Kernel-Based Analysis of Impedance Spectroscopy in a Quartz Crystal Microbalance Biosensor

1st Ceyhun E. Kirimli
Biomedical Engineering (of Acibadem University)
Istanbul, Turkey
ceyhun.kirimli@acibadem.edu.tr

Elcim Elgun
Basic Sciences (of Acibadem University)
Istanbul, Turkey
eelgun@uwaterloo.ca

Pelin Ece Burgun
Computer Engineering (of Acibadem University)
Istanbul, Turkey
pelin.burgun@live.acibadem.edu.tr

Oğulcan Zorba
Computer Engineering (of Acibadem University)
Istanbul, Turkey
ogulcan.zorba@live.acibadem.edu.tr

***Abstract*—Quartz Crystal Microbalance (QCM) biosensors provide real-time, label-free detection by measuring shifts in resonance frequency caused by mass changes on a sensor surface. The behavior of a Quartz Crystal Microbalance (QCM) biosensor was analyzed under a sinusoidally varying concentration regime using impedance spectroscopy. A comprehensive dataset of impedance parameters—resistance, reactance, absolute impedance and admittance, conductance, susceptance, and phase angle—was collected across time-synchronized experiments. These measurements were interpreted through both analytical and statistical frameworks. The Kanazawa equation was employed to derive an expected concentration from the measured frequency shifts associated with fluidic loading, enabling analytical comparison. To evaluate the predictive capacity of the impedance parameters, non-linear feature selection techniques, including Mutual Information (MI) and the Hilbert-Schmidt Independence Criterion (HSIC), were applied. It was demonstrated that impedance spectroscopy, when coupled with both kernel-based analysis and physical modeling, provides a robust platform for real-time biosensing under dynamic input conditions.**

***Index Terms*—QCM biosensor, impedance spectroscopy, Kanazawa equation, regression modeling, Hilbert-Schmidt Independence Criterion, mutual information, kernel methods, resonance analysis, feature selection**

I. Introduction

Quartz Crystal Microbalance (QCM) sensors are established acoustic biosensors widely used for real-time, label-free monitoring of molecular interactions. In these systems, receptor molecules immobilized on a quartz crystal surface bind target analytes from solution, increasing the effective mass and consequently causing a proportional decrease in resonance frequency, as described by Sauerbrey's equation [1].

Compared to other label-free methods such as Surface Plasmon Resonance (SPR), Quartz Crystal Microbalance with Dissipation (QCM-D) provides enhanced sensitivity to small mass changes [2]. QCM-D systems have been extensively used to monitor protein adsorption and cell–surface interactions in real time, demonstrating capability to detect minute mass changes and record viscoelastic responses [3], [4]

Beyond simple frequency measurements, impedance-based QCM techniques analyze the full electrical impedance spectrum of the crystal using equivalent circuit models such as Butterworth–van Dyke [5], extracting parameters including resonant frequency, bandwidth or dissipation, motional resistance, and capacitance. This comprehensive measurement provides deeper insights into surface interactions and viscoelastic properties [6].

Measuring dissipation, in addition to frequency (as in QCM-D), further allows probing of the viscoelastic characteristics of adsorbed layers, resulting in richer characterization of soft films and biomolecular interactions [7], [8]. Such dual-mode measurements help compensate for viscous and viscoelastic effects inherent in liquid-phase sensing.

Alternative methods, such as the ring-down technique, analyze the free decay of crystal oscillations after excitation ceases, providing high-resolution damping and quality-factor (Q-factor) estimation [9], [10]. Although effective, this methods often require specialized hardware and incapable of measuring parameters other than Δf and ΔD.

In contrast, impedance spectroscopy offers direct acquisition of multiple electrical characteristics—resistance, reactance, admittance, and phase—across wide frequency ranges, generating high-dimensional impedance signatures for each sensing event. These richer datasets facilitate advanced analytical and data-driven interpretation, allowing comprehensive sensor-state characterization [11], [12].

Analyzing high-dimensional impedance data requires effective feature selection to identify informative features for accurate analyte prediction [13]–[16]. Classical statistical techniques, such as Pearson's correlation, detect only linear dependencies, potentially missing nonlinear relationships. Mutual Information (MI) can identify nonlinear dependencies by quantifying the reduction in uncertainty about a target variable given a feature, though accurate estimation of MI from limited datasets remains challenging.

Kernel-based methods, such as the Hilbert–Schmidt Independence Criterion (HSIC), overcome this challenge by measuring statistical dependence between variables within a reproducing kernel Hilbert space without explicitly estimat-

ing probability densities [17], [18]. HSIC robustly captures general, nonlinear relationships. By selecting suitable kernels (e.g., Radial Basis Function kernels), HSIC can effectively identify complex, nonlinear functional dependencies, making it particularly suitable for high-dimensional feature selection tasks in sensor data analysis, as demonstrated in recent applications of HSIC-based methods like HSIC-Lasso [19], [20].

In this work, we leverage MI, Pearson correlation, and HSIC to rank high-dimensional QCM impedance features by their relevance to analyte concentration. Glycerol was used as a model analyte to mimic mass loading effect via change in dynamic viscosity of the bulk medium without increasing the mass of the QCM and without creating a thin film on the sensor surface which disrupts the linear relationship between Δf and ΔD as suggested before [16], [21]. In order to minimize noise, a constant flow rate is established using a pneumatic microfluidic pump with 2 pressure regulating channels in conjunction with a passive microfluidic mixer with circular baffles resulting a constant 50 μl/min flow rate and sinusoidal concentration change between 0 and 2% glycerol solution with a period of 1 hour. We further explore linear and nonlinear dimensionality reduction on the feature set. Principal component analysis (PCA) is used to identify orthogonal combinations of features capturing the highest variance; however, PCA is limited to linear relationships and may fail to capture nonlinear structure in the impedance data. We therefore apply kernel PCA, a nonlinear extension of PCA that uses the kernel trick to reveal underlying nonlinear patterns in the data. By comparing PCA and kernel PCA, we assess whether nonlinear structure (e.g. saturation effects, cooperative binding signatures) is present in the impedance response. Finally, we build regression models to predict the analyte concentration from the impedance features (or from transformed features) and compare their performance. The goal is twofold: (1) identify which impedance parameters are most informative ("feature interpretability"), and (2) evaluate how well we can predict concentration using those features or their combinations ("predictive power").

II. Materials and Methods

A. Experimental Setup and Data Collection

Experiments were conducted utilizing a Quartz Crystal Microbalance (QCM) biosensor integrated with a microfluidic passive mixer designed with circular baffles (Figure 1). The microfluidic mixer was fabricated from cast plexiglass using an in-house built CNC with a channel size of 500 μm. An AT-cut quartz crystal operating at a fundamental resonance frequency of 10 MHz was used. Fluid delivery was controlled by a dual-channel Elveflow OB1 MK3+ pressure controller equipped with flow sensors employing Proportional-Integral (PI) feedback control for stable flow rates. Channel 1 supplied deionized water following a sinusoidal flow pattern between 30 µL/min and 50 µL/min, with a period of one hour. Channel 2 delivered a 5% glycerol solution with a complementary, 180-degree phase-shifted sinusoidal flow between 0 µL/min and 20 µL/min. The resulting total flow rate was maintained constant at 50 µL/min.

The mixing process produced a sinusoidal concentration profile of glycerol ranging from 0% to 2%. Impedance spectroscopy was employed to monitor the dynamic QCM response in real time. A Keysight E4990A Impedance Analyzer was utilized to measure and record a comprehensive set of electrical parameters: resistance (R), reactance at first and second resonance peaks $(X1_{peak}, X2_{peak})$, absolute impedance trough and peak $(Z_{abs-trough}, Z_{abs-peak})$, absolute admittance peak and trough $(Y_{abs-peak}, Y_{abs-trough}$, conductance (G), susceptance peaks and troughs (B_{peak}, B_{trough}), and phase angle. For analytical comparison, the Kanazawa equation was applied to calculate a theoretically expected concentration response (Kanazawa-concentration) based on fluid loading properties [22]. Only the fundamental resonance frequency is used for this study.

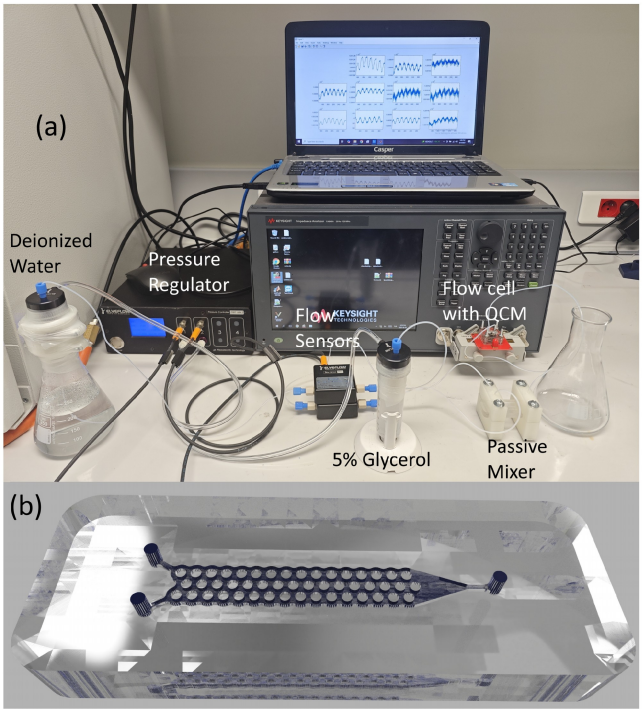

Fig. 1. Experimental Setup.(a) The solutions, pressure regulator, flow sensors, QCM flow cell and passive mixer can be seen (b) The 3d render of the microfluidic mixer is shown

B. Data Preprocessing and Normalization

To ensure comparability across impedance parameters, standardization procedures were applied. Specifically, all impedance parameters, including the Kanazawa-derived concentration parameter, were scaled using z-score normalization, resulting in zero mean and unit variance for each feature. Normalization facilitated direct comparison between features of different magnitudes and units. Kanazawa concentration predictions were calculated from experimental phase angle peak position by converting dynamic viscosities into concentration of glycerol as discussed before [16].

C. Feature Ranking: Mutual Information (MI) and Hilbert-Schmidt Independence Criterion (HSIC)

Feature relevance to glycerol concentration was quantitatively assessed using Mutual Information (MI) and Hilbert–Schmidt Independence Criterion (HSIC), alongside conventional Pearson correlation analysis.

Mutual Information (MI) quantifies the dependency between two variables by measuring how much knowing one variable reduces uncertainty about the other. Mathematically, MI between continuous variables X and Y is defined as

$$MI(X;Y) = \int \int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx\, dy$$

where $p(x,y)$ is the joint probability density function and $p(x)$, $p(y)$ are marginal densities. In this study, MI was directly estimated using a non-parametric k-nearest neighbor (k-NN) approach, which robustly captures nonlinear and non-monotonic dependencies without explicitly calculating probability density functions.

Hilbert–Schmidt Independence Criterion (HSIC) quantifies statistical dependence between variables by evaluating covariance in high-dimensional, nonlinear feature spaces. HSIC relies on embedding the variables into reproducing kernel Hilbert spaces (RKHS), where linear relationships correspond to nonlinear dependencies in the original data space. Formally, HSIC is computed using kernel matrices derived from data points:

$$HSIC(X,Y) = \frac{1}{(n-1)^2} trace(KHLH)$$

where K and L represent kernel matrices for variables X and Y, respectively and $H = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ is a centering matrix ensuring zero mean embedding. In this analysis, Gaussian (Radial Basis Function, RBF) kernels were used, allowing HSIC to capture complex, nonlinear, and arbitrary relationships between impedance parameters and analyte concentration. Due to its kernel-based formulation, HSIC effectively detects dependencies without explicit density estimation, making it particularly suitable for high-dimensional biosensor data exhibiting nonlinear interaction effects.

D. Dimensionality Reduction and Regression Analysis

Principal Component Analysis (PCA) and kernel PCA were employed to reduce dimensionality and to reveal linear and nonlinear patterns in the impedance dataset, respectively. PCA identifies linear combinations of original parameters that capture the highest variance, while kernel PCA extends this by applying nonlinear kernels (e.g., radial basis function kernels) to uncover complex interactions between features.

Regression models—Linear Regression, Polynomial Regression (degree 2), Support Vector Regression (SVR), Random Forest, and XGBoost—were trained using different combinations of impedance features. Model performance was evaluated by computing the coefficient of determination (R^2) through cross-validation, providing quantitative measures of predictive accuracy.

E. Statistical Residue Analysis for Dimensionality Reduction Evaluation

Statistical residue analysis was performed to quantitatively assess how well the original data structure is preserved when reducing dimensionality using both classical PCA and kernel PCA (KPCA) with various kernels. In this analysis, the number of principal components was systematically varied, and the quality of the reconstruction was evaluated for each value of components n. For each method, two metrics were computed: the mean squared error (MSE) between the original data and the reconstructed data, and the Pearson correlation coefficient between the corresponding entries of the pairwise kernel (or covariance) matrices of the original and reduced datasets. By plotting these metrics as a function of the number of components, the extent to which the essential information in the dataset is preserved at each dimensionality was visualized.

III. Results and Discussion

A. Bivariate Analysis of Impedance Parameters

Initially, the relationship between each impedance spectroscopy parameter and glycerol concentration was investigated through bivariate scatter plots (Figure 2). The impedance parameters examined were resistance (R), reactance at first and second peaks $(X1_{peak}, X2_{peak})$, absolute impedance peak and trough $(Z_{abs-peak}, Z_{abs-trough})$, absolute admittance peak and trough $(Y_{abs-peak}, Y_{abs-trough})$, phase angle, conductance (G), and susceptance peaks and troughs (B_{peak}, B_{trough}). Visual inspection of scatter plots suggested strong but varied degrees of correlation, indicating potential for nonlinear dependencies that standard linear analyses might not fully capture.

B. Feature Relevance Ranking: MI, HSIC, and Pearson Correlation

To quantify and compare the strength of relationships between impedance parameters and glycerol concentration, Mutual Information (MI), Hilbert–Schmidt Independence Criterion (HSIC), and Pearson's linear correlation coefficient were computed. The results, shown Figure 3a demonstrated that B_{peak} consistently emerged as the most relevant parameter, exhibiting the highest MI (0.9947), HSIC, and Pearson correlation (0.9310) scores. Conductance (G) and absolute admittance peak $(Y_{abs-peak})$ followed closely, also showing strong associations (MI scores: 0.8784 and 0.8229, Pearson correlations: -0.9234 and 0.9142, respectively). These findings highlighted B_{peak}, G, and $Y_{abs-peak}$ as particularly informative parameters for accurately predicting glycerol concentration. HSIC results complemented MI findings, reinforcing the robustness of these rankings. HSIC, leveraging kernel-based methods, provided additional confirmation of the presence of substantial nonlinear dependencies that linear correlation alone could not fully represent.

C. Dimensionality Reduction: PCA and Kernel PCA

To further explore the impedance parameter space, Principal Component Analysis (PCA) and kernel PCA (kPCA) were

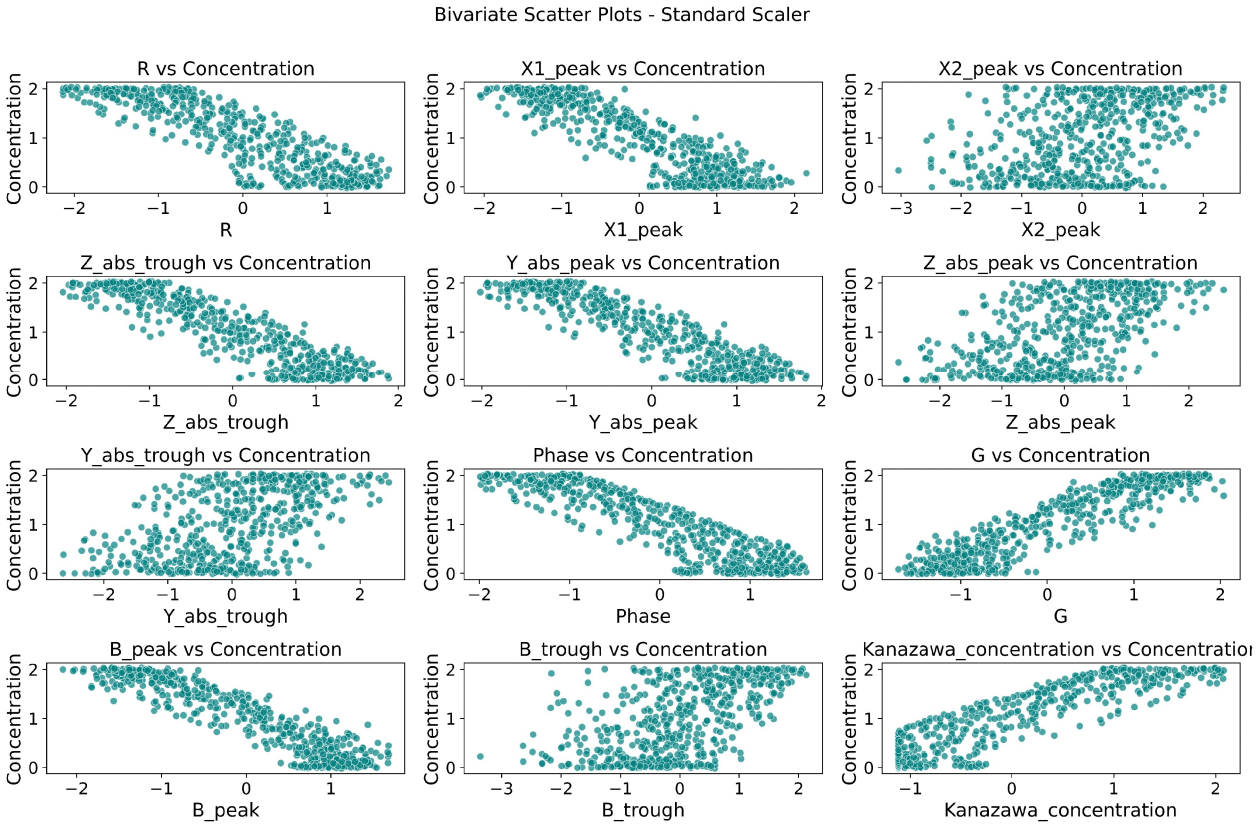

Fig. 2. Bivariate scatter plots of each impedance parameter versus glycerol concentration. Bottom left corner is the scatter plot of voncentration values obtained from Kanazawa equation versus glycerol concentration

applied. PCA revealed that most variance ($> 90\%$) was captured by a few principal components, suggesting inherent redundancy in the original parameter set (Figure 3b). However, kernel PCA using Radial Basis Function (RBF) kernels demonstrated significantly improved feature extraction capability, indicating substantial nonlinear structure in the data. These findings clearly established the presence of nonlinear relationships within the impedance dataset, validating the use of kernel methods to achieve more accurate and interpretable dimensionality reduction and analysis.

D. Regression Modeling and Predictive Performance

Subsequently, regression models—Linear Regression, Polynomial Regression (degree 2), Support Vector Regression (SVR), Random Forest (RF), and XGBoost—were trained using various feature combinations to predict glycerol concentration. Model performance was evaluated via 5-fold cross-validated coefficient of determination (R^2).

When all impedance parameters were used, the Linear Regression model performed well ($R^2 = 0.8550$), while SVR ($R^2 = 0.8389$), RF ($R^2 = 0.8439$), and XGBoost ($R^2 = 0.8171$) demonstrated comparable performance. Polynomial regression exhibited slightly lower performance ($R^2 = 0.7722$), likely due to increased complexity causing diminished generalization.

The analytically derived parameter based on the Kanazawa equation alone provided moderate predictive capability (Linear model $R^2 = 0.7709$), reflecting its analytical validity but also limitations in capturing all dynamic sensor interactions. Combining all impedance parameters with the Kanazawa-derived parameter enhanced performance compared to using impedance parameters alone.

The presented analyses demonstrate the substantial informational richness of impedance spectroscopy data obtained from QCM biosensors under dynamic concentration conditions. Nonlinear feature relevance ranking via MI and HSIC revealed impedance parameters with strong, nonlinear predictive power, surpassing classical linear correlations. Kernel-based dimensionality reduction techniques (kPCA) effectively identified nonlinear interactions between parameters, underscoring the importance of nonlinear modeling approaches in biosensor data interpretation.

Regression model results further support the efficacy of integrating kernel methods into QCM impedance data analysis. SVR and Random Forest, in particular, consistently provided high and stable predictive performance, emphasizing the advantage of nonlinear regression models for complex sensor data.

Finally, while analytical parameters derived from the

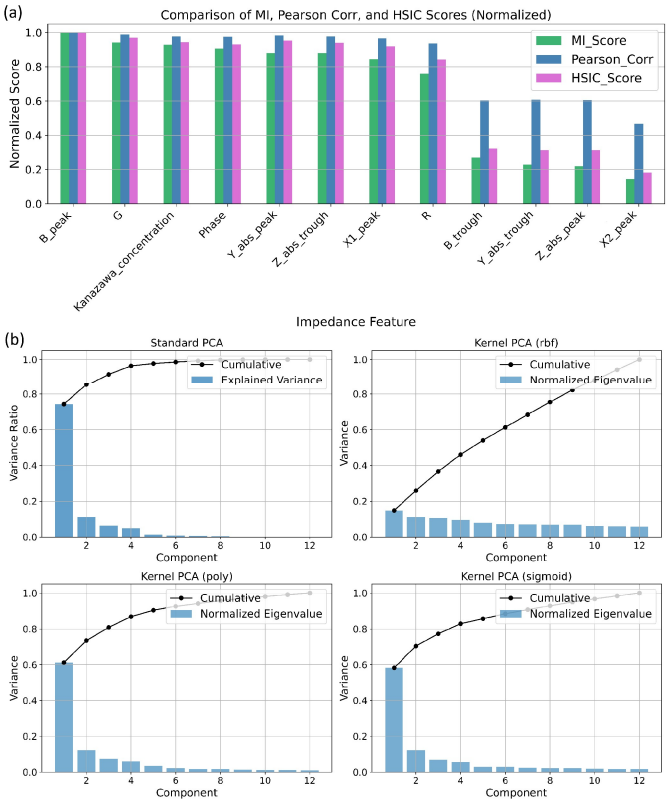

Fig. 3. (a)MI, HSIC and Pearson Correlation Scores represented in different colors (b) PCA and KPCA results on impedance parameters

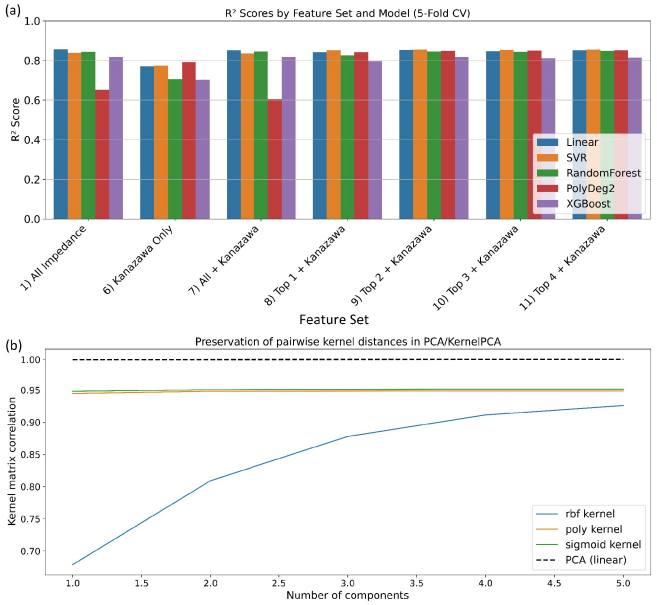

Fig. 4. a) Bar plots illustrating R^2 values of fitted regression models with associated standard deviations as error. b) Residue analysis of components with respect to correlation values

Kanazawa equation proved valuable for theoretical validation, the comprehensive impedance parameter set, analyzed through nonlinear methods, emerged as the optimal approach for accurate and robust concentration prediction.

E. Residue Analysis

The residue analysis method is particularly important for high-dimensional and non-linear datasets, such as those encountered in biosensor impedance measurements, where the preservation of data structure and relationships is critical for reliable interpretation and subsequent analysis. Compared to classical explained variance metrics, the MSE-based residue analysis offers a more direct and interpretable quantification of how much structural information is preserved or lost by different embedding techniques and component choices.

The results (see Figure 4b) demonstrate that "PCA showed an immediate increase in correlation with the first component, while sigmoid and polynomial kernel PCAs stabilized around 0.95 correlation value with the original dataset. The rbf kernel exhibited consistently higher MSE values and hence lower correlation, indicating less effective preservation of data structure for this dataset.

IV. Conclusion

In this study, the applicability of impedance spectroscopy in dynamically modulated Quartz Crystal Microbalance (QCM) biosensing was investigated using a combination of analytical and kernel-based methods. The comprehensive analysis confirmed that impedance parameters, particularly susceptance peak (B_{peak}), conductance (G), and admittance peak $(Y_{abs-peak})$, provided significant nonlinear predictive information regarding analyte concentration. Mutual Information and the Hilbert–Schmidt Independence Criterion effectively identified these nonlinear relationships, exceeding the capabilities of linear correlation analysis alone.

Dimensionality reduction techniques, specifically kernel PCA, revealed essential nonlinear interactions within the impedance data, underscoring the importance of kernel-based methods for accurate feature extraction in QCM data analysis. Regression modeling indicated that nonlinear models such as Support Vector Regression and Random Forest offered robust predictive performance without significant risk of overfitting when trained on the full set of impedance parameters.

Although the analytically derived Kanazawa parameter provided theoretical validation, it showed limited improvement in predictive capability when combined with measured impedance parameters. Overall, the findings demonstrated that integrating impedance spectroscopy data with advanced kernel-based nonlinear methods enhances both interpretability and predictive accuracy, paving the way for more effective real-time biosensing applications.

V. Acknowledgements

This work was supported by the Acibadem University.

References

[1] G. Sauerbrey, "Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung," *Zeitschrift für Physik*, vol. 155, no. 2, pp. 206–222, 1959, doi: 10.1007/BF01337937.

[2] L. Laricchia-Robbio and R. P. Revoltella, "Comparison between the surface plasmon resonance (SPR) and the quartz crystal microbalance (QCM) method in a structural analysis of human endothelin-1," *Biosensors and Bioelectronics*, vol. 19, no. 12, pp. 1753–1758, 2004, doi: 10.1016/j.bios.2003.11.026.

[3] M. C. Dixon, "Quartz crystal microbalance with dissipation monitoring: enabling real-time characterization of biological materials and their interactions," *J. Biomol. Tech.*, vol. 19, no. 3, pp. 151–158, Jul. 2008.

[4] E. Yuca and U. Ö. Ş. Şeker, "Monitoring Molecular Assembly of Biofilms Using Quartz Crystal Microbalance with Dissipation (QCM-D)," *Methods Mol. Biol.*, vol. 2538, pp. 25–33, 2022.

[5] D. Johannsmann, *The Quartz Crystal Microbalance in Soft Matter Research: Fundamentals and Modeling*. Cham: Springer, 2015.

[6] I. Burda, "Advanced Impedance Spectroscopy for QCM Sensor in Liquid Medium," *Sensors (Basel)*, vol. 22, no. 6, Mar. 2022.

[7] D. Johannsmann and I. Reviakine, "Quartz crystal microbalance with dissipation monitoring for studying soft matter at interfaces," *Nature Reviews Methods Primers*, vol. 4, no. 1, p. 63, 2024.

[8] F. Höök and B. Kasemo, "The QCM-D technique for probing biomacromolecular recognition reactions," in *Piezoelectric Sensors*, Springer, 2007, pp. 425–447.

[9] J.-Y. Park and J.-W. Choi, "Electronic circuit systems for piezoelectric resonance sensors," *Journal of The Electrochemical Society*, vol. 167, no. 3, p. 037560, 2020.

[10] D. Johannsmann, A. Langhoff, and C. Leppin, "Studying soft interfaces with shear waves: Principles and applications of the quartz crystal microbalance (QCM)," *Sensors*, vol. 21, no. 10, p. 3490, 2021.

[11] A. Arnau, "A Review of Interface Electronic Systems for AT-cut Quartz Crystal Microbalance Applications in Liquids," *Sensors*, vol. 8, no. 1, pp. 370–411, 2008, doi: 10.3390/s8010370.

[12] A. Alassi, M. Benammar, and D. Brett, "Quartz Crystal Microbalance Electronic Interfacing Systems: A Review," *Sensors*, vol. 17, no. 12, p. 2799, 2017, doi: 10.3390/s17122799.

[13] C. E. Kirimli, E. Elgun, and U. Unal, "Machine learning approach to optimization of parameters for impedance measurements of Quartz Crystal Microbalance to improve limit of detection," *Biosensors and Bioelectronics: X*, vol. 10, p. 100121, 2022, doi: 10.1016/j.biosx.2022.100121.

[14] C. E. Kirimli and E. Elgun, "A Comparison of Impedance and Antenna Analyzers on the Basis of Machine Learning Assisted Limit of Detection Experiments," in *Proc. 2022 Int. Workshop on Impedance Spectroscopy (IWIS)*, 2022, pp. 61-65, doi: 10.1109/IWIS57888.2022.9975132.

[15] C. E. Kirimli and E. Elgun, "Exploratory Data Analysis of Bulk Fluid Viscosity Measurements by Means of Quartz Crystal Microbalance Impedance Analysis," in *Proc. 2024 Int. Workshop on Impedance Spectroscopy (IWIS)*, 2024, pp. 49-53, doi: 10.1109/IWIS63047.2024.10847069.

[16] C. E. Kirimli, E. Elgun, M. M. Yuksel, and S. Y. Tugtag, "An Investigation Into the Impact of Impedance Measurement Parameters on the Limit of Detection of QCM-D Using Machine Learning Model Chaining," *IEEE Sensors Journal*, vol. 25, no. 3, pp. 5688-5696, 2025, doi: 10.1109/JSEN.2024.3511274.

[17] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf, "Measuring Statistical Dependence with Hilbert–Schmidt Norms," in *Proceedings of the 16th International Conference on Algorithmic Learning Theory*, pp. 63–77, 2005.

[18] L. Song, A. Smola, A. Gretton, K. Borgwardt, and J. Bedo, "Supervised Feature Selection via Dependence Estimation," in *Proceedings of the 24th International Conference on Machine Learning*, pp. 823–830, 2007.

[19] M. Yamada, W. Jitkrittum, L. Sigal, E. P. Xing, and M. Sugiyama, "High-Dimensional Feature Selection by Feature-Wise Kernelized Lasso," *Neural Computation*, vol. 26, no. 1, pp. 185–207, 2014.

[20] K. Koyama, K. Kiritoshi, T. Okawachi, and T. Izumitani, "Effective Nonlinear Feature Selection Method based on HSIC Lasso and with Variational Inference," in *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, vol. 151, pp. 10407–10421, 2022.

[21] S. Srimasorn, L. Souter, D. E. Green, L. Djerbal, A. Goodenough, J. A. Duncan, A. R. E. Roberts, X. Zhang, D. Débarre, P. L. DeAngelis, J. C. F. Kwok, and R. P. Richter, "A quartz crystal microbalance method to quantify the size of hyaluronan and other glycosaminoglycans on surfaces," *Scientific Reports*, vol. 12, no. 1, p. 10980, Jun. 2022, doi: 10.1038/s41598-022-14948-7.

[22] K. Kanazawa and J. G. Gordon, "The oscillation frequency of a quartz resonator in contact with liquid," *Analytica Chimica Acta*, vol. 175, pp. 99–105, 1985, doi: 10.1016/S0003-2670(00)82721-X.