Journal of Artificial Intelligence and Data Science (JAIDA)
e-ISSN: 2791-8335
https://dergipark.org.tr/pub/jaida

Vol. 5, No. 1, pp. 01-11, 2025
Research Article
Received: 01/01/2025 Accepted: 25/03/2025

A Comparative Analysis of Unscented Kalman Filter for Smartphone-Based Multi-Parametric Glucose Prediction

Ece AYFER [1,*], Ömer Atılım KOCA [2], Volkan KILIÇ [3]

Abstract

Diabetes mellitus, a chronic disease affecting millions of people worldwide, requires monitoring and management of glucose levels to reduce the risks of hyperglycemia and hypoglycemia. Technological advancements have enabled the development of various digital tools, including continuous glucose monitors (CGMs) for effective management of this disease. However, these tools only provide alerts after glucose levels exceed critical thresholds, which causes delays in taking necessary precautions. To address this issue, various artificial intelligence (AI)-based models have been developed to predict glucose levels in advance. Traditional AI approaches, however, often rely on standardized datasets, limiting their ability to achieve the accuracy required for individualized treatment. Therefore, it is crucial to develop personalized prediction models that can be trained using the individual data of patients. Here, this paper introduces a personalized glucose prediction approach that employs a three-parameter unscented Kalman filter (UKF) to predict future glucose levels using CGM data, as well as basal and bolus insulin values. Experiments on OhioT1DM dataset show the advantage of our proposed approach over the baseline KF and UKF for glucose prediction in terms of Root Mean Square Error. Furthermore, the proposed approach is embedded into a custom-designed cross-platform smartphone application, *GlucoThinker Advance*, capable of providing offline access to the proposed personalized glucose prediction approach to ensure continuous support without requiring an internet connection.

***Keywords**: Glucose Prediction; Multi-Parameter; Personalized; Smartphone Application; Unscented Kalman Filter.*

1. Introduction

Diabetes mellitus is a chronic disease characterized by the inability of the body to regulate blood glucose levels effectively. This disease causes inefficient use of proteins, fats, and carbohydrates due to a lack of insulin production or ineffective insulin utilization. Additionally, diabetes increases the risk of long-term complications such as vision loss, kidney failure, heart disease, disability, and death [1]. According to the International Diabetes Federation, there are an estimated 537 million diabetes patients, which is expected to reach 643 million by 2030, and 783 million by 2045 [2]. Moreover, the World Health Organization reported in 2017 that diabetes will be the seventh cause of death by 2030 [3]. Thus, the development of tools that support self-management is crucial for mitigating complications and reducing the risk of premature mortality associated with the rising prevalence of diabetes [4]. Continuous Glucose Monitoring (CGM) systems have emerged as a key tool, providing real-time glucose data that enables patients to implement more precise and timely intervention strategies for effective diabetes self-management [5]. CGM systems provide alerts at critical glucose levels, but the time delay between receiving these alerts and implementing the necessary actions can negatively impact metabolic processes [6]. These limitations have shifted attention toward advancements in artificial intelligence (AI) and stochastic filtering, with the goal of developing innovative methods for predicting future glucose levels, thereby enabling timely interventions and enhancing diabetes management [7].

Daniels et al. proposed a multi-task network that integrates a shared convolutional neural network with long short-term memory (LSTM) components [8]. This network is designed to capture both inter-subject features and patient-specific feedforward layers. However, a limitation of this model is the requirement for retraining to include new patients, which increases the computational complexity of the network. Zhu et al. introduced a model-agnostic meta-learning technique for the prediction of future glucose levels in type 1 diabetes patients

*Corresponding author

Ece AYFER; İzmir Katip Celebi University, Faculty of Engineering and Architecture, Electrical and Electronics Engineering Department, Türkiye; e-mail: 220403026@ogr.ikcu.edu.tr; 0009-0006-8098-1583

Ömer Atilim KOCA; İzmir Bakircay University, Faculty of Engineering and Architecture, Computer Engineering Department, Türkiye; e-mail: atilim.koca@bakircay.edu.tr; 0009-0007-7286-6785

Volkan KILIÇ; İzmir Katip Celebi University, Faculty of Engineering and Architecture, Electrical and Electronics Engineering Department, Türkiye; e-mail: volkan.kilic@ikcu.edu.tr; 0000-0002-3164-1981

[9]. However, this approach requires an extra personalized fine-tuning phase for each individual patient. This requirement adds to the data collection workload due to the need for manual labeling by medical professionals. Kim et al. conducted an evaluation of the efficiency of a smartphone-based application for glucose management [10]. The results demonstrated the importance of providing well-designed, high-quality tools, such as smartphone applications, for diabetes management. Perez et al. introduced an AI-based network to predict CGM values [11]. The dependence of this network on a single parameter, CGM readings, constrains its capacity to capture the comprehensive range of glucose fluctuations.

Zecchin et al. presented a neural network-based approach that focused on CGM values, which has been demonstrated to have limited effectiveness in reflecting complex changes in glucose levels [12]. Mohebbi et al. explored short-term glucose prediction with LSTM-based recurrent neural network [13]. Alfian et al. proposed multi-layer perceptron approach for CGM values [14]. This approach uses sliding window technique to make paired inputs and outputs. Rabby et al. presented a deep recurrent neural network using stacked LSTMs combined with Kalman smoothing. This approach demonstrated improved accuracy in glucose level predictions compared to approaches that excluded Kalman filtering [15]. McShinsky et al. conducted a comparative analysis of the Kalman filter (KF) and unscented Kalman filter (UKF) for glucose prediction over 30-minute and 60-minute intervals. It was reported that the KF was more effective for 60-minute predictions, whereas the UKF demonstrated superior performance for 30-minute predictions [16]. KF is designed to predict unknown variables by integrating observed measurements. The effectiveness of the KF relies on its ability to update the state of the system and generate predictions at each time step recursively by employing both past measurements and previous predictions. This recursive framework enables the KF to progressively improve its precision over time. The Kalman Filter has limitations in nonlinear systems due to its mathematical framework, which relies on assumptions of linear state transitions and measurement models.

This assumption fails to capture the complex and non-linear dynamics found in real-world scenarios, which can negatively impact the accuracy of predictions [17]. Therefore, the KF has limitations in non-linear systems. To address this limitation, the UKF was developed, specifically designed to handle nonlinear systems more effectively than the standard KF. The recursive nature of these filters allows them to continuously refine their predictions based on new data, making them particularly effective for tracking and predicting glucose levels, which often exhibit complex, non-linear fluctuations [18]. Despite the advances in these AI and filtering-based glucose prediction approaches, their reliance on a single CGM parameter limits their ability to fully capture the complexity of glucose fluctuations. To overcome these limitations, it is essential to incorporate additional parameters, such as basal and bolus insulin, into the prediction approaches [19], [20].

The integration of basal and bolus insulin data with CGM readings offers a more comprehensive understanding of glucose dynamics by considering the correlation between these parameters and glucose levels, resulting in more accurate and reliable predictions. Although multiparametric data integration enhances the understanding of glucose dynamics, the inherent variability among patients presents a significant challenge in glucose prediction approaches [21]. Personalization of these approaches is essential as each patient exhibits unique glycemic patterns, insulin sensitivities, and responses to various physiological and behavioral factors. Standard, non-personalized approaches may fail to capture these individual variations, resulting in insufficient predictions [22]. Moreover, the dynamic nature of glucose fluctuations in patients requires adaptive approaches that can continuously respond to evolving metabolic patterns [23]. Therefore, this study presents a comparative analysis of the KF-based approaches and proposes a multi-parametric personalized glucose prediction approach. This comparative analysis includes the comparison of one, two, three, and four-parameter KF and UKF approaches in predicting glucose levels. The findings indicate that the three-parameter UKF exhibits superior performance, providing predictions 5, 15, 30, 45, 60, 90, and 120 minutes ahead, while dynamically adapting with real-time data to offer personalized diabetes management. This approach effectively captures the unique patterns of each patient, enabling more precise and responsive care. Additionally, proposed three-parameter UKF approach is embedded in a custom-designed cross-platform smartphone application for type 1 diabetes patients, ensuring accessibility even without an internet connection.

The rest of this paper is organized as follows: Section 2 presents the theoretical background of the KF, UKF, data preprocessing, our smartphone application: *GlucoThinker Advance*. Section 3 introduces the dataset, and performance analyses of approaches. Additionally, Section 3.3 shows the comparative results of the approaches, our smartphone application interface. Closing remarks are given in Section 4.

2. Method

This section presents the fundamentals of KF, UKF, data preprocessing, and our smartphone application: *GlucoThinker Advance*.

2.1. Kalman Filter

The KF is designed to predict unknown variables by integrating observed measurements. It estimates the state of a dynamic system over time continuously improves its estimates through recursive updates based on new data. The KF starts with an initial prediction, $\hat{\mathbf{x}}_{0|0}$, and $\mathbf{P}_{0|0}$ where **x** is a state vector, with the dimension of $n_x \times 1$, **P** is an estimate covariance matrix with the dimension of $n_x \times n_x$, and n_x is the number of states in a state vector.

First, the KF predicts the future state of the system, using the state extrapolation equation $\hat{\mathbf{x}}_{t+1|t}$, and covariance extrapolation equation $\mathbf{P}_{t+1|t}$ as follows:

$$\hat{\mathbf{x}}_{t+1|t} = \mathbf{F}\hat{\mathbf{x}}_{t|t} + \mathbf{G}\mathbf{U}_t \tag{1}$$

$$\mathbf{P}_{t+1|t} = \mathbf{F}\mathbf{P}_{t|t}\mathbf{F}^T + \mathbf{Q} \tag{2}$$

where *t* represents the time index, **F** is a state transition matrix with the dimension of $n_x \times n_x$, **G** is a control matrix with the dimension of $n_x \times n_u$, $\mathbf{U}_t$ is an input variable with the dimension of $n_u \times 1$, **Q** is a process noise covariance matrix with the dimension of $n_x \times n_u$, *T* is the transpose, and n_u is the number of elements in the input variable.

The update step is calculated by Kalman gain $\mathbf{K}_t$ with the dimension of $n_x \times n_z$, and n_z is the number of measured states. Additionally, state update $\hat{\mathbf{x}}_{t|t-1}$, and covariance update $\mathbf{P}_{t|t}$ equations are in part of the update step. These equations represent the core of the update step in the KF, where the predicted state, and covariance are corrected based on the measurement information to obtain the Kalman gain, updated state estimate, and covariance update matrix:

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}^T + \mathbf{K}_t)^{-1} \tag{3}$$

$$\hat{\boldsymbol{x}}_{t|t} = \hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{K}_t(\hat{\boldsymbol{z}}_t - \boldsymbol{H}\hat{\boldsymbol{x}}_{t|t-1}) \tag{4}$$

$$\boldsymbol{P}_{t|t} = (\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{H})\boldsymbol{P}_{t|t-1}(\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{H})^T + \boldsymbol{K}_t\boldsymbol{R}_t{\boldsymbol{K}_t}^T \tag{5}$$

where **H** is an observation matrix with the dimension of $n_z \times n_x$, $\mathbf{R}_t$ is a measurement covariance matrix with the dimension of $n_z \times n_z$, $\hat{\mathbf{z}}_\mathrm{t}$ is the measurement vector with the dimension of $n_z \times 1$, and **I** is an identity matrix.

Lastly, covariances are updated by the following equations:

$$\hat{\boldsymbol{z}}_t = \boldsymbol{H}\hat{\boldsymbol{x}}_t \tag{6}$$

$$\boldsymbol{R}_t = \mathbb{E}(\hat{\boldsymbol{v}}_t{\hat{\boldsymbol{v}}_t}^T) \tag{7}$$

$$\boldsymbol{Q}_t = \mathbb{E}\left(\hat{\boldsymbol{w}}_t{\hat{\boldsymbol{w}}_t}^T\right) \tag{8}$$

$$\boldsymbol{P}_t = \mathbb{E}(\boldsymbol{E}_t{\boldsymbol{E}_t}^T) \tag{9}$$

where $\mathbb{E}$ is an expected value, $\hat{\mathbf{v}}_\mathrm{t}$ is a measurement noise vector with the dimension of $n_z \times 1$, $\hat{\mathbf{w}}_\mathrm{t}$ is a process noise vector with the dimension of $n_x \times 1$, and $\mathbf{E}_\mathrm{t}$ is the estimation error matrix.

KF predicts future states in a recursive process, and adjusts them using new measurements to enable efficient tracking of system states in real-time applications despite the presence of noise [24]. The Kalman Filter has been further advanced into the Extended Kalman Filter (EKF) and the UKF to overcome the challenges associated with nonlinear systems [25].

2.2. Unscented Kalman Filter

The UKF provides a solution for handling nonlinear systems, effectively addressing the limitations of the standard KF [26]. Firstly, the UKF generates sigma points from the distribution of the system as follows:

$$\mathcal{X}_0 = \bar{\boldsymbol{x}} \tag{10}$$

$$\mathcal{X}_i = \bar{\boldsymbol{x}} + (\sqrt{(L+\lambda)\boldsymbol{P}})_i \tag{11}$$

$$\mathcal{X}_i = \bar{\boldsymbol{x}} - (\sqrt{(L+\lambda)\boldsymbol{P}})_{i-L} \tag{12}$$

where $\mathcal{X}$ represents the sigma points matrix, $\mathcal{X}_0$ is the initial system matrix, $\bar{\mathbf{x}}$ is the mean of state vector **x**, i is the index of each sigma point generated during UKF operation, L is the dimension of the system, and λ is the scaling parameter that used in the generation of sigma points as follows:

$$\lambda = \alpha^2(L + \kappa) - L \tag{13}$$

where α is a positive scaling parameter between 10^{-4} to 1 that determines the spread of the sigma points around the mean, κ is the secondary scaling parameter between 0 to 3, and it is used for further adjusting of the sigma points. The choice of the parameters α, and κ is crucial for the filter performance.

After selecting the sigma points, UKF propagates the sigma points through the non-linear process model and updates the sigma points with the new parameters. This feature is called unscented transform, and it enables the UKF to maintain updated information, and make accurate predictions by following equations:

$$\bar{\boldsymbol{x}} = \sum_{i=0}^{L} \boldsymbol{W}_i^m \mathcal{X}_i' \tag{14}$$

$$\boldsymbol{P} = \sum_{i=0}^{L} \boldsymbol{W}_i^c (\mathcal{Y}_i - \bar{\boldsymbol{x}})(\mathcal{Y}_i - \bar{\boldsymbol{x}})^T + \boldsymbol{Q} \tag{15}$$

where $\mathbf{W}_\mathrm{i}^\mathrm{m}$ is the weight of the state mean matrix, $\mathbf{W}_\mathrm{i}^\mathrm{c}$ is the weight of the state covariance matrix, and $\mathcal{Y}_\mathbf{i}$ is the predicted sigma points.

After that, the UKF updates itself through the sigma points, measurement mean matrix, measurement covariance matrix, state mean matrix, and state covariance matrix. The sigma points are propagated through the measurement model in following the equations:

$$\bar{\mathbf{z}} = \sum_{i=0}^{L} \mathbf{W}_i^m \mathcal{Z}_i \tag{16}$$

$$\boldsymbol{S} = \sum_{i=0}^{L} \boldsymbol{W}_i^c (\mathcal{Z}_i - \bar{\boldsymbol{z}})(\mathcal{Z}_i - \bar{\boldsymbol{z}})^T + \boldsymbol{R} \tag{17}$$

where **S** is the updated measurement covariance matrix, and $\mathcal{Z}_i$ is the updated sigma points from the model.

$$\boldsymbol{C} = \sum_{i=0}^{L} \boldsymbol{W}_i^c (\mathcal{Y}_i - \bar{\boldsymbol{x}})(\mathcal{Z}_i - \bar{\boldsymbol{z}})^T \tag{18}$$

$$\boldsymbol{K} = \boldsymbol{C}\boldsymbol{S}^{-1} \tag{19}$$

where **C** is the cross-covariance matrix.

Finally, the state mean and covariance matrices are updated as defined by the following equations:

$$\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \boldsymbol{K}(\mathcal{Z}_i - \bar{\boldsymbol{z}}) \tag{20}$$

$$\boldsymbol{P} = \boldsymbol{P} - \boldsymbol{K}\boldsymbol{S}\boldsymbol{K}^T \tag{21}$$

2.3. Data Preprocessing

Time series data preprocessing is an essential step in transforming raw data into a refined format suitable for advanced analysis, modeling and pattern recognition. This preprocessing addresses several challenges inherent in raw time series data, such as missing values, and parameter selection, using various techniques to ensure data integrity and analytical reliability. Missing values are a common issue in time series data, and effective handling is crucial for maintaining the quality of subsequent analysis. A technique for dealing with missing values is linear interpolation, which involves predicting the value of a missing data point based on nearby known data points. This technique creates a smooth transition between the known data points, ensuring a consistent dataset that supports accurate predictions, and analysis.

Parameter selection represents a critical component in modeling, enhancing performance, and reducing complexity by focusing on the most relevant features. A correlation heatmap is a common technique used to visualize the relationships between different variables, enabling the identification of significant patterns, and associations. A correlation heatmap is a color-coded matrix that displays the correlation coefficients between pairs of variables, ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation). The strength, and direction of these relationships are represented with different colors, and intensities, allowing identification of patterns and dependencies within complex multivariate time series data.

2.4. Smartphone Application: GlucoThinker Advance

The proposed approach is embedded in a custom-designed mobile application named *GlucoThinker Advance*. The application is developed in pure Python code, which allows users to make glucose predictions without the need for an internet connection. *GlucoThinker Advance* is developed for both iOS and Android devices using the Python-based Flet library to address the needs of the user, providing a user-friendly interface. This cross-platform compatibility extends the functionality of the application and makes it accessible to a wide range of users. This application incorporates the proposed three-parameter CGM-basal-bolus UKF approach to predict, and update itself to the new glucose values at future intervals of 5, 30, 60, and 120 minutes. The application screen opens with the login screen, allowing users to have different accounts to separate their personalized predictions from the same smartphone device. In *GlucoThinker Advance*, users input their CGM, basal, and bolus insulin values, specify the prediction horizon, and receive predictions generated by the trained UKF for the selected time interval. Additionally, *GlucoThinker Advance* stores data locally, allowing users to input and track CGM value, basal insulin, bolus insulin, result, selected option, and the time period. The application records this information and subsequently uses it to generate real-time graphs of CGM, based on the historical entries. Furthermore, in the occurrence of an incorrect entry, the application allows users to delete previous data. The deleted data is removed from the prediction approach, and updates itself without using the deleted value, ensuring that future predictions remain accurate and uncompromised.

3. Experimental Evaluations

3.1. Dataset

The open-access OhioT1DM dataset is used for the development and evaluation of the proposed approach [27]. This dataset consists of glucose monitoring, insulin, physiological sensor, patient-reported life history glucose level, fingertip blood sampling, basal rate, temporal basal rate, bolus, meal, sleep, work, stress, hypoglycemic event, illness, exercise, heart rate, galvanic skin response, skin temperature, air temperature, number of steps, baseline sleep and acceleration parameters for eight weeks in each of twelve type 1 diabetes patients: 540, 544, 552, 559, 563, 567, 570, 575, 584, 588, 591, and 596. The patients used insulin pumps with CGM sensors throughout the study period. Each patient in the dataset is associated with approximately 14,000 data. Patients used a smartphone application to record daily life events, while a wearable fitness band simultaneously collected physiological measurements to provide a comprehensive view of diabetes management, and lifestyle factors.

3.2. Evaluation Metrics

The developed prediction approach was evaluated, using three statistical metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

RMSE, MAE, and MAPE are statistical metrics frequently used to measure the difference between predicted and actual values. RMSE computes the root mean square of the squared differences. MAE measures the mean of the absolute differences. MAPE calculates the average percentage difference between predicted and actual values.

The formulas for RMSE, MAE, and MAPE can be represented as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{b=1}^{N}(y_b - \hat{y}_b)^2} \tag{22}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{b=1}^{N}|y_b - \hat{y}_b| \tag{23}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{b=1}^{N}\left|\frac{y_b - \hat{y}_b}{y_b}\right| \times 100 \tag{24}$$

where N is the number of the data points, y_b is the b-th measurement and $\hat{y}_b$ is the corresponding prediction in all of the three equations.

3.3. Results and Discussion

This paper presents a comprehensive evaluation of KF, and UKF for their ability to predict glucose levels in real-time with one, two, three, and four parameters. Using the open-access OhioT1DM dataset, our analysis focused on prediction accuracy over different time frames: 5, 15, 30, 45, 60, 90, and 120 minutes.

Figure 1 shows the relationships between the parameters, and CGM values. The results indicate a positive correlation between CGM values, both basal, and bolus insulin values. One of the negative correlations is observed between CGM, and carbohydrate (CHO) values. Therefore, CHO is also included as parameter to assess CHO potential impact on the approach.

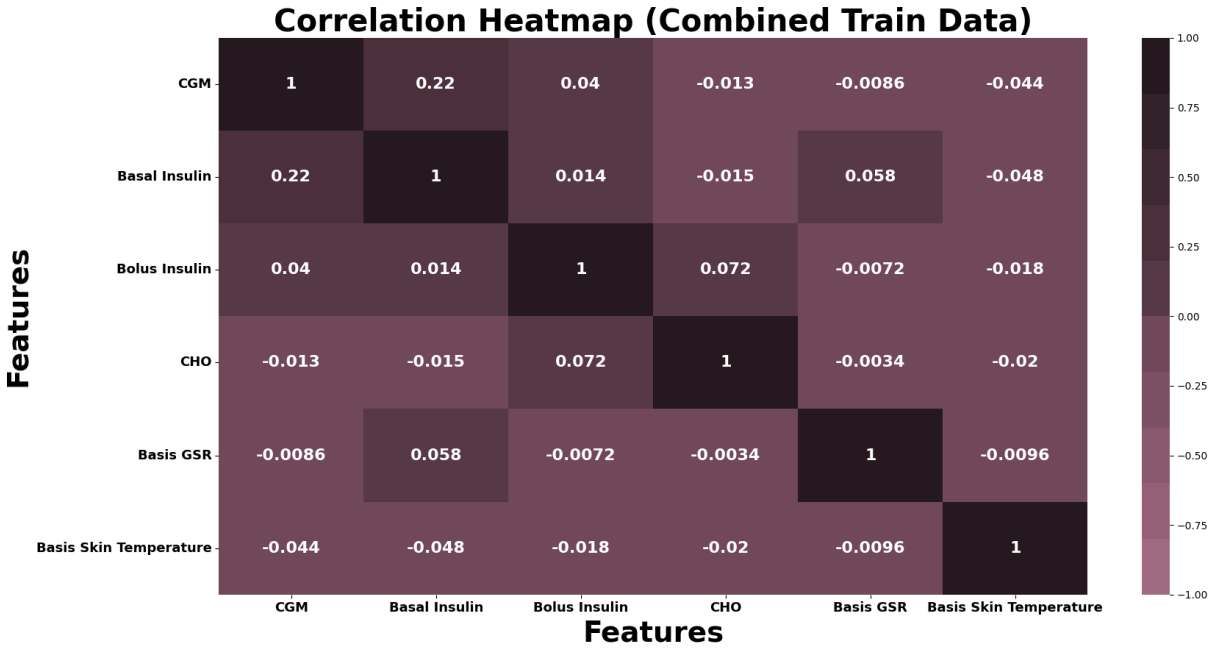


Figure 1. *Correlation Heatmap of Combined Train Data*

Table 1, and Table 2 present a comparative analysis of the performance across one, two, three, and four parameters of KF, and UKF. These parameters include combinations of CGM data, basal insulin, bolus insulin, and CHO data. Table 1 and Table 2 illustrate that the three-parameter UKF, which incorporates CGM, basal, and bolus insulin data, consistently outperforms alternative configurations across all measured intervals in terms of RMSE, MAE, and MAPE. Notably, its capacity to manage nonlinear systems allows for highly accurate and reliable glucose level predictions, achieving an RMSE of 12.76 mg/dL for the 30-minute prediction interval.

Table 1. *Comparison of KF-based approaches across 5, 15, and 30-minute time intervals.*

Method		5-MIN			15-MIN			30-MIN		
Parameter	**Approach**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**
One (CGM)	**KF**	2.92	1.75	1.16	7.96	5.41	3.62	14.14	10.28	6.90
	UKF	2.70	1.61	1.07	7.34	4.98	3.32	13.01	9.45	6.33
Two (CGM-Basal)	**KF**	3.52	2.22	1.47	8.82	6.08	4.01	14.89	10.82	7.20
	UKF	2.66	1.59	1.06	7.20	4.90	3.27	12.79	9.27	6.22
Two (CGM-Bolus)	**KF**	3.53	2.23	1.48	8.84	6.08	4.02	14.88	10.83	7.21
	UKF	2.67	1.60	1.07	7.21	4.91	3.28	12.80	9.28	6.23
Two (CGM-CHO)	**KF**	3.56	2.25	1.50	8.81	6.10	4.03	14.87	10.84	7.22
	UKF	2.68	1.61	1.08	7.22	4.92	3.29	12.81	9.29	6.24
Three (CGM-Basal-Bolus)	**KF**	3.23	2.07	1.40	7.74	5.36	3.55	12.89	9.34	6.28
	UKF	**2.65**	1.58	1.05	**7.19**	4.88	3.26	**12.76**	9.26	6.20
Three (CGM-Basal-CHO)	**KF**	3.24	2.09	1.41	7.73	5.37	3.56	12.90	9.35	6.29
	UKF	2.69	1.62	1.09	7.23	4.93	3.30	12.82	9.30	6.25
Three (CGM-Bolus-CHO)	**KF**	3.22	2.08	1.38	7.72	5.35	3.54	12.88	9.33	6.21
	UKF	2.70	1.62	1.10	7.24	4.94	3.31	12.78	9.31	6.26
Four (CGM-Basal-Bolus-CHO)	**KF**	3.26	2.08	1.39	7.83	5.43	3.60	13.09	9.50	6.34
	UKF	2.71	1.63	1.11	7.25	4.89	3.32	12.77	9.32	6.27

Table 2. *Comparison of KF-based approaches across 45, 60, 90, and 120-minute time intervals.*

Method		45-MIN			60-MIN			90-MIN			120-MIN		
Parameter	**Approach**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**
One (CGM)	**KF**	18.57	14.01	9.49	21.93	16.82	11.43	26.60	20.51	14.02	29.49	22.87	15.58
	UKF	17.11	12.87	8.70	20.26	15.50	10.52	24.68	19.00	12.97	27.46	21.22	14.44
Two (CGM-Basal)	**KF**	19.14	14.34	9.70	22.41	17.09	11.67	26.97	20.69	14.24	29.82	23.14	15.89
	UKF	16.79	12.62	8.53	19.89	15.23	10.33	24.25	18.66	12.75	27.01	20.85	14.19
Two (CGM-Bolus)	**KF**	19.15	14.35	9.71	22.42	17.10	11.68	26.98	20.71	14.25	29.83	23.15	15.91
	UKF	16.82	12.65	8.54	19.92	15.24	10.34	24.28	18.69	12.76	27.04	20.86	14.20
Two (CGM-CHO)	**KF**	19.12	14.32	9.69	22.38	17.06	11.69	26.99	20.70	14.26	29.84	23.13	15.90
	UKF	16.83	12.66	8.55	19.93	15.21	10.35	24.29	18.70	12.77	27.05	20.87	14.21
Three (CGM-Basal-Bolus)	**KF**	17.70	12.39	8.36	20.73	14.91	10.17	24.25	18.42	12.66	27.03	20.88	14.25
	UKF	**16.78**	12.61	8.52	**19.88**	15.20	10.31	**24.24**	18.65	12.73	**27.00**	20.84	14.18
Three (CGM-Basal-CHO)	**KF**	17.71	12.40	8.37	20.72	14.92	10.18	24.26	18.44	12.67	27.06	20.89	14.26
	UKF	16.80	12.63	8.56	19.90	15.22	10.36	24.26	18.67	12.78	27.02	20.90	14.22
Three (CGM-Bolus-CHO)	**KF**	16.69	12.38	8.35	20.71	14.90	10.16	24.12	18.43	12.68	27.07	20.81	14.27
	UKF	16.81	12.64	8.57	19.91	15.25	10.37	24.27	18.68	12.79	27.08	20.92	14.23
Four (CGM-Basal-Bolus-CHO)	**KF**	16.91	12.57	8.48	19.95	15.10	10.32	24.33	18.59	12.79	27.18	20.95	14.36
	UKF	16.84	12.67	8.58	19.94	15.26	10.38	24.30	18.71	12.74	27.09	20.93	14.24

Table 3. *Three-parameter (CGM-basal-bolus) UKF results for twelve patients in terms of RMSE.*

Patient ID	5-MIN	15-MIN	30-MIN	45-MIN	60-MIN	90-MIN	120-MIN
540	2.57	8.42	14.78	19.73	23.91	28.62	31.33
544	1.87	6.00	11.74	15.05	17.37	21.25	22.50
552	**1.79**	**5.53**	**9.76**	**12.11**	**14.76**	**16.81**	**18.25**
559	2.77	7.09	11.66	16.07	19.96	25.62	28.20
563	2.02	5.82	10.64	13.66	16.29	19.02	23.04
567	3.19	8.89	15.98	21.30	25.01	30.01	35.06
570	2.30	5.91	10.89	14.77	18.28	22.46	27.41
575	2.10	6.40	11.75	15.60	19.36	25.13	29.64
584	4.15	9.43	16.32	21.29	24.53	28.95	31.28
588	2.39	7.12	13.08	17.61	19.52	23.29	22.98
591	3.89	9.14	14.57	18.75	21.38	28.23	29.59
596	2.71	6.58	11.94	15.40	18.15	21.51	24.70
Average	2.65	7.19	12.76	16.78	19.88	24.24	27.00

Although the RMSE values of all parameter combinations were similar to each other, the CGM-basal-bolus combination of the UKF gives the lowest error. This demonstrates the importance of accurate health measurements, as even small differences can significantly impact outcomes. The reduced margins of error in short-term predictions are especially valuable for preventing acute glycaemic events. While achieving high accuracy in long-term predictions remains challenging, the UKF shows promise for effective type 1 diabetes management. Table 3 shows the three-parameter CGM, basal, and bolus UKF results for all patients. The results show that patient 552 has the best result in all time intervals. These variations in individual responses demonstrate the advantage of proposed personalized approach that is continuously updated with individual patient data.

Figure 2 illustrates the comparison between the actual, and predicted CGM values of the three-parameter CGM-basal-bolus combination of UKF. The close alignment of these values demonstrates that the approach is performing effectively, as the predicted values are consistently in close proximity to the actual data points.

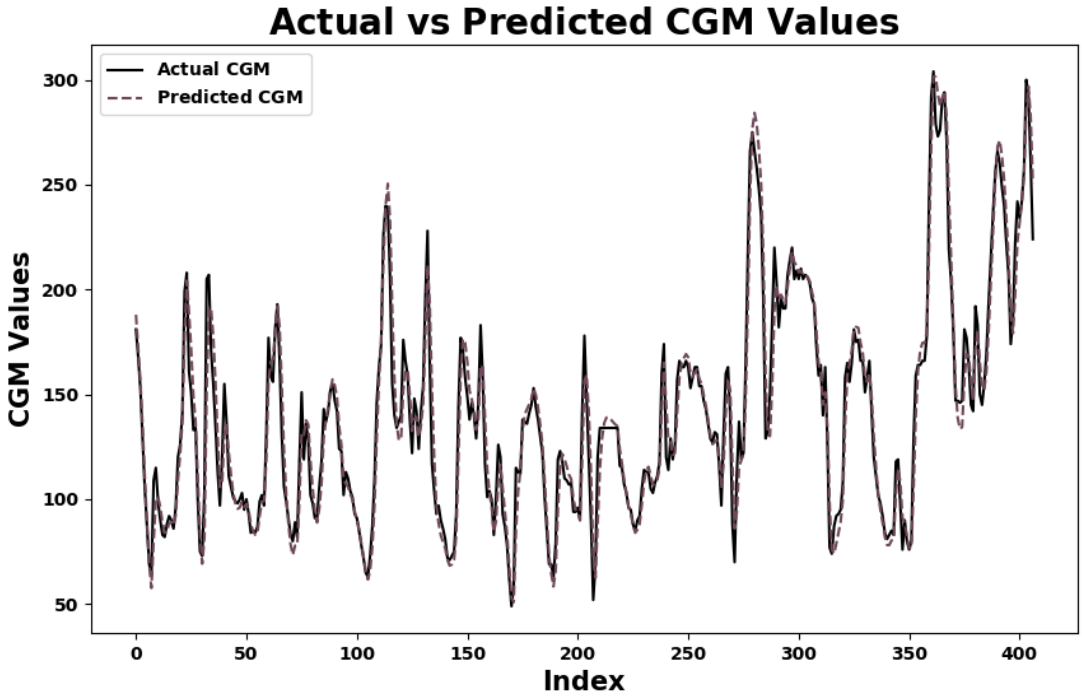
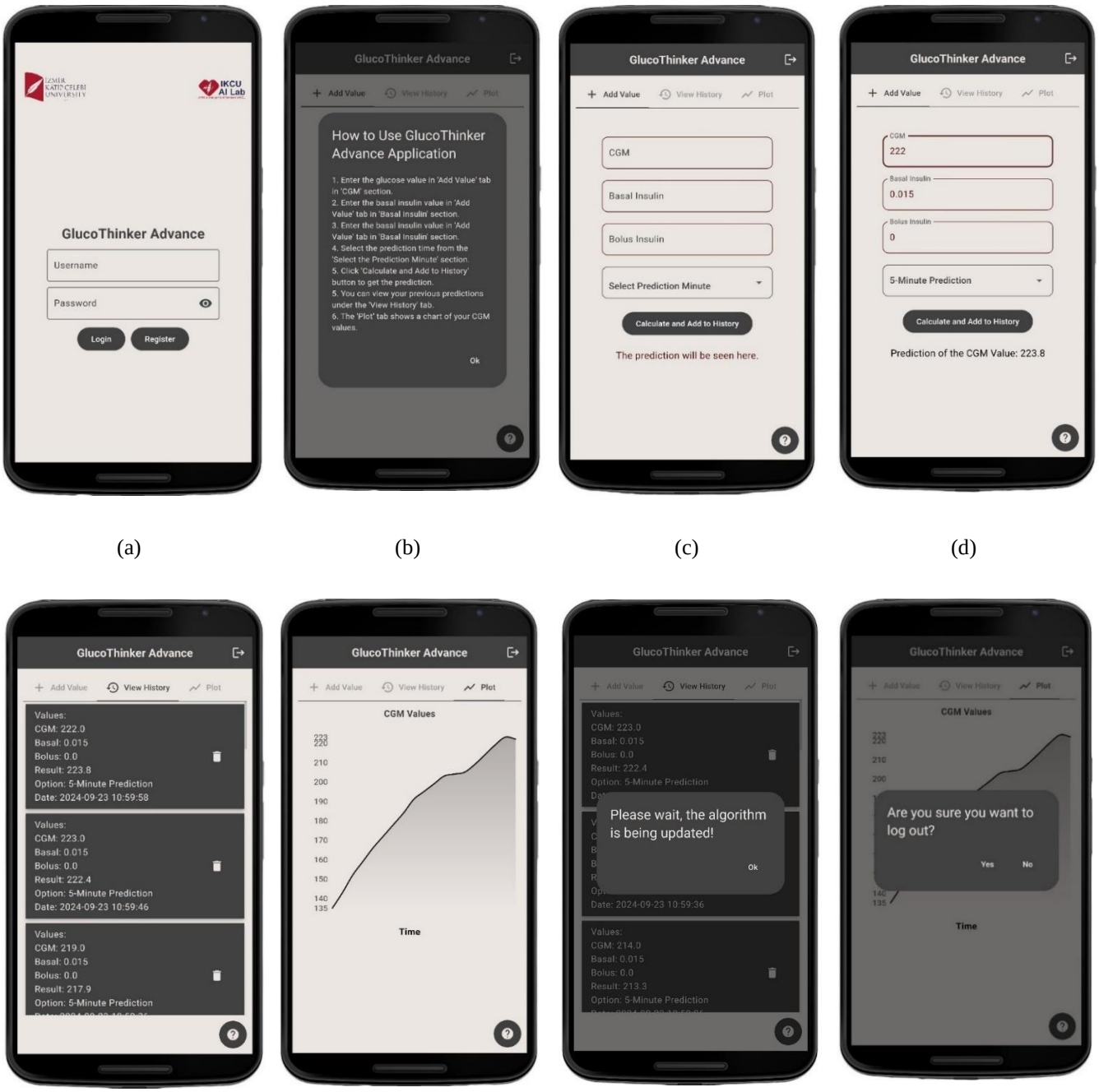


Figure 2. *Actual versus Predicted CGM Values of Patient 552 at 30-Minute*

Furthermore, a personalized smartphone application called *GlucoThinker Advance* has been developed to simplify diabetes management, and help monitor predicted glucose levels for diabetes patients. Figure 3 (a) shows the login interface of the application. This feature allows multiple users to check their CGM values separately on a single smartphone device. Figure 3 (b) presents the help button page, which provides step-by-step instructions on how to use the application. Figure 3 (c) illustrates the user interface, enabling users to add CGM, basal insulin, and bolus insulin values for prediction of the future result. Figure 3 (d) shows the predicted result on the interface. The view history tab, shown in Figure 3 (e), allows users to review their past values in detail. Figure 3 (f) displays a plot of the user entered CGM values, providing a visual representation of glucose levels over time. In addition to the application features, deleting past values also removes those values from proposed approach. Therefore, deleting a value triggers the approach to update itself. During the update process, the user unable to input new values or logout until the update is completed. This feature prevents the errors in the approach and it can be seen in Figure 3 (g). The logout screen interface can be seen in Figure 3 (h).



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 3. *GlucoThinker Advance login interface shown in (a), help page shown in (b), adding value screen shown in (c), displays of the past inputs shown in add value tab (d), view history screen shown in (e), plot screen of CGM values shown in (f), alert dialogue page when UKF is updating screen shown in (g), and logout page shown in (h).*

4. Conclusion

This paper presents a personalized glucose prediction approach, using a three-parameter UKF to predict the next glucose level based on the CGM value of the patient, along with basal and bolus insulin values. Validation of this approach was performed using the OhioT1DM dataset and results showed improved accuracy and reliability in glucose prediction levels at 5, 15, 30, 45, 60, 90, and 120 minutes. The proposed approach showed an RMSE of 2.65 mg/dL for 5-minute predictions, increasing progressively to 7.19 mg/dL, 12.76 mg/dL, 16.78 mg/dL, 19.88 mg/dL, 24.24 mg/dL, and 27.00 mg/dL for 15, 30, 45, 60, 90, and 120-minute predictions, respectively. These results demonstrate the superior performance of the proposed approach compared to KF,

and other parametric UKF. Additionally, the personalized approach facilitates improved diabetes management through increased relevance of predictions for each patient.

The proposed approach offers an advantage over traditional prediction approaches by continuously updating itself with new data, leading to improved glycemic control, and reduced risk of diabetes-related complications. Furthermore, this proposed approach is embedded in our smartphone application *GlucoThinker Advance*, making it useful for type 1 diabetes patients. Our application provides personalized glucose prediction, and enables diabetes to take necessary precautions by predicting glucose levels. The user-friendly interface of *GlucoThinker Advance* ensures that patients can easily navigate the application, and access its features, making glucose management a more accessible part of their daily routine.

In future work, the proposed approach will be focused on predicting potential health risks, taking preventative action by providing timely alerts, and personalized recommendations to the user based on the detection of abnormal changes in glucose levels.

Acknowledgements

This study was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) under the Grant Numbers 222S488 and 1919B012313922. The authors thank to TUBITAK for their supports.

References

[1] S. Gül, E. Ü. Avdal, S. Önal, B. N. Dündar, B. Ö. Pamuk, and Z. Doğan, "Diyabette tıbbi bakım standartlarında değişiklikler," İzmir Katip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi, vol. 5, no. 1, pp. 25-29, 2020.
[2] D. J. Magliano and E. J. Boyko, "IDF Diabetes Atlas," 2022.
[3] W. H. Organization, Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth. World Health Organization, 2017.
[4] Ö. A. Koca, A. Türköz, and V. Kılıç, "Tip 1 Diyabette Çok Katmanlı GRU Tabanlı Glikoz Tahmini," Avrupa Bilim ve Teknoloji Dergisi, no. 52, pp. 80-86, 2023.
[5] L. Heinemann, "Continuous glucose monitoring (CGM) or blood glucose monitoring (BGM): interactions and implications," Journal of Diabetes Science and Technology, vol. 12, no. 4, pp. 873-879, 2018.
[6] Ö. A. Koca, H. Ö. Kabak, and V. Kılıç, "Attention-based multilayer GRU decoder for on-site glucose prediction on smartphone," The Journal of Supercomputing, vol. 80, no. 17, pp. 25616-25639, 2024.
[7] M. M. Kebede and C. R. Pischke, "Popular diabetes apps and the impact of diabetes app use on self-care behaviour: a survey among the digital community of persons with diabetes on social media," Frontiers in Endocrinology, vol. 10, p. 135, 2019.
[8] J. Daniels, P. Herrero, and P. Georgiou, "A multitask learning approach to personalized blood glucose prediction," IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 1, pp. 436-445, 2021.
[9] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "Personalized blood glucose prediction for type 1 diabetes using evidential deep learning and meta-learning," IEEE Transactions on Biomedical Engineering, vol. 70, no. 1, pp. 193-204, 2022.
[10] H.-S. Kim, W. Choi, E. K. Baek, Y. A. Kim, S. J. Yang, I. Y. Choi, K.-H. Yoon, and J.-H. Cho, "Efficacy of the smartphone-based glucose management application stratified by user satisfaction," Diabetes & Metabolism Journal, vol. 38, no. 3, pp. 204-210, 2014.
[11] C. Pérez-Gandía, A. Facchinetti, G. Sparacino, C. Cobelli, E. Gómez, M. Rigla, A. de Leiva, and M. Hernando, "Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring," Diabetes Technology & Therapeutics, vol. 12, no. 1, pp. 81-88, 2010.
[12] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, and C. Cobelli, "A new neural network approach for short-term glucose prediction using continuous glucose monitoring time-series and meal information," in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011: IEEE, pp. 5653-5656.
[13] M. Ali, R. Johansen Alexander, H. Nicklas, E. Christensen Peter, M. Tarp Jens, L. Jensen Morten, B. Henrik, and M. Morten, "Short term blood glucose prediction based on continuous glucose monitoring data," arXiv preprint arXiv:2002.02805, 2020.
[14] G. Alfian, M. Syafrudin, M. Anshari, F. Benes, F. T. D. Atmaji, I. Fahrurrozi, A. F. Hidayatullah, and J. Rhee, "Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features," Biocybernetics and Biomedical Engineering, vol. 40, no. 4, pp. 1586-1599, 2020.
[15] M. F. Rabby, Y. Tu, M. I. Hossen, I. Lee, A. S. Maida, and X. Hei, "Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction," BMC Medical Informatics and Decision Making, vol. 21, pp. 1-15, 2021.
[16] R. McShinsky and B. Marshall, "Comparison of Forecasting Algorithms for Type 1 Diabetic Glucose Prediction on 30 and 60-Minute Prediction Horizons," in KDH@ ECAI, 2020, pp. 12-18.
[17] D. Simon, Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches. John Wiley & Sons, 2006.
[18] S. Haykin, "Kalman filters," Kalman Filtering and Neural Networks, pp. 1-21, 2001.
[19] I. F. Godsland and C. Walton, "Maximizing the success rate of minimal model insulin sensitivity measurement in humans: the importance of basal glucose levels," Clinical Science, vol. 101, no. 1, pp. 1-9, 2001.
[20] H. H. Ko and R. Enns, "Review of food bolus management," Canadian Journal of Gastroenterology and Hepatology, vol. 22, no. 10, pp. 805-808, 2008.
[21] R. Snyderman, "Personalized health care: from theory to practice," Biotechnology Journal, vol. 7, no. 8, pp. 973-979, 2012.
[22] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for T1DM patients," International Journal for Numerical Methods in Biomedical Engineering, vol. 33, no. 6, p. e2833, 2017.

[23] Ö. A. Koca and V. Kılıç, "Multi-Parametric Glucose Prediction Using Multi-Layer LSTM," Avrupa Bilim ve Teknoloji Dergisi, no. 52, pp. 169-175, 2023.
[24] M. St-Pierre and D. Gingras, "Comparison between the unscented Kalman filter and the extended Kalman filter for the position estimation module of an integrated navigation information system," in IEEE Intelligent Vehicles Symposium, 2004: IEEE, pp. 831-835.
[25] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in Signal Processing, Sensor Fusion, and Target Recognition VI, 1997, vol. 3068: Spie, pp. 182-193.
[26] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.
[27] C. Marling and R. Bunescu, "The OhioT1DM dataset for blood glucose level prediction: Update 2020," in CEUR Workshop Proceedings, vol. 2675: NIH Public Access, p. 71.