

Noninvasive Continuous Glucose Monitoring on Aqueous Solutions Using Microwave Sensor with Machine Learning

Saeed M. Bamatraf¹, Maged A. Aldhaeabi², and Omar M. Ramahi^{1, *}

Abstract—In this paper, an electrically-small microwave dipole sensor is used with machine learning algorithms to build a noninvasive continuous glucose monitoring (CGM) system. As a proof of concept, the sensor is used on aqueous (water-glucose) solutions with different glucose concentrations to check the sensitivity of the sensor. Knowledge-driven and data-driven approaches are used to extract features from the sensor's signals reflected from the aqueous glucose solution. Machine learning is used to build the regression model in order to predict the actual glucose levels. Using more than 19 regression models, the results show a good accuracy with Root Mean Square Error of 1.6 and 1.7 by Matern 5/2 Gaussian Process Regression (GPR) algorithm using the reflection coefficient's magnitude and phase.

1. INTRODUCTION

Diabetes is a chronic disease that occurs because of the high levels of glucose in the blood [1]. According to the International Diabetes Federation (IDF), 463 million people are currently living with diabetes, and this number is projected to reach 573 million by 2030 [2]. People with diabetes are encouraged to regularly check their blood glucose levels to monitor any changes (increasing or decreasing) in those levels and to adjust their medications accordingly. This process is called Continuous Glucose Monitoring (CGM), and it is categorized based on the invasiveness as invasive, minimally invasive, and noninvasive.

Invasive techniques are the most widely used because they have the highest accuracy. In invasive CGM, a blood sample is extracted using a lancet; typically, the fingertip is the most widely used human body part from which the blood is extracted. The blood sample is read by a glucometer to measure the glucose level. Due to the pain and inconvenience caused by the frequent finger prick, minimally invasive techniques are introduced to help in reducing the pain. In minimally invasive techniques, measuring the glucose level is done by subcutaneous needle-type electrodes implanted in the body. However, these techniques are still not recommended because they involve direct interaction with tissues and need to be replaced from time to time. In addition, minimally invasive devices are not as accurate as the invasive ones. The last and most recent category is the noninvasive techniques in which there is no need for any blood sample extraction or any implantation of electrodes in the body.

Many research groups are working on developing noninvasive CGM methods using different body fluids. Those fluids include blood [3], saliva [4], sweat [5,6], urine [7], tears [8,9], breath [10], and interstitial fluid [11]. Clearly, there is a need for a noninvasive CGM device, which is simple, pain-free, and relatively inexpensive. In this paper, an electrically-small microwave dipole sensor is used with machine learning algorithms to build a noninvasive CGM system. Initially, the sensor is used on aqueous solutions with different glucose concentrations as a proof of concept and in order to check the sensitivity of the sensor to those concentrations. Machine Learning (ML) techniques are used to analyze the aqueous glucose data and build regression models to predict the glucose levels. The physical principle

Received 9 November 2021, Accepted 20 January 2022, Scheduled 4 February 2022

* Corresponding author: Omar M. Ramahi (oramahi@uwaterloo.ca).

¹ Department of Electrical and Computer Engineering, University of Waterloo, Canada. ² Department of Electronic and Communication Engineering, Hadhramout University, Yemen.

of using microwaves in monitoring glucose is based on the fact that microwave sensors do not produce ionizing radiation, so the molecular structure of the material under test (MUT) remains unaltered. In addition, the reflected signal will contain a signature that depends on the concentration of the glucose in the MUT.

2. METHODOLOGY

In this section, we will show the methodology of the system. The experimental setup consisting of an electrically-small dipole sensor (the sensor used is adopted from [12]), a keysight 8.5 GHz VNA (E5071C), and glucose-water solutions with nine different concentrations as shown in Fig. 1.

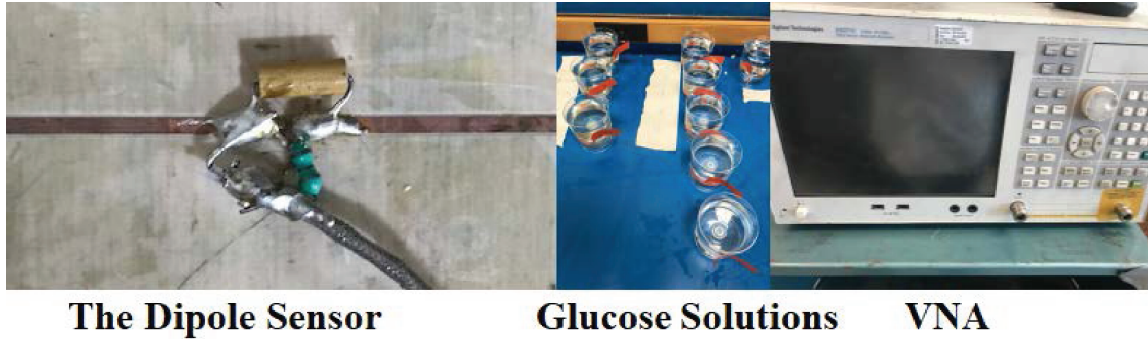


Figure 1. The experimental setup.

The methodology of the system is shown in Fig. 2. First, we prepare the aqueous glucose solutions and place the microwave sensor on top of those solutions to start reading the signals. Then, machine learning is applied by starting with preprocessing to clean and prepare the data for feature extraction. In the feature extraction step, we reduce the dimensionality of the feature space and include the most relevant features only. Next, we train the system to build the regression model which will be used to predict actual glucose concentrations. In the next subsections, we will give a brief explanation of the different components and techniques used by the system.

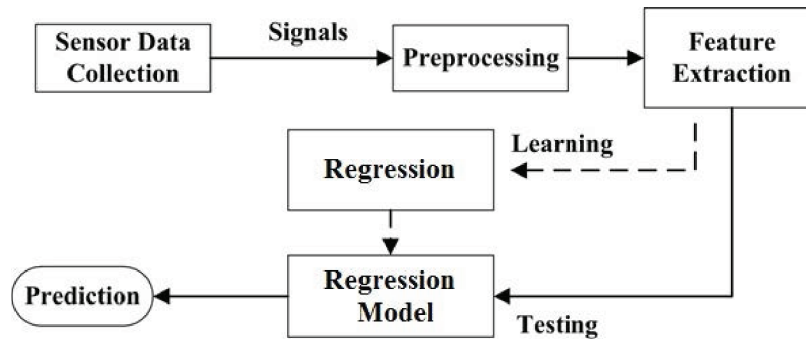


Figure 2. Building blocks of the prediction process.

2.1. The Dipole Sensor

The sensor was designed as a printed dipole of length 92 mm and trace width of 2 mm hosted on a RO4003 Rogers material with a thickness of 1.52 mm and a dielectric substrate of a relative permittivity of $\epsilon_r = 3.38$. The electrical length of the dipole is $\lambda/12$ (where λ is the wavelength in free space) which results in a very low radiation efficiency (corresponding to near unity reflection coefficient (S_{11})).

The dipole sensor is operated over the frequency range of 100–300 MHz, thus providing radiation that penetrates effectively inside the human body. The dipole is tuned to resonate at 200 MHz. We emphasize that the frequency range was only considered based on purely physical considerations, and, therefore, other frequency ranges could have been selected provided that the radiation penetration remains sufficient within the solutions and/or the human body, and also provided that the antenna/probe remains sufficiently small to be used in practical scenarios. (More information about the design of the sensor is available here [12].)

The sensor was placed above the solution with a standoff distance of 5 mm (see Fig. 3). The aqueous solutions are made up of 9 different concentrations: (0 (i.e., water only), 2, 4, 6, 8, 10, 14, 28, 42 mg\dl). The magnitude and phase of the reflection coefficient (S_{11}) of the sensor were then recorded via the VNA at 201 uniformly spaced frequencies spanning the operating frequency range of 100 to 300 MHz.



Figure 3. The Dipole Sensor and its proximity to the solution.

In order to have a clear picture about the nature of the data, in Figs. 4 and 5, we show the responses of the dipole sensor with the water-glucose solutions for the nine different concentrations using the (S_{11}) magnitude and phase. We observe that the range of frequencies that have the most notable discrimination between the responses due to different glucose levels using S_{11} magnitude and phase are 140–240 MHz and 170–200 MHz, respectively, as shown in Figs. 4(b) and 5(b). We observe that the most notable difference between the nine different glucose concentrations occurs around 180 MHz.

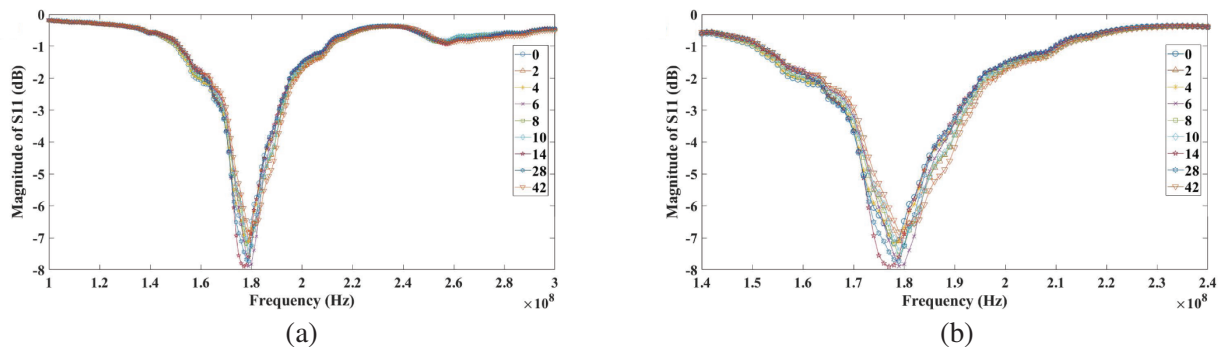


Figure 4. Magnitude of S_{11} for different glucose levels. (a) Entire frequency range. (b) Frequency range of interest.

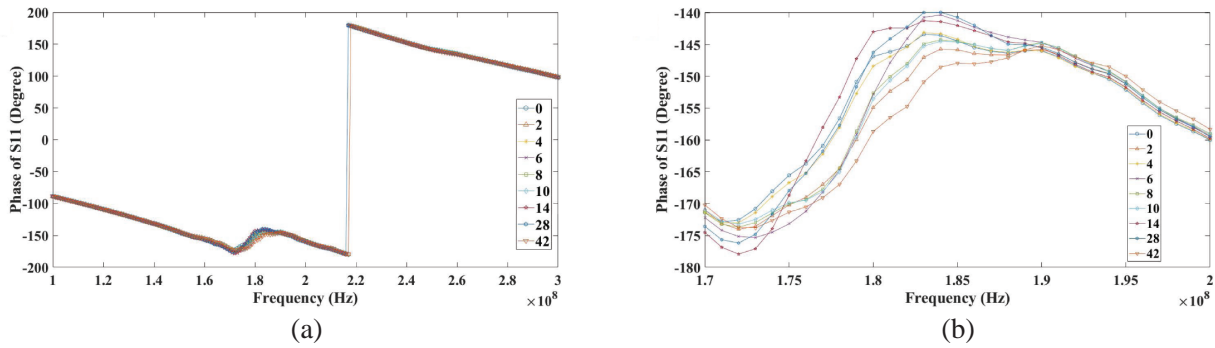


Figure 5. Phase of S_{11} for different glucose levels. (a) Entire frequency range. (b) Frequency range of interest.

Next, we imported and prepared the data using MATLAB [13] to be in a suitable form for feature extraction.

2.2. Feature Extraction

In this subsection, we highlight the step of feature engineering in order to reduce the dimensionality of the feature space by excluding any redundant or irrelevant features and include the most discriminative features only. Our data consisted of three feature vectors: magnitude, phase, and frequency. Each feature vector contains 201 values, which corresponds to the 201 frequencies and the S_{11} magnitude and phase values of each frequency. We adopted two different approaches to extract the feature: a data-driven approach and a domain-knowledge approach. In the data-driven approach, we used Principle Component Analysis (PCA) [14] and selected the highest two Principle Components (PCs) as they preserve 95% of the variance of the whole data. In the domain knowledge-driven approach, we selected the frequency with the minimum S_{11} magnitude and phase (The resonance area). We extracted the minimum two values of S_{11} magnitude and phase. It is worth mentioning here that in both approaches, only two features have been selected from the entire 201 features. This selection of reduced features will help to easily train the regression model in the next step.

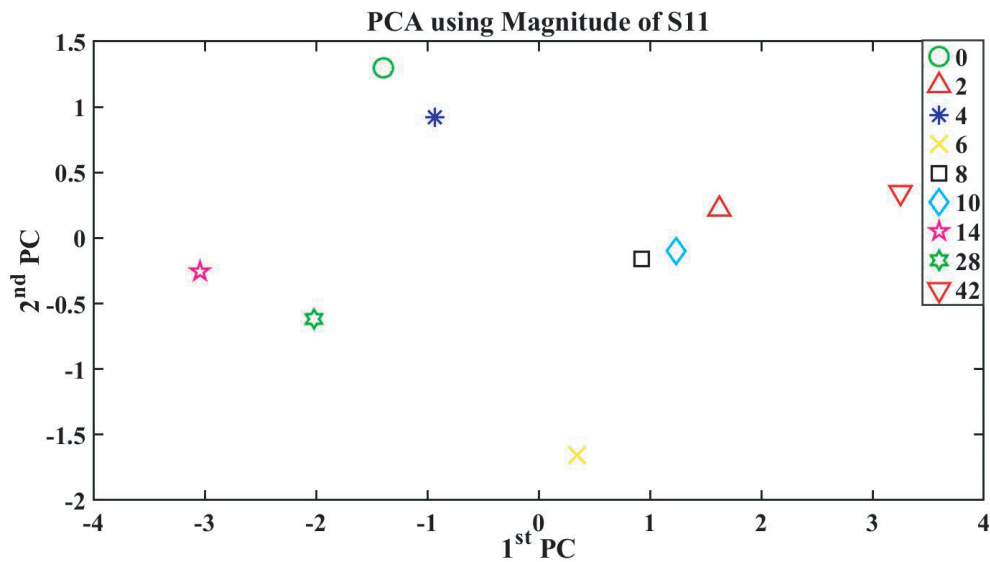


Figure 6. PCA using Magnitude of S_{11} with different glucose levels.

2.3. Regression

Finally, after selecting the most relevant features, they were inputted in the regression model to predict the actual glucose concentrations. A total of 19 different regression models were trained using Matlab regression learning App. RMSE was used as the criteria to select the most accurate regression model (least RMSE).

3. RESULTS AND DISCUSSION

First, we will show the results of the feature extraction step. Here we have four different features: highest 2 PCs using S_{11} magnitude, highest 2 PCs using S_{11} phase, minimum two S_{11} magnitude, and minimum two S_{11} phase. Figs. 6 and 7 show the values of the nine glucose levels representing by the

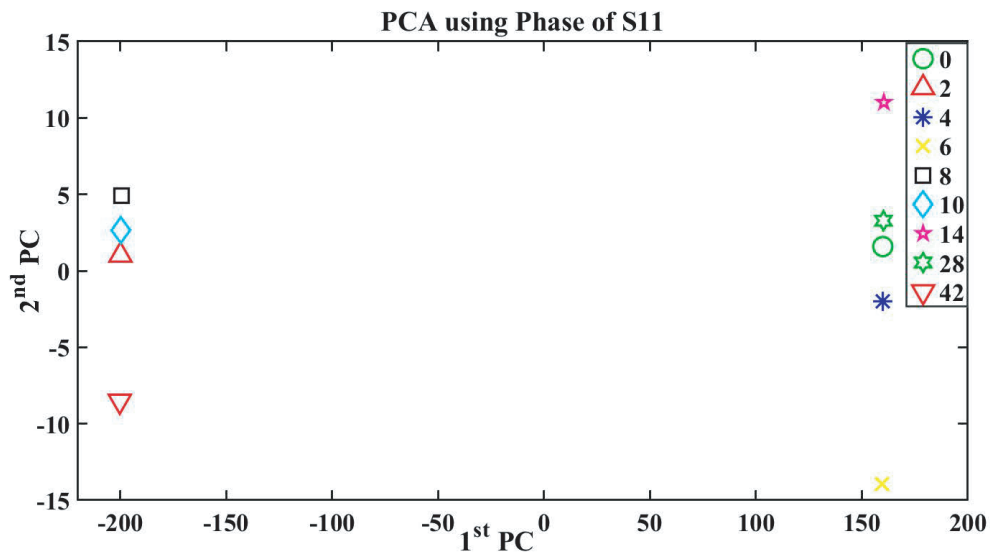


Figure 7. PCA using Phase of S_{11} with different glucose levels.

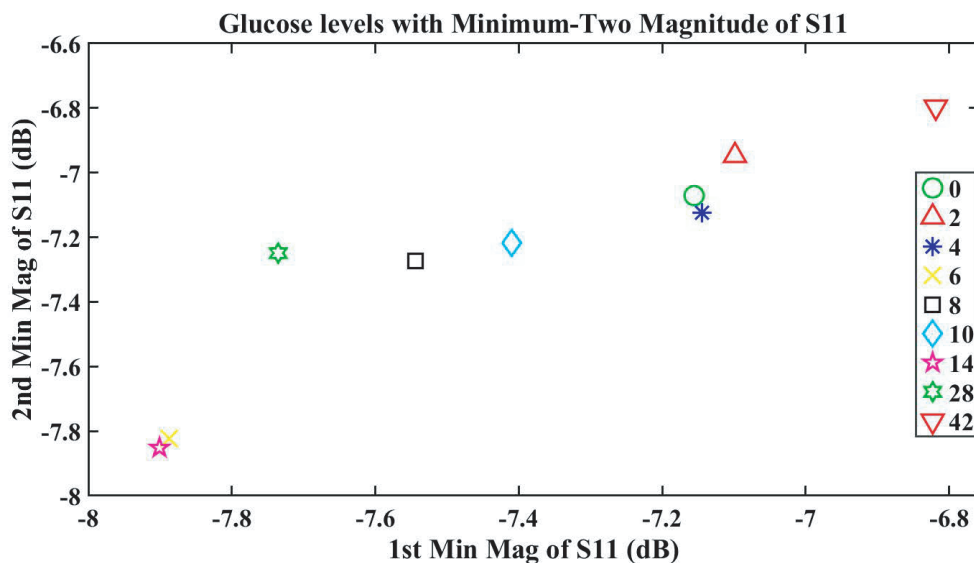


Figure 8. Minimum two Magnitude of S_{11} with different glucose levels.

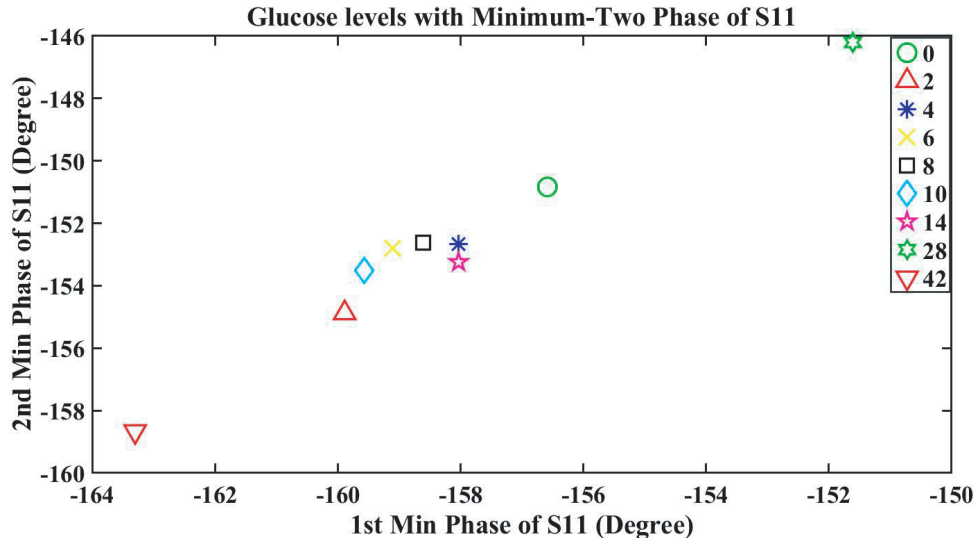


Figure 9. Minimum two Phase of S_{11} with different glucose levels.

highest two PCs using magnitude and phase of S_{11} . Figs. 8 and 9 show the values of the nine glucose levels representing by the minimum two S_{11} magnitude and phase, respectively.

As we can see from all the figures representing the glucose levels with different extracted features, there is a noticeable difference between the glucose levels. However, this difference is not the main goal of the model, and we need to have a good regression function that will easily map each value to its corresponding glucose level with a high accuracy.

Now we will show the results of the best regression models using all the different features approaches (Minimum S_{11} magnitude and phase, and highest 2 PCs using both magnitude and phase). Results from the Matern 5/2 Gaussian Process Regression (GPR) algorithm [15, 16] showed the least RMSE with 1.6 and 1.7 using minimum S_{11} magnitude and phase, respectively. This selection of the minimum S_{11} magnitude and phase is worthy comparing to the available small data size. Figs. 10 and 11 show the response plot of the prediction model by the Matern 5/2 Gaussian Process Regression (GPR) algorithm using the minimum S_{11} magnitude and phase values to predict the actual aqueous glucose levels.

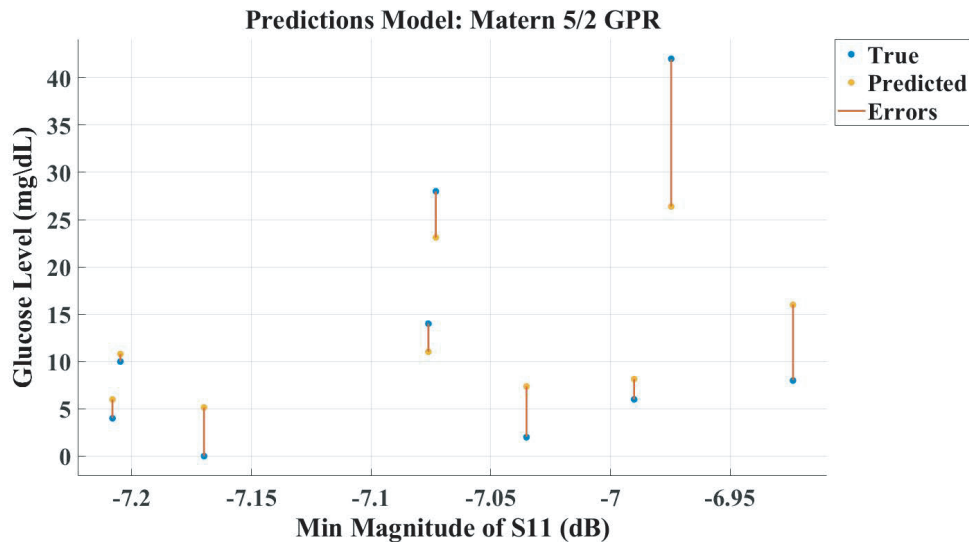


Figure 10. Response plot of the prediction model using minimum Magnitude of S_{11} .

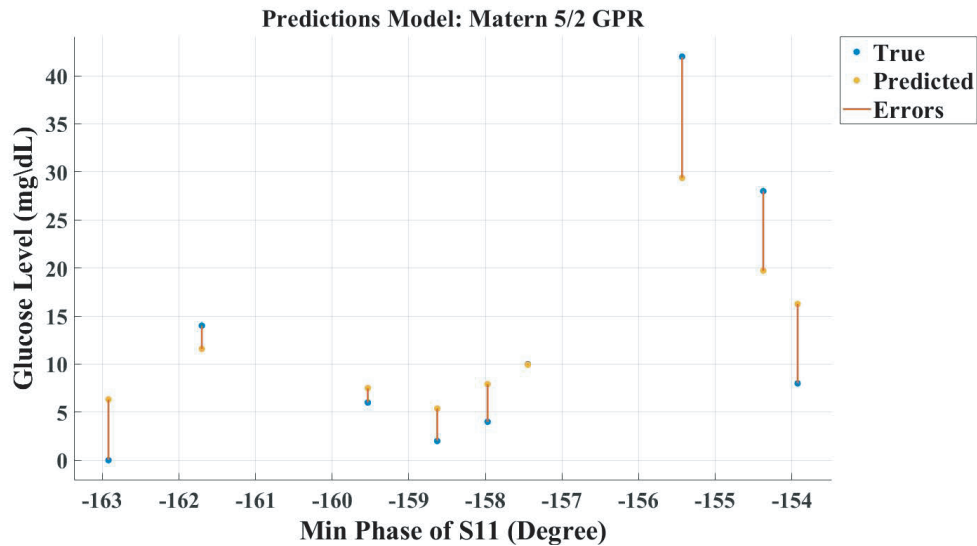


Figure 11. Response plot of the prediction model using minimum Phase of S_{11} .

Results of the regressions using the highest two PCs gave RMSE = 6.7 and 7.7 using S_{11} magnitude and phase, respectively. Those results were obtained using cubic support vector machine (SVM) [17] regression algorithm.

4. CONCLUSION

In this paper, an electrically-small microwave dipole sensor is used with machine learning algorithms to build a noninvasive CGM system. As a proof of concept, the sensor is used on aqueous (water-glucose) solutions with nine different glucose concentrations to check the sensitivity of the sensor to those different glucose concentrations. Feature engineering uses knowledge-driven and data-driven approaches to extract the features from the sensor's reflection coefficient from the aqueous glucose solutions. Machine learning is used to train and build the regression model using those extracted features in order to predict the actual glucose levels. Using more than 19 regression models, the results from minimum S_{11} magnitude and phase showed a good accuracy with Root Mean Square Error (RMSE) around 1.6 by Matern 5/2 Gaussian Process Regression (GPR) algorithm. Regression using the data-driven approach by applying PCA with the highest two PCs gave higher RMSE than knowledge-driven features (minimum S_{11} magnitude and phase). In the future, we will use more data and different types of microwave sensors to improve the accuracy of the prediction. In addition, the sensor will be used with human body tissues in order to build a noninvasive CGM system.

ACKNOWLEDGMENT

The authors acknowledge the Hadhramout Foundation, Hadhramout, Yemen for their support during this research.

REFERENCES

1. DeFronzo, R. A., E. Ferrannini, P. Zimmet, and G. Alberti, *International Textbook of Diabetes Mellitus*, John Wiley & Sons, 2015.
2. Saedi, P., I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, S. Colagiuri, L. Guariguata, A. A. Motala, K. Ogurtsova, et al., "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas," *Diabetes Research and Clinical Practice*, Vol. 157, 107843, 2019.

3. Hanna, J., M. Bteich, Y. Tawk, A. H. Ramadan, B. Dia, F. A. Asadallah, A. Eid, R. Kanj, J. Costantine, and A. A. Eid, "Noninvasive, wearable, and tunable electromagnetic multisensing system for continuous glucose monitoring, mimicking vasculature anatomy," *Science Advances*, Vol. 6, No. 24, eaba5320, 2020.
4. Zhang, W., Y. Du, and M. L. Wang, "Noninvasive glucose monitoring using saliva nano-biosensor," *Sensing and Bio-Sensing Research*, Vol. 4, 23–29, 2015.
5. Olarte, O., J. Chilo, J. Pelegri-Sebastia, K. Barbe, and W. Van Moer, "Glucose detection in human sweat using an electronic nose," *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1462–1465, IEEE, 2013.
6. Heikenfeld, J., "Non-invasive analyte access and sensing through eccrine sweat: Challenges and outlook circa 2016," *Electroanalysis*, Vol. 28, No. 6, 1242–1249, 2016.
7. Mun, P. S., H. N. Ting, Y. B. Chong, and T. A. Ong, "Dielectric properties of glycosuria at 0.2–50 GHz using microwave spectroscopy," *Journal of Electromagnetic Waves and Applications*, Vol. 29, No. 17, 2278–2292, 2015.
8. Yan, Q., B. Peng, G. Su, B. E. Cohan, T. C. Major, and M. E. Meyerhoff, "Measurement of tear glucose levels with amperometric glucose biosensor/capillary tube configuration," *Analytical Chemistry*, Vol. 83, No. 21, 8341–8346, 2011.
9. Yao, H., A. J. Shum, M. Cowan, I. Lähdesmäki, and B. A. Parviz, "A contact lens with embedded sensor for monitoring tear glucose level," *Biosensors and Bioelectronics*, Vol. 26, No. 7, 3290–3296, 2011.
10. Guo, D., D. Zhang, L. Zhang, and G. Lu, "Non-invasive blood glucose monitoring for diabetics by means of breath signal analysis," *Sensors and Actuators B: Chemical*, Vol. 173, 106–113, 2012.
11. Wei, T.-T., H.-Y. Tsai, C.-C. Yang, W.-T. Hsiao, and K.-C. Huang, "Noninvasive glucose evaluation by human skin oxygen saturation level," *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings (I2MTC)*, 1–5, IEEE, 2016.
12. Aldhaeabi, M. A., T. S. Almonneef, A. Ali, Z. Ren, and O. M. Ramahi, "Near field breast tumor detection using ultra-narrow band probe with machine learning techniques," *Scientific Reports*, Vol. 8, No. 1, 1–16, 2018.
13. S. Matlab, "Matlab," The MathWorks, Natick, MA, 2012.
14. Jolliffe, I. T., "Principal components in regression analysis," *Principal Component Analysis*, 167–198, 2002.
15. Rasmussen, C. E., "Gaussian processes in machine learning," *Summer School on Machine Learning*, 63–71, Springer, 2003.
16. Williams, C. K. and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, Vol. 2, No. 3, MIT Press Cambridge, MA, 2006.
17. Cortes, C. and V. Vapnik, "Support-vector networks," *Machine Learning*, Vol. 20, No. 3, 273–297, 1995.