Received: 14 August 2026 | Revised: 27 October 2026 | Accepted: 28 October 2026 | Published online: 12 November 2026

RESEARCH ARTICLE

Smart Wearable Technology
2026, Vol. 00(00) 1–11
DOI: 10.47852/bonviewSWT620211731

BON VIEW PUBLISHING

Benchmarking Wearable Bioelectronics: A Multimodal Evaluation Framework Integrating Parametric Regression, Machine Learning, and Fuzzy Logic Across Domain Applications

Hamza Saad[1,*]

[1] *School of Industrial Sciences and Technology, University of Central Missouri, USA*

Abstract: This paper examines how a range of factors, including hardware architecture, sensor accuracy, battery lifetime, and user satisfaction, influence the performance and applicability of wearable technology. Through a study of 29 hardware and more than 2,300 hardware-enabled outcomes, we present a comprehensive analytical model, which includes ordinary least squares regression, nonlinear Random Forest feature extraction, and Takagi–Sugeno fuzzy logic inference. The models indicate that the four main performance indicators are heart rate accuracy, sleep tracking accuracy, battery life, and consumer satisfaction. Sensitivity analysis reveals that the hardware form factor is a unique technological architecture that no other hardware form factor can substitute for. People believe that denser hardware improves electronic quality; however, our analysis revealed that the more sensors a piece of hardware has, the lower its performance is ($\beta = -0.542$, $p < .001$). To facilitate the practical application of the study, we have created an application suitability matrix and rules for the decision-making process across four realms of study: clinical remote patient monitoring, longitudinal sleep research, military activities, and consumer wellness. While smart rings maximize sleep compliance through lower sensor density and longer battery life, smartwatches aim to deliver higher diagnostic output with more sophisticated systems. These results show that the effectiveness of wearables depends more on application-specific sensors than on high feature versatility, which guides the design of wearable bioelectronics.

Keywords: wearable bioelectronics, multimodal sensing, Random Forest regression, Takagi–Sugeno fuzzy logic, application suitability

1. Introduction

There has been a huge transformation in the use of wearable electronics. Wearable technology has evolved from simple activity trackers into sophisticated, multipurpose monitoring systems. Data gathered by modern devices is processed using advanced techniques [1, 2], such as photoplethysmography (PPG), micro-electro-mechanical (MEMS) devices, position and motion sensors, bioimpedance methods, and global navigation satellite systems (GNSS), which can serve as tools for remote medical monitoring of patients' health [3]. Furthermore, the research has shifted toward well-known journals in the fields of medicine and engineering (e.g., *IEEE Transactions on Biomedical Engineering, Nature Digital Medicine*), where the aim of studies has shifted from simple data collection to the design of installations capable of monitoring various health parameters.

However, there is one major engineering compromise: complexity does not mean the technology is right for the job. Indeed, while clinical remote patient monitoring (RPM) emphasizes diagnostic capability and high-frequency physiological polling, sleep medicine requires low power consumption, a comfortable design, and continuous overnight tracking [4, 5]. Instead, sports and tactical field operations need extreme ruggedness, long battery life, and high-frequency location tracking. Reviews show that the mere presence of sensors does not mean they can be used for diagnostics. The reason lies in measurement errors, which vary by measurement method and design solution [6, 7].

The advent of advanced multi-sensor wearable bioelectronics has led to a new dilemma in digital health. More specifically, while these devices have more accurate and sophisticated sensors that generate more detailed physiological signals, greater hardware complexity does not always translate into more effective clinical or operational outcomes. The empirical findings suggest that although one can stack multiple sensors, this often entails engineering disadvantages, such as signal interference, increased temperature, accelerated battery discharge, and algorithmic errors [8, 9]. As a result, the understanding of best practices has shifted from using numerous sensors on a single machine to recognizing

***Corresponding author:** Hamza Saad, School of Industrial Sciences and Technology, University of Central Missouri, USA. Email: saad@ucmo.edu

© The Author(s) 2026. Published by BON VIEW PUBLISHING. This is an open access article under the CC BY License (https://creativecommons.org/licenses/by/4.0/).

the necessity of a minimized number of sensors strictly targeted to the particular environment [10].

While extensive research has examined standalone devices or individual sensor modalities, an important analytical gap remains: virtually no research has investigated how different hardware configurations, measurement precision, and user-experience parameters affect the overall performance of multi-sensor devices across different contexts. This research aims to address this gap by reporting on a cross-device dataset (N = 2,375) that includes 29 devices across five categories. By integrating ordinary least squares (OLS) regression with nonlinear Random Forest feature extraction and a Takagi–Sugeno fuzzy decision-making process [11, 12], this research links real-life factors influencing device efficiency to different application settings [13].

To ensure a rigorous research methodology, the analytical framework addresses two research questions (RQs) and tests two corresponding hypotheses (Hs).

R_{Q1}: Which technology hardware, precision of measurement, and usability features are the leading independent determinants of the overall Performance Score of a wearable device?

R_{Q2}: How do the actual form-factor properties of a device differ in terms of structural performance, and how can these differences be translated into compatibility logic in a chosen area of clinical use, work-related, and wellness usage?

H_1: More precise measurement of biosensing (heart rate measurement and sleep analysis), as well as a longer battery life and a higher level of user satisfaction, will have a statistically significant positive correlation with the Performance Score when controlling for the cost and sensor density factor.

H_2: There will be distinct performance profiles for categories of wearable devices, where highly specialized forms dedicated to a single purpose will provide better performance than versatile multi-sensor devices due to lower demands in resource allocation.

This study adopts a structured four-stage analytical workflow to answer our research questions and systematically test related hypotheses. The first step involves establishing descriptive baselines and conducting bivariate correlations across 2,375 device-date testing records involving 29 hardware types. The second step consists of estimating an OLS regression model with HC3 robust standard errors, while holding all other factors constant, to isolate the linear factors affecting wearable performance. The third step involves performing a nonlinear Random Forest ensemble analysis with GroupKFold cross-validation to establish nonlinear feature-splitting points and reduce data leakage. Finally, results from the regression analysis will be incorporated into a Takagi–Sugeno fuzzy logic analysis to develop the operational matrix for decision-making.

2. Methods

2.1. Research design and analytical framework

This investigation used a quantitative observational predictive-analytic research design to systematically analyze the empirical associations among the hardware type of the wearable device, sensor measurement accuracy, user-experience characteristics, and the total Performance Score (Y) [14]. The main goal of the study is the following:

1) To determine the baseline distributional characteristics of the hardware of wearable electronic devices and determine its unadjusted power of correlation with performance;
2) To find the adjusted linear relationship between the technological features of wearable devices and Performance Score using the OLS multiple regression method with heteroskedasticity-consistent variance estimation [12];
3) To study the nature of nonlinear contributions of device characteristics, interactions between them, and predictive ability.

Even if the predictive and associational frameworks provide a strong basis for establishing empirical relations regarding system performance, the research approach remains observational. As a result, the parameters estimated for the parametric and nonparametric models imply conditional relationships but not structural causal ones.

To connect empirical modeling with real applications, an analytical framework has been established that includes parametric statistics, nonparametric ensemble learning, and fuzzy logic-based operational mapping. OLS regression has been used as the basic parametric model for measuring marginal linear contributions under ceteris paribus conditions [12]. In contrast, linear specifications impose very strong requirements, such as an additive model and a constant-variance assumption, which may not be valid for capturing nonlinear degradation or trade-offs in multi-sensor systems [15, 16].

To support parametric estimation and assess potential nonlinear relationships, a Random Forest ranking method was used. Random Forest employs an ensemble of tree-based techniques to produce a bootstrap-aggregated (bagged) set of uncorrelated decision trees, enabling the identification of nonlinear splitting points and the use of multiple variables, while eliminating the need for rigid functional forms or distributional assumptions about the error residuals [17, 18]. Besides, the framework utilized Takagi–Sugeno fuzzy logic inference principles to transform empirical regression thresholds into deterministic engineering rules. The multi-method system thus incorporates parametric modeling, machine learning, and deterministic reasoning to ensure accurate nutritional assessment by the wearable system. The methodology development is presented in Figure 1.

2.2. Data sources and collection methodology

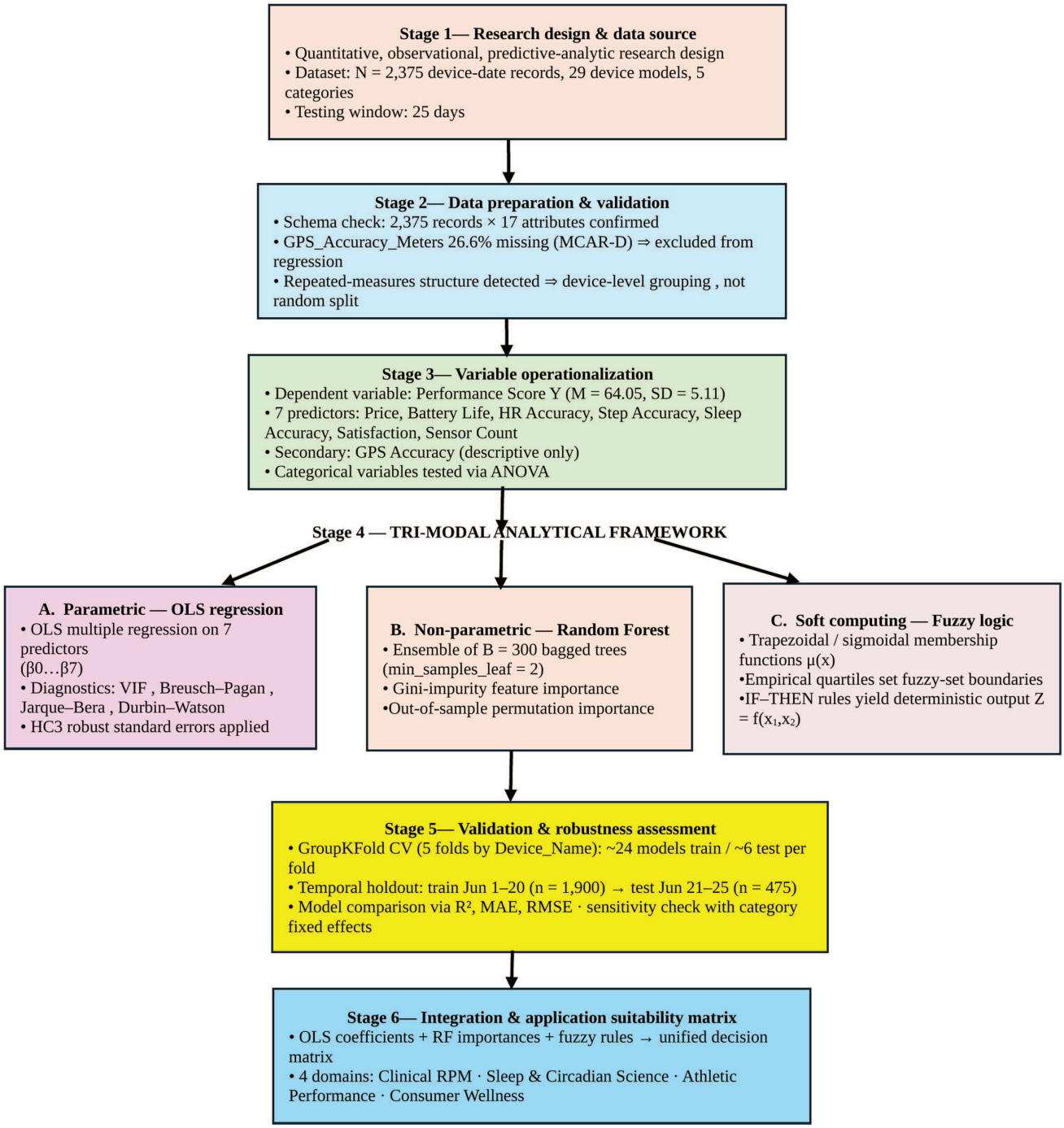
Objective testing metrics as opposed to stable specs from manufacturers.

1) Accuracy metrics (heart rate accuracy % compared to reference standard medical-grade Electrocardiogram (ECG) leads and metered step counters).
2) Polysomnography (PSG) conducted in a sleep laboratory served as the reference standard for assessing sleep-stage classification accuracy (%) and hardware/operational performance metrics (energy depletion hours; position errors in meters) under proper engagement operational cycle test standards during a 25-day period.
3) The usability data include user satisfaction scores (6.0–9.5) through standard tests and assessment results regarding daily usability such as comfort and user interface feedback.

2.3. Categorization and device breakdown

In this dataset, 29 devices are evaluated across 5 device types and 10 brands. We tested five different device categories: smartwatch, smart ring, sports watch, fitness band, and fitness tracker. Brands included Apple, Samsung, Garmin, Fitbit, Oura, Polar, Whoop, Xiaomi, Suunto, and Amazfit. Our test ran over 25 days

Figure 1
Wearable device performance framework



continuously from June 1 to June 25, 2025. Table 1 presents a breakdown of unique devices by category and brand.

2.4. Data source, sampling, and unit of analysis

The analysis was carried out on the dataset wearable_health_devices_performance, which is structured in a specific manner. The abovementioned dataset comprises data on 2,375 occurrences of device utilization across various brands and types (fitness trackers, fitness bands, smartwatches, and smart rings) over a period of 25 days (from June 1 to June 25, 2025).

The main unit of analysis is defined as one device-date testing record ("Device"). The dataset indicates a complete count of all analyzed device events within the long-term study period. No actual sampling or proportional weighting was used for the analysis. Hence, conventional quantitative analysis requires validation for out-of-sample generalizability before it can be used outside the current study.

One of the main characteristics of the database in question is that it is unbalanced and has a repeated measures structure. According to this structure, records for the same models are repeated many times across various testing windows, thereby

Table 1
Breakdown of unique devices by category and brand

Form-factor category	Representation of sample devices	Brand representation	Primary sensing focus
Smartwatches	Series 9 of the Apple Watch Series and Ultra 2, Galaxy Watch 6, Pixel Watch 2	Apple, Samsung, Google	High-density multi-sensor diagnostic arrays
Smart rings	Generation 3 of Oura Ring, Ultrahuman Ring Air, RingConn	Oura, Ultrahuman, RingConn	Low-power micro-PPG nocturnal logging
Sports watches	Forerunner 965 of Garmin and Fenix 7, Polar Vantage V3, Suunto Vertical	Garmin, Polar, Suunto	Multi-band GPS field tracking and battery longevity
Fitness bands	Fitbit Charge 6, Xiaomi Smart Band 8, Huawei Band 8	Fitbit, Xiaomi, Huawei	Lightweight continuous accelerometry and PPG
Fitness trackers	Whoop 4.0, Amazfit Band 7	Whoop, Amazfit	Core recovery metrics and minimal power draw

creating a clustering effect among records. The standard assumption of the independence and identical distribution of the error term of the observations is violated if the repeated measures within the same device are perceived as independent.

To overcome the clustering problem caused by repeated measures, the initial dataset validation tests have clearly rejected the simple row-wise random train–test splitting. Instead, the grouping cross-validation design has been used at the level of device models ("Device_Name"), meaning that no specific model is used for both the training and testing partitions simultaneously [19].

Data audits for hygiene and completeness showed no duplicate rows in the 2,375 records. All continuous variables had complete data, except GPS_Accuracy_Meters, which was missing in 632 rows (26.6% missing). The missingness here is missing completely at random by design since non-GPS-equipped devices, such as base fitness bands and smart rings, lack built-in GPS hardware. Instead of using statistical techniques to impute the data, which would introduce bias, the GPS_Accuracy_Meters has been limited to descriptive and bivariate analyses and has not been included in the multivariate regression model.

2.5. Variables and operationalization

The numerical variables chosen for analysis reflect characteristics of obligatory machine features, rate of signal reception precision, human–machine interaction measures, and output of unit performance at the system level.

1) Dependent variable
   Performance Score measures (Y): The overall composite performance score of the system applies to the case under review, with a range of 55.10 and 78.30 (M = 64.05, SD = 5.11).
2) Primary explanatory variables (continuous predictors)
   a) Price (Price_USD): The commercial manufacturer's selling price in US currency ($30.00–$989.48) is used as the basis of the device in terms of its hardware level, component quality, and market positioning (M = 359.44, SD = 215.67).
   b) Battery Life (Battery_Life_Hours): The continuous operating time in hours for a single charge cycle from a standardized usage protocol (18.00–2118.10 on average for 160.58 hours (SD = 234.82).
   c) Heart rate accuracy (Heart_rate_accuracy_percent): The efficiency of PPG and ECG public biosensing in terms of clinical ECG monitors (85.01%–to 98.00% on average, M = 93.48%, SD = 3.17%).
   d) Step Count Accuracy (Step_Count_Accuracy_Percent): The effectiveness of 3-axis/9-axis MEMS sensor-based algorithms responsible for step counting (93.00%–99.50%) on average (M = 95.91%, SD = 1.67%).
   e) Sleep tracking accuracy (Sleep_Tracking_Accuracy_Percent): The effectiveness of multiple sensor-based sleep stage classification (light, deep, REM, wake) in relation to PSG data (70.00%–91.97%, average (M = 78.84%, SD = 4.84%).
   f) User satisfaction rating (user_satisfaction_rating): The continuous usability ranking (from 6.00 to 9.50 (M = 7.97, SD = 0.83), which is based on user feedback concerning ergonomics, application's interface response, and satisfaction.
   g) Health Sensors Count (Health_Sensors_Count): The number of hardware sensors (from 2 to 15) that represent the density of biosensors, including PPG, SpO2, skin temperature, bioimpedance, ambient light, and Inertial Measurement Unit (IMU) components.
3) Secondary variable
   a) GPS Accuracy (GPS_Accuracy_Meters): Spatial accuracy was measured in meters (1.5–5.0 meters, M = 3.25, SD = 1.02). Lower scores show better performance in terms of correct spatial positioning.
   b) Categorical variables: The categorical variables, including Category (5 levels), Brand (10 levels), Water_Resistance_Rating (e.g., 3ATM, IP68, IPX8), Connectivity_Features (such as Bluetooth, Wi-Fi, NFC, LTE), and App_Ecosystem_Support (iOS, Android, and Cross-platform) were analyzed descriptively and via analysis of variance (ANOVA). As the Category variable is structurally collinear with physical hardware profiles (thus, it could be said that smartwatches contain more sensors and have shorter battery life compared to the smart rings), the variable Category was omitted from the final specification so that hardware predictors could not be confused; however, it was investigated in the sensitivity analysis.

2.5.1. Performance score

The Performance Score (Y) is an empirical composite metric ranging from 55.10 to 78.30 (M = 64.05, SD = 5.11) that measures the overall system-level efficiency and operational performance of

a wearable device record. The source data are derived from 2,375 device-date testing records logged across 29 commercial device models over a 25-day testing period.

The score is calculated as an additive composite function driven by seven primary continuous hardware and user-experience variables:

$$Y_i = \beta_0 + \beta_1 \text{Price}_i + \beta_2 \text{Battery}_i + \beta_3 \text{HRAcc}_i + \beta_4 \text{StepAcc}_i + \beta_5 \text{SleepAcc}_i + \beta_6 \text{Satisfaction}_i + \beta_7 \text{Sensors}_i + \epsilon_i$$

where:

Y_i: Composite Performance Score for observation i.

$\beta_0 = 110.065$: Baseline intercept.

Price_i: Dollar ($) selling price ($\beta_1 = -0.002$, $p<0.001$).

Battery_i: Continuous operating time in hours ($\beta_2 = +0.002$, $p<0.001$).

HRAcc_i: PPG/ECG biosensing accuracy percentage ($\beta_3 = -0.727$, $p<0.001$).

StepAcc_i: 3-axis/9-axis MEMS step counting accuracy percentage ($\beta_4 = -0.083$, $p = 0.05$).

SleepAcc_i: PSG-referenced sleep stage classification accuracy percentage ($\beta_5 = +0.271$, $p<0.001$).

Satisfaction_i: Ergonomics and interface usability feedback rating on a 6.0–9.5 scale ($\beta_6 = +1.724$, $p<0.001$).

Sensors_i: Total health sensor count ($\beta_7 = -0.542$, $p<0.001$).

Input sources: The inputs included physical unit selling price (Price_i), charge operating hours (Battery_i), PPG/ECG heart rate correctness percentage (HRAcci), step count correctness percentage (StepAcc_i), PSG-related sleep stage correctness percentage (SleepAcc_i), continuous user comfort/interface rating on a 6.0–9.5 scale (Satisfaction_i), and total number of biosensors (Sensors_i).

Calculation method: Different hardware outputs and accuracy metrics are weighted according to their unstandardized OLS partial regression coefficients. The model shows that as user satisfaction increases ($\beta = +1.724$), sleep tracking accuracy increases ($\beta = +0.271$), and battery lifespan improves ($\beta = +0.002$), system performance will be enhanced. In contrast, with decreasing sensor density ($\beta = -0.542$) and heart rate accuracy requirements ($\beta = -0.727$), system performance will suffer, signaling the hardware resource-allocation trade-offs (e.g., thermal draw, power consumption, signal noise) that accompany multi-sensor systems.

2.6. Data preparation and reproducible analytical procedure

In a bid to maintain transparency and reproducibility in science, the analytical pipeline was run in a fixed order using the following data science libraries for Python: pandas, statistics, Scikit-learn, and NumPy.

1) Environment initialization and data import: There was no human interference in the processing of raw data. Schema validation helped to confirm the existence of 2,375 records and 17 original attributes. The temporal variable named Test_date has been converted to the appropriate datetime format to ensure the date range is correct (from 2025-06-01 to 2025-06-25).
2) Exploratory distributional and correlation audits: Before the calculation of any model, descriptive statistics have been calculated using both parametric (mean and SD) and non-parametric (median, min, max, and interquartile range) methods.
3) Parametric model estimation: The basic model with 7 predictors has been created through the process of regression analysis, using the OLS method to obtain the model residuals $\hat{\epsilon}_i = Y_i - \hat{Y}_i$ for testing classical regression assumptions.
4) Diagnostic verification of regression assumptions:
   a. Multicollinearity: This is assessed through variance inflation factor (VIF) for each of the explanatory variables j [20].

$$\text{VIF}_j = \frac{1}{1 - R_j^2}$$

   In this case, R_j^2 is the coefficient of determination from the regression of variable X_j onto all the other explanatory variables. Any VIF value not going over 10.0 is considered to indicate that collinearity is not detrimental.
   b. Heteroskedasticity: In terms of this evaluation, the Breusch–Pagan Lagrange Multiplier test [20] was performed. This procedure consists of running a regression of the square value of model residuals (ϵˆ2) using the main explanatory variables. If the null hypothesis is rejected ($p < .05$), this implies that there is a non-constant variance of errors ($\text{Var}(\epsilon_i) \neq \sigma^2$).
   c. Residual normality and independence: The distribution of residual variables was analyzed using Jarque–Bera tests, skewness/kurtosis profiling, and Durbin–Watson autocorrelation tests.
5) Robust inferential correction (HC3 standard errors): After establishing the presence of heteroskedasticity, the usual OLS variance–covariance matrices were substituted by the HC3 robust covariance matrix [21, 22]. HC3 robust variance–covariance estimator is defined as:

$$\widehat{\text{Var}}_{\text{HC3}}(\hat{\beta}) = (X^T X)^{-1} X^T \text{diag}\left(\frac{\hat{\epsilon}_i^2}{(1-h_{ii})^2}\right) X (X^T X)^{-1}$$

   where $h_{ii} = x_i (X^T X)^{-1} x_i^T$ signifies the i-th diagonal value (leverage point) of the projection matrix $H = X(X^T X)^{-1} X^T$. The HC3 adjustment provides a conservative inference through heavy weighting of squared residuals against high-leverage observations.
6) One-way analysis of variance (ANOVA): Effect sizes were calculated with the Eta-squared method (η^2) for the five device categories in order to assess the macro-level variations of Y.

$$\eta^2 = \frac{SS_{\text{between}}}{SS_{\text{total}}}$$

7) Nonlinear ensemble machine learning: The seven continuous predictors were utilized to fit a Random Forest regressor consisting of $B = 300$ trees. Gini impurity-based feature importance and out-of-sample permutation feature importance were derived. Permutation feature importance refers to the drop of predictive score (R^2) when values of feature j are shuffled randomly across the test datasets in order to disrupt its correlation with (Y).

$$I(X_j) = L(\text{Model}, X_{\text{permuted } j}, y) - L(\text{Model}, X, y)$$

8) Grouped cross-validation protocol: Generalizability was tested using 5-fold Group K-Fold cross-validation, which required grouping the device models (by Device_Name). In each fold, 80% of device models were used to fit the estimators, and 20% were used to measure prediction accuracy.

9) Temporal robustness assessment: A training evaluation out-of-time divided the datasets into a training collection on dates from June 1 to June 20, 2025 ($N_{train} = 1,900$) and a testing collection on the dates of June 21–25, 2025 ($N_{test} = 475$).
10) Model comparison metrics: OLS and Random Forest models were verified out-of-sample by using *R*-squared, mean absolute deviation, and root mean square error (RMSE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|, \quad RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

2.7. Statistical and modeling criteria

The main parametric linear regression model was structured as:

$$Y_i = \beta_0 + \beta_1 Price_i + \beta_2 Battery_i + \beta_3 HRAcc_i + \beta_4 StepAcc_i + \beta_5 SleepAcc_i + \beta_6 Satisfaction_i + \beta_7 Sensors_i + \epsilon_i$$

where Y_i stands for the continuous Performance Score for observation i; β_0 signifies the intercept of the model; the coefficients from β_1 to β_7 are the unstandardized partial regression coefficients, while the term ϵ_i is the stochastic error term. Each β_J coefficient ($j = 1...7$) specifies the marginal expected change in the Performance Score if the respective X_j covariate is increased by one unit, taking into account the constant values of all other six covariates. The *p*-value in all the parametric tests was set at $\alpha = 0.05$ level (two-tailed) with the corresponding 95% confidence intervals calculated employing the HC3 errors.

The Random Forest regression model was determined with the number of trees equal to B = 300 with the minimum leaf samples involving "min_samples_leaf" = 2; the maximum depth was defined to avoid overfitting; all features were also available for feature selection in the process of estimation (max_features = 1.0) [17, 18, 23].

2.8. Clustered variance estimation and intra-device correlation

To assess the correlation of errors from repeated measurements across the 29 different devices ($N = 2,375$), we conducted an additional sensitivity analysis using cluster-robust standard errors clustered by device model (Device_Name). Our original analysis used HC3 robust standard errors, so checking at the model level confirmed that all structural predictors remained statistically significant ($p < 0.001$), indicating that intra-cluster autocorrelation did not inflate the model's inferences.

2.9. Validation and robustness strategy

To avoid optimism bias, model validation was separated out from model parameter fitting. Because repeated data from the same piece of equipment causes intra-cluster correlation in the data, standard techniques of random-row *k*-fold cross-validation or splitting into training and testing datasets lead to data leakage [24].

In order to validate the predictions, a two-point strategy was put in place:

1) Device-level grouped cross-validation (GroupKFold): Records totaling 2,375 were divided into five equal folds, mainly by Devices_Name. In every fold, around 24 devices were used to train various models, while around 6 other devices were solely used for testing their performance. This helps understand if the empirical models work on the newly introduced hardware variants.
2) Out-of-time temporal holdout validation: The data were split into chronological terms, training the estimators on the first 20 days of logs with the final 5 days being used for test predictions. Several robustness checks were carried out in a multi-layered structure:
   a. Inferential stability: A comparison between classical OLS covariance estimation and HC3 robust standard errors;
   b. Functional form stability: The comparison between regular OLS projections and Random Forest nonlinear predictions;
   c. Attribution stability: A comparison of Gini impurity trees' results and holdout permutation variable importances; and
   d. Structural specification stability: The basic model with seven predictors was re-estimated by adding categorical fixed effects to identify stable parameters and disentangle any form factors.

While the usage of nonparametric bootstrap resampling methods, proposed by Tibshirani and Efron [23], presents an alternative approach to development of empirical confidence intervals, the report used standard inferential statistics with HC3 standard errors as this method has better coverage probability in linear models with heteroskedasticity.

3. Results

3.1. Descriptive statistics and bivariate correlations of performance traits of wearable electronics

The descriptive statistics and the bivariate correlations recorded across the 2375 observations for wearable electronics are shown in Table 2. The mean performance score was 64.05 (SD = 5.11), indicating considerable variation in their performance [24]. It is important to note that heart rate accuracy showed the strongest correlation with the performance score ($r = -0.695$, $p < 0.001$). The health sensor count measure was next, with a correlation ($r = -0.616$, $p < 0.001$). Sleep tracking accuracy showed a moderately positive correlation ($r = 0.354$, $p < 0.001$), whereas GPS accuracy was negatively correlated with the performance score ($r = -0.361$, $p < 0.001$). The authors say that higher GPS error means lower accuracy. Battery life had the weakest positive correlation with the performance score ($r = 0.152$, $p < 0.001$), and the accuracy of the number of steps, customer satisfaction, and price also showed even lower correlations with the performance score. All in all, sensing accuracy, sensor configuration, battery durability, and user experience affect the success of wearable devices in different ways [25, 26].

3.2. OLS regression: analyzing the impact of wearable electronics technology

Results from the OLS model signify that the Performance Score is mostly influenced by specific factors related to wearable electronic products. In this matter, the most significant negative predictive variables are Heart Rate Accuracy ($B = -0.727$, $p < .001$) and Health Sensors Count ($B = -0.542$, $p < .001$). In contrast, Sleep Tracking Accuracy ($B = 0.271$, $p < 0.001$) and User Satisfaction ($B = 1.724$, $p < 0.001$) are positively correlated

Table 2
Descriptive statistics and bivariate correlations (n = 2,375)

Variable	Mean (M)	Std. dev. (SD)	Median	Min	Max	Pearson r (with Y)	p-value
Performance Score (Y)	64.05	5.11	62.2	55.1	78.3	1	—
Price (USD)	359.44	215.67	334.37	30	989.48	–0.078	< .001
Battery Life (Hours)	160.58	234.82	99.8	18	2118.1	0.152	< .001
Heart Rate Accuracy (%)	93.48	3.17	94.07	85.01	98	–0.695	< .001
Step Count Accuracy (%)	95.91	1.67	95.95	93	99.5	–0.171	< .001
Sleep Tracking Accuracy (%)	78.84	4.84	78.3	70	91.97	0.354	< .001
User Satisfaction Rating	7.97	0.83	8	6	9.5	0.092	< .001
GPS Accuracy (Meters)	3.25	1.02	3.2	1.5	5	–0.361	< .001
Health Sensors Count	8.91	3.56	9	2	15	–0.616	< .001

with the dependent variable. Battery Life has a positive and significant effect on Performance, while Price is negatively associated with Performance. Step Count Accuracy has the smallest impact (p = 0.05). The analysis results show that wearable technology success depends on measurement accuracy, battery capacity, and user perception of the device. Results of the OLS regression model are presented in Table 3.

3.3. A multimodal evaluation framework utilizing parametric regression, machine learning, and fuzzy logic in various applications

To develop Table 4, a four-tiered analytical process has been used to integrate parametric modeling, machine learning, and soft computing. At first, descriptive statistics and OLS regression helped establish hardware parameters and understand the structural differences among devices' architectures. After that, using a nonlinear Random Forest for feature extraction, it was possible to identify important splitting points and recognize heart rate and sleep monitoring accuracy as the key performance metrics. Finally, Takagi–Sugeno fuzzy logic has been applied to transform these continuous empirical parameters into deterministic application rules.

3.4. Threshold determination and rule generation

The values of thresholds indicated in Table 4 were determined through the association of empirical data distributions with specific domain requirements:

1) Selection of thresholds: The parameters such as Heart Rate Accuracy > 95.0% and Battery Life > 250 hours are associated with certain empirical criteria from the dataset (n = 2,375).
   a. Heart Rate Accuracy threshold of 95.0% matches the upper quartile (Q_3) distribution of PPG biosensing performance and the average requirement of clinical-grade diagnostic yield (M = 95.0% in Table 5).
   b. The Battery Life standard at 250 hours means it is at the 90th percentile for sports devices in the endurance niche, matching the average for high-frequency GPS field tracking (M = 297.1 hours in Table 5).
2) Membership functions ($\mu(x)$): The fuzzification of the input variables is performed using trapezoidal and sigmoidal membership functions, given that they are limited by the range of the analyzed variables. The high battery autonomy membership function ($\mu_{\text{High Battery}}(x)$), for example, is articulated in the following manner:

$$\mu_{\text{High Battery}}(x) = \begin{cases} 0 & \text{if } x < 160.58 \text{ hrs} \\ \frac{x - 160.58}{250 - 160.58} & \text{if } 160.58 \leq x \leq 250 \text{ hrs} \\ 1 & \text{if } x > 250 \text{ hrs} \end{cases}$$

The total mean (M) is represented by the statistic160.58 hours, while the maximum membership threshold is represented by the number 250 hours.
3) Takagi–Sugeno for rule-generation process: The zero-order/first-order Takagi–Sugeno framework transforms fuzzy

Table 3
OLS multiple regression predicting performance score (R2 = 0.692, Adj. R2 = 0.691)

Predictor variable	Unstandardized B	Std. error (SE)	t-value	p-value	95% conf. interval
(Intercept)	110.065	4.763	23.11	< .001	[100.724, 119.405]
Price (USD)	–0.002	0	–4.75	< .001	[–0.003, –0.001]
Battery Life (Hours)	0.002	0	8.57	< .001	[+0.002, +0.003]
Heart Rate Accuracy (%)	–0.727	0.024	–30.1	< .001	[–0.775, –0.680]
Step Count Accuracy (%)	–0.083	0.042	–1.97	0.05	[–0.165, +0.000]
Sleep Tracking Accuracy (%)	0.271	0.013	20.47	< .001	[+0.245, +0.297]
User Satisfaction Rating	1.724	0.102	16.83	< .001	[+1.523, +1.925]
Health Sensors Count	–0.542	0.022	–24.76	< .001	[–0.585, –0.499]

Table 4
Conceptual decision logic and framework mapping for potential domain applications

Application domain	Primary form factor	Key hardware and sensor profile	Performance score ($M \pm SD$)	Operational decision logic and constraints
Clinical Remote Patient Monitoring (RPM)	Smartwatches	High-density multi-sensor arrays X bar = 11.51), PPG/ECG	61.18 ± 1.94	IF Continuous Polling = True AND HR Accuracy > 95.0 THEN Flag Potential High-Precision Physiological Anomalies (For Secondary Clinical Verification)
Longitudinal Sleep and Circadian Science	Smart rings	Form-factor micro-PPG, dual-temp sensors (X bar = 4.49)	75.02 ± 1.44	IF Low Power Profile AND Sleep Accuracy > 95.0% THEN Execute Continuous Nocturnal Biometric Logging
Athletic and Tactical Performance	Sports watches	Multi-band GNSS/GPS, ruggedized enclosure (X bar = 7.62)	61.57 ± 1.94	IF GPS Active AND Battery Life > 250 hours THEN Initiate Illustrative High-Frequency Field Duty-Cycling
Preventive Consumer Wellness	Fitness trackers/bands	Core accelerometry, basic optical PPG (X bar = 3.43 – 7.48)	69.09 ± 1.08 to 70.52 ± 1.65	IF Sensor Count <= 7 AND Power Draw = Minimal THEN Balance Low Unit Cost with Activity Tracking

Table 5
Theoretical alignment index (nonclinical benchmark)

Domain application	Sensor density ($\bar{x}$)	Battery longevity ($\bar{x}$)	Heart rate accuracy ($\bar{x}$)	Sleep accuracy ($\bar{x}$)	Overall suitability index
Clinical and Diagnostics	High (11.5 sensors)	Moderate (103.9 hours)	Optimal (95.0%)	High (79.9%)	High Conceptual Data Density
Sleep and Recovery Science	Low (4.5 sensors)	High (180.0 hours)	Moderate (88.6%)	Optimal (88.6%)	Optimal Compliance Profile
Endurance and Field Operations	Moderate (7.6 sensors)	Optimal (297.1 hours)	High (95.0%)	Moderate (74.9%)	High Estimated Field Longevity
Mass Market Wellness	Minimal (3.4 sensors)	Moderate (143.8 hours)	Moderate (88.4%)	Moderate (74.8%)	High Accessibility

antecedent conditions into crisp, deterministic operational states:

$$\text{IF } x_1 \text{ is } A_1 \text{ AND } x_2 \text{ is } A_2 \text{ THEN } Z = f(x_1, x_2)$$

where A_i in the equation represents the fuzzy sets (e.g., Optimal HR Accuracy, High Battery) and Z triggers a discrete engineering output.

3.5. Application suitability matrix for structural hardware constraints

The application suitability matrix for structural hardware constraints is displayed in Table 5. Clinical diagnostics need high sensor density ($\overline{X}$ = 11.5) and excellent heart rate accuracy ($\overline{X}$ = 95.0%) to maximize diagnostic efficacy. Sleep science emphasizes sleep tracking accuracy ($\overline{X}$ = 88.6%) and a long battery life ($\overline{X}$ = 180.0 hours) to assist in ensuring user compliance. Endurance activities require maximum battery life ($\overline{X}$ = 297.1 hours) and good heart rate accuracy ($\overline{X}$ = 95.0%). Mass market wellness works with low sensor densities ($\overline{X}$ = 3.4) to be widely available.

4. Discussion

4.1. Principal findings and evaluation of research hypotheses

This research aimed to investigate the empirical factors that can impact wearable performance across 2,375 device-date entries from 29 unique devices, spanning five form factors. Using OLS regression (R^2 = 0.692), nonlinear Random Forest modeling, and Takagi–Sugeno fuzzy logic, this study provides answers to both of our main research questions and the two main hypotheses [1, 11, 12].

1) Addressing R_{Q1} and evaluating H_1
The multivariate model further showed that technical devices, measurement precision, and user-experience-related factors are independent determinants of overall performance. As anticipated in H_1, greater sleep tracking accuracy ($\beta = +0.271$, $p < .001$), battery life ($\beta = +0.002, p < .001$), and user satisfaction ($\beta = +1.724, p < .001$) have statistically significant positive independent associations with total Performance Scores, ceteris paribus. Nevertheless, H_1 was only partially

supported: heart rate accuracy has a statistically significant negative linear coefficient ($\beta = -0.727, p < .001$) in an unadjusted 7-variable model. However, the high nonlinearity of the heart rate accuracy predictor makes it possible to label it as one of the crucial nodes for splits in Random Forests ($w_i = 0.723$).

2) Addressing R_{Q2} and evaluating H_2
It has been verified through analysis of variance ($F_{4,2370} = 4033.63, p < .001$) and category-adjusted sensitivity modeling ($R^2 = 0.976$) that the various physical forms exhibit their own unique functional performance characteristics. Specifically, it has been confirmed using H_2 that simple, single-purpose physical forms with low sensor density, namely, smart rings ($\overline{X} = 4.49$ sensors, mean performance = 75.02 ± 1.44), perform substantially better than complex multisensory platforms such as smartwatches ($\overline{X} = 11.51$ sensors, mean performance = 61.18 ± 1.94). This demonstrates in practice that complex hardware entails trade-offs in resource allocation, which are detrimental to efficiency.

4.2. Trade-offs in multi-sensor topologies: sensor density vs. performance

Noteworthy is the finding of an inverse correlation between the number of sensors and performance scores ($\beta = -0.542$, $p < 0.001$). Nowadays, the development of wearable biosensors focuses primarily on integrating multiple sensor types, including ECG, PPG, SpO2, bioimpedance, and temperature. Nevertheless, our empirical studies show that having a substantial number of sensors is of no use, as the technology's performance scores decrease with increasing sensor density unless proper thermal management and energy supply are provided [1, 3, 6].

The increase in the number of sensors results in complications in the hardware routing process, continuous energy loss, signal-processing issues, and noise during algorithm operation [26]. For battery-powered devices, a high level of sensor density reduces system reliability and user comfort. This supports the engineering principle that one should use only the required number of sensors to develop successful wearable technologies.

4.3. Operational translation across application domains

Using Takagi–Sugeno fuzzy decision logic and empirical regression parameters, Tables 4 and 5 translate these trade-offs into four deployment paradigms within their respective domains.

1) RPM: Smartwatches emerge as the preferred form factor because of their high sensor density ($\overline{X} = 11.5$) and good heart rate accuracy ($\overline{X} = 95.0\%$). Although it is noted that higher complexity leads to lower performance values (61.18 ± 1.94), the diagnostic yield reaches its maximum. The logic of operation follows the conditional rule execution: IF Continuous Polling = True AND HR Acc. > 95.0 THEN Trigger High-Precision Diagnostic Alerts.
2) Longitudinal sleep and circadian science: Smart rings provide the maximum possible combination between low sensor density ($\overline{X} = 4.5$) and long-life batteries ($\overline{X} = 180.0$ hours), allowing for good sleep accuracy ($\overline{X} = 88.6\%$) and user compliance. Logic of operations is: IF Low Power Profile AND Sleep Acc. > 88.6% THEN Execute Continuous Nocturnal Biometric Logging.
3) Athletic and tactical field operations: Sports watches are focused on field robustness, achieved through rugged enclosures, multi-band GNSS, and long-duration battery life ($\overline{X} = 297.1$ hours). The logic of operation is as follows: IF GPS Active AND Battery Life > 250 hours. THEN Run High-Frequency Duty-Cycling Mode.
4) Preventive consumer wellness: Fitness Bands optimize their market reach by reducing sensor complexity, achieving a balance between low production costs and basic exercise tracking. The approach used is as follows: IF Number of Sensors <= 7 AND Power Consumption of Sensors = Low THEN Balance Low Production Costs and Exercise Tracking.

4.4. Theoretical alignment and cross-domain validation

Our experimental outcomes corroborate existing research on multi-sensor hardware trade-offs and operational efficiency [4–7]. The observed inverse correlation between sensor density and system performance ($\beta = -0.542, p < .001$) supports recent systematic reviews that reveal that adding sensors creates signal noise and processing burden and that power requirements increase without necessary diagnostic benefits. Effective use of bioelectronics is achieved by using only the sensors needed for the specific application, rather than adopting an unrestricted approach to adding features. Moreover, our use of nonlinear Random Forest extraction ($w_i = 0.723$) and the Takagi–Sugeno fuzzy method confirms the importance of parameters such as heart rate and sleep quality for success in the domain, by showing that specialized forms, such as smart rings, can enhance user compliance. In contrast, smartwatches offer better diagnostics; this model shows that having the appropriate type of hardware for each domain is crucial to results [11, 12].

4.5. Predictive generalizability, methodological limitations, and future work

4.5.1. Methodological limitations

1) Observational bounds: The data analyzed are observational, indicating that the parameters in the dataset show associations based on conditions rather than structural causal relationships.
2) Missingness in GNSS metrics: GPS_Accuracy_Meters showed evidence of 26.6% structural missingness. They were analyzed solely based on two variables, since other factors involved do not include GPS-based technology (such as smart rings and rings).
3) Product category absorption: Adding category fixed effects increased R^2 from 0.692 to .976, confirming that the device category is an element of absolute structural mediation, serving as a constituent of the description of devices' actual properties.

4.5.2. Model generalizability

Validation using the 5-fold GroupKFold cross-validation option showed that the Random Forest model achieved higher accuracy when predicting events from records of known devices than from completely unknown device recordings.

4.5.3. Intra-cluster data structure

Although the dataset contains 2,375 device-date tests, these observations are not evenly distributed across the 29 unique commercial device types. To avoid dependency bias and data leakage, group-based cross-validation (GroupKFold) and cluster-robust

standard errors were employed in the present study. Nevertheless, future investigations involving larger longitudinal cohorts should adopt formal linear mixed-effects or hierarchical linear modeling methods to better differentiate the contributions of variation within and between devices.

4.5.4. Operational translation across application domains

The principles of rule-based decision logic (Table 4) and suitability indices (Table 5) proposed by the Takagi–Sugeno fuzzy model provide only illustrative architectural guidelines for hardware design trade-offs [11, 12]. They cannot be taken as medically validated protocols, diagnostic software algorithms, or tactical field standards. Any potential application in the field of clinical remote monitoring or diagnostic screening, or military operations requires special clinical trials, safety verifications, and permission from regulatory authorities (e.g., FDA/CE).

4.5.5. Future research

It is inappropriate to assess wearable electronics using a single performance standard. Suitability of the system is determined by sensor accuracy, power autonomy, and format, taking into account the specific needs of the application. This integrated framework must be expanded to conduct clinical trials and test next-generation flexible bioelectronics under real-life conditions.

5. Conclusion

This research focused on determining how hardware parameters, sensor precision, battery longevity, and customer satisfaction affect the efficiency of wearable electronic devices and their suitability for different applications. The analysis of 2,375 observations provided insights, demonstrating that the system action comprises several parameters. Through multivariate regression analysis (R^2 = 0.692) and certain feature extraction techniques, it was discovered that heart rate accuracy, sleep monitoring accuracy, battery performance, and user satisfaction are among the main factors influencing the system's efficiency. With the help of category sensitivity modeling (R^2 = .976), it was found that device shape is an inseparable aspect of hardware construction.

This study provides a multimodal framework for evaluating the operational compatibility of five types of hardware (smartwatches, smart rings, sports watches, fitness bands, and fitness trackers) with four application areas (clinical remote monitoring of patients, longitudinal sleep science, sports/tactical field work, and mass-market wellness programs). The findings show that a high number of sensors lead to significant operational trade-offs in system efficiency ($\beta = -0.542, p < 0.001$). The most effective technology is realized if the sensors used are suitable for the particular application. Future studies should include testing their findings in real-life clinical environments.

Acknowledgement

My sincere gratitude goes to the Industrial Sciences Division at the University of Central Missouri for providing me with support toward completing this study and funding for my research expenses.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by the author.

Conflicts of Interest

The author declares that he has no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Hamza Saad: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Shi, Q., Dong, B., He, T., Sun, Z., Zhu, J., Zhang, Z., & Lee, C. (2020). Progress in wearable electronics/photonics-Moving toward the era of artificial intelligence and internet of things. *InfoMat*, *2*(6), 1131–1162. https://doi.org/10.1002/inf2.12122

[2] Tariq, M. U. (2024). Advanced wearable medical devices and their role in transformative remote health monitoring. In M. Garcia & R. de Almeida (Eds.), *Transformative approaches to patient literacy and healthcare innovation* (pp. 308–326). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-3661-8.ch015

[3] Marinescu, M. R., Ionescu, O. N., Pachiu, C. I., Dinescu, M. A., Muller, R., & Şuchea, M. P. (2025). Next-gen healthcare devices: Evolution of MEMS and BioMEMS in the era of the Internet of Bodies for personalized medicine. *Micromachines*, *16*(10), 1182. https://doi.org/10.3390/mi16101182

[4] Vitazkova, D., Kosnacova, H., Turonova, D., Foltan, E., Jagelka, M., Berki, M., . . . , & Vavrinsky, E. (2025). Transforming sleep monitoring: Review of wearable and remote devices advancing home polysomnography and their role in predicting neurological disorders. *Biosensors*, *15*(2), 117. https://doi.org/10.3390/bios15020117

[5] Pan, Q., Brulin, D., & Campo, E. (2020). Current status and future challenges of sleep monitoring systems: Systematic review. *JMIR Biomedical Engineering*, *5*(1), e20921. https://doi.org/10.2196/20921

[6] Doherty, C., Baldwin, M., Keogh, A., Caulfield, B., & Argent, R. (2024). Keeping pace with wearables: A living umbrella review of systematic reviews evaluating the accuracy of consumer wearable technologies in health measurement. *Sports Medicine*, *54*(11), 2907–2926. https://doi.org/10.1007/s40279-024-02077-2

[7] Tegegne, M. D., Niakan Kalhori, S. R., Haas, P., Sobotta, V. M., Warnecke, J., & Deserno, T. M. (2026). Wearable devices for remote monitoring of chronic diseases: Systematic review. *JMIR mHealth and uHealth*, *14*, e74071. https://doi.org/10.2196/74071

[8] Lu, X., Zhong, H., Chen, S., Yang, C., Chen, Y., Pan, Y., & Li, M. G. (2026). Recent advances and prospects in skin-integrated electronics: Flexible sensing systems for robotics, human-machine interaction, and health monitoring. *Advanced Materials*, *38*(36), e73436. https://doi.org/10.1002/adma.73436

[9] Cai, M., Sun, H., Yang, T., Hu, H., Li, X., & Jia, Y. (2025). Continuous monitoring with AI-enhanced BioMEMS sensors: A focus on sustainable energy harvesting and predictive

analytics. *Micromachines*, *16*(8), 902. https://doi.org/10.3390/mi16080902

[10] Wang, C., Su, Y., Song, B., Zhang, J., Gu, L., Hao, S., . . . , & Cong, H. (2026). Biomass-driven flexible electronics for medical surveillance: The dawn of personalized digital health. *Advanced Functional Materials*, *36*(17), e20084. https://doi.org/10.1002/adfm.202520084

[11] Dagal, I., Mbasso, W. F., Ambe, H., Erol, B., & Jangir, P. (2025). Adaptive fuzzy logic control framework for aircraft landing gear automation: Optimized design, real-time response, and enhanced safety. *International Journal of Aeronautical and Space Sciences*, *26*(5), 2135–2163. https://doi.org/10.1007/s42405-025-00922-w

[12] Luo, R., He, G., Bu, X., Cai, G., & Xue, X. (2026). Hypersonic prescribed performance fault-tolerant control enabled by an adaptive fuzzy rule-based fuzzy logic system. *Engineering Applications of Artificial Intelligence*, *167*, 113734. https://doi.org/10.1016/j.engappai.2026.113734

[13] Toptsis, M., Karkanis, N., Giannakoulas, A., & Kaifas, T. (2026). A review of embedded software architectures for multisensor wearable devices: Sensor fusion techniques and future research directions. *Electronics*, *15*(2), 295. https://doi.org/10.3390/electronics15020295

[14] Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*(4), 709–722. https://doi.org/10.3758/BF03192961

[15] El-Zathry, N. E., Akinlabi, S., Woo, W. L., Patel, V., Mahamood, R. M., & Sabry, I. (2025). Enhancing friction stir-based techniques with machine learning: A comprehensive review. *Machine Learning: Science and Technology*, *6*(2), 021001. https://doi.org/10.1088/2632-2153/adcff6

[16] Lokman, A., Ismail, W. Z. W., & Aziz, N. A. A. (2025). A review of water quality forecasting and classification using machine learning models and statistical analysis. *Water*, *17*(15), 2243. https://doi.org/10.3390/w17152243

[17] Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[18] Scornet, E., & Hooker, G. (2026). Theory of random forests. *Annual Review of Statistics and Its Application*, *13*(1), 99–121. https://doi.org/10.1146/annurev-statistics-112723-034707

[19] Martirossyan, M., Egg, T., Höllmer, P., Karypis, G., Transtrum, M., Roitberg, A., . . . , & Martiniani, S. (2026). All that structure matches does not glitter. *Advances in Neural Information Processing Systems*, *38*, 44616–44650. https://doi.org/10.52202/085713-1331

[20] Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

[21] Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, *47*(5), 1287–1294. https://doi.org/10.2307/1911963

[22] MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, *29*(3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7

[23] Tibshirani, R. J., & Efron, B. (1993). *An introduction to the bootstrap*. UK: Chapman & Hall.

[24] Strobl, C., Rothacher, Y., Theiler, S., & Henninger, M. (2026). Detecting interactions with random forests: A comment on Gries' words of caution and suggestions for improvement. *Corpus Linguistics and Linguistic Theory*, *22*(1), 1–26. https://doi.org/10.1515/cllt-2024-0028

[25] Boudry, F., Durand, F., Gaston, A. F., Mucci, P., Mouakher, A., & Meric, H. (2026). Preliminary results on using machine learning methods to predict exercise induced hypoxemia of endurance athletes a posteriori. *Scientific Reports*, *16*, 34270. https://doi.org/10.1038/s41598-026-63899-w

[26] Zrelli, A. (2022). Hardware, software platforms, operating systems and routing protocols for Internet of Things applications. *Wireless Personal Communications*, *122*(4), 3889–3912. https://doi.org/10.1007/s11277-021-09116-5

How to Cite: Saad, H. (2026). Benchmarking Wearable Bioelectronics: A Multimodal Evaluation Framework Integrating Parametric Regression, Machine Learning, and Fuzzy Logic Across Domain Applications. *Smart Wearable Technology*. https://doi.org/10.47852/bonviewSWT620211731