers with a safe environment to test different BGC control algorithms without putting any patients in potential risks for experiencing hypo- or hyperglycemic episodes.
In addition, simulation studies can facilitate the comparison of different control schemes on the same group of patients, and since simulation studies are easy to expand in scale, it becomes an effective way to evaluate the robustness of tested control algorithms against the interindividual variability before *in vivo* clinical trials. As a result, this review will focus on human subject studies as well as simulation studies.

Currently, clinical trials from different research groups range from inpatient to outpatient studies. These clinical studies can last from only night hours to a few full days or even weeks. As a practice, all their results are reported in terms of percentage of time within certain glucose range. Most of these studies are called hybrid closed-loop studies since they require meal announcements from the patients and bolus insulin is usually manually administered. The studies that are completely devoid of human intervention are called fully closed-loop. Some of studies have physical activities such as trend mill test, and some do not (or as free-living studies that do not restrict life style of patients). In contrast, simulation studies mostly feature control runs that last from a few hours around meals to a few days with meal announcements, and without any involvement of other disturbances such as physical activities.

By control strategy, there are mainly four categories in AP:

**PID Controller**

PID controller is a traditional feedback controller that has been used in many applications across industry. Its general controller is shown in Eq.1 below.

\[
    u(t) = K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{de(t)}{dt} \tag{1}
\]

where \( e(t) \) represents the deviation of measured blood glucose from its set point (in unit of mg/dl), and \( u(t) \) represents the signal for insulin action required to bring down the elevated blood
glucose. Eq. 1 is for fully feedback control. It only reacts to deviation from the set point of blood glucose level. The PID controller is often compared with Sensor-Augmented Pump (SAP) therapy (SAP mainly consists of CGM and Insulin pump). The control scheme in SAP is a simple insulin suspension system. That is, the insulin pump will shut off as soon as it detects patient’s blood glucose level below a certain threshold.

In clinical trials, research on PID control for diabetes patients shows this type of control algorithm does not appear to improve glucose control over SAP therapy significantly [21].

**MPC Controller**

The MPC controller is the most popular controller in AP community, since MPC can potentially adjust for future blood glucose changes to minimize the delay in insulin action. According to the model types used in MPC, there are two major subgroups. The most popular one is the physiological based compartment model. Usually can be viewed as an extension of Bergman’s Minimal model [22] using insulin and glucose dynamics only as demonstrated in Eqs. 2 and 3,

\[
\frac{dG}{dt} = -[S_G + X(t)]G \tag{2}
\]

\[
\frac{dX}{dt} = p_2 I(t) - p_3 X(t) \tag{3}
\]

where \( G \) represents plasma glucose concentration (mg/dl), \( I \) represents plasma insulin (unit/dl), \( X(t) \) denotes remote insulin actions (i.e. the insulin effect is organs that are not in quick equilibrium with blood stream) in unit of (1/min), \( S_G \) denotes glucose effectiveness (1/min), \( p_2 \) represents fractional appearance rate of insulin in interstitial fluids, and \( p_3 \) represents fractional clearance rate of insulin in interstitial fluids. Although Bergman’s model only considered the dynamics of glucose after an injection of glucose into blood, Equations 2 and 3 forms the basis
of modern physiological based compartment models. Those compartment models predict BGC excursion based on carbohydrate contents in food (that is reason why hybrid closed-loop experiments require meal announcements), and rapid-acting insulin [23-26]. While these control schemes can demonstrate their improvement over SAP therapy, night time during sleep without many unmeasured disturbances is easier to achieve tight control [27].

There are also research projects that make use of glucagon in maintaining BGC control (dual-hormone models) [28, 29]. However, on one hand, the usage of glucagon will increase the burden on patients, and on the other hand, glucagon infusion/glucagon online measurement device is not commercially available right now.

For insulin and glucose models discussed above, there are exercise components included in some models but most models do not have exercise as inputs. Also, most of them have not applied fully subject-specific modelling methods, and only certain individual information is collected (e.g. subject body weight, BMI, daily insulin dosage).

Another subgroup within the MPC controller includes empirical based models. A typical empirical models used is autoregressive moving average with exogenous inputs (ARMAX) model where output BGC data is used in predicting BGC [30-32]. These algorithms use prediction models that depend on output (i.e. measured BGC). In this case, the contribution to insulin infusion calculation could be mostly due to output measurements, which indicates the cause and effect relationship may not be well captured in the models. The results tend to show these MPC algorithms had improved control during night time and did not achieve significant better control during daytime due to disturbances from daily activities.
Exercise data in autoregressive moving average (ARMA) or ARMAX models is often used for constraints. Linear or low order models are used in modeling, where those linear or low order terms cannot provide proper physiological structure for BGC dynamics.

Since models used in MPC fail to or can only partially capture dynamic behavior of BGC and blood insulin, in order to adapt to BGC/blood insulin dynamics, MPC controllers often use receding horizon strategy in which parameters are updated whenever new measurements are available.

**Fuzzy Logic/MD-Logic Controller**

This type controller is to mimic the thinking process of the physicians. It does not use dynamic BGC structure to calculate insulin infusion. This type of controller had some success in improving BGC control during night time [33,34].

As mentioned before, all the results in AP research being reported in terms of percentage of time within certain range (e.g. 70-180 mg/dl) of BGC makes it difficult to tell if the insulin infusion rate calculation is based more on dynamic BGC structure or more on previous BGC measurements. If the previous BGC measurements are contributing more to insulin infusion rate, then the change of correlation structure under life pattern change could damage the predictability of the controller.

**Feedforward Controller**

Feedforward control is a relatively less touched category of control approach in diabetes literature. Marchetti et al. proposed a Feedforward-feedback control algorithm [35,36]. It used meal based glucose contents as measured disturbance. In terms of transfer function, first order transfer function was used for the approximation of the process from Laplace transformations of the disturbance, D(s) to BGC denoted by Y(s) as shown in Eq. 4. A second order transfer
function is to approximate the process from Laplace transformations of insulin infusion rate \( U(s) \) to \( Y(s) \) as in Eq.5.

\[
\frac{Y(s)}{D(s)} = G_d(s) = \frac{K_d}{\tau_d s + 1} \tag{4}
\]

\[
\frac{Y(s)}{U(s)} = G_p(s) = \frac{Ke^{-\theta}}{\tau_1 s + 1)(\tau_2 s + 1)} \tag{5}
\]

\[
U(s) = G_f(s)D(s) \tag{6}
\]

\[
G_f(s) = -\frac{G_d(s)}{G_p(s)} = -\frac{K_d}{K} \frac{(\tau_1 s + 1)(\tau_2 s + 1)e^{-\theta}}{\tau_d s + 1} \tag{7}
\]

\[
G_f(s) = -\frac{G_d(s)}{G_p(s)} = -K_f \frac{\tau_3 s + 1}{\tau_4 s + 1} \tag{8}
\]

The feedforward control law is dictated by Eqs. 6 and 7. However, the problem with this controller is that it is unrealizable (i.e. the controller will depend on future values of disturbances). A typical way to solve this issue is to approximate the exponential term with a lead-lag term [49]. Hence, Eq. 7 becomes Eq. 8, where \( K_f = \frac{K_d}{K} \), and \( \tau_3 = \tau_1 + \tau_2 + \theta \), and \( \tau_4 = \tau_d \).

The tuning approach is to minimize the discrete cost function, as in Eq. 9.

\[
J = \sum_k [G_{sp} - \hat{G}] \Delta t \tag{9}
\]

where \( G_{sp} \) is the BGC set point and \( \hat{G} \) is an estimator of BGC.
This research was conducted *in silico*. The patient model used was Cambridge model [20]. The simulation study shows that Feedforward control can successfully reduce post-meal glucose by set point reduction.

Alternatively, Abu-Rmileh and Garcia-Gabin evaluated Feedforward control with scheduled gain approach in simulation [37]. State space models were used in the control scheme. This study showed decent controller performance in a fasting condition with meal challenges.

*In silico* studies for feedforward controller can provide initial evaluations on the performance. But these studies failed to recognize that free living environment cannot be realized by current simulation methods, and also traditional approximation in unrealizable controller could negatively affect the precision of the controller. This type of approximation does not prevent large swings of BGC from happening.

*In silico* studies

The history of AP can date back to 1977 when the first commercialized AP Biostator was used for inpatients [38]. However, its development was not a smooth sail. Even up until the end of 2016, the first commercialized AP has just been approved by FDA for people over 14 years of age with insulin-dependent diabetes [39]. It can automatically correct basal insulin infusion rate, but people still need to manually determine the values of the bolus insulin. The most important consideration in clinical trials is the safety of patients, which adds additional constraints to clinical trials and the development of the AP. With the consideration of patient safety, extreme conditions cannot be tested in clinical trials. Also with the cost of large scale inpatient studies, the pace of AP development has been slow. Therefore, simulation studies become an effective substitute to preclinical trials, animal trials and clinical trials to a certain degree. Although the simulation environment is not as realistic as environment in clinical trials, *in silico* studies
expedite the development of AP projects since it is easy for simulators to work under extreme conditions that cannot be done for in vivo studies, as well as to scale up studies to extract information on interindividual variability in test sample.

The objective of diabetes simulation studies is to demonstrate the proposed control algorithm can maintain BGC within desired region and reduce the occurrence of hypoglycemic episodes [40]. Diabetes simulators have aided in the demonstration that closed-loop control can help people tighten their BGC than open-loop control [41], can help maintain BGC overnight [42], and should be a valuable replacement for animal trials that precede clinical trials [20]. Recent simulation studies show that meal announcement is an essential part in tighter control of BGC during meal challenges [43]. If meal announcement or record is missing, meal detection and meal size estimation algorithms have been developed and can effectively reduce the mean BGC when combined with MPC in closed-loop control [44]. Also, MPC can provide robust performance under different physiological conditions with a low order state space model [45], and outperform traditional PID control [43,45]. Linear time varying MPC has also been applied in BGC control and it shows great potential as well [46]. Other dynamic tuning algorithms are also being demonstrated in simulation settings [30]. There are studies that involve investigation of the performance of PID control with intravenous pumps [47], and control algorithm that makes use of glycemic index (i.e. the composition of food based on categories of carbohydrate) [48]. Most simulation studies last less than 3 days, and usually features 3 meals or less per day.

**Literature Cited**


CHAPTER 3: THE DEVELOPMENT OF A VIRTUAL SENSOR IN GLUCOSE MONITORING FOR NON-INSULIN DEPENDENT PEOPLE

A paper published in Journal of Bioinformatics and Diabetes

Derrick K. Rollins\textsuperscript{1,2*}, Lucas Beverlin\textsuperscript{2}, Yong Mei\textsuperscript{1}, Kaylee Kotz\textsuperscript{1}, David Andre\textsuperscript{3}, Nisarg Vyas\textsuperscript{3}
Greg Welk\textsuperscript{4} and Warren D. Franke\textsuperscript{4}

\textsuperscript{1}Department of Chemical and Biological Engineering, Iowa State University, Ames, Iowa 50011
\textsuperscript{2}Department of Statistics, Iowa State University, Ames, Iowa 50011
\textsuperscript{3}BodyMedia Inc., 4 Smithfield St., Pittsburgh, Pennsylvania 15222
\textsuperscript{4}Department of Kinesiology, Iowa State University, Ames, Iowa 50011

*Corresponding Author

Abstract

Continuous-time glucose monitoring (CGM) effectively improves glucose control, as oppose to infrequent glucose measurements (i.e. using Lancet Meters), by providing frequent blood glucose concentration (BGC) to better associate this variation with changes in behavior. Currently, the most widely used CGM devices rely on a sensor that is inserted invasively under the skin. Because of the invasive nature and also the replacement cost of sensors, the primary users of current CGM devices are insulin dependent people (type 1 and some type 2 diabetics). Most non-insulin dependent diabetics use only lancet glucose measurements. The ultimate goal of this research is the development of CGM technology that overcomes these limitations (i.e. invasive sensors and their cost) in an effort to increase CGM applications among non-insulin dependent people. To meet this objective, this preliminary work has developed a methodology to mathematically infer BGC from measurements of non-invasive input variables which can be thought of as a “virtual” or “soft” sensor approach. In this work virtual sensors are developed and evaluated on 20 subjects using four BGC measurements per day and eight input variables representing meals, activity, stress, and clock time. Up to four weeks of data are collected for each subject. One evaluation consists of 3 days of training and up to 25 days of testing data. The
second one consists of one week of training, one week of validation, and 2 weeks of testing data. The third one consists two weeks of training, one week of validation and one week of testing data. Model acceptability is determined on an individual basis based on the fitted correlation to CGM testing data. For 3 day, 1 week, and 2 weeks training studies, 35%, 55% and 65% of the subjects, respectively, met the Acceptability Criteria that we established based on the concept of usefulness.

**Keywords:** Virtual Sensor, Wiener Modeling, Block-Oriented Modeling, Type 2 Diabetes.

**Introduction**

Recent research suggests that real-time, frequent, glucose monitoring can improve blood glucose control over infrequent monitoring provided through the use of lancet glucose meters for both insulin dependent [1]-[6], and non insulin dependent diabetics [7]. Frequent glucose measurement capability is referred to as continuous-glucose monitoring (CGM); although not really continuous, current devices can deliver on-line glucose measurements as often as every one to five minutes [8]. Nonetheless, this is a substantial improvement over lancet monitoring that only produces a few values (e.g. four values) per day, at best. CGM therefore improves the user’s ability to achieve better glucose control by providing frequent, real-time, glucose concentration levels that enables correlation with activity and food consumption. For example, a user is able to see with a high frequency display rate the extent to which the size of a meal affects glucose changes. Currently, the most widely used and effective CGM devices rely on a sensor that is inserted invasively under the skin. Sensors cost from $35 to $60 and last 3 days to a week. Thus, two significant drawbacks of these devices are comfort and cost [9]. Given these drawbacks, these devices are not widely used except by insulin dependent diabetics that rely heavily on a fast sampling rate for better control. For this reason, these devices are less likely to
be used by non-insulin dependent people, including non-diabetic, pre-diabetics and diet-controlled type 2 diabetics.

**Fig. 1.** The SenseWear® Armband of BodyMedia, Inc.

Hence, the motivation of this work is the development of a useful, non-invasive, subject-specific (personalized), continuous monitoring system in an effort to increase CGM among non-insulin dependent people.

To achieve this goal we seek to develop a low maintenance, high frequency monitoring system with an accuracy that is high enough to be useful for non-insulin dependent people. Moreover, this preliminary work proposes an inferential (i.e., virtual) sensor approach for predicting blood glucose concentration (BGC) from noninvasive inputs. This virtual sensor updates at the same rate as conventional physical sensor CGM devices. The model is developed from lancet BGC measurements that are obtained at a rate of four measurements per day. Since each sensor is calibrated from user data, the model developed for each person is said to be “subject-specific.” While inferential modeling of BGC has been done by a number of researchers [19]-[24], [29], [31] particularly in type 1 diabetic applications using frequent glucose measurements, this is the first approach that we are aware of that seeks to develop an inferential model for non-insulin dependent subjects using infrequent lancet measurements from the subject’s personal lancet glucose meter. Our approach to achieve this goal is to use a novel
modeling method to infer glucose concentration using non-invasive input measurements for each subject from variables representing food, activity, clock time\cite{10}-\cite{12}, and stress\cite{13},\cite{14}.

Methods

The main physical component of this system is a BodyMedia\textsuperscript{®} armband of the type shown in Fig. 1. This device is a multi-sensor monitoring device that provides accurate estimates of physical activity data using accelerometers, heat related sensors and galvanic skin response (GSR)\cite{15}. GSR is the conductivity of the wearer’s skin that varies due to physical and emotional stimuli. For more details see\cite{27},\cite{28}. Given that the armband currently uses complex algorithms (e.g., for pattern recognition) it should also be able to incorporate our proposed BGC prediction algorithm. However, this research is beyond the scope of this article which is focused on the development of the modeling methodology.

The most critical challenge in this highly complex, non-linear, multiple-input, highly underdetermined modeling problem is the estimation of a large set of dynamic and static parameters from a very small set of BGC data, with a sampling frequency of only 4 values per day. To achieve accuracy under these conditions is a significant advancement over the work of Rollins et al. and a unique accomplishment. Other challenges include adequately guarding against over-fitting, the lack of initial steady state data, low quality meal information that uses a designation of small, medium and large, and frequent and arbitrary removal of the armband monitor. Through novel modifications of the Rollins et al.\cite{16} approach, this work demonstrates an ability to overcome these challenges, and thus, has promising potential to develop an effective inferential continuous-time BGC sensor for the target population of non-insulin dependent people. The details of the proposed modeling approach are now described.
The Modeling Approach

The basic objective of this work is the development of a subject-specific “soft sensor” or “virtual” sensor methodology that provides “useful” information to help individuals monitor and control their glucose more effectively than with lancet glucose meters. The most critical and challenging objective in this highly underdetermined problem is that the model must be developed from a BGC sampling rate of only four samples per day. These samples will come from the lancet meter of the subject and the idea is to transform these measurements to a CGM display frequency during the period of the day that the subject is not sleeping. This virtual sensor approach is an inferential model that is developed from measured variables that are termed inputs. This virtual sensor idea has seen wide applications in process monitoring and control applications in recent years [17], [18] due to advancements in computer hardware, software, and measurement technology. Note that to distinguish the type of sensor, i.e., “virtual” versus “physical,” we will use the terms “virtual-sensor” and “physical-sensor.” In addition, it should be noted that our use of “monitoring” include both the use of a virtual-sensor or physical-sensor although virtual sensors do not measure the process variable being monitored directly. This major challenge in this work is the frequency of BGC data for model building (in this research, 4 times per day) is much less than the virtual measurement rate of 5 minutes. This limitation means that the information available for model identification, i.e., parameter estimation, is quite limited and could thus, severely impact accuracy.

The information for the development of a virtual sensor comes from two sources -- the response data set and the input data set. Since the information content of lancet BGC is quite limited, the proposed approach strongly relies on the input data set for information on glucose behavior. More specifically, this data set consists of meal size with three levels, six (6) variables from the BodyMedia armband, and the time of day (TOD) in minutes on the 24 hour clock. The inputs that we selected for this study from the armband are those selected by Rollins et al [16].
We eliminated near body temperature as we determined it was not contributing significantly to glucose behavior for any of the subjects. The inputs are shown in Table 1.

The ability to map the available input/output information to accurate sensor measurements depends on the model structure, the model building procedure, and the inferential algorithm that we are calling the “Inferential Engine.” The model structure consists of the mathematical functions and the network that tie these functions together. The model building (i.e., identification) procedure is the process of using input/output information to estimate the values of unknown parameters in the mathematical functions. The Inferential Engine is the equation used to obtain the virtual sensor measurements at the desired sampling frequency. This equation represents input selection, parameter estimates, and the use of lancet glucose measurements to enhance reliability. The purpose of this section is to describe these three components of the proposed technique in detail.

**Table 1. Input variables: Meal Size (1), Armband (2-7), and TOD (8).**

<table>
<thead>
<tr>
<th>Input</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Meal Size Index</td>
</tr>
<tr>
<td>2.</td>
<td>Transverse accel – peaks</td>
</tr>
<tr>
<td>3.</td>
<td>Heat flux – average</td>
</tr>
<tr>
<td>4.</td>
<td>Longitudinal accel – average</td>
</tr>
<tr>
<td>5.</td>
<td>Transverse accel – MAD</td>
</tr>
<tr>
<td>6.</td>
<td>GSR – average</td>
</tr>
<tr>
<td>7.</td>
<td>Energy expenditure</td>
</tr>
<tr>
<td>8.</td>
<td>Time of day (TOD)</td>
</tr>
</tbody>
</table>
Modeling Structure

The modeling structure of this application must permit accurate parameter estimation under a small number of sampling times ($n$), effectively handling several inputs with different dynamic behavior, and mild extrapolation. The proposed modeling network is what we call the Coupled Dynamic Insulin (CDI) network (its structure is given in Fig. 2). As shown, the first input, meal size ($x_1$), enters both a linear dynamic food block ($G_1$) and a linear dynamic unmeasured insulin block ($G_I$). The output from the unmeasured dynamic insulin block ($v_I$) enters a pseudo blood insulin block which is coupled with the food block ($G_1$) which produces the dynamic food input ($v_1$) to the pseudo BGC block. Then, the unmeasured output from the coupled food block is the dynamic glucose ($G_f$) input due to food consumption. Each of the other inputs (e.g. inputs 2-8) enters a separate linear dynamic block and the outputs from these blocks are collected into non-observable variables ($v_i$) and together with $G_f$ are passed through a static block which can be any type of function. The CDI model simulates the process where food digestion is responsible for the rise of blood glucose after each meal, while the secretion of insulin is responsible for the fall of blood glucose level a period of time later after the meal. The CDI network is defined by the attributes of allowing separate dynamic behavior for each input and the use of variables for unmeasured insulin generation ($v_I$) and unmeasured blood insulin concentration ($I$). To our knowledge, this is first application of unmeasured pseudo insulin in modeling blood glucose concentration. This idea is a key reason for the success of our modeling approach in this application of infrequent BGC measurements.

The dynamic functions for $G_i$, $i = 1, \ldots, p$, $I$ (the $I$ is for insulin), follow the modeling work of Rollins et al. [16] and are second order differential equations of the form:

$$
\tau_i^2 \frac{d^2y(t)}{dt^2} + 2\tau_i \nu \frac{dy(t)}{dt} + y(t) = \tau_u \frac{dx(t)}{dt} + x(t)
$$

(1)
where $x_i(t)$ is the $i$th input, $i$ varies from 1 to $p$, $p$ is the total number of inputs, $\tau_{ai}$ is the lead parameter, $\tau_i$ is the time constant, and $\zeta_i$ is the damping coefficient, with $x_1 = \text{meal size input variable}$, $x_i$, $i = 2, \ldots, p-1$, are armband input variables, and $v_p$ is the TOD input variable.

Using backward difference finite derivative approximations, Eq. (1) gives (Rollins et al., [16])

$$\frac{v_{ij}}{\Delta t} = \frac{2\xi_i \tau_i \Delta t}{\tau_i^2 + 2\xi_i \tau_i \Delta t + \Delta t^2} v_{ij-\Delta t} + \frac{-\Delta t^2}{\tau_i^2 + 2\xi_i \tau_i \Delta t + \Delta t^2} v_{ij-2\Delta t}$$

$$+ \tau_{ai} \Delta t + \Delta t^2 x_{ij-\Delta t} + \frac{-\tau_{ai} \Delta t}{\tau_i^2 + 2\xi_i \tau_i \Delta t + \Delta t^2} x_{ij-2\Delta t}$$

with

$$v_{ij} = \delta_{ij} v_{i,j-\Delta t} + \delta_{ij} v_{i,j-2\Delta t} + \alpha_{ij} x_{i,j-\Delta t} + \omega_{ij} x_{i,j-2\Delta t}$$

such that $\omega_{i,2} = 1 - \delta_{i,1} - \delta_{i,2} - \omega_{i,1}$. This constraint is used to impose a unity gain restriction for the linear dynamic blocks. $\Delta t$ is the sampling time for the inputs. In the Laplace domain, the linear dynamic functions are

$$G_i(s) = \frac{V_i(s)}{X_i(s)} = \frac{\tau_{ai} s + 1}{\tau_i^2 s^2 + 2\tau_i \xi_i s + 1}$$

Note that the number of dynamic parameters associated with each input is three. This small number is a strength that we exploit to obtain parameter estimates under limited sampling, as discussed below. The CDI model for food alone is represented by the following coupled Eqs. (5) and (6):

$$\frac{dG_j(t)}{dt} = \alpha_1 v_j(t) - \alpha_2 G_j(t) I(t)$$

$$\frac{dI(t)}{dt} = \alpha_3 v_j(t) - \alpha_4 I(t)$$
where $v_1$ and $v_2$ are outputs from dynamic blocks $G_1$ and $G_2$ respectively, and $\alpha_1$ to $\alpha_4$ are the “coupled” model parameters.

We also use backward difference finite derivative approximation on Eqs. (5) and (6) to give

$$G_{f,t} = \frac{\alpha_1 v_{1,t} \Delta t + G_{f,t-\Delta t}}{1 + \alpha_2 I_{t} \Delta t} \tag{7}$$

$$I_{t} = \frac{\alpha_3 v_{1,t} \Delta t + I_{t-\Delta t}}{1 + \alpha_4 \Delta t} \tag{8}$$

Note there are four additional parameters ($\alpha_1$, $\alpha_2$, $\alpha_3$, and $\alpha_4$) that need to be identified.

The function $f(V)$ is called “the static function” and is a function of all of inputs. This function can theoretically be of any form. For effectiveness under mild extrapolation and minimum parameter estimation (as discussed below) we have chosen a first order linear regression model of the form:

$$y_i = \eta_i + \epsilon_i = a_0 + G_{f,t} + a_2 v_{2,t} + \ldots + a_p v_{p,t} + \epsilon_i \tag{9}$$

where $\epsilon_i$ is the error term and assumed to be independently normally distributed with mean 0 and variance $\sigma^2$ for all $t$, and $a_i$’s are static parameters.

As stated in Rollins et al. [16], the modeling objective is simply to maximize the true but unknown correlation coefficient between measured and fitted BGC. This quantity is represented by $\rho_{y,y}$ and estimated by $r_{fit}$. Thus, under this criterion, as a minimum, a model is considered useful, if, and only if,

$$\rho_{y,y} > 0 \tag{10}$$
Since the degree of usefulness increases with \( \rho_{\hat{y}, \hat{y}} \), the goal is to obtain the largest (as close to the upper limit of 1) value as possible. Due to the highly complex mapping of the parameters into the response space of \( r_{\beta_t} \), the following indirect criterion is used in obtaining the parameter estimates as described in Rollins et al. [16].

\[
\text{Maximize } r_{\beta_t} \text{ by Minimizing } \quad \text{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]

Subject to: \( \zeta_i > 0, \tau_i > 0, \theta_i \geq 0 \quad \forall i \)  

Note that only training data are used to compute SSE under Eq. (11).

**Model Identification Procedure**

We use the CDI network with Eqs. (1)-(9) and developed a procedure that can accurately estimate the \( 3(p+1) \) dynamic parameters, 4 coupled parameters (\( \alpha_1, \alpha_2, \alpha_3, \) and \( \alpha_4 \)) and the \( p \) static parameters even when the number of sampling times (\( n \)) is much less than \( 4p + 7 \), the total number of parameters. This procedure requires each input to have a separate set of dynamic parameters as uniquely met by the Wiener network but not by other common networks (e.g. such as the **Auto Regressive Moving Average with eXogenous (ARMAX)** variables network) [16].

Let \( G_{f,t} = 0 \) and \( v_{l,t} = 0 \) for all \( i \) in Eq. (9) except for one value of \( i = j \), i.e., \( v_i = v_j \neq 0 \), for one value of \( i \), \( i = 2, \ldots, p \). Thus, with only one input variable \( v_i = v_j \), Eq. (9) becomes a simple linear regression model (SLRM). To distinguish this SLM from Eq. (9), the fitted form is written as
where \( \hat{y}_{i,t} \) is the fitted BGC for the one input \( i \) at \( t \); \( \hat{y}_{oi} \) and \( \hat{y}_{i} \) are the estimated intercept and slope parameters, respectively, for the SLRM for input \( i \); and \( \hat{w}_{i,t} \) is the SLRM estimate of \( v_{i,t} \).

Note that for fitting the SLRM, only five (5) parameters (the temporary static parameters \( \gamma_0 \) and \( \gamma_i \), and the permanent dynamic parameters \( \tau_i, \zeta_i \) and \( \tau_{ai} \)) are estimated each time which, as necessary, is less than \( n = 12 \) for three days of data collection, for example.

In Appendix A, a proof is given to show that for the SLRM, \( r_{fit} = r_{y,v_i} \). More specifically, for the SLRM, \( r_{fit} \) is determined by \( \hat{w}_{i,t} \), only and not by the static model coefficients, \( \hat{y}_{oi} \) and \( \hat{y}_i \). Thus, for the SLRM, since \( v_i \) only depends on the dynamic parameters for input \( i \), one can find the set of dynamic parameters that results in the best \( r_{fit} \) for each input \( i \) separately (i.e., \( \tau_i, \zeta_i \) and \( \tau_{ai} \)). We exploit this result by decomposing the modeling problem into separate sub-problems that will be identified in 3 steps: 1. the dynamic parameters for each input \( i \), \( i = 2, \ldots, p \), under Eq. (12) (five parameters are estimated for each \( i \)); 2. the insulin and food dynamic parameters under Eqs. (7),(8), (13) (eleven parameters are estimated; one temporary intercept parameter, four coupled parameters for initial values to be used in Step 3, and six permanent dynamic parameters); and 3. the permanent static and coupled parameters with all the inputs included under Eq. (9) (at most \( p + 4 \) coupled parameters are estimated).
In Step 1, our current procedure is to manually adjust the dynamic parameters one input at a time to find the “best” set of values for each input. Our definition of “best” will be given momentarily. In Step 2, the following reduced form of Eq. (9) is applied:

\[ y_t = \lambda_0 + G_{f,t} + \epsilon_t \]  \hspace{1cm} (13)

where \( \lambda_0 \) is a temporary parameter only used in Step 2. This step is the most challenging. With a given set of initial values, either some or all the parameters are estimated simultaneously using an effective nonlinear regression algorithm. This process is the most iterative and time consuming as some parameters are manually set and fixed and the rest are estimated using the optimization algorithm. This process is iteratively repeated until no more improvement can be made in \( r_{fit} \). The only input involved in Step 2 is food, i.e., meal size. If an adequate \( r_{fit} \) (\( r_{fit} \) for training should be positive in agreement with Eq (10)) is not found in this step, the modeling procedure is terminated, and we conclude that the procedure failed to find an adequate model for this subject from the given data sets. Step 3 is completed in one estimation trial when the dynamic and couple modeling parameters from Steps 1 and 2 are used since Eq. (9) becomes a first order linear regression model. However, if one desires to estimate the couple modeling parameters also, Eq. (9) becomes a nonlinear regression model and the estimation process can be more challenging. Note that, for example, with \( n = 12 \), or three days of data collection and \( p = 8 \), at most twelve parameters are estimated in Step 3, which is still not exceeding \( n \). At the end of all three estimation steps, with \( p = 8 \), \( 4p + 7 \) or 39 total parameters have been uniquely estimated in a highly nonlinear modeling problem from at least \( n =12 \) or three days of data collection, for example. This ability is a critical novelty and a powerful benefit of this approach.

The “best” set of modeling parameters is determined for two given scenarios. The first one only uses a Training set of data. In this scenario, the goal is to maximize \( r_{fit} \) of the training set. Consequently, the procedure is to reach convergence at the global minimum for the least
squares objective criterion. This estimation procedure is “unsupervised” training (note this is a different definition from that of T. Hastie’s book [25] and A.J. Izenman’s book [30]). The second scenario is when there are both Training and Validation sets of data. In this scenario, the goal is to determine the largest $r_{fit}$ for the Validation data set with a “close” value of $r_{fit}$ for the Training data set. Here, convergence for the Training set may not be reached and the Validation set determines when the iterative process terminates. Since the Validation results determine when the optimization process terminates, this is a type of “supervised” training. This procedure is used to guard against overfitting, (i.e., fitting BGC behavior in the Training set that is not due to true variation in BGC). The success of both types of training is evaluated through the use of an additional set of data called the “test set” which had no influence on the model identification process (i.e., the parameter estimates). The first scenario is used when $n$ is small, say 12, the number after 3 days, whereas the second scenario is used otherwise, e.g., when $n$ is 24, the number after 6 days. Parameter estimation was done using the Excel® Solver Routine.

Successful model identification relies on effective selection of initial conditions and starting values for model parameters and the dynamic inputs (i.e., the $v_i$’s). The following procedure is given under a protocol where the armband is worn nearly 24 hours a day and removed only for showering. The initial steady state is chosen during a period of slow change, commonly early in the morning. The set of initial values in our procedure are $\tau_i = 1.1$, $\zeta = 0.9$, $\alpha_2 = 20$, $\alpha_4 = 0.1$, and all other initial parameter values are equal to zero. The initial values for the $v_i$’s have to be determined iteratively. When the dynamic parameters are set to values so are the $v_i$’s as shown by Eq. (2). Our procedure is to set the initial values of the $v_i$’s to their average values over the training data. These values are to remain fixed during estimation and changed after estimation of dynamic parameters. The estimation process for a set of dynamic parameters
is completed when the “best” $r_{fit}$ is obtained with initial values of $v_i$’s close to their average values for the training data.

For the missing data due to removal of armband[32]: if data missing lasts for a short period of time (e.g. no more than one hour of missing data), the missing data were interpolated with the average value of the two sides of the missing data interval. If data missing lasts longer than one hour, we set the missing data to its initial value.

**Development of the Inferential Engine**

After obtaining a full set of parameter estimates, the proposed model development procedure has two more refinements. The first one is elimination of any armband inputs that adversely affect the value of $r_{fit}$. This is done by setting each, and only one, $a_i$ (for $i = 2, \ldots, 7$) to zero at a time, and observing $r_{fit}$. If $r_{fit}$ increases for the Training set in Scenario 1 or for the Validation set for Scenario 2, this input is removed. After this process is completed for each input, all the remaining static parameters are estimated under Step 3 for a final time.

The final refinement involves the use of lancet glucose to help to reduce model bias. Since these measurements are infrequent and are not measured at a constant rate, it is not possible to build a correction model based on the correlation of residuals. The correction equation that we use comes from Rollins et al. [16] where only the most recent measurement, at $t = t^*$, is used. This equation, which represents the proposed virtual sensor, is given as:

$$
\hat{y}_t = \hat{\eta}_t + (y_{t^*} - \hat{\eta}_{t^*})\lambda \frac{t - t^*}{\Delta t}
$$

subject to: $t > t^*$ and $0 < \lambda < 1$, where $\lambda$ is an adjustable constant, $y_{t^*}$ is the lancet BGC measurement at $t = t^*$, $\hat{\eta}_t$ = the estimated BGC at time $t$ under the Eq. (9) model, $\hat{\eta}_{t^*}$ = the estimated BGC at time $t = t^*$ under the Eq. (9) model, and $\hat{y}_t$ = the virtual (i.e., soft) sensor value for the proposed method at time $t$. Note that $y_{t^*} - \hat{\eta}_{t^*}$ represents that amount of correction and this correction diminishes as time increases based on the value of $\lambda$ which is close to 1. Thus, by the
time the next lancet measurement is taken, usually $\hat{y}_i = \tilde{\eta}_i$. This means that at $t = t^*$, $\hat{y}_i = \tilde{\eta}_i$; at $t = t^* + \Delta t$, $\hat{y}_i = y_i$; and for $t = t^* + k\Delta t$, with $k \gg 1$ and before the next lancet measurement, $\hat{y}_i = \tilde{\eta}_i$.

That is, at the time of the lancet measurement, the proposed virtual monitor would display a value close to $\tilde{\eta}_i$, the next value would be close to the lancet measurement, and as time proceeded, the lancet value would have less corrective influence as the predictor would rely more on the model to infer BGC. When two sets are used to estimate model parameters, $\lambda$ can be set to give the most accurate values in the validation set. When only a training set is used to estimate the model parameters, a default value can be used based on results from modeling several subjects.

**Clinical Study for 22 Subjects**

For the proposed method, the development of a virtual-sensor requires 4 lancet measurements per day spread as evenly as possible over the time the subject is awake in about a 14 hour period. We did not have access to data meeting this requirement. However, from a previous study, we had physical-sensor CGM data sets which were collected with Institutional Review Board (IRB) approval, and the data sets were used to develop and evaluate the methodology. Thus, these data sets played two roles. First, for each subject, they played the role of a surrogate person, i.e., the true BGC for the purpose of evaluation. Secondly, they played the role of the lancet sampled data, i.e., the data used to build the virtual-sensors.

Using 22 test subjects (see Table 2) with 4 weeks of data collection (in most cases and slightly under 4 weeks in other cases except for Subject 1 and 8 which had only about 3 weeks of data due to loss data), we have obtained results to support the modeling viability. As just stated, these data sets were collected for another study (see Beverlin et al. [26], [32]). Modifications had to be made to these data sets for use in this study. First, food quantities, which were in grams of carbohydrates, fats and proteins, had to be converted to a food index representing meal sizes
with 0 for no meal, 1 (two time stamps) for a small meal, 2 (three time stamps) for a medium meal and 3 (four time stamps) for a large meal. In practice the time stamps will be entered by the user pressing the time stamp button on the armband at the start of a meal. The conversion we used was based on the grams of carbohydrates only with less than 20 grams being a small meal, more than 100 grams being a large meal and all other amounts considered a medium size meal.

Secondly, infrequent BGC measurements were not obtained from a lancet meter but converted from a continuous glucose monitoring system (CGMS) at a sampling rate of only four values per day at particular and fixed times (i.e. only 4 values per day out of the continuous readings from CGMS were used) to mimic infrequent lancet sampling. CGMS values were taken only at 8 am, noon, 4 pm and 8 pm; if data were unavailable, then the nearest value was taken with no more than 4 values used per day. The monitoring period was taken to be from 8 am to 10 pm daily which means that this was the only period that virtual BGC were reported. Thus, the period from 10 pm to 8 am was taken to be a non-monitoring period in order to mimic that monitoring is not required during the sleeping period.

Note that, the original data sets contain meal information in terms of grams of carbohydrates, fats and proteins. The amounts were calculated from self reporting logs of the type and quantities of food eaten. Hence, the errors of these quantities are likely quite high at times and it is likely that a significant number of meals were not recorded or logged at the proper times. When we converted the quantities to an index value for meal size for this study (i.e. “1” represents small meal size, “2” for medium size, and “3” for large size), we applied the same conversion equation to all of the subjects. Thus, the quality of food information that we developed our models from in this study is quite poor. Therefore, since these results are obtained under poor food information they indicate the robustness of the technique to low quality food information.
Before evaluating model acceptability, subject 21 and 22 were rejected due to poor food information. As a result, subject 21 and 22 were removed from this research from this point on.

**Measures of Performance**

Model acceptability will be determined on an individual subject basis given that the models are subject-specific and each individual will only be concerned about model accuracy as it pertains to model developed for them. Thus, this study is evaluated based on the number of subject-models that meet a particular Acceptability Criteria. But we state and justify this criteria momentarily, after we present the statistics they it uses.

The first one is called the averaged error (AE) and is simply the average value of the residuals:

\[
AE = \frac{\sum_{i=1}^{m} y_i - \hat{y}_i}{m}
\]

where \( m \) is the number of terms being averaged.

The second one is called the averaged absolute error and is similar to Eq. (15) except that the absolute difference is used for the term in the summation as follows:

\[
AAE = \frac{\sum_{i=1}^{m} | y_i - \hat{y}_i |}{m}
\]

A scaled AAE value to adjust for spread is used called the relative AAE (RAAE). This measure of performance is determined by dividing Eq. (16) by the standard deviation of the values used to calculate AAE as follows:

\[
RAAE = \frac{AAE}{\sqrt{\sum_{i=1}^{m} (y_i - \bar{y}_i)^2 / m}} = \frac{AAE}{\text{stdev}(y_i)}
\]
RAAE is a relative AAE statistic that accounts for large spread in the glucose variation of subjects. For replicated lancet measurements, the study in Rollins et al. [16] determined RAAE to be about 0.60. Thus, we will assume that a value around 0.60 is comparable to the performance of a glucose lancet meter. However, lancet accuracy or even repeatability can vary widely from individual to individual due to the accuracy of the device, inherent variability in the measurement protocol, and human error. Nonetheless, since this is the only result that likely exists in this type of study (i.e., four weeks of data collection under the protocol of this study) we will use it in our criteria. It is also noted that, given that the models in this study are developed from three discrete levels of food size and not from the three types of consumed quantities, and from a much lower frequency of glucose data when comparing to typical CGMS values (i.e., four values per day versus 12 values per hour), we expected RAAE to be higher and allow for slightly higher values in the Acceptability Criteria.

The last statistic or performance measure is $r_{fit}$. Based on the results in Rollins et al. [16] for a type 2 diabetic and in Beverlin et al. [26] for the 20 subjects used here, we set a minimum acceptable value for $r_{fit}$ of 0.40. Using these three measures of performance, the Model Acceptability Criteria (MAC) for this study is given as:

$$r_{fit} = \begin{cases} 0.6 \\ \text{or } \max(0.4, 0.05 \text{ AAE} - 0.4) \\ \text{or } \max(0.4, \text{ RAAE} - 0.2) \end{cases}$$ (18)

As the MAC shows, a model with an $r_{fit}$ of at least 0.6 is considered acceptable based on this value alone. However, if this value is between 0.4 and 0.6, the fitted model must meet a certain
level of performance on both r_{fit} and AAE or both r_{fit} and RAAE. For example, if AAE and RAAE are 18 mg/dL and 0.75, respectively, the AAE sub-criterion is r_{fit} ≥ 0.5 and the RAAE one is r_{fit} ≥ 0.55. Thus, for this example, the fitted model will meet the MAC if, and only if, r_{fit} ≥ 0.5.

As this example illustrate, when r_{fit} is in this range, the MAC is written to require r_{fit} values to be higher than 0.4 with this requirement increasing as AAE and RAAE increase. Note that on the upper boundary of r_{fit} = 0.6, AAE must be < 20 mg/dL or RAAE must be < 0.8, and on the lower boundary of r_{fit} = 0.4, AAE < 16 mg/dL or RAAE < 0.6. The lower boundary was defined based on the results in Rollins, et al. [16] and the upper boundary was established from an examination of fitted models in this work. See Fig. 3 for plots of three fitted models as they compare with CGM measured responses at the limits and middle of the MAC.

Table 2. Characteristic information on the 22 subjects used in this study.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender</th>
<th>Age</th>
<th>BMI</th>
<th>BGC Average (mg/dL)</th>
<th>BGC Standard Deviation (mg/dL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>44</td>
<td>25.8</td>
<td>106.7</td>
<td>26.5</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>40</td>
<td>36.7</td>
<td>107.3</td>
<td>19.5</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>55</td>
<td>26.8</td>
<td>110.5</td>
<td>19.3</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>51</td>
<td>22.5</td>
<td>112.9</td>
<td>14.9</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>55</td>
<td>24.0</td>
<td>115.0</td>
<td>22.0</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>57</td>
<td>27.3</td>
<td>124.9</td>
<td>17.0</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>42</td>
<td>31.9</td>
<td>154.6</td>
<td>33.1</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>55</td>
<td>26.6</td>
<td>132.3</td>
<td>24.5</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>55</td>
<td>26.0</td>
<td>131.8</td>
<td>27.0</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>47</td>
<td>30.8</td>
<td>135.7</td>
<td>32.6</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>55</td>
<td>42.5</td>
<td>143.6</td>
<td>61.1</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>49</td>
<td>23.1</td>
<td>103.3</td>
<td>14.8</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>40</td>
<td>29.0</td>
<td>122.1</td>
<td>23.2</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>56</td>
<td>23.4</td>
<td>105.9</td>
<td>19.6</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>35</td>
<td>39.5</td>
<td>137.6</td>
<td>28.4</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>53</td>
<td>22.1</td>
<td>107.1</td>
<td>22.2</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>60</td>
<td>29.0</td>
<td>180.5</td>
<td>61.5</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>56</td>
<td>23.5</td>
<td>113.9</td>
<td>23.1</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>43</td>
<td>25.5</td>
<td>122.8</td>
<td>23.4</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>52</td>
<td>27.8</td>
<td>117.7</td>
<td>18.3</td>
</tr>
<tr>
<td>21</td>
<td>F</td>
<td>54</td>
<td>36.3</td>
<td>232.2</td>
<td>60.4</td>
</tr>
<tr>
<td>22</td>
<td>M</td>
<td>47</td>
<td>28.0</td>
<td>120.5</td>
<td>18.6</td>
</tr>
</tbody>
</table>
Results

The results of this study are given in Tables 3-5. Each table represents a different training period. There are two types of predictions in these tables; \( \hat{\eta} \), the values of the fitted model at the time the lancet measurements were taken and \( \hat{y} \), fitted model at the sampling rate of the CGMS, i.e., every five minutes. The model parameters are estimated using \( \hat{\eta} \), but the MAC is applied to the testing results for \( \hat{y} \) to determine model acceptability on an individual basis which is shown in the tables. In addition, summary results for the performance measures are given for all the subjects and for the set of subjects meeting the MAC.

Table 3 gives the modeling results are for three (3) days of training under Eq. (13) (i.e., for food only). Since training stop at convergence under the least squares criterion given by Eq. (11), the remaining days consisted of the test set. In addition, for all these subjects, \( \zeta_1 = 0.2 \), \( \tau_{a1} = 0 \), \( \zeta_I = 0.8 \), and \( \tau_{aI} = 0 \). This was done to increase the degrees of freedom to estimate the more critical parameter \( \tau_1 \), \( \tau_I \) and the four coupled parameters and to simplify the optimization. The best choice for these values is future research work. As Table 3 shows, the results indicate that 35% of the cases met the MAC. This is really quite promising as a minimum initial calibration period given that the number of data points, \( n \), used is only 12. In practice, it appears that a significant number of subjects could have successful calibration after three days and as more data are collected this number would grow. This conclusion is supported by the results in the next two tables.

Table 4 contains results under Eq. (9) for one week of training, one week of validation and two weeks of testing. As shown, 55% of these cases met the MAC. In addition, for this group that meets MAC versus all the cases in table 4 as whole, the average values of AAE and RAAE dropped considerably from 19.8 mg/dL and 0.76 to 13.5 mg/dL and 0.71, respectively, while \( r_{fit} \) increased from 0.47 to 0.55. These values are excellent. Table 5, also under Eq. (9),
contains results for two weeks of training, one week of validation and one week of testing. As shown, the number meeting the MAC further to 65% (it is a promising result given the strict MAC) with very good average results for this group with AAE = 15.0 mg/dL, RAAE = 0.74 and \( r_{fit} = 0.54 \).

We found that using the armband inputs increases \( r_{fit} \) for \( \hat{y}_t \) by 0.1 over using just food alone. (These cases are not shown for space considerations). Thus, both food and the armband inputs are to obtain the results presented in this section. The robustness to poorer food quality is supported by similar \( r_{fit} \) values for this study as compared to the ones in Beverlin et al. [26] and Beverlin [32] where food quantities were used on these same data sets.

**Concluding Remarks**

This article presented preliminary work on the development of a virtual sensor for BGC with the objective of developing a noninvasive CGM system that could increase CGM among non-insulin dependent people. This device would require users to wear a readily available armband monitor and manually entering meal sizes through the use of a button on the armband. This device would require four (4) lancet measurements per day as most current invasive CGMSs require.

The modeling methodology presented in this work is quite powerful. It takes on the challenge of modeling BCG in a highly complex, non-linear, multiple-input, highly underdetermined problem. As illustrated in this work, it is able to develop useful multiple-input dynamic models for BGC under free living, outpatient, data collection from just four glucose measurements per day and from as little as three days of data. In addition, these results are achieved with minimal food information of only three discrete levels. This ability stems from a number of innovative ideas to overcome several challenges in this complex modeling problem as
follows. First, the use of the coupled structure allows for the inclusion of inferential blood insulin concentration and leads to insulin and glucose interaction in the blood. This structure is a significant advancement over a straight Wiener network and contributes significantly to the accuracy and ability to obtain adequate fitting for acceptable model usefulness. Secondly, the result in Appendix A provided the knowledge that produced the idea to decompose the modeling problem into a dynamic part and a static part. Added to this idea is the inspiration of determining the dynamic parameters for each input, one input at a time. Once the dynamic parameters are determined for each input, they are fixed. Note that, from the use of a validation set we are able to control over-fitting and by controlling $r_{fit}$ to be about the same in the training set and validation set for each input separately, we have found that this helps the final $r_{fit}$ in all the data sets (Training, Validation, and Testing) to be quite similar. After obtaining the dynamic parameters, the low number of static parameters is then obtained separately as a linear regression model. Thirdly, as the results show, the correction provided by Eq. (14) contributes strongly to the accuracy of the proposed method in the case of continuous glucose monitoring (CGM). While this correction does contribute significantly to the reduction in bias, it also contributes majorly in the reduction in AAE. This can only occur if there is a significant positive correlation for the fitted response, as the correction brings it close at the times infrequent measurements occur, but the correlation determines the direction from these points. If the correlation is not positive, the trend will be in the wrong direction and accuracy would suffer tremendously. This is why subjects must have a significant positive correlation for acceptability. In practice if this is not achieved, the device would simply give a calibration error and not report measurements. Lastly, we developed the new Model Acceptability Criteria (MAC) which instead of evaluating separate aspects of performance such as correlation or bias, and is able to evaluate all of them and also can give a summary statistics (MAC Passing Rate) on the sample population.
This work applied and improved the methodology from Rollins et al. [16]. It was not the purpose of this work to improve on this approach by investigating the impact of other armband variables, or the time variant nature of model parameters, as well as other model improvement issues. To develop the best model for a specific subject, these issues could be considered but they add to the modeling overhead, which is already very high given the small amount of information. The value of this work lies in that the modeling methodology shows great potential in modeling BGC, and provided powerful tools for statistical learning in real life scenario. Nonetheless, these are issues that can be addressed in future research.

Future work will involve running clinical studies under the protocol that subjects will follow when wearing the device such as time stamping for meal size and using only their glucose meter to collect data. If these studies are successful, we plan to develop a prototype armband and evaluate it on several subjects. We envision this device collecting input and output data into the armband where the model will reside. After a sufficient number of lancet measurements have been collected, the model will be built from these data automatically for calibration of the device. After successful calibration, the armband will collect input data, infrequent output data, and display BGC continuously over time on a watch type display or smart phone. Transmission of data from the armband to the display monitor may utilize Bluetooth technology.

We have overcome many challenges such as the use of a food index, the lack of initial conditions, frequent and long term removal of the armband and multiple inputs, subject-specific, modeling under infrequent sampling. However, as this work is only preliminary, there are still several challenges to overcome. This includes finding novel ways to improve the accuracy that leads to a higher percent of users meeting the MAC. In addition, the model procedure is quite complex as it requires advanced modeling experience and consists of several steps. One way we plan to improve accuracy is by gaining a better understanding on the bounds of each parameter. To address the model identification issue, we plan to development an estimation algorithm that
identifies parameters automatically. This program will reside in the armband and will be used to calibrate the virtual sensor from on-line data. These are areas of future research that we have begun and the results are quite promising.

Fig. 3. Graphical examples during the testing period for three subjects meeting the MAC: Subject 2 (strongly meeting the MAC); Subject 6 (weakly meeting the MAC) and; Subject 20 (moderately meeting the MAC).
### Table 3. Modeling results of 3 days training and up to 25 days of testing data

<table>
<thead>
<tr>
<th>Subject</th>
<th>( \hat{\eta}_t )</th>
<th>( \hat{y}_t )</th>
<th>Meeting Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Testing</td>
<td>Testing</td>
</tr>
<tr>
<td></td>
<td>AE</td>
<td>AAE</td>
<td>RAAE</td>
</tr>
<tr>
<td>1</td>
<td>0.0</td>
<td>11.6</td>
<td>0.48</td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>10.9</td>
<td>0.55</td>
</tr>
<tr>
<td>3</td>
<td>0.0</td>
<td>7.5</td>
<td>0.66</td>
</tr>
<tr>
<td>4</td>
<td>0.0</td>
<td>6.8</td>
<td>0.45</td>
</tr>
<tr>
<td>5</td>
<td>0.0</td>
<td>10.2</td>
<td>0.57</td>
</tr>
<tr>
<td>6</td>
<td>0.0</td>
<td>14.7</td>
<td>0.99</td>
</tr>
<tr>
<td>7</td>
<td>0.0</td>
<td>20.2</td>
<td>0.51</td>
</tr>
<tr>
<td>8</td>
<td>0.0</td>
<td>6.4</td>
<td>0.51</td>
</tr>
<tr>
<td>9</td>
<td>0.0</td>
<td>14.0</td>
<td>0.57</td>
</tr>
<tr>
<td>10</td>
<td>0.0</td>
<td>15.0</td>
<td>0.58</td>
</tr>
<tr>
<td>11</td>
<td>0.0</td>
<td>38.5</td>
<td>0.64</td>
</tr>
<tr>
<td>12</td>
<td>0.0</td>
<td>9.2</td>
<td>0.70</td>
</tr>
<tr>
<td>13</td>
<td>0.0</td>
<td>20.5</td>
<td>0.75</td>
</tr>
<tr>
<td>14</td>
<td>0.0</td>
<td>19.3</td>
<td>0.69</td>
</tr>
<tr>
<td>15</td>
<td>0.0</td>
<td>33.3</td>
<td>0.88</td>
</tr>
<tr>
<td>16</td>
<td>0.0</td>
<td>9.0</td>
<td>0.71</td>
</tr>
<tr>
<td>17</td>
<td>0.0</td>
<td>37.5</td>
<td>0.45</td>
</tr>
<tr>
<td>18</td>
<td>0.0</td>
<td>19.9</td>
<td>0.70</td>
</tr>
<tr>
<td>19</td>
<td>0.0</td>
<td>14.8</td>
<td>0.55</td>
</tr>
<tr>
<td>20</td>
<td>0.0</td>
<td>10.3</td>
<td>0.56</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0</td>
<td>16.5</td>
<td>0.62</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0</td>
<td>9.7</td>
<td>0.14</td>
</tr>
</tbody>
</table>

For cases meeting the criteria

| Mean    | 0.0  | 12.6 | 0.64 | 0.47 | -4.0  | 18.8 | 0.80 | 0.27 | -1.5 | 15.2 | 0.69 | 0.50 | 35%   |
| Stdev   | 0.0  | 5.8  | 0.09 | 0.25 | 9.5   | 3.7  | 0.12 | 0.08 | 5.7  | 1.9  | 0.10 | 0.07 |       |
Table 4. Modeling results of 1 week training, 1 week validation and 2 weeks testing data

<table>
<thead>
<tr>
<th>Subject</th>
<th>( \hat{\eta}_t )</th>
<th>( \hat{y}_t )</th>
<th>Meeting Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Validation</td>
<td>Testing</td>
</tr>
<tr>
<td></td>
<td>AE  AAE RAAE</td>
<td>AE  AAE RAAE</td>
<td>AE  AAE RAAE</td>
</tr>
<tr>
<td></td>
<td>( r_{\hat{\eta}} )</td>
<td>( r_{\hat{\eta}} )</td>
<td>( r_{\hat{\eta}} )</td>
</tr>
<tr>
<td>1</td>
<td>-1.0 14.9 0.64 0.60</td>
<td>-5.8 17.3 0.80 0.42</td>
<td>7.9 24.3 0.74 0.40</td>
</tr>
<tr>
<td>2</td>
<td>-1.1 12.1 0.62 0.60</td>
<td>-2.0 12.7 0.57 0.63</td>
<td>7.7 17.6 0.65 0.54</td>
</tr>
<tr>
<td>3</td>
<td>-2.8 8.6 0.59 0.64</td>
<td>-2.5 11.3 0.64 0.57</td>
<td>13.1 15.7 0.94 0.45</td>
</tr>
<tr>
<td>4</td>
<td>-0.8 9.0 0.69 0.42</td>
<td>-0.6 7.8 0.71 0.47</td>
<td>13.1 11.1 0.93 0.32</td>
</tr>
<tr>
<td>5</td>
<td>2.9 16.1 0.63 0.50</td>
<td>-4.2 18.8 0.71 0.37</td>
<td>1.6 15.6 0.75 0.33</td>
</tr>
<tr>
<td>6</td>
<td>-0.5 17.2 0.72 0.45</td>
<td>-10.1 14.7 1.10 0.36</td>
<td>-7.8 15.8 0.81 0.50</td>
</tr>
<tr>
<td>7</td>
<td>0.3 25.2 0.74 0.42</td>
<td>1.9 13.7 0.60 0.52</td>
<td>2.6 15.1 0.71 0.51</td>
</tr>
<tr>
<td>8</td>
<td>0.0 13.7 0.67 0.38</td>
<td>6.1 23.6 0.71 0.44</td>
<td>0.0 22.0 0.72 0.40</td>
</tr>
<tr>
<td>9</td>
<td>0.2 16.3 0.67 0.63</td>
<td>-2.8 14.4 0.70 0.68</td>
<td>-0.3 18.6 0.70 0.49</td>
</tr>
<tr>
<td>10</td>
<td>1.3 15.2 0.40 0.80</td>
<td>-4.8 23.2 0.69 0.49</td>
<td>-7.9 24.1 0.69 0.53</td>
</tr>
<tr>
<td>11</td>
<td>0.0 42.3 0.65 0.55</td>
<td>-16.1 37.8 0.97 0.54</td>
<td>-24.3 45.3 0.80 0.43</td>
</tr>
<tr>
<td>12</td>
<td>0.0 8.7 0.64 0.38</td>
<td>9.7 15.5 1.01 0.29</td>
<td>3.6 11.2 0.89 0.27</td>
</tr>
<tr>
<td>13</td>
<td>3.1 19.1 0.91 0.49</td>
<td>14.5 19.8 1.30 0.53</td>
<td>16.9 24.4 0.72 0.52</td>
</tr>
<tr>
<td>14</td>
<td>-1.1 14.4 0.54 0.74</td>
<td>-7.6 13.7 0.83 0.43</td>
<td>-7.6 18.0 0.77 0.53</td>
</tr>
<tr>
<td>15</td>
<td>0.3 21.0 0.78 0.40</td>
<td>21.1 24.3 1.01 0.45</td>
<td>0.7 23.3 0.68 0.43</td>
</tr>
<tr>
<td>16</td>
<td>-0.2 21.4 0.70 0.39</td>
<td>-26.1 26.1 1.68 0.52</td>
<td>-24.3 25.1 1.56 0.42</td>
</tr>
<tr>
<td>17</td>
<td>4.4 51.9 0.64 0.62</td>
<td>-4.5 36.1 0.66 0.59</td>
<td>16.6 41.1 0.73 0.59</td>
</tr>
<tr>
<td>18</td>
<td>0.1 13.0 0.55 0.69</td>
<td>8.5 14.8 0.64 0.63</td>
<td>-5.3 16.8 0.95 0.33</td>
</tr>
<tr>
<td>19</td>
<td>-9.2 14.9 0.69 0.56</td>
<td>-4.3 16.1 0.65 0.50</td>
<td>-5.7 19.7 0.90 0.14</td>
</tr>
<tr>
<td>20</td>
<td>0.0 13.9 0.56 0.64</td>
<td>-3.7 11.0 0.78 0.43</td>
<td>-12.4 16.9 0.98 0.38</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.2 18.4 0.65 0.55</td>
<td>-1.6 18.8 0.84 0.49</td>
<td>-0.9 21.2 0.83 0.42</td>
</tr>
<tr>
<td>Stdev</td>
<td>2.7 10.8 0.10 0.13</td>
<td>10.4 7.8 0.28 0.10</td>
<td>11.4 8.6 0.20 0.11</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>3.8 15.5 0.75 0.47</td>
<td>1.8 16.9 0.83 0.38</td>
</tr>
<tr>
<td></td>
<td>Stdev</td>
<td>8.4 5.7 0.16 0.12</td>
<td>7.5 4.2 0.13 0.08</td>
</tr>
</tbody>
</table>

For cases meeting the criteria
Table 5. Modeling results of 2 week training, 1 week validation and 1 week testing data

<table>
<thead>
<tr>
<th>Subject</th>
<th>( \hat{y}_{t^*} )</th>
<th>Testing</th>
<th>( \hat{y}_t )</th>
<th>Meeting Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AE</td>
<td>AAE</td>
<td>RAAE</td>
<td>( r_{fit} )</td>
</tr>
<tr>
<td>1</td>
<td>1.7</td>
<td>16.2</td>
<td>0.72</td>
<td>0.47</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>11.6</td>
<td>0.56</td>
<td>0.66</td>
</tr>
<tr>
<td>3</td>
<td>-1.8</td>
<td>10.0</td>
<td>0.61</td>
<td>0.59</td>
</tr>
<tr>
<td>4</td>
<td>-0.5</td>
<td>8.5</td>
<td>0.70</td>
<td>0.45</td>
</tr>
<tr>
<td>5</td>
<td>6.7</td>
<td>17.6</td>
<td>0.68</td>
<td>0.41</td>
</tr>
<tr>
<td>6</td>
<td>-3.7</td>
<td>16.1</td>
<td>0.80</td>
<td>0.41</td>
</tr>
<tr>
<td>7</td>
<td>-2.3</td>
<td>22.0</td>
<td>0.70</td>
<td>0.43</td>
</tr>
<tr>
<td>8</td>
<td>2.5</td>
<td>18.4</td>
<td>0.67</td>
<td>0.45</td>
</tr>
<tr>
<td>9</td>
<td>0.7</td>
<td>14.0</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>10</td>
<td>-5.1</td>
<td>19.7</td>
<td>0.55</td>
<td>0.66</td>
</tr>
<tr>
<td>11</td>
<td>-2.6</td>
<td>35.7</td>
<td>0.65</td>
<td>0.51</td>
</tr>
<tr>
<td>12</td>
<td>0.8</td>
<td>12.3</td>
<td>0.85</td>
<td>0.10</td>
</tr>
<tr>
<td>13</td>
<td>9.1</td>
<td>17.0</td>
<td>0.90</td>
<td>0.54</td>
</tr>
<tr>
<td>14</td>
<td>-2.9</td>
<td>15.2</td>
<td>0.62</td>
<td>0.64</td>
</tr>
<tr>
<td>15</td>
<td>10.1</td>
<td>21.2</td>
<td>0.78</td>
<td>0.48</td>
</tr>
<tr>
<td>16</td>
<td>9.9</td>
<td>18.1</td>
<td>0.73</td>
<td>0.27</td>
</tr>
<tr>
<td>17</td>
<td>2.3</td>
<td>48.1</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>18</td>
<td>1.3</td>
<td>12.2</td>
<td>0.52</td>
<td>0.69</td>
</tr>
<tr>
<td>19</td>
<td>-1.0</td>
<td>18.7</td>
<td>0.81</td>
<td>0.28</td>
</tr>
<tr>
<td>20</td>
<td>6.5</td>
<td>13.2</td>
<td>0.65</td>
<td>0.59</td>
</tr>
<tr>
<td>Mean</td>
<td>1.6</td>
<td>18.3</td>
<td>0.69</td>
<td>0.50</td>
</tr>
<tr>
<td>Stdev</td>
<td>4.6</td>
<td>9.1</td>
<td>0.10</td>
<td>0.15</td>
</tr>
</tbody>
</table>

For cases meeting the criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Passing Rate</th>
<th>65%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.1 0.70 0.46</td>
<td>4.7 17.6 0.79 0.40</td>
</tr>
<tr>
<td>Stdev</td>
<td>4.6 7.7 0.11 0.17</td>
<td>9.4 5.9 0.14 0.15</td>
</tr>
</tbody>
</table>

Literature Cited


The purpose of this appendix is to provide a mathematical proof that $r_{fit}$, under the simple linear regression, i.e., Eq. 27 with one input. Let $\hat{\eta}_i = \hat{a}_0 + \hat{a}_i \hat{v}_{i,j}$, in this context, $r_{fit}$ is mathematically given by

$$r_{fit} = r_{y_i, \hat{y}} = \frac{\sum_{j=1}^{n} (y_j - \bar{y})(\hat{\eta}_j - \bar{\eta})}{\sqrt{\sum_{j=1}^{n} (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^{n} (\hat{\eta}_j - \bar{\eta})^2}} = \frac{\sum_{j=1}^{n} (y_j - \bar{y}) (\hat{a}_0 + \hat{a}_i \hat{v}_{i,j} - \hat{a}_0 - \hat{a}_i \bar{v}_j)}{\sqrt{\sum_{j=1}^{n} (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^{n} (\hat{a}_0 + \hat{a}_i \hat{v}_{i,j} - \hat{a}_0 - \hat{a}_i \bar{v}_j)^2}}$$

$$= \frac{\hat{a}_i \sum_{j=1}^{n} (v_j - \bar{y})(\hat{v}_{i,j} - \bar{v}_j)}{\sqrt{\hat{a}_i^2} \cdot \sqrt{\sum_{j=1}^{n} (v_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^{n} (\hat{v}_{i,j} - \bar{v}_j)^2}} = \frac{\hat{a}_i}{|\hat{a}_i|} r_{v_i, \hat{v}_{i,j}}$$

Thus, with $\hat{a}_i > 0$, $r_{fit} = r_{y_i, \hat{y}}$, and for $\hat{a}_i < 0$, $r_{fit} = -r_{y_i, \hat{y}}$. This result means that if the correlation of measured blood glucose concentration (BGC) and $\hat{v}_{i,j}$ is positive, $\hat{a}_i$ can be set at any positive value and $r_{fit}$, which will be $> 0$, will depend only of the behavior of $\hat{v}_{i,j}$ which is independently
controlled by the values of the dynamic parameters associated with \( v_{i,t} \). Conversely, if the correlation of BGC and \( \hat{v}_{i,t} \) is negative, \( \hat{a}_i \) can be set at any negative value and \( r_{fit} \) will be $> 0$ and independently controlled by the values of the dynamic parameters associated with \( v_{i,t} \).
CHAPTER 4: DYNAMIC MODELING WITH CORRELATED INPUTS: THEORY, METHOD AND EXPERIMENTAL DEMONSTRATION

A paper published in Industrial & Engineering Chemistry Research

Derrick K. Rollins, Sr.\textsuperscript{1,2*}, Amy Roggendorf\textsuperscript{1}, Yiying Khor\textsuperscript{1}, Yong Mei\textsuperscript{1}, Peggy Lee\textsuperscript{1} and Stephanie Loveland\textsuperscript{1}

\textsuperscript{1}Department of Chemical Engineering, 2114 Sweeney Hall, Iowa State University, Ames, IA 50011, United States;
\textsuperscript{2}Department of Statistics, 102 Snedecor Hall, Iowa State University

*Corresponding author

Abstract

When modeling dynamic processes for several inputs with freely existing data, such as data collected with normal process operations, the ability to accurately model the output response for a given input change is impeded when inputs are cross- (i.e., pair-wised) correlated as this adversely affects accurate estimation of the causative effects of inputs on the response variable. The causative effects of the inputs can be evaluated functionally and analytically via the Jacobian Matrix which is done in this work for NARMAX and Wiener structures that are linear and nonlinear in model parameters. This analysis shows that the Wiener structure with physically-based nonlinear parameterization is superior. This conclusion is also supported in this work by a modeling study on a real distillation column consisting of eight test runs over a period of three years.

Key Words

Wiener Modeling, NARMAX Modeling, ARMAX Modeling, Modeling Plant Data, Modeling Freely Existing Data
Introduction

In chemical processes, output variables are determined by the values of input variables, which may or may not be measured and some input variables may not even be known. Mathematical modeling is the process of mapping measured variables to output variables using a mathematical formulation. In mathematical modeling a structure must be selected and any unknown coefficients (i.e., parameters) must be estimated. For a general model structure, let its expectation be represented as $\eta_i = f(X_i; \theta)$, where $\eta_i$ is the expected value of the response (i.e., output) at the $i^{th}$ sampling time, $i = 1, \ldots, n$; $X_i$ is the vector of input values at the $i^{th}$ sampling time; and $\theta$ is the vector of unknown model parameters with $\theta = [\theta_1 \ldots \theta_q]^T$. Therefore, the element of its Jacobian Matrix, $J_{nxq}$, in the $i^{th}$ row and $j^{th}$ column is \[ J_{ij} = \frac{\partial \eta_i}{\partial \theta_j}, \] that is, \[ J = \begin{bmatrix} \frac{\partial \eta_1}{\partial \theta_1} \\ \vdots \\ \frac{\partial \eta_n}{\partial \theta_q} \end{bmatrix}. \]

Moreover, the $j^{th}$ column represents $\theta_j$ and its column vector represents the change in the response space as $\theta_j$ changes for the set of experimental conditions. If two columns, say $j$ and $k$, are orthogonal, their correlation coefficient is zero, and the information to estimate $\theta_j$ is decoupled (i.e., separate or independent) from the information to estimate $\theta_k$, and vice versa. The advantage of this is that causative relationships of inputs on the response can be obtained and standard estimation errors of parameters are minimum [1]. Correlated columns in the Jacobian Matrix arise from pairwise correlation of inputs. Thus, obtaining orthogonal columns for each parameter necessitates setting $X_i$ according to some predetermined statistical experimental design. However, running a statistically designed experiment is not always practical or possible. The ability to accurately model the output response for given inputs is impeded when inputs are cross-correlated [2,3]. Consequently, model identification often involves the use of experimental data with correlated inputs that result in columns of $J$ that are not orthogonal (i.e., correlated).
Notwithstanding, for a given set of experimental data, with pairwise correlated inputs, the best model (in the sense of evaluating causative relationship between inputs and response and extracting scientific knowledge from parameters) will be the one that produces columns of $J$ with the smallest pair-wise correlation. Therefore, the goal of this work is to evaluate dynamic model structures based on pair-wise correlation of the columns of $J$ when the inputs are pair-wise correlated. This scope is restricted to transfer function models that are linear in the time-dependent process variables. More specifically, this work compares structures that are developed from transfer functions that are applicable to Nonlinear Autoregressive Moving Average with eXogenous variables (NARMAX) and Wiener models.

The basic difference of NARMAX and Wiener transfer functions is the characteristic equation which is the same for each input for NARMAX but can be different for each input for Wiener. With discrete-time modeling, this work will evaluate linear and nonlinear regression transfer function models of NARMAX and Wiener networks. The engineering literature typically defines linear models based on the form of the time-dependent variables in the differential equations of transfer functions. However, this scope is parameter estimation with focus on $J$ which is based on the behavior of the parameters. Hence, this article adopts the statistical definition of linear and nonlinear models in parameter estimation. More specifically, a linear model in this work is one that is linear in parameters and a nonlinear model is one that is nonlinear in parameters [1]. All parameters in this work are estimated with the least squares criterion.

From a search in the process identification literature, we discovered that for NARMAX models and its subclasses, linear forms in parameters are widely, if not exclusively, used (see the following articles for examples: Baldacchino, Anderson & Kadirkamanathan, 2013 [10]; Chiuso
It also revealed that the common practice of Wiener Modeling is the use of transfer functions that are linear in the estimated parameters (see the following articles for examples: Giri & Bai, 2010 [6]; Hagenblad, Ljung & Wills, 2008 [7]; Norquay, Palazoglu & Romagnoli, 1998 [8]; Pearson & Pottmann, 2000 [9]). The only discrete-time Wiener approach that we found to use nonlinear transfer functions is the one developed by Rollins et al. [5] The nonlinear parameterized NARMAX structure will be developed in this work following this approach.

This article is organized as follows. The Method and Theory Section will give the details linear and nonlinear parameterized discrete-time NARMAX and Wiener structures to be evaluated. This section will also compare and evaluate these structures based on $J$. The next section will evaluate these structures in a study on a pilot distillation column with several unmeasured disturbances. Models with nine inputs for both model structures are developed from one run of training data and one run of validation data. These models are evaluated on eight test cases (i.e., runs of the column) spanning a period of about three years. The last section gives concluding remarks and comments on future work.

**Methods and Theory**

The main purpose of this section is to give the $J$ analysis for each model. These models will need to be derived first. Since the Wiener Model (WM) is most general, and since the NARMAX (NM) can be a subclass, the WM is derived first. After the models are given, a $J$ analysis will be given for each one.

**The Wiener Structures**

Figure 1 is a block diagram with unity gain linear dynamic blocks (i.e., linear differential equation in the time dependent variables), $G_i$, and a static gain block $f(V)$, where $x_i$ is the input to
transfer function $G_i$, $v_i$ is its output, vector $V = [v_1, v_2, \ldots, v_p]^T$, $f(V)$ is an unrestricted function, and $i = 1, \ldots, p$. Therefore, Fig. 1 is representative of the Wiener network and as we will show later in this section, it is also representative of a NARMAX network with the restrictions that the $G_i$'s have unity gain and the same characteristic equations. This section derives discrete-time linear and nonlinear regression transfer function model structures for the $G_i$'s.

![Block diagram for a general transfer function network](image)

**Fig. 1.** Block diagram for a general transfer function network with $p$ inputs and one output. Each input, $x_i$, is passed through their own unity gain linear dynamic block, $G_i$, after which these unobservable intermediate outputs are collected and passed through a single unrestricted static gain function, $f(V)$, to produce the output, $y$. This is a NARMAX network when the $G_i$'s have the same CEs and a Wiener network when they can have different CEs.

Rollins et al. [5] presented a discrete-time WM structure for $G_i$ with a second-order-plus-lead-plus-dead time (SOPLPDT) form that estimates the physically-based dynamic parameters of the $v_i$'s directly with a highly non-linear structure including physical constraints. The form of this differential equation (dead-time is excluded for simplicity) is

$$
\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i \zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t)}{dt} + x_i(t)
$$

(1)

where $i = 1, \ldots, p$, $p$ is the total number of inputs ($x_i$'s), $\tau_i$ is the time constant, $\zeta_i$ is the damping coefficient, and $\tau_{ai}$ is the lead parameter. Note that a zero dead time assumption is not made; dead
time is not shown in the Eq. 1 only to simplify the written form. With a parameter for dead time included, the inputs in the equations would shift in time by the amount of dead time [5]. Using a backward difference approximation \( \frac{dv_i(t)}{dt} = \frac{v_{i,t} - v_{i,t-\Delta t}}{\Delta t} \) applied to a sampling interval of \( \Delta t \), one will obtained the following approximate discrete-time form of Eq. 1 [5]:

\[
v_{i,t} = \delta_{i,1} v_{i,t-\Delta t} + \delta_{i,2} v_{i,t-2\Delta t} + \omega_{i,1} x_{i,t-\Delta t} + \omega_{i,2} x_{i,t-2\Delta t}
\]

(2)

where

\[
\delta_{i,1} = \frac{2 \tau_i^2 + 2 \tau_i \zeta_i \Delta t}{\tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2}
\]

(3)

\[
\delta_{i,2} = \frac{-\tau_i^2}{\tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2}
\]

(4)

\[
\omega_{i,1} = \frac{(\tau_{ai} + \Delta t) \Delta t}{\tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2}
\]

(5)

and

\[
\omega_{i,2} = 1 - \delta_{i,1} - \delta_{i,2} - \omega_{i,1}
\]

(6)

to satisfy the constraint of unity gain, and \( x_{i,t} \) is the value of the \( i \)th input at \( t \). Two additional physical constraints are \( \tau_i > 0 \) and \( \zeta_i > 0, \forall i \). Note that Eq. 2 is linear in the \( \delta \)'s and \( \omega \)'s. Thus, Eq. 2 represents the linear regression form of the WM with the \( \delta \)'s and \( \omega \)'s estimated directly. Notwithstanding, the nonlinear regression form of the WM is obtained by the substitution of Eqs. 3-6 into Eq. 2 and estimating the \( \tau \)'s, \( \zeta \)'s, and \( \tau_a \)'s (i.e., the physically-based dynamic parameters) directly with \( v_{i,t} \) given as a highly nonlinear function of the estimated parameters. Physical constraints also strengthen estimation by adding more true structure. Notably, these are: Eq. 6, \( \tau_i > 0 \) and \( \zeta_i > 0, \forall i \). Another advantage from using physically-based dynamic parameters is the ability to give sensible starting values for estimates.

In general Eq. 2 can be given as
\[ v_{i,t} = \delta_{i,1} v_{i,i-\Delta t} + \ldots + \delta_{i,r} v_{i,i-r\Delta t} + \omega_{i,1} x_{i,i-\Delta t} + \ldots + \omega_{i,s} x_{i,i-s\Delta t} \]  

(7)

From Eq. 7, \( G_{i,t} \) is obtained as follows:

\[ v_{i,t} - \delta_{i,1} v_{i,i-\Delta t} - \ldots - \delta_{i,r} v_{i,i-r\Delta t} = \omega_{i,1} x_{i,i-\Delta t} + \ldots + \omega_{i,s} x_{i,i-s\Delta t} \]

\[ \left(1 - \delta_{i,1} B - \ldots - \delta_{i,r} B^r\right)v_{i,t} = \left(\omega_{i,1} B + \ldots + \omega_{i,s} B^s\right)x_{i,t} \]

\[ \Rightarrow G_{i,t} = \frac{v_{i,t}}{x_{i,t}} = \frac{\omega_{i,1} B + \ldots + \omega_{i,s} B^s}{1 - \delta_{i,1} B - \ldots - \delta_{i,r} B^r} = \frac{\omega(B)}{\delta(B)} \]  

(8)

(9)

where \( s \) and \( r \) are the orders in the numerator and denominator, respectively, and \( B \) denotes the backwards shift operator (i.e., \( B^n x_t = x_{t-m\Delta t} \), where \( \Delta t \) is the sampling time). Note that in Eq. 2, \( r = s = 2 \). From Eq. 9,

\[ v_{i,t} = G_{i,t} x_{i,t} = \frac{\omega(B)}{\delta(B)} x_{i,t} \]  

(10)

The general discrete-time, “white noise,” WM is given as

\[ y_t = f(V) + \varepsilon_t = \eta_t + \varepsilon_t \]  

(11)

where

\[ \varepsilon_t \sim N(0, \sigma^2) \] independently \( \forall t \)  

(12)

the common assumption with least squares estimation and \( y_t \) is the measured output at time \( t \).

Note that the expected response at time \( t \) is \( \eta_t = E[y_t] \), since \( E[\varepsilon_t] = 0 \). While \( f(V) \) can be any function, in the work, for simplicity, it is given as

\[ \eta_t = f(V) = a_1 v_{1,t} + \ldots + a_p v_{p,t} \]  

(13)

where \( a_i \) = the steady state gain associated with input \( x_{i,t} \). Substituting Eqs. 2 and 6 into Eq. 13 gives

\[ \eta_t = a_1 \delta_{i,1} v_{1,i-\Delta t} + a_1 \delta_{i,2} v_{1,i-2\Delta t} + a_1 \omega_{i,1} x_{i,i-\Delta t} + a_1 \left(1 - \delta_{i,1} - \delta_{i,2} - \omega_{i,1}\right)x_{i,i-2\Delta t} + \ldots \]

\[ + a_p \delta_{p,1} v_{p,i-\Delta t} + a_p \delta_{p,2} v_{p,i-2\Delta t} + a_p \omega_{p,1} x_{p,i-\Delta t} + a_p \left(1 - \delta_{p,1} - \delta_{p,2} - \omega_{p,1}\right)x_{p,i-2\Delta t} \]  

(14)
Eq. 14 will be the form that we will use in the \( \mathbf{J} \) analysis.

The estimator for \( \eta \), is found using Eq. 11 as follows:

\[
\varepsilon_t = y_t - \eta_t = 0
\]  
(15)

Therefore, the optimal estimator of \( \eta \) with “white noise” is given as

\[
\hat{y}_t = \hat{\eta}_t = \hat{\alpha}_1 \hat{v}_{1,t} + \ldots + \hat{\alpha}_p \hat{v}_{p,t}
\]  
(16)

where the symbol “\(^\wedge\)” denotes estimator.

**The NARMAX Structures**

With “linear” defined as a linear dynamic system, i.e., one of linear differential equations, in Eq. 16.6 in this book [4], Nelles defines “the general linear model” as

\[
A(q)y(k) = \frac{B(q)}{F(q)} u(k) + \frac{C(q)}{D(q)} v(k)
\]  
(17)

Eq. 17 is the basic structure containing the transfer function form for the NARMAX family of models (i.e., NARMAX and all of its subclasses) as shown in Nelles [4] in Chapter 16 for Autoregressive Moving Average with eXogenous variables (ARMAX) models and its subclasses and in Chapter 17 for NARMAX in general. Rewriting Eq. 17 in our preferred discrete-time notation with Eq. 13 for \( \eta \), gives, for one input:

\[
\delta(B)\eta = \omega^*(B) x_t
\]  
(18)

since \( E[\varepsilon_t] = E[v(k)] = 0 \), where

\[
\delta(B) = \left(1 - \delta_1 B - \ldots - \delta_p B^p\right)
\]  
(19)

and

\[
\omega^*(B) = a\left(\omega_1 B + \ldots + \omega_p B^p\right) = \omega_1 * B + \ldots + \omega_p * B^p
\]  
(20)

More specifically, Eq. 18 is the expected response for the family of ARMAX models. For \( p \) inputs, Eq. 18 becomes
\[ \delta(B)\eta_i = \omega_i^t * (B)x_{i,t} + \ldots + \omega_p * (B)x_{p,t} \]  
where \[ \omega_i^t(B) = a_i^t (B) B + \ldots + a_s^t (B) B^s \]  
A general NARMAX equation for the expected response is given as

\[ \delta(B)\eta_i = f^t * (x_{i,-\Delta t}, \ldots, x_{i,-s\Delta t}, \ldots, x_{p,-\Delta t}, \ldots, x_{p,-s\Delta t}) = f^t (X) \]  
where \( f^t(X) \) is an unrestricted function of all inputs from \( t - \Delta t \) to \( t - s\Delta t \). The expected response for ARMAX is a special case for NARMAX with Eq. 13 with

\[ v_{i,t} = G_{i,t} x_{i,t} = \frac{\omega_i(B)}{\delta(B)} x_{i,t} \]  
Note that the denominator in Eq. 24 is the same for each transfer function meaning that

NARMAX models use the same characteristic equation for each one, which is the primary difference from WM. With Eq. 24, at \( t = n\Delta t \), Eq. 14 becomes

\[ \eta_{n\Delta t} = a_i \delta_i v_{1,(n-1)\Delta t} + a_i \delta_2 v_{1,(n-2)\Delta t} + a_i a_{i,1} x_{1,(n-1)\Delta t} + a_i \left( 1 - \delta_i - \delta_2 - a_{i,1} \right) x_{1,(n-2)\Delta t} + \ldots \]
\[ + a_p \delta_i v_{p,(n-1)\Delta t} + a_p \delta_2 v_{p,(n-2)\Delta t} + a_p a_{p,1} x_{p,(n-1)\Delta t} + a_p \left( 1 - \delta_i - \delta_2 - a_{p,1} \right) x_{p,(n-2)\Delta t} + \ldots \]
\[ = \delta_i \eta_{(n-1)\Delta t} + \delta_2 \eta_{(n-2)\Delta t} + a_{i,1} a_i x_{1,(n-1)\Delta t} + a_i \left( 1 - \delta_i - \delta_2 - a_{i,1} \right) x_{1,(n-2)\Delta t} + \ldots \]
\[ + a_{p,1} a_p x_{p,(n-1)\Delta t} + a_p \left( 1 - \delta_i - \delta_2 - a_{p,1} \right) x_{p,(n-2)\Delta t} \]
\[ = (1 - \delta(B)) \eta_i + f^t * (x_{1,(n-1)\Delta t}, x_{1,(n-2)\Delta t}, \ldots, x_{p,(n-1)\Delta t}, x_{p,(n-2)\Delta t}) = (1 - \delta(B)) \eta_i + f^t (X) \]  
Note that, the form of Eq. 25 is consistent with that of Eq. 23. This is the linear parameterized NARMAX form that will be used in the J Analysis. The nonlinear form that will be used in the J consists in the substitution of Eqs. 26-28, given below, into Eq. 25.

\[ \delta_i = \frac{\tau_i^2 + 2\tau_i \zeta \Delta t}{\tau_i^2 + 2\tau_i \zeta \Delta t + \Delta t^2} \]  
\[ \delta_2 = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i \zeta \Delta t + \Delta t^2} \]  
\[ \omega_{i,1} = \frac{\tau_{i,1} + \Delta t}{\tau_i^2 + 2\tau_i \zeta \Delta t + \Delta t^2} \]
The J Analysis

Using the general notation from the Introduction Section, $J$ with $t = k\Delta t$ to $n\Delta t$, is represented by Eq. 29 below.

$$
\begin{bmatrix}
\frac{\partial \eta_{k\Delta t}}{\partial \theta_1} & \frac{\partial \eta_{k\Delta t}}{\partial \theta_2} & \cdots & \frac{\partial \eta_{k\Delta t}}{\partial \theta_q} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial \eta_{n\Delta t}}{\partial \theta_1} & \frac{\partial \eta_{n\Delta t}}{\partial \theta_2} & \cdots & \frac{\partial \eta_{n\Delta t}}{\partial \theta_q}
\end{bmatrix}
$$

(29)

Although, technically, we will use the ARMAX expected response function in the analysis in this subsection, the analysis applies in general to NARMAX. With $r = s = 2$, the unconstrained linear parameterized ARMAX expected response represented in Eq. 23, from $t = 3\Delta t$ to $n\Delta t$, becomes

$$
\eta_{k\Delta t} = \delta_1 \eta_{(k-1)\Delta t} + \delta_2 \eta_{(k-2)\Delta t} + \omega_{1,1} x_{1,(k-1)\Delta t} + \omega_{1,2} x_{1,(k-2)\Delta t} + \omega_{p,1} x_{p,(k-1)\Delta t} + \omega_{p,2} x_{p,(k-2)\Delta t}
$$

(30)

Thus, $J$ for Eq. 30, from $t = 3\Delta t$ to $n\Delta t$ is

$$
\begin{bmatrix}
\frac{\partial \eta_{2\Delta t}}{\partial \delta_1} & \frac{\partial \eta_{2\Delta t}}{\partial \delta_2} & \frac{\partial \eta_{2\Delta t}}{\partial \omega_{1,1}^*} & \frac{\partial \eta_{2\Delta t}}{\partial \omega_{1,2}^*} & \cdots & \frac{\partial \eta_{2\Delta t}}{\partial \omega_{p,1}^*} & \frac{\partial \eta_{2\Delta t}}{\partial \omega_{p,2}^*} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial \eta_{n\Delta t}}{\partial \delta_1} & \frac{\partial \eta_{n\Delta t}}{\partial \delta_2} & \frac{\partial \eta_{n\Delta t}}{\partial \omega_{1,1}^*} & \frac{\partial \eta_{n\Delta t}}{\partial \omega_{1,2}^*} & \cdots & \frac{\partial \eta_{n\Delta t}}{\partial \omega_{p,1}^*} & \frac{\partial \eta_{n\Delta t}}{\partial \omega_{p,2}^*}
\end{bmatrix}
$$

(31)

As Eq. 31 shows, the first two columns representing $\delta_1$ and $\delta_2$ are clearly highly correlated and the correlation of the inputs translate directly into the columns representing the $\omega^*$s. Thus, when the inputs are significantly correlated, an unconstrained linear parameterized NARMAX model (NM) would not likely fit the data acceptably due to the inflation of the standard estimation errors of the parameters.
With \( r = s = 2 \), the constrained linear parameterized ARMAX expected response is given by Eq. 25. From Eq. 25, for this case, the columns of \( J \) will come from Eqs. 33-36 below.

\[
\frac{\partial \eta_i}{\partial \delta_1} = \eta_{i-2\Delta t} - a_1x_{i-2\Delta t} - \cdots - a_px_{p,i-2\Delta t} \quad (33)
\]

\[
\frac{\partial \eta_i}{\partial \delta_2} = \eta_{i-2\Delta t} - a_1x_{i-2\Delta t} - \cdots - a_px_{p,i-2\Delta t} \quad (34)
\]

\[
\frac{\partial \eta_i}{\partial \omega_i} = a_1x_{i-2\Delta t} - a_1x_{i,i-2\Delta t} \quad \forall i \quad (35)
\]

\[
\frac{\partial \eta_i}{\partial a_i} = \omega_i x_{i-2\Delta t} + \left(1 - \delta_1 - \delta_2 - \omega_i\right)x_{i,i-2\Delta t} \quad \forall i \quad (36)
\]

As shown by Eqs. 33 and 34, the first two columns representing \( \delta_1 \) and \( \delta_2 \), respectively, are clearly highly correlated as they only differ by one time step in \( \eta \). As Eqs. 35 and 36 show, the correlation of the inputs translates directly into the columns representing the \( \omega_i \)'s and \( a_i \)'s. Thus, when the inputs are significantly correlated, a constrained linear parameterized NM would also not likely fit the data acceptably.

With \( r = s = 2 \), the constrained linear parameterized WM expected response is given by Eq. 14 and the columns of \( J \) will come from Eqs. 37-40 below.

\[
\frac{\partial \eta_i}{\partial \delta_{1,i}} = a_{i,i}v_{i,j-2\Delta t} - a_{i,i}x_{i,i-2\Delta t} \quad \forall i \quad (37)
\]

\[
\frac{\partial \eta_i}{\partial \delta_{2,i}} = a_{i,i}v_{i,j-2\Delta t} - a_{i,i}x_{i,i-2\Delta t} \quad \forall i \quad (38)
\]

\[
\frac{\partial \eta_i}{\partial \omega_{i,i}} = a_{i,i}x_{i,j-2\Delta t} - a_{i,i}x_{i,i-2\Delta t} \quad \forall i \quad (39)
\]

\[
\frac{\partial \eta_i}{\partial a_i} = v_{i,i} = \delta_{1,i}v_{i,j-2\Delta t} + \delta_{2,i}v_{i,j-2\Delta t} + \omega_{i,i}x_{i,i-2\Delta t} + \left(1 - \delta_1 - \delta_2 - \omega_i\right)x_{i,i-2\Delta t} \quad \forall i \quad (40)
\]

As shown by Eqs. 37-39, the columns representing the linear dynamic parameters will be strongly correlated for reasons mentioned previously. However, the correlation of columns
representing the static parameters, i.e., the \( a_i \)'s, will depend on the cross correlation of the outputs from the transfer function, i.e., the \( v_i \)'s. While the cross correlation of the \( x_i \)'s will influence the cross correlation of the \( v_i \)'s, it can be significantly different, depending on the dynamic behavior of the transfer functions. Nonetheless, when the inputs are significantly correlated, the linear parameterized WM suffers from the limitations as all linear structures in the dynamic parameters and would also not likely fit the data acceptably.

The nonlinear parameterized NM in this work is based on the Eqs. 26-28 for the dynamic parameters. With Eq. 13, the columns of \( J \), would come from Eqs. 41-42 below:

\[
\frac{\partial \eta_i}{\partial \theta_j} = a_1 \frac{\partial v_{i,t}}{\partial \theta_j} + \cdots + a_p \frac{\partial v_{p,t}}{\partial \theta_j}; \quad \theta_i = \tau, \zeta, \text{ or } \tau_{ai} \tag{41}
\]

\[
\frac{\partial \eta_i}{\partial a_j} = v_{i,t} = \delta_{i} v_{i,j-2\Delta t} + \omega_{i} x_{i,j-2\Delta t} + \left(1 - \delta_{i} - \omega_{i}\right) x_{i,j-2\Delta t} + \left(1 - \delta_{i} - \omega_{i}\right) x_{i,j-2\Delta t}

\]

\[
\frac{\tau^2 + 2\tau \zeta \Delta t}{\tau^2 + 2\tau \zeta \Delta t + \Delta t^2} v_{i,j-2\Delta t} + \frac{-\tau^2}{\tau^2 + 2\tau \zeta \Delta t + \Delta t^2} v_{i,j-2\Delta t} + \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau^2 + 2\tau \zeta \Delta t + \Delta t^2} x_{i,j-2\Delta t} \tag{42}
\]

For \( \tau_i \) and \( \zeta_i \), respectively, Rollins et al. [5] derived the following equations (the equations for \( \tau_{ai} \) were not given but would be similar to the ones below):

\[
\frac{\partial v_{i,t}}{\partial \tau_i} = \frac{2\Delta t \left( \frac{\tau^2 \zeta_i + 2\tau_i \zeta \Delta t + \zeta \Delta t^2}{\tau^2 + 2\tau_i \zeta \Delta t + \Delta t^2} \right) v_{i,j-2\Delta t} + \left( \frac{2\tau^2 + 2\tau \zeta \Delta t}{\tau^2 + 2\tau \zeta \Delta t + \Delta t^2} \right) \frac{\partial v_{i,j-2\Delta t}}{\partial \tau_i}}{\tau^2 + 2\tau \zeta \Delta t + \Delta t^2} \tag{43}
\]

For \( \zeta_i \), respectively, Rollins et al. [5] derived the following equations (the equations for \( \tau_{ai} \) were not given but would be similar to the ones below):
\[
\frac{\partial v_{i,j}}{\partial \zeta_i} = \left( \frac{2 \tau_i \Delta t (\Delta t^2 - \tau_i^2)}{\left( \tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2 \right)^2} \right) v_{i,j-\Delta t} + \left( \frac{2 \tau_i^2 + 2 \tau_i \zeta_i \Delta t}{\left( \tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2 \right)^2} \right) \frac{\partial v_{i,j-\Delta t}}{\partial \zeta_i} \\
- \left( \frac{2 \tau_i^3 \Delta t}{\left( \tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2 \right)^2} \right) v_{i,j-2\Delta t} \frac{\partial v_{i,j-2\Delta t}}{\partial \zeta_i}
\]
(44)

With \( \tau_i = \tau \) and \( \zeta_i = \zeta \), for \( i = 1, \ldots, p \) (i.e., the characteristic equation (CE) is the same for each transfer function) substituting Eqs. 43-44 into Eq. 41 indicates that the columns of \( J \) for the dynamic parameters are highly nonlinear and thus, are not likely to be significantly correlated.

Eq. 42 indicates that the correlation of the columns for the static parameters will depend on the cross correlation of the \( v_i \)’s as in the linear parameterized NM structure with unity gain transfer functions. However, since the CEs are the same, the pairwise correlation of the \( v_i \)’s can be greater than that of the \( x_i \)’s. This is seen mathematically from the similarity for different \( i \) in Eqs. 43 and 44. The only difference is in \( \tau_{ai} \). Consequently, although the nonlinear parameterized structure appears to be better than the linear structure for the dynamic parameters, this strength could be offset by greater pairwise correlation in the \( v_i \)’s than in the \( x_i \)’s, resulting in large standard estimation errors for the static parameters.

For the nonlinear parameterized WM, the CEs can all be different. Thus, Eq. 41 is
\[
\frac{\partial \eta_i}{\partial \theta_i} = a_i \frac{\partial v_{i,j}}{\partial \theta_i} + \cdots + a_p \frac{\partial v_{p,j}}{\partial \theta_i}, \quad \theta_i = \tau_i, \zeta_i, \text{ or } \tau_{ai}
\]
(45)

This is the critical difference that makes the WM structure superior to the NM structure. More specifically, this strength results in less pairwise correlation of the \( v_i \)’s and thus more accurate static parameters. In summary, this \( J \) analysis supports the nonlinear parameterized Wiener structure as the most superior (i.e., less pairwise correlation of the columns of \( J \)). In modeling
data, this superiority can be evaluated by comparing the pairwise correlation matrix of the \( v_i \)'s, which should be significantly better for the WM when the pairwise correlation of the \( x_i \)'s is large. In the next section, this evaluation is made in modeling data from a distillation column.

**Modeling Study**

In this section the nonlinear parameterized NM and WM structures formulated in the previous section will be compared and evaluated. The structure of these models is exactly the same except for the CE which is the same for the NM and allowed to be different for the WM. The hypothesis is that the correlation matrix of the \( v_i \)'s will be significantly better (i.e. less pairwise correlation) for the WM with correlated input variables and this will result in the best fit as supported by the \( J \) analysis in the previous section.

Since this article is focusing on modeling inputs, the models in the study are developed using inputs only. Thus, \( k \)-steps-ahead-modeling is outside the scope of this work as the use of outputs can dominate the fit to the degree that the contribution of the inputs to the model accuracy can be hidden or not directly obtainable. Two applications where models must be built using inputs only are feedforward control (FFC) and virtual sensor development. In FFC the model uses input measurements at the current time to determine the value of the manipulated variable needed to compensate for these input changes. This type of model will be developed and evaluated in Part 1 of this study. More specifically, in Part 1 the models will be developed from data where the output and inputs are available at each sampling instant. In Part 2, models will be developed as if for a virtual sensor application. More specifically, these models will be developed from limited output data to mimic the situation where the output is sampled infrequently and irregularly and the model is built off line and then used to provide frequent, constant sampled online virtual measurements.
The output in this study is the top tray temperature of a pilot distillation column that separates methanol and water. The details of this process will be given first. Next, the models will be developed from data that are split into a Training Set and Validation Set. Results are collected on eight (8) test sets or runs over a period of three years.

The Process

The distillation process is a pilot-scale methanol/water distillation column consisting of 12 trays, with an inside diameter of 6 inches. Feed was introduced at Tray 4 and had a concentration of 15% (mole) methanol. A process instrumentation diagram of the column is shown in Fig. 2 (Loveland and Rosa, 2005 [15]). The column is connected to an industrial type distributed control system. The nine (9) input variables were feed flow rate, feed temperature set point, reboiler level, reboiler steam pressure, reflux flow rate, column pressure, bottoms product flow rate, distillate product flow rate and overhead condensate temperature. The output response of the process was the top tray (Tray 12) temperature. These variables were sampled at a rate, $\Delta t = 1/12$ minute (m).

Results – Part 1

The Training and Validation Sets consisted of one run of the distillation column that was nearly evenly split into the first 121 minutes and the last 122 minutes of the run as Training and Validation data, respectively. We did attempt to build a linear parameterized NM using MATLAB and Excel. Not surprising, based on the J analysis in the previous section, the fits in both cases were very poor and essentially showed no fitted correlation to the test data sets. The specific results are not worth reporting and thus, are not included in the tables to conserve space. After building the NMs and the WMs they were tested on the eight (8) test sets. All the model
results for Part 1 are given in Table 1 for three statistics. The first one, the average error (AE), is defined as

$$AE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)$$ (46)

AE is an estimate of systematic model bias. While it is informative to obtain this estimate in Part 1, model bias is an irrelevant metric in the evaluation of a FFC model since the purpose of a FFC model is to estimate the amount of change needed in the manipulated variable to offset changes in inputs and not to estimate output response. The average absolute error (AAE) is defined as

$$AAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$ (47)

While this statistic gives a measure of closeness to the measured output, it is also affected by model bias and thus, is also not relevant to the evaluation of a FFC model. Both AE and AAE have units of degrees Celsius ($^\circ$C). The last statistic in Table 1, the correlation of $y_i$ and $\hat{y}_i$, $r_{fit}$, is the premier measure of performance for a model meant for a FFC application.
Fig. 2. Process instrumentation diagram of the distillation process.
### Table 1. Part 1 results for NM and WM: Training from 100% of the output data**.

<table>
<thead>
<tr>
<th>Case</th>
<th>Date</th>
<th>Type</th>
<th>Duration (Min)</th>
<th>WM</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AE</td>
<td>AAE</td>
</tr>
<tr>
<td>A</td>
<td>03/08/08</td>
<td>Training</td>
<td>123</td>
<td>0.00</td>
<td>0.09</td>
</tr>
<tr>
<td>B</td>
<td>03/08/08</td>
<td>Validate</td>
<td>122</td>
<td>-0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>C</td>
<td>03/22/08</td>
<td>Testing</td>
<td>568</td>
<td>0.11</td>
<td>0.37</td>
</tr>
<tr>
<td>D</td>
<td>04/13/08</td>
<td>Testing</td>
<td>391</td>
<td>0.17</td>
<td>0.24</td>
</tr>
<tr>
<td>E</td>
<td>06/05/08</td>
<td>Testing</td>
<td>230</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>F</td>
<td>09/10/08</td>
<td>Testing</td>
<td>142</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>G</td>
<td>09/24/08</td>
<td>Testing</td>
<td>245</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>H</td>
<td>05/22/09</td>
<td>Testing</td>
<td>205</td>
<td>0.31</td>
<td>0.32</td>
</tr>
<tr>
<td>I</td>
<td>02/27/11</td>
<td>Testing</td>
<td>130</td>
<td>-0.05</td>
<td>0.18</td>
</tr>
<tr>
<td>J</td>
<td>03/24/11</td>
<td>Testing</td>
<td>105</td>
<td>-0.14</td>
<td>0.19</td>
</tr>
</tbody>
</table>

| Testing Absolute Mean | 0.23 | 0.31 | 0.84 | 2.41 | 2.47 | 0.28 |
| Testing Absolute Stdev | 0.21 | 0.16 | 0.10 | 1.86 | 1.82 | 0.30 |

**Note: AE and AAE results are in °C.

Before we discuss the results in Table 1, it is informative to compare the correlation matrix of the inputs with the correlation matrix of the \( v_{i,t} \)'s that will be called the “dynamic inputs” since they are the dynamic counterparts of the \( x_{i,t} \)'s. The training data correlation matrices for the inputs, the NM dynamic inputs, and the WM dynamic inputs are given in Tables 2-4, respectively. Results with absolute values of 0.5 or greater are in bold and red text. As shown in Table 2, the absolute value of three results are much greater than 0.5. For these three pairs, as shown in Tables 3 and 4, both methods have similar significantly smaller results. However, both methods have higher and similar correlation for variables 2 and 3, and higher (0.79 for NM and 0.63 for WM) but not similar correlation for variables 3 and 8. While this difference is notable, the important difference in these results for the two methods is the much greater numbers for NM greater than 0.5 in absolute value. More specifically, while the inputs (Table 2) and the WM (Table 4) only had the same five pairs, the NM had these same five pairs and 10 more pairs (Table 3), which is 3 times higher. Since the only difference between NM and
WM is the same CE for NM and different ones for WM, it appears that use of the same CE is causing greater cross-correlation in the dynamic inputs which is consistent with the J analysis. We will now examine the modeling results to see how this difference is impacting the fit of the models.

Table 2. Training Set correlation matrix for the input variables

<table>
<thead>
<tr>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>x5</th>
<th>x6</th>
<th>x7</th>
<th>x8</th>
<th>x9</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>1</td>
<td>0.01</td>
<td>0.14</td>
<td>-0.14</td>
<td>0.13</td>
<td>-0.03</td>
<td>0.15</td>
<td>-0.11</td>
</tr>
<tr>
<td>x2</td>
<td>1</td>
<td>0.65</td>
<td>0.01</td>
<td>-0.09</td>
<td>0.04</td>
<td>0.02</td>
<td>0.36</td>
<td>-0.18</td>
</tr>
<tr>
<td>x3</td>
<td>1</td>
<td>0.01</td>
<td>0.07</td>
<td>0.08</td>
<td>0.01</td>
<td>0.56</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>x4</td>
<td>1</td>
<td>-0.16</td>
<td>0.93</td>
<td>0.91</td>
<td>0.11</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x5</td>
<td>1</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.10</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x6</td>
<td>1</td>
<td>0.88</td>
<td>0.11</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x7</td>
<td>1</td>
<td>0.10</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x8</td>
<td>1</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Training Set correlation matrix for the NM dynamic input variables

<table>
<thead>
<tr>
<th>v1</th>
<th>v2</th>
<th>v3</th>
<th>v4</th>
<th>v5</th>
<th>v6</th>
<th>v7</th>
<th>v8</th>
<th>v9</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1</td>
<td>1</td>
<td>0.52</td>
<td>0.59</td>
<td>0.21</td>
<td>0.51</td>
<td>-0.07</td>
<td>-0.02</td>
<td>0.72</td>
</tr>
<tr>
<td>v2</td>
<td>1</td>
<td>0.80</td>
<td>0.12</td>
<td>0.30</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.85</td>
<td>0.70</td>
</tr>
<tr>
<td>v3</td>
<td>1</td>
<td>0.16</td>
<td>0.44</td>
<td>0.05</td>
<td>-0.03</td>
<td>0.79</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>v4</td>
<td>1</td>
<td>0.20</td>
<td>0.81</td>
<td>0.85</td>
<td>0.22</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v5</td>
<td>1</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.60</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v6</td>
<td>1</td>
<td>0.78</td>
<td>0.03</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v7</td>
<td>1</td>
<td>-0.01</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v8</td>
<td>1</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Training Set correlation matrix for the WM dynamic input variables

<table>
<thead>
<tr>
<th></th>
<th>v1</th>
<th>v2</th>
<th>v3</th>
<th>v4</th>
<th>v5</th>
<th>v6</th>
<th>v7</th>
<th>v8</th>
<th>v9</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1</td>
<td>1</td>
<td>0.08</td>
<td>0.09</td>
<td>0.24</td>
<td>0.15</td>
<td>0.11</td>
<td>-0.03</td>
<td>0.08</td>
<td>0.36</td>
</tr>
<tr>
<td>v2</td>
<td>1</td>
<td>0.79</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.47</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>v3</td>
<td>1</td>
<td>0.02</td>
<td>0.22</td>
<td>0.03</td>
<td>0.02</td>
<td>0.63</td>
<td>-0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v4</td>
<td>1</td>
<td>0.14</td>
<td>-0.86</td>
<td>-0.73</td>
<td>0.23</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v5</td>
<td>1</td>
<td>0.04</td>
<td>-0.01</td>
<td>0.27</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v6</td>
<td>1</td>
<td>0.80</td>
<td>-0.21</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v7</td>
<td>1</td>
<td>-0.33</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v8</td>
<td>1</td>
<td>-0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 1, the training and validation results for both methods are excellent with NM ($r_{fit}$ equal to 0.92 and 0.92, respectively) being only slightly worse than the WM ($r_{fit}$ equal to 0.96 and 0.97, respectively). However, the testing results for the NM are considerably worse with an average $r_{fit}$ of 0.28 versus 0.84 for the WM. Thus, it appears that using the same CE is causing greater pairwise correlation for the dynamic inputs that are translating into an adverse effect on causative modeling, resulting in a significantly worse fit. Therefore, with conditions of cross-correlation of the inputs, the nonlinear parameterized WM should be selected over the nonlinear parameterized NM and any linear parameterized model. A representative case (Case H) is given in Fig. 3 where measured data and the fits for both models are plotted. As shown, the WM fits (i.e., correlates with) the data considerably better than the NM.
Fig. 3. The response plot for Case H in Table 1 for measured, NM and WM fitted output values.

Results – Part 2

Using uniformly distributed random sampling, we took ten percent of the training and validation output data, each, and obtained a fit for the WM. This case is made to mimic an infrequent sampling situation where the sampling rate is not constant and values for a variable are needed at some specified higher sampling rate that can be provided by modeling a set of measured variables. These results are given in Table 5 along with the corresponding results from Table 1. As shown, AAE is about the same with 0.32 °C and 0.31 °C for 10% and 100% sampling, respectively. The average $r_{fit}$ for 10% sampling is slightly lower at 0.79 versus 0.84, but still quite high. In some of the specific cases $r_{fit}$ is not lower or only slightly lower (e.g., Cases F, G and J). The most $r_{fit}$ dropped was for Case H which is significantly more than for any of the other cases. Also note that the training and validation results are only slightly less for 10% sampling.

With 10% of the sampled values available, the WM can be used to provide good accuracy and fit of the data since accuracy is an important performance measure for a virtual sensor. These
measured output values are used to reduce model bias by applying feedback correction (FBC) as given by Eq. 48 below.

\[
\hat{y}_t = \hat{\eta}_t + \left( \hat{\eta}_t^* - \hat{\eta}_t \right) 
\]

(48)

where \( t^* \) is the time of the most recently measured output.

The WM 10% sampling results with FBC are given in Table 6 along with the 10% sampling results from Table 5. As shown, with FBC, AAE dropped from 0.32 to 0.18 or by 44%, which is substantial. The average \( r_{fit} \) also improved to the level with 100% sampling. However for one, Case E, this value actually dropped from 0.79 to 0.74. This can happen when FBC is reducing bias but not improving fitted correlation. However, since in this application accuracy is more important than correlation, this loss would be acceptable. As a representative case to illustrate the improvement from FBC, response plots for Case E are shown in Fig. 4 without FBC (Fig.4A) and with FBC (Fig.4B). As shown, FBC greatly improves the agreement of the fitted WM model with the measured values.

**Table 5. WM 10% sampling results compared to 100% sampling results from Table 1**.

<table>
<thead>
<tr>
<th>Case</th>
<th>AE</th>
<th>AAE</th>
<th>AAE -- 100% Sampled Model</th>
<th>( r_{fit} )</th>
<th>( r_{fit} ) -- 100% Sampled Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.00</td>
<td>0.01</td>
<td>0.09</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td>B</td>
<td>0.00</td>
<td>0.01</td>
<td>0.09</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>C</td>
<td>0.09</td>
<td>0.39</td>
<td>0.37</td>
<td>0.54</td>
<td>0.61</td>
</tr>
<tr>
<td>D</td>
<td>0.18</td>
<td>0.28</td>
<td>0.24</td>
<td>0.74</td>
<td>0.82</td>
</tr>
<tr>
<td>E</td>
<td>0.48</td>
<td>0.50</td>
<td>0.54</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>F</td>
<td>0.44</td>
<td>0.45</td>
<td>0.54</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>G</td>
<td>-0.05</td>
<td>0.18</td>
<td>0.11</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>H</td>
<td>0.22</td>
<td>0.28</td>
<td>0.32</td>
<td>0.77</td>
<td>0.90</td>
</tr>
<tr>
<td>I</td>
<td>-0.10</td>
<td>0.24</td>
<td>0.18</td>
<td>0.84</td>
<td>0.89</td>
</tr>
<tr>
<td>J</td>
<td>-0.16</td>
<td>0.23</td>
<td>0.19</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td>Testing Absolute Mean</td>
<td>0.22</td>
<td>0.32</td>
<td>0.31</td>
<td>0.79</td>
<td>0.84</td>
</tr>
<tr>
<td>Testing Absolute Stdev</td>
<td>0.16</td>
<td>0.11</td>
<td>0.16</td>
<td>0.12</td>
<td>0.10</td>
</tr>
</tbody>
</table>

**Note: AE and AAE results are in °C.**
Table 6. WM 10% sampling results with FBC compared to results without (w/o) FBC**.

<table>
<thead>
<tr>
<th>Case</th>
<th>AAE FBC</th>
<th>AAE -- w/o FBC</th>
<th>$r_{fit}$ FBC</th>
<th>$r_{fit}$ -- w/o FBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.36</td>
<td>0.39</td>
<td>0.67</td>
<td>0.54</td>
</tr>
<tr>
<td>D</td>
<td>0.15</td>
<td>0.28</td>
<td>0.88</td>
<td>0.74</td>
</tr>
<tr>
<td>E</td>
<td>0.21</td>
<td>0.50</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>F</td>
<td>0.10</td>
<td>0.45</td>
<td>0.95</td>
<td>0.91</td>
</tr>
<tr>
<td>G</td>
<td>0.14</td>
<td>0.18</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>H</td>
<td>0.12</td>
<td>0.28</td>
<td>0.87</td>
<td>0.77</td>
</tr>
<tr>
<td>I</td>
<td>0.18</td>
<td>0.24</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td>J</td>
<td>0.15</td>
<td>0.23</td>
<td>0.89</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>0.18</td>
<td>0.32</td>
<td>0.84</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Stdev</strong></td>
<td>0.08</td>
<td>0.11</td>
<td>0.09</td>
<td>0.12</td>
</tr>
</tbody>
</table>

**Note: AE and AAE results are in °C.

Fig. 4. The response plots for Case E in Table 6 with all the measured outputs, 10% sampled outputs, and the WM fit with 10% sampling. The left plot (4A) is the fitted model without FBC and the right plot (4B) is with FBC. The reduction in model bias from using FBC is evident.
Concluding Remarks

The objective of this work is the identification and evaluation of an effective multiple-input process identification method with significant input crossed-correlation. Specifically, this work compared NM structures to WM structures that were linear and nonlinear in model parameters. These structures were formulated based on differential equations for the transfer functions with a scope of discrete-time modeling. The $J$ analysis of this study showed clearly the unacceptability of any structure that is linear in the dynamic model parameters. This study also showed why the WM is superior to the NM which is the use of different CEs for the transfer functions. This difference results in less pairwise correlation in the outputs from the transfer function in WM structures, which results in less pairwise correlation in the column of $J$ associated with the static parameters. The ability of the WM to have less pairwise correlation in the dynamic variables was demonstrated using data from a distillation study by examining the correlation matrix. In this nonlinear model parameterization study, the only difference in the NM and WM is the CEs, which is the same for the NM and allowed to be different for the WM. Using one fitted model and testing it on eight test sets spanning a period of about three years, the WM greatly outperformed the NM and gave excellent fitted correlation coefficients in two types of applications – FFC and soft sensor development.

This work shows the importance of using nonlinear parameterized physically-based model structures with unique parameterization for transfer functions when the inputs are significantly cross-correlated. Currently, the WM of Rollins et al. [5] is the only one that we know with these attributes. Thus, future work could be the expansion of methods that develop, improve and extend approaches with these characteristics.
Literature Cited


CHAPTER 5: MULTIPLE-INPUT SUBJECT-SPECIFIC MODELING OF PLASMA GLUCOSE CONCENTRATION FOR FEEDFORWARD CONTROL

A paper published in Industrial & Engineering Chemistry Research

Kaylee Kotz\(^1\), Ali Cinar\(^2\), Yong Mei\(^1\), Amy Roggendorf\(^\text{d}\), Elizabeth Littlejohn\(^3\), Laurie Quinn\(^4\) and Derrick K. Rollins, Sr. \(^1,5\)*

\(^1\)Department of Chemical and Biological Engineering, Iowa State University, Ames, IA, \(^2\)Department of Chemical and Biological Engineering, Illinois Institute of Technology, Chicago, IL, \(^3\)Institute for Endocrine Discovery and Clinical Care, University of Chicago Medicine, Chicago, IL, \(^4\)College of Nursing, University of Illinois at Chicago, Chicago, IL, \(^5\)Department of Statistics, Iowa State University, Ames, IA

*Corresponding Author

Abstract

The ability to accurately develop subject-specific, input causation models, for blood glucose concentration (BGC) for large input sets can have a significant impact on tightening control for insulin dependent diabetes. More specifically, for Type 1 diabetics (T1Ds), it can lead to an effective artificial pancreas (i.e., an automatic control system that delivers exogenous insulin) under extreme changes in critical disturbances. These disturbances include food consumption, activity variations, and physiological stress changes. Thus, this article presents a free-living, outpatient, multiple-input, modeling method for BGC with strong causation attributes that is stable and guards against over-fitting to provide an effective modeling approach for feedforward (FF) control. This approach is a Wiener block-oriented methodology, which has unique attributes for meeting critical requirements for effective, long-term, FF control.

Key words: Type 1 diabetes, artificial pancreas, Wiener modeling, block-oriented modeling, predictive modeling, model predictive control.
**Introduction**

Diabetes is characterized by an inability to synthesize, secrete, and/or, in some cases, respond to insulin. Without this vital hormone, cells and tissues cannot absorb glucose, and the patients’ cells can starve to death, despite high levels of glucose in the bloodstream. Among the two major types of diabetes, Type 1 diabetes is characterized by the inability to produce insulin. Type 1 diabetics often experience extreme variations in BGC which can have adverse long- and short-term effects such as severe hypoglycemia, hyperglycemia and organ destruction. Studies have established that there is a need to maintain glucose levels within a normal range (e.g. 80–150 mg/dL) to avoid complications caused by diabetes [1-4]. Therefore, Type 1 diabetics require daily exogenous insulin infusion for survival. Current injection treatment usually involves an insulin pump with manually controlled bolus infusion and pre-programmed basal infusion. However, often times the patient is still not able to mimic a normally occurring insulin profile using insulin pumps or/and insulin injections, which leads to inadequate regulation of blood glucose concentration (BGC) possibly causing hyperglycemia, hypoglycemia or various complications [5,6].

Consequently, what is needed is an automatic insulin delivery system (i.e., artificial pancreas) with the ability to determine continuously the amount of insulin required to provide optimum closed-loop glucose control (i.e., to minimize the variability around a desired glucose level) and to eliminate the individual from the insulin dosage decision making in this control loop.

The development of a closed-loop artificial pancreas has the potential to simultaneously reduce the risks of hypoglycemia and hyperglycemia while also enabling individuals with Type 1 diabetes mellitus to maintain a normal lifestyle [7]. To create a closed-loop artificial pancreas,
three crucial components are needed: a continuous glucose sensor, an insulin pump, and a robust controller [7-11].

For effective long-term control of BGC, the control system must be capable of tight control under critical disturbances with extreme changes such as food, activity and stress. While feedback control (FBC) and model predictive control (MPC) have shown promise under mild changes (e.g., overnight) in disturbances [12-15] these approaches have not shown strong promise for long-term tight control under extreme changes in disturbances. Due to recent technological advancements of body monitoring devices [30,31], activity-, stress-, circadian rhythm-related disturbances [16-19] can be monitored in real time, which makes feedforward control (FFC) a possibility. Given that FFC directly models the relationship between disturbances and the control variable, BGC in this context, an accurate modeling approach that can produce stable causation relationships between critical disturbances and BGC has the potential to make a significant advancement in the development of an effective long-term artificial pancreas.

Hence, the focus of this article is strictly model development for effective FFC. The maximization of cause-effect relationship between critical disturbances and BGC is the goal of this model. Mathematically, a viable and general FFC law based on the model \( \hat{y}_t = f_x(x_t; \hat{\theta}) \) is given below by Eq. 1:

\[
f_x(x_t; \hat{\theta}) - Y_{\text{set}} - B = f_x(x_t; \hat{\theta}) - Y_{\text{set}} - \left( f_x(\theta; \hat{\theta}) - Y_{\text{set}} \right) = f_x(x_t; \hat{\theta}) - f_x(\theta; \hat{\theta}) = 0
\]

(1)

where \( B \) is constant systematic model biased such that \( B = f_x(\theta; \hat{\theta}) - Y_{\text{set}} \), \( Y_{\text{set}} \) is the target value of the controlled variable (i.e., the set point); \( \hat{y}_t \) is the modeled estimate of BCG at current time \( t \);
*f* is the fitted function; *x* is a vector of measured input variables at *t*; \( \hat{\Theta} \) is the vector of estimated parameters; and *x* \(_{i,j}\) is the insulin infusion rate at *t* that is required to satisfy Eq. 1 (i.e., the inlet flow rate needed so that Eq. 1 is satisfied at *t*). Eq. 1 gives the estimated insulin infusion rate to compensate for the all modeled input changes. Thus, the goal and scope of this work is to obtain a model for *f* that is able to significantly tighten BGC in an automatic FFC scheme. Note that under Eq. 1 large systematic modeling bias does not impede effective FF control since \( B \) cancels out as shown. Physically, this means that Eq. 1 estimates the amount of insulin infusion at each time instant, i.e., \( x_{i,j} \), needed to dynamically compensate for deviations of modeled inputs from their initial values where the model was at the target BGC level. Modeling errors in estimating \( x_{i,j} \) will exist and will be compensated for under FB control. Note that any modeling approach that contains outputs, such as \( k\)-steps-ahead prediction models, does not meet the requirement of *f* in Eq. 1 and are, therefore, not in the scope of this work. Moreover, the only types of models that have relevance to our scope are those that depend on inputs only.

There are a number of studies in the literature involving the development of models in BGC for real type 1 diabetic subjects [20-26]. There are models that used measured BGC only [20-23], ones that use BGC and food consumption, namely carbohydrates only [24], and ones that use BGC, food and activities variables [25,26]. However, we have not found any approach that gives modeling results for only inputs and thus, the results reported in these articles are not in this scope of this work as they do not meet our criteria under Eq. 1.

Therefore, the goal of this work is the development of a FF modeling approach that has the characteristics mentioned above under Eq. 1. The outline for this article is as follows. Specific mathematical details of the proposed approach are given in the next section. Next, the details of the study in this article to evaluate the proposed method are given. Following this
section are the results of this study on 15 two-week outpatient data collection cases, which is then followed by concluding remarks including future work.

**Modeling Methodology**

The proposed modeling approach is a critical advancement over the modeling method proposed by Rollins et al. [27] which is an extension of the Wiener method developed by Rollins and Bhandari [28]. The Rollins et al. technique was developed in the context of non-insulin dependent Type 2 diabetics. For modeling Type 1 diabetics, it becomes necessary to refine this approach due to the incorporation of insulin infusion as an input. This refinement or extension involves the development of a new parameter estimation procedure that guards better against over-fitting and is better able to handle a large input set. Before introducing this new procedure, we present the modeling equations under a general Wiener framework that are the foundation to this approach.

**Mathematical Models**

Wiener modeling follows a block-oriented model structure formed by a series and/or parallel arrangement of unrestricted static functions and linear dynamic blocks. A block diagram with $p$ inputs and one output is given in Fig. 1.

The inputs, $x_i$ for $i = 1, \ldots, p$, of the Wiener network are the measured noninvasive variables or disturbances (i.e., food, activity, and stress) and the output, $y$, is BGC. Each input has its own linear dynamic block, $G_i$, and each dynamic block has an intermediate unobservable, output $v_i$, which represents the independent dynamic response of its corresponding input. All the intermediate $v_i$’s are collected and passed through a nonlinear static gain block, $f(V)$, to produce the final output, $y$. The linear dynamic blocks are essentially linear ordinary differential equations; a second-order-plus-lead with dead time (SOPLDT) form as shown in Eq. 2.
where \( i = 1, \ldots, p, p \) is the total number of inputs, \( \tau_i \) is the time constant, \( \zeta_i \) is the damping coefficient, \( \tau_{ai} \) is the lead parameter and \( \theta_i \) is the dead time. Using a backward difference approximation (e.g., \( \frac{dv_i(t)}{dt} \approx \frac{v_{i,t} - v_{i,t-\Delta t}}{\Delta t} \)), applied to a sampling interval of \( \Delta t \), Rollins et al. [27] obtained an approximate discrete-time form of Eq. 2.

\[
v_{i,t} = \delta_{1,i}v_{i,t-\Delta t} + \delta_{2,i}v_{i,t-2\Delta t} + \omega_{1,i}x_{i,t-(\Delta t-\theta)} + \omega_{2,i}x_{i,t-(2\Delta t-\theta)}
\]

where, to satisfy the unity gain constraint, \( \omega_{2,i} = 1 - \delta_{1,i} - \delta_{2,i} - \omega_{1,i} \) and

\[
\delta_{1,i} = \frac{2\tau_i^2 + 2\tau_i\zeta_i\Delta t}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2} \quad (4)
\]

\[
\delta_{2,i} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2} \quad (5)
\]

\[
\omega_{1,i} = \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2} \quad (6)
\]

As described in Rollins et al. [27], this discrete form provides several strengths. First, the function form does not change as values of the parameters change, unlike the continuous form that can change as \( \zeta_i \) changes. Secondly, one does not have to be concerned about applying a fading memory algorithm that is needed for the continuous form to truncate terms after a certain period in the past. Thirdly, Eqs. 4-6 are highly nonlinear in the dynamic parameters \( (\tau_{ai}, \tau_i, \text{and} \zeta_i) \) and these intelligent complex structures aid in strengthening input-causation relationships by restricting parameters estimates to regions that are phenomenologically sound.
Lastly, the physical constraints, namely unity gain, $\zeta_i > 0$ and $\tau_i > 0$, also provide intelligence towards physiologically sound structure.

The engineering community tends to define a linear model based on the time dependent variables in the model. For example, Eq. 2 is linear in $v(t)$ and the transfer functions represented by this equation are said to be “linear” (but in the variables). However, this article defines a nonlinear model based on the statistical definition, which in Bates and Watts [34] is defined as “at least one of the derivatives of the expectation function with respect to the parameters depends on at least one of the parameters.” Thus, the statistical definition is based on the form of the parameters in the model and not the variables. If the proposed approach estimated the parameters in Eq. 3 directly, the model would be linear and would fall in the scope of the article by Garnier et al. [33] for linear multiple-input, single output (MISO) structures. However, because the parameters estimated in this work are the ones given in Eqs. 4-6, i.e., the dynamic parameters, the proposed model is nonlinear, and not in the scope of the models in Garnier et al. [33] which is deliberate and is a unique strengthen of this approach.

Figure 1. Block diagram for a general Wiener network with $p$ inputs and one output. Each input, $x_i$, is passed through their own linear dynamic block, $G_i(s)$, after which these unobservable intermediate outputs are collected and passed through a single unrestricted static gain function, $f(V)$, to produce the output, $y$. 
After obtaining Eq. 3 for each $i$, the modeled glucose value is determined by substituting these results into the static function, $f(V)$, such as a second order regression form shown below:

$$\eta_t = f(V) = a_0 + a_1v_{1,t} + \cdots + a_pv_{p,t} + b_1v_{1,t}^2 + \cdots + b_pv_{p,t}^2 + c_{1,2}v_{1,t}v_{2,t} + \cdots + c_{p-1,p}v_{p-1,t}v_{p,t}$$  \hspace{1cm} (7)

where $a_i$, $b_i$, and $c_{i,j}$ denote the linear, quadratic and interaction parameters for $i = 1, \ldots, p-1$ and $j = 2, \ldots, p$. The measurement model is given as

$$y_t = \eta_t + \varepsilon_t$$  \hspace{1cm} (8)

where $y_t$ is the modeled glucose concentration at time instant $t$, $\varepsilon_t$ is the error term under the assumptions of independence, normality and constant variance (i.e., $\varepsilon_t \sim N(0, \sigma^2)$, $\forall t$). Under these assumptions Rollins et al. [27] proposed the following estimator for BGC under this measurement model:

$$\hat{y}_t = \hat{\eta}_t = \hat{a}_0 + \hat{a}_1v_{1,t} + \cdots + \hat{a}_pv_{p,t} + \hat{b}_1v_{1,t}^2 + \cdots + \hat{b}_pv_{p,t}^2 + \hat{c}_{1,2}v_{1,t}v_{2,t} + \cdots + \hat{c}_{p-1,p}v_{p-1,t}v_{p,t}$$  \hspace{1cm} (9)

Thus, Eq. 9, along with Eqs. 3-6 give the functional form with its supporting equations for the proposed FF controller under this work. Later, in the Results Section, this control law is given as a differential equation.

**Modeling Procedure**

The proposed modeling procedure is a novel approach to maximize input-causation, guard against over-fitting, and maximize long-term stability. As discussed above, we attempt to maximize input-causation throughout the use of highly nonlinear structures and physical constraints. Cross-validation, in a novel fitting strategy, is used to guard against over-fitting. We do not use a $k$-fold cross validation procedure with the testing data randomly split into $k$ equal
groups as this is not a realistic evaluation since in practice the model can only be applied to data collected after the model is built. Thus, our cross-validation procedure uses only Testing data obtained after any data (i.e., Training and Validation data) used to influence model building. We seek to enhance long-term stability by obtaining consistent performance under significant changes in unmeasured disturbances. Our cross-validation procedure aids in this goal by seeking to achieve similar fitting results on all data sets. In addition, we evaluate the models using Testing data several days after the Training data so that unmeasured disturbances are more likely to be correlated differently with measured inputs.

The proposed modeling methodology is an extension and enhancement of the procedure proposed by Rollins et al. [27] due to the larger number of inputs (13 variables) including the addition of two exogenous insulin inputs and the additional complexities they bring. For simplicity and to provide the best fit for mild extrapolation, this work used a reduced form of Eq. 9 that is given by Eq. 10, below. As shown, Eq. 10 eliminates all second order and interaction terms of Eq. 9 and consists only of the first order, $a$, terms.

$$
\hat{y}_t = \hat{\eta}_t = \hat{a}_0 + \hat{a}_1 \hat{v}_{1,t} + \cdots + \hat{a}_{13} \hat{v}_{13,t}
$$

(10)

where $\hat{v}_{i,t}$ is the estimate of $v_{i,t}$ obtained by substituting the estimated dynamic parameters, i.e., $\hat{\tau}_i, \hat{\tau}_w, \hat{\tau}_s$, and $\hat{\theta}_i$, into Eqs. 3-6 for all $i$. Note that the linear form of Eq. 10 makes this particular network structure equivalent to a general class of transfer functions where the gains for each one is contained in Eq. 10 by $a_i, i=1, \ldots, 13$. Also, note that a model is completely specified when the dynamic parameters in Eqs. 4 to 6 have estimates for obtaining Eq. 3; then for each input these equations are incorporated into Eq. 10 along with estimates for the $a$’s.

As stated in Rollins et al. [27], the modeling objective is simply to maximize the true but unknown correlation coefficient between measured and fitted BGC. This quantity is represented
by $\rho_{y,\hat{y}}$ and estimated by $r_{fit}$. Thus, under this criterion, as a minimum, a model is considered

**useful**, if, and only if,

$$\rho_{y,\hat{y}} > 0 \quad (11)$$

Since the degree of usefulness increases with $\rho_{y,\hat{y}}$, the goal is to obtain the largest (as close to the upper limit of 1) value as possible. Due to the highly complex mapping of the parameters into the response space of $r_{fit}$, the following criterion is used in obtaining the parameter estimates:

\[
\text{Minimize } \text{Training } \text{SSE} = \sum_{i=1}^{n_{data}} (y_i - \hat{y}_i)^2 \quad (12)
\]

Subject to: $\zeta_i > 0, \tau_i > 0, \theta_i > 0$, maximizing $r_{fit}$ for Training and Validation, $\forall i$

The objective criterion, Eq. 12, is written to address the effect of unmeasured disturbances which is an artifact of real modeling as opposed to hypothetical simulated data modeling without unmeasured disturbances. The validation data set helps to guard against fitting the training data to unmeasured input variables that are correlated with measured input variables during training but differently during validation. This is done by seeking to obtain similar fitted correlation for training and validation sets which is a goal of the proposed modeling approach. The systematic bias that appears in the validation set is largely due to level changes in unmeasured inputs. However, as Eq. 1 shows, the proposed approached is not affected by this type of systematic bias and thus, can be effective in the presence of level changes for unmeasured disturbances. This is a critical attribute for long-term effective FF control.

For FF model evaluation, of the statistics commonly used for evaluating model fit, $r_{fit}$ is premier as supported by the discussion earlier in regards to Eq. 1. As discussed, a FF model needs to be accurate for the **change** in inputs and $r_{fit}$ is the best statistical measure of this ability. Statistics that are affected by model bias are not relevant as measures of performance in this context as discussed earlier.
The novelty of the proposed modeling procedure lies in a two level decomposition of the parameter estimation problem. The first level decomposes the static and dynamic problem. That is, the dynamic parameters, i.e., the parameters in Eq. 2, are estimated first and separately from the static parameters, i.e., the $\hat{a}_i$'s in Eq. 10. The second level decomposes dynamic parameter estimation into $p = 13$ separate (i.e., sub-) problems, one for each input. For this approach to be possible, a modeling structure must allow these decompositions. Under a SOPLDT dynamic model structure, the Wiener network is the only one that does as opposed to other common networks like the autoregressive moving-average exogenous input (ARMAX) model [27,29].

For this approach to be effective, with only one input, $x_i$, in each dynamic estimation problem, $r_{fit}$ must depend only on the dynamic parameters to obtain the best set. That is, the value of $r_{fit}$ must be solely controlled by the values of the dynamic parameters irrespective of the values of static parameters. Fortunately, this is the case because with only one input, $x_i$, $r_{fit} = r_{y,\hat{y}_i}$, the correlation coefficient for $y_t$ and $\hat{y}_{i,t}$ (see the Appendix for the mathematical proof). Thus, for the simple linear regression model (SLRM) (i.e., Eq. 10 with one input), since $\hat{y}_{i,t}$ depends only on the dynamic parameters, $r_{fit}$ depends only on the dynamic parameters. Although a formal proof is given in the Appendix, one can prove this in practice quite easily by changing the static parameters for a fixed set of dynamic parameters and verifying that the value of $r_{fit}$ does not change by observation.

In the proposed procedure, models are developed from Training and Validation data sets. The Training set is used to determine the value of SSE and in adjusting the values of the parameters directly to minimize this value. The Validation set calculates $r_{fit}$ for each adjustment on the values of the parameters, and is used to stop the minimization process for SSE if $r_{fit}$ for the Training set increases significantly and causes a significant drop in $r_{fit}$ for the Validation set. This
is the practical way that we feel that cross validation is done in practice as mentioned above. In addition, the proposed procedure includes a more stringent condition -- that is, $r_{fit}$ for both sets to be close to one another. Thus, for each input, the goal is not just for high values of $r_{fit}$ for Training and Validation but also for their values to be close. Moreover, this procedure will lower the value of $r_{fit}$ in the training data to bring its value close to the validation data and vice versa. We impose this condition because we have found from modeling many cases that when this condition is met that the final fit of the static model at the fixed set of dynamic values produces $r_{fit}$ values in Training, Validation and Testing sets that are very close together to minimize overfitting and maximum long-term stability as discussed above.

After finding the dynamic parameters, the next step in the procedure is to obtain the static parameters under Eq. 10. With the dynamic parameters fixed, this becomes a linear regression problem which has a global minimum as the solution. However, since we have a validation set, we observe its $r_{fit}$ performance under an iterative approach to the global minimum using an iterative optimization process. We have found that most of the time, the global minimum is the optimal solution but sometimes we find a slightly better solution based on the validation set that is not too far away from the global minimum.

The final process in the proposed procedure is the elimination of inputs that adversely affect the final model when fitting the combined set of inputs under the static model. Each input is removed from the model with all the other inputs kept in the model. If $r_{fit}$ increases when an input is removed, this input is taken out of the final model. After completing this process for all the inputs, the inputs that passed this test are used to obtain the fit of the final model under the static model. It should be noted that in most cases all the activity inputs were retained in the model and if any were eliminated this number was only a few. Rollins et al. obtained this set
from an extensive study in type 2 modeling involving all 22 of the armband inputs. Given that most of the inputs were retained in this work for each subject, this set appears to be quite acceptable. Also, note that the final set of inputs for a given model is not of concern in this work since we have no use for the models beyond model building to evaluate this approach.

The Study

Subjects in this study followed a two week free-living outpatient protocol in which no constraints or conditions were placed on their daily diet or lifestyle. The subjects in this study were all healthy young adults from the ages of 18 to 25 with type 1 diabetes and on insulin pump therapy with a body mass index (BMI) from 20.8 to 27.6. To obtain a sufficiently fast sampling rate necessary for discrete-time (DT) dynamic glucose modeling, the iPro™ Continuous Glucose Monitor (CGM) (Medtronic MiniMed, Inc., Northridge, California) was used to provide glucose measurements. Use of the CGM requires the insertion of a short flexible sensor (by needle) into the subcutaneous tissue of the abdominal/supra-iliac area (i.e., between the umbilicus and the hip). The sensor samples the surrounding interstitial glucose, which is then used to infer an individual’s blood glucose levels with a reporting frequency of every five minutes. Following FDA recommendations the sensors were replaced every three days. A period of one to two hours of missing measurements resulted during initialization of the new sensor after each insertion. To maximize sensor reading accuracy the sensor must be calibrated with at least four finger-stick measurements daily from the subject’s personal blood glucose lancet meter.

Activity information was collected using the SenseWear® Pro3 Body Monitoring System (BodyMedia Inc., Pittsburgh, Pennsylvania) shown in Fig. 2, which is worn on the triceps of the subject’s arm. The SenseWear® armband utilizes pattern detection algorithms [30,31] that employ physiologic signals from a unique combination of sensors to generate values for twenty
activity variables. The armband collects data using a two-axis accelerometer and four sensors that are used to determine heat flux, skin temperature, near body temperature, and galvanic skin response (GSR). The two-axis accelerometer provides information about body position and tracks upper arm movement. The heat flux sensor calculates the amount of heat being dissipated from the body by measuring the amount of heat lost along a thermally conductive path between the skin and a vent on the side of the armband. Skin temperature and near-body temperature are measured by sensitive thermistors and GSR is measured via the conductivity of the subject’s skin as it varies due to physical and emotional stimuli [31]. The SenseWear® armband samples at a rate of once per minute, however, measurements at five minute intervals were used to match the sampling rate of the CGM used in this study. The armband was typically only removed once a day while the subject was showering. Finally, to represent circadian rhythm (i.e. the body’s internal clock) we used a variable that we called the time of day (TOD) which is simply 24 hour clock time.

Figure 2. BodyMedia, Inc. SenseWear® Pro3 Body Monitoring System

Food information was collected using food logs. For Subjects 1-6, and 11 detailed food logs were kept on the actual food consumed and for Subjects 7-10 meals were logged based on the size of the meal with small = 1, medium = 2 and large = 3. As part of the detailed protocol,
subjects recorded the approximate serving size and the time they started eating, for all of the food they consumed, into a PDA, which used Weightmania® Pro software (Edward A. Greenwood, Inc., Cambridge, Massachusetts) to determine the carbohydrate, fat, and protein content of their meals. In addition, the subjects’ insulin pumps were downloaded on biweekly basis to retrieve their daily bolus and basal insulin infusion rate data. The thirteen variable input set for this study is given in Table 1.

Table 1. Input variables.

<table>
<thead>
<tr>
<th>Food</th>
<th>Activity</th>
<th>Circadian Rhythm</th>
<th>Insulin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7. Near Body Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. Transverse Accel – MAD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9. GSR – average</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10. Energy Expenditure</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on results in Rollins et al. [27] for a Type 2 subject and in Beverlin et al. [32] for 20 non-insulin dependent Type 2 diabetics, a goal was set for $r_{fit}$ to be greater than 0.40 with a value greater than 0.60 considered excellent. Note that this goal of $r_{fit}$ may not seem very high. However, one must recognize that this is an application in FF model development for inputs only and not a model prediction application requiring high model accuracy. Our objective function, Eq. 11, is defined in terms of usefulness, which essentially means any model that has the potential to significantly tighten BGC for a given subject. For a given subject, it is not likely necessary for $\rho_{y,y}$ to be too high to achieve usefulness when input model causation is strong. However, ultimately, the only way to truly evaluate the effectiveness of a given model is its use
in FFC. Models with strong input-causation fitting can actually do better than models with weaker input-causation and higher $r_{fit}$ results.

The first half of this study involved splitting the data into a week of training data and a week of validation data. The second half of the study split the data into a week of training data, four days of validation data, and three days of test data.

While correlation is the premier performance statistic as mentioned above, three other statistics were determined that are affected by model bias. Since bias can be neutralized as shown in Eq. 1, these statistics are irrelevant as FF model performance measures, but are included to give an indication of how well $r_{fit}$ holds its level in Validation and Testing data under conditions when model bias can be significant. These additional statistics are the average deviation (AD), the average absolute deviation (AAD) and mean relative absolute deviation (MRAD). The AD is simply the average of the residuals and is an estimate of model bias, $B$, as shown below:

$$\hat{B} = AD = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)$$

(13)

where $n$ is the number of observed blood glucose measurements in the statistic being calculated. A check that the aforementioned convergence criterion is met is the AD value equaling 0.0 mg/dL. A model with a significantly large absolute value of AD or model bias will tend to raise AAD values. The equation for AAD is similar to AD except AAD takes the absolute value of the residuals before finding the average.

$$AAD = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

(14)

The spread in BCG can vary widely among T1Ds and AAD will tend to be large when the BGC spread is large. Therefore, MRAD, a relative ADD value, is defined as follows:
Results

Tables 2 and 3 contain results for 15 subject-specific models and Tables 4 and 5 for 11 subject-specific models. The data for Subjects 1-11 were collected by an experienced graduate student that left the project after completing the first phase of the study. The other ones (Subjects 12-15) were collected by a new, less experienced, graduate student in the second phase of the study. This lack of experienced is revealed in the tables by the number of days of modeling data. Although data collection for each subject was about 2 weeks, the amount of useful modeling data was much less for Subjects 12-15 because of missing BGC data that the CGMS did not give. Thus, for a given subject, the data in Tables 2 and 3 had a split of ½ for Training and ½ for Validation and in Tables 4 and 5 the split was ½ Training, 2/7th Validation, and 3/14th Testing.

As mentioned, Tables 2 and 3 equally splits the data into Training and Validation sets. In Table 2 all the inputs are included in the models and in Table 3 the armband inputs are not included. Thus, a comparison of these tables indicates the modeling improvement from use of the armband inputs. The average $r^2_{fit}$ with and without the armband for Validation are 0.62 and 0.52, respectively, indicating a very significant improvement from use of the armband. The Training and Validation $r^2_{fit}$ values are in general quite close together and thus, supporting the effectiveness of the proposed modeling approach to obtain similar values to guard against over-fitting. Note that the mean Training and Validation values are 0.60 and 0.62, respectively, and 0.50 and 0.52, respectively, for Tables 2 and 3, respectively. While the three biased indicating statistics are quite larger for the Validation results in several cases, $r^2_{fit}$ is quite consistent given that it is an estimate of fitted correlation with significant standard error. Thus, the approach

$$MRAD = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}$$  \hspace{1cm} (15)
appears to be maintaining its level of fit quite well from Training to Validation results. The fit of
model for Subject 11 is given in Fig. 3. As shown, the input-only model tracks the observed
BGC quite well in terms of correlation. The systematic bias of the Validation fit is quite evident
as well as the apparent shift in the average BGC level from Training to Validation data.

Table 2. Model results for one week (or 7/14th) of training and one week (or 7/14th) of validation
with all inputs included. AE and AAE values are in mg/dL.

| Subject | Days | Training | |  |  | Validation | |  |  |
|---------|------|----------|----------|----------|----------|----------|----------|----------|
|         | AD   | AAD      | MRAD     | r_fit    | AD       | AAD      | MRAD     | r_fit    |
| 1       | 14.0 | 0.00     | 44.8     | 0.44     | 0.61     | 20.4     | 50.5     | 0.33     | 0.68     |
| 2       | 13.0 | 0.00     | 72.1     | 0.62     | 0.49     | -18.0    | 68.5     | 0.72     | 0.51     |
| 3       | 13.9 | 0.00     | 48.0     | 0.37     | 0.68     | -0.1     | 49.3     | 0.33     | 0.66     |
| 4       | 10.7 | 0.00     | 31.6     | 0.38     | 0.53     | 33.4     | 48.4     | 0.39     | 0.55     |
| 5       | 14.0 | 0.00     | 62.6     | 0.47     | 0.56     | 15.1     | 73.6     | 0.53     | 0.55     |
| 6       | 13.9 | 0.00     | 50.1     | 0.31     | 0.67     | 24.9     | 45.6     | 0.23     | 0.68     |
| 7       | 14.0 | 0.00     | 46.7     | 0.43     | 0.69     | 37.1     | 56.5     | 0.36     | 0.64     |
| 8       | 14.0 | 0.00     | 32.7     | 0.36     | 0.45     | 10.8     | 43.2     | 0.42     | 0.43     |
| 9       | 13.9 | 0.00     | 51.8     | 0.37     | 0.63     | -35.2    | 64.2     | 0.62     | 0.56     |
| 10      | 16.8 | 0.00     | 47.4     | 0.30     | 0.57     | 14.0     | 46.8     | 0.30     | 0.73     |
| 11      | 15.1 | 0.00     | 33.7     | 0.23     | 0.72     | -24.3    | 47.2     | 0.43     | 0.79     |
| 12      | 8.9  | 0.00     | 56.3     | 0.25     | 0.63     | -33.0    | 83.0     | 0.57     | 0.72     |
| 13      | 8.2  | 0.30     | 55.4     | 0.47     | 0.54     | 52.5     | 76.9     | 0.41     | 0.58     |
| 14      | 7.9  | -0.10    | 48.2     | 0.39     | 0.56     | -30.1    | 48.6     | 0.59     | 0.61     |
| 15      | 13.6 | 0.00     | 23.1     | 0.21     | 0.56     | 20.7     | 45.2     | 0.35     | 0.56     |
| Mean    | 12.8 | 0.00     | 47.4     | 0.37     | 0.60     | 7.1      | 56.5     | 0.44     | 0.62     |

The mean Validation r_fit for the detailed food logged cases, 1-6, and 11-15 is 0.63. The mean Validation r_fit for the non-detailed food logged cases, 7-10 is 0.59.

Tables 4 and 5 include test data. In Table 4 all the inputs are included in the models and
in Table 5 the armband inputs are not included. The average r_fit with and without the armband for
Testing are 0.59 and 0.51, respectively. These results are very similar to the previous ones in
Tables 2 and 3 and support a significant improvement from use of the armband. The Training,
Validation, and Testing r_fit values are in general quite close together and thus, supporting the
effectiveness of the proposed modeling approach to obtain similar values to guard against over-fitting. Note that the mean Training, Validation, and Testing values are 0.58, 0.59 and 0.59, respectively, and 0.50, 0.54, and 0.51, respectively, for Tables 4 and 5, respectively. The AAD and MRAD values in these tables are similar to the ones in Tables 2 and 3 and the analysis of model bias is the same as before. That is, while they are larger for the Validation and Testing results in several cases, \( r_{fit} \) is quite consistent, Thus, the approach appears to be maintaining its level of \( r_{fit} \) quite well from Training to Validation to Testing results.

**Table 3.** Model results for one week (or 7/14\(^{th}\)) of training and one week (or 7/14\(^{th}\)) of validation without armband inputs included. AE and AAE values are in mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Days</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AD</td>
<td>AAD</td>
<td>MRAD</td>
</tr>
<tr>
<td>1</td>
<td>14.0</td>
<td>0.0</td>
<td>49.2</td>
</tr>
<tr>
<td>2</td>
<td>13.0</td>
<td>0.0</td>
<td>65.4</td>
</tr>
<tr>
<td>3</td>
<td>13.9</td>
<td>0.0</td>
<td>61.0</td>
</tr>
<tr>
<td>4</td>
<td>10.7</td>
<td>0.0</td>
<td>32.2</td>
</tr>
<tr>
<td>5</td>
<td>14.0</td>
<td>0.0</td>
<td>66.1</td>
</tr>
<tr>
<td>6</td>
<td>13.9</td>
<td>0.0</td>
<td>62.8</td>
</tr>
<tr>
<td>7</td>
<td>14.0</td>
<td>0.0</td>
<td>58.3</td>
</tr>
<tr>
<td>8</td>
<td>14.0</td>
<td>0.0</td>
<td>35.8</td>
</tr>
<tr>
<td>9</td>
<td>13.9</td>
<td>0.0</td>
<td>59.3</td>
</tr>
<tr>
<td>10</td>
<td>16.8</td>
<td>0.0</td>
<td>48.8</td>
</tr>
<tr>
<td>11</td>
<td>15.1</td>
<td>0.0</td>
<td>33.5</td>
</tr>
<tr>
<td>12</td>
<td>8.9</td>
<td>0.2</td>
<td>58.4</td>
</tr>
<tr>
<td>13</td>
<td>8.2</td>
<td>0.0</td>
<td>61.5</td>
</tr>
<tr>
<td>14</td>
<td>7.9</td>
<td>-0.1</td>
<td>48.2</td>
</tr>
<tr>
<td>15</td>
<td>13.6</td>
<td>0.0</td>
<td>25.2</td>
</tr>
</tbody>
</table>

The mean Validation \( r_{fit} \) for the detailed food logged cases, 1-6, and 11-15 is 0.52. The mean Validation \( r_{fit} \) for the non-detailed food logged cases, 7-10 is 0.50.

In comparing \( r_{fit} \) results of detailed food log cases (Subjects 1-6, 11-15) versus meal size (non-detailed) food log cases (Subjects 7-10), it is not conclusive how much the detailed food logs improve the fit, if at all. Averaged \( r_{fit} \) Validation results in Tables 2 and 3 for detailed and
non-detailed cases are 0.63 versus 0.59 and 0.52 versus 0.50, respectively. In Tables 4 and 5, averaged \( r_{fit} \) Testing results for detailed and non-detailed cases are 0.62 versus 0.53 and 0.50 versus 0.55, respectively. If Subject 8 is removed in Table 4, the averaged \( r_{fit} \), for the 3 remaining subjects, increases from 0.53 to 0.60, and it is very close to the detailed result of 0.62. Thus, it seems that if detailed food log are improving the fit, it does not appear to be very significant.

**Table 4.** Model results for one week (or 7/14\(^{th}\)) of training, 4 days (or 4/14\(^{th}\)) of validation, and 3 days (or 3/14\(^{th}\) of testing with all inputs included. AD and AAD values are in mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Days</th>
<th>7 Days Training</th>
<th>4 Days Validation</th>
<th>3 Days Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AD</td>
<td>AAD</td>
<td>MRAD</td>
<td>( r_{fit} )</td>
</tr>
<tr>
<td>1</td>
<td>14.0</td>
<td>0.00</td>
<td>45.0</td>
<td>0.45</td>
</tr>
<tr>
<td>2</td>
<td>13.0</td>
<td>0.00</td>
<td>67.2</td>
<td>0.55</td>
</tr>
<tr>
<td>3</td>
<td>13.9</td>
<td>0.00</td>
<td>50.9</td>
<td>0.39</td>
</tr>
<tr>
<td>4</td>
<td>10.7</td>
<td>0.00</td>
<td>31.9</td>
<td>0.38</td>
</tr>
<tr>
<td>5</td>
<td>14.0</td>
<td>0.00</td>
<td>63.3</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>13.9</td>
<td>0.00</td>
<td>52.8</td>
<td>0.33</td>
</tr>
<tr>
<td>7</td>
<td>14.0</td>
<td>0.00</td>
<td>56.4</td>
<td>0.53</td>
</tr>
<tr>
<td>8</td>
<td>14.0</td>
<td>0.00</td>
<td>36.5</td>
<td>0.41</td>
</tr>
<tr>
<td>9</td>
<td>13.9</td>
<td>0.00</td>
<td>53.2</td>
<td>0.38</td>
</tr>
<tr>
<td>10</td>
<td>16.8</td>
<td>0.00</td>
<td>44.9</td>
<td>0.29</td>
</tr>
<tr>
<td>11</td>
<td>15.1</td>
<td>0.00</td>
<td>30.4</td>
<td>0.21</td>
</tr>
<tr>
<td>Mean</td>
<td>13.9</td>
<td>0.00</td>
<td>48.4</td>
<td>0.40</td>
</tr>
</tbody>
</table>

The mean Testing \( r_{fit} \) for the detailed food logged cases, 1-6, and 11 is 0.62. The mean Testing \( r_{fit} \) for the non-detailed food logged cases, 7-10 is 0.53.
Table 5. Model results for one week (or 7/14th) of training, 4 days (or 4/14th) of validation, and 3 days (or 3/14th) of testing without armband inputs. AD and AAD values are in mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Days</th>
<th>7 Days Training</th>
<th>4 Days Validation</th>
<th>3 Days Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AD</td>
<td>AAD</td>
<td>MRAD</td>
</tr>
<tr>
<td>1</td>
<td>14.0</td>
<td>0.00</td>
<td>49.2</td>
<td>0.47</td>
</tr>
<tr>
<td>2</td>
<td>13.0</td>
<td>0.00</td>
<td>68.0</td>
<td>0.58</td>
</tr>
<tr>
<td>3</td>
<td>13.9</td>
<td>0.00</td>
<td>60.6</td>
<td>0.48</td>
</tr>
<tr>
<td>4</td>
<td>10.7</td>
<td>0.00</td>
<td>32.1</td>
<td>0.38</td>
</tr>
<tr>
<td>5</td>
<td>14.0</td>
<td>0.00</td>
<td>66.0</td>
<td>0.51</td>
</tr>
<tr>
<td>6</td>
<td>13.9</td>
<td>0.00</td>
<td>63.5</td>
<td>0.40</td>
</tr>
<tr>
<td>7</td>
<td>14.0</td>
<td>0.00</td>
<td>60.7</td>
<td>0.58</td>
</tr>
<tr>
<td>8</td>
<td>14.0</td>
<td>0.00</td>
<td>33.6</td>
<td>0.38</td>
</tr>
<tr>
<td>9</td>
<td>13.9</td>
<td>0.00</td>
<td>63.8</td>
<td>0.48</td>
</tr>
<tr>
<td>10</td>
<td>16.8</td>
<td>0.00</td>
<td>48.3</td>
<td>0.31</td>
</tr>
<tr>
<td>11</td>
<td>15.1</td>
<td>0.00</td>
<td>31.5</td>
<td>0.22</td>
</tr>
</tbody>
</table>

The mean Testing r-fit for the detailed food logged cases, 1-6, and 11 is 0.50. The mean Testing r-fit for the non-detailed food logged cases, 7-10 is 0.55.

Figure 3. Fitted and observed BGC versus time for Subject 11 in Table 2.
The FF model can be obtained by application of Eq. 1. Note that from Eq. 10 at the initial steady state when \( x_1 = v_1 = 0 \) and \( \hat{y} = Y^{\text{set}} + \hat{B} \), then \( \hat{a}_o = Y^{\text{set}} + \hat{B} \). For simplicity, we use only two inputs, one for say, carbohydrates, for example, \( x_1 \), and one for insulin infusion, say \( x_2 \). Thus, Eq. 1 in continuous time is:

\[
\hat{a}_1 \hat{v}_1(t) + \hat{a}_2 \hat{v}_2(t) = 0
\]  

(16)

In the Laplace domain Eq. 16 becomes

\[
\hat{a}_1 \hat{V}_1(s) + \hat{a}_2 \hat{V}_2(s) = \hat{a}_1 X_1(s) \hat{G}_1(s) - \hat{a}_2 X_2(s) \hat{G}_2(s) = 0
\]  

(17)

By taking \( \hat{\theta}_i = 0 \), for simplicity, in the \( s \)-domain, Eq. 2 becomes

\[
\hat{G}_i(s) = \frac{\hat{r}_{ar} s + 1}{\hat{r}_{ar}^2 s^2 + 2 \hat{r}_{ar} \hat{\xi}_s s + 1}
\]  

(18)

Substituting Eq. 18 into Eq. 17, rearranging and writing as a differential equation, one gets the following form for the proposed FF controller:

\[
\hat{\xi}_{a2} \hat{\xi}_1^2 \frac{d^3 x_1(t)}{dt^3} + \hat{\xi}_1 \left( 2 \hat{\xi}_{a2} \hat{\xi}_1 + \hat{\xi}_1 \right) \frac{d^2 x_1(t)}{dt^2} + \left( \hat{\xi}_{a2} + 2 \hat{\xi}_1 \hat{\xi}_2 \right) \frac{dx_1(t)}{dt} + \hat{a}_1 x_1(t) = \\
-\frac{\hat{a}_1 \hat{\xi}_{al} \hat{\xi}_2^2 \frac{d^3 x_1(t)}{dt^3}}{\hat{a}_2} - \frac{\hat{a}_2 \hat{\xi}_{al} \hat{\xi}_2 \frac{d^2 x_1(t)}{dt^2}}{\hat{a}_2} - \frac{\hat{a}_2 \hat{\xi}_1 \frac{dx_1(t)}{dt}}{\hat{a}_2} - \frac{\hat{a}_2 x_1(t)}{\hat{a}_2}
\]  

(19)

As shown by Eq. 19, the FF control law contains the numerator and denominator dynamics of both the load variable and the manipulated variable, the insulin infusion rate. This equation can be solved numerically using a technique such as Euler’s method to give the insulin infusion rate at each time instant to satisfy Eq. 1.

For an ARMAX model, the FF controller would be (the derivation is not shown for space considerations)

\[
\hat{\xi}_{a2} \frac{dx_1(t)}{dt} + x_2(t) = -\frac{\hat{a}_1 \hat{\xi}_{al} \frac{dx_1(t)}{dt}}{\hat{a}_2} - \frac{\hat{a}_1 x_1(t)}{\hat{a}_2}
\]  

(20)
Note that this controller only has numerator dynamics for the inputs. Since inputs in this context will have very different denominator dynamics (e.g., very different residence times of carbohydrates and fats), it is not reasonable to use a model with this restriction when it is unnecessary. Thus, ARMAX and its related structures such as ARX are not considered based on the goals of this research work.

**Concluding Remarks**

This work proposed an input-only, multiple-input, outpatient free-living, modeling methodology for T1D subjects for FFC. We have not found any input-only models for real T1D subjects in the literature and thus, none that meet the requirements for FF controller development under Eq. 1, our scope. The proposed methodology extends the one developed by Rollins et al. [27] for Type 2 diabetic subjects to include insulin infusion. It decomposes the static and dynamic parameter estimation problems and then decomposed the dynamic parameter estimation problem into a separate one for each input. This strategy seeks to guard against over-fitting and to strengthen long-term stability by producing Training, Validation and Testing fitted correlations ($r_{fit}$) that are similar as evidence of achieving these goals. The activity inputs provided by the armband were shown to be quite valuable in improving model fit. For several subjects, the fits were excellent ($r_{fit} \geq 0.6$) even though they were developed from free-living data and totally from non-invasive inputs. This work makes a major step towards the goal of the development of a long-term automatic insulin delivery system for T1Ds.

The goal of a FF controller is to determine the insulin infusion rate that will cancel the effects of measured input changes on BGC. Thus, the FF model used to build the FFC system can have large model bias and be quite effective as long as it is able to accuracy determine the
insulin infusion rate to cancel out changes of measured and modelled inputs. Consequently, a FF model can be quite biased and still very effective in this application.

Given that of model bias does not matter in this application, the premier performance measure is $r_{fit}$ as it is not affected by model bias and gives an indication of model fit. However, high correlation does not necessarily mean high causation. Moreover, under free-living data collection, and this context of modeling real subjects, input model causation cannot be determined and can only, ultimately, be evaluated under real subject FFC. Nonetheless, our approach has attempted to strengthen input causation in model building by using highly structure nonlinear models with physically interpretable dynamic parameters and a model identification strategy to minimize over fitting. We sought to accomplish the latter by a cross-validation strategy that used sequential data for training, validation, and testing sets and obtains similar values of $r_{fit}$ for all the data sets. In addition, by using a sequential cross-validation approach as opposed to a $k$-fold approach, ones is able to evaluate how $r_{fit}$ maintains it level under a model with changing bias due to changes in unmeasured disturbances and this is more realistic in terms of practice since the model will be used in practice on data that is collected after the model is built.

In the future, we will continue to improve the accuracy of the method and evaluate its suitability for FF control in real data studies. While the Wiener structure has unique strengths, it is still limited. Consequently, we will continue to look for other types of structures that have better phenomenological attributes, especially for the incorporation of unmeasured blood insulin. We feel this accomplishment has the potential for a significant advancement in model-based FFC applications as it should provide the insulin infusion rate to compensate for multiple and simultaneous input changes in a dynamic fashion.
Acknowledgements

The authors would like to thank BodyMedia, Inc. of use of the SenseWear® Pro3 Body Monitoring System and Nisarg Vyas from BodyMedia, Inc. for his assistance with the study. Financial support from the National Institutes of Health (Grant 1R01DK085611 ) through Illinois Institute of Technology and the Juvenile Diabetes Research Foundation (Grant 5-2010-7) is gratefully acknowledged.

Literature Cited


Appendix

The purpose of this appendix is to provide a mathematical proof that \( r_{fit} \), under the simple linear regression, i.e., Eq. 27 with one input. Let \( \hat{\eta}_i = \hat{a}_0 + \hat{a}_i \hat{v}_{i,t} \), in this context, \( r_{fit} \) is mathematically given by

\[
\begin{align*}
    r_{fit} &= r_{y_i, \hat{y}_i} = \frac{\sum_{j=1}^{n} (y_j - \bar{y})(\hat{\eta}_j - \bar{\eta})}{\sqrt{\sum_{j=1}^{n} (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^{n} (\eta_j - \bar{\eta})^2}} \\
    &= \frac{\sum_{j=1}^{n} (y_j - \bar{y})(\hat{a}_0 + \hat{a}_i \hat{v}_{i,j} - \hat{a}_0 - \hat{a}_i \bar{v}_i)}{\sqrt{\sum_{j=1}^{n} (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^{n} (\hat{a}_0 + \hat{a}_i \hat{v}_{i,j} - \hat{a}_0 - \hat{a}_i \bar{v}_i)^2}} \\
    &= \frac{\hat{a}_i \sum_{j=1}^{n} (y_j - \bar{y})(\hat{v}_{i,j} - \bar{v}_i)}{\sqrt{\hat{a}_i^2} \sqrt{\sum_{j=1}^{n} (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^{n} (\hat{v}_{i,j} - \bar{v}_i)^2}} \\
    &= \frac{\hat{a}_i}{|\hat{a}_i|} r_{y_i, \hat{v}_{i,t}}
\end{align*}
\]  

Thus, with \( \hat{a}_i > 0 \), \( r_{fit} = r_{y_i, \hat{v}_{i,t}} \) and for \( \hat{a}_i < 0 \), \( r_{fit} = -r_{y_i, \hat{v}_{i,t}} \). This result means that if the correlation of measured blood glucose concentration (BGC) and \( \hat{v}_{i,t} \) is positive, \( \hat{a}_i \) can be set at any positive value and \( r_{fit} \), which will be \( > 0 \), will depend only of the behavior of \( \hat{v}_{i,t} \) which is independently controlled by the values of the dynamic parameters associated with \( v_{i,t} \). Conversely, if the
correlation of BGC and $\hat{v}_{i,t}$ is negative, $\hat{a}_i$ can be set at any negative value and $r_{fit}$ will be $> 0$ and independently controlled by the values of the dynamic parameters associated with $v_{i,t}$. 
CHAPTER 6: AN EXTENDED STATIC AND DYNAMIC FEEDFORWARD CONTROL ALGORITHM FOR INSULIN DELIVERY IN THE CONTROL OF BLOOD GLUCOSE LEVEL

A paper published in Industrial & Engineering Chemistry Research

Derrick K. Rollins, Sr. 1,2*, Christina E. Goeddel1, Shana L. Matthews3, Yong Mei1, Amy Roggendorf1, Elizabeth Littlejohn4, Laurie Quinn5 and Ali Cinar6

1Department of Chemical and Biological Engineering, Iowa State University, Ames, IA; 2Department of Statistics, Iowa State University, Ames, IA; 3Department of Computer Science, Iowa State University, Ames, IA; 4Institute for Endocrine Discovery and Clinical Care, University of Chicago Medicine, Chicago, IL; 5College of Nursing, University of Illinois at Chicago, Chicago, IL; 6Department of Chemical and Biological Engineering, Illinois Institute of Technology, Chicago, IL

*Corresponding Author

Abstract

The potential for successful automatic control of blood glucose concentration (BGC) has entered a new era due to recent technological advancements in insulin pumps and blood glucose sensors. However, a critical advancement necessary for full automation and long-term use is a control algorithm that can effectively maintain tight control of BGC under extreme variation of important disturbances such as activity, stress, and food consumption. Since feedforward control (FFC) models disturbances directly it has the potential to eliminate the effects of disturbances completely. A Wiener-type feedforward control law is limited to the inclusion of only input (i.e., modeled disturbances and the manipulated variable) dynamics. Using a semi-coupled modeling network that includes pseudo-blood insulin concentration, this work presents a more phenomenological FFC law that includes input dynamics, blood insulin and blood glucose dynamics and blood glucose levels. Modeling results on fifteen adults with type 1 diabetes mellitus for the proposed method are nearly identical to Wiener modeling results.
Key words: Type 1 diabetes, artificial pancreas, Wiener modeling, block-oriented modeling, predictive modeling, model predictive control.

Introduction

In a person without diabetes, several systems such as the metabolic, endocrine, cardiovascular, etc. function collectively to maintain homeostasis. However in a person with diabetes, their inherent glucose regulation mechanism is dysfunctional. Glucose levels are affected by the state of the metabolic-physiological-endocrine system (consisting of factors such as insulin, stress, physical activity, hormonal levels, fatigue, etc.) [1-5]. The effects of all these factors on blood glucose concentration (BGC) are highly complex and inter-related [6]. In addition, factors such as food intake can cause glucose levels to change greatly and make glucose regulation and health management more difficult.

Insulin therapy involves multiple daily doses of insulin before meals or after meals to correct high blood glucose, with the amount either pre-recommended by a physician or decided by the patient on the basis measured blood glucose concentration (BGC) and number of carbohydrates to be ingested at the time of the meal. This protocol is not only inconvenient but also unreliable; often resulting in hypoglycemic and hyperglycemic episodes, both of which can be life-limiting and life-threatening [7]. Consequently, what is needed is automatic delivery of insulin that results in minimal variability around the desired glucose target.

The potential for successful automatic insulin delivery has entered a new era due to recent technological advancements of insulin pumps and blood glucose sensors. However, for full automation and control twenty four hours a day/seven days a week (24/7) the control algorithm must be capable of tight control for major disturbances such as meals, various from of activity and stress. Theoretically, the superiority of feedforward control (FFC) over all other
control systems is that corrective action can be taken to cancel the effects of disturbances on the control variable (i.e., BGC) [8]. While Feedback Control (FBC) [9, 10, 11] and some model-based algorithms [12-15] have shown promise and progress in real studies, there does not appear to be FFC clinic studies on real subjects in the literature. Actually, even from a search in the process control literature the success implementation of FFC on real systems appears to be quite limited as we were only able to find a few articles [16, 17]. We believe that this is due to the difficulty of developing accurate causative relationship of inputs on the control variable for real processes because of the existence of unmeasured disturbances and pairwise correlation of the inputs. Therefore, the objective of this work is the development a subject-specific FFC modeling methodology for maintaining tight BGC under free-living data collection (i.e., without any restrictions on the subject’s eating or activity) to effectively compensate for changes in meals, activity, and stress.

The general FFC law used in this work is determined as follows. Let $\hat{y}_t = \text{estimated value of the output (i.e., BGC) at time } t$ and given as

$$\hat{y}_t = f_x(X_t, \hat{\theta}) = B + \hat{\eta}_t$$  \hspace{1cm} (1)$$

with

$$\hat{y}_0 = f_x(X_0, \hat{\theta}) = B + \text{Y}^{set} = \hat{\eta}_0$$  \hspace{1cm} (2)$$

where $f_x(X_t, \hat{\theta})$ is a fitted function of input variables only (more specifically, no outputs); $X_t$ is a matrix of measured input variables; $\hat{\theta}$ is the vector of estimated parameters; $B$ is the model bias; $Y^{set}$ is the target value of the controlled variable (i.e., the set point); $B$, is determined by

$$B = f_x(X_0, \hat{\theta}) - \text{Y}^{set}$$

$\hat{\eta}_t$ is the bias corrected estimated BCG at $t$; and $\hat{\eta}_t$ is the estimate of the expected value of BGC at $t$. For the input variables that are measured, with a perfect model for this set of inputs, the FFC model determines the value of the manipulated variable, in this context
The insulin infusion rate at $t$, to offset all these input changes at $t$. Mathematically, using Eqs. 1 and 2 a general FFC law is given as:

$$\hat{y}_t - \hat{y}_0 = f_x\left(X_t; \hat{\theta}\right) - f_x\left(X_0; \hat{\theta}\right) = B + \hat{y}_t^e - (B + Y^\text{set}) = \hat{y}_t^e - Y^\text{set} = 0 \quad (3)$$

Thus, at each time instant, Eq. 3 determines the value of $x_{t,I}$ for $f_x\left(X_t; \hat{\theta}\right)$ to remain at its initial value of $f_x\left(X_0; \hat{\theta}\right)$, the set point plus its systemic bias. The FFC law is stated via Eq. 3 because $f_x\left(X_t; \hat{\theta}\right)$ is unrestricted. The common way the FFC law is given in process control textbook is for simple block diagrams [8]. However, in this work the block diagram is not simple but a semi-coupled network. Using a highly nonlinear coupled network in a simulation study of a continuous stirred tank reactor (CSTR), the implementation of Eq. 3 as a control algorithm is illustrated in this work.

Note that $f_x\left(X_t; \hat{\theta}\right)$ is a function of inputs only which is a necessity to satisfy Eq. 3 for FFC. Therefore, models that use outputs, such as $k$-steps-ahead models, cannot satisfy Eq. 3 for FFC. For example, for $k$-steps-ahead models, Eq. 3 becomes

$$\hat{y}_t - \hat{y}_0 = f_x\left(X_t; \hat{\theta}\right) + f_y\left(Y_t; \check{\phi}\right) - f_x\left(X_0; \hat{\theta}\right) - f_y\left(Y^*; \check{\phi}\right) = 0$$

$$\Rightarrow f_x\left(X_t; \hat{\theta}\right) = f_y\left(Y^*; \check{\phi}\right) - f_y\left(Y_t; \check{\phi}\right) \neq 0 \quad (4)$$

where $f_x\left(Y_t; \check{\phi}\right)$ is the portion of fitted model dependent on the past outputs; $Y_t$ is a vector of measured output variables; $\check{\phi}$ is a vector of estimated residual parameters; and $t^*$ is the earliest time $t$ that model estimates the fitted value. Comparing the FFC law with Eq. 4, it is evident the use of outputs overcompensates by $f_y\left(Y^*; \check{\phi}\right) - f_y\left(Y_t; \check{\phi}\right)$. Finally, it should be noted in Eq. 3 that the systematic bias, $B$, cancels and thus, does not adversely affect the determination of $x_{t,I}$. A modification of Eq. 3 for non-constant $B$ will be given later in the article.
Subject-specific modeling of BGC involving inputs have been limited to short data collection periods [19-22] for model development, mostly one or two disturbances (carbohydrates and exercise) [23, 24], and heavy reliance on previously measured BGC [25-28]. The only multiple-input subject-specific modeling of BGC meeting the requirements of Eq. 3 known to us is the work of Kotz et al [18] where a Wiener modeling approach was taken. In this work the FFC law was given in terms of model parameters as a differential equation. The limitation of this equation is that it is only a function of the static and dynamic parameters of the inputs which is a limitation of the Wiener structure. In this work, using a pseudo-variable (i.e., unmeasured variable) for blood insulin concentration (BIC) and a semi-coupled network for BGC and BIC is developed that leads to a FFC law that depends on static and dynamic input parameters, blood insulin and blood glucose dynamic parameters and the level of BGC. Thus, the proposed model is theoretically statically and dynamically superior to the Wiener model and thus, has the potential to provide a significant improvement in the FFC of BGC. A comparison study with the Wiener model on the 15 subjects [18] is given in this work.

**Modeling Methodology**

As mentioned above, the proposed semi-coupled modeling approach is an advancement over the Wiener modeling method (WMM) [18] in development of a FFC algorithm to address static and dynamic behavior better. While the WMM approach can produce feedforward (FF) controllers that contain input-specific numerator and denominator dynamics, its structural limitations impact its ability to maintain high accuracy over a broad range of BGC, especially at the extreme limits. The proposed method also allows for different input dynamics as in the Wiener approach but extends the ability to address other types of static and dynamic behavior via the use of a semi-coupled network that has an unmeasured pseudo-blood insulin concentration.
(BIC) variable. The use of this variable allows modeling on the blood insulin directly and the inclusion of the interaction of glucose and insulin in the blood. The net result of this novelty is the ability to specify a FF controller that not only has input-specific numerator and denominator dynamics, but also blood glucose and pseudo-blood insulin dynamics as well as a dependence on the BGC level. This section gives a brief review of the WMM and then the proposed extension of the coupled network.

**The Wiener Modeling Method**

In general, Wiener modeling follows a block-oriented model structure formed by a series and/or parallel arrangement of nonlinear static and linear dynamic blocks. A block diagram with $p$ inputs and one output is given in Fig. 1.

![Fig. 1. Block diagram for a general multiple-input, single-output Wiener network.](image)

The inputs, $x_i$ for $i = 1, \ldots, p$, of the Wiener network are the measured noninvasive variables or disturbances (i.e., food, activity, and stress) and the output, $y$, is BGC. Each input has its own linear dynamic block, $G_i$, and each dynamic block has an intermediate, unobservable output, $v_i$, which represents the independent dynamic response of its corresponding input. All the intermediate $v_i$'s are collected and passed through an unrestricted static gain block, $f(V)$, to
produce the final output, $y$. The linear dynamic blocks are essentially linear ordinary differential equations; a second-order-plus-lead-plus-dead-time (SOPLDT) form is shown in Eq. 5.

$$\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i \zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t)}{dt} + x_i(t)$$

(5)

where $i = 1, \ldots, p$, $p$ is the total number of inputs, $\tau_i$ is the time constant, $\zeta_i$ is the damping coefficient, $\tau_{ai}$ is the lead parameter and $\theta_i$ is the dead time. Using a backward difference approximation applied to a sampling interval of $\Delta t$, an approximate discrete-time form of Eq. 5 given below by Eq. 6 [29]:

$$v_{i,t} = \delta_{i,1} v_{i,t-\Delta t} + \delta_{i,2} v_{i,t-2\Delta t} + \omega_{i,1} x_{i,t-\Delta t} + \omega_{i,2} x_{i,t-2\Delta t}$$

(6)

where

$$\delta_{i,1} = \frac{2\tau_i^2 + 2\tau_i \zeta_i \Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}$$

(7)

$$\delta_{i,2} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}$$

(8)

$$\omega_{i,1} = \frac{2\tau_i \delta_{i,1}}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}$$

(9)

and

$$\omega_{i,2} = 1 - \delta_{i,1} - \delta_{i,2} - \omega_{i,1}$$

(10)

to satisfy the constraint of unity gain, and $x_{i,t}$ is the value of the $i$th input at $t$. Two additional physical constraints are $\tau_i > 0$ and $\zeta_i > 0$, $\forall i$.

The substitution of Eqs. 7-10 into Eq. 6 provide an advantage in modeling the independent input effects when using free-living outpatient data over the use of the linear form of Eq. 6 which would estimate the directly $\delta$’s and $\omega$’s directly [29]. More specifically, for each $v_{i,t}$, the parameters (i.e., $\delta_{1,i}$, $\delta_{2,i}$, $\omega_{1,i}$, $\omega_{2,i}$) in Eq. 6 are determined from highly non-linear functions of the continuous-time dynamic parameters, $\tau_{ai}$, $\tau_i$, and $\zeta_i$, via Eqs. 7-10. After
obtaining Eq. 6 for each \( i \), the modeled glucose value is determined by substituting these results into the specific static function, \( f(V) \), such as a second-order regression form shown below:

\[
\eta_i = f(V) = a_0 + a_1 v_{1,i} + \cdots + a_p v_{p,i} + b_1 v_{1,i}^2 + \cdots b_p v_{p,i}^2 \\
+ c_{1,2} v_{1,i} v_{2,i} + \cdots + c_{p-1,p} v_{p-1,i} v_{p,i}
\]  

(11)

where \( a_i, b_i, \) and \( c_{i,j} \), denote the linear, quadratic and interaction parameters for \( i = 1, \ldots, p-1 \) and \( j = 2, \ldots, p \). Note that Eq. 11 depends only on inputs as required by Eq. 3. The measurement model that corresponds to Eq. 11 is

\[
y_i = \eta_i + \epsilon_i
\]  

(12)

where \( y_i \) is the modeled BGC at time instant \( t \), \( \epsilon_i \) is the error term under the assumptions of independence, normality and constant variance (i.e., \( \epsilon_i \sim N(0, \sigma^2), \forall t \)). Under these assumptions the WMM estimator for BGC is given as [29]:

\[
\hat{y}_i = \hat{\eta}_i = \hat{a}_0 + \hat{a}_1 \hat{v}_{1,i} + \cdots + \hat{a}_p \hat{v}_{p,i} + \hat{b}_1 \hat{v}_{1,i}^2 + \cdots \hat{b}_p \hat{v}_{p,i}^2 \\
+ \hat{c}_{1,2} \hat{v}_{1,i} \hat{v}_{2,i} + \cdots + \hat{c}_{p-1,p} \hat{v}_{p-1,i} \hat{v}_{p,i}
\]  

(13)

The Proposed Approach

While the Wiener network is an excellent choice for the input transfer function model, in the response space, it is limited in representing the interaction of insulin and glucose in the blood. This drawback limits the ability to develop an accurate fit for insulin infusion rate, which is critical to controller performance. However, to model blood insulin and glucose interaction, one needs BIC at the sampling rate of BGC (i.e., every five minutes with the glucose sensor used in our studies). There is no such sensor in existence currently. To circumvent this need, we developed a semi-coupled network for BIC and BGC that was inspired by the compartment modeling work of Topps et al. [30] as illustrated by the following two equations for this work:

\[
\frac{dG(t)}{dt} = R_0(t) - [E_{G0}(t) + S_I(t)]G(t)
\]  

(14)
\[
\frac{dI(t)}{dt} = \frac{\beta(t)G(t)^2}{\alpha + G(t)^2} - kI(t)
\]  

(15)

where \(G\) is the BGC; \(I\) is the blood insulin concentration (BIC); \(R_0\) is the net rate of production at zero glucose; \(E_{G0}\) is the total glucose effectiveness at zero insulin; \(S_I\) is the total insulin sensitivity; \((\beta)\) is the mass of pancreatic beta cells (measured in mg); \((\alpha)\) is a coefficient in the Hill function that describes a sigmoid ranging from 0 to 1; and \(k\) is the clearance constant representing the combined insulin uptake at the liver, kidneys, etc.

BIC is not a measured variable in the proposed approach. As a result, the proposed approach uses a “pseudo” or “latent” BIC variable that may or may not even be observable in the body; hence, the use of the description “pseudo” insulin. This variable is allowing the use of a modeling structure that is hypothesized to provide characteristics for BGC that are more physiologically correct than a Wiener modeling approach. To our knowledge, this is the first use of such a variable in this context of subject-specific clinical modeling of BGC consisting only of inputs in free-living data collection. As this work will show, this novel change made a substantial advancement in obtaining a better phenomenological model and in providing a dynamic relationship between insulin infusion rate, consumed nutrients, and BGC. Our proposed Wiener/Semi-Coupled network is shown in Fig. 2.

As shown in Fig. 2, all of the inputs, \(x_i\), pass through a linear dynamic block to produce the unobservable dynamic output variables \(v_i\) as in the WMM. Note that, \(i = A_i, \ldots, A_p\), for the \(p\) activity inputs, \(C\) for carbohydrates, \(F\) for fats, \(P\) for proteins, and \(I\) for insulin. A dynamic mass balance on \(G_{FI}\) block in Fig. 2, which represents the BGC due to food and insulin only, gives

\[
\frac{dV_{FI}(t)}{dt} = a_cV_c(t) + a_FV_F(t) + a_PV_P(t) - a_{FI}I(t)V_{FI}(t)
\]  

(16)
where $a_C$, $a_F$, $a_P$, and $a_{FI}$, are estimable model parameters; $I(t)$ is the unmeasured BIC at time $t$; and $V_{FI}(t)$ is the BGC due to food and insulin changes only. Similarly, a dynamic mass balance on the $G_{BI}$ block, which represents BIC, gives

$$\frac{dI(t)}{dt} = a_I v_I(t) - a_{BI} I(t)$$

(17)

where $a_I$ and $a_{BI}$, are estimable model parameters. Note the similarities of Eqs. 16 and 17 with Eqs. 14 and 15, respectively.

The function $f(V)$ is called “the static function.” This function can theoretically be of any form. For effectiveness under mild extrapolation, in modeling real BGC data we will use a first-order linear regression structure, given in discrete form as:

$$y_i = f(V) = G_i \eta_i + \varepsilon_i = a_0 + V_{FI} + a_{AI} v_{AI} + \ldots + a_{AP} v_{AP} + \varepsilon_i$$

(18)
where \( e \) is the error term assumed to be independently normally distributed with mean 0 and variance \( \sigma^2 \) for all \( t \), and \( a_0 \), \( a_i \)'s are static estimable model parameters. The modeling approach with the network defined by Fig. 2, i.e., the proposed method, will be called the “coupled modeling method” (CMM) in this article, although strictly speaking, Fig. 2 is a semi-coupled network.

**Model Identification Procedure**

The use of free-living data presents the challenge of not over-fitting the data to behavior that is correlated with the BGC, the response. The common way to address this challenge is to use cross-validation, which splits the data into three sets: Training, Validation, and Testing. The model parameters are estimated using the Training data, which is usually the largest set. The Validation set is used to guard against over-fitting, as the final set of parameters must maximize the fit in this set. Since these two sets influence the parameter estimation process, the Testing set is used as a final check on model fit, as the data in this set has no influence on the values of the estimated parameters (called “process or system identification”). To achieve the best fit possible, this work follows the cross-validation process of the WMM [18] which seeks to maximize fit in the training data set and obtain comparable (i.e., similar) fit in the other two sets. Following the cross-validation strategy of the WMM [18], this work developed a novel procedure to estimate the model parameters in the CMM under the least squares criterion that decomposes the problem as follows. First, all the coupled parameters, i.e., the ones in Eqs. 14-15, are estimated under a fixed set of dynamic parameters for the nutrient variables obtained from Eq. 6 and Eq. 16 with \( a_{A1} \) to \( a_{Ap} \) set to zero (i.e., without activity variables). Next, the dynamic parameters for each activity variable are estimated individually under Eqs. 6 and 16 with \( V_{FI} \) set to 0. After obtaining
the dynamic parameters for each activity variable, the static model coefficients of Eq. 16 are estimated.

**The Studies**

There are two fundamental contributions of this work towards the goal of the development of an effective artificial pancreas. The first one is the semi-coupled modeling approach as represented in Fig. 2 with the nonlinear parameterized approach as given by Eqs. 5-10. The first study in this section will evaluate the proposed approach based on its ability to develop an improved FFC model. These models will be developed from data of real people with type 1 diabetes. The second study will evaluate the control algorithm based on the FFC law given by Eq. 3 to directly implement a highly nonlinear model structure of a complex network such as Fig. 2 without using a linear approximation. Since controller data does not exist for this control approach, this study will be given using simulated data from a continuous stirred tank reactor (CSTR). This study consists of three inputs and will model one measured state variable (the controlled variable) and one unmeasured state variable in a coupled network. The fitted model is highly nonlinear in time dependent process variables and in the dynamic and static model parameters. In this study feedback control (FBC) will be compared with feedback-feedforward control (FBFFC) to assess the strengths of the proposed control algorithm.

**Modeling Real Subjects with Type 1 Diabetes**

**The Data Sets**

The data sets for evaluating the coupled modeling method (CMM) of Fig. 2 are taken from Kotz et al. [18], where the WMM was evaluated on 15 data sets consisting of two weeks of data each. (Note that for Subjects 1–6, and 11, detailed food logs were kept on the actual food consumed and for Subjects 7–10 only meal size (small, medium and large) was logged.) In this
section, the CMM will be compared directly with the WMM [18]. Given that the modeling
objective is the determination of the insulin infusion rate to offset changes in measured input
variables on BGC, i.e., the development of FFC models, a model is useful when it achieves this
objective to any significant degree. Thus, in this context model usefulness is not determined by
prediction accuracy but by the ability of the model to explain output behavior strictly from
measured input changes. The effectiveness of a FF controller for any modeled disturbance will
depend on the accuracy of its causative modeled relationship on BGC and its contribution to the
variability of BGC for the given subject. Therefore, for a given set of inputs: 1. their combined
contribution to the variability of BGC is at a particular level; 2. this level represents the limit of
the FF controller on the reduction of variability for the subject modeled, and; 3. the actual or
effective reduction of variability will depend on how well the model has obtain the causative
input relationships on BGC. Given the consistency between the Training, Validation, and Testing
results of the WMM on these 15 data sets, it appears that the fitted models in this work captured
quite adequately the variability of BGC for the set of inputs modeled. Thus, the CMM is not
expected to significantly improve the fit over the WMM and the similar fit is an indication that it
was also able to capture quite adequately the amount of variation in BGC due to set of modeled
inputs.

**Performance Statistic**

A FF controller will be useful when it can significantly reduce the variation of BGC. A
FF controller will be useful when it is able to effectively manipulate insulin infusion rate to
compensate for the variables that are in its input set. Since its effectiveness is based on its ability
to model the relationship of the measured inputs on BGC and model bias can be neutralized, the
premier performance statistic is the correlation between the fitted input-only model and the measured BGC, $r_{fit}$. This work also reports AD, the average of the difference between the fitted and measured BGC and AAD, the average of the absolute difference between the fitted and measured BGC. It is recognized that correlation does not imply causation. The only way to truly evaluate model accuracy under causation is to independently manipulate the inputs and determine the fit of the model. This is not possible using free living data which is inherently cross correlated. However, after presenting the $r_{fit}$ results for this study, we will interrogate the modeling approaches for physical soundness and potential for success in FFC applications.

**Inputs**

The 13 variable input set is given in Table 1 and consists of three (3) food nutrients, seven (7) activity variables including one to measure stress, 24-hour clock time for circadian rhythm, and two (2) for insulin infusion. Modeling 13 inputs is quite a large set in this context and we know of no work beyond these data sets that has modeled nearly this many variables in clinical data collection. Note that for any particular case, the final form of the reduced model is irrelevant to our objective since we have no use for the fitted models beyond determining their $r_{fit}$ values in this study.

The activity variables were collected using the SenseWear® Pro3 Body Monitoring System (BodyMedia Inc., Pittsburgh, Pennsylvania) shown in Fig. 3, which is worn on the triceps of the subject’s arm. The SenseWear® armband utilizes pattern detection algorithms [31, 32] that employ physiologic signals from a unique combination of sensors to generate values for twenty activity variables. The armband collects data using a two-axis accelerometer and four sensors that are used to determine heat flux, skin temperature, near body temperature, and galvanic skin response (GSR). The two-axis accelerometer provides information about body
position and tracks upper arm movement. The heat flux sensor calculates the amount of heat being dissipated from the body by measuring the amount of heat lost along a thermally-conductive path between the skin and a vent on the side of the armband. Skin temperature and near-body temperature are measured by sensitive thermistors, and GSR is measured via the conductivity of the subject’s skin as it varies due to physical and emotional stimuli [32]. The SenseWear® armband samples at a rate of once per minute, however, measurements at five minute intervals are used to match the sampling rate of the Medtronic continuous glucose monitor system (CGMS).

**Results**

Modeling results for the WMM and the CMM are given in Tables 2-7 for the 15 subject-specific models. Results are given for three different sets of data. The “Training” data set is used to estimate model parameters. The “Validation” data set is used to stop the convergence process of the estimation procedure on the Training data and to guard against over-fitting the model. The goal is to stop the estimation process when the highest $r_{fit}$ is obtained on the Validation data given that its value was close to the Training value. The “Testing” data set represents data not used in any way to influence the fitting of the model.
Table 1. Input variables.

<table>
<thead>
<tr>
<th>Food</th>
<th>Activity</th>
<th>Circadian Rhythm</th>
<th>Insulin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7. Near Body Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. Transverse Accel – MAD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9. GSR – average</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10. Energy Expenditure</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data collection was done in two parts. Part 1 consisted of Subjects 1-11 and was done in the first year of the study and run by an experienced graduate student. The data collection was not completed in the first year because we were not able to fully recruit all the needed subjects. This student graduated and new, inexperienced students were used the second year in Part 2 for Subjects 12-15.

Our performance goal for the CMM was an $r_{fit}$ similar to or better than the WMM. This goal was reached as shown by the results in Tables 2-7. In addition, our modeling procedure guarded well against over-fitting as $r_{fit}$ values for Training and Validation are very similar. Table 2 contains results for Part 1 using all the inputs with one week of training and one week of validation. Table 3 represents the same conditions as Table 2 except for excluding the armband inputs. As one can see, the armband appears to contribute significantly in improving the fit; from an average $r_{fit}$ of 0.54 to 0.64 for the CMM. Tables 4 and 5 are for Subjects 1-11 with one week for Training, 4 days for Validation and 3 days for Testing. The averaged $r_{fit}$ values vary only slightly for the three data sets in both tables and for the WMM and the CMM. Again, the
armband appears to significantly help as the averaged $r_{fit}$ for the Testing data sets go from 0.53 to 0.61. The results for Part 2, i.e., Subjects 11-15, are very similar to the ones in Part 1. The only observation worth mentioning is that for the CMM, the armband did not help quite as much with $r_{fit}$, going from an average value of 0.60 to 0.64. However, there is less confidence in this size of improvement due to the significantly smaller number of cases.

From the form of Eq. 13, its fundamental limitation is revealed as an inadequate structure in determining a dynamic food to insulin infusion relationship to achieve a target value of BGC. This limitation is due to the separate additive nature of Eq. 13 for food consumption and insulin infusion. While its use can be effective in estimating BGC under a particular correlation relationship among all the inputs, it is not capable of providing a food consumption/insulin infusion coupling relationship for BGC. A simple test of inadequacy is to set all food inputs to zero with non-zero insulin infusion and vice versa. With all food inputs set to zero (under an assumption that BGC only changes for food consumption) $\hat{y}$ should drop to zero over time. And with insulin infusion set to zero, $\hat{y}$ should continue to increase over time. Figure 4 illustrates this inadequacy on one of the fitted cases, Subject 11. The top plot gives the fit with all the variables included. As shown, the fit is quite good given that it was produced using only input variables. The bottom left plot shows what happens when all the food coefficients in Eq. 13 are set to zero. As shown, its discrepancy with the measured BGC increases, but it is stable and not decreasing as it should be doing. Similarly, the bottom right plot shows what happens when the insulin coefficient in Eq. 13 is set to zero. As before, its discrepancy increases, but its level is stable and not increasing as it should be doing. Thus, while the WMM is capable of providing accuracy for fitting BGC, its structure is not adequate to give a correct physiological coupled relationship between food intake and insulin infusion on BGC.
Results to evaluate the CMM for inadequacy are shown in Fig. 5 for Subject 11. The top plot gives the fit with all the variables included. As shown, the fit is quite good and similar to the WMM shown in Fig. 4. The bottom left plot in Fig. 5 shows what happens when all the food coefficients in Eq. 10 are set to zero. As shown, the fitted BGC drops steadily towards zero as it should. The response is oscillatory and does not drop to a fixed value of zero because, as seen by Eq. 18, the activity variables have a Wiener-type structure and are thus, additive with respect to $V_{FI}$ and also contribute to BGC. Similarly, the bottom right plot shows what happens when the insulin infusion coefficient in Eq. 15 is set to zero. Here, as it should be, BGC rises steadily over time. Thus, not only can the CMM provide a good fit, but it also overcomes the WMM deficiencies in regards to zero insulin and zero food intake scenarios. Another critical advantage of the CMM worth noting is the use of BIC even though this output variable is unmeasured, which is another novel improvement.

While it is only possible to truly evaluate the cause-and-effect ability of our fitted models under real FFC for the subject modeled, we found, however, a way to evaluate the soundness of our CCM in its ability to provide realistic insulin infusion rates. We did this by obtaining the steady state relationships between carbohydrates ($x_C$), BGC ($v = V_{FI} = G$), and insulin infusion rates ($x_I$) (see Fig. 2) as follows. Note that in practice this experiment could not be done as it would not be practical or safe, but having a model allows this virtual evaluation. Using a backwards difference finite derivative, in discrete form, Eq. 16 becomes

$$\frac{V_{FI,t} - V_{FI,t-\Delta t}}{\Delta t} = a_c v_{C,t} + a_p v_{P,t} + a_p v_{P,t} - a_{FI} I_{I,t} V_{FI,t}$$  \hspace{1cm} (19)$$

and solving for $V_{FI,t}$ gives

$$V_{FI,t} = \frac{(a_c v_{C,t} + a_p v_{F,t} + a_p v_{P,t}) \Delta t + V_{FI,t-\Delta t}}{1 + a_{FI} I_{I,t} \Delta t}$$  \hspace{1cm} (20)$$
Setting $a_0 = a_F = a_P = a_A1 = ... = a_{Ap} = 0$, Eq. 20 becomes

$$G_t = V_{Fi,t} = \frac{a_C v_{i,t} \Delta t + G_{i,t}}{1 + a_{Fi} / \Delta t}$$

(21)

Table 2. WMM and CMM Part 1 Study results with all the inputs under one week of training and one week of validation. The units for AD and AAD are mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>Training</th>
<th>$r_{fit}$</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AD</td>
<td>AAD</td>
<td>$r_{fit}$</td>
</tr>
<tr>
<td>1</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>44.8</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>WMM</td>
<td>13.0</td>
<td>0.0</td>
<td>72.1</td>
<td>0.49</td>
</tr>
<tr>
<td>3</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>48.0</td>
<td>0.68</td>
</tr>
<tr>
<td>4</td>
<td>WMM</td>
<td>10.7</td>
<td>0.0</td>
<td>31.6</td>
<td>0.53</td>
</tr>
<tr>
<td>5</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>62.6</td>
<td>0.56</td>
</tr>
<tr>
<td>6</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>50.1</td>
<td>0.67</td>
</tr>
<tr>
<td>7</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>46.7</td>
<td>0.69</td>
</tr>
<tr>
<td>8</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>32.7</td>
<td>0.45</td>
</tr>
<tr>
<td>9</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>51.8</td>
<td>0.63</td>
</tr>
<tr>
<td>10</td>
<td>WMM</td>
<td>16.8</td>
<td>0.0</td>
<td>47.4</td>
<td>0.57</td>
</tr>
<tr>
<td>11</td>
<td>WMM</td>
<td>15.1</td>
<td>0.0</td>
<td>33.7</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>47.4</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WMM</td>
<td>14.0</td>
<td>1.1</td>
<td>43.5</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>13.0</td>
<td>-0.4</td>
<td>57.8</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>13.9</td>
<td>-0.1</td>
<td>51.7</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>10.7</td>
<td>-0.1</td>
<td>29.9</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>14.0</td>
<td>0.3</td>
<td>63.3</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>13.9</td>
<td>-0.5</td>
<td>49.9</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>14.0</td>
<td>0.3</td>
<td>44.8</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>14.0</td>
<td>0.9</td>
<td>32.8</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>13.9</td>
<td>6.3</td>
<td>54.3</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>16.8</td>
<td>0.0</td>
<td>43.5</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>CMM</td>
<td>15.1</td>
<td>0.2</td>
<td>29.8</td>
<td>0.82</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>WMM</td>
<td>13.9</td>
<td>0.9</td>
<td>45.4</td>
<td>0.65</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>CMM</td>
<td>13.9</td>
<td>0.9</td>
<td>45.4</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Table 3. WMM and CMM Part 1 Study results without armband inputs under one week of training and one week of validation. The units for AD and AAD are mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>Training AD</th>
<th>Training AAD</th>
<th>$r_{fit}$</th>
<th>Validation AD</th>
<th>Validation AAD</th>
<th>$r_{fit}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>49.2</td>
<td>0.52</td>
<td>26.9</td>
<td>56.8</td>
<td>0.57</td>
</tr>
<tr>
<td>2</td>
<td>WMM</td>
<td>13.0</td>
<td>0.0</td>
<td>65.4</td>
<td>0.58</td>
<td>-13.1</td>
<td>64.4</td>
<td>0.56</td>
</tr>
<tr>
<td>3</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>61.0</td>
<td>0.46</td>
<td>18.2</td>
<td>56.6</td>
<td>0.42</td>
</tr>
<tr>
<td>4</td>
<td>WMM</td>
<td>10.7</td>
<td>0.0</td>
<td>32.2</td>
<td>0.48</td>
<td>32.6</td>
<td>32.6</td>
<td>0.49</td>
</tr>
<tr>
<td>5</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>66.1</td>
<td>0.50</td>
<td>11.3</td>
<td>81.7</td>
<td>0.44</td>
</tr>
<tr>
<td>6</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>62.8</td>
<td>0.39</td>
<td>13.9</td>
<td>50.8</td>
<td>0.35</td>
</tr>
<tr>
<td>7</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>58.3</td>
<td>0.52</td>
<td>21.3</td>
<td>51.5</td>
<td>0.57</td>
</tr>
<tr>
<td>8</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>35.8</td>
<td>0.25</td>
<td>15.1</td>
<td>46.6</td>
<td>0.34</td>
</tr>
<tr>
<td>9</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>59.3</td>
<td>0.52</td>
<td>-21.8</td>
<td>64.4</td>
<td>0.43</td>
</tr>
<tr>
<td>10</td>
<td>WMM</td>
<td>16.8</td>
<td>0.0</td>
<td>48.8</td>
<td>0.55</td>
<td>13.3</td>
<td>49.8</td>
<td>0.67</td>
</tr>
<tr>
<td>11</td>
<td>WMM</td>
<td>15.1</td>
<td>0.0</td>
<td>33.5</td>
<td>0.70</td>
<td>-23.5</td>
<td>48.3</td>
<td>0.77</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>52.0</td>
<td>0.50</td>
<td>19.2</td>
<td>54.9</td>
<td>0.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>Training AD</th>
<th>Training AAD</th>
<th>$r_{fit}$</th>
<th>Validation AD</th>
<th>Validation AAD</th>
<th>$r_{fit}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CMM</td>
<td>14.0</td>
<td>4.0</td>
<td>49.4</td>
<td>0.55</td>
<td>26.8</td>
<td>66.4</td>
<td>0.51</td>
</tr>
<tr>
<td>2</td>
<td>CMM</td>
<td>13.0</td>
<td>13.8</td>
<td>60.2</td>
<td>0.68</td>
<td>-31.6</td>
<td>64.1</td>
<td>0.66</td>
</tr>
<tr>
<td>3</td>
<td>CMM</td>
<td>13.9</td>
<td>1.0</td>
<td>60.3</td>
<td>0.44</td>
<td>-5.0</td>
<td>55.9</td>
<td>0.41</td>
</tr>
<tr>
<td>4</td>
<td>CMM</td>
<td>10.7</td>
<td>-2.3</td>
<td>31.5</td>
<td>0.58</td>
<td>65.9</td>
<td>68.7</td>
<td>0.52</td>
</tr>
<tr>
<td>5</td>
<td>CMM</td>
<td>14.0</td>
<td>0.2</td>
<td>68.4</td>
<td>0.50</td>
<td>0.1</td>
<td>75.9</td>
<td>0.53</td>
</tr>
<tr>
<td>6</td>
<td>CMM</td>
<td>13.9</td>
<td>-0.1</td>
<td>58.9</td>
<td>0.51</td>
<td>20.9</td>
<td>49.3</td>
<td>0.51</td>
</tr>
<tr>
<td>7</td>
<td>CMM</td>
<td>14.0</td>
<td>3.8</td>
<td>54.2</td>
<td>0.59</td>
<td>-53.2</td>
<td>69.4</td>
<td>0.58</td>
</tr>
<tr>
<td>8</td>
<td>CMM</td>
<td>14.0</td>
<td>-0.1</td>
<td>34.9</td>
<td>0.35</td>
<td>14.5</td>
<td>46.3</td>
<td>0.30</td>
</tr>
<tr>
<td>9</td>
<td>CMM</td>
<td>13.9</td>
<td>1.1</td>
<td>62.7</td>
<td>0.50</td>
<td>-53.4</td>
<td>78.0</td>
<td>0.48</td>
</tr>
<tr>
<td>10</td>
<td>CMM</td>
<td>16.8</td>
<td>0.0</td>
<td>45.0</td>
<td>0.65</td>
<td>23.6</td>
<td>52.2</td>
<td>0.64</td>
</tr>
<tr>
<td>11</td>
<td>CMM</td>
<td>15.1</td>
<td>1.1</td>
<td>29.7</td>
<td>0.81</td>
<td>-30.4</td>
<td>58.2</td>
<td>0.81</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>CMM</td>
<td>13.9</td>
<td>2.6</td>
<td>50.5</td>
<td>0.56</td>
<td>29.6</td>
<td>62.2</td>
<td>0.54</td>
</tr>
</tbody>
</table>
Table 4. WMM and CMM Part 1 Study results with all the inputs under one week of training, four days of validation, and 3 days of testing. The units for AD and AAD are mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>7 Day Training</th>
<th>4 Day Validation</th>
<th>3 Day Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>AD</td>
<td>AAD</td>
<td>$r_{fit}$</td>
</tr>
<tr>
<td>1</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>45.0</td>
<td>0.60</td>
</tr>
<tr>
<td>2</td>
<td>WMM</td>
<td>13.0</td>
<td>0.0</td>
<td>67.2</td>
<td>0.57</td>
</tr>
<tr>
<td>3</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>50.9</td>
<td>0.64</td>
</tr>
<tr>
<td>4</td>
<td>WMM</td>
<td>10.7</td>
<td>0.0</td>
<td>31.9</td>
<td>0.53</td>
</tr>
<tr>
<td>5</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>63.3</td>
<td>0.55</td>
</tr>
<tr>
<td>6</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>52.8</td>
<td>0.65</td>
</tr>
<tr>
<td>7</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>56.4</td>
<td>0.56</td>
</tr>
<tr>
<td>8</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>36.5</td>
<td>0.21</td>
</tr>
<tr>
<td>9</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>53.2</td>
<td>0.61</td>
</tr>
<tr>
<td>10</td>
<td>WMM</td>
<td>16.8</td>
<td>0.0</td>
<td>44.9</td>
<td>0.62</td>
</tr>
<tr>
<td>11</td>
<td>WMM</td>
<td>15.1</td>
<td>0.0</td>
<td>30.4</td>
<td>0.79</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>48.4</td>
<td>0.58</td>
</tr>
<tr>
<td>1</td>
<td>CMM</td>
<td>14.0</td>
<td>18.2</td>
<td>45.9</td>
<td>0.71</td>
</tr>
<tr>
<td>2</td>
<td>CMM</td>
<td>13.0</td>
<td>2.3</td>
<td>57.7</td>
<td>0.73</td>
</tr>
<tr>
<td>3</td>
<td>CMM</td>
<td>13.9</td>
<td>-4.9</td>
<td>52.0</td>
<td>0.63</td>
</tr>
<tr>
<td>4</td>
<td>CMM</td>
<td>10.7</td>
<td>0.7</td>
<td>30.2</td>
<td>0.62</td>
</tr>
<tr>
<td>5</td>
<td>CMM</td>
<td>14.0</td>
<td>1.0</td>
<td>51.9</td>
<td>0.71</td>
</tr>
<tr>
<td>6</td>
<td>CMM</td>
<td>13.9</td>
<td>0.8</td>
<td>58.0</td>
<td>0.53</td>
</tr>
<tr>
<td>7</td>
<td>CMM</td>
<td>14.0</td>
<td>0.6</td>
<td>47.4</td>
<td>0.68</td>
</tr>
<tr>
<td>8</td>
<td>CMM</td>
<td>14.0</td>
<td>0.4</td>
<td>33.3</td>
<td>0.44</td>
</tr>
<tr>
<td>9</td>
<td>CMM</td>
<td>13.9</td>
<td>3.4</td>
<td>55.3</td>
<td>0.61</td>
</tr>
<tr>
<td>10</td>
<td>CMM</td>
<td>16.8</td>
<td>0.4</td>
<td>45.2</td>
<td>0.63</td>
</tr>
<tr>
<td>11</td>
<td>CMM</td>
<td>15.1</td>
<td>-0.1</td>
<td>30.9</td>
<td>0.79</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>CMM</td>
<td>13.9</td>
<td>3.0</td>
<td>46.2</td>
<td>0.64</td>
</tr>
</tbody>
</table>
Table 5. WMM and CMM Part 1 Study results without the armband inputs under one week of training, four days of validation, and 3 days of testing. The units for AD and AAD are mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>7 Day Training</th>
<th>4 Day Validation</th>
<th>3 Day Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>AD</td>
<td>AAD</td>
<td>$r_{fit}$</td>
</tr>
<tr>
<td>1</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>49.2</td>
<td>0.52</td>
</tr>
<tr>
<td>2</td>
<td>WMM</td>
<td>13.0</td>
<td>0.0</td>
<td>68.0</td>
<td>0.56</td>
</tr>
<tr>
<td>3</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>60.6</td>
<td>0.47</td>
</tr>
<tr>
<td>4</td>
<td>WMM</td>
<td>10.7</td>
<td>0.0</td>
<td>32.1</td>
<td>0.49</td>
</tr>
<tr>
<td>5</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>66.0</td>
<td>0.49</td>
</tr>
<tr>
<td>6</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>63.5</td>
<td>0.38</td>
</tr>
<tr>
<td>7</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>60.7</td>
<td>0.47</td>
</tr>
<tr>
<td>8</td>
<td>WMM</td>
<td>14.0</td>
<td>0.0</td>
<td>35.8</td>
<td>0.25</td>
</tr>
<tr>
<td>9</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>63.8</td>
<td>0.44</td>
</tr>
<tr>
<td>10</td>
<td>WMM</td>
<td>16.8</td>
<td>0.0</td>
<td>48.3</td>
<td>0.56</td>
</tr>
<tr>
<td>11</td>
<td>WMM</td>
<td>15.1</td>
<td>0.0</td>
<td>31.5</td>
<td>0.77</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>WMM</td>
<td>13.9</td>
<td>0.0</td>
<td>52.7</td>
<td>0.49</td>
</tr>
<tr>
<td>1</td>
<td>CMM</td>
<td>14.0</td>
<td>37.5</td>
<td>65.2</td>
<td>0.63</td>
</tr>
<tr>
<td>2</td>
<td>CMM</td>
<td>13.0</td>
<td>-3.6</td>
<td>65.7</td>
<td>0.69</td>
</tr>
<tr>
<td>3</td>
<td>CMM</td>
<td>13.9</td>
<td>0.9</td>
<td>60.3</td>
<td>0.44</td>
</tr>
<tr>
<td>4</td>
<td>CMM</td>
<td>10.7</td>
<td>0.7</td>
<td>30.2</td>
<td>0.62</td>
</tr>
<tr>
<td>5</td>
<td>CMM</td>
<td>14.0</td>
<td>0.4</td>
<td>63.3</td>
<td>0.57</td>
</tr>
<tr>
<td>6</td>
<td>CMM</td>
<td>13.9</td>
<td>0.8</td>
<td>58.0</td>
<td>0.53</td>
</tr>
<tr>
<td>7</td>
<td>CMM</td>
<td>14.0</td>
<td>-0.1</td>
<td>55.5</td>
<td>0.57</td>
</tr>
<tr>
<td>8</td>
<td>CMM</td>
<td>14.0</td>
<td>-1.2</td>
<td>35.9</td>
<td>0.36</td>
</tr>
<tr>
<td>9</td>
<td>CMM</td>
<td>13.9</td>
<td>-0.1</td>
<td>61.5</td>
<td>0.45</td>
</tr>
<tr>
<td>10</td>
<td>CMM</td>
<td>16.8</td>
<td>0.7</td>
<td>47.4</td>
<td>0.60</td>
</tr>
<tr>
<td>11</td>
<td>CMM</td>
<td>15.1</td>
<td>0.1</td>
<td>32.1</td>
<td>0.73</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>CMM</td>
<td>13.9</td>
<td>4.2</td>
<td>52.3</td>
<td>0.56</td>
</tr>
</tbody>
</table>
Table 6. WMM and CMM Part 2 Study results with all the inputs under one week of training and one week of validation. The units for AD and AAD are mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>AD</td>
<td>AAD</td>
</tr>
<tr>
<td>12</td>
<td>WMM</td>
<td>8.9</td>
<td>0.0</td>
<td>56.3</td>
</tr>
<tr>
<td>13</td>
<td>WMM</td>
<td>8.2</td>
<td>0.3</td>
<td>55.4</td>
</tr>
<tr>
<td>14</td>
<td>WMM</td>
<td>7.9</td>
<td>-0.1</td>
<td>48.2</td>
</tr>
<tr>
<td>15</td>
<td>WMM</td>
<td>13.6</td>
<td>0.0</td>
<td>23.1</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>WMM</td>
<td>9.7</td>
<td>0.1</td>
<td>45.8</td>
</tr>
</tbody>
</table>

|         |        |      | AD | AAD | $r_{fit}$ | AD | AAD | $r_{fit}$ |
| 12      | CMM    | 8.9  | 1.3 | 55.6 | 0.66 | -39.8 | 72.8 | 0.76 |
| 13      | CMM    | 8.2  | 0.0 | 49.4 | 0.67 | 48.5  | 74.3 | 0.61 |
| 14      | CMM    | 7.9  | -0.2| 45.3 | 0.60 | -5.4  | 41.9 | 0.61 |
| 15      | CMM    | 13.6 | 2.3 | 23.4 | 0.56 | 21.2  | 46.0 | 0.57 |
| Avg of Absolute Value | CMM | 9.7  | 0.9 | 43.4 | 0.62 | 28.7  | 58.7 | 0.64 |

Table 7. WMM and CMM Part 2 Study results without armband inputs under one week of training and one week of validation. The units for AD and AAD are mg/dL.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Method</th>
<th>Days</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>AD</td>
<td>AAD</td>
</tr>
<tr>
<td>12</td>
<td>WMM</td>
<td>8.9</td>
<td>0.2</td>
<td>58.4</td>
</tr>
<tr>
<td>13</td>
<td>WMM</td>
<td>8.2</td>
<td>0.0</td>
<td>61.5</td>
</tr>
<tr>
<td>14</td>
<td>WMM</td>
<td>7.9</td>
<td>-0.1</td>
<td>48.2</td>
</tr>
<tr>
<td>15</td>
<td>WMM</td>
<td>13.6</td>
<td>0.0</td>
<td>25.2</td>
</tr>
<tr>
<td>Avg of Absolute Value</td>
<td>WMM</td>
<td>9.7</td>
<td>0.1</td>
<td>48.3</td>
</tr>
</tbody>
</table>

|         |        |      | AD | AAD | $r_{fit}$ | AD | AAD | $r_{fit}$ |
| 12      | CMM    | 8.9  | 0.0 | 56.9 | 0.65 | -52.5 | 81.4 | 0.71 |
| 13      | CMM    | 8.2  | 0.0 | 59.0 | 0.47 | 68.0  | 87.5 | 0.56 |
| 14      | CMM    | 7.9  | -0.2| 45.3 | 0.60 | -5.4  | 41.9 | 0.61 |
| 15      | CMM    | 13.6 | -2.6| 24.6 | 0.54 | 10.9  | 45.8 | 0.50 |
| Avg of Absolute Value | CMM | 9.7  | 0.7 | 46.4 | 0.57 | 34.2  | 64.1 | 0.60 |
Fig. 4. Representative WMM fitted response plots for Subject 11 with one week of training and one week of validation. The graph on the top has all the inputs. The graph on the bottom left has no food consumption. The graph on the bottom right has no insulin infusion.
Again by using a backwards difference finite derivative, in discrete form, Eq. 17 becomes

\[ I_t = \frac{a_{ij}v_{ij}\Delta t + I_{t-\Delta t}}{1 + a_{hi}\Delta t} \]  

(22)

At steady state, \( V_{C,t} = X_{C,t} = X_C \), \( V_{I,t} = X_{I,t} = X_i \), \( G_t = G_{t-\Delta t} = \ldots = \bar{G} \), and \( I_t = I_{t-\Delta t} = \ldots = \bar{I} \), and Eq. 22 becomes
\[ \bar{I} = \frac{a_J x_J \Delta t + \bar{I}}{1 + a_{bl} \Delta t} \]  
(23)

and then from simplifying Eq. 23,

\[ \bar{I} = \frac{a_J}{a_{bl}} x_J \]  
(24)

Similarly, at steady state, Eq. 21 becomes

\[ \bar{G} = \frac{a_c x_c \Delta t + \bar{G}}{1 + a_{fJ} \bar{I} \Delta t} \]  
(25)

Solving for \( \bar{G} \) and substituting in Eq. 24, Eq. 25 becomes

\[ \bar{G} = \frac{a_c x_c \Delta t}{a_{fJ} \bar{I} \Delta t} = \frac{a_c}{a_{fJ} \bar{I}} x_c = \frac{a_c a_{bl}}{a_{fJ} a_J} x_c \]  
(26)

Now from Eq. 26

\[ x_I = \frac{a_c a_{bl}}{a_J a_{fJ} \bar{G}} x_c \]  
(27)

Soundness can be seen from a close examination of Eq. 27. First, \( x_I \) is correctly proportional to \( x_C \). Secondly, \( x_I \) is correctly inversely proportional to \( \bar{G} \). Since this effect cannot be demonstrated using real people, this relationship may not be immediately intuitive; but it implies that for a constant consumption rate of carbohydrates, the insulin infusion rate to maintain a constant BGC level (i.e., \( \bar{G} \)), decreases as \( \bar{G} \) increases. For example, a higher infusion rate is needed to maintain a BGC level of 100 mg/dL than to maintain a level of 300 mg/dL. This makes intuitive sense because additional insulin is needed to drop the level from 300 to 100 mg/dL. By substituting the estimated coefficients for a subject into Eq. 27, an evaluation of this subject can be made on the basis of the “realistic” results. This was done for Subject 11 and the results are plotted in Fig. 6. As shown, at \( x_C = 60 \) gm, \( x_I \) ranges from 2.4 units
to 7.3 units of insulin for $\bar{G}$ equal to 300 mg/dL to 100 mg/dL, respectively. These values are very practical and thus, Eq. 27 produces values that are quite realistic for this subject. Note that when the model is actually implemented in a real control setting, the dynamic form will be manipulating insulin levels for some target value of BGC and thus, this ratio will vary dynamically based on a number of conditions at each time instant and be specific to the subject’s personal model. The analysis presented is a very practical check on the soundness of a modeling approach and, perhaps equally as important, on the soundness of a specific subject’s model. Thus, an important contribution of this work is this practical evaluation and we recommend it for evaluation of FFC models.

If the modeling bias is not constant, i.e., $B = B_i$, which is likely the case, then Eq. 3 becomes

\[
0 \frac{B_i}{\theta_X} = 0 
\]  

\[
(28)
\]

\[
\frac{d}{dt} \begin{pmatrix} x_i(t); \dot{\theta} \end{pmatrix}_{x_i, \theta} = \left( f \left( X_i; \dot{\theta} \right) \right)_{x_i, \theta} - \left( f \left( X_0; \dot{\theta} \right) \right)_{x_i, \theta} = 0
\]

**Fig. 6.** The steady state relationship in the CMM for insulin infusion ($x_i$) and food consumption ($x_C$) at constant BGC ($G$).
If the model is causative for the inputs on BGC, for any subject, one can evaluate the impact of that particular FFC model to reduce variability from a target BGC upon its implementation into FFC. This will be illustrated for Subject 11 and its fit given in Table 4 using all the data. As shown by AAD, model bias varies considerably between the data sets. The initial time, $t = 0$, was selected the first night when the measured glucose was fairly stable and the subject was resting. At the selected time, $y_0 = Y^{\text{set}} = 140$ mg/dL. Thus, $\hat{B}_0 = \hat{y}_0 - 140$, where “\(^{\hat{}}\)” is used to represent estimate. For $B$, a very simple algorithm was used to estimate its value. More specifically, $\hat{B}_t = \hat{y}_{t-N} - y_{t-N}$. With $y_{\text{FFC},t}$ and $y_{\text{NAC},t}$ as the observed BGC with FFC and the observed BBC with no automatic control, respectively, then

$$y_{\text{FFC},t} = y_{\text{NAC},t} - (\hat{y}_t - \hat{y}_0) + \hat{B}_0 - \hat{B}_t$$

$$= y_{\text{NAC},t} - (\hat{y}_t - \hat{y}_0) + (\hat{y}_0 - y_{\text{NAC},0}) - (\hat{y}_{t-N} - y_{\text{NAC},t-N})$$  \hfill (29)

Note that $(\hat{y}_t - \hat{y}_0)$ is the amount that the model “compensates” by insulin infusion when the bias is constant (see Eq. 3). The results of applying Eq. 29 to Subject 11 are shown in Fig. 7 for the full two weeks of data. The plot on the left is $y_{\text{NAC},t}$ and the one on the right is $y_{\text{FFC},t}$. As shown, the variability is substantially reduced, with the standard deviations of $y_{\text{NAC},t}$ ($\sigma_{\text{NAC}}$) and $y_{\text{FFC},t}$ ($\sigma_{\text{FFC}}$) equal to 73.9 mg/dL and 8.7 mg/dL, respectively, that is a decrease of 88.3%. Thus, this level of the fitted model has the potential to greatly tighten BGC in an automatic FFC system. Also note how stable the response is in maintaining variation about the target of $Y^{\text{set}} = 140$ mg/dL. This stability supports the adaptive effectiveness of the algorithm to estimate model bias at each time instant over long periods of time.
Simulation Study of the Control Law

The Process

The process to for this part is a mathematical model for a CSTR from Smith and Corripio [34] and shown in Fig. 8. A dynamic model of the process was developed using first principles on the mass, species and energy balances of the process for the generating artificial data. Conditions included the following: constant densities and heat capacities of the tank and jacket contents; constant volumes in the tank and jacket; perfect mixing in the tank; negligible thermal capacitance of the tank wall and jacket wall; for this liquid system constant volume heat capacities \( c_v, c_{vc} \) and constant pressure heat capacities at \( c_p, c_{pc} \) are approximately equal and constant for both the reactor and jacket contents, respectively; and the energy due to flow streams is adequately described by the enthalpy. A list of variables and initial values are given in Table 8.

Fig. 7. The potential reduction in BGC variability from use of the proposed approach. The plot of the left represents the two weeks of BGC data for Subject 11 and plots \( y_{NAC,t} \) versus time in this analysis. The plot on the right is \( y_{FFC,t} \) versus time in this analysis for the fitted model in Table 4 for Subject 11.
The Coupled Network

For this study, this process has two loads \( F \) and \( C_{Ai} \), three state variables \( T, T_j, \) and \( C_A \), where \( T \), the controlled variable, is measured, \( T_j \) is not measured but estimated (i.e., identified), and \( C_A \) is not measured or estimated. We chose to identify \( T_j \) in the model as it is coupled with \( T \) and is directly affected by the manipulated variable just as BIC directly affects BGC and is directly affected by the manipulated variable in this process, i.e., \( M \). The coupled network for this system that we developed to model this process is given in Fig. 9. Note that, since this network is coupled, it is more complex that the semi-coupled network of Fig. 2.

Fig. 8. A flow diagram for the CSTR process.

The CSTR process is described mathematically by Eqs. 30-34 below:

\[
\frac{dC_A}{dt} = \frac{F}{V} (C_{Ai} - C_A) - kC_A^2 \tag{30}
\]

\[
\frac{dT}{dt} = \frac{F}{V} (T_T - T_T) - \frac{\Delta H_R}{\rho C_p} kC_A^2 - \frac{UA}{\rho V C_p} (T_T - T_C) \tag{31}
\]
\[
\frac{dT_C}{dt} = \frac{UA}{V_c \rho_c C_p C_T} (T_T - T_C) - \frac{F}{V_c} (T_C - T_{C_i})
\] (32)

\[
k = k_0 \exp \left( -\frac{E}{R(T_r + 273.16)} \right)
\] (33)

\[
F_C = F_{C_{\text{max}}} \times a^{(-M)}
\] (34)

Table 8. Definition of variables and initial steady-state values for the CSTR process.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>SS value (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Heat transfer area</td>
<td>5.40 (m²)</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>Control valve rangeability parameter</td>
<td>50 (none)</td>
</tr>
<tr>
<td>(C_A)</td>
<td>Concentration of species A in reactor</td>
<td>1.0302 (kgmol/m³)</td>
</tr>
<tr>
<td>(C_{A_i})</td>
<td>Concentration of species A in inlet stream</td>
<td>2.88 (kgmol/m³)</td>
</tr>
<tr>
<td>(c_p)</td>
<td>Heat capacity of feed and product streams</td>
<td>1.815x10⁵ (J/kgmol-°C)</td>
</tr>
<tr>
<td>(c_{pc})</td>
<td>Heat capacity of coolant</td>
<td>4184 (J/kg-°C)</td>
</tr>
<tr>
<td>(\Delta H_R)</td>
<td>Heat of reaction</td>
<td>-9.86x10⁷ (J/kgmol)</td>
</tr>
<tr>
<td>E</td>
<td>Activation energy</td>
<td>1.182x10⁷ (J/kgmol)</td>
</tr>
<tr>
<td>F</td>
<td>Feed flow rate</td>
<td>0.45 (m³/s)</td>
</tr>
<tr>
<td>(F_C)</td>
<td>Coolant flow rate</td>
<td>0.44 (m³/s)</td>
</tr>
<tr>
<td>(F_{C_{\text{max}}})</td>
<td>Maximum flow rate of coolant through control valve</td>
<td>1.2 (m³/s)</td>
</tr>
<tr>
<td>K</td>
<td>Reaction rate constant</td>
<td>0.09 (m³/s-kgmol)</td>
</tr>
<tr>
<td>(k_o)</td>
<td>Arrhenius frequency parameter</td>
<td>0.0744 (m³/s-kgmol)</td>
</tr>
<tr>
<td>M</td>
<td>Input signal to the valve</td>
<td>0.26 (none)</td>
</tr>
<tr>
<td>R</td>
<td>Gas law constant</td>
<td>8314.39 (J/kgmol-K)</td>
</tr>
<tr>
<td>(\rho)</td>
<td>Density of reactor contents</td>
<td>19.2 (kgmol/m³)</td>
</tr>
<tr>
<td>(\rho_c)</td>
<td>Density of coolant</td>
<td>1000 (kg/m³)</td>
</tr>
<tr>
<td>(T_c)</td>
<td>Coolant temperature in the jacket</td>
<td>50.48 (°C)</td>
</tr>
<tr>
<td>(T_{C_i})</td>
<td>Coolant inlet temperature</td>
<td>27 (°C)</td>
</tr>
<tr>
<td>T</td>
<td>Reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>(T_m)</td>
<td>Measured reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>U</td>
<td>Overall heat transfer coefficient</td>
<td>2.13x10⁵ (J/s-m²-°C)</td>
</tr>
<tr>
<td>(V_C)</td>
<td>Cooling jacket volume</td>
<td>1.82 (m³)</td>
</tr>
<tr>
<td>V</td>
<td>CSTR volume</td>
<td>7.08 (m³)</td>
</tr>
</tbody>
</table>

From Fig. 9, we get:

\[
T_i = f_1(V_{i,t}) = a_{D_i}v_{i,t} + a_{F_i}v_{F,t} + a_{C_{A_i}}v_{C_{A_i},t}
\] (35)

\[
T_{j_i} = f_2(V_{j,t}) = a_{M_i}v_{M,t} + a_{F}v_{F,t}
\] (36)
The Fitted Model

For the three inputs, M, F and $C_{Ai}$, the training data were generated using sequential step tests from a Box-Behnken design with one center point. This gave a total of 13 sequential step tests (100 seconds each). This input sequence is shown in Fig. 10, along the measured response of reactor temperature. Measurement noise was added to the true reactor temperature, $T$, according to Eq. 12 with $\sigma = 0.20 \, ^{\circ}C$, which is quite substantial as revealed in Fig. 10.

![Block diagram of the coupled network of the model for CSTR.](image)

**Fig. 9.** Block diagram of the coupled network of the model for CSTR.
Fig. 10. The input training sequences plotted on the left and the fitted and measured tank temperature responses plotted on the right.

During the fitting process, the dynamics of the $T$ and $T_j$ were found to be negligible allowing the fitted equations of Eqs. 35 and 36 to be expressed as

$$\hat{T}_t = \hat{a}_T \hat{T}_j + \hat{a}_F \hat{F}_{t-\Delta t} + \hat{a}_{CAi} \hat{C}_{Ai,t}$$

(37)

$$\hat{T}_j = \hat{b}_M \hat{v}_{M,t} + \hat{b}_T \hat{T}_{t-\Delta t}$$

(38)

Extending Eq. 37 to a second order regression structure to address the nonlinear static behavior of this process and from substitution of Eq. 38 into Eq. 37 gives:

$$\hat{T}_t = \hat{a}_0 + \hat{a}_T \left( \hat{b}_M \hat{v}_{M,t} + \hat{b}_T \hat{T}_{t-\Delta t} \right) + \hat{a}_F \hat{v}_{F,t} + \hat{a}_{CAi} \hat{v}_{CAi,t} + \hat{a}_{MP} \hat{v}_{M,t} \hat{v}_{F,t} + \hat{a}_{MCAi} \hat{v}_{M,t} \hat{v}_{CAi,t} + \hat{a}_{FF} \hat{v}_{F,t}^2 + \hat{a}_{CAiCAi} \hat{v}_{CAi,t}^2$$

(39)

Second order dynamics, based on Eq. 5, was modeled for the transfer functions for $F$ and $C_{Ai}$ and $M$ was modeled using first order dynamics leading to the discrete-time derivation of $\hat{v}_{M,t}$ as shown below:
To fit Eq. 39 the dynamic parameters (using Eqs. 5-10 and 40) and the static parameters (using Eq. 39) were estimated simultaneously using the method of nonlinear regression [33]. The fitted correlation coefficient for $T_m$ and $\hat{T}$ (i.e., $r_{fit}$) is 0.999. This excellent fit is also shown in Fig. 10.

Cross validation is unnecessary in this simulation study because there were no unmeasured disturbances and a statistical experimental design was used insuring causative modeling of the inputs (due to their orthogonality) on the output behavior. Nonetheless, a different test sequence was generated to evaluate the fit of the model to verify its testing ability. This sequence along with the fitted response plot is shown in Fig. 11. As expected, the fit is excellent. The estimated values of the model parameters are given in Table 9.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>-0.008</td>
<td>$a_{CAlCAl}$</td>
<td>0.164</td>
</tr>
<tr>
<td>$a_F$</td>
<td>2.009</td>
<td>$b_M$</td>
<td>3.815</td>
</tr>
<tr>
<td>$a_{Tf}$</td>
<td>0.654</td>
<td>$b_T$</td>
<td>1.406</td>
</tr>
<tr>
<td>$a_{CAi}$</td>
<td>1.347</td>
<td>$\zeta_F$</td>
<td>6.996</td>
</tr>
<tr>
<td>$a_{MF}$</td>
<td>-2.440</td>
<td>$\zeta_F$</td>
<td>1.022</td>
</tr>
<tr>
<td>$a_{MCAi}$</td>
<td>0.639</td>
<td>$\zeta_{Al}$</td>
<td>-3.907</td>
</tr>
<tr>
<td>$a_{FCAi}$</td>
<td>0.957</td>
<td>$\zeta_{CAi}$</td>
<td>6.004</td>
</tr>
<tr>
<td>$a_{MM}$</td>
<td>2.171</td>
<td>$\zeta_{CAi}$</td>
<td>1.363</td>
</tr>
<tr>
<td>$a_{FF}$</td>
<td>-4.868</td>
<td>$\tau_M$</td>
<td>13.339</td>
</tr>
</tbody>
</table>
Fig. 11. The input testing sequences plotted on the left and the fitted and measured tank temperature responses plotted on the right.

The temperature of the jacket contents, $T_j$, is not measured and used in the model and thus, the model is not fitting $T_j$. However, one of the purposes of this simulation study was to determine the importance of accurate predictions for the unmeasured state variable on the accuracy of the modeled stated variable that is measured. For the training and testing sets, the model predictions and the true values of $T_j$ are shown in Fig. 12. As shown, the predictions are not close to the true values quite often, especially for the training data. However, given the high degree of fit of the measured variable as shown in Figs. 10 and 11, the poor predictions of jacket temperature are not translating into poor fit of tank temperature. In relating this to the proposed BGC model, this suggests that for whatever the pseudo-BIC is physically representing, its accuracy can be poor and the fit of BGC can still be quite acceptable. Moreover, pseudo-BIC is playing a role to provide a structure and its accurate determination does not appear to be crucial.

The Feedback Controller (FBC)

A typical proportional-integral (PI) controller was implemented to control the reactor temperature. The manipulated variable chosen to maintain the reactor temperature is the coolant flow rate through the jacket of the CSTR vessel and is varied by changing the controller signal
sent to the valve, i.e., $M$. It was tuned to give the best possible response to the sequence of the two load changes in Fig. 11. For this controller $K_c = 1.40, \tau_l = 11.0$ and $M = M_{fb}$, where $M_{fb}$ is the signal from the PI controller to the valve. The response of the PI controller is shown in Fig. 13.

**Fig. 12.** The true and predicted responses of $T_j$ for the training data (left graph) and the testing data (right graph).

**Fig. 13.** Controller results; The FB controller response is on the left and the FBFF controller response is on the right. The addition of FF control of the proposed approach reduced the standard deviation of the controlled variable from set point by nearly 75%.

### The FBFF Controller

The coupled model was implemented into the FBFF controller in conjunction with the FB PI controller. Since $\hat{T}$ is a quadratic function as shown by Eq. 39, by application of the FFC law
given by Eq. 3, the roots were found at each time $t$ using the quadratic equation. The correct root at each time instant was

$$
\hat{v}_{M,t} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}
$$

where $a = \hat{a}_{MM}$, $b = \hat{a}_{TM} \hat{b}_M + \hat{a}_{MP} \hat{v}_{F,t} + \hat{a}_{MC,At} \hat{v}_{CAU}$, and

$$
c = \hat{a}_0 + \hat{a}_{TF} \hat{v}_{F,t} + \hat{a}_{CA} \hat{v}_{CAU} + \hat{a}_{FCM} \hat{v}_{F,t} \hat{v}_{CAU} + \hat{a}_{FF} \hat{v}_{F,t}^2 + \hat{a}_{CA,ICA} \hat{v}_{CAU}^2.
$$

By making $M_t$ the dependent variable in Eq. 40, $M_{ff}$ was obtained at each time instant from Eq. 42 below.

$$
M_{ff,t} = M_t = \hat{v}_{M,t} \frac{-\hat{v}_{M,t-N} + \hat{v}_{M,t}}{\Delta t} + \hat{v}_{M,t}
$$

Therefore, for this FBFF controller, $M = M_{fb} + M_{ff}$. The response of the FBFF controller to the same input sequence as the FB controller is also given in Fig. 13. As shown, the addition of this FFC system greatly reduced the deviations of the controlled variable, $T_m$, from its set point temperature of 88 °C. For the FB and FBFF controllers, the standard deviations of $T_m$ from its set point were 0.509 °C and 0.128 °C, respectively, or a reduction of 74.9%. Thus, the proposed coupled model FFC approach appears to be very promising.

Concluding Remarks

This work proposed a modeling methodology that extends the WMM [18] approach for multiple-input modeling of BGC in free-living, subject-specific, outpatient data collection. The performance objective of each model is to obtain the best cause-and-effect fit possible from the given set of inputs since they are being obtained for FFC. To meet this objective, a semi-coupled modeling network was developed that uses an unmeasured pseudo-blood insulin concentration
(i.e., latent) variable to allow the inclusion of the product BGC and BIC in the dynamic glucose balance.

In real data sets such as these where the model horizon covers a long time period, unmeasured disturbances are likely to be significantly correlated with measured inputs. In addition, this type correlation can change significantly from one period to the next. This characteristic challenges the ability to obtain accurate causative input modeling and can lead to significant over-fitting. To guard against the latter, we use the cross-validation procedure developed for the WMM [18] that decomposes dynamic and static modeling in a novel technique that results in similar fits between Training, Validation, and Testing data sets, as demonstrated by the results on the 15 case studies in that work. In this work the proposed approach was evaluated on the same data sets and maintained similarity in fit for all three data sets with a slight improvement over the WMM. Thus, we feel that the proposed CMM captured a large percent of the information on glucose behavior from the given measured input data sets. In this work, this translated into an $r_{fit}$ value that averaged from 0.60 to 0.65 with a high for Validation of 0.85 and a high for testing of 0.77. This level of fit has the potential to greatly nullify measured inputs and thereby, significantly reduce BGC as demonstrated on Subject 11 were the potential decrease is almost 90%.

The only way to truly know how effective a model developed under correlated inputs can be for FF control is to test it in actual control studies. The only way to do this at this stage of this work is in inpatient clinical studies under hospital observation and care. However, in this work we included an analysis to evaluate the physical soundness of the CMM results. This was first done by showing expected behavior with carbohydrate consumption without insulin infusion, and without carbohydrate consumption and with insulin infusion, which showed (in Fig. 5)
continual increase in BGC and continual drop in BGC to the zero level, respectively. In contrast, the WMM did not show this expected behavior (in Fig. 4). Another evaluation in this work consisted of examining the steady state relationship for the amount of insulin infusion at a given level of carbohydrate consumption, as given in Eq. 25 and shown in Fig. 6 for one of the subjects. These numbers are quite reasonable, and the relationships between BGC, insulin infusion and carbohydrate consumption are sound. Thus, when proposing modeling methods for use in FFC, it is our recommendation that these types of preliminary analyses be done to evaluate model soundness as a prior step to actually using them to build a FF controller.

To evaluate the coupled approach with an unmeasured state variable and its implementation into FFC via the control law given by Eq. 3, a simulation study was given using a mathematical CSTR. It was demonstrated that a complex network with a highly nonlinear model meeting the requirement of Eq. 3 can be effectively implemented in the time domain without linearization of the model. In addition, it appears that the estimates of the unmeasured pseudo-state variable do not have to be very accurate to estimate the controlled variable accurately. This means that accuracy of pseudo-BIC is not a concern. It is playing a role that allows a physiologically sound structure and the addition of pseudo-BIC did not cause a reduction in accuracy in controlled variable as compared to the WMM results. Finally, the simulation study demonstrated the effectiveness of the FFC approach by a 75% reduction in the standard deviation from the set point for the FBFF controller as compared to FB controller.

Acknowledgements

The authors would like to thank BodyMedia, Inc. of use of the SenseWear® Pro Body Monitoring System and Nisarg Vyas from BodyMedia, Inc. for his assistance with the study. Financial support from the National Institutes of Health (Grant 1R01DK085611) through Illinois
Institute of Technology and the Juvenile Diabetes Research Foundation (Grant 5-2010-7) is gratefully acknowledged.

**Literature Cited**


glucose control system for overnight glucose control in adolescents and young adults with type 1 diabetes, Diabetes Care, Nov, Vol.35(11)(2012), p.2182(6).


CHAPTER 7: BLOCK-ORIENTED FEEDFORWARD CONTROL WITH DEMONSTRATION TO NONLINEAR PARAMETRIZED WIENER MODELING

A paper published in Chemical Engineering Research and Design

Derrick K. Rollins\textsuperscript{1,2*}, Yong Mei\textsuperscript{1}, Stephanie D. Loveland\textsuperscript{1} and Nidhi Bhandari\textsuperscript{1}

\textsuperscript{1}Department of Chemical and Biological Engineering, Iowa State University, Ames, Iowa 50011
\textsuperscript{2}Department of Statistics, Iowa State University, Ames, Iowa 50011

*Corresponding Author

Abstract

Block-oriented modeling (BOM) is a multiple-input, multiple-output modeling approach for nonlinear dynamic processes. Current implementation of BOM into feedforward control (FFC) results in linearization of the model and decomposition into separate components for each input. This work presents a multiple-input BOM FFC approach that does not linearize and decompose the BOM into separate components for each input. This implementation uses a new FFC law that uses the complete BOM in the time domain. The approach is demonstrated with a Wiener model for a simulated continuous stirred tank reactor (CSTR) with four (4) measured inputs. The Wiener model is nonlinear in the physically-based dynamic parameters of the transfer functions and linear in the static parameters of the static gain function. The static gain function has a second order linear regression form with interaction and quadratic terms. The Wiener model is built under open-loop conditions using a Box-Behnken statistical experimental design consisting of 27 sequential step tests. Under a sequence of multiple input changes, the addition of this feedforward controller to the feedback controller reduced the standard deviation of the controlled variable from its set point by 70\% in comparison to the response with only feedback control.
Introduction

There are basically two types of control approaches – feedback control (FBC) and feedforward control (FFC). FBC is any control approach that determines the settings for the manipulated variable based on the deviation of the controlled variable from its target or set point ($y_{set}$). This deviation can be at the current time instant, as in common FBC or a more sophisticated one like a Smith Predictor, or it can be a predicted deviation at some future time as in model predictive control (MPC). FFC differs from FBC in that it changes the manipulated variable based on the values of input variables and not the deviation from set point. More specifically, the FFC control objective is to maintain the output of the FFC model, which consists of measured inputs only, at a constant value by changing the manipulated variable to “compensate” for changes in the measured inputs. Thus, to implement FFC effectively on the measured set of inputs, the model must be capable of accurate determination of the manipulated variable to “offset” the measured input changes while in an automatic control scheme. Approximations of the model to incorporate it into the control algorithm can also adversely affect control performance.

Some models for nonlinear gain behavior have been proposed for model-based controllers including radial basis functions (RBF) [4,5], genetic algorithms (GA) [6], Nonlinear Auto Regressive Models And eXogenous inputs (NARMAX) models [7-9], and block-oriented models (BOMs) [10-15]. An important advantage of NARMAX and BOMs is that they can use transfer functions, i.e., linear dynamic equations with physically interpretable parameters. However, a limitation of the NARMAX structure is that all of its transfer functions have the
same characteristic equation or denominator dynamics [2]. BOMs use the outputs from blocks of
dynamic (transfer) functions that are linear (L) differential equations as inputs to functions that
can be nonlinear (N) with respect to static gain parameters. The simplest of the BOMs is the
Hammerstein network (NL), which has an N block followed by an L block and the Wiener
network (LN), which reverses the order of these two blocks. More complicated block-oriented
structures include sandwich models such as an LNL network, which has linear dynamic blocks,
followed by a nonlinear static block, followed by a second linear dynamic block. When the
inputs can have different dynamic behavior, the Wiener network is the preferred choice over
Hammerstein and is superior to NARMAX because the inputs can have completely different
dynamic structures [2] as shown by its block diagram in Fig. 1.

A number of researchers have studied the identification of model parameters for the
Hammerstein and Wiener networks, including Greblicki [12,16], Eskinat et al. [17], Shi and Sun
[18], and Al-Duwaish and Naeem [6]. Perhaps, due to its complexity, the LNL network has not
gotten as much attention, but some have proposed methods for its parameter identification [19].
There has been much progress over the last decade in the identification of BOMs [19-25] and
recently, by taking a nonlinear parametrized approach [26] for estimation of the dynamic
parameters, Rollins et al. [2] demonstrated accurate Wiener modeling using nine (9) inputs on a
real distillation process with large variation due to unmeasured disturbances and with highly
pairwise cross-correlation of the inputs. While there has been progress in the use of BOMs in
model based control [27-29], progress of FFC using BOMs appears to have been limited to single
input models [30-31].

To implement a $p$-input nonlinear model structure the current FFC approach would
linearize the structure, transfer it to the Laplace domain and decompose it into $p$ FFC blocks.
Additional approximations would be required for FFC blocks that are physically unrealizable to make them physically realizable. A numerical procedure (which is another approximation) such as Euler’s method would be needed to obtain the output from each FFC block at each time instant. For the approach proposed in this work, a FFC law and methodology is presented that eliminates all these approximations. More specifically, the value of the manipulated variable to “offset” input changes at each time instant is obtained directly from the FFC model in the time domain in a root solving procedure and passing this solution through the inverse of the process transfer function. Since this procedure is done completely in the time domain the inverse of the process transfer function uses backward difference approximations for derivatives. Thus, an approximation to achieve a physically realizable transfer function is not needed for the proposed approach.

Thus, the objective of this work is the development of a general FFC framework for multiple-input BOMs in FFC with nonlinear static gain behavior. The BOM structure and identification will be demonstrated using a nonlinear parametrized Wiener model [1,2,26] with a second-order static gain structure on a simulated CSTR. The Wiener model is built under open-loop conditions using a Box-Behnken statistical experimental design consisting of 27 runs or sequential step tests. The FFC model will be implemented using the proposed FFC approach discussed above. Under a sequence of multiple input changes, the addition of this feedforward controller to the feedback controller reduced the standard deviation of the controlled variable from its set point by 70% in comparison to the response with only feedback control.

**Methodology**

This section gives the methods used to develop and evaluate the BOM FFC approach with specific application to a CSTR. The general BOM FFC law is presented first and then its
form specific to the Wiener modeling approach used in this work. This includes specification of the Wiener model (WM), its FFC law, and the parameter estimation approach. Lastly details of the CSTR are given in this section.

![Block diagram for the Wiener network with $p$ inputs and one output.](image)

**Fig. 1.** Block diagram for the Wiener network with $p$ inputs and one output. Each input, $x_i$, is passed through their own unity gain linear dynamic block, $G_i$, after which these unobservable intermediate outputs are collected and passed through a single unrestricted static gain function, $f(V)$, to produce the output, $y$.

**General FFC Law**

The concept of FFC appears to have been applied as early as 1925 to level control systems for boiler drums [32]. It allows for theoretically perfect control of a process system because it corrects for input disturbances before the process outputs deviate from their desired values. However, this requires timely and efficient measurement of all possible process disturbances, which is not likely in most applications, so it is commonly used in conjunction with FBC. The addition of FBC compensates for any deviation of the process output from its set point, regardless of the cause of the deviation. For each input, its FFC law is typically found independently of the other inputs from a function in the Laplace domain that uses an approximation to meet the requirement of physically reliability when necessary [32]. The joint FFC law for all the inputs is a sum of the individual ones, i.e., is linear and additive [33].
The FFC law that this work proposes for BOM is derived by modeling all inputs simultaneously without restrictions on linearity and additivity; it is also not necessary to work in the Laplace domain or to incorporate an approximation to meet the requirement of being physically realizable. The proposed FFC law is given as follows. Let \( x_{i,t} = \) the \( i \)th measured input at time \( t \), \( i = 1, \ldots, p \) and \( v_{i,t} = \) the output from the \( i \)th transfer function in Fig. 1 with \( X_t = [x_{1,t}, \ldots, x_{p,t}]^T \) and \( V_t = [v_{1,t}, \ldots, v_{p,t}]^T \) with the manipulated variable as \( x_{p,t} \). With the process initially at steady state and at \( Y_{\text{set}} \), and with no model bias, a general BOM FFC law, when the first block is N (e.g., Hammerstein model), is given as

\[
0 = f(X_t; \hat{\theta}) - Y_{\text{set}}
\]

where \( f(X_t; \hat{\theta}) \) is an unrestricted function of input variables only (more specifically, no outputs) for the first N block and \( \hat{\theta} \) is the vector of estimated parameters. Note that “\(^\wedge\)” is used for estimate throughout this article. Thus, at each time instant, Eq. 1 determines the value of \( x_{p,t} \) (i.e., the output from the FFC system) for \( f(X_t; \hat{\theta}) \) to remain at \( Y_{\text{set}} \). Similarly, with the process initially at steady state and at \( Y_{\text{set}} \), and with no model bias, a general BOM FFC law, when the first block is L (e.g., a Wiener network), is given as

\[
0 = f(V_t; \hat{\theta}) - Y_{\text{set}}
\]

such that

\[
x_{p,t} = G_p^{-1}(\hat{v}_{p,t})
\]

Thus, at each time instant, Eq. 2 determines the value of \( \hat{v}_{p,t} \) for \( f(V_t; \hat{\theta}) \) to remain at \( Y_{\text{set}} \) and then determines the output from the FFC system, \( x_{p,t} \), from Eq. 3. Note that this FFC law does not require linearization of \( f(V_t; \hat{\theta}) \) or its transformation to the Laplace domain and back to the time domain to obtain \( \hat{v}_{p,t} \). Furthermore, for discrete-time approximation with a sampling time
of $\Delta t$, $x_{p,t}$ can be obtained from Eq. 3 by inverting the discrete time approximation function of the differential equation and using backward difference approximations for derivatives. For example, if $G_p$ is a first order process, then

$$\tau_p \frac{dy_p(t)}{dt} + y_p(t) = \hat{\tau}_p \frac{\hat{y}_{p,t} - \hat{y}_{p,t-\Delta t}}{\Delta t} + \hat{\nu}_{p,t} = x_{p,t}$$

(4)

For a simple process with two inputs and linear gains ($K_1$ and $K_2$) such as the ones commonly appearing in process control textbooks [32], Eq. 2 can be shown to be in agreement with these textbooks as follows:

$$Y_{set} + K_1 v_1(t) + K_2 v_2(t) - Y_{set} = K_1 v_1(t) + K_2 v_2(t) = 0$$

$$\Rightarrow K_1 V_1(s) + K_2 V_2(s) = K_1 X_1(s) G_1(s) + K_2 X_2(s) G_2(s) = 0$$

$$\Rightarrow \frac{X_2(s)}{X_1(s)} = -\frac{K_1 G_1(s)}{K_2 G_2(s)}$$

(5)

As shown, Eq. 5 agrees with Eq. 15-21 in Seborg et al. [32] since $G_2(s)$ is a product of the load transmitter, valve and process transfer functions and gains of $G_1(s)$ and $G_2(s)$ are one here. A general BOM feedback feedforward (FBFF) block diagram representing this approach is given Fig. 2.

**The Wiener Model FFC Law**

This article applies and evaluates the discrete-time Wiener modeling approach of Rollins et al. [26] that was evaluated modeling the blood glucose levels of real subjects with type 2 [26,34] and type 1 [1,3] diabetes and the top tray temperature of a real pilot distillation column [2]. This Wiener model (WM) structure for $G_i$ in Fig. 1 is a second-order-plus-lead-plus-dead-time (SOPLPDT) form that estimates the physically-based dynamic parameters of the $v_i$’s
directly with a highly non-linear structure including physical constraints. The form of this differential equation (dead-time is excluded for simplicity) is

\[ \tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2 \tau_i \xi_i \frac{d v_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{d x_i(t)}{dt} + x_i(t) \]  \hspace{1cm} (6)

Fig. 2. A general BOM FBFF block diagram shown with \( m \) loads and \( p \) FFC variables. If the first block is N, Eq. 1 is applicable, the output of the BOM block is \( x_p \) and \( G_{FFC} = 1 \). If the first blocks are L, Eq. 2 is applicable, the output of the BOM block is \( v_p \) and \( G_{FFC} = G_p^{-1} \).

where \( i = 1, \ldots, p, p \) is the total number of inputs (\( x_i \)'s), \( \tau_i \) is the time constant, \( \xi_i \) is the damping coefficient, and \( \tau_{ai} \) is the lead parameter. Dead time is not shown in Eq. 6 since the process simulation in this work has no dead time. In a parameter estimation problem such as this one, a discrete-time approach has the advantage over a continuous-time approach of the not needing to change form as the dampening coefficient changes during parameter search process [26].

Furthermore, even after obtaining the set of parameters, implementation of a continuous-time function will be difficult because inputs change at each sampling instant that will result in a very complex input sequence that will require truncation after some time in the past and thus, will also
be an approximation. Thus, unless a continuous-time method can be justified on the basis of significantly better accuracy, a discrete-time method is preferred in this application. 

Using a backward difference approximation \( \left( \text{e.g., } \frac{dV_i(t)}{dt} \approx \frac{V_{i,t} - V_{i,t-\Delta t}}{\Delta t} \right) \), the following approximate discrete-time form of Eq. 6 is obtained [4]:

\[
v_{i,t} = \delta_{i,1} V_{i,t-\Delta t} + \delta_{i,2} V_{i,t+2\Delta t} + \omega_{i,1} x_{i,t-\Delta t} + \omega_{i,2} x_{i,t+2\Delta t}
\]  

(7)

where

\[
\delta_{i,1} = \frac{2 \tau_i^2 + 2 \tau_i \zeta_i \Delta t}{\tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(8)

\[
\delta_{i,2} = \frac{-\tau_i^2}{\tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(9)

\[
\omega_{i,1} = \frac{(\tau_i + \Delta t)\Delta t}{\tau_i^2 + 2 \tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(10)

and

\[
\omega_{i,2} = 1 - \delta_{i,1} - \delta_{i,2} - \omega_{i,1}
\]  

(11)

to satisfy the constraint of unity gain. Two additional physical constraints are \( \tau_i > 0 \) and \( \zeta_i > 0, \forall i \). Note that Eq. 7 is linear in the \( \delta \)'s and \( \omega \)'s. Thus, Eq. 7 represents a linear regression form with the \( \delta \)'s and \( \omega \)'s estimated directly. Notwithstanding, the nonlinear regression form is obtained by the substitution of Eqs. 8-11 into Eq. 7 and estimating the \( \tau \)'s, \( \zeta \)'s, and \( \tau_a \)'s (i.e., the physically-based dynamic parameters) directly with \( v_{i,t} \) given as a highly nonlinear function of the estimated parameters. Physical constraints also strengthen estimation by adding more true structure. Notably, these are: Eq. 6, \( \tau_i > 0 \) and \( \zeta_i > 0, \forall i \). Another advantage from using physically-based dynamic parameters is the ability to give sensible starting values for estimates.

The general discrete-time, “white noise,” WM is given as

\[
y_i - Y_i = y_i' = f(V_i) + \epsilon_i = \eta_i - Y_i + \epsilon_i
\]  

(12)

where
\[ e_i \sim N(0, \sigma^2) \text{ independently } \forall t \]  

(13)

the common assumption with least squares estimation and \( y_i \) is the measured output at time \( t \).

Note that the expected response at time \( t \) is \( \eta_t = E[y_t] \), since \( E[\epsilon_i] = 0 \). While \( f(V_t) \) can be any function, in the work it is a second order linear regression function of \( p \) variables as shown below:

\[
\eta_t - Y^{\text{set}} = f(V_t) = a_0 + a_1 v_{1,t} + \cdots + a_{p-1} v_{p-1,t} + a_p v_{ff,t} + a_{11} v_{1,2,t} + \cdots
\]

(14)

\[ + a_{p-1,p} v_{p-1,p} v_{ff,t} + a_{11} v_{1,2,t}^2 + \cdots + a_{p-1,p} v_{p-1,p}^2 + a_{pp} v_{ff,t}^2 \]

where \( v_{p,t} = v_{ff,t} \). With \( f(\hat{V}_t; \hat{\theta}) = f(\hat{V}_t) + Y^{\text{set}} \) and Eqs. 12-14, Eq. 2 becomes

\[
f(\hat{V}_t; \hat{\theta})_v - Y^{\text{set}} = f(\hat{V}_t) = \hat{a}_0 + \hat{a}_1 \hat{v}_{1,t} + \cdots + \hat{a}_{p-1} \hat{v}_{p-1,t} + \hat{a}_p \hat{v}_{ff,t} + \hat{a}_{11} \hat{v}_{1,2,t} + \cdots
\]

(15)

\[ + \hat{a}_{p-1,p} \hat{v}_{p-1,p} \hat{v}_{ff,t} + \hat{a}_{11} \hat{v}_{1,2,t}^2 + \cdots + \hat{a}_{p-1,p} \hat{v}_{p-1,p}^2 + \hat{a}_{pp} \hat{v}_{ff,t}^2 \]

\[ = a \hat{v}_{ff,t}^2 + b \hat{v}_{ff,t} + c = 0 \]

where \( a = \hat{a}_{pp}, b = \hat{a}_{1p} \hat{v}_{1,t} + \cdots + \hat{a}_{p-1,p} \hat{v}_{p-1,t} \) and \( c = \hat{a}_0 + \hat{a}_1 \hat{v}_{1,t} + \cdots + \hat{a}_{p-1} \hat{v}_{p-1,t} + \hat{a}_{11} \hat{v}_{1,2,t} + \cdots \). For the CSTR in this work, a first order transfer function is used giving, with the solution to Eq. 15, the FFC system output as:

\[
\tau_p \frac{dv_{ff}(t)}{dt} + v_{ff}(t) = x_{ff}(t)
\]

(16)

\[ \Rightarrow \hat{x}_{ff,t} = \tau_p \frac{\hat{v}_{ff,t} - \hat{v}_{ff,t-\Delta t}}{\Delta t} + \hat{v}_{ff,t} = M_{ff} \]

Note that for estimation of \( \tau_p \), a backwards differences approximation is used and in this case Eq. 16 is written as
\[
\hat{\mathbf{E}}_p \left( \hat{\mathbf{v}}_{ff,i} - \hat{\mathbf{v}}_{ff,i-\Delta t} \right) + \hat{\mathbf{v}}_{ff,i} = \mathbf{x}_{ff,i-\Delta t} \Delta t \\
\Rightarrow \hat{\mathbf{v}}_{ff,i} = \frac{\hat{\mathbf{E}}_p}{\hat{\mathbf{E}}_p + \Delta t} \hat{\mathbf{v}}_{ff,i-\Delta t} + \frac{\Delta t}{\hat{\mathbf{E}}_p + \Delta t} \mathbf{x}_{ff,i-\Delta t}
\]

(17)

Thus, in estimating the model parameters, Eq. 17 is used and in determining the output of the FFC system, Eq. 16 is used. The FBFF block diagram specific to this WM approach is shown in Fig. 3.

Fig. 3. A Wiener FBFF block diagram based on Fig. 2 with \( m \) loads and \( p \) FFC variables; \( e = \) estimate, e.g., \( \mathbf{v}^e_p = \hat{\mathbf{v}}_{pp,i} = \mathbf{v}_{ff,i}^e \).

The Mathematical Model for the CSTR

The process to evaluate the WM FFC method is a simulated CSTR from Smith and Corripio [35] and shown in Fig. 4. A dynamic model of the process was developed using first
principles on the mass, species and energy balances of the process for the simulation. The assumptions made included the following: constant densities and heat capacities of the tank and jacket contents; constant volumes in the tank and jacket; perfect mixing in the tank; negligible thermal capacitance of the tank wall and jacket wall; for this liquid system constant volume heat capacities \( (c_v, c_{vc}) \) and constant pressure heat capacities at \( (c_p, c_{pc}) \) are approximately equal and constant for both the reactor and jacket contents, respectively; and the energy due to flow streams is adequately described by the enthalpy. A list of variables and initial values is given in Table 1.

The CSTR process is described mathematically in Eqs. 18-22 below:

\[
\frac{dC_A}{dt} = \frac{F}{V} (C_{Ai} - C_A) - \frac{k}{V} C_A^2
\]

\[
\frac{dT}{dt} = \frac{F}{V} (T_i - T) - \frac{\Delta H_B}{\rho C_p} k C_A^2 - \frac{UA}{\rho V C_p} (T - T_C)
\]

\[
\frac{dT_C}{dt} = \frac{UA}{V_c \rho C_{pc}} (T - T_C) - \frac{F_C}{V_c} (T_C - T_{ci})
\]

\[
k = k_0 \exp \left( \frac{-E}{R(T + 273.16)} \right)
\]

\[
F_C = F_{C_{max}} \times \alpha (-M)
\]
Table 1. Definition of variables and initial steady-state values for the CSTR process.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>SS value (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Heat transfer area</td>
<td>5.40 (m$^2$)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Control valve rangeability parameter</td>
<td>50 (none)</td>
</tr>
<tr>
<td>C$_A$</td>
<td>Concentration of species A in reactor</td>
<td>1.0302 (kgmol/m$^3$)</td>
</tr>
<tr>
<td>C$_{A,i}$</td>
<td>Concentration of species A in inlet stream</td>
<td>2.88 (kgmol/m$^3$)</td>
</tr>
<tr>
<td>$c_p$</td>
<td>Heat capacity of feed and product streams</td>
<td>1.815x10$^7$ (J/kgmol)</td>
</tr>
<tr>
<td>$c_{pc}$</td>
<td>Heat capacity of coolant</td>
<td>4184 (J/kgmol-°C)</td>
</tr>
<tr>
<td>$\Delta H_R$</td>
<td>Heat of reaction</td>
<td>-9.86x10$^7$ (J/kgmol)</td>
</tr>
<tr>
<td>E</td>
<td>Activation energy</td>
<td>1.182x10$^7$ (J/kgmol)</td>
</tr>
<tr>
<td>F</td>
<td>Feed flow rate</td>
<td>0.45 (m$^3$/s)</td>
</tr>
<tr>
<td>F$_C$</td>
<td>Coolant flow rate</td>
<td>0.44 (m$^3$/s)</td>
</tr>
<tr>
<td>F$_{C,max}$</td>
<td>Maximum flow rate of coolant through control valve</td>
<td>1.2 (m$^3$/s)</td>
</tr>
<tr>
<td>K</td>
<td>Reaction rate constant</td>
<td>0.09 (m$^3$/s-kgmol)</td>
</tr>
<tr>
<td>$k_o$</td>
<td>Arrhenius frequency parameter</td>
<td>0.0744 (m$^3$/s-kgmol)</td>
</tr>
<tr>
<td>M</td>
<td>Input signal to the valve</td>
<td>0.26 (none)</td>
</tr>
<tr>
<td>R</td>
<td>Gas law constant</td>
<td>8314.39 (J/kgmol-K)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Density of reactor contents</td>
<td>19.2 (kgmol/m$^3$)</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>Density of coolant</td>
<td>1000 (kg/m$^3$)</td>
</tr>
<tr>
<td>T$_j$</td>
<td>Coolant temperature in the jacket</td>
<td>50.48 (°C)</td>
</tr>
<tr>
<td>T$_{CI}$</td>
<td>Coolant inlet temperature</td>
<td>27 (°C)</td>
</tr>
<tr>
<td>T</td>
<td>Reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>T$_{m}$</td>
<td>Measured reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>U</td>
<td>Overall heat transfer coefficient</td>
<td>2.13x10$^5$ (J/s-m$^2$-°C)</td>
</tr>
<tr>
<td>V$_C$</td>
<td>Cooling jacket volume</td>
<td>1.82 (m$^3$)</td>
</tr>
<tr>
<td>V</td>
<td>CSTR volume</td>
<td>7.08 (m$^3$)</td>
</tr>
</tbody>
</table>
Wiener Modeling Results

This study consisted of the following four inputs: $F(x_1)$, $TCI(x_2)$, $CAi(x_3)$, and $M(x_4)$. The training data were generated using sequential step tests from a Box-Behnken design with three center points and four inputs. This gave a total of 27 sequential step tests (100 s each) or times that input changes occurred. This input sequence is shown in Fig. 5, and the response of reactor temperature to this series of input changes is given in Fig. 6. Measurement noise was added to the true reactor temperature, $T$, according to Eq. 13 with $\sigma = 0.03$ °C.

![Fig. 5. The input training sequences (a) and the fitted and measured tank temperature response (b).](image)

The final form of the reduced fitted model is given by Eq. 23 below.

$$f(\mathbf{\hat{v}}_t) = \hat{a}_0 + \hat{a}_1\mathbf{v}_{1,t} + \hat{a}_2\mathbf{v}_{2,t} + \hat{a}_3\mathbf{v}_{3,t} + \hat{a}_4\mathbf{v}_{4,t} + \hat{a}_{12}\mathbf{v}_{1,t}\mathbf{v}_{2,t} + \hat{a}_{13}\mathbf{v}_{1,t}\mathbf{v}_{3,t} + \hat{a}_{14}\mathbf{v}_{1,t}\mathbf{v}_{4,t} + \hat{a}_{23}\mathbf{v}_{2,t}\mathbf{v}_{3,t} + \hat{a}_{24}\mathbf{v}_{2,t}\mathbf{v}_{4,t} + \hat{a}_{34}\mathbf{v}_{3,t}\mathbf{v}_{4,t} + \hat{a}_{11}\mathbf{v}_{1,t}^2 + \hat{a}_{22}\mathbf{v}_{2,t}^2 + \hat{a}_{33}\mathbf{v}_{3,t}^2 + \hat{a}_{44}\mathbf{v}_{4,t}^2$$  \hspace{1cm} (23)

To fit Eq. 23 the dynamic parameters (using Eqs. 6-11 and 17) and the static parameters (using Eq. 23) were estimated simultaneously using the method of nonlinear regression [36]. For a static model parameter to be retained, its individual $P$-value had to be less than 0.05. Two parameters
did not meet this criterion, $a_{23}$ and $a_{22}$, as shown by comparing Eq. 15 with Eq. 23. $R^2$, which measures the amount of “explained variation,” is 99.9% for the static model. All the retained parameter estimates are given in Table 2. The fitted correlation coefficient for $T_m$ and $\hat{T}_m$ (i.e., $r_{fit}$) for Eq. 23 is 0.998. This excellent fit is shown in Fig. 5.

The input sequences for the test set are given in Fig. 6 along with a fit of the model. As shown, the fit is also excellent to this test sequence.

**The FB Controller**

A typical proportional-integral (PI) controller was implemented to control the reactor temperature. The manipulated variable chosen to maintain the reactor temperature is the coolant flow rate through the jacket of the CSTR vessel and is varied by changing the controller signal sent to the valve, i.e., $M$. It was tuned to give the best possible response to the sequence of the three load changes in Fig. 6. For this controller $K_c = 1.40, \tau_i = 11.0$ and $M = M_{fb}$, where $M_{fb}$ is the signal from the PI controller to the valve. The response of the PI controller is shown in Fig. 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>0.005</td>
<td>$a_{33}$</td>
<td>0.678</td>
</tr>
<tr>
<td>$a_1$</td>
<td>26.027</td>
<td>$a_{44}$</td>
<td>26.758</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.337</td>
<td>$\tau_1$</td>
<td>7.1</td>
</tr>
<tr>
<td>$a_3$</td>
<td>16.287</td>
<td>$\zeta_1$</td>
<td>1.054</td>
</tr>
<tr>
<td>$a_4$</td>
<td>30.507</td>
<td>$\tau_{a1}$</td>
<td>-5.02</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.506</td>
<td>$\tau_2$</td>
<td>5.0</td>
</tr>
<tr>
<td>$a_{1,3}$</td>
<td>7.816</td>
<td>$\zeta_2$</td>
<td>1.384</td>
</tr>
<tr>
<td>$a_{1,4}$</td>
<td>-33.405</td>
<td>$\tau_3$</td>
<td>7.2</td>
</tr>
<tr>
<td>$a_{2,4}$</td>
<td>-0.338</td>
<td>$\zeta_3$</td>
<td>1.096</td>
</tr>
<tr>
<td>$a_{34}$</td>
<td>8.276</td>
<td>$\tau_4$</td>
<td>14.5</td>
</tr>
<tr>
<td>$a_{11}$</td>
<td>-64.756</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The FBFF Controller

The WM was implemented into the FBFF controller in conjunction with the FB PI controller. Since \( f(V_t) \) is a quadratic function as shown by Eq. 15, the roots were found at each time \( t \) using the quadratic equation. The correct root at each time instant was

\[
\hat{\dot{y}}_{ff,t} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}
\]

(24)

Using the result from Eq. 24, \( M_{ff} \) was obtained from Eq. 16. Therefore, for this FBFF controller, \( M = M_{fb} + M_{ff} \). The response of the WM FBFF controller to the same input sequence as the FB controller is also given in Fig. 7. As shown, the addition of this FFC system greatly reduced the deviations of the controlled variable, \( T_m \), from its set point temperature of 88 °C. For the FB and FBFF controllers, the standard deviations of \( T_m \) from its set point were 0.523 °C and 0.158 °C, respectively, or a reduction of 69.8%. Thus, the proposed WM FFC approach appears to have great promise as an effective FFC approach.
As discussed above the inverse of the model for Eq. 16 is not physically realizable. Consequently, we also evaluated this approach using the following approximation (where $\tau_{av}$ is a small number relative to $\tau_p$) which is physically realizable.

$$\tau_p \frac{dv_{ff}(t)}{dt} + v_{ff}(t) = \tau_{av} \frac{dx_{ff}(t)}{dt} + x_{ff}(t)$$  \hspace{1cm} (25)

We evaluated Eq. 25 using Euler’s method to solve for $x_{ff,t}$ at each time instant with $\tau_{av} = 0.09$.

The standard deviation of $T_m$ from its set point with this approach was found to be 0.244 °C or a reduction of 52.3% from the FBC system, which is significantly smaller than the result using Eq. 16 of 69.8%.

**Concluding remarks**

In this work, we have proposed a new approach for BOM FFC. This approach is represented graphically in a block diagram (Fig. 2) and mathematically by Eqs. 1-3. To
determine the FF controller output signal, all the inputs are used simultaneously in the complete nonlinear structure of the FFC model, linearization of the model is not necessary or transformation into the Laplace domain, as in the current approach with multiple inputs.32. Moreover, the FFC law of this approach is rather general and should be applicable to FFC models of any type not just BOM.

The primary objective was proposal of the new BOM FFC approach, a secondary objective was its demonstration using the nonlinear parameterized Wiener modeling approach of Rollins et al.26 on a simulated CSTR. A FBFF block diagram specific to a Wiener network was derived from the general one in Fig. 2 and given in Fig. 3. As Fig. 3 shows, the WM produces an estimate of $v_{yf}$ and another step is needed to determine $x_{yf}$. In this work solving the discrete-time equation for the process transfer function for $x_{yf}$ in terms of $v_{yf}$ resulted in a greater reduction (69.8%) in the standard deviation from set point relative to the FB controller than an approach to approximate this transfer function to meet the physically realizable criterion (52.3%). More work will be needed to verify the superiority of this approach over the common practice of adding terms in the denominator to meet the physical realizable criterion. Nonetheless, for this example this FFC approach reduced the variation from set point substantially and appears to be a viable contribution to model-based control methodologies.

When model bias can change over time, its value will need be approximated at each time instant. Modeling real data using this WM approach has shown that the level of model bias is likely to persist for several times periods [1,2,26] due to shifts in unmeasured disturbances. When this is the case, model bias at each time $t$ should be approximated well by $y_{t-\Delta t} - \hat{y}_{t-\Delta t}$.

Future work would also consist of applying the proposed BOM FFC approach to real processes. Our plan in the near future is to apply this work to the distillation process in Rollins et
al. [2] and in development of FFC for disturbances that affect the blood glucose levels of people that rely on externally infused insulin for control of their blood sugar [1].

**Literature Cited**


CHAPTER 8: A FEEDFORWARD CONTROL APPROACH WHEN THE TIME DELAY OF THE MANIPULATED VARIABLE IS GREATER THAN THE INPUT VARIABLE

A paper to be submitted to Chemical Engineering Research and Design

Derrick K. Rollins, Sr. \textsuperscript{1,2} and Yong Mei\textsuperscript{1}

\textsuperscript{1}Department of Chemical and Biological Engineering, Iowa State University, Ames, IA;  
\textsuperscript{2}Department of Statistics, Iowa State University, Ames, IA

\textsuperscript{*}Corresponding Author

\textbf{Abstract}

Feedforward control (FFC) is the only theoretically perfect control scheme. Potentially, it has the ability for tight control in real applications. However, when the dead time of the manipulated variable is greater than a modeled load disturbance (i.e., measured input), the result is a FFC algorithm that is physically unrealizable since it requires future information of the input. Notwithstanding, when future behavior of these types of inputs is available, FFC is possible without approximations to achieve a physically unrealizable controller. This work presents a multiple-input FFC framework with this capability. The motivation for this framework is the development of an automatic insulin delivery device for the control of blood glucose concentration (BGC) for people with type 1 diabetes. In this application, the manipulated variable, the exogenous insulin feed rate, has a large dead time that is likely to be greater than that of any measured variable that is a candidate for FFC. This novel approach is demonstrated using a multiple-input mathematically simulated continuous-stirred-tank-reactor (CSTR).

\textbf{Key words:} feedforward control, time delay, \textit{in silico}, closed-loop control, artificial pancreas, CSTR
**Introduction**

There is great need of tight control in Blood Glucose Concentration (BGC) for people with diabetes. The artificial pancreas (AP) is a tool that was created to help insulin-dependent people with diabetes to automatically control their BGC. The AP has the following control components: continuous glucose sensor, insulin pump, and control algorithm. The control algorithm takes input information such as measured BGC, processes it through an algorithm that manipulates the insulin feed rate (IFR) with the control object to minimize deviation from the BGC target called the “set point”. For the AP to achieve the tight control necessary for long term, i.e., 24 hour a day control, the control algorithm must significantly advance.

Current AP algorithms include traditional feedback control (FBC) [1], model predictive control (MPC) [2-11], fuzzy logic control (FLC) [12,13], and Feedforward control (FFC) [14-16]. Although researchers have made considerable advancements in inpatients/outpatients BGC closed-loop control, daytime automatic control has only marginally improved over manual subject control. As sensor technology continues to improve for disturbances that affect BGC, with effective mapping of these disturbances to BGC behavior (i.e., modeling), the potential to greatly tighten BGC using model-based control algorithms increases. The potentially most powerful control algorithm is FFC due to its proactive theoretically perfect property. However, to develop an effective FFC system in this application, at least four major challenges need to be overcome. The first one is the development of an effective multiple-input modeling method with strong cause and effect input relationships to the output, BGC. While improvements are still needed, our research team has made considerable progress in recent years [17,23]. The second one is the development of an effective multiple-input FFC algorithm. We feel that we have addressed this challenge adequately in Rollins et al [18]. The third one is the advancement of
sensor technology for measuring the critical inputs such as meals, activity and stress. Sensor technology has been rapidly improving in the last decade and many of these sensors currently exists or will likely to be invented in a few years [23,24]. The last one is a control algorithm that overcomes the physically unrealizable limitation of FFC models in this context without making approximations to the model [11, 13]. Overcoming these challenges will improve the model in cancelling out disturbances more effectively. Therefore, the purpose of this work is to propose a new FFC approach that addresses the physically unrealizable property due to dead time. In this approach, a predictive model is used to make current changes in the manipulated variable to effectively cancel future effects of input variables. We have termed this approach feedforward predictive control (FFPC). We demonstrate its effectiveness using the CSTR in Rollins et al [18]. While there are diabetes simulators that we could use in this study, they are all limited to just one input variable – meal size, as well as other limitations. Using the CSTR for this study allows the demonstration of FFPC algorithm in a multiple-input process study. In later work, we plan to evaluate our proposed FFPC approach using an FDA approved diabetes simulator.

Methodology

Simulated CSTR

A CSTR process is used in this initial development and evaluation of FFPC for the following reasons: 1. the currently available FDA approved UVA-padova simulator in type 1 diabetes does not support a multiple-disturbance control approach [20,21]; 2. it is easy to implement relatively long time delays for this process and; 3. the dynamics of a CSTR are well understood theoretically as the process model is developed completely from conservation laws of material and energy balances. Thus, the model can be used with confidence that it represents true
dynamic behavior of a chemical process. This CSTR process [18] is illustrated in Fig. 1 and described by the following equations:

\[
\frac{dC_A}{dt} = \frac{F}{V}(C_{Ai} - C_A) - kC_A^2 
\]

(1)

\[
\frac{dT_r}{dt} = \frac{F}{V}(T_r - T_T) - \frac{\Delta H_E}{\rho C_p}kC_A^2 - \frac{UA}{\rho V C_p}(T_r - T_c) 
\]

(2)

\[
\frac{dT_c}{dt} = \frac{UA}{V C_p C_T}(T_r - T_c) - \frac{F}{V_c}(T_c - T_{ci}) 
\]

(3)

\[
k = k_0 \exp\left(\frac{-E}{R(T_r + 273.16)}\right) 
\]

(4)

\[
F_c = F_{c_{max}} \times \alpha^{(-M)} 
\]

(5)

The definitions and values of the variables are shown in Table 1.

In this study, three input variables are used: input signal to valve, \(M(x_1)\), temperature of entering coolant, \(T_{ci}(x_2)\), and concentration of component A in inlet, \(C_{Ai}(x_3)\). The one output for this study is the measured tank temperature, \(T_m\). The CSTR is altered to have relatively long dead times \((\theta)\) in \(T_{ci}\), \(C_{Ai}\), and the input signal to valve \((M)\), specified as \(\theta_i = k_i \Delta t\), where \(k_i\) is a positive integer, \(i = 1, 2, 3\), for \(M\), \(T_{ci}\), and \(C_{Ai}\), respectively. The dead time for \(M\) is the largest.

**Modeling Methodology**

The Wiener network from [22] is used for modeling the CSTR process. This Multiple-Input, Single-Output (MISO) network is shown in Fig. 2. Each input \((x_i)\) has its own linear dynamic block, \(G_i\), and each dynamic block has an intermediate unobservable, output \(v_i\), which represents the independent dynamic response of its corresponding input. All the intermediate \(v_i\)'s are collected and passed through a nonlinear static gain block, \(f(V)\), to produce the final measured output, \(y\).
Fig. 1. Simulated CSTR process for this study.

The dynamic blocks are linear ordinary differential equations. \( M \) has first-order-plus-dead-time dynamics given by Eq. 6 below:

\[
\tau_1 \frac{dv_1(t)}{dt} + v_1(t) = x_i(t - \theta_1)
\]  

(6)

where \( \tau_1 \) is the time constant and \( \theta_1 \) is the dead time. Using a backward difference approximation \( \left( e.g., \frac{dv_1(t)}{dt} \approx \frac{v_{i+1} - v_{i-\Delta t}}{\Delta t} \right) \), applied to a sampling interval of \( \Delta t \), an approximate discrete-time form of Eq. 6 is:

\[
v_{i,j} = \delta_{i,1} v_{i,j-\Delta t} + \left(1 - \delta_{i,1}\right) x_{i,j-\Delta t}
\]  

(7)

where

\[
\delta_{i,1} = \frac{\tau_1}{\tau_1 + \Delta t}
\]  

(8)
**Table 1.** Definitions of variables and initial steady-state values for the CSTR process.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>SS value (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Heat transfer area</td>
<td>5.40 (m²)</td>
</tr>
<tr>
<td>α</td>
<td>Control valve rangeability parameter</td>
<td>50 (none)</td>
</tr>
<tr>
<td>CA</td>
<td>Concentration of species A in reactor</td>
<td>1.0302 (kgmol/m³)</td>
</tr>
<tr>
<td>CAi</td>
<td>Concentration of species A in inlet stream</td>
<td>2.88 (kgmol/m³)</td>
</tr>
<tr>
<td>c_p</td>
<td>Heat capacity of feed and product streams</td>
<td>1.815x10⁷ (J/kgmol-°C)</td>
</tr>
<tr>
<td>c_pc</td>
<td>Heat capacity of coolant</td>
<td>4184 (J/kg-°C)</td>
</tr>
<tr>
<td>ΔHR</td>
<td>Heat of reaction</td>
<td>-9.86x10⁷ (J/kgmol)</td>
</tr>
<tr>
<td>E</td>
<td>Activation energy</td>
<td>1.182x10⁷ (J/kgmol)</td>
</tr>
<tr>
<td>F</td>
<td>Feed flow rate</td>
<td>0.45 (m³/s)</td>
</tr>
<tr>
<td>FC</td>
<td>Coolant flow rate</td>
<td>0.44 (m³/s)</td>
</tr>
<tr>
<td>F_Cmax</td>
<td>Maximum flow rate of coolant through control valve</td>
<td>1.2 (m³/s)</td>
</tr>
<tr>
<td>K</td>
<td>Reaction rate constant</td>
<td>0.09 (m³/s·kgmol)</td>
</tr>
<tr>
<td>k_o</td>
<td>Arrhenius frequency parameter</td>
<td>0.0744 (m³/s·kgmol)</td>
</tr>
<tr>
<td>M</td>
<td>Input signal to the valve</td>
<td>0.26 (none)</td>
</tr>
<tr>
<td>R</td>
<td>Gas law constant</td>
<td>8314.39 (J/kgmol-K)</td>
</tr>
<tr>
<td>ρ</td>
<td>Density of reactor contents</td>
<td>19.2 (kgmol/m³)</td>
</tr>
<tr>
<td>ρ_c</td>
<td>Density of coolant</td>
<td>1000 (kg/m³)</td>
</tr>
<tr>
<td>Tj</td>
<td>Coolant temperature in the jacket</td>
<td>50.48 (°C)</td>
</tr>
<tr>
<td>Ti</td>
<td>Coolant inlet temperature</td>
<td>27 (°C)</td>
</tr>
<tr>
<td>T</td>
<td>Reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>T_m</td>
<td>Measured reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>U</td>
<td>Overall heat transfer coefficient</td>
<td>2.13x10⁵ (J/s·m²·°C)</td>
</tr>
<tr>
<td>V_C</td>
<td>Cooling jacket volume</td>
<td>1.82 (m³)</td>
</tr>
<tr>
<td>V</td>
<td>CSTR volume</td>
<td>7.08 (m³)</td>
</tr>
</tbody>
</table>

**Fig. 2.** Block diagram for a general Wiener Network with \( p \) inputs and one (1) output. Each input, \( x_i \), is passed through their own unity gain linear dynamic block, \( G_i \), after which these
unobservable intermediate outputs are collected and passed through a single unrestricted static gain function, \( f(V) \), to produce the output, \( y \).

Similarly, \( T_{ci} \) and \( C_{Ai} \) have a second-order-plus-dead-time-plus-lead (SOPDTPL) form as shown in Eq. 9 below:

\[
\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i \zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t-\theta)}{dt} + x_i(t-\theta)
\]  

(9)

where \( i = 2, 3 \), \( \tau_i \) is the time constant, \( \zeta_i \) is the damping coefficient, \( \tau_{ai} \) is the lead parameter and \( \theta_i \) is the dead time. Using a backward difference approximation to a sampling interval of \( \Delta t \), an approximate discrete-time form of Eq. 9 is:

\[
v_{i,\Delta t} = \delta_{1,\Delta t} v_{i,\Delta t-\Delta t} + \delta_{2,\Delta t} v_{i,\Delta t-2\Delta t} + \omega_{1,\Delta t} x_{i,\Delta t-\Delta t} + \omega_{2,\Delta t} x_{i,\Delta t-(k+1)\Delta t}
\]  

(10)

where \( \omega_{2,\Delta t} = 1 - \delta_{1,\Delta t} - \delta_{2,\Delta t} - \omega_{1,\Delta t} \) to satisfy the unity gain constraint with

\[
\delta_{1,\Delta t} = \frac{2\tau_i^2 + 2\tau_i \zeta_i \Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(11)

\[
\delta_{2,\Delta t} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(12)

\[
\omega_{1,\Delta t} = \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(13)

After obtaining \( v_{i,\Delta t} \) for each input \( i \), the modeled output value is determined by substituting these results into \( f(V_i) \).

The general discrete-time, “white noise,” Wiener model (WM) is given as [18]:

\[
y_i = y^{set} = y'_i = f(V_i) + \epsilon_i = \eta_i - Y^{set} + \epsilon_i
\]  

(14)

where

\[
\epsilon_i \sim N\left(0, \sigma^2\right) \text{ independently } \forall t
\]  

(15)
the common assumption with least squares estimation and $y_t$ is the measured output at time $t$.

Note that the expected response at time $t$ is $\eta_t = E[y_t]$, since $E[\epsilon_t] = 0$. While $f(V_t)$ can be any function, in the work it is a second order linear regression function of three (3) variables as shown below:

$$\eta_t - Y_{set} = f(V_t) = a_0 + a_1 v_{1,t} + \cdots + a_3 v_{3,t} + b_1 v_{1,t}^2 + \cdots + b_3 v_{3,t}^2 + c_{1,2} v_{1,t} v_{2,t} + \cdots + c_{2,3} v_{2,t} v_{3,t}$$  \hspace{1cm} (16)

where $a_i$, $b_i$, and $c_{j,k}$, denote the linear, quadratic and interaction parameters for $i = 1, 2, 3$, $j = 1$, 2 and $k = j+1, \ldots, 3$.

**FFPC Law**

The FFPC objective is to keep $\eta_t$ on the set point, $Y_{set}$, by changing $x_{i,t}$. More specifically, from Eq. 16, setting the predicted tank temperature $k_1$ time steps in the future to the set point gives

$$\eta_{t+k_1\Delta t} - Y_{set} = a_0 + a_1 v_{1,t+k_1\Delta t} + \cdots + a_3 v_{3,t+k_1\Delta t} + b_1 v_{1,t+k_1\Delta t}^2 + \cdots + b_3 v_{3,t+k_1\Delta t}^2 + c_{1,2} v_{1,t+k_1\Delta t} v_{2,t+k_1\Delta t} + \cdots + c_{2,3} v_{2,t+k_1\Delta t} v_{3,t+k_1\Delta t}$$

$$= a v_{1,t+k_1\Delta t}^2 + b v_{1,t+k_1\Delta t} + c = 0$$  \hspace{1cm} (17)

where $a = b_1$, $b = c_{1,2} v_{2,t+k_1\Delta t} + c_{1,3} v_{3,t+k_1\Delta t}$ and

$$c = a_0 + a_2 v_{2,t+k_1\Delta t} + a_3 v_{3,t+k_1\Delta t} + b_2 v_{2,t+k_1\Delta t}^2 + b_3 v_{3,t+k_1\Delta t}^2 + c_{2,3} v_{2,t+k_1\Delta t} v_{3,t+k_1\Delta t},$$

$$v_{2,t+k_1\Delta t} = \delta_{1,2} v_{2,t+(k_1-1)\Delta t} + \delta_{2,2} v_{2,t+(k_1-2)\Delta t} + \omega_{1,2} x_{2,t+(k_1-1)\Delta t} + \omega_{2,2} x_{2,t+(k_1-2)\Delta t}$$

$$v_{3,t+k_1\Delta t} = \delta_{1,3} v_{3,t+(k_1-1)\Delta t} + \delta_{2,3} v_{3,t+(k_1-2)\Delta t} + \omega_{1,3} x_{3,t+(k_1-1)\Delta t} + \omega_{2,3} x_{3,t+(k_1-2)\Delta t}$$

For $k_1$ time steps in the future, Eq. 7 becomes

$$v_{1,t+k_1\Delta t} = \delta_{1,1} v_{1,t+(k_1-1)\Delta t} + (1 - \delta_{1,1}) x_{1,t}$$  \hspace{1cm} (20)

where $a_i$, $b_i$, and $c_{j,k}$, denote the linear, quadratic and interaction parameters for $i = 1, 2, 3$, $j = 1$, 2 and $k = j+1, \ldots, 3$. Thus, from Eqs. 8 and 20,
Results

Fit of the WM

For this study $k_1 \Delta t$, $k_2 \Delta t$ and $k_3 \Delta t$ are 10 seconds, 5 seconds, and 0 seconds, respectively, with $\Delta t = 0.1$ seconds. The training input sequences for $x_1 - x_3$ are shown in Fig. 3. The fit for this training data is shown in Fig. 4. As shown, the fit is nearly perfect. A similar testing sequence resulted in a fit that was very similar.

Closed-loop Control

With the model parameters identified, the effects of FFPC, in this case, which are the effects of input variables announcements will be examined. The goal in this study is to minimize the variation of tank temperature around its set point (88 °C) by manipulating input signal to valve, $M$. This study was divided into two parts. First part has only one disturbance $T_{ci}$. In the second part, results for both disturbances $T_{ci}$ and $C_{ai}$ are given.

\[
M_{g,t} = x_{1,t} = \frac{v_{l,t+k_3 \Delta t} - \delta_{l,1} v_{l,t+(k_1-1) \Delta t}}{1 - \delta_{l,1}} = \tau_1 \frac{v_{l,t+k_3 \Delta t} - v_{l,t+(k_1-1) \Delta t}}{\Delta t} + v_{l,t+k_1 \Delta t} \quad (21)
\]

Fig. 3. Input sequences for input signal to valve, $M$ ($x_1$), Temperature of inlet coolant flow, $T_{ci}$ ($x_2$), and Concentration of component A in inlet flow, $C_{ai}$ ($x_3$).
A proportional-integral (PI) feedback controller (FBC) was implemented in conjunction with the FFC (no announcement of disturbances) and FFPC (announcement of the future behavior of disturbances) systems. For FBC, $K_C = 1.40$, $\tau_f = 11.0$ and $M_{fb}$ is the signal from the FBC system to the valve. The combined signal is given as

$$M_t = M_{fb,t} + M_{ff,t}$$  \hspace{1cm} (22)

In Fig. 5, $T_{ci}$ is the only disturbance. In this figure, the top plot shows the tank temperature response for FBC, the bottom left plot for FBC with FFC and the bottom right plot for FBC with FFPC. As shown in Fig. 5, the variation of $T_m$ around its set point (e.g., its standard deviation) is reduced with FFC even without $T_{ci}$ announcement, from 0.1891°C to 0.1494°C, or about 21%. However, with FFPC, it is much greater, from 0.1891°C to 0.0686°C, a 63.7% reduction.
In Fig. 6, $T_{ci}$ and $C_{ai}$ are disturbances. In this figure, the top plot shows the tank temperature response for FBC, the bottom left plot for FBC with FFC for $C_{ai}$ (no announcement) and FFPC for $T_{ci}$ (announcement) and the bottom right plot for FBC with FFPC (announcement for both disturbances). As shown in Fig. 6, the variation of $T_m$ around its set point (given by its standard deviation) is only slightly reduced without announcement for both disturbances from 0.4352°C to 0.4003°C, or about 8%. However, with full FFPC, the reduction is from 0.4352°C to 0.1131°C, a 74% reduction with both $T_{ci}$ and $C_{ai}$ announcements.

**Fig. 5.** The controller response plots with $T_{ci}$ as the only disturbance; FBC (top plot), FBC with FFC (bottom left plot), and FBC with FFPC (bottom right plot).
Fig. 6. The controller response plots with $T_{ci}$ and $C_{di}$ as disturbances; FBC (top plot), FBC with announcement for $T_{ci}$ only (bottom left plot), and FBC with full FFPC (bottom right plot).

**Concluding Remarks**

This article proposed a novel feedforward predictive control approach when the manipulated variable has significant dead time as in the case for exogenous insulin infusion. It extends the work in [17, 18]. As demonstrated, the use of this approach has the potential to greatly impact control application that can benefit from FFPC such as control of BGC for people that depend on insulin. Thus, future work will focus on evaluating the methodology on an FDA approved diabetes simulator.
Acknowledgements

The authors thank Tiffany Camarillo for her assistance in preparing the manuscript.

Financial support from the Juvenile Diabetes Research Foundation (Grant 1-INo-2016-329-S-B) is gratefully acknowledged.

Literature Cited


CHAPTER 9: DEMONSTRATION OF THE EFFECTIVENESS OF FEEDBACK PREDICTIVE MODELING AND CONTROL METHODOLOGY IN PROCESSES WITH SIGNIFICANT TIME DELAY AND UNMEASURED DISTURBANCES

A paper to be submitted to Industrial & Engineering Chemistry Research

Derrick K. Rollins, Sr. \(^1,2^*\) and Yong Mei\(^1\)

\(^1\)Department of Chemical and Biological Engineering, Iowa State University, Ames, IA;
\(^2\)Department of Statistics, Iowa State University, Ames, IA

*Corresponding Author

Abstract

The development of an automatic insulin delivery system (i.e., the so called, “artificial pancreas (AP)”) is a very active area of research. The challenge has been to develop a control system that can substantially tighten blood glucose concentration (BGC) around the clock and under normal changes of measured and unmeasured disturbances. The unmeasured disturbances have a large effect on the variation of BGC. This effect, along with the large time delay of the manipulated variable, insulin feed rate (IFR), and the requirement for an accurate model of IFR, are three of the most significant challenges to overcome to achieve the goal of tight control of BGC. Model predictive control (MPC), a feedback control approach, is the most widely accepted control algorithm in AP research. In this work, a promising predictive modeling and control methodology is proposed that addresses the aforementioned challenges and thus, has considerable promise as an effective AP. Unlike MPC, for the proposed approach, a model for manipulated variable is not required. The proposed predictive modeling approach has a novel way of addressing time delay in the manipulated variable that does not use an approximation for the physically unrealizable attribute. It addition, it uses a novel noise structure to accurately predict future BGC in the presence of unmeasured disturbances. These strengths of the proposed
method are demonstrated on a mathematical continuous-stirred tank reactor (CSTR) in a comparative study with MPC.

**Key words:** Artificial pancreas, model predictive control, feedback control, feedforward control, dynamic modeling.

**Introduction**

There is great need of tight control in blood glucose concentration (BGC) for people with diabetes. The artificial pancreas (AP) is the tool that was created to help insulin-dependent people with diabetes to automatically control their BGC. The AP has the following control components: continuous glucose sensor, insulin pump, and control algorithm. The control algorithm takes input information such as measured BGC, processes it through an algorithm that manipulates the insulin feed rate (IFR) with the control object to minimize deviation from the BGC target called the “set point.” For the AP to achieve the tight control necessary for long term 24 hour a day control, the effectiveness of AP control algorithm must significantly be improved.

Current AP algorithms include traditional feedback control (FBC) [1], model predictive control (MPC) [2-11], fuzzy logic control (FLC) [12, 13], and Feedforward control (FFC) [14-16]. Although considerable achievements have been made for inpatients/outpatients BGC closed-loop control, daytime automatic control has only marginally improved over manual subject control. As sensor technology continues to improve for disturbances that affect BGC, with effective mapping of these disturbances to BGC behavior (i.e., modeling), the potential to greatly tighten BGC using model-based control algorithms increases.

There are two basic components of a predictive model-based control strategy – the model and the control algorithm. More specifically, the model must be capable of accurate predictions of the controlled when the future predictions are most needed. For the AP this is the time distant
into the future when insulin will start to affect BGC after a change in IFR, i.e., the dead time $D$ for this variable. In addition, the model must be capable of accurate predictions this distance into the future in the presence of disturbances that significantly change BGC, and maintain high accuracy during close-loop control. A multiple-input modeling method that has strong cause-and-effect mapping of the inputs to the response space inclusive of phenomenologically sound input relationships, is also required. We believe that the modeling approach developed in [17, 23] strongly meets this criterion. The proposed control method presented in this article builds on these modeling strengths.

Model predictive control (MPC) [19] and feedforward predictive control (FFPC) [24,25] are control approaches that require accurate models for the manipulated variable in addition to accurate predictions $\theta$ time into future. It is very difficult to obtain an accurate model of the manipulated variable using any current modeling method for real BGC data. The approach that we are proposing in this work does not require a model of the manipulated variable, which is a critical advantage over MPC and FFPC. The proposed method uses a model to predict the value for the controlled variable $\theta$ time into the future, uses this value to determine the “feedback error” and then uses a classical feedback control (FFC) algorithm to change the IFR at the current time. Thus, we call this approach “feedback predictive control (FBPC).” As long as sufficiently accurate predictions are obtainable $\theta$ time into future under automatic control and with good tuning parameters, FBPC can be quite effective. In this work we evaluate FBPC against MPC using a simulated CSTR in the presence of very large unmeasured disturbance behavior. A CSTR is used because we can have large unmeasured disturbances as well as multiple-input modeling. Excluding the manipulated variable, no FDA approved diabetes
simulator has more than one input (meals) and the variation due to unmeasured disturbances is very small [24], if not, negligible.

Methodologies

Simulated CSTR

The CSTR process is used in this initial development and evaluation of the FBPC algorithm for the following reasons: 1. the currently available FDA approved UVA-padova simulator in type 1 diabetes does not support a multiple-disturbance control approach [20, 21]; 2. it is easy to implement the desired time delays for this process; 3. the dynamics of a CSTR are well understood theoretically as the process model is developed completely from conservation laws of material and energy balances and; 4. unmeasured disturbances can be implemented as desired. Thus, the model can be used with confidence that it represents true dynamic behavior of a chemical process and includes critical attributes of real BGC data. The CSTR process [18, 25] is illustrated in Fig. 1 and described by the following equations:

\[
\frac{dC_A}{dt} = \frac{F}{V}(C_{A_i} - C_A) - kC_A^2 \tag{1}
\]

\[
\frac{dT_r}{dt} = \frac{F}{V}(T_r - T_r) - \frac{\Delta H_r}{\rho C_p} kC_A^2 - \frac{UA}{\rho V C_p} (T_r - T_c) \tag{2}
\]

\[
\frac{dT_c}{dt} = \frac{UA}{V_c \rho C_p} (T_r - T_c) - \frac{F}{V_c} (T_c - T_c) \tag{3}
\]

\[
k = k_0 \exp \left( \frac{-E}{R(T_r + 273.16)} \right) \tag{4}
\]

\[
F_c = F_{c_{\text{max}}} \times \alpha^{(-M)} \tag{5}
\]

The definitions and values of the variables are shown in Table. 1.
In this study, three input variables are used in the model. They are the input signal to valve, \( M(x_1) \), temperature of entering coolant, \( T_{ci}(x_2) \), and concentration of component A in inlet, \( C_{Al}(x_3) \). There is one unmeasured disturbance, the feed flow rate \( F \). There is one output for this study, the measured tank temperature, \( T_m \). The CSTR is altered to include dead time \( \theta_i \) in \( T_{ci}, C_{Al} \), and the input signal to valve \( (M) \), specified as \( \theta_i = k_i \Delta t \), where \( k_i \) is a positive integer, \( i = 1, 2, \) or \( 3 \), for \( M, T_{ci}, \) and \( C_{Al}, \) respectively. The dead time for \( M \) is the largest.

Fig. 1. Simulated CSTR process.

**Modeling Methodology**

For FBPC to be successful the model must give accurate predictions in closed-loop control \( k_I \) time steps into the future, that is, when the current change in the manipulated variable affects the controlled variable. In practice, this is difficult because models are developed with one correlation structure (e.g., open loop) and used with another one (closed-loop). In addition, the models must have *cause-and-effect* input mapping to the outputs, maintain accuracy as unmeasured disturbances change, have low measure bias, and have physically sound phenomenological properties.
The Wiener network from [22] is used for modeling the CSTR process. This Multiple-Input, Single-Output (MISO) network is shown in Fig. 2. Each input \((x_i)\) has its own linear dynamic block, \(G_i\), and each dynamic block has an intermediate unobservable, output \(v_i\), which represents the independent dynamic response of its corresponding input. All the intermediate \(v_i\)'s are collected and passed through a nonlinear static gain block, \(f(V)\), to produce the final measured output, \(y\).

### Table 1. Definitions of CSTR process parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>SS value (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Heat transfer area</td>
<td>5.40 (m²)</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>Control valve rangeability parameter</td>
<td>50 (none)</td>
</tr>
<tr>
<td>(C_A)</td>
<td>Concentration of species A in reactor</td>
<td>1.0302 (kgmol/m³)</td>
</tr>
<tr>
<td>(C_{A_i})</td>
<td>Concentration of species A in inlet stream</td>
<td>2.88 (kgmol/m³)</td>
</tr>
<tr>
<td>(c_p)</td>
<td>Heat capacity of feed and product streams</td>
<td>1.815x10⁵ (J/kgmol·°C)</td>
</tr>
<tr>
<td>(c_{pc})</td>
<td>Heat capacity of coolant</td>
<td>4184 (J/kg·°C)</td>
</tr>
<tr>
<td>(\Delta H_R)</td>
<td>Heat of reaction</td>
<td>-9.86x10⁷ (J/kgmol)</td>
</tr>
<tr>
<td>E</td>
<td>Activation energy</td>
<td>1.182x10⁷ (J/kgmol)</td>
</tr>
<tr>
<td>F</td>
<td>Feed flow rate</td>
<td>0.45 (m³/s)</td>
</tr>
<tr>
<td>(F_C)</td>
<td>Coolant flow rate</td>
<td>0.44 (m³/s)</td>
</tr>
<tr>
<td>(F_{C_{max}})</td>
<td>Maximum flow rate of coolant through control valve</td>
<td>1.2 (m³/s)</td>
</tr>
<tr>
<td>K</td>
<td>Reaction rate constant</td>
<td>0.09 (m³/s·kgmol)</td>
</tr>
<tr>
<td>(k_o)</td>
<td>Arrhenius frequency parameter</td>
<td>0.0744 (m³/s·kgmol)</td>
</tr>
<tr>
<td>M</td>
<td>Input signal to the valve</td>
<td>0.26 (none)</td>
</tr>
<tr>
<td>R</td>
<td>Gas law constant</td>
<td>8314.39 (J/kgmol·K)</td>
</tr>
<tr>
<td>(\rho)</td>
<td>Density of reactor contents</td>
<td>19.2 (kgmol/m³)</td>
</tr>
<tr>
<td>(\rho_c)</td>
<td>Density of coolant</td>
<td>1000 (kg/m³)</td>
</tr>
<tr>
<td>(T_j)</td>
<td>Coolant temperature in the jacket</td>
<td>50.48 (°C)</td>
</tr>
<tr>
<td>(T_{C_i})</td>
<td>Coolant inlet temperature</td>
<td>27 (°C)</td>
</tr>
<tr>
<td>T</td>
<td>Reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>(T_{m})</td>
<td>Measured reactor temperature</td>
<td>88 (°C)</td>
</tr>
<tr>
<td>U</td>
<td>Overall heat transfer coefficient</td>
<td>2.13x10⁵ (J/s·m²·°C)</td>
</tr>
<tr>
<td>(V_C)</td>
<td>Cooling jacket volume</td>
<td>1.82 (m³)</td>
</tr>
<tr>
<td>V</td>
<td>CSTR volume</td>
<td>7.08 (m³)</td>
</tr>
</tbody>
</table>
Mathematical Model

The dynamic blocks are linear ordinary differential equations. The manipulated variable \( M \) has first-order-plus-dead-time dynamics given by Eq. 6 below:

\[
\tau_1 \frac{dv_1(t)}{dt} + v_1(t) = x_i(t - \theta_1)
\]  

where \( \tau_1 \) is the time constant and \( \theta_1 \) is the dead time. Using a backward difference approximation \( e.g., \frac{dv_1(t)}{dt} \approx \frac{v_{1,t} - v_{1,t-\Delta t}}{\Delta t} \) applied to a sampling interval of \( \Delta t \), an approximate discrete-time form of Eq. 6 is:

\[
v_{1,t} = \delta_{1,1} v_{1,t-\Delta t} + (1 - \delta_{1,1}) v_{1,t-\delta_{1,1}\Delta t}
\]  

where

\[
\delta_{1,1} = \frac{\tau_1}{\tau_1 + \Delta t}
\]
Similarly, $T_{ci}$ and $C_{Ai}$ have second-order-plus-dead-time-plus-lead (SOPDTPL) form as shown in Eq. 9 below:

$$
\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i \zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx(t-\theta)}{dt} + x_i(t-\theta)
$$

(9)

where $i = 2, 3$, $\tau_i$ is the time constant, $\zeta_i$ is the damping coefficient, $\tau_{ai}$ is the lead parameter and $\theta_i$ is the dead time. Using a backward difference approximation applied to a sampling interval of $\Delta t$, an approximate discrete-time form of Eq. 9 is:

$$
v_{i,t} = \delta_{i,1} v_{i,t-k\Delta t} + \delta_{i,2} v_{i,t-2\Delta t} + \omega_{i,1} x_i,t-k\Delta t + \omega_{i,2} x_i,t-(k+1)\Delta t
$$

(10)

where $\omega_{2,t} = 1 - \delta_{1,t} - \delta_{2,t} - \omega_{1,t}$ to satisfy the unity gain constraint with

$$
\delta_{i,1} = \frac{2\tau_i^3 + 2\tau_i \zeta_i \Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
$$

(11)

$$
\delta_{i,2} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
$$

(12)

$$
\omega_{i,j} = \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
$$

(13)

After obtaining $v_{i,t}$ for each input $i$, the modeled output value is determined by substituting these results into $f(V_i)$.

$$
\eta_i = f(V) = a_0 + a_1 v_{1,i} + \cdots + a_p v_{p,i} + b_1 v_{1,i}^2 + \cdots + b_p v_{p,i}^2 + c_{1,2} v_{1,i} v_{2,i} + \cdots + c_{p-1,p} v_{p-1,i} v_{p,i}
$$

(14)

Modification of Eq. 14 for predicting $k_i$ time steps into the future with $p = 3$ gives

$$
\eta_{i,i+k_i\Delta t} = a_0 + a_1 v_{1,i+k_i\Delta t} + \cdots + a_3 v_{3,i+k_i\Delta t}
$$

$$
+ b_1 v_{1,i+k_i\Delta t}^2 + \cdots + b_3 v_{3,i+k_i\Delta t}^2 + c_{1,2} v_{1,i+k_i\Delta t} v_{2,i+k_i\Delta t} + \cdots + c_{2,3} v_{2,i+k_i\Delta t} v_{3,i+k_i\Delta t}
$$

(15)

where $a_i$, $b_i$, and $c_{i,j}$ denote the linear, quadratic and interaction parameters for $i = 1, 2, 3$ and $j = 2$ and 3,
To model unmeasured disturbances and bias, we use the noise model structure in [26] and apply it to Eq. 14 as follows:

\[ y_t = \eta_t + N_t, \]  

where

\[ N_t = \frac{\theta_1(B)}{\phi_p(B)} a_t = \frac{1-\theta_1B-\theta_2B^2-\ldots-\theta_pB^p}{1-\phi_1B-\phi_2B^2-\ldots-\phi_pB^p} a_t = \frac{a_t}{\Phi(B)} \quad \forall t \]  

\[ a_t \sim N(0, \sigma^2) \]  

\[ y_t \] is the measured tank temperature, \( T_m \), at \( t \) and \( B'x_t = x_{t-y_0t} \). Then

\[ y_t [1-\phi_1B-\phi_2B^2-\ldots] = \eta_t [1-\phi_1B-\phi_2B^2-\ldots] + a_t \]  

\[ \Rightarrow y_t = \eta_t \phi_1(y_t-\eta_t) + \phi_2(y_t-2\Delta t-\eta_t) + \cdots + a_t \]  

Modification of Eq. 22 for predicting \( k \) time steps into the future gives

\[ \hat{y}_{t+k\Delta t} = \hat{\eta}_{t+k\Delta t} + \hat{\phi}_1(y_t-\hat{\eta}_t) + \hat{\phi}_2(y_t-2\Delta t-\hat{\eta}_t) + \cdots \]  

where \(^\hat{}\) is used for estimate. Our identification of Eq. 23 (i.e., estimation of the unknown parameters) follows [26].

**Control Algorithms**

The MPC and FFPC control algorithms we use in this study are given in detail in [25].

For these details, we refer the reader to this work. However, the FBPC approach will be described in detail here.

The classical PID controller equation is given by Eq. 24 below.
\[ M_{fb}(t) = x_i(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} \]  \hspace{1cm} (24)

where

\[ e(t) = y_{set} - \hat{y}_{t+k_i\Delta t} \]  \hspace{1cm} (25)

\( M_{fb}(t) \) is the FBC signal to the insulin pump at time \( t \), \( y_{set} \) is the set point for the tank temperature, \( \hat{y}_{t+k_i\Delta t} \) is the predicted tank temperature \( k_i \) time steps into the future, with \( k_i \Delta t \) as the dead time for manipulated variable, \( e(t) \) is the “feedback error” at \( t \) and; \( K_p, K_i, \) and \( K_d \) are the proportional, integral, and derivative tuning parameters, respectively, for the proportional, integral, and derivative (PID) controller in this work. FBPC applies the network of the FBC algorithm with modifications, to use the predicted \( T_m \left( \hat{y}_{t+k_i\Delta t} \right) \) in the feedback error shown by Eq. 25 and Fig. 3.

Fig. 3. Illustration of FBPC scheme

Results

Fit of the WM

The training input sequences for \( x_1-x_3 \) are shown in Fig. 4. The fits for this training data are shown in Fig. 5. First one can see inverse response behavior for the feed rate in the response plots of \( T_m \). There are two fits in Fig. 5. The first one is the model that fits the inputs and the
serial correlation structure of the noise, the unmeasured disturbances and the model bias as given by Eq. 19. As shown, this fit is excellent. This is the model that we used in the control algorithms. The fit on the right is a fit of the inputs on only. Thus, it does not fit the unmeasured disturbance, and as a result this fit is very poor given the large changes in the feed rate. Note that plot on the right reveals the large impact from the unmeasured disturbances, feed rate.

![Image of input sequences for input signal to valve, M (x1), Temperature of inlet coolant flow, Tci (x2), and Concentration of component A in inlet flow, CAi (x3).]

**Fig. 4.** Input sequences for input signal to valve, $M(x_1)$, Temperature of inlet coolant flow, $T_{ci}$ ($x_2$), and Concentration of component A in inlet flow, $C_{Ai}$ ($x_3$).

![Image of fitted measured tank temperature $T_m$. The blue dotted line denotes observed measured tank temperature, and black solid line denotes fit of the model. The plot on the left is for a serial correlated noise structure (i.e. $N_t$ is given by Eq. 19) and the plot on the right is for a “white” noise structure (i.e. $N_t = a_t$ as given by Eq. 20).]

**Fig. 5.** Fitted measured tank temperature $T_m$. The blue dotted line denotes observed measured tank temperature, and black solid line denotes fit of the model. The plot on the left is for a serial correlated noise structure (i.e. $N_t$ is given by Eq. 19) and the plot on the right is for a “white” noise structure (i.e. $N_t = a_t$ as given by Eq. 20).
Closed-loop Control

For this study $k_1\Delta t$, $k_2\Delta t$ and $k_3\Delta t$ are 10 seconds, 5 seconds, and 0 seconds with $\Delta t = 0.1$ seconds. Thus, the predictions for each control method is at least 100 time steps into the future.

The control objective is to minimize the variation around the set point (i.e. the standard deviation in the control run about its mean (Stdev), the set point (88 °C)) by manipulating the input signal to valve, $M$. This study examined five different cases for the control run as shown in Table 2: No control, FFPC, MPC with FFPC (for $J = 10$ and $J = 15$, where $J$ is the MPC tuning parameter and represents the time steps beyond $k_i$ time steps in the future), and FBPC with FFPC. As indicated by Table 2, the FBPC case significantly outperformed the MPC cases which have Stdev values 35.3 and 37.7% higher for $J$ equals 10 and 15, respectively. Note that the Stdev for MPC decreases as $J$ decreases. However, the aggressiveness of the manipulated variable increases as $J$ decreases and for these values of $J$, $M$ is significantly more aggressive and FBPC is still much better than MPC.

<table>
<thead>
<tr>
<th>Control Method</th>
<th>Prediction Horizon (s)</th>
<th>Stdev (°C)</th>
<th>% decrease from No Control</th>
<th>% Increase from FBPC/FFPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Control</td>
<td>0</td>
<td>3.68</td>
<td>0.0</td>
<td>429.4</td>
</tr>
<tr>
<td>FFPC</td>
<td>100</td>
<td>1.69</td>
<td>53.9</td>
<td>143.9</td>
</tr>
<tr>
<td>MPC/FFPC</td>
<td>110</td>
<td>0.94</td>
<td>74.4</td>
<td>35.3</td>
</tr>
<tr>
<td>MPC/FFPC</td>
<td>115</td>
<td>0.96</td>
<td>74.0</td>
<td>37.7</td>
</tr>
<tr>
<td>FBPC/FFPC</td>
<td>100</td>
<td>0.69</td>
<td>81.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Fig. 6. The control run without any control system manipulating the coolant flow rate.

Fig. 7. The control runs for FBPC and MPC.
Fig. 8. Responses for the manipulated variable ($M$) for FBPC versus MPC for $J = 10$ (left plot) and $J = 15$ (right plot). As expected, the variability of $M$ increases as $J$ decreases. At these values of $J$ the MPC is significantly more variable than FBPC.

**Concluding Remarks**

FBPC appears to have a lot of promise as an AP control scheme when future predictions are accurate. Since it uses a classical FBC system to determine with the predicted feedback error, it can provide tight control without the need for a model of the manipulated variable like MPC and FFC. This is a critical advantage in BGC application because it is difficult to obtain accurate models for the manipulated variable when modeling real data. With a model that can accurately predict BGC $k_f$ time steps in the future under unmeasured disturbances, FBPC should do quite well in terms of control performance. While for a PID FBPC, determining three turning parameters may appear to be a disadvantage in comparison to MPC that only has one, it is actually an advantage due to having more degrees of freedom or ways to optimize feedback control. Therefore, as demonstrated in this work, the use of this approach has the potential to greatly impact control application that can benefit from FBPC such as control of BGC for people that depend on insulin. Thus, future work will focus on evaluating the methodology on an FDA approved Diabetes simulator.


CHAPTER 10: AN IN-SILICO STUDY OF FEEDFORWARD PREDICTIVE CONTROL IN BLOOD GLUCOSE CONCENTRATION FOR PEOPLE WITH TYPE 1 DIABETES

The paper to be submitted to Journal of Diabetes Science & Technology
Yong Mei* and Derrick K. Rollins, Sr. 1,2

1Department of Chemical and Biological Engineering, Iowa State University, Ames, IA; 
2Department of Statistics, Iowa State University, Ames, IA

*Corresponding Author

Abstract

Type 1 diabetes is the condition where the pancreas is unable to produce insulin. The so-called “artificial pancreas” (automatic insulin delivery control system) that typically includes an insulin pump, a blood glucose sensor, and control algorithm, has the potential to help people with diabetes to tightly control their blood glucose concentration. This work applies and evaluates a feedforward predictive control (FFPC) algorithm using the first five subjects from a diabetes simulator. For each of these five cases, control was as tight or tighter than the most widely used approach, model predictive control (MPC).

Key words: Artificial pancreas, model predictive control, feedforward control, dynamic modeling

Introduction

Tight control of blood glucose concentration (BGC) is critical for people with diabetes. For people without diabetes, their BGC typically varies between 70 – 180 mg/dL (fasting glucose is around 70-130 mg/dL, and postprandial glucose level is typically less than 180mg/dL) [1]. For people with Type 1 diabetes (T1D), their BGC can go far above 180mg/dL. High levels of BGC can cause nerve, kidney and eyesight damage, in addition to many other health problems. At the other extreme, low levels of BGC can cause insufficient energy supply to the brain and can be immediately life threatening.
To survive, people with T1D must use exogenous insulin. The timing and amount of insulin to compensate for meals and other changes that affect BGC (i.e., disturbances) determines how BGC varies over time. The control objective is the tight variation around the target BGC. Because of the dynamic and uniquely complex effects of the factors that cause BGC variation (i.e., the inputs), tight control of BGC is not likely to be achieved by the subject (manual control) and it highly challenges the automatic control systems. Perhaps the most critical factor is the large time delay associated with the use of exogenous insulin, the manipulated variable. While this time is subject-specific and variable, it can take as much as an hour for insulin to begin to lower BGC after injection [2]. This property of the manipulated variable is unique to this application and is not common in industrial process control. Because of this property, effective BGC can only be achieved by making insulin changes ahead of any disturbance with a smaller time delay so that their effects on BGC are synchronous. For example, if carbohydrates have a dead time of 30 minutes and exogenous insulin has a dead time of one hour, insulin should be infused 30 minutes before eating so that one hour after that time when BGC will rise from the meal, the insulin is there to counteract this rise. Thus, optimal control also requires future knowledge for the times when these disturbances change (e.g., meal announcements) and the knowledge to predictively change insulin infusion rate to effectively cancel out the changes of disturbances (i.e., a control algorithm or model). The most popular approach that incorporates this type of optimality for automatically controlling BGC is model predictive control (MPC) [3-10]. Other approaches such as traditional feedback control is also being used in this application [11]. While MPC has seen success in automatic control of BGC, substantially more progress is needed in artificial pancreas research to achieve an acceptable and widespread control around the clock [12,13].
Automatic control of BGC can possibly be enhanced by the addition of an effective feedforward predictive control (FFPC) strategy in concert with an effective predictive feedback control (FBC) approach, such as MPC. Thus, the objective of this work is to evaluate the FFPC approach in Rollins [14] using artificially generated data from an FDA approved UVA/Padova diabetes simulator [15]. The use of this simulator allows the independent evaluation of FFPC because we are able to achieve a nearly perfect fit of the model using its only inputs, carbohydrates and insulin infusion rate, which is also an indication that unmeasured disturbances have small effect on the variation of BGC. Thus, in this work we evaluate FFPC without the addition of FBC. This allows us to examine the correctness of the proposed FFPC algorithm in the application of controlling BGC. Because of its proactive nature and control objective to keep the model fixed, in this study, the FFPC algorithm has the potential of outperforming MPC, given the low amount of unexplained variation of by the model. Thus, if the proposed FFPC algorithm can be effective in this application, the expectation is that in this study it will perform as well and maybe even better than MPC.

This article is organized as follows. In the next section, we will present the modeling methodology for this application, introduce the FFPC approach, discuss MPC in the context of this work, and the manual bolus (MB) method we use in this study. Following this section, the next one gives the protocols for the virtual patients and discusses the conditions of the study. After this section, the next gives results of the control run on virtual patients. The final section gives concluding remarks and ideas for future work.
Modeling Methodology

The model for BGC is the Wiener Model (WM) found in Kotz [16] and evaluated on real subject data. This approach is used due to its simplicity over the semi-coupled model [17] and because it is able to obtain excellent causative and physically correct fits to the subject data in the simulator.

The WM used in this study features multiple parallel dynamic inputs and one output. Figure 1 illustrates its block diagram. As shown, the inputs are first passed through dynamic, nonlinear, blocks where they are transformed into unmeasurable intermediate variables, and then all the intermediate variables are collected are passed through a function representing the static gain block.

![Block diagram](image)

**Figure 1.** Block diagram for a general wiener network with $p$ inputs and one (1) output.

In the context of diabetes, the inputs, $x_i$ for $i = 1,\ldots, p$, of the Wiener network are the measured noninvasive variables (i.e., meal components, physical activity, and emotional stress) and the output, $y$, is BGC. Each input has its own linear dynamic block, $G_i$, and each dynamic block has an intermediate unobservable, output $v_i$, which represents the independent dynamic response of its corresponding input. All the intermediate $v_i$’s are collected and passed through a nonlinear static gain block, $f(V)$, to produce the final measured output, $y$. The linear dynamic
blocks are essentially linear ordinary differential equations; a second-order-plus-lead with dead time (SOPLDT) form [18,19] as shown in Eq. 1. 

\[
\tau_i^2 \frac{d^2v_i(t)}{dt^2} + 2\tau_i\zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t-\theta)}{dt} + x_i(t-\theta) 
\]  

where \( i = 1, \ldots, p, p \) is the total number of inputs, \( \tau_i \) is the time constant, \( \zeta_i \) is the damping coefficient, \( \tau_{ai} \) is the lead parameter and \( \theta_i \) is the dead time. Using a backward difference approximation (e.g., \( \frac{dv_i(t)}{dt} \approx \frac{v_i(t)-v_i(t-\Delta t)}{\Delta t} \)), applied to a sampling interval of \( \Delta t \), Rollins [19] obtained an approximate discrete-time form of Eq. 2 as:

\[
v_{i,t} = \delta_{1,i}v_{i,t-\Delta t} + \delta_{2,i}v_{i,t-2\Delta t} + \omega_{1,i}x_{i,t-k_i,\Delta t} + \omega_{2,i}x_{i,t-(k_i+1)\Delta t}
\]  

where \( \theta_i = k_i\Delta t, k_i \) is an integer, and to satisfy the unity gain constraint, \( \omega_{2,i} = 1 - \delta_{1,i} - \delta_{2,i} - \omega_{1,i} \) with

\[
\delta_{1,i} = \frac{2\tau_i^2 + 2\tau_i\zeta_i\Delta t}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2}
\]

\[
\delta_{2,i} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2}
\]

\[
\omega_{1,i} = \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2}
\]

After obtaining Eq. 2 for each input \( i \), the modeled glucose value is determined by substituting these results into \( f(V) \) such as a second order regression form shown below:

\[
\eta_i = f(V) = a_0 + a_1v_{i,1} + \cdots + a_pv_{p,i} + b_1v_{i,1}^2 + \cdots + b_pv_{p,i}^2 + c_{1,2}v_{i,1}v_{2,i} + \cdots + c_{p-1,p}v_{p-1,i}v_{p,i}
\]

where \( a_i, b_i, \) and \( c_{i,j} \), denote the linear, quadratic and interaction parameters for \( i = 1, \ldots, p-1 \) and \( j = 2,\ldots, p \). This study uses only linear terms in the v’s, as this is sufficient to give the best fit.
In this simulation study, there are only two inputs: \( x_{1,t} \) is the grams of carbohydrate consumed at time \( t \) and \( x_{2,t} \) is the bolus injection/insulin infusion rate in unit of insulin unit per minute at \( t \). Thus, in this study,

\[
\eta_t = f(V) = a_0 + a_1 v_{1,t} + a_2 v_{2,t}
\]  

(7)

The measurement model is

\[
y_t = \eta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \ \forall t
\]  

(8)

where \( y_t \) is the measured glucose concentration at \( t \), \( \varepsilon_t \) is the error term under the assumptions of independence, normality and constant variance. Our identification of this model (i.e., estimation of the unknown parameters) follows Kotz [16]. In this study \( \theta_1 \) and \( \theta_2 \) are \( m \Delta t \) and \( n \Delta t \), respectively, with \( m \) and \( n \) as positive integers.

To model unmeasured disturbances and bias, we use the noise model structure in [22] and apply it to Eq. 8 as follows:

\[
y_t = \eta_t + N_t,
\]  

(9)

where

\[
N_t = \frac{\theta_q(B)}{\phi_p(B)} a_t = \frac{1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q}{1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p} a_t = \frac{a_t}{\Phi(B)} = \frac{a_t}{1 - \phi_1 B - \phi_2 B^2 - \ldots} \ \forall t
\]  

(10)

\[
a_t \sim N(0, \sigma^2)
\]  

(11)

\( y_t \) is the measured BGC at \( t \) and \( B^r x_i = x_{i-r} \). Then

\[
y_t \left[1 - \varphi_1 B - \varphi_2 B^2 - \ldots\right] = \eta_t \left[1 - \varphi_1 B - \varphi_2 B^2 - \ldots\right] + a_t
\]  

(12)

\[
\Rightarrow \ y_t = \eta_t + \varphi_1 (y_{t-\Delta t} - \eta_{t-\Delta t}) + \varphi_2 (y_{t-2\Delta t} - \eta_{t-2\Delta t}) + \cdots + a_t
\]  

(13)

Modification of Eq. 13 for predicting \( n \) time steps into the future gives
\[
\hat{y}_{t+n\Delta} = \hat{\eta}_{t+n\Delta} + \phi_1(y_t - \hat{\eta}_t) + \phi_2(\hat{y}_{t-\Delta} - \hat{\eta}_{t-\Delta}) + \cdots \\
= \hat{\eta}_{t+n\Delta} + \phi_1 \epsilon_t + \phi_2 \epsilon_{t-\Delta} + \cdots
\] (14)

where ‘\(^\hat{}\)’ is used for estimate. Our identification of Eq. 14 (i.e., estimation of the unknown parameters) follows [22].

**Control Algorithms**

**FFPC**

FFPC is the proposed control strategy in this study. Its control objective is to keep \(\eta_t\) on the set point, \(y_{set}\), by changing \(x_{2,t}\). More specifically, from Eq. 7, setting the predicted BGC \(n \Delta t\) time steps in the future to the set point gives

\[
y_{set} = \eta_{t+n\Delta} = a_0 + a_1 v_{1,t+n\Delta} + a_2 v_{2,t+n\Delta}
\] (15)

Rearranging Eq. 15 and using Eq. 2 gives

\[
v_{2,t+n\Delta} = \frac{y_{set} - a_0 + a_1 v_{1,t+n\Delta}}{a_2} = \delta_{1,2} v_{1,t+(n-1)\Delta \Delta} + \delta_{2,2} v_{2,t+(n-2)\Delta \Delta} + \omega_{1,2} x_{2,t} + \omega_{2,2} x_{2,t-\Delta}
\] (16)

Solving for the current bolus insulin rate to satisfy Eq. 15 gives from Eq. 16

\[
x_{2,t} = \frac{y_{set} - a_0 + a_1 v_{1,t+n\Delta} - \delta_{1,2} v_{2,t+(n-1)\Delta \Delta} - \delta_{2,2} v_{2,t+(n-2)\Delta \Delta} - \omega_{1,2} x_{2,t-\Delta}}{\omega_{1,2}}
\] (17)

where

\[
v_{1,t+n\Delta} = \delta_{1,1} v_{1,t+(n-1)\Delta \Delta} + \delta_{2,1} v_{2,t+(n-2)\Delta \Delta} + \omega_{1,1} x_{1,t+(n-1)\Delta \Delta} + \omega_{2,1} x_{1,t+(n-1)\Delta \Delta}
\] (18)

Given that negative values can occur from Eq. 17 and that it is possible that \(\eta_{t+n\Delta} < y_{set}\), the FFPC control law that we implemented is this study is as follows:

If \(\eta_{t+n\Delta} < y_{set}\), then set \(x_{2,t} = 0\), otherwise,
where $\beta$ is a tuning parameter between 0 and 1 to assist in hypoglycemia levels of BGC. The application of feedforward control here is unique because changes in the manipulated variable does not have the ability to increase BGC as always the case in chemical process control. Thus, the proposed FFPC is written to address this limitation of the manipulated variable in this application.

**MPC**

The MPC algorithm that we used in this study is now given as adapted from Seborg [20]. It uses the same WM as FFPC. First, a unit step change response in BGC, denoted as $S_j$, with respect to one unit step change of bolus feed rate is approximated from the second order equation given by Eq. 1, where $j$ denotes the tuning parameter that represents the number of steps the predicted model would take to reach the set point beyond $n \Delta t$ time steps in the future.

Then, with $k$ representing current time instant, and $n$ representing time delay for insulin, at $j$ steps ahead, based on Eq.14, prediction of BGC without current or future control action is given by:

$$
\hat{y}^0_{k+n+j} = a_0 + a_1 \hat{v}^0_{(k+n+j)} + a_2 \hat{v}^0_{2(k+n+j)} + \varphi_1(y_{k-1} - \eta_{k-1}) + \varphi_2(y_{k-2} - \eta_{k-2}) + \cdots 
$$

where according to Eq. 2 - 5, for $j \geq 2$,

$$
\hat{v}^0_{p,(k+j)} = \delta_{1,p} \hat{v}^0_{p,(k+j-1)} + \delta_{2,p} \hat{v}^0_{2p,(k+j-2)} + \omega_{1,p} x_{p,(k-1)} + \omega_{2,p} x_{p,(k-2)}
$$

and for $j = 1$,

$$
\hat{v}^0_{p,(k+1)} = \delta_{1,p} \hat{v}^0_{p,k} + \delta_{2,p} \hat{v}^0_{2p,(k-2)} + \omega_{1,p} x_{p,(k-1)} + \omega_{2,p} x_{p,(k-2)}
$$
where $p = 1, 2$, representing the inputs of meal and insulin infusion rate.

And the change of control action (manipulated variable) is given by,

$$\Delta x_k = \frac{y_{set} - y_{k+n+j}^0}{S_j}$$

(23)

Therefore, the manipulated variable at time instant $k$ (current time) will be,

$$x_k = x_{k-1} + \Delta x_k$$

(24)

And $j$ becomes the sole tuning parameter for MPC, which keeps it a simple form suited for subject specific tuning in the study.

**MB**

The third control strategy tested in this study provides a baseline for closed loop BGC control. This manual bolus (MB) control method calculated insulin infusion rate base on carb/insulin ratio of the subject only at the time of meal. In the aspect of administration frequency, MB is equivalent to open loop insulin bolus injection. Difference is that insulin is administrated automatically. And it can provide the baseline for the performance of closed loop BGC control since it is already implemented into commercialized insulin pumps.

The manipulated variables $x(t)$ from FFPC, MPC and MB are constrained by the bound of $0 U \leq x(t) \leq 100 U$. That is, whenever $x$ falls below 0 from FFPC or MPC, it will be set to 0, and if values of $x$ is beyond 100, it will be set to 100.

Both the FFPC and MPC are applied on virtual patients from a diabetes simulator with WM used. Since in this article, the proposed WM structure is stable and does not drift in a five-day control run, Time-varying version of MPC (e.g. MPC-LTV) [9] is not necessary for this study. Also, by pre-training the model before control runs and have the parameters in both FFPC and MPC fixed, the computational cost in running the algorithm can be greatly reduced.
Training Performance Statistics

Three measures of performance are used in this study for evaluating the fit of the model. The first one is the correlation coefficients between the fitted and measured BGC ($r_{fit}$);

$$r_{fit} = \frac{\sum_{i=1}^{n_t} y_i \hat{y}_i - \left( \frac{\sum_{i=1}^{n_t} y_i}{n_t} \right) \left( \frac{\sum_{i=1}^{n_t} \hat{y}_i}{n_t} \right)}{\sqrt{\sum_{i=1}^{n_t} y_i^2 - \left( \frac{\sum_{i=1}^{n_t} y_i}{n_t} \right)^2} \sqrt{\sum_{i=1}^{n_t} \hat{y}_i^2 - \left( \frac{\sum_{i=1}^{n_t} \hat{y}_i}{n_t} \right)^2}}$$

(25)

where $n_t$ is the number of pairs of values in the set, and $y_i$ and $\hat{y}_i$ are the $i^{th}$ measure and fitted values in the set, respectively. The closer this statistic is to 1.0 the better the model fits the measured response data without consideration of model bias. The second statistic, the “averaged error (AE),” gives a measure of model bias. As shown in the equation below, it is the average of deviation between for the differences of $y_i$ and $\hat{y}_i$.

$$AE = \frac{1}{n_t} \sum_{i=1}^{n_t} (y_i - \hat{y}_i)$$

(26)

The final statistic is the “averaged absolute error (AAE),” is a measure of the average closeness of $y_i$ and $\hat{y}_i$. Its formula is:

$$AAE = \frac{1}{n_t} \sum_{i=1}^{n_t} |y_i - \hat{y}_i|$$

(27)

A common benchmark used in this area is the percent of values in a range, e.g. 70-180 mg/dL suggested by American Diabetes Association [21], with values below 70mg/dL commonly considered as the hypoglycemic region and above 180mg/dL commonly considered as the hyperglycemic region.
Protocols of the *In-Silico* trials

The simulation environment used in this study is the diabetes simulator developed by UVA/Padova, which is approved by FDA for substitution of animal trials. As in this pilot study, five (5) virtual subjects were used in this study. The virtual patient information is summarized below in Table 1.

**Table 1. Summary of patient information**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Body Weight (kg)</th>
<th>Fasting BGC (mg/dL)</th>
<th>Basal (U/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>36.16</td>
<td>115</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>50.88</td>
<td>123</td>
<td>0.54</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>54.49</td>
<td>127</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>41.47</td>
<td>123</td>
<td>0.74</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>59.34</td>
<td>118</td>
<td>0.53</td>
</tr>
</tbody>
</table>

This study contains two phases: phase one is the Open Loop Phase (OLP) where individualized model parameters are identified based on two weeks of data with experimentally designed orthogonal inputs of meal and bolus injections in training period. Phase two is the Closed-Loop Phase (CLP) which is a five-day controlled run for each of the three control systems: FFPC, MPC and MB. In CLP, three meals were provided each day; meal times were 7am, 12pm and 8pm, respectively, and each meal contained 50 gm of carbohydrate. Meal announcements were provided 10 mins prior to a meal (i.e., $m = 10$ since the sampling frequency was every one (1) minute), and each meal lasted 15 minutes. Time delays associated with meal ingestion and insulin infusion rate were taken as 10 and 20 minutes ((i.e., $n = 20$), respectively. The values were determined during modeling the subjects. Under closed-loop, the control systems provided continuous insulin infusion at the changing frequency of every minute. The
virtual pumps supplied each subject with a pre-determined basal insulin (U/hr) according to patient information that kept BGC balanced during fasting.

**Results and Discussion**

In OLP, the experiment design for the input changes, which are orthogonal, for the nine (9) days of Training data, are given in Table 2. Table 3 gives the input changes to generate the two (2) days of Validation and three (3) days of Testing data.

The modeling results for each subject is given in Table 4. As shown, \( r_{fit} \) is excellent for all five subjects and for the full set of data with values of 0.99 or better. Similarly, the AE and AAE results are also excellent averaging 0.217 and 1.685 mg/dL, respectively on the Testing data. A graphical example of these excellent fits under CLP is given in Fig. 2 for Subject 4.

**Table 2.** Input changes used to generate the Training data.

<table>
<thead>
<tr>
<th>Time (t) (hr)</th>
<th>Meal Size ((x_{1,t})) (grams)</th>
<th>Insulin Bolus ((x_{2,t})) (U/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>54</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>79</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>103</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>127</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>151</td>
<td>0.01</td>
<td>2</td>
</tr>
<tr>
<td>175</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>199</td>
<td>50</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3.** Input changes used to generate the Validation and Testing data.

<table>
<thead>
<tr>
<th>Time (t) (hr)</th>
<th>Meal Size ((x_{1,t})) (grams)</th>
<th>Insulin Bolus ((x_{2,t})) (U/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation</td>
<td>223</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>247</td>
<td>40</td>
</tr>
<tr>
<td>Testing</td>
<td>271</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>295</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>319</td>
<td>50</td>
</tr>
</tbody>
</table>
Table 4. Model fitting results in mg/dL

<table>
<thead>
<tr>
<th>Subject</th>
<th>Subject</th>
<th>$r_{fit}$</th>
<th>AE (mg/dL)</th>
<th>AAE (mg/dL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Validation</td>
<td>Testing</td>
<td>Testing</td>
</tr>
<tr>
<td>1</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.275</td>
</tr>
<tr>
<td>2</td>
<td>0.997</td>
<td>0.996</td>
<td>0.997</td>
<td>0.433</td>
</tr>
<tr>
<td>3</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.219</td>
</tr>
<tr>
<td>4</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.068</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.999</td>
<td>0.999</td>
<td>0.088</td>
</tr>
<tr>
<td>Mean</td>
<td>0.999</td>
<td>0.998</td>
<td>0.999</td>
<td>0.217</td>
</tr>
</tbody>
</table>

Figure 2. Model fitness for subject 4 under CLP, where the orange line denotes BGC sensor measurements and blue line represents model prediction.
Figure 3. Panels (a) – (e) represent BGC measurements under different control algorithms for subject 1 – 5, respectively. Black lines denote BGC measurements under FFPC, blue dashed lines are for MPC, and red dotted lines are for MB.
Table 5. Results, in mg/dL, of control runs under FFPC, MPC and MB for all 5 subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Set Point</th>
<th>FFPC Mean (mg/dL)</th>
<th>FFPC Stdev (mg/dL)</th>
<th>FFPC 70-180 mg/dL</th>
<th>MPC Mean (mg/dL)</th>
<th>MPC Stdev (mg/dL)</th>
<th>MPC 70-180 mg/dL</th>
<th>MB Mean (mg/dL)</th>
<th>MB Stdev (mg/dL)</th>
<th>MB 70-180 mg/dL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>115</td>
<td>119.7</td>
<td>41.0</td>
<td>86.6%</td>
<td>108.2</td>
<td>42.4</td>
<td>69.8%</td>
<td>149.6</td>
<td>45.8</td>
<td>66.4%</td>
</tr>
<tr>
<td>2</td>
<td>113.3</td>
<td>112.3</td>
<td>32.0</td>
<td>93.6%</td>
<td>110.8</td>
<td>32.7</td>
<td>94.2%</td>
<td>153.3</td>
<td>36.7</td>
<td>74.9%</td>
</tr>
<tr>
<td>3</td>
<td>116.3</td>
<td>116.3</td>
<td>23.1</td>
<td>100.0%</td>
<td>114.8</td>
<td>23.6</td>
<td>100.0%</td>
<td>139.1</td>
<td>25.9</td>
<td>99.3%</td>
</tr>
<tr>
<td>4</td>
<td>129.5</td>
<td>129.5</td>
<td>40.6</td>
<td>85.0%</td>
<td>130.0</td>
<td>41.3</td>
<td>84.2%</td>
<td>149.9</td>
<td>42.6</td>
<td>64.3%</td>
</tr>
<tr>
<td>5</td>
<td>101.8</td>
<td>101.8</td>
<td>20.7</td>
<td>100.0%</td>
<td>108.1</td>
<td>21.9</td>
<td>100.0%</td>
<td>117.8</td>
<td>21.6</td>
<td>100.0%</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>116.1</td>
<td>31.5</td>
<td>93.0%</td>
<td>114.4</td>
<td>32.4</td>
<td>89.7%</td>
<td>141.9</td>
<td>34.5</td>
<td>81.0%</td>
</tr>
</tbody>
</table>

The CLP results for all three control algorithms are given graphically in Fig. 3 and numerically in Table 5. The tuning parameter $j$ for MPC ranges from 250 to 700. The results for FFPC and MPC are very similar. The mean values for the MB results are consistently higher than FFPC and MPC. Since the mean levels for FFPC and MPC can be shifted by changing the set point, the most important result is the standard deviation about the mean level, given as Stdev in Table 5. In each of the five cases, FFPC is better than MPC. On the average, FFPC has a higher percent in the 70 to 180 mg/dL range, 93.0% versus 89.7% for MPC. Therefore, the proposed FFPC algorithm appears to be capable of contributing to better close-loop control of BGC for people with T1D.

Concluding Remarks

FFPC appears to have the potential to enhance automatic control of BGC. Although in this evaluation FFPC and MPC were treated as separate control systems, in practice they would operate in concert with each other as one control system. FFPC would use the predictive model to change insulin infusion to keep the model inputs at a fixed value proactively at the current time instant. A predictive FBC method such as MPC, at the next future sampling time, would change insulin fusion rate to correct any predicted deviations from set point which would largely be due to unmeasured disturbances. Thus, FFPC control would compensate earlier for the
modeled disturbances and the MPC would compensate for unmeasured disturbances. The algorithm for FFPC is fundamentally different than MPC or any other control algorithm. In general feedforward control is a powerful control approach because it is theoretically perfect control. In this application, FFPC does not have this theoretical capability because manipulated variable, insulin, can only lower BGC. To get the most out of FFPC, an effective and safe way must be found to raise and to lower BGC. Future work will involve the usage of glucagon and insulin infusion to evaluate the impact of even tighter control for the subjects in this simulator.

In practice, any control system that relies on an accurately modeled relationship for the manipulated variable, e.g., FFPC and MPC, will suffer in performance considerably when the relationship is inaccurate. Currently, there is no modeling methodology, to our knowledge, that can develop this relationship accurately enough to implement an effective control system in practice over several days or weeks. Thus, given the current limitations in modeling accuracy for insulin infusion rate in real subject applications, an effective predictive control algorithm is needed to effectively change the insulin infusion rate without using an accurate model for manipulated variables. Our research group has developed such a predictive feedback modeling method that will be evaluated using this simulator and will compare its results with MPC. Our hypothesis is that this method will be significantly better than these methods and thus, is more likely to be effective in modeling real subjects.

**Literature Cited**


CHAPTER 11: A NOVEL FEEDBACK PREDICTIVE CONTROL ALGORITHM APPLIED IN BLOOD GLUCOSE CONCENTRATION CONTROL FOR PEOPLE WITH TYPE 1 DIABETES

The paper to be submitted to Journal of Diabetes Science & Technology

Yong Mei¹* and Derrick K. Rollins, Sr. ¹,²

¹Department of Chemical and Biological Engineering, Iowa State University, Ames, IA; ²Department of Statistics, Iowa State University, Ames, IA

*Corresponding Author

Abstract

Type 1 Diabetes (T1D) is a prevalent health disease. To better control blood glucose concentration and ease the burden for people with T1D, the closed-loop insulin-delivery system (i.e., artificial pancreas), which includes a glucose sensor, a control algorithm, and an insulin pump, appears to have much promise. In this work, a promising predictive modeling methodology is proposed. It uses a powerful modeling methodology to accurately predict BGC in the future that is equivalent to the dead time of exogenous insulin. At this future time, it determines the feedback error and uses this value in a feedback controller at the present time to achieve optimal distance in the future. This feedback predictive control (FBPC) algorithm is introduced and evaluated against model predictive control (MPC) and manual bolus (MB) using five subjects from a diabetes simulator. The reduction in BGC standard deviation around its mean significantly smaller than these three approaches.

Key words: Artificial pancreas, model predictive control, feedforward control, feedback control, dynamic modeling
Introduction

For people with type 1 diabetes (T1D), their pancreatic beta cells are destroyed and fail to produce insulin to maintain glycemic homeostasis. In that case, blood glucose concentration (BGC) cannot be balanced at a desired level. Their postprandial BGC can reach the hyperglycemia region (i.e. BGC > 180mg/dL), where untreated hyperglycemia status could cause irreversible damage to organs, eye sight, or limbs [1,2]. On the other hand, typical treatment for people with T1D is insulin therapy by injection or infusion via a pump. BGC could be brought to hypoglycemia region (BGC < 70mg/dL) if insulin delivery is too high. This status has immediate effects on people and since insufficient BGC supply will affect brain functionality, hypoglycemia can be immediately life-threatening [3]. Thus, tight control of BGC for people with T1D has become an important subject in diabetes research. An effective control algorithm can greatly impact the health for people with T1D.

There is a long history in management of BGC for people with T1D. A typical approach is, as mentioned before, insulin injection. In typical manual (i.e., open-loop) control, a person measures their BGC, and based on the value and recorded/anticipated carbohydrate ingestion, for example, the amount of injected insulin is calculated and administrated. This method requires a restricted life style and its effectiveness greatly depends on the management skills of the patients. Recent research shows that, due to advances in continuous glucose sensors, insulin pumps, fast acting insulin, automatic (i.e., closed-loop) control has much promise in tightening BGC. As a result, artificial pancreas (AP) research is a highly active field [4-14].

While automatic control is a highly effective and mature practice in industrial processes, its application to BGC control is relatively new. There are at least three critical reasons that automatic control of BGC is a unique and much more challenging application in automatic
control. The first one is the large dead time of exogenous insulin. In chemical processes, the manipulated variable, like the flow rate of coolant to the jacket around a reactor, is located close to the controlled variable, so its changes are observed relatively quickly. In contrast, when insulin is administered it can take as much as an hour [1] before the BGC begins to drop. The second one is that, unlike coolant flow rate to the jacket, for example, that is able to increase or decrease the controlled variable by its flow rate, changes in insulin flow rate is unidirectional, i.e., can only decrease BGC. The third one is that the input space for the generation of modeling data from humans is more restricted than the typical chemical process for not only safety reasons but also by personal limits and desires of the individual. In this article, the focus is the development of a control approach that overcomes the first challenge, large dead time in the manipulated variable.

Model predictive control (MPC) [15] and feedforward predictive control (FFPC) [16] are control approaches that require accurate models for the manipulated variable in addition to accurate predictions of \( \theta \) time into future, where \( \theta \) is the dead time of the manipulated variable. It is very difficult to obtain an accurate model of the manipulated variable using any current modeling method for real BGC data. The approach that we are proposing in this work does not require a model of the manipulated variable, which is a critical advantage over MPC and FFPC. The proposed method uses a model to predict BGC \( \theta \) time into future, using this value to determine the “feedback error” and then uses a classical feedback control (FBC) algorithm to change the insulin infusion rate at the current time. Thus, we call this approach feedback predictive control (FBPC). As long as sufficiently accurate BGC \( \theta \) time into future under automatic control are obtained and with good tuning parameters, FBPC can be quite effective. In this work, we evaluate FBPC against MPC using the FDA approved UVA/Padova diabetes
FBPC with FFPC was introduced in Rollins et al. [18] and compared MPC with FFPC in a simulation study using a continuous stirred tank reactor (CSTR). There are two critical advantages of that process over the diabetes simulator in this study – unmeasured disturbances and multiple inputs. In that study, FBPC was far superior to MPC. However, this study uses a diabetes simulator, which allows for a comparison with some of the unique challenges of BGC control mentioned above that cannot be mimicked in the CSTR simulation study.

Methodology

Control Algorithms

The MPC control algorithms we use in this study are given in detail in Seborg et al. and Mei and Rollins [15,16]. For these details, we refer the reader to these works. However, the FBPC approach will be described in detail here.

The classical PID controller equation is given by Eq. 1 below.

\[
x(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt}
\]  

(1)

where

\[
e(t) = y_{set} - \hat{y}_{t+n\Delta t}
\]  

(2)

\(x(t)\) is the controller signal to the insulin pump at time \(t\), \(y_{set}\) is the set point for BGC, \(\hat{y}_{t+n\Delta t}\) is the predicted BGC \(n\) time steps into the future, with \(n\Delta t\) as the dead time for insulin feed rate, \(e(t)\) is the “feedback error” at \(t\) and; \(K_p\), \(K_i\), and \(K_d\) are the proportional, integral, and derivative, tuning parameters, respectively, for the proportional, integral, and derivative (PID) controller in this.
work. FBPC applies the network of the FBC algorithm with modifications, to use the predicted BGC ($\hat{y}_{t+nM}$) in the feedback error shown in by Eq. 2 and Fig. 1.

**Fig. 1. Illustration of FBPC scheme**

**Modeling Methodology**

For FBPC to be successful, the model must give accurate predictions in closed-loop control $n$ time steps into the future, that is, when the current change in insulin infusion rate affects BGC. In practice, this is difficult because models are developed with one correlation structure (e.g., open loop) and used with another one. In addition, the models must have cause-and-effect input mapping to the outputs, maintain accuracy as unmeasured disturbances change, have low measure bias, and have physically sound phenomenological properties. For the diabetes simulator in this work, the Wiener method (WM) used in Mei and Rollins [16] was shown that it met these conditions quite well and will also be used in this study.

The WM used in this study features multiple parallel dynamic inputs and one output. Figure 2 illustrates its block diagram. As shown, the inputs are first passed through dynamic, nonlinear, blocks where they are transformed into unmeasurable intermediate variables, and then
all the intermediate variables are collected are passed through a function representing the static gain block.

Fig. 2. Block diagram for a general Wiener Network with p inputs and one (1) output.

In the context of diabetes, the inputs, \( x_i \) for \( i = 1, \ldots, p \), of the Wiener network are the measured noninvasive variables (i.e., meal components, physical activity, and emotional stress) and the output, \( y \), is BGC. Each input has its own linear dynamic block, \( G_i \), and each dynamic block has an intermediate unobservable, output \( v_i \), which represents the independent dynamic response of its corresponding input. All the intermediate \( v_i \)'s are collected and passed through a nonlinear static gain block, \( f(V) \), to produce the final measured output, \( y \). The linear dynamic blocks are essentially linear ordinary differential equations; a second-order-plus-lead with dead time (SOPLDT) form [19,20] as shown in Eq. 3.

\[
\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i \zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t-\theta)}{dt} + x_i(t-\theta)
\]  

(3)

where \( i = 1, \ldots, p \), \( p \) is the total number of inputs, \( \tau_i \) is the time constant, \( \zeta_i \) is the damping coefficient, \( \tau_{ai} \) is the lead parameter and \( \theta_i \) is the dead time. Using a backward difference
approximation \( \left( \text{e.g., } \frac{dv_i(t)}{dt} \simeq \frac{v_{i,t} - v_{i,t-\Delta t}}{\Delta t} \right) \) applied to a sampling interval of \( \Delta t \), an approximate discrete-time form of Eq. 3 is:

\[
\tau_i^2 \frac{d^2v_i(t)}{dt^2} + 2\tau_i \zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t - \theta_i)}{dt} + x_i(t - \theta_i)
\]  

(4)

where \( \theta_i \) is the dead time, and to satisfy the unity gain constraint, \( \omega_{2,i} = 1 - \delta_{1,i} - \delta_{2,i} - \omega_{1,i} \) with

\[
\delta_{1,i} = \frac{2\tau_i^2 + 2\tau_i \zeta_i \Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(5)

\[
\delta_{2,i} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(6)

\[
\omega_{1,i} = \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau_i^2 + 2\tau_i \zeta_i \Delta t + \Delta t^2}
\]  

(7)

After obtaining Eq. 4 for each input \( i \), the modeled glucose value is determined by substituting these results into \( f(V) \) such as a second order regression form shown below:

\[
\eta_i = f(V) = a_0 + a_{i,v_{1,i}} + \cdots + a_{p-1,p,t} v_{p,t}^2 + \cdots + b_i v_{1,i}^2 + \cdots + b_{p-1,p,t} v_{p,t} + c_{i,2} v_{1,i} v_{2,t} + \cdots + c_{p-1,p,1,v_{p-1,t}} v_{p,t}
\]  

(8)

where \( a_i, b_i, \) and \( c_{i,j} \), denote the linear, quadratic and interaction parameters for \( i = 1, \ldots, p-1 \) and \( j = 2, \ldots, p \). In this study, linear terms are sufficient for the modeling of BGC.

In this simulation study, there are only two inputs: \( x_{1,t} \) is the grams of carbohydrate consumed at time \( t \) and \( x_{2,t} \) is the bolus injection/insulin infusion rate in unit of insulin unit per minute at \( t \). Thus, in this study,

\[
\eta_i = f(V) = a_0 + a_{1,v_{1,t}} + a_{2,v_{2,t}}
\]  

(9)

In this study, \( \theta_1 \) and \( \theta_2 \) are \( m \Delta t \) and \( n \Delta t \), respectively, with \( m \) and \( n \) as positive integers.
To model unmeasured disturbances and bias, we use the noise model structure in Rollins et al. [21] and apply it to Eq. 9 as follows:

\[ y_t = \eta_t + N_t, \quad (10) \]

where

\[ N_t = \frac{\theta_1(B)}{\phi_p(B)} a_t = \frac{1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q}{1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p} a_t = \frac{a_t}{\Phi(B)} = \frac{a_t}{1 - \phi_1 B - \phi_2 B^2 - \ldots} \quad \forall t \quad (11) \]

\[ a_t \sim N(0, \sigma^2) \quad (12) \]

\( y_t \) is the measured BGC at \( t \) and \( B^r x_t = x_{t-yr} \). Then,

\[ y_t\left[1 - \phi_1 B - \phi_2 B^2 - \ldots\right] = \eta_t\left[1 - \phi_1 B - \phi_2 B^2 - \ldots\right] + a_t \quad (13) \]

\[ \Rightarrow y_t = \eta_t + \phi_1 (y_{t-\Delta t} - \eta_{t-\Delta t}) + \phi_2 (y_{t-2\Delta t} - \eta_{t-2\Delta t}) + \cdots + a_t \quad (14) \]

Modification of Eq. 14 for predicting \( n \) time steps into the future gives

\[ \hat{y}_{t+n\Delta t} = \hat{\eta}_{t+n\Delta t} + \hat{\phi}_1 (y_t - \hat{\eta}_t) + \hat{\phi}_2 (\hat{y}_{t-\Delta t} - \hat{\eta}_{t-\Delta t}) + \cdots \]

\[ = \hat{\eta}_{t+n\Delta t} + \hat{\phi}_1 e_t + \hat{\phi}_2 e_{t-\Delta t} + \cdots \quad (15) \]

where “\( \hat{\cdot} \)” is used for estimate. Our identification of Eq. 15 (i.e., estimation of the unknown parameters) follows Rollins et al. [21].

**Protocols of In-silico Study**

The diabetes simulator used in this study is developed by UVA/Padova, and is approved by FDA [17]. It is a valid substitute for animal trials in the early stages for development of AP. In addition, it provides us a safe tool to test different BGC control algorithms before moving them onto real patient study.
Table 1. Summary of patient information

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Body Weight (kg)</th>
<th>Fasting BGC (mg/dL)</th>
<th>Basal (U/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>36.16</td>
<td>115</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>50.88</td>
<td>123</td>
<td>0.54</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>54.49</td>
<td>127</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>41.47</td>
<td>123</td>
<td>0.74</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>59.34</td>
<td>118</td>
<td>0.53</td>
</tr>
</tbody>
</table>

The pool for this study contains 5 virtual patients from the simulator. Their characteristics are summarized in Table 1. To minimize the effect of measurement delay, both the virtual insulin pump and virtual BGC sensor were selected as intravenous (IV) type. As for the scenario tested in this study, three meals were served each day at 7am, 12pm and 8pm, and all the serving sizes contain 50 grams of carbohydrate. To examine the robustness of the proposed FBPC method, a longer than usual period of five-day closed loop control run was done. At the start of each trial, a two-day run-in period was implemented. In total, a seven-day trial was run for each virtual patient. A total of 3 different control algorithms were used in this in-silico study:

- Feedback predictive control (FBPC). Insulin was administrated for one hour before meal time at every minute.
- Model predictive control (MPC). As in Mei and Rollins [16] insulin was administrated at every minute.
- Manual bolus (MB). Traditional bolus insulin that was determined by insulin/carb ratio as explained in Mei and Rollins [16].

All four control algorithms were supplemented by basal insulin infusion (at the rate of U/hr) that is customized to each subject as shown in Table 1. The set point for each subject was set to their fasting BGC before the start of the trial. In addition, to prevent hypoglycemia episode, a pump shut-off protocol was implemented as when the predicted BGC was below set point – that is, the insulin pump would be turned off.
Evaluation Metrics

The evaluation of the effectiveness of a BGC control algorithm is often in the forms of percentage of time spent within 70-180mg/dL. Additionally, there are other metrics that can describe the different aspects of the effects of the control algorithms. The standard deviation around mean indicates the spread. Also, low blood glucose index (LBGI) and high blood glucose index (HBGI) can be used to gauge the risks in hypo-/hyperglycemia for each subject. The higher values of LBGI or HBGI, the higher probability of hypo- or hyperglycemia.

Table 2. Interpretations for values of LBGI and HBGI

<table>
<thead>
<tr>
<th>Risk</th>
<th>LBGI</th>
<th>HBGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimal</td>
<td>≤ 1.1</td>
<td>&lt; 5.0</td>
</tr>
<tr>
<td>low</td>
<td>1.1 – 2.5</td>
<td>5.0 – 10.0</td>
</tr>
<tr>
<td>medium</td>
<td>2.5 – 5.0</td>
<td>10.0 – 15.0</td>
</tr>
<tr>
<td>high</td>
<td>&gt; 5.0</td>
<td>&gt; 15.0</td>
</tr>
</tbody>
</table>

Training Performance Statistics

Three measures of performance are used in this study for evaluating the fit of the model. The first one is the correlation coefficients between the fitted and measured BGC ($r_{fit}$);

$$r_{fit} = \frac{\sum_{i=1}^{n_t} y_i \hat{y}_i - \left(\frac{\sum_{i=1}^{n_t} y_i}{n_t}\right) \left(\frac{\sum_{i=1}^{n_t} \hat{y}_i}{n_t}\right)}{\sqrt{\sum_{i=1}^{n_t} y_i^2 - \left(\frac{\sum_{i=1}^{n_t} y_i}{n_t}\right)^2} \sqrt{\sum_{i=1}^{n_t} \hat{y}_i^2 - \left(\frac{\sum_{i=1}^{n_t} \hat{y}_i}{n_t}\right)^2}} \quad (16)$$

where $n_t$ is the number of pairs of values in the set, and $y_i$ and $\hat{y}_i$ are the $i^{th}$ measure and fitted values in the set, respectively. The closer this statistic is to 1.0, the better the model fits the measured response data without consideration of model bias. The second statistic, the “averaged
error (AE),” gives a measure of model bias. As shown in the equation below, it is the average of deviation between for the differences of \( y_i \) and \( \hat{y}_i \).

\[
AE = \frac{1}{n_t} \sum_{i=1}^{n_t} (y_i - \hat{y}_i) 
\]

(17)

The final statistic is the “averaged absolute error (AAE)”, which is a measure of the average closeness of \( y_i \) and \( \hat{y}_i \). Its formula is:

\[
AAE = \frac{1}{n_t} \sum_{i=1}^{n_t} |y_i - \hat{y}_i| 
\]

(18)

A common benchmark used in this area is the percent of values in a range, e.g. 70-180 mg/dL suggested by American Diabetes Association [22], with values below 70mg/dL commonly considered as the hypoglycemic region and above 180mg/dL commonly considered as the hyperglycemic region.

**Results and Discussion**

Training of the models follows the protocols describes in Rollins et al. [20]. The data set for identifying all parameters is the same as in Mei and Rollins [16]; specifically, 9 days of training, 2 days of validation and 3 days of testing. Modeling results are summarized in Table 3.

The modeling results for each subject is given in Table 3. As shown, \( r_{fit} \) is excellent for all five subjects and for the full set of data with values of 0.99 or better. Similarly, the AE and AAE results are also excellent averaging 0.22 and 1.7 mg/dL, respectively on the Testing data. To demonstrate the excellent ability of the model to fit well \( n \) time steps into the future under automatic control, an example of the fit for FBPC is shown in Fig. 3 for Subject 4. The \( r_{fit} \) between \( n \) steps ahead predictor and measurements is 0.9855, AE is 1.1771, and AAE is 3.8764.
<table>
<thead>
<tr>
<th>Subject</th>
<th>$r_{fit}$</th>
<th>AE (mg/dL)</th>
<th>AAE (mg/dL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Validation</td>
<td>Testing</td>
</tr>
<tr>
<td>1</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>2</td>
<td>0.997</td>
<td>0.996</td>
<td>0.997</td>
</tr>
<tr>
<td>3</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>4</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>Mean</td>
<td>0.999</td>
<td>0.998</td>
<td>0.999</td>
</tr>
</tbody>
</table>

**Table 3.** Results of model fitting

![Figure 3](image.png)

**Figure 3.** Model predictions under FBPC for subject 4. The orange line represents BGC measurements under FBPC, and blue dotted line represents model prediction of 20 mins into the future.

In this study, five subjects were used and their results are shown in Fig. 4(a)–(e). For each subject, the proposed FBPC approach outperforms the other ones, as the black lines for FBPC have smaller variation around mean than all other methods tested in this study. For more specific pairwise comparison, the results are summarized in the following Table 4.
Fig. 4. Panel (a)-(e) represent BGC measurements for subject 1-5. Black line represents results under proposed FBPC, blue dashed line denotes MPC, and yellow dotted line is for MB.
Table 4. Summary results of control runs with different algorithms

<table>
<thead>
<tr>
<th>Subject</th>
<th>Set Point</th>
<th>FBPC</th>
<th>MPC</th>
<th>MB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (mg/dL)</td>
<td>Stdev (mg/dL)</td>
<td>70-180 mg/dL</td>
</tr>
<tr>
<td>1</td>
<td>115</td>
<td>114.4</td>
<td>36.6</td>
<td>92.9%</td>
</tr>
<tr>
<td>2</td>
<td>123</td>
<td>124.0</td>
<td>26.4</td>
<td>94.9%</td>
</tr>
<tr>
<td>3</td>
<td>127</td>
<td>122.1</td>
<td>20.7</td>
<td>100.0%</td>
</tr>
<tr>
<td>4</td>
<td>123</td>
<td>124.0</td>
<td>34.1</td>
<td>98.8%</td>
</tr>
<tr>
<td>5</td>
<td>118</td>
<td>116.9</td>
<td>13.6</td>
<td>100.0%</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>120.3</td>
<td>26.3</td>
<td>97.3%</td>
</tr>
</tbody>
</table>

For pairwise comparison, standard deviation reduction with FBPC is 19% and 23%, compared to MPC and MB, respectively. Furthermore, the increase of time spent within 70 – 180 mg/dL with FBPC against MB is 21%, and on average, 97.3% of time was spent within desired BGC region.

For evaluation of risks for hypo-/hyperglycemia episode, the statistics of LBGI and HBGI are used. They serve as alerts and boundaries for control methods tested. The goal for LBGI and HBGI is to maintain both risks at minimal level according to Table. 2, whereas LBGI weighs more in our consideration since hypoglycemia has more immediate life-threatening effects on people with T1D. Results for LBGI and HBGI are demonstrated in Tables 5 and 6. For LBGI, all cases under FBPC have minimal risks for hypoglycemic episode, and 2 cases under MPC have risks that are greater than minimal. As for HBGI, all cases are under the values of 5, which indicates the concern for hyperglycemia is minimum for all control algorithms tested in this study.
Table 5. LBGI for different control algorithms

<table>
<thead>
<tr>
<th>Subject</th>
<th>FBPC</th>
<th>MPC</th>
<th>MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.07</td>
<td>3.54</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.12</td>
<td>1.19</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.12</td>
<td>0.76</td>
<td>0.02</td>
</tr>
<tr>
<td>4</td>
<td>0.7</td>
<td>0.92</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>0.94</td>
<td>0.32</td>
</tr>
<tr>
<td>Mean</td>
<td>0.422</td>
<td>1.47</td>
<td>0.108</td>
</tr>
</tbody>
</table>

Table 6. HBGI for different control algorithms

<table>
<thead>
<tr>
<th>Subject</th>
<th>FBPC</th>
<th>MPC</th>
<th>MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1.29</td>
<td>4.68</td>
</tr>
<tr>
<td>2</td>
<td>1.63</td>
<td>0.88</td>
<td>4.3</td>
</tr>
<tr>
<td>3</td>
<td>1.77</td>
<td>0.68</td>
<td>2.49</td>
</tr>
<tr>
<td>4</td>
<td>3.49</td>
<td>2.38</td>
<td>4.82</td>
</tr>
<tr>
<td>5</td>
<td>1.78</td>
<td>0.35</td>
<td>0.68</td>
</tr>
<tr>
<td>Mean</td>
<td>2.134</td>
<td>1.116</td>
<td>3.394</td>
</tr>
</tbody>
</table>

Concluding Remarks

This article demonstrated the potential superiority of FBPC to tighten BGC control for people with T1D over current methods. In addition to this advantage, FBPC just requires a model that can predict accurately a time distant into the future equal to the dead time of the insulin infusion rate. This is a considerable advantage over model based methods, such as MPC and FFPC that require an accurate cause-and-effect model for the manipulated variable as the input to determine BGC. Modeling methods to implement FBPC should focus on multiple-input cause-and-effect [19], phenomenologically sound structure [23], and high accuracy under automatic control conditions. The pre-whitening procedure we applied in this work will effectively handle unmeasured disturbances and model bias as demonstrated in Rollins and Mei [18] and in Kotz et al. [24] on real Type 1 diabetes data with $r_{fits}$ as high as 0.9 for several subjects. In practice, since it is a FBC method it would be combined with FFPC when accurate models for the manipulated
variable can be obtained. Since the method has shown it superiority over MPC here using a diabetes simulator and in Rollins and Mei [18], the next step is to evaluate it on real subjects.

**Literature Cited**


CHAPTER 12: FUTURE WORKS

The successful development and demonstration of the effectiveness of feedback predictive control (FBPC) algorithm, has paved the way for real subject clinical studies in the future. Additional work needs to be done to further polish this methodology. For example, physiologically sound semi-coupled structure should be implemented into FBPC to capture more cause-and-effect relationships between meals, insulin, and BGC. Plus, more measured disturbances such as variables that concern physical activities and emotional stress should be incorporated into the model. Last but not least, FBPC could work in concert with other control algorithms and potentially achieve tighter BGC control.

In addition, Further investigation will be conducted on model predictive control (MPC), feedforward predictive control (FFPC) and other modeled based control algorithms, with the development of sensing technology in blood glucose concentration (BGC) and plasma insulin, and the commercialization of other hormones such as glucagon.

For people with type 2 diabetes, virtual BGC sensing algorithm can be incorporated into lifestyle management system residing on portable devices that can provide valuable information on how people should manage their diets and exercises.