Digital Signal Processing and Artificial Intelligence
for Automatic Learning

DLINE JOURNALS

ISSN:2583-5009

DSP 2026: 5 (3)
https://doi.org/10.6025/dspaial/2026/5/3/135-167

# Comprehensive Signal Processing and Anomaly Detection of Multivariate Sensor Signals Using Time - Frequency and Entropy - based Methods

Hathairat Ketmaneechairat
Department of Information and Production Technology Management
College of Industrial Technology
King Mongkut's University of Technology North Bangkok
Thailand
hathairatk@kmutnb.ac.th

## ABSTRACT

*Low cost Micro Electromechanical Systems (MEMS) sensors are widely deployed in embedded systems, the Internet of Things (IoT), and cyber physical environments; however, their raw output under nominally static conditions may exhibit substantial temporal, spectral, distributional, and statistical complexity that cannot be captured by a single analytical representation. This study proposes and applies an integrated multidomain signal processing and unsupervised anomaly detection framework for the comprehensive characterization of multivariate static MEMS accelerometer data. The framework combines time domain statistics, Fast Fourier Transform (FFT) and Welch power spectral density (PSD) analysis, short time Fourier transform spectrograms, wavelet domain energy decomposition, autocorrelation and rolling root mean square (RMS) analysis, Pearson inter channel correlation, signal to noise ratio (SNR) estimation, Shannon and spectral entropy measures, rolling statistical change point detection, and two complementary unsupervised anomaly detection methods Isolation Forest and PCA reconstruction based detection. The framework is applied to a four channel dataset acquired from a stationary MPU-6050 accelerometer, comprising three raw acceleration axes and a derived percentage error channel, containing approximately 37,459 observations. The results reveal pronounced heterogeneity among the sensor channels: Sensors 1, 2, and 4 exhibit strongly non Gaussian, heavy tailed amplitude distributions with extreme skewness and excess kurtosis, whereas Sensor 3 displays substantially greater amplitude variability, strong temporal persistence, and concentrated spectral organization. Frequency domain analysis identifies a distinctive approximately 21 sample periodic component in Sensor 2, while time frequency and wavelet analyses expose localized transient events not captured by global spectral summaries. Dual entropy analysis demonstrates that amplitude domain and frequency domain complexity are distinct and independently informative properties. Change point detection identifies two statistical regime transitions, and both Isolation Forest and PCA reconstruction consistently identify 24 anomalous windows (anomaly rate H" 1.028%) in the absence of predefined labels. The integrated findings demonstrate that no single analytical domain is sufficient to characterize static MEMS sensor behavior, and that a multidomain framework is necessary to distinguish persistent nominal structure, transient anomalous deviations, and sustained statistical regime changes. The proposed framework provides a structured, reproducible, and label-free analytical foundation for sensor monitoring and anomaly detection in resource constrained and IoT-enabled systems.*

**Keywords**: MEMS Accelerometer, Multivariate Time Series Analysis, Unsupervised Anomaly Detection, Time Frequency Analysis, Entropy Based Signal Characterization, Isolation Forest, PCA Reconstruction, Power Spectral Density, Sensor Signal Processing, IoT Monitoring

**Received:** 17 December 2025, Revised 9 April 2026, Accepted: 30 April 2026

**Copyright: DLINE**

## 1. Introduction

The rapid adoption of multi sensor technologies has enabled the acquisition of increasingly complex and high-dimensional time series data across a wide range of application domains, including healthcare (HC), human activity recognition (HAR), industrial control systems (ICS), the Internet of Things (IoT), and other cyber–physical environments. These systems continuously generate multivariate streams that capture the temporal dynamics and interactions among multiple sensing channels. As the volume and complexity of such data increase, reliable identification of abnormal or non representative behavior becomes an important requirement for effective monitoring, diagnosis, and decision making.

## 2. Background and Early Studies

Unsupervised anomaly detection has consequently emerged as an important research area for multivariate sensor time series analysis. A central challenge is to identify generalized patterns of normal system behavior while simultaneously capturing the spatial and temporal correlations that exist among multiple sensor channels. This challenge is particularly important because noisy observations may become intertwined with the training data. Such contamination can mislead anomaly detection models and make it difficult to distinguish among normal observations, genuine abnormal events, and noise [1]. Therefore, effective analysis of multivariate sensor signals requires approaches that can characterize signal behavior at multiple levels, including temporal variations, frequency domain characteristics, inter channel relationships, and signal complexity.

Traditional anomaly detection techniques have been extensively investigated across several application areas. Chandola et al. [2] provide a comprehensive review of conventional anomaly detection methods applied to domains such as intrusion detection, fraud detection, and fault detection. These methods provide an important foundation for identifying deviations from expected patterns, although the increasing dimensionality and complexity of modern sensor systems create additional challenges. In parallel, deep learning based approaches have attracted substantial attention because of their ability to learn complex representations directly from data. Chalapathy et al. [3] provide an overview of deep learning approaches for anomaly detection, including Autoencoders, recurrent neural networks, and convolutional neural networks. These approaches demonstrate the potential of representation learning for identifying complex and nonlinear patterns in high dimensional data.

The relevance of anomaly detection is particularly evident in IoT-enabled systems, where sensors continuously monitor dynamic environments. Cook et al. [4] reviewed time series anomaly detection in IoT environments and discuss the major challenges involved in developing anomaly detection solutions for dynamic systems monitored by IoT technologies. Similarly, Pang et al. [5] present a comprehensive survey of deep anomaly detection and provide a taxonomy based on the underlying assumptions, advantages, and limitations of different deep neural network architectures. In the specific context of sensor systems, Erhan et al. [6] review state of the art anomaly detection approaches addressing applications such as cybersecurity, predictive maintenance, fault prevention, and industrial automation. Choi et al. [7] provide further background on anomaly detection for multivariate time series data and comparatively examine deep learning based anomaly detection models using several benchmark datasets.

### 2.1 Multivariate Signal Processing for Anomaly Detection

Data driven anomaly detection methods frequently rely on statistical relationships among distributed measurements. Such relationships can provide information about the normal operating state of a system and can therefore be used to identify deviations that may indicate abnormal behavior. Posenato et al. [8] compared several signal processing techniques using strain time histories and demonstrated that moving principal component analysis (MPCA) showed considerable promise for anomaly detection. In this context, changes identified through the analysis can indicate changes in structural performance or the presence of damage. Nevertheless, MPCA has limitations in detecting changes when damage is not sufficiently close to the sensor location or when the damage severity is not sufficiently high [9].

These limitations highlight the importance of considering multivariate relationships rather than relying exclusively on localized or individual sensor responses. Multivariate signal analysis enables the integrated examination of time series representing several parameters simultaneously. Such approaches are particularly valuable for large and complex structures containing numerous sensors, where individual channel analysis may not adequately capture system level behavior. One possible strategy is to cluster sensors according to their correlations or other relevant statistical measures and subsequently examine changes in the correlations within and among these clusters. Changes in these relationships can provide evidence of anomalous system behavior and may also assist in identifying the potential locations of anomalies [10, 11].

With the growing emphasis on data driven decision making, multisensor systems increasingly acquire high-dimensional data streams that capture complex temporal dynamics. Consequently, multivariate time series anomaly detection has become highly relevant across numerous application domains. A major advantage of data driven approaches is their ability to exploit relationships among multiple measurements. However, conventional supervised and semi supervised approaches generally depend on labeled data, which may be unavailable or difficult to obtain in many practical scenarios. This limitation motivates the development and application of unsupervised and data driven signal processing approaches for anomaly identification.

In this context, M. A. Belay et al. [12] investigated the effectiveness of four signal processing techniques for supporting a data-driven anomaly detection strategy based on correlations between bridge responses and temperature distributions. Their work demonstrates the potential of exploiting relationships between multiple measured variables to identify changes in system behavior [13]. Such findings further emphasize the importance of analyzing both temporal signal characteristics and inter variable relationships when developing robust multivariate anomaly detection methodologies.

### 2.2 Entropy-Based Analysis of Multivariate Sensor Signals

In addition to conventional statistical and signal processing techniques, entropy provides an important framework for characterizing the complexity and information content of time series data. Entropy is a fundamental concept originating in statistical physics, where the second law of thermodynamics indicates the difficulty of decreasing the entropy of a macroscopic system [14]. Shannon subsequently introduced entropy into information theory as a measure for quantifying the information associated with a system [15, 16]. Within this framework, an element's contribution to a system's entropy reflects its importance, with more essential elements contributing more substantially to the overall entropy [17].

The ability of entropy measures to characterize data structure and algorithmic behavior has encouraged their application in a wide range of data analysis and machine learning problems. Researchers have consequently explored additional entropy based measures and proposed variants designed for specific analytical requirements [18]. In signal analysis, entropy can be used to characterize the properties and complexity of signals and can therefore contribute to the identification of anomalous or non representative behavior [19].

Entropy is particularly relevant to time series analysis because it can act as a nonlinear indicator of signal complexity. Changes in entropy may reflect changes in the underlying dynamics of a monitored system and can therefore provide complementary information to conventional time-domain and frequency domain features. This makes entropy based analysis potentially valuable for anomaly detection in multivariate sensor environments.

Several multivariate entropy methods have been developed and applied in signal analysis. However, existing approaches can have limitations in their ability to simultaneously capture information within individual channels and relationships across different channels [20]. This limitation is important in multisensor systems because anomalous behavior may manifest not only as a change in the complexity of a single signal but also as a change in the relationships between multiple sensor channels. Consequently, combining entropy based measures with multivariate and time frequency signal analysis provides a potentially useful framework for more comprehensive characterization of sensor system behavior.

### 2.3 Anomaly Detection in Complex Industrial and Cyber–Physical Systems

Anomaly detection is a critical interdisciplinary research area concerned with identifying deviations from standard or expected operational patterns in complex systems [21, 22]. The ability to recognize such deviations is essential for monitoring system health, identifying faults, preventing failures, and supporting reliable operation across a variety of industrial and cyber physical environments.

The importance of anomaly detection extends across several interconnected domains of industrial systems engineering. These include Cyber Physical Systems (CPSs) [23], Industrial Control Systems (ICSs) [24], Intrusion Detection Systems (IDSs) [25], and the Internet of Things (IoT) [26]. In each of these environments, sensor observations provide information about system states and operational conditions. Detecting deviations from established patterns can therefore help identify faults, abnormal operating conditions, security related events, and other system level anomalies.

The increasing availability of multisensor data provides opportunities for developing more comprehensive anomaly detection strategies. Rather than examining individual measurements independently, multivariate approaches can exploit temporal dependencies, cross channel relationships, and changes in signal complexity. Time frequency analysis can provide information about changes occurring at different temporal and spectral scales, while entropy based measures can characterize changes in the complexity and information content of signals. Integrating these complementary perspectives can provide a broader characterization of system behavior than any single analytical approach.

### 2.4 Research Motivation and Analytical Perspective

The literature indicates that anomaly detection in multivariate sensor systems remains challenging because normal behavior can be highly dynamic, measurements may be noisy, and anomalies may appear at different temporal and spatial scales. Conventional approaches can provide useful statistical and structural information, but their effectiveness may depend on the nature, location, and severity of an abnormal event [8, 9]. At the same time, supervised and semi-supervised approaches can be constrained by the availability of labeled observations. These challenges reinforce the importance of unsupervised and data-driven methodologies capable of learning meaningful characteristics of complex sensor signals without requiring extensive labeled datasets.

A comprehensive signal processing framework can therefore benefit from combining complementary representations of multivariate sensor data. Time domain analysis can characterize temporal behavior, frequency domain analysis can reveal spectral patterns, and time frequency methods can identify changes occurring across different scales. Entropy based analysis can further quantify signal complexity and provide information about changes in the underlying dynamics. At the multivariate level, relationships among sensor channels can be incorporated to identify changes that may not be apparent from individual signals.

Accordingly, the combined use of signal processing, time frequency, multivariate, entropy based, and anomaly-detection approaches provides a coherent basis for investigating complex sensor systems. Such an integrated perspective is particularly relevant to applications involving HC, HAR, ICS, CPS, IDS, and IoT environments, where large scale multisensor observations must be interpreted to distinguish normal system behavior from noise and meaningful anomalies.

Overall, the reviewed literature establishes the importance of moving beyond isolated sensor analysis toward integrated approaches capable of capturing temporal dynamics, spectral characteristics, cross channel relationships, and nonlinear signal complexity. Taken together, these perspectives provide a strong analytical foundation for comprehensive anomaly detection in multivariate sensor time series data.

## 3. Research Problem and Research Design

### 3.1 Research Problem

Modern Micro Electromechanical Systems (MEMS) sensors are widely employed in embedded systems, robotics, Internet of Things (IoT) devices, and cyber physical systems because of their low cost, compact size, and ability to provide continuous multidimensional measurements. However, the raw output of low cost MEMS accelerometers may contain bias, thermal drift, quantization effects, electrical interference, mechanical disturbances, and other forms of measurement variability. Consequently, even measurements obtained under nominally static conditions may exhibit substantial temporal, spectral, distributional, and statistical complexity rather than behaving as simple Gaussian noise. The reviewed literature similarly emphasizes the need to characterize multivariate sensor behavior across multiple analytical dimensions when abnormal observations are difficult to distinguish from normal variability and noise [1–9]. The present study addresses this problem using a static triaxial accelerometer dataset acquired from an MPU-6050 MEMS sensor together with a derived percentage error channel. The sensor was maintained in a stationary configuration, with the Y-axis oriented downward along the gravitational direction. Despite the absence of intentional dynamic motion, the recorded signals exhibit substantial differences in amplitude variability, distri- butional shape, temporal persistence, spectral organization, and signal quality.

Sensors 1, 2, and 4 show pronounced skewness and excess kurtosis, whereas Sensor 3 exhibits substantially greater temporal persistence. These heterogeneous characteristics indicate that the sensor channels cannot be adequately described through a single statistical representation.

The problem is further complicated by the absence of predefined anomaly labels. Conventional supervised classification is therefore not directly applicable as the primary anomaly detection strategy. In addition, each analytical representation captures only a particular aspect of sensor behavior. Time domain statistics describe amplitude and distributional characteristics but cannot identify the frequency or temporal scale of a disturbance. Frequency domain analysis identifies global spectral characteristics but provides limited information about when spectral events occur. Time frequency analysis addresses temporal localization but does not directly quantify inter channel relationships or signal complexity. Similarly, correlation analysis describes cross channel linear dependence but does not characterize temporal persistence, spectral structure, or anomalous behavior.

Accordingly, the research problem is formulated as follows:

How can heterogeneous temporal, spectral, time frequency, complexity, signal quality, and multivariate characteristics of static MEMS sensor signals be jointly characterized to identify abnormal observations and changes in sensor behavior without relying on predefined anomaly labels?

The study addresses this problem through an integrated analytical framework that combines descriptive signal statistics, Fourier analysis, power spectral density, time frequency analysis, temporal dependence analysis, inter sensor correlation, signal to noise ratio, entropy, change point detection, and unsupervised anomaly detection. The objective is not simply to identify isolated outliers, but to establish a multidimensional representation of sensor behavior capable of distinguishing persistent nominal structure, transient deviations, and changes in statistical regime.

### 3.2 Research Questions

The research is guided by the following questions:

**RQ1.** What are the principal distributional and amplitude characteristics of the four MEMS sensor channels under static operating conditions?

**RQ2.** What dominant spectral components and frequency dependent characteristics are present in the sensor signals?

**RQ3.** How does sensor behavior vary across time and frequency, and are transient or localized signal energy changes present?

**RQ4.** To what extent do the sensor channels exhibit temporal persistence and cross channel linear dependence?

**RQ5.** How do the channels differ in estimated signal quality and in amplitude and spectral complexity characteristics?

**RQ6.** Can change point analysis identify temporal transitions in the statistical behavior of the sensor signals?

**RQ7.** Can unsupervised methods, specifically Isolation Forest and PCA reconstruction based anomaly detection, identify unusual multivariate sensor windows in the absence of labeled anomalies?

**RQ8.** How can the complementary findings from the different analytical domains be integrated into a coherent characterization of static MEMS sensor behavior?

Together, these questions provide a progression from basic signal characterization to multivariate interpretation and unsupervised detection.

### 3.3 Research Design

The study adopts a quantitative, observational, exploratory, and unsupervised research design based on secondary analysis of a continuously recorded multivariate sensor dataset. It is quantitative because the investigation relies on numerical observations and mathematically derived signal features. It is observational because the analysis examines measurements collected under a static experimental condition without imposing a dynamic stimulus or intervention during the analyzed recording. It is exploratory because the objective is to identify previously unspecified temporal, spectral, distributional, complexity related, and anomalous structures. Finally, it is unsupervised because the dataset contains no predefined anomaly labels and the anomaly-detection stage therefore identifies deviations from the dominant signal structure rather than learning from known abnormal classes.

The research design follows a sequential but interconnected analytical strategy. First, the raw multivariate signals are examined using time domain statistics to establish their amplitude distributions and variability. Frequency-domain analysis is then performed using the Fast Fourier Transform (FFT) and Welch power spectral density (PSD) to identify dominant spectral components and energy distributions. Time frequency analysis subsequently examines whether the spectral characteristics remain stable over the observation period. Temporal dependence is assessed through autocorrelation and rolling RMS analysis. Pearson correlation is used to characterize linear relationships among the four channels. Signal to noise ratio and entropy measures provide complementary information regarding signal quality and complexity. Finally, change point analysis and unsupervised anomaly detection are used to identify statistical regime transitions and unusual multivariate windows.

The analytical sequence is deliberately designed so that each stage addresses a limitation of the preceding representation. In this way, the study does not treat the individual analyses as independent endpoints. Instead, they form complementary components of a multidomain characterization framework.

### 3.4 Research Architecture

The research architecture comprises five conceptual layers:

Layer 1 — Sensor Data and Experimental Context:
Raw triaxial accelerometer observations and the derived percentage-error channel are obtained from the static MPU-6050 acquisition.

Layer 2 — Signal Characterization:
Time-domain statistics, distributions, FFT, PSD, spectrograms, wavelets, autocorrelation, and rolling RMS are used to characterize amplitude, spectral, time frequency, and temporal behavior.

Layer 3 — Multivariate and Complexity Characterization:
Pearson correlation, SNR, Shannon entropy, and spectral entropy are used to examine cross-channel relationships, signal quality, and signal complexity.

Layer 4 — Deviation and Regime Detection:
Change-point analysis identifies temporal transitions, while Isolation Forest and PCA reconstruction identify unusual multivariate windows.

Layer 5 — Integrated Interpretation:
The outputs from the preceding layers are jointly interpreted to distinguish persistent nominal structure, transient deviations, and sustained statistical changes.

The resulting architecture can be represented as:

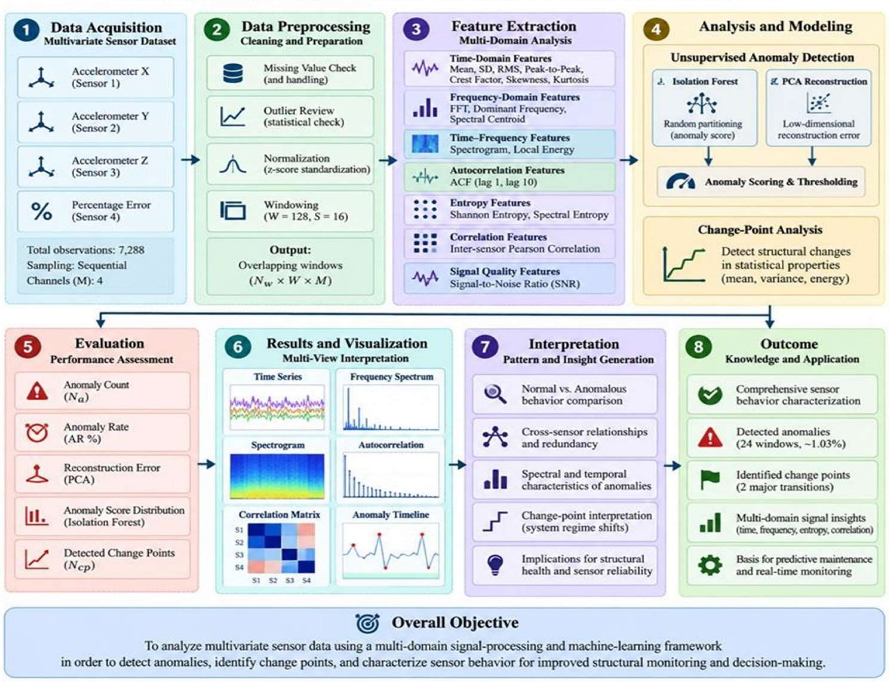


Figure 1. Research architecture of the integrated multidomain signal-processing and unsupervised anomaly detection framework for static MEMS sensor data

The architecture reflects the central methodological premise that no single representation is sufficient to characterize the observed sensor behavior. Time domain analysis captures amplitude and distributional structure; FFT and PSD identify global spectral characteristics; spectrogram and wavelet analysis reveal localized time frequency behavior; autocorrelation and rolling RMS characterize temporal persistence and local signal energy; correlation captures inter channel relationships; SNR and entropy characterize signal quality and complexity; and change point and anomaly detection methods identify deviations and temporal regime transitions.

# 4. Dataset and Methodology

## 4.1 Dataset and Experimental Configuration

The study uses the Accelerometer Sensor Data MPU6050 dataset containing measurements acquired from an MPU-6050 six degree of freedom MEMS motion tracking device. Although the hardware integrates a three-axis accelerometer and a three axis gyroscope, only the accelerometer measurements are analyzed in the pres- ent study. Data acquisition was performed using a Raspberry Pi Pico microcontroller based on the RP2040 dual core ARM Cortex M0+ processor. Communication between the sensor and microcontroller was established through the Inter Integrated Circuit ($I^2C$) protocol, while data were transferred through a serial USB connection.

The sensor was maintained in a stationary configuration throughout the acquisition period. The Y-axis was oriented downward toward the ground and therefore aligned with the gravitational acceleration vector. Under ideal static conditions, the $X$- and $Z$-axis measurements are expected to represent the axes orthogonal to gravity, whereas the $Y$-axis represents the gravity loaded component. This configuration provides a controlled setting for examining sensor bias, noise, temporal variability, spectral structure, and anomalous observations without intentional dynamic motion.

Four channels were analyzed: the raw $X$-axis acceleration measurement, the raw $Y$-axis acceleration measurement, the raw $Z$-axis acceleration measurement, and a calculated percentage-magnitude error associated with the gravity affected $Y$-axis. The analysis portion of the manuscript reports 37,459 observations across the four channels. However, another dataset description portion of the source document reports 37,458 observations. This discrepancy must be checked against the original sensor_data.csv file and corrected consistently throughout the manuscript before publication.

## 4.2 Data Structure and Preprocessing

The sensor observations were treated as a multivariate time series,

[ = [x_{t, j}]_{N M}, ]

where ($N$) denotes the number of observations, ($M$=4) denotes the number of analyzed channels, ($t$) represents the temporal index, ($j$) represents the sensor channel index, and ($x\_\{t, j\}$) denotes the observation from channel ($j$) at time ($t$).

The four channels were analyzed both independently and jointly. Independent analysis was used to identify channel specific statistical, temporal, and spectral characteristics, whereas joint analysis was used to examine inter channel relationships and multivariate deviations.

The source dataset does not provide a documented physical sampling frequency. Therefore, the temporal reference was retained as the original sample index, and all frequency domain quantities were expressed in cycles per sample rather than Hertz. Similarly, periods were expressed in samples. This avoids introducing an unsupported conversion between normalized frequency and physical frequency.

No dynamic stimulus was introduced during the analysis, and the purpose of the preprocessing stage was therefore characterization rather than motion segmentation. For anomaly detection, the multivariate observations were subsequently segmented into overlapping windows of 128 samples using a stride of 16 samples. The resulting windows served as the basic units for multivariate anomaly analysis.

## 4.3 Time-Domain Analysis

The first stage of analysis characterized the amplitude and distributional properties of each sensor channel. The calculated descriptive measures included the mean, standard deviation, variance, minimum, first quartile, median, third quartile, maximum, root mean square (RMS) amplitude, peak to peak amplitude, skewness, excess kurtosis, and crest factor. These measures provide complementary information about central tendency, variability, extreme observations, distributional asymmetry, tail behavior, and signal amplitude.

For a signal ($x_i$) containing ($N$) observations, the RMS amplitude was calculated as

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$$

The peak to peak amplitude was obtained as

$$A_{P\text{-}P} = max(x) - min(x)$$

while the crest factor was calculated as

$$\text{CF} = \frac{max(|x|)}{\text{RMS}}$$

Skewness and excess kurtosis were used to determine whether the observed distributions departed substantially from symmetric Gaussian behavior. This stage was particularly important because the dataset contains pronounced heavy tailed behavior in several channels. The extremely high skewness and kurtosis values reported for Sensors 1, 2, and 4 provide evidence that conventional Gaussian assumptions may not adequately describe the measurements.

### 4.4 FFT and PSD

Frequency domain analysis was performed using the Fast Fourier Transform (FFT) to identify dominant spectral components and periodicities in the four signals. For a discrete signal ($x[n]$), the Fourier transform can be represented as

$$X[k] = \sum_{n=0}^{N-1} x\,[n]\, e^{-j\,2\pi\, kn/N}$$

The magnitude spectrum was examined to identify dominant frequency components. The principal frequency-domain descriptors included dominant frequency, dominant period, spectral centroid, spectral entropy, and total spectral power.

The dominant period was determined from

$$\text{T} = \frac{1}{f_d}$$

where ($f_d$) represents the dominant frequency in cycles/sample. The spectral centroid was used to characterize the center of mass of the frequency distribution, while spectral entropy was used to quantify the degree to which spectral energy was distributed across frequency components.

Because the sampling frequency is unavailable, these quantities were explicitly retained in normalized sample-based units. The analysis identified a particularly distinctive component in Sensor 2 at approximately 0.0472 cycles/sample, corresponding to a period of approximately 21.18 samples.

To complement the FFT analysis, Welch’s method was employed to estimate the power spectral density (PSD). PSD analysis provides a more robust representation of signal power distribution across frequency than a single global FFT spectrum because it averages spectral estimates obtained from overlapping segments.

The PSD analysis was used to determine whether sensor energy was concentrated at low frequencies or distributed over a broader frequency range. This analysis revealed that most channels contained substantial low frequency energy, while Sensor 2 exhibited a more distinct spectral component around 0.047 cycles/sample.

### 4.5 Time–Frequency Analysis

Short time Fourier transform (STFT)-based spectrogram analysis was employed to investigate temporal changes in spectral energy. Unlike global FFT analysis, which provides an overall spectral representation, the spectrogram permits frequency content to be examined as a function of time.

For a windowed signal, the short time Fourier transform can be represented as

$$\mathrm{STFT}(\tau, f) = \sum_{n} x[n]\, w[n-\tau]\, e^{-j2\pi fn}$$

where ($w$[$n$-]) is a localized analysis window centered at time ().

The resulting spectrograms were examined for persistent frequency components, localized energy concentrations, and transient spectral events. This analysis is particularly relevant because the signals contain both relatively stable low frequency behavior and isolated high energy events. The results indicate temporal variation in spectral energy, especially for Sensor 3, while Sensor 4 contains localized high energy events corresponding to transient spikes.

Wavelet decomposition was employed to complement Fourier based methods by characterizing signal energy across multiple temporal scales. Whereas FFT analysis is particularly effective for identifying global periodic components, wavelet analysis is better suited to localized and transient signal structures.

Wavelet domain energy was examined across scales to determine how signal activity was distributed between persistent low frequency behavior and localized high frequency or transient components. This was particularly appropriate for the present dataset because Sensors 1, 2, and 4 exhibit strong non Gaussian behavior and Sensor 4 contains isolated high amplitude spikes.

### 4.6 Temporal Dependence

Temporal dependence was investigated using autocorrelation functions (ACFs). The ACF measures the relationship between a signal and lagged versions of itself and therefore provides an indication of temporal memory or persistence.

The autocorrelation at lag ($k$) was considered in the general form

$$\rho(k) = \frac{\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2}$$

ACF values at lag 1 and lag 10 were extracted to provide comparable short and longer lag measures of temporal dependence. Sensor 3 demonstrated exceptionally strong persistence, with ACF values of approximately 0.94 at lag 1 and 0.93 at lag 10, whereas Sensors 1 and 2 showed substantially weaker temporal dependence.

Rolling RMS analysis was also employed to identify temporal variations in signal energy. For a moving window of length (W), dependence, correlation, signal quality, entropy, change point, and anomaly detection information, respectively.

The purpose of this integration is to distinguish three broad forms of sensor behavior:

1. persistent nominal signal structure;

2. transient anomalous deviations; and

3. sustained changes in statistical regime.

The resulting framework combines amplitude behavior, spectral structure, temporal persistence, localized time frequency activity, multivariate relationships, signal quality, complexity, anomalous behavior, and regime transitions within one analytical workflow.

## 5. Results

### 5.1 Overall Sensor-Signal Characteristics

The four channels exhibit clearly different temporal behaviors. Sensors 1 and 2 maintain comparatively stable baseline levels but contain occasional abrupt deviations. Sensor 3 exhibits substantially larger amplitude fluctuations across the observation period, whereas Sensor 4 generally remains within a narrow range but contains isolated high amplitude spikes. These differences demonstrate that the four channels possess heterogeneous signal characteristics even though the sensor was maintained in a static configuration.

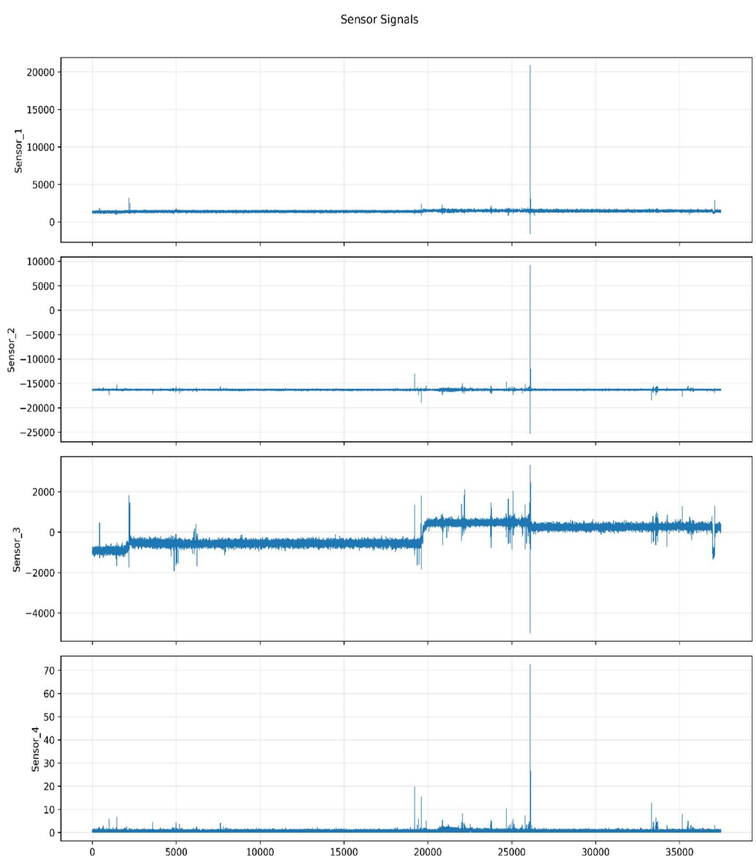


Figure 2. Sensor Signals

$$RMS_t = \sqrt{\frac{1}{W}\sum_{i=0}^{W-1} x_{t-i}^2}$$

Elevated rolling RMS values were interpreted as periods of increased signal amplitude or transient disturbance and were subsequently considered relevant for anomaly characterization.

### 4.7 Multivariate Correlation

Pearson correlation analysis was performed to quantify linear relationships among the four sensor channels. The Pearson correlation coefficient between variables ($X$) and ($Y$) was calculated as

$$r_{XY} = \frac{\sum (x_i - x)(y_i - y)}{\sigma_x \sigma_y}$$

The resulting correlation matrix was visualized using a heatmap to facilitate comparison of inter sensor relationships. The analysis identified moderate positive relationships between Sensors 1 and 2 (r) and Sensors 2 and 4 (r), while Sensor 3 showed weaker relationships with the remaining channels.

This analysis establishes whether individual channels contain redundant information or whether sensor fusion can provide complementary information for system level characterization and anomaly detection.

### 4.8 Signal-to-Noise Ratio

Signal-to-noise ratio (SNR) was estimated to quantify the relative strength of structured signal components compared with noise or unwanted variability. The SNR was expressed in decibels as

$$SNR_{db} = 10 \log_{10} \frac{P_{\text{signal}}}{P_{\text{noise}}}$$

The resulting values were compared across channels to assess relative signal quality. Sensor 3 exhibited the highest estimated SNR (13.81 dB), followed by Sensor 2 (8.41 dB), Sensor 4 (5.53 dB), and Sensor 1 (4.19 dB).

SNR was not interpreted independently; instead, it was considered jointly with temporal persistence, spectral concentration, entropy, and anomaly characteristics.

### 4.9 Entropy

Two complementary entropy measures were calculated: Shannon entropy and spectral entropy. Shannon entropy was used to characterize the diversity and uncertainty of the amplitude distribution, whereas spectral entropy was used to characterize the distribution of energy across frequency components.
For a discrete probability distribution ($p\_i$), Shannon entropy was defined as

$$H(X) = -\sum_i p_i \log_2(p_i)$$

Spectral entropy was calculated from the normalized spectral-power distribution, thereby measuring the degree of spectral concentration or dispersion.

The analysis showed that Sensor 3 had the highest Shannon entropy (3.237), indicating greater amplitude-distribution diversity, but comparatively low spectral entropy. In contrast, Sensors 1, 2, and 4 exhibited relatively high spectral entropy. This dual analysis permits amplitude complexity to be distinguished from frequency domain complexity.

### 4.10 Change-Point Detection

Change-point analysis was used to identify temporal locations at which the statistical characteristics of the signals changed sufficiently to indicate a potential transition between different regimes.

A rolling-statistical procedure was applied to the multivariate sensor observations. Conceptually, if ($y\_t$) represents a derived signal statistic, a change point () separates two regimes:

[ $y\_t$=]

The objective was to identify sustained changes in statistical behavior rather than isolated observations.

Change point detection was deliberately distinguished from anomaly detection. An anomaly represents a localized observation or window that deviates from the dominant signal structure, whereas a change point represents a transition toward a different statistical regime. The source analysis reports two detected change points.

### 4.11 Anomaly Detection

Because the dataset contains no predefined anomaly labels, anomaly detection was formulated as an unsupervised learning problem. The four channel time series was segmented into overlapping windows of 128 samples with a stride of 16 samples. Window level multivariate characteristics were then used to identify observations that departed from the dominant signal structure.

Two complementary approaches were used

#### Isolation Forest

Isolation Forest identifies observations that can be isolated relatively easily through recursive partitioning of the feature space. Observations requiring shorter isolation paths are considered more unusual than observations requiring longer paths. This approach is appropriate for unlabeled anomaly detection because it does not require predefined normal and abnormal classes.

#### PCA Reconstruction-Based Detection

Principal Component Analysis (PCA) was used as a linear reconstruction based approach. The multivariate feature representation was projected into a lower dimensional principal component space and subsequently reconstructed. For an input feature vector (_m), its reconstruction can be represented as

[ _m = _K_K^{T}_m, ]

where (_K) contains the retained principal-component loading vectors.

The reconstruction error was calculated as

[ E_m= | _m-_m |_2^2. ]

Windows exhibiting comparatively large reconstruction errors were considered anomalous.

Importantly, this method is a PCA-based linear reconstruction surrogate, not a nonlinear deep-learning autoencoder. Accordingly, its results are interpreted as linear reconstruction-based anomaly detection rather than as evidence from a deep neural autoencoder.

### 4.12 Integrated Analytical Framework

The individual methods were integrated into a multidomain framework rather than treated as independent analytical endpoints. The framework is represented conceptually by

[ = [ *{TD}*, {FD}, *{TF}*, {TEMP}, *{COR}*, {SNR}, *{ENT}*, {CP}, _{AD}], ]

where the component groups represent time-domain, frequency-domain, time–frequency, temporal-

Time-series representation of the four sensor channels across the complete observation period. The figure illustrates differences in baseline level, amplitude variability, transient excursions, and temporal behavior among the sensor channels.

The temporal differences observed in Figure 1 provide the basis for the subsequent distributional analysis. In particular, the isolated extreme excursions visible in Sensors 1, 2, and 4 suggest that the underlying amplitude distributions may deviate substantially from Gaussian behavior.

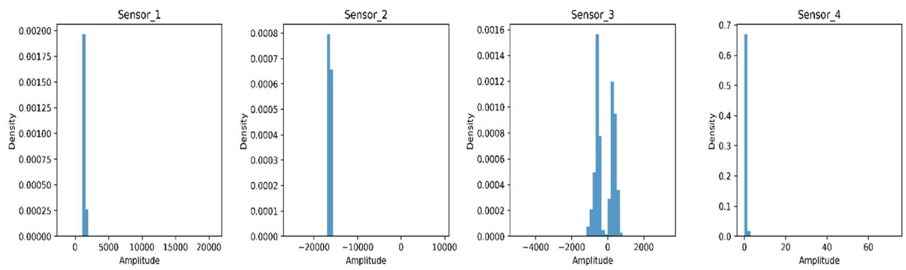


Figure 3. Signal Distributions

Amplitude distributions of the four sensor channels showing pronounced asymmetry and heavy tailed behavior in Sensors 1, 2, and 4 and comparatively broader, more symmetric behavior in Sensor 3.

The distributional findings are important for the subsequent anomaly-detection analysis. Because the sensor signals are strongly non-Gaussian, simple parametric thresholding based on normality assumptions may incorrectly classify legitimate extreme observations as anomalies or fail to distinguish transient events from the dominant signal structure.

The distributional analysis further confirms this heterogeneity. Sensors 1, 2, and 4 exhibit pronounced asymmetry and heavy-tailed behavior, while Sensor 3 presents a comparatively broader and more symmetric distribution. The distributional observations are consistent with the higher order statistics reported in Table 1.

## 5.2 Time-Domain Statistical Results

| **Sensor** | **Mean** | **SD** | **Variance** | **RMS** | **Peak-to-Peak** | **Skewness** | **Excess Kurtosis** | **Crest Factor** |
|---|---|---|---|---|---|---|---|---|
| Sensor 1 | 1431.77 | 142.71 | 20,364.92 | 1438.86 | 22,416 | 66.99 | 9115.77 | 14.48 |
| Sensor 2 | -16301.33 | 187.82 | 35,275.42 | 16302.41 | 34,428 | 82.41 | 10804.07 | 1.55 |
| Sensor 3 | -154.58 | 493.23 | 243,278.46 | 516.88 | 8,316 | 0.03 | “0.84 | 9.65 |
| Sensor 4 | 0.59 | 0.68 | 0.46 | 0.90 | 72.56 | 49.92 | 4587.57 | 80.93 |

Table 1. Time-domain statistical characteristics of the four sensor channels

The time-domain results show substantial differences among the channels. Sensor 3 has the largest standard deviation (493.23) and variance (243,278.46), indicating the greatest overall amplitude variability. Sensors 1 and 2 have extremely large positive skewness values of 66.99 and 82.41, respectively, while Sensor 4 has a skewness of 49.92. Excess kurtosis is also extremely high for Sensors 1, 2, and 4, demonstrating strongly heavy tailed distributions and the presence of rare extreme observations.

Sensor 4 is particularly notable because its average amplitude is very small while its crest factor is approximately 80.93. This combination indicates that the channel is normally characterized by low amplitude but contains isolated excursions that are very large relative to its typical signal level. Sensor 2 has the largest RMS value, primarily because its signal has a large absolute baseline magnitude.

Collectively, these results establish that the sensor measurements cannot be adequately represented using a simple Gaussian noise assumption and provide the statistical basis for subsequent spectral, temporal, complexity, and anomaly analyses.

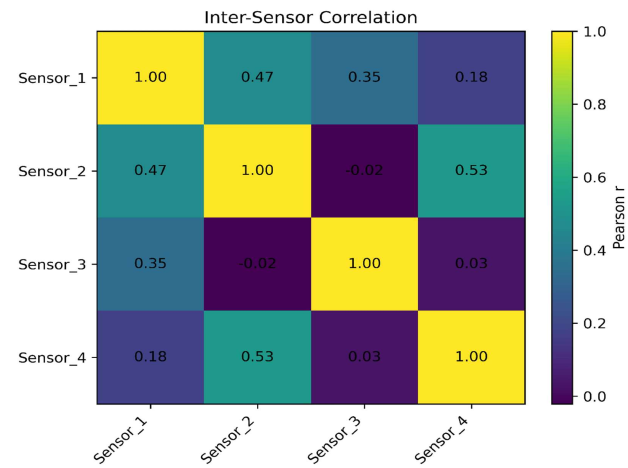


Figure 4. Inter-Sensor Correlation Heatmap

Pearson correlation coefficients among the four sensor channels. Moderate relationships are observed primarily between Sensors 1–2 and Sensors 2–4, while several channel pairs exhibit weak linear dependence.

The absence of uniformly strong correlations indicates that the channels contain substantial complementary information rather than simply duplicating the same signal behavior. This is particularly relevant to anomaly detection because unusual behavior may be detectable only when multiple sensor dimensions are considered jointly.

## 5.3 Frequency-Domain Results

| **Sensor** | **Dominant Frequency (cycles/ sample)** | **Dominant Period (samples)** | **Spectral Centroid (cycles/ sample)** | **Spectral Entropy*** | **Total Spectral Power** |
|---|---|---|---|---|---|
| Sensor 1 | (5.34^{-5}) | 18,729.5 | 0.2559 | 13.2374 | (1.29962^{13}) |
| Sensor 2 | 0.047225 | 21.175 | 0.2070 | 13.7841 | (2.47275^{13}) |
| Sensor 3 | (5.34^{-5}) | 18,729.5 | 0.0412 | 5.8037 | (6.40214^{13}) |
| Sensor 4 | (5.34^{-5}) | 18,729.5 | 0.1682 | 13.4317 | (3.20827^{8}) |

Table 2. Principal frequency-domain characteristics

*Spectral entropy values in this table correspond to the FFT-based frequency-domain feature extraction reported in the source analysis.

The frequency domain results reveal distinct spectral signatures among the four channels. Sensor 2 is the most distinctive, with a dominant component at approximately 0.0472 cycles/sample, corresponding to a period of approximately 21.18 samples. In contrast, Sensors 1, 3, and 4 are dominated by an extremely low frequency component of approximately (5.34^{-5}) cycles/sample, corresponding to approximately 18,729.5 samples per cycle.

Sensor 3 has the lowest spectral centroid and substantially lower FFT-based spectral entropy than the other channels, indicating relatively concentrated spectral organization despite its comparatively large time domain variability. The total spectral power also differs substantially among channels, demonstrating that signal amplitude and spectral energy are not distributed uniformly across the sensor system.

Because the acquisition sampling frequency is unavailable, these frequency components are interpreted in cycles/sample rather than converted into Hertz.

## 5.4 Power Spectral Density

Figure 5 presents the Welch power spectral density (PSD) estimates for the four sensor signals. Most of the signal energy is concentrated at relatively low frequencies, indicating slowly varying behavior. Sensor 2, however, exhibits a substantially more distinct spectral peak near 0.047 cycles/sample, corresponding to a periodicity of approximately 21 samples.

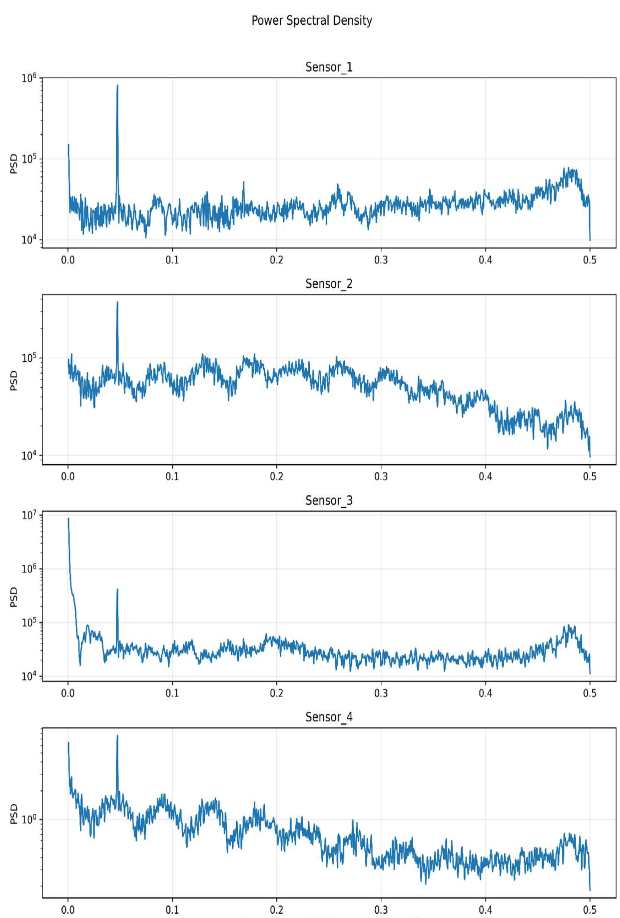


Figure 5. Power Spectral Density (PSD)

Welch PSD estimates for the four sensor channels showing predominantly low-frequency energy and a distinctive spectral component in Sensor 2.

The concentration of energy at low frequencies in Sensors 1, 3, and 4 suggests slowly evolving signal components or long-term trends. Sensor 2 is different because its distinct peak indicates a more recognizable short period component. Since the dataset does not provide a physical sampling frequency, all reported frequencies are expressed in cycles/sample rather than Hertz.

The PSD results establish that the channels differ not only in amplitude statistics but also in their underlying spectral organization. This observation motivates the use of time frequency analysis to determine whether the spectral characteristics remain constant throughout the observation period.

### 5.5 Time–Frequency Results

Figure 5 presents spectrograms obtained using short-time Fourier analysis. The representation demonstrates that the frequency content is not completely stationary over the observation period. Sensor 3 exhibits noticeable temporal variation in spectral energy, while Sensors 1 and 2 show relatively stable low frequency behavior with intermittent localized concentrations of energy. Sensor 4 contains isolated high energy events that correspond to the transient spikes observed in the raw time domain signal.

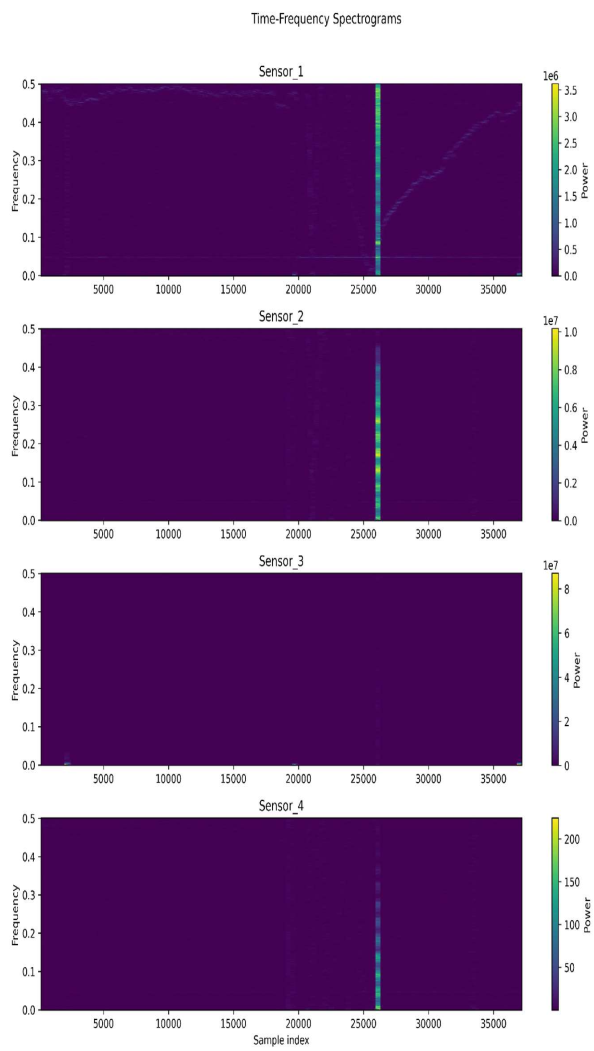


Figure 6. Time–Frequency Spectrograms

Short time Fourier representations showing the temporal evolution of spectral energy across the four sensor channels.

The spectrogram results demonstrate why global frequency-domain analysis alone is insufficient. A global PSD can identify the frequencies at which energy is concentrated but cannot determine whether a component persists throughout the recording or appears only during a localized interval. The time frequency representation therefore provides an important bridge between the global spectral findings and the subsequent temporal dependence and anomaly analyses.

The spectrogram analysis shows that the signal energy is not completely stationary over time. Sensor 3 exhibits noticeable temporal variation in spectral energy, while Sensors 1 and 2 show relatively stable low-frequency behavior with intermittent localized energy concentrations. Sensor 4 exhibits localized high-energy events corresponding to the isolated spikes observed in the raw signal.

The wavelet representation provides complementary scale dependent information. The sensor channels differ in the distribution of signal energy across temporal scales, and the localized representation is particularly relevant for the transient events and heavy tailed behavior identified in the time domain analysis. Thus, the time frequency results demonstrate that the observed signal complexity includes both persistent components and localized events that may not be fully represented by global spectral measures.

## 5.6 Temporal Dependence and Local Signal Energy

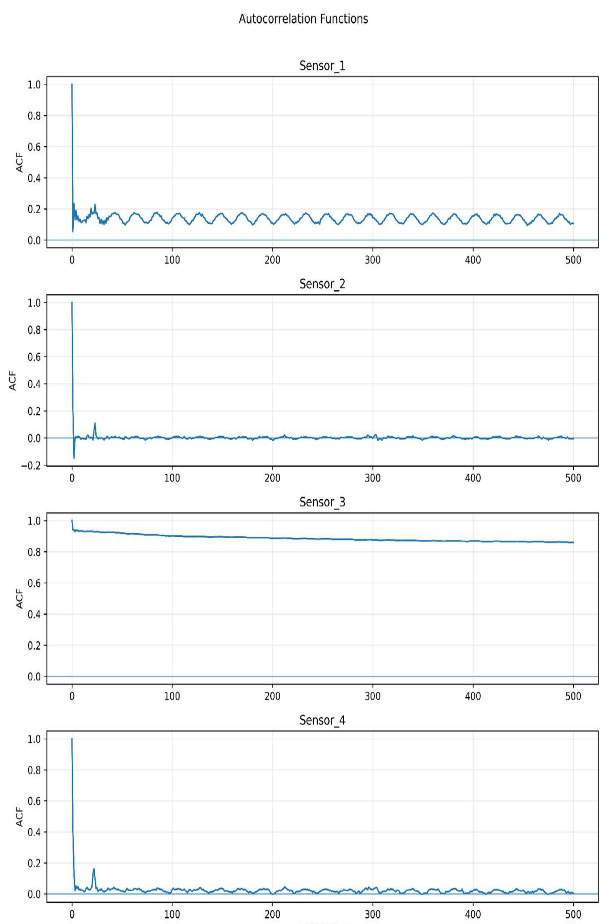


Figure 7. Autocorrelation Functions

Figure 7 presents the autocorrelation functions of the four sensor channels. Sensor 3 exhibits exceptionally strong temporal persistence, with autocorrelation values of approximately 0.94 at lag 1 and 0.93 at lag 10. Sensors 1 and 2 show substantially weaker temporal dependence, while Sensor 4 exhibits moderate short-term correlation that decreases rapidly with increasing lag.

Autocorrelation functions showing markedly stronger temporal persistence in Sensor 3 compared with the other sensor channels.

| **Sensor** | **ACF at Lag 1** | **ACF at Lag 10** |
|---|---|---|
| Sensor 1 | 0.0541 | 0.1154 |
| Sensor 2 | 0.2123 | -0.0003 |
| Sensor 3 | 0.9401 | 0.9339 |
| Sensor 4 | 0.4004 | 0.0252 |

Table 3. Autocorrelation coefficients at selected lags

Sensor 3 exhibits exceptionally strong temporal persistence, with autocorrelation remaining above 0.93 at both lag 1 and lag 10. Sensor 2 shows moderate short term dependence at lag 1 but essentially no dependence at lag 10. Sensor 1 exhibits weak short term dependence, while Sensor 4 exhibits moderate short term dependence that decreases substantially by lag 10.

These results demonstrate pronounced heterogeneity in temporal dynamics. In particular, Sensor 3 evolves substantially more smoothly over time than the other channels. The present findings establish strong temporal persistence but do not, by themselves, establish formal long memory behavior.

The rolling RMS analysis further indicates substantial differences in local signal energy. Sensor 3 consistently exhibits higher rolling energy because of its larger amplitude fluctuations, whereas Sensor 4 generally exhibits low energy except during isolated spike events.

### 5.7 Rolling RMS and Signal Energy

Figure 7 presents the rolling RMS signal energy. Sensor 3 consistently exhibits the greatest local signal energy, reflecting its substantially larger amplitude fluctuations. Sensor 4 generally remains at a low energy level, but exhibits pronounced local increases associated with isolated spike events.

Moving-window RMS estimates illustrating temporal variations in local signal energy and highlighting transient high energy events.

The rolling RMS results complement the autocorrelation analysis. Whereas autocorrelation describes temporal dependence between observations at different lags, rolling RMS captures local changes in signal intensity. Their combination provides a more complete representation of temporal sensor behavior.

### 5.8 FFT Peak Analysis

Figure 9 presents the dominant spectral components identified through Fast Fourier Transform (FFT) analysis. Sensor 2 is the most distinctive channel, exhibiting a dominant frequency of approximately 0.0472 cycles/sample, corresponding to a period of approximately 21.18 samples.

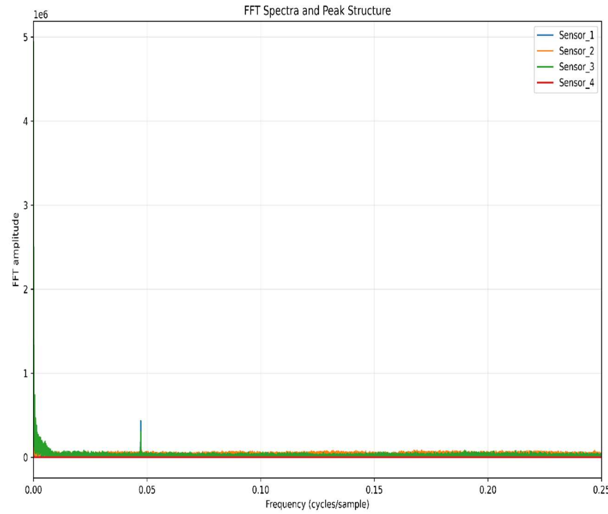


Figure 9. FFT Peak Analysis

Dominant FFT spectral components for the four sensor channels. Sensor 2 exhibits a distinct approximately 21-sample periodic component, whereas Sensors 1, 3, and 4 are dominated by extremely low-frequency components.

Sensors 1, 3, and 4 are dominated by a very low-frequency component of approximately (5.34\times10^{-5}) cycles/sample, corresponding to approximately 18,729.5 samples per cycle. This suggests slow variations, long-term trends, or low frequency system dynamics rather than a clearly defined short period oscillation.

The FFT results are consistent with the PSD findings and provide a more explicit identification of the dominant spectral component in Sensor 2. As the dataset does not provide a physical sampling frequency, these frequencies cannot be converted reliably into Hertz.

## 5.9 Multivariate Correlation Results

| | **Sensor 1** | **Sensor 2** | **Sensor 3** | **Sensor 4** |
|---|---|---|---|---|
| **Sensor 1** | 1.000 | 0.471 | 0.346 | 0.177 |
| Sensor 2 | 0.471 | 1.000 | “0.021 | 0.530 |
| Sensor 3 | 0.346 | “0.021 | 1.000 | 0.032 |
| Sensor 4 | 0.177 | 0.530 | 0.032 | 1.000 |

Table 4. Pearson correlation matrix

The strongest linear association occurs between Sensor 2 and Sensor 4 (($r$)), followed by Sensor 1 and Sensor 2 (($r$)). Sensor 1 and Sensor 3 exhibit a weaker to moderate positive association (($r$)), whereas Sensor 2 and Sensor 3 are essentially uncorrelated linearly (($r$)). Sensor 3 and Sensor 4 also show negligible linear association.

The absence of uniformly strong correlations indicates limited linear redundancy among the channels. However, the results should be interpreted as evidence of limited linear dependence rather than statistical independence. The heterogeneous correlations support the use of multivariate analysis because individual channels contain different temporal and statistical characteristics.

## 5.10 Wavelet Energy Analysis

Figure 10 provides a time-scale representation of signal energy obtained through wavelet decomposition. The results demonstrate that signal energy is distributed differently across temporal scales among the four sensor channels.

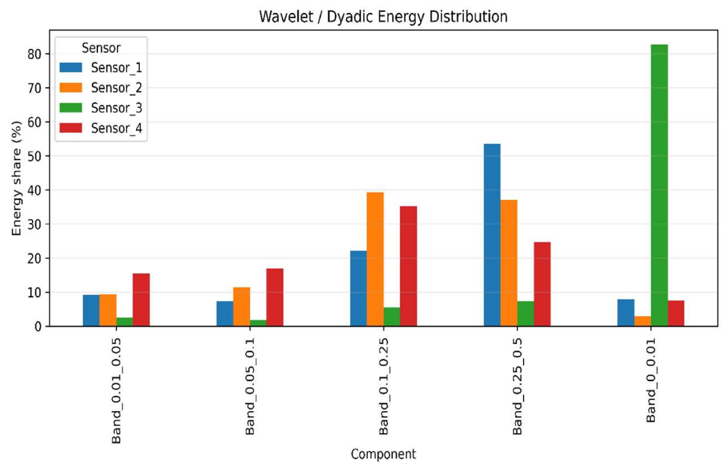


Figure 10. Wavelet Energy

Wavelet-domain representation of signal energy across temporal scales for the four sensor channels.

The wavelet representation complements the FFT and PSD because it can identify localized signal behavior across different scales. This is particularly important for the present dataset because Sensors 1, 2, and 4 exhibit strongly non Gaussian distributions and Sensor 4 contains isolated high amplitude events.

Consequently, wavelet domain energy provides an additional mechanism for distinguishing persistent low-frequency behavior from localized transient events. The result reinforces the observation from Figure 5 that the sensor signals contain both persistent and transient components.

## 5.11 Signal-to-Noise Ratio

Figure 11 compares the estimated SNR across the four sensor channels. Sensor 3 achieves the highest estimated SNR of 13.81 dB, followed by Sensor 2 at 8.41 dB, Sensor 4 at 5.53 dB, and Sensor 1 at 4.19 dB.

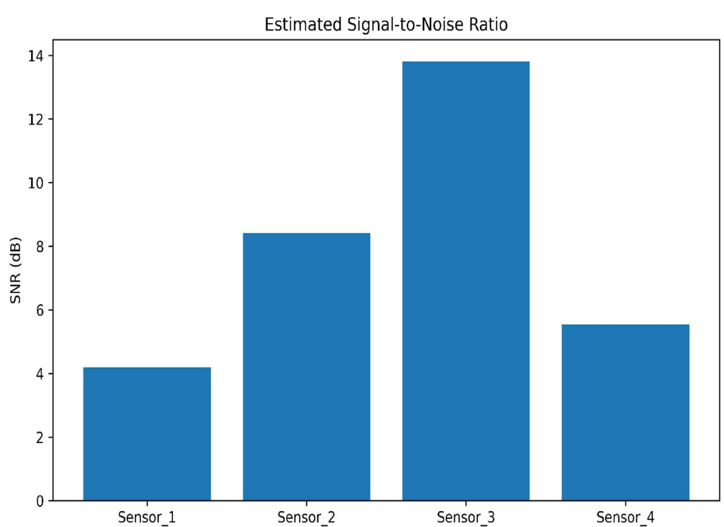


Figure 11. Signal-to-Noise Ratio. Comparative estimated SNR across the four sensor channels

The substantially higher SNR of Sensor 3 indicates that its signal structure is comparatively more distinguishable from noise. This finding agrees with its strong temporal persistence and relatively concentrated spectral organization. Conversely, Sensor 1 has the lowest estimated SNR, indicating a comparatively larger contribution from noise or unwanted variability.

The SNR results should nevertheless be interpreted comparatively rather than as absolute hardware specifications because the dataset does not provide an independently measured noise reference.

| **Sensor** | **Estimated SNR** |
|---|---|
| Sensor 1 | 4.19 dB |
| Sensor 2 | 8.41 dB |
| Sensor 3 | 13.81 dB |
| Sensor 4 | 5.53 dB |

Table 5. Estimated signal-to-noise ratio

Sensor 3 has the highest estimated SNR at 13.81 dB, followed by Sensor 2 at 8.41 dB, Sensor 4 at 5.53 dB, and Sensor 1 at 4.19 dB. Thus, Sensor 3 exhibits the strongest estimated separation between signal and noise, whereas Sensor 1 has the lowest estimated signal quality.

The SNR results complement the preceding temporal and spectral analyses. Sensor 3 combines high SNR with strong temporal persistence and concentrated spectral characteristics, whereas Sensor 1 combines low SNR with highly skewed and heavy-tailed amplitude behavior. Nevertheless, these SNR values should be interpreted comparatively because an independently measured noise reference is not available in the source dataset.

## 5.12 Entropy-Based Complexity Analysis

Figure 12 compares Shannon entropy and spectral entropy for the four sensor channels. Sensor 3 is particularly distinctive because it simultaneously exhibits the highest Shannon entropy (3.237) and the lowest spectral entropy (8.060).

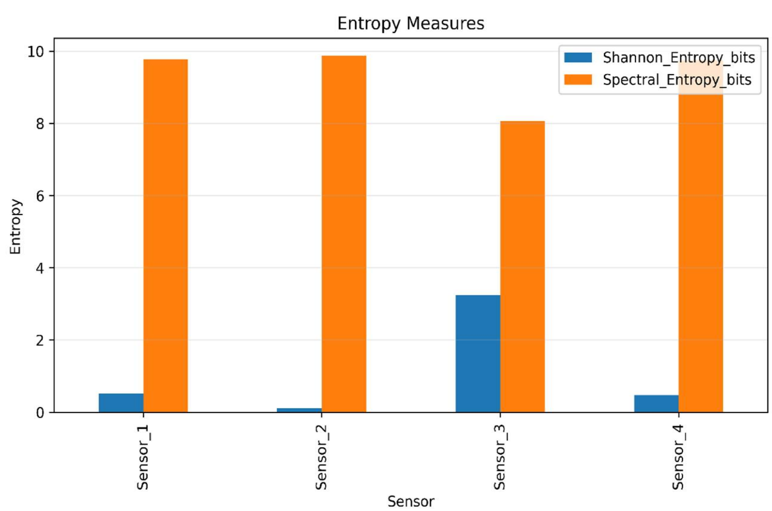


Figure 12. Entropy Analysis

Comparison of amplitude-based Shannon entropy and spectral entropy across the four sensor channels.

| Sensor | Shannon Entropy | Spectral Entropy |
|---|---|---|
| Sensor 1 | 0.515 | 9.768 |
| Sensor 2 | 0.113 | 9.865 |
| Sensor 3 | 3.237 | 8.060 |
| Sensor 4 | 0.474 | 9.745 |

Table 6. Shannon and spectral entropy

Sensor 3 exhibits the highest Shannon entropy (3.237), indicating substantially greater diversity in its amplitude distribution. This result is consistent with its relatively large standard deviation and broad amplitude range. Sensor 2 has the lowest Shannon entropy (0.113), indicating a comparatively concentrated amplitude distribution despite its large absolute signal magnitude.

The spectral entropy results provide a different perspective. Sensors 1, 2, and 4 exhibit comparatively high spectral entropy values, whereas Sensor 3 has a substantially lower spectral entropy of 8.060. Thus, Sensor 3 simultaneously exhibits high amplitude distribution complexity and relatively concentrated spectral organization. This distinction demonstrates that amplitude variability and frequency-domain complexity are not equivalent properties.

The combination of entropy measures therefore reinforces the multidomain character of the sensor signals. A signal may exhibit considerable amplitude variability while maintaining relatively organized spectral behavior, as observed for Sensor 3.

### 5.13 Change-Point Results

Figure 13 presents the temporal locations of the change points identified using the rolling-statistical change-point procedure. Two change points were detected across the sensor signals.

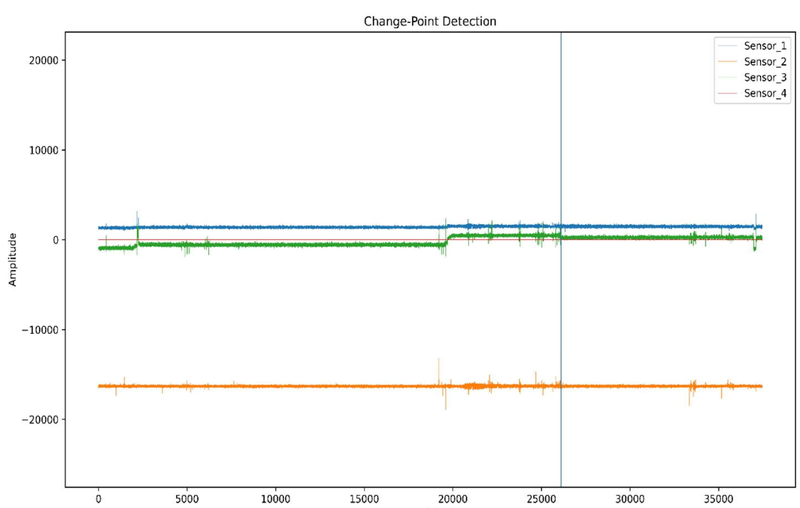


Figure 13. Change Points

Temporal locations of the two change points detected using the rolling-statistical procedure.

The detected change points indicate locations at which the statistical characteristics of the observed signals changed sufficiently to represent transitions between different regimes. These results complement the anomaly-detection findings but represent a different type of deviation.

An anomaly corresponds to an observation or window that is unusual relative to the dominant signal behavior, whereas a change point represents a transition in the statistical regime. Consequently, an isolated transient spike may be detected as an anomaly without representing a sustained regime change, while a change point may correspond to a persistent change in mean, variance, or another rolling statistical characteristic

The rolling statistical change point analysis identified two change points across the sensor signals. These locations represent temporal positions at which the statistical characteristics of the signals changed sufficiently to be identified as transitions between different regimes.

The change point results complement the anomaly detection analysis. An anomalous window represents a localized departure from the dominant multivariate structure, whereas a change point represents a transition in the statistical behavior of the signal. Consequently, the two methods provide different forms of temporal information: isolated transient deviations may appear as anomalies without producing a sustained regime change, whereas a change point indicates a more persistent alteration in statistical behavior.

The exact sample locations and affected channels of the two change points should be reported in the final manuscript alongside the corresponding figure so that the results can be independently interpreted.

### 5.14 Unsupervised Anomaly-Detection Results

Figure 12 presents the anomaly-detection results obtained from 128-sample windows using a stride of 16 samples. Both Isolation Forest and PCA reconstruction based detection identified 24 anomalous windows, corresponding to an anomaly rate of 1.028%.

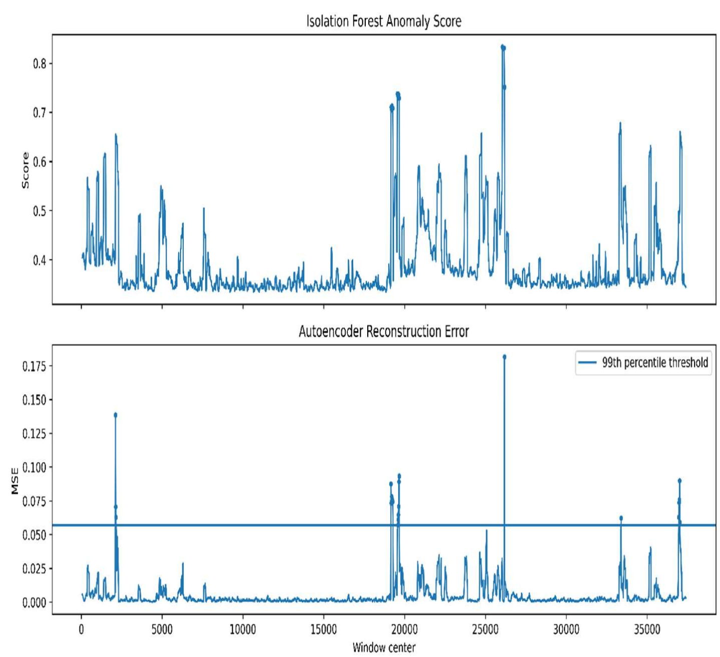


Figure 14. Unsupervised Anomaly Detection

Anomalous windows identified using Isolation Forest and PCA reconstruction-based anomaly detection.

Anomaly detection was performed using 128 sample windows with a stride of 16 samples. Both Isolation Forest and PCA reconstruction based detection identified 24 anomalous windows, corresponding to an anomaly rate of approximately 1.028%.

| Detection method | Anomalous windows | Anomaly rate |
|---|---|---|
| Isolation Forest | 24 | 1.028% |
| PCA reconstruction / linear surrogate | 24 | 1.028% |

Table 7. Unsupervised anomaly-detection results

The agreement between the two approaches indicates consistency in the number of windows identified as unusual despite the different principles underlying the methods. Isolation Forest identifies observations that are relatively easy to isolate in the feature space, whereas PCA reconstruction identifies windows that are poorly represented by the dominant low dimensional structure. Their agreement therefore provides complementary evidence that the identified windows represent unusual multivariate signal patterns rather than being exclusively dependent on one detection mechanism.

However, because the dataset contains no independently labeled anomalies, agreement between the methods cannot be treated as definitive validation of the physical origin of the detected events. The anomalous windows should therefore be interpreted as statistically unusual multivariate observations requiring further physical or experimental validation.

### 5.15 Integrated Results

The integrated findings demonstrate that the static MEMS sensor data contain multiple layers of signal structure. Time domain analysis identifies strong non Gaussian and heavy tailed behavior in Sensors 1, 2, and 4, while Sensor 3 exhibits the greatest amplitude variability. Frequency domain analysis identifies a distinctive approximately 21-sample periodic component in Sensor 2 and comparatively concentrated spectral organization in Sensor 3. Time frequency analysis demonstrates localized and temporally varying energy patterns that are not fully captured by global spectral summaries. Temporal analysis further shows exceptionally strong persistence in Sensor 3 and much weaker persistence in the remaining channels.

The multivariate correlation results indicate that the channels are not strongly linearly redundant, while the SNR results demonstrate substantial differences in estimated signal quality. Entropy analysis further distinguishes amplitude distribution complexity from spectral complexity, with Sensor 3 exhibiting the notable combination of the highest Shannon entropy and comparatively low spectral entropy. Finally, the detection of 24 anomalous windows and two change points demonstrates that the dataset contains both localized deviations and temporal transitions in statistical behavior.

The findings collectively support the proposed multidomain analytical framework. Rather than relying on a single measure of sensor quality or abnormality, the results demonstrate the value of jointly considering amplitude distributions, temporal dependence, spectral structure, time frequency localization, cross channel relationships, signal quality, complexity, anomalies, and statistical regime transitions.

## 6. Discussion

### 6.1 Interpretation of Heterogeneous Static Sensor Behavior (RQ1)

The time domain results demonstrate that a nominally static MEMS accelerometer does not produce signals that conform to simple Gaussian noise assumptions. Sensors 1, 2, and 4 exhibit extreme positive skewness (66.99, 82.41, and 49.92, respectively) and excess kurtosis values exceeding 4,500, confirming the presence of rare but extreme amplitude excursions superimposed on otherwise bounded baseline behavior. Sensor 3, by contrast, displays near-zero skewness (0.03) and negative excess kurtosis (-0.84), indicating a comparatively symmetric and platykurtic amplitude distribution. These findings are consistent with the observations reported by Posenato et al. [8] and Kromanis [9], who emphasized that raw sensor responses in structural and industrial monitoring contexts often contain measurement artifacts, transient disturbances, and non stationary components that cannot be captured through a single parametric model.

The crest factor of Sensor 4 (approximately 80.93) is particularly informative. Because the mean amplitude of this channel is extremely small (0.59) while isolated excursions reach magnitudes on the order of 72.56 peak-to peak, the channel alternates between near-quiescent behavior and abrupt transient events. This pattern is characteristic of quantization-induced spikes, electrical interference, or intermittent mechanical coupling in low cost MEMS devices, as discussed in the broader sensor characterization literature. The implication for anomaly detection is significant: thresholding strategies based on global mean and standard deviation would either miss these events or generate excessive false alarms if applied without accounting for the underlying heavy tailed structure.

### 6.2 Spectral Organization and Channel Distinctiveness (RQ2)

The frequency domain analysis reveals that the four channels possess fundamentally different spectral signatures. Sensor 2 exhibits a clearly defined periodic component at approximately 0.0472 cycles/sample (period $\approx$ 21.18 samples), whereas Sensors 1, 3, and 4 are dominated by extremely low frequency components near $5.34 \times 10^{-5}$ cycles/sample (period $\approx$ 18,729.5 samples). This distinction indicates that Sensor 2 contains a recognizable short period oscillatory structure that is absent from the remaining channels.

The origin of this periodic component cannot be definitively attributed from the dataset alone; however, plausible explanations include residual electronic oscillation in the sensor readout circuitry, a periodic interference source in the acquisition environment, or a coupling effect specific to the *Y*-axis orientation. The Welch PSD analysis confirms that the energy associated with this component is spectrally concentrated rather than broadly distributed, supporting the interpretation of a quasi periodic phenomenon rather than broadband noise.

Sensor 3 exhibits the lowest spectral centroid (0.0412 cycles/sample) and the lowest FFT-based spectral entropy (5.8037), indicating that its spectral energy is concentrated within a narrow low frequency band. This finding is noteworthy because Sensor 3 simultaneously possesses the largest time domain variance (243,-278.46). The coexistence of high amplitude variability and concentrated spectral organization suggests that Sensor 3 undergoes large but slowly varying fluctuations rather than rapid oscillatory behavior. This interpretation is reinforced by the exceptionally high autocorrelation values observed for this channel.

Because the dataset does not provide a documented physical sampling frequency, all spectral quantities are necessarily expressed in cycles/sample. This limitation prevents conversion to Hertz and therefore restricts direct comparison with physically specified frequency bands. Nevertheless, the relative spectral comparisons among channels remain valid and informative.

### 6.3 Time–Frequency Non-Stationarity and Transient Events (RQ3)

The spectrogram and wavelet analyses demonstrate that the spectral characteristics of the sensor signals are not strictly stationary over the observation period. Sensor 3 exhibits noticeable temporal variation in spectral energy, while Sensor 4 contains localized high-energy events corresponding to the transient spikes identified in the time-domain analysis. Sensors 1 and 2 show relatively stable low-frequency behavior punctuated by intermittent localized energy concentrations.

These findings underscore a fundamental limitation of global spectral analysis. A single FFT or averaged PSD representation cannot determine whether a spectral component persists throughout the recording or appears only during a localized interval. The time frequency results therefore provide essential complementary information: they confirm that the dominant low frequency components identified by FFT are broadly persistent, while simultaneously revealing transient events that would be averaged out in a global spectral estimate.

The wavelet domain energy analysis further supports this interpretation by demonstrating that signal energy is distributed differently across temporal scales. The localized high frequency energy observed in Sensor 4 during spike events is consistent with the heavy tailed amplitude distribution and high crest factor reported in the time domain analysis. This multi scale perspective aligns with the broader signal processing literature emphasizing that transient and persistent signal components require different analytical representations for adequate characterization.

### 6.4 Temporal Persistence and Cross-Channel Dependence (RQ4)

The autocorrelation analysis reveals a striking asymmetry in temporal dynamics. Sensor 3 maintains autocorrelation values above 0.93 at both lag 1 and lag 10, indicating that successive observations are highly predictable from their immediate predecessors. This strong persistence is consistent with the low spectral centroid and concentrated spectral organization of this channel: a signal that evolves slowly in time will naturally exhibit high short lag autocorrelation.

Sensors 1 and 2, by contrast, show weak to moderate short term dependence that decays rapidly. Sensor 2 exhibits moderate autocorrelation at lag 1 (0.2123) but essentially zero correlation at lag 10 (“0.0003), suggesting that its distinctive 21-sample periodic component does not translate into strong monotonic temporal persistence. Sensor 4 shows moderate short term correlation (0.4004 at lag 1) that decreases substantially by lag 10 (0.0252), consistent with a signal that is normally quiescent but subject to intermittent disruptions.

It is important to note that strong autocorrelation does not, by itself, establish formal long memory or fractional integration behavior. The present findings demonstrate strong short to moderate lag persistence in Sensor 3 but do not extend to the formal statistical tests (e.g., Hurst exponent estimation or rescaled range analysis) that would be required to confirm long range dependence.

The Pearson correlation matrix indicates that the four channels are not strongly linearly redundant. The strongest associations occur between Sensors 2 and 4 ($r \approx 0.530$) and Sensors 1 and 2 ($r \approx 0.471$), while Sensor 3 shows weak or negligible linear correlation with the remaining channels. This result has practical implications for sensor fusion and anomaly detection: because the channels carry partially complementary information, multivariate analysis can detect anomalies that would not be apparent from any single channel in isolation. However, the moderate correlation values also indicate that linear dependence alone does not fully characterize the inter channel relationships, and nonlinear or information theoretic measures may provide additional insight in future work.

### 6.5 Signal Quality and Complexity Characterization (RQ5)

The estimated SNR values reveal a clear ordering of signal quality: Sensor 3 (13.81 dB) > Sensor 2 (8.41 dB) > Sensor 4 (5.53 dB) > Sensor 1 (4.19 dB). The high SNR of Sensor 3 is consistent with its strong temporal persistence and concentrated spectral organization, suggesting that its signal structure is more readily distinguishable from noise. Conversely, Sensor 1 combines the lowest SNR with extreme skewness and kurtosis, indicating that its measurements contain a comparatively larger contribution from noise or unwanted variability relative to its structured signal component.

The entropy analysis provides a complementary and particularly informative perspective. Sensor 3 simultaneously exhibits the highest Shannon entropy (3.237) and the lowest spectral entropy (8.060). This combination indicates that while the amplitude distribution of Sensor 3 is diverse and broadly distributed (high Shannon entropy), its spectral energy is concentrated within a relatively narrow frequency band (low spectral entropy). In other words, the signal varies substantially in amplitude but does so in a spectrally organized manner. Sensors 1, 2, and 4 show the opposite pattern: relatively low Shannon entropy but higher spectral entropy, indicating more concentrated amplitude distributions but more dispersed spectral energy.

This distinction between amplitude domain complexity and frequency domain complexity is a key contribution of the multidomain framework. It demonstrates that a single entropy measure cannot adequately characterize signal behavior: a signal may be complex in one domain while remaining relatively organized in another. The dual entropy analysis therefore provides a more nuanced characterization than either measure alone.

### 6.6 Statistical Regime Transitions (RQ6)

The rolling-statistical change-point analysis identified two temporal locations at which the statistical characteristics of the sensor signals changed sufficiently to represent transitions between different regimes. These change points are conceptually distinct from the anomalous windows identified by the unsupervised detection methods. A change point indicates a sustained shift in the underlying statistical properties (e.g., mean, variance, or higher order structure) of the signal, whereas an anomaly represents a localized departure from the dominant pattern.

The detection of two change points in a nominally static acquisition suggests that the sensor system underwent subtle but persistent changes in operating conditions during the recording period. Possible explanations include gradual thermal drift in the MEMS element, slow changes in the electrical environment of the acquisition hardware, or minor mechanical settling of the sensor mounting. Regardless of the physical origin, the presence of regime transitions indicates that the assumption of strict stationarity over the entire observation period is not fully justified, and that anomaly detection methods should account for potential non stationarity in the background signal.

### 6.7 Unsupervised Anomaly Detection and Methodological Agreement (RQ7)

Both Isolation Forest and PCA reconstruction based detection identified 24 anomalous windows out of the total segmented windows, corresponding to an anomaly rate of approximately 1.028%. The agreement between these two methodologically distinct approaches provides meaningful evidence that the identified windows represent genuinely unusual multivariate patterns rather than artifacts of a single detection mechanism.

Isolation Forest operates by recursively partitioning the feature space and identifying observations that require fewer splits to isolate, making it sensitive to observations that are sparse or distant in the multivariate feature representation. PCA reconstruction, by contrast, identifies observations that are poorly represented by the dominant low dimensional linear structure of the data. The fact that both methods converge on the same number of anomalous windows despite their different underlying principles strengthens confidence in the detection results.

However, several important caveats apply. First, the dataset contains no independently labeled anomalies, and therefore the detected windows cannot be validated against a known ground truth. The agreement between methods provides internal consistency but does not constitute external validation. Second, the PCA-based approach used in this study is a linear reconstruction surrogate and should not be interpreted as equivalent to a nonlinear deep learning autoencoder. The linear assumption may limit the ability of this method to capture complex nonlinear anomaly structures. Third, the windowing parameters (128-sample window, 16-sample stride) influence the granularity of detection, and different parameter choices may yield different anomaly counts.

The anomalous windows should therefore be interpreted as statistically unusual multivariate observations that warrant further physical or experimental investigation rather than as confirmed fault events.

### 6.8 Integrated Multidomain Characterization (RQ8)

The central methodological contribution of this study is the demonstration that no single analytical representation is sufficient to characterize the observed sensor behavior. Each domain captures a different aspect of the signal:

- Time domain statistics reveal amplitude distributions, asymmetry, and tail behavior.

- FFT and PSD identify global spectral components and energy distributions.

- Spectrogram and wavelet analysis expose temporal variation in spectral content and localized transient events.

- Autocorrelation and rolling RMS characterize temporal persistence and local signal energy.

- Pearson correlation quantifies linear inter channel relationships.

- SNR provides a comparative measure of signal quality.

- Entropy measures distinguish amplitude domain complexity from frequency domain complexity.

- Change-point detection identifies sustained statistical transitions.

- Unsupervised anomaly detection identifies localized multivariate deviations.

The integration of these perspectives enables the distinction among three broad forms of sensor behavior: persistent nominal structure, transient anomalous deviations, and sustained changes in statistical regime. This multidomain framework addresses a gap identified in the reviewed literature, where individual analytical methods are often applied in isolation without explicit consideration of their complementary strengths and limitations.

### 6.9 Comparison with Existing Literature

The present findings are consistent with several themes in the existing anomaly detection literature. The observation that low cost MEMS sensors produce non-Gaussian, heavy tailed measurements under static conditions aligns with the challenges discussed by Cook et al. [4] and Erhan et al. [6] regarding the difficulty of distinguishing noise from genuine anomalies in IoT and industrial sensor environments.

The finding that multivariate relationships provide complementary information beyond individual-channel analysis supports the arguments of Gul et al. [10] and Shi et al. [11] regarding the value of cross-channel correlation analysis for system-level anomaly detection.

The use of entropy as a complexity measure for sensor signals is consistent with the broader application of entropy based methods in signal analysis and anomaly detection, as discussed by Aguayo Tapia [19] and Ding [20]. The present study extends this perspective by demonstrating that amplitude domain and frequency-domain entropy capture different aspects of signal behavior, and that their joint consideration provides a more complete characterization than either measure alone.

The unsupervised detection approach addresses the practical limitation highlighted by Belay et al. [12] and Kromanis [13], namely that labeled anomaly data are often unavailable in real sensor monitoring scenarios. The use of Isolation Forest and PCA reconstruction as complementary unsupervised methods provides a practical and computationally efficient alternative to supervised or deep-learning approaches that require extensive labeled training data.

### 6.10 Limitations

Several limitations of the present study should be acknowledged.

First, the dataset does not provide a documented physical sampling frequency. All frequency domain quantities are therefore expressed in cycles/sample rather than Hertz, which limits direct comparison with physically specified frequency bands and with results reported in other studies that use physical frequency units.

Second, the dataset contains no predefined anomaly labels. While this motivates the unsupervised approach adopted in the study, it also means that the detected anomalous windows cannot be validated against a known ground truth. The agreement between Isolation Forest and PCA reconstruction provides internal consistency but does not constitute external validation.

Third, the PCA reconstruction method used in this study is a linear technique. It may not adequately capture nonlinear anomaly structures that could be identified by nonlinear methods such as deep autoencoders or kernel-based approaches. The results should therefore be interpreted as evidence from a linear reconstruction surrogate rather than from a comprehensive nonlinear detection framework.

Fourth, the study analyzes a single static acquisition from a single MPU-6050 sensor. The findings may not generalize to other sensor types, dynamic operating conditions, or different MEMS hardware platforms. Replication across multiple sensors and operating conditions would be necessary to establish the broader applicability of the proposed framework.

Fifth, the windowing parameters used for anomaly detection (128-sample window, 16-sample stride) were selected based on practical considerations. A systematic sensitivity analysis of these parameters was not performed, and different choices may influence the number and location of detected anomalous windows.

Sixth, a minor data inconsistency exists in the source documentation, where one section reports 37,459 observations and another reports 37,458 observations. This discrepancy should be resolved against the original data file before publication.

Finally, the change-point analysis reports two detected change points, but the exact sample locations and affected channels are not fully specified in the current draft. These details should be included in the final manuscript to enable independent interpretation and replication.

### 6.11 Practical Implications

The findings have several practical implications for sensor monitoring and anomaly detection in embedded and IoT systems. The demonstration that static MEMS sensor signals exhibit substantial non Gaussian behavior, temporal heterogeneity, and transient events suggests that simple thresholding or Gaussian based monitoring strategies are likely to be inadequate for reliable anomaly detection in such systems. Instead, multidomain characterization frameworks that jointly consider amplitude, spectral, temporal, and complexity features are more likely to provide robust and informative monitoring.

The moderate inter channel correlations observed in this study suggest that multivariate analysis can provide additional detection sensitivity beyond single-channel monitoring, but that the channels are not sufficiently redundant to permit simple substitution of one channel for another. This finding supports the use of sensor fusion strategies that exploit complementary information across channels.

The entropy based distinction between amplitude complexity and spectral complexity provides a practical diagnostic tool. A sensor channel that exhibits high amplitude variability but concentrated spectral organization (as observed for Sensor 3) may be experiencing large but structured variations, whereas a channel with low amplitude variability but dispersed spectral energy may be subject to broadband noise or interference. This distinction can inform maintenance and calibration decisions in practical sensor networks.

## 7. Conclusion

This study presented a comprehensive multidomain signal processing and anomaly detection framework for static multivariate MEMS sensor data. The framework integrates time domain statistics, Fourier and power spectral density analysis, time frequency representations, temporal dependence analysis, inter sensor correlation, signal to noise ratio estimation, entropy based complexity measures, change point detection, and unsupervised anomaly detection into a unified analytical workflow. The framework was applied to a four channel static accelerometer dataset acquired from an MPU-6050 MEMS sensor, comprising three raw acceleration axes and a derived percentage error channel.

The principal findings of the study are as follows. First, the four sensor channels exhibit strongly heterogeneous statistical characteristics under nominally static conditions. Sensors 1, 2, and 4 display pronounced skewness, extreme excess kurtosis, and heavy-tailed amplitude distributions, while Sensor 3 exhibits a comparatively symmetric distribution with substantially greater amplitude variability. These findings demonstrate that static MEMS sensor measurements cannot be adequately described by simple Gaussian noise models.

Second, frequency domain analysis identified a distinctive periodic component in Sensor 2 at approximately 0.0472 cycles/sample (period ≈ 21.18 samples), while Sensors 1, 3, and 4 are dominated by extremely low-frequency components. Spectral entropy and spectral centroid analysis further distinguished the concentrated spectral organization of Sensor 3 from the more dispersed spectral characteristics of the remaining channels.

Third, time frequency analysis revealed that the spectral properties of the signals are not strictly stationary. Localized energy concentrations and transient spectral events were identified, particularly in Sensors 3 and 4, demonstrating that global spectral summaries alone are insufficient for complete signal characterization.

Fourth, temporal dependence analysis showed that Sensor 3 exhibits exceptionally strong autocorrelation (above 0.93 at lags 1 and 10), while the remaining channels display substantially weaker persistence. Rolling RMS analysis further confirmed differences in local signal energy and identified transient high energy events.

Fifth, multivariate correlation analysis indicated moderate linear relationships between selected channel pairs but no strong redundancy, supporting the value of multivariate analysis for capturing complementary information across sensor dimensions.

Sixth, SNR and entropy analyses provided complementary characterizations of signal quality and complexity. Sensor 3 exhibited the highest estimated SNR (13.81 dB) and the highest Shannon entropy (3.237) combined with the lowest spectral entropy (8.060), demonstrating that amplitude domain and frequency domain complexity are distinct and independently informative properties.

Seventh, change point analysis identified two temporal transitions in the statistical behavior of the sensor signals, indicating that the assumption of strict stationarity over the entire observation period is not fully justified.

Eighth, unsupervised anomaly detection using both Isolation Forest and PCA reconstruction based methods identified 24 anomalous windows (anomaly rate ≈ 1.028%). The agreement between these two methodologically distinct approaches provides evidence that the detected windows represent genuinely unusual multivariate patterns, although the absence of labeled ground truth precludes definitive external validation.

Taken together, these findings demonstrate that comprehensive characterization of multivariate sensor signals requires the joint consideration of multiple analytical domains. No single representation whether time domain, frequency domain, time frequency, temporal, multivariate, or complexity based is individually sufficient to capture the full range of signal behavior observed in static MEMS sensor data. The proposed multidomain frame work addresses this limitation by integrating complementary analytical perspectives into a coherent workflow capable of distinguishing persistent nominal structure, transient anomalous deviations, and sustained statisticalregime changes.

The study also highlights important practical considerations for sensor monitoring in embedded, IoT, and cyber physical systems. The non Gaussian, heavy tailed, and temporally heterogeneous nature of low cost MEMS sensor measurements implies that simple parametric thresholding strategies are inadequate for reliable anomaly detection. Instead, multidomain and multivariate approaches that exploit complementary signal representations are better suited to the complexity of real sensor data.

Future work should address several directions. First, the framework should be validated on datasets with labeled anomalies to enable quantitative assessment of detection performance. Second, the analytical methodology should be extended to dynamic operating conditions and to a broader range of sensor types and hardware platforms. Third, nonlinear anomaly detection methods, including deep autoencoders and kernel based approaches, should be incorporated to complement the linear PCA reconstruction used in the present study. Fourth, the effect of windowing parameters on anomaly detection performance should be systematically investigated. Fifth, information theoretic inter channel measures such as mutual information and transfer entropy should be explored to capture nonlinear dependencies that are not reflected in Pearson correlation. Finally, the framework should be evaluated in real time or near real time monitoring scenarios to assess its computational feasibility for embedded and edge deployed sensor systems.

Overall, the study contributes a structured, reproducible, and multidomain analytical framework for the comprehensive characterization and unsupervised anomaly detection of multivariate MEMS sensor signals. By integrating complementary signal processing, time frequency, entropy based, and machine learning methods within a single coherent architecture, the framework provides a foundation for more reliable and informative sensor monitoring in applications spanning healthcare, human activity recognition, industrial control, and the Internet of Things.

## References

[1] Zhang, Y., Chen, Y., Wang, J., Pan, Z. (2023). Unsupervised deep anomaly detection for multi-sensor time-series signals. *IEEE Transactions on Knowledge and Data Engineering, 35*(2), 2118–2132. https://doi.org/10.1109/TKDE.2021.3102110.

[2] Chandola, V., Banerjee, A., Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys, 41*(3), Article 15. https://doi.org/10.1145/1541880.1541882.

[3] Chalapathy, R., Chawla, S. (2019). Deep learning for anomaly detection: A survey. *arXiv*. https://arxiv.org/abs/1901.03407.

[4] Cook, A. A., Misirli, G., Fan, Z. (2020). Anomaly detection for IoT time-series data: A survey. *IEEE Internet of Things Journal, 7*(7), 6481–6494. https://doi.org/10.1109/JIOT.2020.2988181.

[5] Pang, G., Shen, C., Cao, L., Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys, 54*(2), Article 38. https://doi.org/10.1145/3439950.

[6] Erhan, L., Ndubuaku, M., Di Mauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., Liotta, A. (2021). Smart anomaly detection in sensor systems: A multi-perspective review. *Information Fusion, 67*, 64–79. https://doi.org/10.1016/j.inffus.2020.10.001.

[7] Choi, K., Yi, J., Park, C., Yoon, S. (2021). Deep learning for anomaly detection in time-series data: Review, analysis and guidelines. *IEEE Access, 9*, 120043–120065. https://doi.org/10.1109/ACCESS.2021.3108975.

[8] Posenato, D., Kripakaran, P., Inaudi, D., Smith, I. F. C. (2010). Methodologies for model-free data interpretation of civil engineering structures. *Computers Structures, 88*(7–8), 467–482. https://doi.org/10.1016/j.compstruc.2009.12.001.

[9] Kromanis, R., Kripakaran, P. (2013). Support vector regression for anomaly detection from measurement hist ories. *Advanced Engineering Informatics, 27*(4), 486–495. https://doi.org/10.1016/j.aei.2013.03.002.

[10] Gul, M., Catbas, F. N. (2011). Structural health monitoring and damage assessment using a novel time series analysis methodology with sensor clustering. *Journal of Sound and Vibration, 330*(6), 1196–1210. https://doi.org/10.1016/j.jsv.2010.09.024.

[11] Shi, H., Worden, K., Cross, E. J. (2018). A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring. *Mechanical Systems and Signal Processing, 103*, 381–397. https://doi.org/10.1016/j.ymssp.2017.10.013.

[12] Belay, M. A., Rasheed, A., Rossi, P. S. (2024). Multivariate time series anomaly detection via low rank and sparse decomposition. *IEEE Sensors Journal, 24*(21), 34942–34952. https://doi.org/10.1109/JSEN.2024.3467221.

[13] Kromanis, R., Kripakaran, P. (2021). Performance of signal processing techniques for anomaly detection using a temperature based measurement interpretation approach. *Journal of Civil Structural Health Monitoring, 11*, 15–34. https://doi.org/10.1007/s13349-020-00435.

[14] Liu, X., Sun, D., Wei, W. (2022). Alleviating the over smoothing of graph neural computing by a data augmentation strategy with entropy preservation. *Pattern Recognition, 132*, Article 108951. https://doi.org/10.1016/j.patcog.2022.108951.

[15] Bromiley, P., Thacker, N., Bouhova Thacker, E. (2004). Shannon entropy, Renyi entropy, and information. *Statistical Inference Series, 9*, 2–8.

[16] Wei, W., Li, H., Zhou, S., Li, B., Liu, X. (2025). Attention with system entropy for optimizing credit assignment in cooperative multi-agent reinforcement learning. *IEEE Transactions on Automation Science and Engineering, 22*, 14775–14787. https://doi.org/10.1109/TASE.2024.3491803.

[17] Wei, G., Liu, Y. (2025). A graph contrastive learning method for enhancing genome recovery in complex microbial communities. *Entropy, 27*(9), Article 921. https://doi.org/10.3390/e27090921.

[18] Sepúlveda-Fontaine, S. A., Amigó, J. M. (2024). Applications of entropy in data analysis and machine learning: A review. *Entropy, 26*(12), Article 1126. https://doi.org/10.3390/e26121126.

[19] Aguayo-Tapia, S., Avalos-Almazan, G., Rangel-Magdaleno, J. d. J. (2024). Entropy based methods for motor fault detection: A review. *Entropy, 26*(4), Article 299. https://doi.org/10.3390/e26040299.

[20] Ding, R., et al. (2025). Multivariate multiscale variations dispersion entropy and multivariate multiscale variations fuzzy dispersion entropy for time-series analysis. *IEEE Sensors Journal, 25*(8), 12676–12687. https://doi.org/10.1109/JSEN.2025.3544242.

[21] Rafique, S. H., Abdallah, A., Musa, N. S., Murugan, T. (2024). Machine learning and deep learning techniques for internet of things network anomaly detection Current research trends. *Sensors, 24*(6), Article 1968. https://doi.org/10.3390/s24061968.

[22] Li, Q., Tang, Y., Chu, L. (2024). Generative adversarial networks for prognostic and health management of industrial systems: A review. *Expert Systems with Applications, 253*, Article 124341. https://doi.org/10.1016/j.eswa.2024.124341.

[23] Rani, S., Kataria, A., Kumar, S., Karar, V. (2025). A new generation cyber-physical system: A comprehensive review from security perspective. *Computers Security, 148*, Article 104095. https://doi.org/10.1016/j.cose.2024.104095.

[24] Rawat, H., Rawat, A., Rajavat, A., Rawat, R. (2025). Literature review analysis for cyberattacks at management applications and industrial control systems. *In Information Visualization for Intelligent Systems* (p. 461–487). Wiley. https://www.wiley.com/en-us/Information+Visualization+for+Intelligent+Systems+-p-9781394305797.

[25] Rahman, M. M., Al Shakil, S., Mustakim, M. R. (2025). A survey on intrusion detection system in IoT networks. *Cyber Security and Applications, 3*, Article 100082. https://doi.org/10.1016/j.csa.2025.100082.

[26] Mishra, R., Mishra, A. (2025). Current research on Internet of Things (IoT) security protocols: A survey. *Computers Security, 151*, Article 104310. https://doi.org/10.1016/j.cose.2025.104310.