Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Kalman filter and sliding mode observer in artificial pancreas: an in-silico comparison ⋆

I. Sala-Mira * **M. Siket** ** **Gy. Eigner** ** **J. Bondia** *,*** **L. Kovacs** **

* *Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, 46022 Valencia, Spain (e-mail: ivsami@upv.es, jbondia@isa.upv.es)*
** *Physiological Controls Research Center, Research and Innovation Center, Obuda University, Budapest, Hungary (email: {siket.mate,eigner.gyorgy,kovacs.levente}@nik.uni-obuda.hu*
*** *Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Madrid, Spain.*

Abstract: Observers are essential in artificial pancreas systems, either for states or disturbance estimation. Kalman filters based on the Hovorka model are largely applied for this purpose. However, simpler approaches can be used too. We intend to analyze whether the observer structure, the applied model and the individualization of the model parameters affect the estimation accuracy. We perform an in-silico comparison between two Kalman filters and a sliding mode observer, as observer structures previously proposed in the field. All observers are implemented in population and individualized versions of the Hovorka and the simpler Identifiable Virtual Patient models. To tune the Kalman filters, a genetic algorithm based framework was developed. The results indicate that the choice of the model has a larger effect on the outcome than the choice of the observer structure. Finally, the observer based on the Hovorka model does not always perform the most accurate estimation given the higher difficulty during the identification of its parameters.

Keywords: type 1 diabetes, artificial pancreas, state and disturbance observers, Kalman filters, sliding mode observers.

1. INTRODUCTION

Type 1 diabetes mellitus is a condition in which the insulin-producing β-cells of the pancreas are destroyed during an autoimmune reaction not known completely yet. Due to the subsequent deficiency in insulin, exogenous insulin injections are required to keep the balance of the metabolic homeostasis. The automatic glucose regulation – the so-called artificial pancreas (AP)– has been investigated as an alternative to traditional basal-bolus therapies, to improve the glucose control, while reducing the burden for the patients. Unlike the glucose concentration that can be measured in real-time, the unavailability of the remaining state variables makes the use of observers a must in AP. The applications of observers range from state feedback control (Kovacs et al. (2019)) to insulin limitation for hypoglycemia avoidance (Hajizadeh et al. (2018)) or meal detection (Ramkissoon et al. (2018)).

The design of Kalman filters (KF) using Hovorka's model in Hovorka et al. (2004) is well-accepted for state estimation in AP (Bondia et al. (2018)). This research intends to analyze if the complexity of both the Kalman-type observer and the Hovorka model justify their wide application compared to other simpler approaches, such as the sliding mode observer based on the Identifiable Virtual Patient model (IVP, Kanderian et al. (2009)) presented in (Sala-Mira et al. (2019)). In this regard, an in-silico comparison is carried out in Section 3 in the UVA/PADOVA simulator (Dalla Man et al. (2014)) to evaluate the differences in terms of estimation error between three observers presented in Section 2 – Dual KF (DKF), Joint KF (JKF) and the nonlinear sliding mode observer (NSMO) – when they are designed with either the IVP model or the Hovorka model. As an additional goal, the effect on the observer accuracy of model individualization compared to the use of average models is explored in this analysis. In order to analyze all these aspects, we developed a genetic algorithm based framework for parametric identification and tuning of the KFs, where our goal was to provide uniform conditions for both the models and the observers.

⋆ This work was supported by the Spanish Ministry of Economy, Industry and Competitiveness (MINECO) [grant number DPI2016-78831-C2-1-R]; the European Union [FEDER funds]; Generalitat Valenciana [grant numbers ACIF/2017/021,BEFPI/2019/077 and and BEST/2019/173]. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 679681) Corresponding author: J. Bondia (jbondia@isa.upv.es).

2. METHODS

2.1 Model identification

To investigate the effect of model complexity on the performance of observers, the Hovorka model in Hovorka et al. (2004) was compared with the simpler IVP model in Kanderian et al. (2009). Before the identification of the models, the GenSSi software (Ligon et al. (2018)) was employed to perform the identifiability analysis. Both models were locally structurally identifiable, if the plasma insulin (Ip) was considered as a known signal, besides the glucose and the insulin infusion. Moreover, the corresponding carbohydrate absorption models were not included during the identification of the models, since the observers do not employ such information. All the parameters not appearing in the absorption model were considered during the identification of the IVP model: the EGP (endogenous glucose production), $GEZI$ (glucose effectiveness at zero insulin), p_2 (related to the delay in insulin action), C_I (insulin clearance), S_I (insulin sensitivity) and the insulin pharmacokinetics time constants τ_1 and τ_2. Given the large number of parameters in the Hovorka model, a subset of them was selected according to the global parameter sensitivity analysis performed with the Elementary Effect Test (Pianosi et al. (2015)). The selected parameters were: V_I (insulin distribution volume), V_G (glucose distribution volume), k_{12} (transfer rate between glucose compartments), t_{maxI} (time-to-maximum insulin absorption), k_e (fractional elimination rate of insulin) the constants k_{b1}, k_{b2}, k_{a1} and k_{a2} defining the insulin sensitivity. The remaining parameters were set to the values in Hovorka et al. (2004). The identification utilized the default settings of the genetic algorithm in Matlab 2018b. The cost function consisted of two terms: the sum of the normalized root-mean-squared error (NRMSE) of the glucose measurement (CGMS) and the Ip. The reference signals that were used to compute these metrics, as well as the insulin infusion and the rate of glucose appearance (Ra), were collected from a 3-meal simulation of the average patient in the UVA/Padova simulator (Dalla Man et al. (2014)).

To account for the inter-patient variability, in the next stage, the most sensitive parameters of the previous models were individualized. Based on sensitivity analysis, the S_I and C_I were selected for the IVP case, and k_{a1} and k_e for the Hovorka model. The selected parameters were identified for the 10 adult patients in the academic version of the UVA/PADOVA simulator using the same procedure applied above.

2.2 Investigated observers

We have compared three observers, the first one was the NSMO applied in (Sala-Mira et al. (2019)). This observer allows a robust state estimation despite the presence of matching disturbances if the discontinuous gain is larger than the bound of the disturbance. Since the number of outputs in this study coincides with the number of disturbances, unmeasurable states are left in open loop (Wu et al. (2012)). Moreover, the observer provides a simple estimation of the disturbance based on a low pass filter. To reduce the chattering phenomena, an implicit Euler discretization was implemented (Sala-Mira et al. (2019)).

Additionally, the JKF and DKF were both realized (Haykin (2001)). Both of the KFs employed the discretized quasi-Linear Parameter Varying state-space representation of the corresponding models (Gáspár et al. (2017)). However, their discretization techniques are different, the DKF uses exact discretization while the JKF uses Euler method as described in Toth et al. (2010). The system matrix augmented with the disturbance – which is considered to be static – is non-invertible, hence exact discretization cannot be applied in the JKF. Notable difference from tuning point of view is in the Q covariance matrix. Adding one parameter as a variable to be estimated increases the dimension of the Q matrix by one in the JKF, which means one more parameter to tune even if we take into account only the main diagonal. For the parameter estimation, the DKF has a separate Q covariance matrix and a λ forgetting factor, resulting in two more variables to tune (for further details see Haykin (2001)).

2.3 Kalman filter tuning

Two types of KFs, two models and three tuning scenarios were considered in this study. This resulted in a total of 12 different tunings. Henceforth, we refer to the tuning scenarios as "30Diag", "30Ra" and "UVADiag" followed by the corresponding abbreviation of the filter (DKF or JKF) and model (IVP or Hovorka). The three tuning scenarios are different in terms of parameter uncertainty, structural difference and Q covariance matrix complexity. Two configurations ("30Diag" and "30Ra") had no structural mismatch between the model propagated by the observer and the virtual patient. Dissimilarity in these cases was achieved by applying fixed +30% variability in all the model parameters. Based on the literature reviews of Hovorka et al. (2002) and Kanderian et al. (2009) the 30% variability in the model parameters are in physiologically relevant ranges. The "30Diag" took into account every element in the main diagonal of the Q matrix, the "30Ra" only the elements corresponding to the blood glucose level and to the Ra. The investigation was complemented by the latter one in order to see the effect when the estimation of the inner state variables converge to open-loop operation estimating only the Ra. This KF setup resembles the operation of the applied NSMO in the sense that the unmeasurable state variables are estimated in open-loop in the sliding mode observer. The third configuration ("UVADiag") includes structural mismatch between the model of the observer and the virtual patient. The observers propagated the average model (IVP or Hovorka) – their identification process is described in Subsection 2.1 –, while the reference was one specific patient of the UVA/PADOVA simulator. The "UVADiag" setups utilized the full main diagonal of the Q covariance matrix.

The basic idea behind the tuning of the KFs was to provide uniform conditions and to minimize the effect of the human factor. To this end, a tuning framework was developed. The framework is fundamentally an optimization done by the genetic algorithm of Matlab 2018b. During the tuning process, the optimization horizon was a 24-hour-long simulation with one meal, but we only considered the

transients and neglected steady-states. The parameters to tune – genes of an individual – were the elements in the main diagonal of the Q covariance matrix. The number of parameters in the optimization ranged between 2 and 10. The lower bounds of the parameters were set to zero. The upper bounds were chosen by fixing each diagonal element of Q to be the only nonzero element separately, and setting them large enough so that the KF converges to the measurements.

The cost functions of the genetic algorithm were defined as the weighted sum of the NRMSE between the steady-state-offset-compensated estimations and the actual virtual patient. There was a need to compensate for the steady state offset, because the filter cannot alter the steady state, hence it will modify the transient in order to reduce the cost. Meanwhile, the dynamic behavior and the steady state offset together affect the NRMSE. This could result in unfavorable, distorted transient, so only the dynamic behavior was taken into account while calculating the cost function. To give a tangible example, the x_2 state variable of the Hovorka model is shown in Fig. 1, where the case "A" refers to the tuning without considering steady-state, while case "B" refers to the tuning considering steady-state. The estimation denoted by "B" provides lower NRMSE compared to the estimation denoted by "A", however, the insulin decreases significantly after the meal intake for case "B", thus case "A" shows a more realistic waveform.

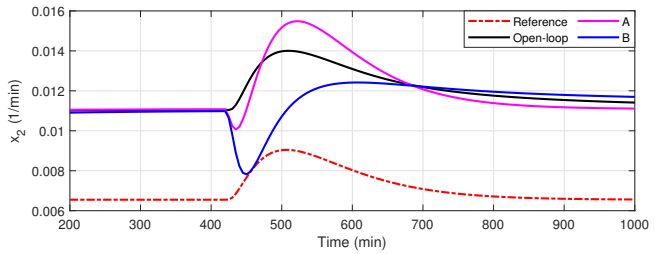

Fig. 1. Effect of not compensating the steady state offset.

Of note, this tuning method is applicable in-silico, since there could be virtual or non-measurable state variables. Our goal was not to provide a tuning algorithm, but to provide uniform tuning method during the investigation of the observers and models. There is a significant difference in the number of state variables between the models, the Dalla Man model has 20 state variables, the Hovorka has 8 state variables and the IVP has 4 state variables. There are three equivalent variables amongst the models (with different units of measurement), namely the CGMS, Ip and Ra. The weights are introduced in the cost function (3)-(5) discussed later to reduce the distortion effect caused by the increased number of terms in the Hovorka model. By applying the coefficients $\frac{9}{5}$ and $\frac{3}{5}$ in (4), the blood glucose level, plasma insulin and Ra contribute the same amount of cost in the Hovorka and in the IVP model.

Furthermore, a measure of oscillation is formulated. It has been observed that the genetic algorithm tended to overfit the estimation, hence small-amplitude oscillation could occur. In a practical sense direction changes were calculated during the simulations in the following way:

$$x_{osc} = \sum_{i=3}^{n} \mid \mathrm{sgn}(x_i - x_{i-1}) - \mathrm{sgn}(x_{i-1} - x_{i-2}) \mid, \quad (1)$$

$$k = \begin{cases} 10, & \text{if } x_{osc} > threshold \\ 1, & \text{otherwise} \end{cases} \quad (2)$$

where n is the number of samples and the thresholds were dependent on the number of meals. The k penalizing coefficient increases the cost by an order of magnitude, pushing the genetic algorithm to find a solution under the oscillation threshold. We found that the CGMS noise does not affect significantly the final result, thus the generated measurement data was the true blood glucose values of one of the models. Taking into account the aforementioned factors, the following cost functions have been developed for each virtual patient model:

$$J_{IVP} = k(G + I_{eff} + I_p + I_{sc} + Ra), \quad (3)$$

$$J_{Hovorka} = k\Big(\frac{9}{5}(G + I_p + Ra) + \frac{3}{5}(Q_2 + x_1 + x_2 + x_3 + S_1 + S_2)\Big), \quad (4)$$

$$J_{UVA} = k(G + I_p + Ra), \quad (5)$$

where the notations of the model state variables (see the corresponding notations in Hovorka et al. (2004) and Kanderian et al. (2009)) represent their corresponding steady-state-offset-compensated NRMSE, k is the oscillation penalizing coefficient.

2.4 In-silico comparison

The purpose of the comparison is threefold: 1) to analyze the differences between observers and their tunings, 2) to determine if a simpler model could be used in the observer design, and 3) to study the feasibility of population models in the observer design. From a statistical point of view, this is equivalent to finding the significance of the factors "Observer", "Model" and "Personalization". To that end, the UVA/PADOVA simulator, extended by different sources of variability, was employed to simulate a 3-meal scenario for the 10 virtual adults in the simulator, for every combination of the factor levels. The scenario included measurement noise – according to the default built-in CGMS model – and variability in insulin sensitivity and insulin absorption. Moreover, the initial condition of the observers was set to -20% of the basal conditions.

The comparison was based on the RMSE of the CGMS, Ip and the Ra. The references for RMSE computation were collected from the Dalla-Man model which is implemented in the simulator (Dalla Man et al. (2014)), while the estimated signal units were modified accordingly to match the Dalla-Man ones.

To determine whether the factors are significant, the linear mixed-effect model (LMM) approach was applied. This technique is able to deal with the lack of independence in the data, i.e., the same cohort was used in each simulation. The LMM is a generalization of the linear model in which the regression coefficients could depend on a specific factor (Winter (2013)). In this analysis, only the intercept was made dependent on the subjects, i.e., random intercept. The fixed effects are the "Observer", the "Model" and the "Personalization", as well as, their pairwise interactions.

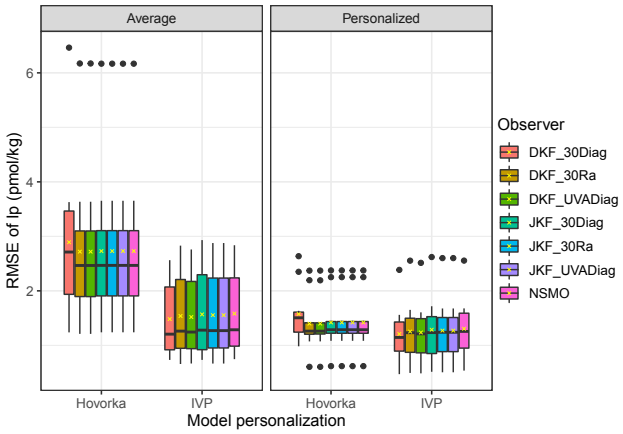

Fig. 2. Grouped boxplot for the plasma insulin RMSE. The length of the box corresponds to the interquartile range, the black solid line is the median and the yellow cross is the mean.

Higher order interactions were not considered here, because they were not found to be statistically significant. Three LMM's were fitted for the RMSE of the three considered signals. The general LMM reads as,

$$\begin{aligned} RMSE_{sub} = \alpha_{sub} &+ \beta_1 \cdot Model \\ &+ \beta_2 \cdot Personalization + \beta_3 \cdot Observer \\ &+ \beta_4 \cdot Personalization \cdot Model \\ &+ \beta_5 \cdot Model \cdot Observer \\ &+ \beta_6 \cdot Personalization \cdot Observer + \epsilon \end{aligned} \tag{6}$$

where α_{sub} is the random intercept dependent on the subject *sub*, the β's terms are the coefficients for the fixed effects, and ϵ, the normally distributed residuals.

Given the non-normality of the residuals, the coefficients of the LMM's in (6) were fitted by the *rlmer* function provided in the *robustlmm* package in R (Koller (2016)). Equation (6) determines the statistical significance of the factors: A factor (or interaction) was considered significant if the corresponding β coefficient included the 0 value in the 95% Wald confidence interval (Akdur et al. (2016)). Finally, a pairwise comparison between the significant factors was completed by using the Wilcoxon signed rank test for paired data with the Benjamini-Hochberg p-value correction (Lee and Lee (2018)). The significance level was set to 0.05.

3. RESULTS

3.1 Plasma insulin estimation

Fig. 2 overviews the plasma insulin RMSE for each combination of the factor levels. The observers based on the average Hovorka model lead to the largest RMSE. As shown in Table 1, either the use of the IVP model or the application of personalization, significantly improve the RMSE by a margin of 0.9 pmol/kg. The personalization of the IVP model also decreases the RMSE regarding the average case, but this reduction is less remarkable compared to the improvement of the personalization in the Hovorka model. Thus, the interaction of "Model" and "Personalization" is also significant. According to the pairwise comparison of this interaction, the personalization of the Hovorka

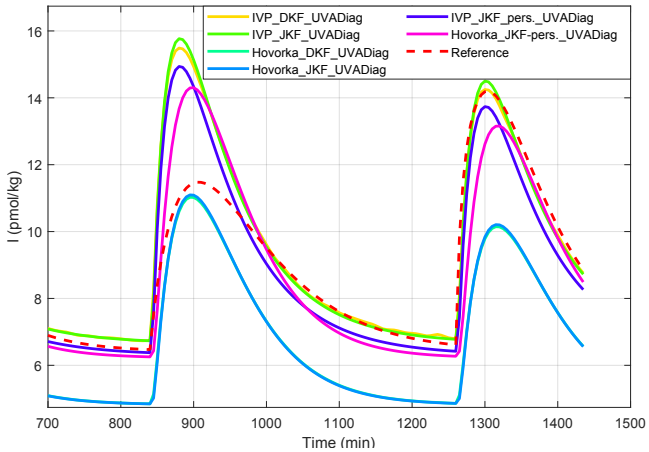

Fig. 3. Effect of the model type and model personalization in the plasma insulin estimation

model underperforms the IVP personalized model, since the former has lower mean than the latter and both groups have been found to be statistically different at 0.05 level. However, no difference in terms of RMSE has been shown between the personalized Hovorka and the average IVP model. The lower error in the IVP model and in the personalized Hovorka models are related to a reduction of the offset in the steady state (see Fig. 3). The personalized Hovorka model, despite reducing this offset, yields a slower estimation compared to the IVP model. That indicates the high dependency of the estimation accuracy on the model, but not on the observer structure. Indeed, the "Observer" factor has not been found to be significant (Table 1). The reason is that the observer gains are near 0 – exactly 0 in the NSMO case –, therefore, the unmeasurable states behave like the open-loop model. The only exception are the DKF_30Diag observers using the Hovorka model. These observers have larger gains, but they also tend to underperform the other observers (see DKF_30Diag boxplot in the left pannel of Fig. 2), although without statistical significance. As stated, for example, in Wu et al. (2012), given a system with disturbances affecting only the measurable states, the addition of an output correction gain would feed those disturbances into the unmeasurable states. Therefore, increasing those gains could amplify the effect of the disturbance on the unmeasurable states.

3.2 Rate of glucose appearance

The individualization of the model reduces the variability in the Ra compared to the average case (Fig. 4), but there is not enough statistical evidence to state that it improves significantly the RMSE, independently of the model or the observer. Conversely, the observer, the model and their interaction results to be significant. The observers using the IVP model tend to have larger RMSE than the ones using the Hovorka model, with the exception of the JKF_UVADiag and the DKF_30Diag. For example, changing the observer to DKF_30Ra, the improvement over the intercept is about 0.35 mg/dL/min (see coefficient Observer.DKF_30Ra in Table 1). However, using the IVP model and the DKF_30Ra the improvement is only 0.08 mg/dL/min (see coefficient Model + Model:Observer.DKF_30Diag in Table 1). This result is

Table 1. Summary of the fixed effects coefficients of the plasma insulin(Ip), rate of glucose appearance (Ra) and glucose measurement (CGMS) root mean-squared error. $\hat{\beta}$ denotes the estimated coefficient value of LMM (6), CI the 95%-Wald confidence interval and t the t-statistic of the test. The intercept corresponds to the DKF observer tuned with the 30Diag and for the average Hovorka model

	Ip			Ra			CGMS		
Coefficients name	$\hat{\beta}$	t	CI	$\hat{\beta}$	t	CI	$\hat{\beta}$	t	CI
Intercept	2.46	11.74	[2.05, 2.87]	2.12	16.67	[1.87, 2.37]	0.57	133.31	[0.56, 0.58]
Model	-0.99	-6.13	[-1.30, -0.67]	-0.45	-3.59	[-0.70, -0.21]	-0.44	-110.65	[-0.45, -0.43]
Personal.	-0.90	-5.60	[-1.22, -0.59]	-0.05	-0.39	[-0.30, 0.20]	-0.00	-0.46	[-0.01, 0.01]
Observer.DKF_30Ra	-0.24	-1.29	[-0.60, 0.12]	-0.35	-2.38	[-0.63, -0.06]	-0.57	-124.96	[-0.58, -0.56]
Observer.DKF_UVADiag	-0.24	-1.29	[-0.60, 0.12]	-0.34	-2.38	[-0.63, -0.06]	-0.57	-124.88	[-0.58, -0.56]
Observer.JKF_30Diag	-0.23	-1.22	[-0.59, 0.14]	-0.30	-2.04	[-0.58, -0.01]	-0.57	-124.95	[-0.58, -0.56]
Observer.JKF_30Ra	-0.22	-1.21	[-0.58, 0.14]	-0.31	-2.12	[-0.59, -0.02]	-0.57	-124.97	[-0.58, -0.56]
Observer.JKF_UVADiag	-0.22	-1.21	[-0.58, 0.14]	0.96	6.59	[0.67, 1.24]	-0.57	-124.97	[-0.58, -0.56]
Observer.NSMO	-0.22	-1.22	[-0.59, 0.14]	-0.40	-2.76	[-0.68, -0.12]	-0.13	-28.82	[-0.14, -0.12]
Model:Personal.	0.58	5.13	[0.36, 0.81]	0.03	0.29	[-0.15, 0.20]	-0.00	-0.07	[-0.01, 0.01]
Model:Observer.DKF_30Ra	0.26	1.24	[-0.15, 0.68]	0.37	2.18	[0.04, 0.69]	0.50	94.11	[0.49, 0.51]
Model:Observer.DKF_UVADiag	0.25	1.16	[-0.17, 0.66]	0.36	2.15	[0.03, 0.69]	0.52	97.41	[0.51, 0.53]
Model:Observer.JKF_30Diag	0.28	1.30	[-0.14, 0.69]	0.44	2.60	[0.11, 0.76]	0.61	115.54	[0.60, 0.62]
Model:Observer.JKF_30Ra	0.26	1.23	[-0.15, 0.68]	1.08	6.43	[0.75, 1.41]	0.44	83.99	[0.43, 0.45]
Model:Observer.JKF_UVADiag	0.26	1.23	[-0.15, 0.68]	-0.29	-1.74	[-0.62, 0.04]	0.44	83.99	[0.43, 0.45]
Model:Observer.NSMO	0.29	1.38	[-0.12, 0.71]	0.54	3.24	[0.21, 0.87]	0.20	37.56	[0.19, 0.21]
Personal.:Observer.DKF_30Ra	0.04	0.20	[-0.37, 0.46]	-0.09	-0.55	[-0.42, 0.24]	0.00	0.36	[-0.01, 0.01]
Personal.:Observer.DKF_UVADiag	0.04	0.21	[-0.37, 0.46]	-0.09	-0.54	[-0.42, 0.24]	0.00	0.36	[-0.01, 0.01]
Personal.:Observer.JKF_30Diag	0.05	0.25	[-0.36, 0.47]	0.04	0.22	[-0.29, 0.37]	0.00	0.36	[-0.01, 0.01]
Personal.:Observer.JKF_30Ra	0.05	0.23	[-0.37, 0.47]	0.05	0.30	[-0.28, 0.38]	0.00	0.36	[-0.01, 0.01]
Personal.:Observer.JKF_UVADiag	0.05	0.23	[-0.37, 0.47]	-0.00	-0.02	[-0.33, 0.33]	0.00	0.36	[-0.01, 0.01]
Personal.:Observer.NSMO	0.05	0.24	[-0.37, 0.47]	-0.09	-0.56	[-0.42, 0.24]	0.00	0.71	[-0.01, 0.01]

consistent with Fig. 5 which shows how the IVP model-based observers underestimate the Ra, while the Hovorka model-based observers present a quicker response with a more accurate postprandial peak estimation. A possible reason for this improvement could be the use of a 2-compartment model in the Hovorka model to explain the glucose metabolism more accurately.

Regarding the comparison between observers, the pairwise analysis reveals that within a model, there is not any statistical difference in terms of RMSE between the DKFs (with the exception of the DKF_30Diag in the Hovorka model) and the NSMO. However, the JKFs exhibit signif-

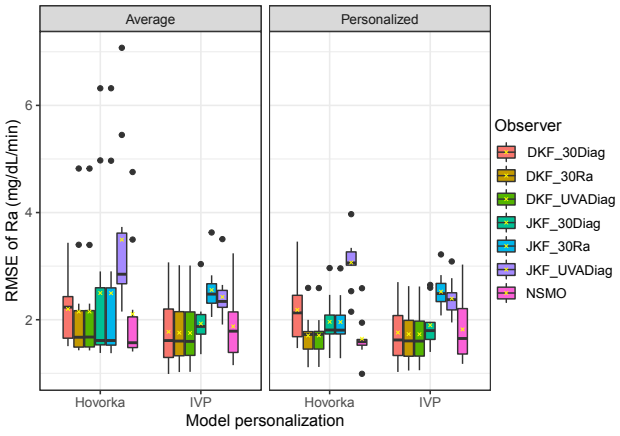

Fig. 4. Grouped boxplot for the rate of glucose appearance root mean-squared error. The length of the box corresponds to the interquartile range, the black solid line is the median and the yellow cross is the mean.

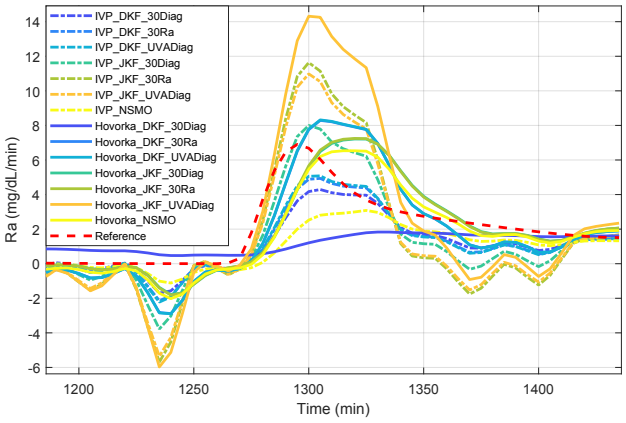

Fig. 5. Effect of the model type and model personalization in the plasma insulin estimation

icantly larger error. Moreover, while there are no remarkable differences between the DKF tunings, there is more variability within the JKF tunings. That might be related to the extra degree of freedom – the forgetting factor – used by the DKF to estimate the disturbance.

3.3 Glucose measurement

The model, the observer and their interaction are significant as observed from the confidence intervals in Table 1. The pairwise comparison determines that all the combinations in the interaction Model-Observer are significantly different. However, the difference in mean between the largest RMSE (Hovorka-based DKF_30Diag) and the lowest one (Hovorka-based DKF_30Ra) is under 0.6 mg/dL, which is negligible from the clinical point of view.

4. CONCLUSIONS

A comparison between three observers (NSMO, DEKF, and JEFK) and the effect in the estimation accuracy of model complexity (IVP vs. Hovorka) and parameter individualization has been addressed. The applied model proved to be a determinant factor in the estimation of plasma insulin, rate of glucose appearance and glucose. The results cannot prove that a more complex model leads always to a more accurate estimation. On the one hand, a simpler model like the IVP could be preferable in the estimation of the plasma insulin since all the parameters can be included in the identification and more accurate identification of the average model can be achieved. On the other hand, increasing the complexity in glucose metabolism could be beneficial in a quicker and more accurate estimation of the rate of glucose appearance. In addition, the parameter individualization improves the accuracy in the plasma insulin estimation. That result could justify further work on real-time parameter estimation. Finally, the differences between observers and their tunings are less significant because of the effect of the model. An example of that is that the "Observer" factor was not significant in the estimation of plasma insulin. That means that the use of the genetic algorithm to tune the covariance matrix did not manage to take advantage of the larger degrees of freedom that the Kalman filters have, compared to the sliding mode with a simpler tuning. We have found differences between observers in the rate of glucose estimation: the JEKF tends to overestimation, while the DEKF and sliding mode perform similarly in terms of RMSE. The major differences between observers were found in the glucose estimation. However, these deviations were not remarkable from the clinical point of view.

ACKNOWLEDGEMENTS

Authors want to thank Dr. Tamás Ferenci (Óbuda University) for his valuable recommendations regarding the statistical analysis. Also, authors are grateful to the reviewers and associate editor for their constructive comments.

REFERENCES

Akdur, H.T.K., Özonur, D., and Bayrak, H. (2016). A Comparison of Confidence Interval Methods of Fixed Effect in Nested Error Regression Model. *SDÜ Fen Bil Enst Der*, 20(2).

Bondia, J., Romero-Vivó, S., Ricarte, B., and Díez, J.L. (2018). Insulin Estimation and Prediction: A Review of the Estimation and Prediction of Subcutaneous Insulin Pharmacokinetics in Closed-Loop Glucose Control. *IEEE Control Syst.*, 38(1), 47–66.

Dalla Man, C., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., and Cobelli, C. (2014). The UVA/PADOVA Type 1 Diabetes Simulator: New Features. 10.

Gáspár, P., Szabó, Z., Bokor, J., and Németh, B. (2017). Modeling of LPV Systems. In P. Gáspár, Z. Szabó, J. Bokor, and B. Nemeth (eds.), *Robust Control Design for Active Driver Assistance Systems: A Linear-Parameter-Varying Approach*, Advances in Industrial Control, 11–70. Springer International Publishing, Cham.

Hajizadeh, I., Rashid, M., Samadi, S., Feng, J., Sevil, M., Hobbs, N., Lazaro, C., Maloney, Z., Brandt, R., Yu, X., Turksoy, K., Littlejohn, E., Cengiz, E., and Cinar, A. (2018). Adaptive and Personalized Plasma Insulin Concentration Estimation for Artificial Pancreas Systems. *J Diabetes Sci Technol*, 12(3), 639–649.

Haykin, S.S. (ed.) (2001). *Kalman filtering and neural networks*. Adaptive and learning systems for signal processing, communications, and control. Wiley, New York.

Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T.R., Schaller, H.C., Schaupp, L., Vering, T., and Wilinska, M.E. (2004). Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol. Meas.*, 25(4), 905–920.

Hovorka, R., Shojaee-Moradie, F., Carroll, P.V., Chassin, L.J., Gowrie, I.J., Jackson, N.C., Tudor, R.S., Umpleby, A.M., and Jones, R.H. (2002). Partitioning glucose distribution/transport, disposal, and endogenous production during IVGTT. *Am. J. Physiol. Endocrinol. Metab.*, 282(5), E992–1007.

Kanderian, S.S., Weinzimer, S., Voskanyan, G., and Steil, G.M. (2009). Identification of Intraday Metabolic Profiles during Closed-Loop Glucose Control in Individuals with Type 1 Diabetes. *J Diabetes Sci Technol*, 3(5), 1047–1057.

Koller, M. (2016). robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. *J. Stat. Soft.*, 75(6).

Kovacs, L., Eigner, G., Siket, M., and Barkai, L. (2019). Control of Diabetes Mellitus by Advanced Robust Control Solution. *IEEE Access*, 7, 125609–125622.

Lee, S. and Lee, D.K. (2018). What is the proper way to apply the multiple comparison test? *Korean J Anesthesiol*, 71(5), 353–360.

Ligon, T.S., Fröhlich, F., Chiş, O.T., Banga, J.R., Balsa-Canto, E., and Hasenauer, J. (2018). GenSSI 2.0: multi-experiment structural identifiability analysis of SBML models. *Bioinformatics*, 34(8), 1421–1423.

Pianosi, F., Sarrazin, F., and Wagener, T. (2015). A Matlab toolbox for Global Sensitivity Analysis. *Environmental Modelling & Software*, 70, 80–85.

Ramkissoon, C., Herrero, P., Bondia, J., and Vehi, J. (2018). Unannounced Meals in the Artificial Pancreas: Detection Using Continuous Glucose Monitoring. *Sensors*, 18(3), 884.

Sala-Mira, I., Díez, J.L., Ricarte, B., and Bondia, J. (2019). Sliding-mode disturbance observers for an artificial pancreas without meal announcement. *Journal of Process Control*, 78, 68–77.

Toth, R., Heuberger, P., and Den Hof, P.V. (2010). Discretisation of linear parameter-varying state-space representations. *IET Control Theory Applications*, 4(10), 2082–2096.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv:1308.5499 [cs]*. ArXiv: 1308.5499.

Wu, T.C., Peng, C.C., and Chen, C.L. (2012). An Adaptive Sliding Mode Observer for Nonlinear Systems Subject to Mismatched Uncertainties. *IFAC Proceedings Volumes*, 45(13), 796–801.