Prediction of postprandial glucose excursions in type 1 diabetes using control-oriented process models

D. Adelberger * **F. Reiterer** ** **P. Schrangl** ***
Ch. Ringemann **** **T. Huschto** **** **L. del Re** *

* *Institute for Design and Control of Mechatronical Systems, Johannes Kepler University Linz, Altenberger Strasse 69, A-4040 Linz, Austria (e-mail: daniel.adelberger@jku.at)*
** *Nemak Linz GmbH, A-4030 Linz, Austria*
*** *Magna Powertrain Engineering Center Steyr GmbH & Co KG, A-4300 St. Valentin, Austria*
**** *Roche Diabetes Care GmbH, D-68305 Mannheim, Germany*

Abstract: Reliable prediction of future blood glucose (BG) values is of high relevance for diabetes patients, since it enables the use of predictive glucose alarms (warning the patient about impending situations with dangerously low or high BG), as well as of model-based algorithms for smart glucose control. Control-oriented graybox process models have proven very suitable for such tasks, especially when identified on data from clinical trials under well-defined conditions. The current paper analyzes how such models can also be reliably parametrized using outpatient data of patients on multiple daily injection (MDI) therapy. A dedicated preprocessing algorithm is presented to look for suitable (*i.e.* complete and sensible) data segments that allow for a reliable system identification. The focus of the current paper is on the prediction of postprandial glucose trajectories, more specifically on predictions made exactly at the time of meal ingestion. This corresponds to a particularly challenging task, but one with high importance for the model-based optimization of insulin doses. It is demonstrated that the identified process models are a suitable choice for predicting such postprandial glucose excursions.

Keywords: Diabetes, Identification, Grey box modeling, Prediction, Validation

1. INTRODUCTION

Patients with type 1 diabetes mellitus (T1DM) require insulin injections in order to mitigate the long–term effects of a chronically increased blood glucose (BG) level. However, too much insulin leads to hypoglycemia which is a potentially life threating situation. Managing BG by injecting a suitable amount of insulin is a difficult task and a considerable burden for patients with diabetes, especially seen that there is a large day-to-day variability of BG dynamics and a myriad of factors that influence BG. Therefore, the design of smart control algorithms that help patients with managing their BG has kept control engineers busy for many years already. These research efforts have resulted in the design of numerous algorithms for assisting patients with the control of their glucose levels, *e.g.* by warning patients about impeding hypoglycemia or hyperglycemia, by giving insulin dosing advice via a decisions support systems, or by providing closed-loop glucose control via an artificial pancreas. Most of these algorithms rely at their core on a mathematical model for the prediction of future BG levels. A whole variety of models has been proposed for this purpose. Typically these can be categorized either as physiological models or data-based models. Seen that they are far easier to parametrize/personalize, the vast majority of publications on the topic uses data-based models for this purpose, see Oviedo et al. (2017) for an overview. A sub-class of these data-based prediction models that have proven especially suitable for control-oriented tasks are graybox process models. Over the years different process model structures have been proposed for describing the glucose response to carbohydrate intakes and bolus insulin injections, see Percival et al. (2010); van Heusden et al. (2011); Kirchsteiger et al. (2014); Cescon et al. (2014); Tárník et al. (2015); Toffanin et al. (2018). Such models have been successfully applied for the identification of bolus calculator settings as in Kirchsteiger and del Re (2014); Reiterer et al. (2015a,b); Bock et al. (2015), closed loop glucose control in an artificial pancreas as in Gondhalekar et al. (2016); Boiroux et al. (2017) and the prediction of impending hypoglycemia as in Toffanin et al. (2018). The main advantages of process models are the straightforward interpretation of the model parameters and the simple parametrization (seen that the models are restricted to few parameters). Model individualization strategies described in the literature are either based on simple tuning rules as in van Heusden et al. (2011); Boiroux et al. (2017); Tárník et al. (2015) or based on optimizing model parameters based on the data from suitable postprandial glucose responses, either using high quality data from clinical trials as in Percival et al. (2010); Kirchsteiger et al. (2014); Bock et al. (2015); Cescon et al. (2014) or manually selected

data segments that are deemed suitable as in Toffanin et al. (2018). The current paper explores the use of the control-oriented process model proposed in Kirchsteiger et al. (2014) for the short and medium-term prediction of glucose values after meal intakes. Compared to previously published works with a similar focus, the novelty of this publication is the following:

- Models are identified from outpatient data of patients on multiple daily injection (MDI) therapy recorded under real-life conditions. Information about insulin dosing and meal intakes are based solely on diary entries of patients, which are often erroneous or missing. A data preprocessing algorithm is proposed here to automatically screen for suitable data segments which allows for a reliable model identification even under such challenging conditions.
- The paper analyzes the reliability of process models for the prediction of future glucose values by analyzing their performance as a function of the prediction horizon.
- Different options for describing the effect of the initial state on the predicted postprandial glucose trajectory are explored.

2. METHODOLOGY

2.1 The Kirchsteiger Model

A process model of the following structure is considered in the current paper for describing the glucose response to carbohydrate intakes, as well as to bolus insulin injections:

$$CGM(s) = \frac{K_1}{(1+sT_1)^2 s} \cdot D(s) + \frac{K_2}{(1+sT_2)^2 s} \cdot U(s) \quad (1)$$

In this formula, $CGM(s)$ describes the glucose level recorded via a continuous glucose monitoring (CGM) system, $D(s)$ the carbohydrates of meal intakes and $U(s)$ the bolus insulin injections, all in the Laplace domain. Other influencing inputs like basal insulin, stress, sports, mixed meal composition etc. are not incorporated into this model structure. The parameters in (1) have an easy to grasp physiological interpretation: Whereas K_1 describes the effect of 1 gram of carbohydrates on glucose levels, K_2 corresponds to the effect of 1 IU of bolus insulin (both for $t \to \infty$). Time constants T_1 and T_2 are proportional to the response time to carbohydrate and insulin inputs. Parameter K_2 has the same interpretation as the insulin sensitivity factor ISF, whereas the ratio K_2/K_1 tells how many grams of carbs are compensated by 1 IU, just as it is the case for the carbohydrate-to-insulin ratio CIR in Advanced Carbohydrate Counting (ACC), see Reiterer and Freckmann (2019). The process model structure used here was first reported in Kirchsteiger et al. (2011) and has already been applied extensively (and successfully) in previous works, among others for the parametrization of bolus calculator settings in ACC, see Kirchsteiger and del Re (2014); Reiterer et al. (2015a,b).

2.2 Framework for the Identification of Model Parameters and the Prediction of Future Glucose Values

Two different frameworks are considered for the identification of model parameters of (1), as well as for the later-on prediction of postprandial glucose responses. These frameworks differ in how they model the effect of the initial state on the predicted glucose trajectory. In both cases, however, the core of the model identification corresponds to the minimization of the following cost function:

$$J(\theta) = \sum_{d=1}^{d_{tot}} \left(\sum_{k=k_0(d)}^{k_N(d)} f(y_k, \hat{y}_k(\theta))^2 \right) \quad (2)$$

with

$$f(y_k, \hat{y}_k(\theta)) = \begin{cases} y_k - \hat{y}_k(\theta) & \text{if } y_k < 100 \text{ mg/dl} \\ 100 \cdot \frac{y_k - \hat{y}_k(\theta)}{y_k} & \text{if } y_k \geq 100 \text{ mg/dl} \end{cases} \quad (3)$$

In this cost function y_k corresponds to the measured output, *i.e.* the CGM data, whereas $\hat{y}_k$ is the model output. Each data segment d (d_{tot} in total) used in the identification segment is characterized by a starting index k_0 and an end index k_N. The model output $\hat{y}_k$ is computed using model parameters θ and an estimate of the initial state $\hat{x}_0$. This initial state estimate corresponds to the state of the system at the start of the identification segment k_0. For each data segment of a patient the same parameter vector θ is used, representing the best average patient-specific model parametrization.

Process Model combined with Kalman filter (KF-PM):

In the first setting a Kalman filter is used to estimate the initial state. This initial state is then used in model (1) together with the information about meal size and injected bolus quantity to simulate the postprandial glucose trajectory. However, in order not to create any bias in the predicted glucose values, it is necessary to already consider the Kalman filter when optimizing the model parameters. For the system identification the Kalman filter starts with its state estimation 6 hours before the starting point of the identification data segment and computes estimates for the state for each of those time points up to the start of the identification segment. The last estimate of $\hat{x}$ (just at the start of the identification data segment) corresponds to the initial state for the system identification. The model inside the Kalman filter (derived from the process model with parameters θ) is updated in every iteration step of the optimization using the latest estimate of the model parameters θ, *i.e.* the process model used for prediction and the (same) process model used inside the Kalman filter are optimized simultaneously. A new Kalman filter is thus designed in each iteration step and is subsequently used to compute the state estimates $\hat{x}(t_{k_0})$ at each time t_{k_0} based on input (d, u) and output $(\Delta y = y - G_b)$ data up to time t_{k_0} (with G_b, the estimate of the patient's basal glucose level). Using the estimated state $\hat{x}(t_{k_0})$ at time t_{k_0} as initial condition, a simulation with the process model k steps into the future can be computed ($y_{\text{sim},k}(t_{k_0})$). A prediction for time $t_{k_0} + kT_S$ (with T_S, the sample time) can thus be calculated as:

$$\hat{y}(t_{k_0} + kT_S | t_{k_0}) = y_{\text{sim},k}(t_{k_0}) + G_b \quad (4)$$

Hybrid Process and Autoregressive Model (AR-PM):

$\hat{x}_0$ is assumed to be 0 for the process model (no impact of the initial state on the process model output), but the effect of the initial state is captured by an autoregressive (AR) model. The predicted glucose output of this hybrid model corresponds to the sum of the process model prediction and the prediction by the AR model. A population

mean AR model with parameter values identified from data during the night period (*i.e.* without any influence of meals) is used for this purpose. Let $\hat{y}_{\mathrm{AR}}(t_{k_0} + kT_{\mathrm{S}}|t_{k_0})$ be the prediction of time $t_{k_0} + kT_{\mathrm{S}}$ given information up to time t_{k_0}. The combined model output for the predicted glucose trajectory is then calculated according to the following equation:

$$\hat{y}(t_{k_0} + kT_{\mathrm{S}}|t_{k_0}) = y_{\mathrm{sim},k}(t_{k_0}) + \hat{y}_{\mathrm{AR}}(t_{k_0} + kT_{\mathrm{S}}|t_{k_0}) \quad (5)$$

where $y_{\mathrm{sim},k}(t_{k_0})$ is the output of the process model simulated with the assumption of $\hat{x}(t_{k_0}) = 0$ as initial condition.

2.3 Reference Prediction Models

Besides the process model (1), additional prediction models are used in order to facilitate the assessment of the achieved prediction performance. The following prediction models are used for this purpose:

- *Zero Order Hold:* The zero order hold (ZOH) model computes the prediction $\hat{y}$ at time $t + kT_{\mathrm{s}}$ based on the CGM data y up to time t by keeping the latest available value constant, *i.e.*

$$\hat{y}(t + kT_{\mathrm{s}}) = y(t). \quad (6)$$

- *Global AR Model:* A simple, second order global AR model of the form

$$\Delta\hat{y}_{\mathrm{AR}}(t + kT_{\mathrm{s}}) = a_k \Delta y(t) + b_k \Delta y(t - T_{\mathrm{S}}) \quad (7)$$

$$\Delta y(t) = y(t) - G_b \quad (8)$$

$$\hat{y}_{\mathrm{AR}}(t + kT_{\mathrm{s}}) = \Delta\hat{y}_{\mathrm{AR}}(t + kT_{\mathrm{s}}) + G_b \quad (9)$$

  is used (where G_b is again the patient-specific estimate of the basal glucose level). The parameters (a_k, b_k) are optimized using least squares (LS) optimization for each prediction horizon k.
- *Alternative Process Model Structure:* To put the prediction performance of model (1) into context, additionally, an alternative process model structure without integrator terms is analyzed:

$$CGM(s) = \frac{K_1}{(1 + sT_1)^2} \cdot D(s) + \frac{K_2}{(1 + sT_2)^2} \cdot U(s) \quad (10)$$

  This model structure has *e.g.* already been successfully used in Boiroux et al. (2017). To facilitate the further discussion, in the following model (1) will be referred to as PM1, whereas the alternative process model structure (10) is going to be called PM2. As for PM1, two parameter sets of PM2 are identified for each patient, using either a Kalman filter for estimating the initial state at the start of a prediction (KF-PM2), or the combination with a global AR model (AR-PM2).

2.4 Data Pre-Selection

To find suitable data segments for the identification of process models (PM1 and PM2), a preprocessing procedure is used that tries to evaluate the quality of the recorded data. An algorithm checks for a multitude of possible flaws in the data and marks the corresponding areas. The following flaws in the data are automatically screened for:

- *Double-entries:* In case multiple entries with equal amount of carbohydrates or insulin are recorded within 5 minutes in time, these are marked in the data as suspicious.
- *Unrealistically big carbohydrate and insulin inputs:* Carbohydrate and bolus insulin intakes with a value bigger than a specific threshold are marked as suspicious in the data. The threshold is calculated as 1.5 times the interquantile range above the 75% quantile of all carbohydrate and insulin values of the specific patient.
- *Unrealistic carbohydrate to insulin ratios:* The base for this check is an estimate for the expected bolus insulin amount (BI) which is computed according to the following formula from Walsh and Roberts (2013):

$$BI_{\mathrm{expected}} = \frac{CHO}{CIR} + \frac{\Delta G}{ISF} \quad (11)$$

  with the following elements:
  · Carbohydrate content of meal (CHO)
  · Deviation of the CGM value from a nominal target BG value
  ($\Delta G = CGM - BG_{\mathrm{target}}$, with BG_{target} set to 110 mg/dl)

  The CIR and ISF values in (11) are the ones the patients rely upon for computing their bolus insulin needs. The data segment is marked as suspicious in case the actually injected bolus insulin amount for a meal BI differs more than 40% from the expected value BI_{expected}, *i.e.* if the BI is outside the following bounds:

$$0.6 \cdot BI_{\mathrm{expected}} < BI < 1.4 \cdot BI_{\mathrm{expected}} \quad (12)$$

- *CGM rise without any meal input:* In order to detect an invalid CGM rise, *i.e.* a CGM rise without any meal input (which is an indication of an incomplete dataset) the following algorithm is used:
  (1) Filter the original CGM time signal $y_{\mathrm{cgm}}(t)$ using a Savitzky-Golay-Filter (SGF) with filter parameters d and w. This results in a filtered signal $y(t)$.
  (2) Detect all minima and maxima in the filtered signal $y(t)$.
  (3) For each rising segment (minimum until next maximum), consider the following steps:
    (a) Consider the following points in time (see Fig. 1):
      · t_{min}: time where a local minimum occurs in $y(t)$
      · t_{max}: time where the next local maximum occurs in $y(t)$ after t_{min}
      · t_1: time where $\dot{y}(t)$ has largest value in the interval $[t_{\mathrm{min}}, t_{\mathrm{max}}]$
      · t_2: time where $\ddot{y}(t)$ has largest value in the interval $[t_{\mathrm{min}}, t_{\mathrm{max}}]$
      · $t_{\mathrm{start}} = t_{\mathrm{min}} - \Delta T$
      · $t_{\mathrm{end}} = t_2 + \Delta T$
      · $\Delta y = y(t_{\mathrm{max}}) - y(t_{\mathrm{min}})$
    (b) The carbohydrates input signal is denoted $c(t)$
    (c) If the condition

$$\Delta y > \Delta y_{\mathrm{min}} \wedge \dot{y}(t_1) > \dot{y}_{\mathrm{min}} \wedge \sum_{t=t_{\mathrm{start}}}^{t_{\mathrm{end}}} c(t) \leq c_{\mathrm{min}} \quad (13)$$

    is satisfied, the time interval $[t_{\mathrm{min}}, t_{\mathrm{max}}]$ is considered an invalid rise in CGM and the data segment is marked as suspicious.

  Table 1 lists the parameters of this algorithm.

Besides the set of rules used to detect incomplete or erroneous data segments, additionally, a check is implemented to identify suitable starting points for the model identification. Such a valid starting point is a meal with at least

20 g of carbohydrates and with a simultaneous injection of an appropriate amount of bolus insulin (as determined via the checks described before). Data segments used for the model identification commence with a valid starting point and end with a data point that is marked as suspicious, a hole in the CGM trace (longer than 30 minutes), the next marker for a valid starting point, or after 2 hours (whatever occurs first). Such data segments must be 30 minutes or longer in order to be considered for the process model identification.

Table 1. Parameters of algorithm to detect CGM rises w/o meal input.

Parameter	Symbol	Unit	Value
SGF degree	d	1	3
Window size of SGF	w	min	120
Rate-of-change threshold	$\dot{y}_{\min}$	mg/dl/min	0.75
Threshold for min. CGM rise	$\Delta y_{\min}$	mg/dl	30
Time shift	ΔT	min	30
Min. amount of carbs	$c_{\min}$	g	0

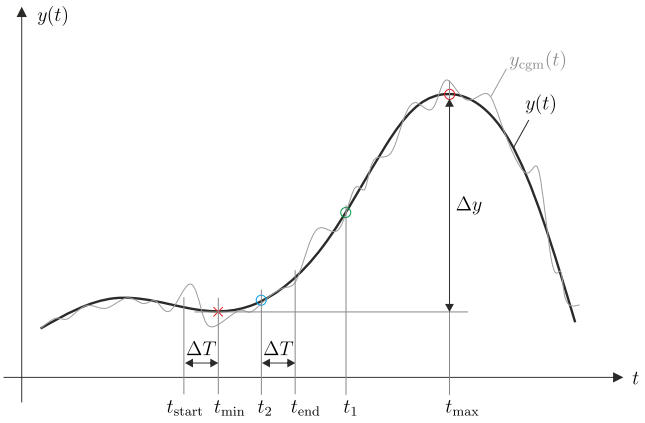


Fig. 1. Illustration of algorithm used to detect invalid rises in CGM data.

3. EXPERIMENTAL SETUP

3.1 Data

For the current work, data of 175 patients collected during an 8 week outpatient clinical trial are used. All patients had T1DM and were on MDI therapy. During the time of the trial patients were equipped with a Dexcom G5 CGM (sampling time T_s: 5 minutes) and kept an electronic diary about their meal intakes (including timing and estimated carbohydrate content), as well as about their injections of basal and bolus insulin (timing and insulin amount). Furthermore, information about CIR and ISF values used by the patients to compute their bolus insulin needs have been documented.

3.2 Model Identification

The data of each patient is split into half. Only the first 28 days of data are used for model identification, whereas the remaining length of the dataset (typically 28 days of data as well) are used for testing and validation of the identified prediction models. Patient-specific process models PM1 and PM2 are identified for each patient using the procedures outlined in Sec. 2.2. Preliminary test computations revealed that prediction performance of process models can be increased significantly by using a different set of model parameters for different times of the day. Therefore, separate model parameter sets are identified for breakfast, lunch and dinner time. Breakfast models are identified from data segments with a starting point between 5:30am and 10:30am, lunch models from those with a starting time between 10:30am and 2:30pm, and dinner models from those with a starting time between 5:00pm and 9:00pm. In order to obtain not only good prediction results, but also sensible model parameters for PM1 and PM2, the values of K_1, T_1, K_2 and T_2 are restricted in the optimization to a physiologically meaningful range. The corresponding parameter limits used for this purpose are summarized in Table 2. Furthermore, the initial guess for the parameter values as used in the optimization procedure can be found in the same table. CIR and ISF in the table refer to the values actually used by the patients during the trial to compute their bolus insulin needs. For the fine tuning of the Kalman filters for KF-PM1 and KF-PM2 matrices Q and R have to be chosen. For this purpose diagonal matrices are used with diagonal entries 10^{10} (R, both settings), 30 (Q for KF-PM1) and $3 \cdot 10^6$ (Q for KF-PM2), respectively. Furthermore, for computing predictions with PM1 and PM2, as well as with the global AR model an estimate of the basal glucose level G_b is required. Here, a patient-specific estimate is computed as the mean CGM value of the first 28 days of data of each patient.

Table 2. Constraints and initial guesses for PM1 and PM2.

		PM1	PM2
K_1 (mg/dl/g)	lower bound	0.5	100
	upper bound	10	6000
	initial guess	CIR/ISF	$271.83 \cdot$ ISF/CIR
T_1 (min)	lower bound	10	20
	upper bound	60	200
	initial guess	29.5	100
K_2 (mg/dl/IU)	lower bound	-100	-80000
	upper bound	-10	-1300
	initial guess	$-$ISF	$-489.29 \cdot$ ISF
T_2 (min)	lower bound	25	50
	upper bound	250	300
	initial guess	56.7	180

3.3 Model Validation and Testing

From the 28 days of the identification dataset 12 sets of patient-specific process model parameters are identified (KF-PM1, KF-PM2, AR-PM1 and AR-PM2, each with parameters identified from breakfast, lunch and dinner data), as well as a second order global AR model from the combined dataset of all patients. These identified models are subsequently tested on the remaining data of each patient. The testing is restricted to the prediction of postprandial glucose excursions, with the prediction made at the time of the meal ingestion. Only those postprandial data segments are considered that fulfill the quality criteria specified in Sec. 2.4. Results are evaluated for breakfast, lunch and dinner events separately. For the process models

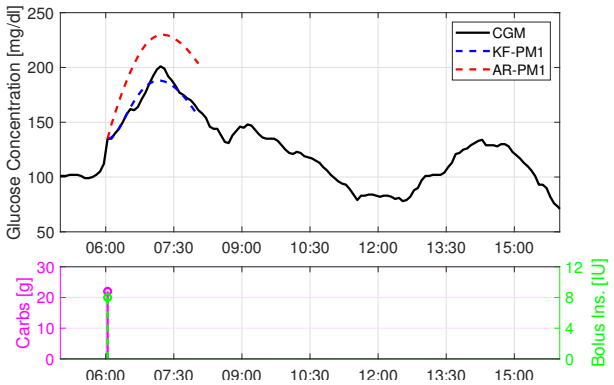


Fig. 2. Example for a predicted postprandial glucose trajectory (Patient 043, day 37).

the predicted glucose trajectories are computed using the set of model parameters identified for the corresponding meal type. For each meal event considered in the validation part of the data the postprandial glucose trajectory is evaluated up to the next system input, up to the detection of an invalid CGM rise (as described in Sec. 2.4) or up to two hours after the meal event (whatever occurs first). For the assessment of the considered prediction models the performance indicator $P_{15/15}$ is calculated. For a specific patient and prediction horizon $P_{15/15}$ is computed as follows:

$$p_k = \begin{cases} 1 & \text{if } \begin{array}{l} y_k \leq 100\text{mg/dl and } \hat{y}_k \leq 100\text{mg/dl} \\ \text{and } |y_k - \hat{y}_k| \leq 15\text{mg/dl} \end{array} \\ 1 & \text{if } \begin{array}{l} (y_k > 100\text{mg/dl or } \hat{y}_k > 100\text{mg/dl}) \\ \text{and } 0.85 y_k \leq \hat{y}_k \leq 1.15 y_k \end{array} \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

$$P_{15/15} = \frac{100}{N} \sum_{k=1}^{N} p_k \quad (15)$$

with y_k the measured CGM value, $\hat{y}_k$ the corresponding predicted glucose value and N the total number of points. $P_{15/15}$ corresponds to the percentage of predicted glucose points within ±15% of the measured CGM value (or within ±15 mg/dl below 100 mg/dl).

4. PREDICTION RESULTS AND DISCUSSION

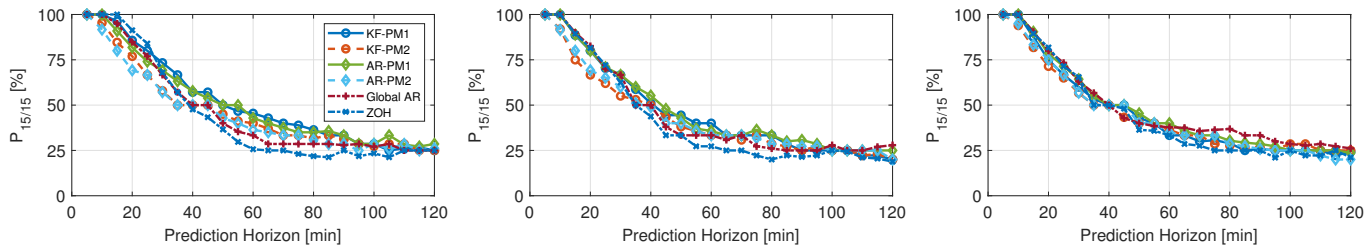
An example for predicted glucose trajectories for the PM1 and a comparison with the corresponding postprandial CGM measurements from the validation section of the clinical data can be seen in Fig. 2. The prediction obtained with setting KF-PM1 is shown in blue, whereas the one for setting AR-PM1 is displayed in red. The bottom panel shows the carbohydrate (in magenta) and bolus insulin (in green) intakes for the considered meal event. The trajectory for KF-PM1 matches pretty well with the CGM data, whereas there are some discrepancies for setting AR-PM1, see Fig. 2. In this case the predicted trajectory of AR-PM1 is outside the boundaries of $P_{15/15}$ for most of the points. It should be noted though that the postprandial excursion for the meal event depicted in this plot is especially difficult to predict, seen that there is some inflection in the CGM data just around the time of the meal intake. A summary of the overall performance of the analyzed process models, as well as of the chosen reference prediction models can be seen in Fig. 3. In the three subplots the median $P_{15/15}$ of all models (median over all 175 patients) is shown as a function of the prediction horizon for breakfast, lunch and dinner separately. Furthermore, for easier analysis, Table 3 lists the median performance of the analyzed models for three specific prediction horizons. It can be seen from the results that the main advantage of the analyzed process models is for medium prediction horizons between 30 and 90 minutes. It is also interesting to see that the process model performance is very different for the three analyzed meal types. By far the biggest advantage of the process models over simpler prediction models can be achieved for the breakfast events. For lunch there is only a relatively small performance gain for medium prediction horizons. For dinner events on the other hand no real advantage of the analyzed process models can be achieved as compared to the simple global AR model. There are several explanations for this observation. First of all, mixed meal composition at breakfast tends to be the least variable across days, with many people eating the same food items for breakfast every day. Secondly, because of their high contribution of fast absorbing carbohydrates, breakfasts tend to give the most pronounced postprandial peaks which is advantageous at the model identification step. Thirdly, before breakfast the system's initial state is the closest to basal, which makes it easier to perform a prediction of the postprandial glucose trajectory. Comparing the performance of PM1 and PM2, results tend to be somewhat better for PM1. This might be surprising at first, seen that the integrating behavior of PM1 is strictly speaking not physiological. However, for the analyzed prediction horizons up to 2 hours this effect of PM1 is not really relevant for its prediction outcomes. Interestingly, the performance of PM2 is especially low for very short prediction horizons. Among the two analyzed options for predicting the impact of the initial state, the use of a Kalman filter tends to give slightly better outcomes than the hybrid process and AR modeling approach, but differences are small.

Table 3. Median $P_{15/15}$ results of considered prediction models. The highest values in the respective category are displayed in bold.

		KF		AR			ZOH
		PM1	PM2	PM1	PM2	Global	
breakfast	45 min	**57.1**	50.0	53.9	50.0	50.0	43.3
	60 min	**45.5**	40.0	42.9	36.4	33.3	25.7
	90 min	**33.3**	31.3	**33.3**	28.6	27.9	25.0
lunch	45 min	44.4	43.3	**47.9**	40.0	38.0	33.3
	60 min	**40.0**	33.3	35.7	33.3	33.3	27.3
	90 min	25.0	25.0	**30.8**	27.3	25.0	21.4
dinner	45 min	**50.0**	43.3	**50.0**	**50.0**	42.9	48.1
	60 min	33.3	36.4	**40.0**	35.0	37.8	33.3
	90 min	25.8	26.7	**28.6**	25.8	33.3	25.0

5. CONCLUSIONS AND OUTLOOK

The current paper analyzes the identification of control-oriented graybox process models and their application for the prediction of postprandial glucose trajectories. More specifically, the focus is on predictions made just at the time of the meal ingestion. Two different process model structures are considered for this purpose, the well-



(a) Breakfast events (b) Lunch events (c) Dinner events

Fig. 3. Median performance ($P_{15/15}$) of considered prediction models for different meal times.

known Kirchsteiger model PM1, as well as an alternative process model without integrating behavior PM2. The model identification is done using outpatient data from T1D patients on MDI therapy, relying on manual diary entries as information about the system inputs, which leads to a more realistic data quality compared to studies under well-controlled clinical conditions. In order to allow for a reliable model identification from such error-prone and incomplete data, a dedicated preprocessing algorithm is proposed here to screen the data for segments that look complete and sensible and are therefore suitable for the identification task. Analyzing the prediction performance of the identified process models and comparing it to that of simple reference prediction models (ZOH and second order AR), it is found that the considered process models perform especially well for breakfast events, and less so for lunch and dinner. Among the two analyzed process models, PM1 performs better at predicting postprandial glucose values than PM2.

ACKNOWLEDGEMENTS

This work has been supported by Roche Diabetes Care.

REFERENCES

Bock, A., François, G., and Gillet, D. (2015). A therapy parameter-based model for predicting blood glucose concentrations in patients with type 1 diabetes. *Computer Methods and Programs in Biomedicine*, 118, 107–123.

Boiroux, D., Duun-Henriksen, A.K., Schmidt, S., Nørgaard, K., Poulsen, N.K., Madsen, H., and Jørgensen, J.B. (2017). Adaptive control in an artificial pancreas for people with type 1 diabetes. *Control Engineering Practice*, 58, 332–342.

Cescon, M., Johansson, R., Renard, E., and Maran, A. (2014). Identification of individualised empirical models of carbohydrate and insulin effects on t1dm blood glucose dynamics. *International Journal of Control*, 87(7), 1438–1453.

Gondhalekar, R., Dassau, E., and Doyle III, F.J. (2016). Periodic zone-mpc with asymmetric costs for outpatient-ready safety of an artificial pancreas to treat type 1 diabetes. *Automatica*, 71, 237–246.

Kirchsteiger, H. and del Re, L. (2014). A model based bolus calculator for blood glucose control in type 1 diabetes. In *Proceedings of the 2014 American Control Conference (ACC)*, 5465–5470. doi: 10.1109/ACC.2014.6858980.

Kirchsteiger, H., Castillo Estrada, G., Pölzer, S., Renard, E., and Del Re, L. (2011). Estimating interval process models for type 1 diabetes for robust control design. In *Proceedings of the 18th IFAC World Congress*, 11761–11766.

Kirchsteiger, H., Johansson, R., Renard, E., and del Re, L. (2014). Continuous time interval model identification of blood glucose dynamics for type 1 diabetes. *International Journal of Control*, 87(7), 1454–1466.

Oviedo, S., Vehí, J., Calm, R., and Armengol, J. (2017). A review of personalized blood glucose prediction strategies for t1dm patients. *International Journal for Numerical Methods in Biomedical Engineering*, 33(6), e2833.

Percival, M.W., Bevier, W.C., Wang, Y., Dassau, E., Zisser, H.C., Jovanovič, L., and Doyle III, F.J. (2010). Modeling the effects of subcutaneous insulin administration and carbohydrate consumption on blood glucose. *Journal of Diabetes Science and Technology*, 4(5), 1214–1228.

Reiterer, F., Kirchsteiger, H., Assalone, A., Freckmann, G., and del Re, L. (2015a). Performance assessment of estimation methods for CIR/ISF in bolus calculators. In *Proceedings of the 9th IFAC Symposium on Biological and Medical Systems (BMS)*, 231–236.

Reiterer, F., Kirchsteiger, H., Freckmann, G., and del Re, L. (2015b). Identification of diurnal patterns in insulin action from measured CGM data for patients with T1DM. In *Proceedings of the 2015 European Control Conference (ECC)*, 1–6. doi: 10.1109/ECC.2015.7330516.

Reiterer, F. and Freckmann, G. (2019). Advanced carbohydrate counting: An engineering perspective. *Annual Reviews in Control*. doi:10.1016/j.arcontrol.2019.06.003.

Tárník, M., Bátora, V., Jørgensen, J., Boiroux, D., Miklovičová, E., Ludwig, T., Ottinger, I., and Murgaš, J. (2015). Remarks on models for estimating the carbohydrate to insulin ratio and insulin sensitivity in T1DM. In *Proceedings of the 2015 European Control Conference (ECC)*, 31–6. doi:10.1109/ECC.2015.7330521.

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

van Heusden, K., Dassau, E., Zisser, H.C., Seborg, D.E., and Doyle III, F.J. (2011). Control-relevant models for glucose control using a priori patient characteristics. *IEEE Transactions on Biomedical Engineering*, 59(7), 1839–1849.

Walsh, J. and Roberts, R. (2013). *Pumping Insulin*. Torrey Pines Press, San Diego, CA, 5.1 edition.