Tampere University

Maria Islam Smrity

CHALLENGES AND METHODS IN BIOMEDICAL SIGNAL QUALITY EVALUATION

A Multimodal Approach

Faculty of Medicine and Health Technology (MET)
Master's Thesis
Examiners: Mark van Gils,
Saana Seppälä,
Alpo Värri
November 2025

ABSTRACT

Maria Islam Smrity : Challenges and Methods in Biomedical Signal Quality Evaluation: A Multimodal Approach
Master's Thesis
Tampere University
Master's Degree Programme in Biomedical Science and Engineering
November 2025

Evaluating Bio-signal quality is essential for accurate analysis and diagnosis because these signals are often affected by noise and artifacts. This thesis examines a method developed by Idrobo-Ávila et al. (2024) that uses a multimodal approach to assess the quality of physiological signals. Their method was originally tested on two EEG channels, ECG, and respiratory signals across two datasets.

The goal is to evaluate the multimodal method proposed in the referenced article, testing whether it still performs well when using more EEG channels, comparing it to simpler single-signal approaches, and understanding its strengths and weaknesses in real sleep data.

In this study, the same method was applied to the Siesta dataset, utilizing six EEG channels along with ECG and respiratory signals, combining the Signal Quality Index (SQI), rule-based techniques, and statistical indicators to detect poor-quality signals.

The signals were analyzed using continuity, envelope stability, statistical, and entropy-based SQIs, averaging a discontinuity of 0.41%, which indicates good integrity. The ECG gave the best signal quality, and one EEG channel showed the poorest result, whereas the modality-specific statistics verified the presence of stable physiological patterns. Overall, the multimodal framework efficiently distinguished reliable from poor-quality data and demonstrated that multimodal SQIs are a more robust basis for adaptive, automated health monitoring.

Keywords: Biomedical Signal Processing, Signal Quality Assessment, Multimodal Biosignals Analysis, EEG, ECG, Respiratory Signals, SQIs, Rule-Based Methods, Statistical Features, Entropy, Zero-Crossing Rate, Sleep Data, SIESTA Dataset, Machine Learning Limitations

USE OF AI

I have utilized AI tools in my thesis:

☐ NO

☒ Yes

The AI tools utilized in my thesis and their purposes are described below:

ChatGPT5: AI was used to assist in improving the clarity, readability, and formality of the language. Sometimes, for LaTeX code clarification. Also used AI during scripting to assist on development, bug fixing, and merging PNG images (histograms).

Sections where AI tools were used: AI tools were used throughout the entire thesis.

No AI tools were used in generating or analyzing research content or results.

I acknowledge that I am fully responsible for the entire content of my thesis, including the parts generated by AI, and accept accountability for any violations of ethical standards in publications.

PREFACE

Alhamdulillah. By the grace of Almighty, I have reached to the end of my thesis. I am grateful to Mark van Gils, Saana Seppälä, and Alpo Värri for their guidance during my thesis work. I would like to extend my gratitude to Mark and Saana for helping me in writing my thesis.

Lastly, I would like to thank my parents, husband, and children, whose constant support and well-wishes have made it possible for me to finish my master's thesis.

Tampere, 26 November 2025

Maria Islam Smrity

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviation	Meaning
EEG	Electroencephalogram — brain electrical activity signal
ECG	Electrocardiogram — cardiac electrical activity signal
EOG	Electrooculogram — eye movement signal
EMG	Electromyogram — muscle activity signal
PPG	Photoplethysmography — optical blood-volume signal
PSG	Polysomnography — a sleep study, records brain waves, the oxygen level in blood, and heart rate and breathing during sleep
Resp	Respiratory airflow signal
SQI	Signal Quality Index — quantitative measure of signal reliability
SD	Standard Deviation
SD_upper / SD_lower	Standard deviation of Hilbert envelope upper / lower bounds
ZCR	Zero-Crossing Rate — number of times a signal changes sign per unit time
SNR Signal-to-Noise Ratio	
RSA	Respiratory Sinus Arrhythmia — interaction between breathing and heart rhythms
EDF	European Data Format — standard file type for biosignal storage
ML	Machine Learning
ML-based SQI	Signal Quality Index estimated using machine learning models
CSV	Comma-Separated Values — data storage format
Hz	Hertz — unit of frequency
µV	Microvolt — unit of electrical potential
s	Second — time unit

List of Mathematical Symbols

Symbol	Description
$x(t)$	Signal amplitude as a function of time
$\bar{x}$	Mean value of signal x
σ	Standard deviation of the signal
f_s	Sampling frequency
n	Number of samples
t	Time variable

1. INTRODUCTION

1.1 Background and Motivation

Modern health monitoring is based on bio-signals, which affect physical functions and conditions. Among all, ECG, PPG, EEG, breathing, and capnography are widely used methods and are considered useful in common wearable devices, as well as in medical settings. However, in areas such as real-life situations like sleep study, home care, or just movement, very often a signal becomes technically unreliable due to noise, movement, or natural body movements, causing a wrong reading or error in diagnosis. [1, 2, 3, 4].

Many health monitors nowadays take into consideration only a single type of signal, and this sometimes can be a drawback if that particular signal goes faulty while others remain intact. One way to go around it would be the combining of multiple signals (multimodal approach) for Cross verification, therefore signals can confirm each other's accuracy, also it allows for Smarter analysis, which means if one signal is suspiciously wrong, others could help decide whether this is true or just a temporary fault. Hence, the reliability of physiological data in multimodal systems depends not only on the individual sensor quality but also on cross-modal validation and contextual information [3].

This thesis explores a smarter way to assess bio-signals, one that looks at how they relate to each other to ensure accurate and reliable health monitoring.

1.2 Problem Statement

Traditional Signal Quality Indices rely only on a single signal metric to detect noise or artifact, often ignoring contextual and physiological consistency with other modalities. This one-dimensional approach leads to several vulnerabilities. Such as, it can result in overly conservative rejection, where usable data segments may be discarded only for minor errors. In some cases, it may cause false inclusions, retaining signal segments that are improbable or physiologically implausible. In addition, it

might lead to misclassifications, in which genuine abnormal physiological events are mistakenly labeled as poor-quality artifacts.

Moreover, there is no established framework to evaluate a bio-signal's suitability—its accuracy, reliability, and relevance for specific analytical tasks (heart rate estimation, sleep staging, or respiratory assessment).

Thesis Focus: This work addresses the lack of a multimodal, task-sensitive approach—combining rule-based and statistical methods—to assess bio-signal suitability. Specifically, it investigates:

> How can we apply and extend a multimodal framework to reliably determine the usability of individual signals—such as EEG, ECG, and respiration—in real-world, noise-prone sleep data?

1.3 Objectives

This thesis evaluates a published multimodal method for assessing the suitability of physiological signals, aiming to determine whether bio-signals are appropriate for multimodal analysis.

In this study, the method is applied to the Siesta dataset only, using a broader set of signals, including 8 EEG channels, ECG, and respiratory data. The main goals are to validate whether the method performs as reported when used with more EEG channels, to compare the results to simpler, single-signal approaches, and to assess the strengths and limitations of the method in the context of real-world sleep data. This evaluation helps determine the method's reliability, scalability, and practical usefulness in multimodal bio-signal analysis.

2. LITERATURE REVIEW

2.1 Bio-signals in Physiological Monitoring

Biosignals are measurable physiological indicators that reflect underlying biological processes such as electrical, chemical, or mechanical in the human body.-[1, ?] They serve as critical data streams in a variety of medical applications, including clinical diagnostics, continuous patient monitoring, sleep analysis, and wearable health technologies. [2] They can come from breathing (like airflow) or the nervous system (EEG), or the heart (ECG). Specialists can check how organs are working, identify problems, and decide how to treat people by analyzing these signals.

2.1.1 Example of Common Biosignals

There are many types of biosignals those are frequently used in physiological monitoring. These signals can be in forms like an electric pulse, pressure, rhythm, and so on. Biosignals provide valuable insights into our body's functioning mechanisms and help us adapt ourselves into the environment we live in.

Every signal has a narrative to convey. From muscle contraction captured in EMG to the waving of hands or to blood flow via the PPG, all signals have unique details. The understanding of signals is not only about data analysis; rather, it is about merging technology with biology to better humanity. As Webster said in 2009, biosignals are the backbone of biomedical instrumentation, mapping physiology with digital insight [5]. Examples of common Bio-signals are:

Electrocardiography (ECG) ECG records the electrical activity of the heart by placing electrodes on the skin. It is widely used to diagnose arrhythmia, monitor heart health during surgery, or in other situations, and to find ischemic changes. A typical ECG wave consists of P-wave, QRS complex and T-waves, each of which is a distinct phase of the heart's electrical activity during a cardiac cycle.[4].

Electroencephalography (EEG) EEG measures the electrical activity of the brain by using the electrodes placed on the skull. This is necessary to monitor anesthesia depth during surgery, along with epilepsy, sleep disorders, and brain damage. EEG signals are analyzed in separate frequency bands - Delta, Theta, Alpha, Beta, and Gamma - each of which has a meaning regarding specific neural processes and functional states of the brain. [2].

Respiratory Signals Respiratory monitoring can be performed using strategies consisting of airflow volume through a nasal cannula, chest impedance, or capnography. These recordings are important for detecting respiratory patterns, identifying apnea events, and monitoring air flow effectiveness in the course of anesthesia. Airflow signals quantify the amount of air moving in and out of the lungs, whilst capnography measures the concentration of CO_2 in exhaled breath [3].

Other Biosignals Beyond ECG, EEG, and respiratory statistics, many other biosignals are utilized in medical practice, together with photoplethysmography (PPG) for pulse rate and oxygen saturation, electromyography (EMG) for muscle activity, and blood pressure tracking [1]. These signals may be analyzed individually or in aggregate to offer a greater whole picture of a patient's physiological state.

2.1.2 Challenges in Biosignal Monitoring

Observing biosignals, such as the electrocardiography (ECG), electroencephalography (EEG), and respiration, is required to know a patient's physiological state; however, gathering and analyzing these signals in real-world clinical environments is not an easy task. Numerous challenges have impacted the accuracy, reliability, and usability of recorded data.

1. **Noise and Artifacts** Biosignals are contaminated by noise, coming from various sources. For instance, motion artifacts may be generated by the patient's movements or the motion of the surgical staff adjusting equipment.Signals from one physiological process can overlap with one another. For example, electromyography (EMG) activity from facial muscles can be seen in EEG recordings, especially during patient's movement or speaking. Similarly, ECG artifacts can be present in EEG traces if electrodes are placed close to the scalp overlying arteries [4]. A power-line interference can result in a constant

50/60 Hz humming in the signal. Other matters, such as poor contact of sensors or loose electrodes, may set up distorted or flat signals that are of little diagnostic value [6, 1].

2. **Signal Dropouts and Missing Data** A drop in the signal or missing data points would occur during the long-term monitoring processes when the sensor would lose contact with the skin temporarily, or some cable might be pulled out during patient repositioning. Data dropouts usually result in gaps in the data acquisition or in segments of constant value. If these dropouts are not handled carefully, they could affect the reliability of any automated analysis [3].

3. **Large Data Volumes** Classical monitoring systems acquire simultaneously from multiple channels at a high frequency. A single operation may carry gigabytes. Inspecting each segment for quality manually would be impractical. It is a requirement for automated signal quality assessment methods [4].

4. **Lack of Standardized Quality Indicators** While some signal quality indicators (SQIs) exist for specific signals such as ECG or photoplethysmography (PPG), there are no widely adopted multimodal SQIs that can assess the overall quality of several signals together. This limits the ability to perform reliable multimodal analysis in real-time [3].

5. **Variations in Setup** Biosignal recordings vary across institutions, across devices, datasets, etc. Various systems of electrode placement, sensor technologies, and sampling rates can induce subtle shifts in the signal characteristics, presenting another hurdle toward comparing, or even joining, data from various sources [2].

6. **Environmental Factors**

   Changes in temperature, humidity, or mechanical vibrations in the room can also affect sensor's stability. In some cases, condensation in respiratory monitoring equipment can cause short time false readings [2].

These challenges require that biosignal datasets should be appropriately preprocessed and evaluated prior to their use in making clinical decisions or research. The possibility of developing automated tools for quality assessment must be explored; such tools would guarantee that analyses can only be made on reliable data.

2.2 Signal Quality Indices (SQIs)

Signal Quality Indices, or SQIs, are quantitative measures that measure the reliability of a physiological signal. They serve to determine if a given signal segment can be used for clinical interpretation or the basis of automated analysis. By means of an SQI, a researcher can exclude noisy or corrupt signal segments automatically rather than inspect every signal manually, which becomes especially useful when working with large datasets[1].

2.2.1 Definition of Signal Quality Indices

An SQI is a metric or collection of metrics used to assess how well a recorded biosignal captures physiological activity while controlling for undesired noise or artifacts. Various aspects of the signal, including amplitude, frequency content, morphology, or statistical characteristics, may be the focus of the SQIs. A measurement to ascertain whether the signal has stayed constant over a time window is one of the simplest examples of SQIs. Others measure the power ratio in specific frequency bands or examine entropy [4].

2.2.2 Common SQIs in Practice

Three types of SQIs are frequently used:

- **Rule-Based Indicators:** These look for instances of physically unrealistic drifts, out-of-range amplitudes, or constant values.
- **Statistical Indicators:** It may reveal anomalous signal patterns including skewness and kurtosis, zero-crossing rate, and spectral entropy.
- **Physiological Indicators:** These are predicated on established physiological limitations, such as the typical ranges for heart rate and nearly every possible EEG frequency [3].

Combinations of these SQIs have been used in clinical datasets such as VitalDB and SIESTA to evaluate whether respiratory, ECG, and EEG signals are suitable for multimodal analysis [3].

2.2.3 Limitations of SQIs

Despite the widespread use of Signal Quality Indices for assessing the reliability of physiological signals, some limitations impede their application in real-world scenarios. These barriers can be technical, methodological, or practical, eventually hindering fully automated biosignal monitoring.

1. **Unimodal Focus**

   Most SQIs in use today are unimodal. This is a drawback because multimodal datasets are more useful to take into account, such as those that combine respiratory, EEG, and ECG signals. If one signal is corrupted but others are intact, unimodal SQIs may incorrectly classify the entire segment as unusable. There is still a lack of multimodal SQIs that jointly assess quality across different biosignal types [3].

2. **Signal-Specific Nature of SQIs**

   The core idea behind most of the SQIs is that each is specific to the signals under consideration, e.g., ECG or PPG. An ECG SQI might require QRS-marks on the signal for calculation, whereas a PPG SQI might look at pulse waveform characteristics. These indices cannot be adapted to other modalities like EEG or respiratory signals since the characteristics of the signals change. This specificity reduces their flexibility when it comes to multimodal applications where various signals have to be inspected simultaneously [4].

3. **Lack of Universally Accepted Standards**

   An exact, universal definition regarding a ”good” or ”bad” signal does not exist. Various research groups or clinical applications may employ various thresholds for SQIs. For instance, one study could consider an ECG signal acceptable if the heart rate can be reliably extracted, whereas another study could require clean morphology for all ECG components. Such inconsistency makes it unable to compare results across various datasets or replicate findings between studies [1].

4. **Sensitivity to Noise and Recording Conditions**

   Some SQIs detect specific noise sources such as motion artifacts or the detachment of electrodes but may not be able to tackle more complex distortions involving overlapping physiological signals (say, EMG contamination in EEG). Further, SQIs generally assume stable conditions for recording. What lies in a real clinical environment is a backdrop that may vary greatly from one patient

to another, from one medical device to the other, or even from one monitoring setup to another. Thus an SQI that yields good results in one environment may perform rather poorly in the other [6].

5. **Computational and Practical Constraints**

   Some SQIs are computationally demanding, particularly those based on entropy or advanced statistical measures. This may preclude their application in real-time surveillance systems that require prompt decision-making. Second, SQI findings will not always be as helpful in a busy clinical setting unless they are well documented and presented to the clinician in an understandable manner because clinical staff members may not always understand their technical meaning [1].

6. **Variability Across Datasets**

   SQIs are subjected to performance variations owing to dataset characteristics such as the sampling rate, type of sensor, and patient population. For example, entropy measures are dependent on sampling frequency. Because of this, it becomes difficult to compare results across datasets, say between sleep laboratory data and surgical monitoring data. This variability works against the generalization of findings, thus opening the door for quality metrics that are independent of datasets [3].

7. **Limited Validation Across Populations**

   Many SQIs are developed and validated under controlled settings or within certain patient groups (i.e., young adults or cardiac patients). Nonetheless, biosignal quality may differ greatly in elderly patients, pediatric populations, or persons afflicted with chronic illnesses. Without broad validation, it is uncertain whether SQIs remain accurate across diverse clinical groups [4].

SQIs are needed to handle the big dataset of biosignals but paradoxically remain limited by their signal-specific nature, lack of standardization, sensitivity to conditions around recording, uni-dimensional view of signal degradation, and computational or complexity demands and validation. Closing these gaps demands the development of robust, multimodal, and standardized SQIs that can generalize across datasets, patient groups, and clinical environments.

2.3 Multimodal Approach in Biosignal Analysis

The human body is an interconnected, complex network of the vascular, respiratory, and cerebral systems, among others. Unimodal methodologies have traditionally

been used to assess and measure the human body and its biosignals, such as the ECG or EEG. Although these approaches are useful, their narrow scope may not be in accordance with the complex systemic dynamics that are fundamental to a comprehensive understanding of health and disease. However, multimodal systems can revolutionize data-driven research by simultaneously acquiring and analyzing multiple physiological signals, offering a more comprehensive understanding of a patient's condition. [7, 8].

2.3.1 Definition and Concept

The term "multimodal" means the study of different physiological signals, such as ECGs, EEGs, and the respiratory system. The signals occurring simultaneously are aligned and jointly analyzed to observe inter-system relationships and patterns that may be less visible in others, providing better physiological assessments of a patient. For example, one would study cardiorespiratory coupling by combining ECG and respiratory data, whereas brain-heart interactions during sleep and anesthesia are studied by combining EEG with cardiovascular measurements [9, 10].

More broadly, multimodal analysis implies the recording of several signals, but it does not stop there. The signals must be carefully synchronized, preprocessed, and integrated so that they can be meaningfully compared. This might include the alignment of signals with different sampling rates, noise artifact compensation peculiar to each modality, and the extraction of features that represent physiological processes on the same time scale [11, 3].

2.3.2 Importance of Multimodal Analysis in Biomedical Signal Analysis

The importance of considering several modalities at a time can be listed as below:

Capturing Inter-System Interactions

Interacting systems is another definition for inter-system interactions. One pertinent example of this would be cardiorespiratory coupling as it arranges mechanisms by which breathing modulates heart rate, such as respiratory sinus arrhythmia. Analogously, as EEG activity patterns during sleep are associated with variations in respiration and cardiac rhythms, these cross-modality interactions further elucidate autonomic regulation, stress response, and homeostatic balance—an understanding which the isolated analysis of either modality alone cannot provide [12, 9].

Improving Diagnostic Accuracy

In diagnostic settings, multimodal approaches have proven to increase diagnostic reliability. In sleep medicine, for example, the diagnosis of sleep apnea requires the integration of EEG (the sleep stage), airflow (interruption of breathing), and oxygen saturation (physiological impact). A unidimensional approach would fail in giving the whole picture and may result in false positives or false negatives. In much the same way, in ICU, a combination of ECG, blood pressure, and respiration would detect the onset of sepsis or shock earlier than would be possible by tracking just a single parameter [1, 2, 13].

Ensuring Patient Safety and Monitoring

Another element of utmost importance is monitoring safety in real time. For instance, during surgery, anesthesiologists would also consider multimodal monitoring—it could be any combination of an ECG, EEG, oxygen saturation, or capnography—to find the appropriate depth of anaesthesia and detect any complications that could arise. The use of pure EEG would mostly end in misinterpretation of an artifact; on the other hand, when the system is integrated with cardiovascular and respiratory measurements, it drastically reduces such risk of errors. Therefore, this redundancy delimits the optimization of monitoring and decreases false alarms that accumulate and constitute alarm fatigue in the clinical setting [6].

Supporting Advanced Computational Models

Multimodal biosignal analysis supports the developments of machine learning and AI-based models for healthcare. Deep learning architectures such as CNNs and RNNs tend to perform better when trained on multiple complementary signals. For instance, combining ECG and photoplethysmography (PPG) improves blood pressure estimation, whereas the combination of EEG and ECG features helps in the automatic classification of sleep stages. Thus, multimodal analysis assists greatly in enriching interpretability and serves as a source of ingenuity for predictive analysis and decision support systems. [14, 15, 11]

Robustness Against Noise and Artifacts

Lastly, multimodal analysis can ensure system robustness against noise contamination. In real-world applications, particular signal channels could very well be contaminated due to situations such as sensor detachment or electromagnetic interference. By checking one modality against another, the system can still come up with useful information and can avoid making a critical error. For example, a

flat-line ECG resulting from a detached electrode would be interpreted as an artifact if respiration and EEG remained normal. This redundancy is the most persuasive argument for using multimodal approaches in both research and clinical monitoring [16, 15]. However, to utilize the full array of benefits, it is necessary to face technical challenges in areas such as synchronization, standardization, and quality assessment [17, 18].

2.3.3 Benefits for Suitability Assessment

Perhaps one of the most powerful aspects of multimodal biosignal analysis is in assessing data suitability and quality. In large-scale datasets, not every single segment of recorded signal is usable due to noise, artifacts, or other technical issues. In the standard scenario, the signal would be inspected manually in order to assess datum suitability for analysis, but with thousands upon thousands of hours of recordings, such an approach becomes implausible. The multimodal approach can improve and make this assessment more reliable by harnessing the advantage of redundancy across modalities [1].

It benefits multimodal suitability assessment to have redundancy across signals, to be capable of detecting physiological inconsistencies, to have higher labeling confidence, to be scalable to larger datasets, and to be useful for clinical decision support.

Benefits of Multimodal approach

The single best advantage that we gain from multimodal evaluation is the ability to check multimodal physiological coherence. To illustrate with a few examples:

- A flat line on the ECG accompanied by active EEG and respiration suggests an electrode detachment as opposed to a cardiac arrest. The use of this kind of cross-modal redundancy has been highlighted in network physiology. [7, 19]

- Atypical breathing patterns without accompanying EEG arousal or ECG variability indicate faulty airflow sensors instead of a true apnea event.

- Simultaneous eruptions on the EEG, heart rate, and respiration markers are multimodal confirmed arousal events.

Multimodal evaluation is unique in its capability of distinguishing true physiology from its artifacts and is impossible using a single evaluation system. It represents the principles of network physiology, which views the human body as a single system, as opposed to separate and independent units. [7]

Multimodal analysis, hence, becomes vital toward more reliable biosignal processing for both research and clinical practice. Redundancy Across Modalities, Physiological Inconsistency Detection, Higher Confidence in Data Quality, Enabling Large-Scale Automated Screening, and Improved Clinical Decision Support are some of the key sections where multimodal suitability effects drastically. Despite the tremendous advantages of multimodal analysis, they can only be realized if standardized multimodal SQIs and computational methods suitable for large interdatasets are developed. [16, 14]

2.3.4 Challenges in Implementing Multimodal Approaches

Despite the fact that multimodal biosignal analysis has great advantages compared to unimodal approaches, it still has a complex practical implementation. The problems arise from technical issues like synchronization and data integrity but also from the limitations posed by the different methodological and ethical constraints in deploying the technology on a large scale.

1. **Signal Synchronization and Temporal Drift**

   One of the most persistent barriers is the synchronization of multimodal recordings. Differences in sampling rates, hardware latency, or device-specific delays can affect the physiological coupling measures. Modern datasets such as VitalDB have attempted standardized timing [20], yet minor offsets still disrupt the temporal coherence between EEG, ECG, and respiratory channels. Recently, edge-AI systems, such as BioGAP-Ultra [16], have shown integrated synchronization modules, yet cross-device harmonization at sub-millisecond resolution is still unresolved.

2. **Handling Missing and Heterogeneous Data**

   In real-world recordings, sensors frequently fail or detach, resulting in partial loss of data. Although unimodal analyses can just drop such segments, multimodal frameworks risk to loss large-scale of data. Foundation-model-based methods are now trying to infer or recover for the missing modalities [21, 15], but the challenge of maintaining accurate results under very fluctuating conditions, such as motion or perspiration, still exists as a key limitation.

3. **Lack of Standardized Multimodal SQIs**

   There is no accepted SQIs for multimodal data at present although there are many SQIs for the single signals like ECG or PPG. Such lack of standard

makes it challenging to measure the global quality of multimodal datasets. A common approach consists in combining unimodal SQIs and consider a segment as suitable if all the modalities cross their thresholds. Although this is reasonable, it may be too stringent and some informative data parts are discarded. The development of universal multimodal SQIs is necessary to capture the integrated quality of multiple signals better [1].

4. **Computational Burden and Real-Time Feasibility**

   Multimodal analysis is much more computationally demanding than unimodal methods. High-definition information in many pipelines will then produce large data volumes in a short time. For instance, a temporally long surgery may yield gigabytes of pools of multimodal recordings. With the real-time processing of these signals for clinical uses such as anesthesia monitoring or intensive care, efficiency in the algorithms and hardware architecture is required. By contrast, bottlenecks in computing can limit the use of multimodal approaches for hospitals with limited resources [16, 4].

5. **Variability Across Clinical Environments**

   Another important problem is the variability of signals obtained in multiple clinical settings. For example, sleep lab signals tend to be far less noisy than those from the operating room, as patients are generally immobile there; the surgical data, on the other hand, suffers through electrocautery, equipment contamination, and acute patient movements. This discrepancy implies that an algorithm trained on a given dataset might be inefficient on another and thereby impair multimodal methods from generalizing [6].

6. **Ethical and Privacy Concerns**

   Another crucial issue with multimodal datasets is that they may contain identifiable patient data. Sharing datasets for research presents ethical and privacy concerns, particularly as further physiological signals are integrated. This increases the possibility of pinpointing an individual patient, so adherence to data-security regulations like GDPR is a must. Striking a balance between protecting the patients' all-important privacy as opposed to allowing for data-sharing for research purposes remains a challenge [22, 23, 24].

2.4 Gaps in Existing Work

Although multimodal biosignal analysis has gained significant attention in recent years, however, the area still has no unified theory, and it is methodologically inconsistent, which results in crucial scientific and practical gaps. Researchers in the

past have associated the quality of the signal with mere absence of noise, overlooking the fact that the signal might be carrying some physiological information that reflects the true state of the interconnected systems. In such a comprehensive framework as Idrobo-Ávila [3] the concept of the appropriateness of a signal was mainly linked with statistical limits and not with systemic coherence among the signals coming from the brain, heart, and lungs. This shows that the requirement for definitions of "suitability" grounded in physiology is still there and it has to be broader than the technicality of signal purity. Furthermore, advances in machine learning and multimodal fusion have been really remarkable, but still, the existing methods are largely based on correlation and very rarely attempt to explain the causal or mechanistic basis of observed inter-signal relationships. The new area of network physiology [25, 8] reveals that the physiological systems act as networks that are dynamically coupled with each other, however, most of the multimodal pipelines ignore these controlling connections and treat the multimodal fusion as a mere computational task instead of a biological modeling problem. Furthermore, there is still poor knowledge about the way uncertainty is passed from one modality to others in the case where there is a degradation of one modality—for example, if an ECG segment is not reliable, only a few frameworks will attempt to quantify the effect this has on related EEG or respiratory inferences. Just a few recent works [14, 15] address the issue of cross-modal quality dependency; thus, the question of redundancy and resilience in multimodal datasets remains unanswered.

Another gap arises from sleep-related signals and long-duration recordings, which is another source of this disparity. The annotation is simple in the case of short, event-based or laboratory acquisitions (emotion recognition, surgical monitoring) where most of the literature stays, but such SIESTA-type multi-hour physiological recordings show dynamics that cause slow drifts, intermittent sensor detachment, and state transitions. Therefore, the current SQIs remain incapable of capturing these phenomena to a good extent [13, 26]. The clinical adoption gap has also become bigger alongside algorithmic innovation: high-performance architectures like PHemoNet [11] and edge-AI systems such as BioGAP-Ultra [16] have not yet been translated into interpretable, trustworthy clinical tools, which indicates that the barrier is more methodological than computational—i.e., it is all about linking automated decision-making with human interpretability and workflow integration. Moreover, ethical and semantic inconsistencies exist; multimodal datasets not only differ greatly in the way they are annotated, but also in the metadata they contain and the level of privacy they offer, thus making it difficult to reproduce and share securely. While new initiatives like Rammohan et al. (2025) [27] and Schouten et al. (2024) [28] advocate for open-source annotation frameworks and more manage-

able governance models, the community is still without a consensus on the ontology of ”quality” and ”suitability” for biosignals. Taking these points into account, the shortcomings still indicate that even with the advancement of technology, the research on multimodal biosignals is still dependent on non-unified physiologic definitions, mechanistic modeling of inter-system coupling, empirical understanding of cross-modal quality propagation, validated adaptation to continuous real-world data, and standardized ethical frameworks for data stewardship that the present thesis addresses.

3. METHODOLOGY

3.1 Data Collection

The present research was done on the SIESTA project, which was a European endeavor charted during the 1990s to set up a unified database of sleep research. [29] To avoid the "first-night effect," every subject had two nights of full polysomnography (PSG) in a controlled laboratory environment. The dataset comprises 197 healthy individuals and 97 patients suffering from widely spread sleep and neurological disorders like sleep apnea, Parkinson's disease, depression, and anxiety (according to the ICD-10 classifications). The study kept patients on unstable psychotropic medicines out to ensure the physiological validity of the results.

The participants were given medical and psychological examinations to check if they were fit for the study. During the recordings, EEG, EOG, and EMG channels, as well as an ECG channel, were used to record the activity of the brain, eyes, muscles, and heart, respectively. Respiration was monitored through the use of airflow sensors, belts placed around the thorax and abdomen, and pulse oximetry.

SIESTA not only gathered physiological data but also collected sleep diaries, actigraphy for 14 days, and psychometric assessments of mood, attention, and memory. The use of the multimodal structure made it possible to analyze the physiological, behavioral, and psychological factors of sleep.

3.2 Tools and Software Used

For the research in this thesis, open-source software packages, self-developed Python scripts, and an appropriately configured computational setup were utilized as a mixture of tools. As with other forms of data analysis, the processing of biomedical signals has its own set of requirements that stem from the domain itself, as well as from numerics. For this purpose, the needs of reliable implementation of biomedical file formats, effective signal processing, and large-scale tabular data management needed to be met. The selection of tools was guided by the principles of transparency, reproducibility, and extensibility.

Python as the Core Environment

All data analysis was carried out in the Python programming language (version 3.11.9). This is frequently employed in biomedical informatics studies, given that it is open-source as well as compatible with numeric computation, AI, and signal processing libraries. Python can manage, from start to finish, all the steps of the workflow, which includes the importing of raw data and the final computation of the SQI, and this entire process could be documented, versioned, and rerun.

MNE-Python for Biosignal Import and Management

Handling the raw polysomnography (PSG) recordings from the SIESTA dataset required specialized software capable of reading EDF (European Data Format) [30] files. For that, the MNE-Python [31] library was imported. MNE is a well-established and highly utilized tool in neuroscience and biomedical engineering that supports EEG, ECG, and other multimodal physiological data in a comprehensive manner.

NeuroKit2 for ECG Signal Quality Indices

Though MNE is effective for the general handling of biosignals, ECG-specific analysis required additional features. The NeuroKit2 [32] library was used for this purpose. NeuroKit2 incorporates implementations of popular QRS detection algorithms such as Pan-Tompkins (1985) [33], Hamilton (2002) [34] and Christov (2004) [35] methods. Each of these algorithms performs well under different noise levels and ECG signal morphologies.

NumPy, Pyarrow and Pandas for Computation and Data Handling

To perform large-scale SQI computations, efficient use of arrays and tabular data was necessary. NumPy was useful, particularly for its vectorized numerical operations, which streamlined the calculation of standard deviations, variances, and even kurtosis and skewness over signal windows. NumPy's array-oriented design empowered the computation of SQIs on thousands of overlapping windows and made feasible the computation of SQIs without prohibitive compute time.

Pandas served as the central framework for data organization. After the computations of SQIs were finished, the results were extracted into long-format data frames and directly exported to tables in the results chapter.

Pandas was also used for exporting results into CSV files, which was central to the workflow. However, for rule-based calculations, Pyarrow was used to save large files

in **parquet** format, rather than CSV format, which was essential to handle such huge files.

Custom Python Scripts for Signal Quality Indices

Not every SQI of concern was available in the libraries, necessitating several bespoke Python implementations. All the measures were implemented in Python for ease of integration, although the logic behind them was based on thresholds and methods proposed in earlier research. The coding process of these SQIs made it possible to modify parameters, add new features, and reshape the whole process to the SIESTA dataset.

Data Export and Organization

The final outputs were organized at three levels:

Window-level CSV and parquet files: Containing detailed SQI values for every 10-second segment.

Per-record CSV files: Containing a summary for each EDF file, such as the portion of suitable windows.

Summary CSV files: Providing dataset aggregates expressed as mean ± std.

PNG files: Histograms were saved as PNG-type image files.

Because of this structure, it was possible to check detailed information and also provide summary findings. Additionally, it was easier to fix problems: if there was an issue with the summary, it was possible to go deeper into the per-record or to the window-level outputs for further examination.

Computational Infrastructure

The analysis was completed using Python and VS Code on a Windows 10 computer(32GB RAM). This SIESTA dataset work, consisting of hundreds of PSG recordings, required considerable computational power, and the workstation was sufficient, yet time-consuming. The use of light open-source software allowed the work to be done on multiple devices, if needed, without any complications in the licensing or the use of proprietary software.

3.3 Preprocessing Steps

To compute signal quality indices (SQIs), the biosignals from the SIESTA dataset had to go through a preprocessing step. Preprocessing serves as a significant phase in the processing of biomedical signals which is mainly because the raw data has a lot of issues, including inconsistent channel names, sporadic missing data, and changing conditions of the recording. Biometric data quality metrics and algorithms face difficulties with such issues and even a small problem can cause significant undetectable systematic bias in the outcome. The described problems were the same difficulties that the preprocessing pipeline intended to tackle in this research. Unlike artifact removal, the author did not intend to "clean" the signals but rather the opposite; she wanted to keep the honest richness of the data and make very small transformations in order to maintain consistency.

Steps taken for reading and inspecting EDF Files:

1. Loading the EDF File:

Used the MNE library to load the polysomnography (PSG) data from an EDF file. (16-channel polysomnography)

Extracted available channel names to select specific EEG channels for analysis. Description [36] of channels is given in Table 3.1

These channels refer to different types of physiological measurements typically used in sleep studies or polysomnography.

2. Channel Selection:

Choose the 'Fp1-M2' to O2-M1, 6 EEG channels, ECG channel, and Respiratory channel for analysis.

3. Signal Extraction:

Extracted the signal data of the selected channels using MNE's data extraction methods.

Verified the extracted signal by checking its shape and plotting the first few samples.

Challenges with reading data: A major problem noticed at the beginning was irregularities in the channel names. As an example, some of the EDF headers had either null characters (\x00) or trailing spaces that caused an anticipated "ECG" channel to be recognized as "ECG(\x00"). This was a serious issue because a later

Channel(s)	Description
Fp1-M2, Fp2-M1, C3-M2, C4-M1, O1-M2, O2-M1	EEG (Electroencephalography) channels measuring brain activity. The letters and numbers indicate specific electrode placements on the scalp following the international 10–20 system.
M2-M1	Reference channel for EEG, with M1 and M2 placed on the left and right mastoid processes behind the ears.
Pos8-M1, Pos18-M1	EOG (Electrooculography) channels measuring eye movements.
EMGmental, EMGleft	EMG (Electromyography) channels measuring muscle activity, specifically from the mental (chin) and left (leg) areas.
ECG	Electrocardiography channel measuring heart electrical activity.
Airflow	Measures the airflow through the nose and mouth.
Chest, Abdomen	Respiratory effort channels measuring the movement of the chest and abdomen.
SaO2	Measures oxygen saturation in the blood using a pulse oximeter.

Table 3.1 *Description of Channels in the SIESTA Dataset*

part of the pipeline would depend on the automated selection of certain channels, such as Airflow, ECG, or EEG_1–EEG_6. To fix this, every file was fixed to be subjected to channel renaming right after import:

```
raw.rename_channels(
    lambda name: name.strip().replace('\x00', '')
)
```

The step above seems minor, but it was indispensable for proper downstream analyses. It standardized all channel names in the dataset. Such a small error in labeling a channel could have had immense consequences such as skipping a file or processing it incorrectly.

It is important to mention that most of those files are saved in folders named with the alphabet 'M', which were unreadable through Python and Matlab. Although

the online EDF viewer was able to plot those files, assuming they have used more forgiving parsing, rather than using strict parsing as Python or Matlab. However, with the Hex viewer, it was noticeable that the header format was not as exact as an EDF should be. As the number of files was insignificant compared to the whole database, these files were ignored.

Selection of Channels and Maintaining Consistency

The SIESTA recordings have a number of channels, including EEG derivations, eye movements, muscle activity, respiratory effort, airflow, ECG, and oxygen saturation. A targeted subset was labeled for this study:

1. EEG_1 to EEG_6 (to capture cortical electrical activity)

2. ECG (always found at index 11 after cleaning)

3. Airflow (to track respiratory patterns)

Concentrating on these essential channels ensured equal treatment for all subjects as well as to avoid the problems caused by the presence of optional or auxiliary channels, as these were not available in all recordings. The choice to use fixed channel indices (especially for ECG) instead of heuristic channel picking made the process easier to manage and guaranteed the same results every time.

Segmentation into Windows

Once the pertinent signals were loaded and selected, a continuous recording was segmented into 10-second windows with a 5-second hop size, equating to a 50% overlap. The segmentation approach was governed by the trade-off between resolution in time and computational practicality. A 10-second window suffices to capture several cycles of relevant physiological activity (e.g., multiple heartbeats in ECG or several breaths in airflow) while still being short enough to localize transient noise or artifacts.

To mitigate the occurrence of "blind spots," overlapping windows were implemented: if an artifact or signal dropout happened in the middle of a window, it might not be detected with a non-overlapping approach. The 50% overlap ensured coverage such that each part of the recording was contained in at least two windows, improving robustness in subsequent aggregation. The segmentation produced thousands of windows per subject, which constituted the basis for SQI computation.

Handling Missing and Invalid Data

Although the SIESTA dataset is of very high quality, occasional missing or invalid samples still occur, which is also important for this research to recognize. This can be due to electrode disconnection, sensor malfunction, or technical errors during recording. In the preprocessing pipeline, any detected NaN (Not-a-Number) values were automatically replaced with zeros (boolean False) to prevent crashes in downstream functions:

if np.isnan(signal).any(): signal = np.nan_to_num(signal, nan=0.0)

Because SQIs look for unusable windows (including flat-line signals), turning NaNs into zeros ensured those windows still got marked as low quality, allowing the pipeline to continue. In essence, bad data were flagged as "bad" instead of being smoothed over or skipped entirely.

Preparing for SQI's

Once the signals were segmented and cleaned, the next logical move was to compute the SQI. Preprocessing at this point, however, needed to have completed both the general and the channel-specific procedures. For the ECG signal, NeuroKit2 was used to run multiple QRS detection algorithms in parallel. For the EEG and airflow signals, envelope extraction and statistical computations to prepare NumPy arrays were performed.

In addition, preprocessing ensured that every window contained metadata tags for the file name, channel, and time boundaries. This was essential for handling the outputs: window-level SQIs could be filtered by channel or aggregated by subject. The decision to add metadata to each window prevented errors like those seen in extensive datasets, where files can become separated from their source.

Consistency with Literature Protocols

Finally, it is worth noting that these steps were thoughtfully chosen based on existing methodology. For example, there is a specific explainer paper where 10-second windows with 5-second overlaps are ideal in certain windows of analysis. Specifically, Idrobo-Ávila [3] made a valid comparison between data collected from surgeries and sleep data. In addition, there are expanding discourses alongside the biomedical signal quality literature, where physiologically plausible thresholds, alongside standardizing channel labels, are assumed norms. Integrating the preprocessing steps with what is already well-known and empirically grounded helped the study deliver results that are reliably replicable and can be integrated with the body of literature at hand.

3.4 Signal Quality Metrics Used

After segmenting the biosignals from the SIESTA database into overlapping windows, the next step in this study was to analyze their appropriateness with signal quality indices (SQIs). The SQIs utilized in this study fall into three main categories: rule-based SQIs, statistical SQIs, and physiological plausibility checks. Combined, these ensured checks from different vantage points to the signal quality of the EEG, ECG, and respiratory channels.

Rule-Based Signal Quality Indices

Rule-based SQIs are implemented using simple thresholds based on known sensor physical constraints or physiological ranges. Their main benefit is their transparency and low computational cost, which is crucial for processing large datasets such as SIESTA.

Constant signal :

Any window in which the signal did not change for a duration of 500 milliseconds or more was marked as a poor-quality or non-continuous signal. These segments were identified by checking whether consecutive samples within a given time window remained identical. Formally, this condition is expressed as :

$$x_i = x_{i+1} = \cdots = x_{i+w_{\text{length}}-1}, \tag{3.1}$$

which holds for any sample index $i = 0, 1, \ldots, N - w_{\text{length}}$, where N denotes the total number of samples in the signal x, and w_{length} represents the number of samples corresponding to a 500 ms time window. This procedure was implemented in accordance with the method described in [37].

Out-of-range :

Another rule-based measure was the amplitude check. For ECG, signals exceeding ±3.5 mV were marked as invalid [38], while for EEG, the threshold was ±110 μV. [39] This was designed based on accepted ranges and physiological norms from the literature. Values outside these limits typically reflect artifacts, amplifier saturation, or external interference. In the case of airflow, similar constraints were applied to airflow to make sure it represented a plausible respiratory pattern and not some other sensor noise.

All these checks, based on rules, were intended to be the first quality-control filters. They were intended to be conservative in the sense that it was not required that

windows passing the checks be clean of artifacts, but it was guaranteed that any window failing the check was not useful.

Statistical Signal Quality Indices:

Alongside rule-based checks, several SQIs were generated from the signal's statistical properties. Such indices help multiple degrees of nuance in signal variability and complexity and often identify issues that multiple thresholds fail to highlight.

Standard deviation of the upper and lower envelopes:

For ECG and EEG, the 10-second signal was divided into 2-second sub-windows while for respiration 5-second sub-windows were used. Maxima and minima were obtained within each sub-window to capture an envelope of the signal for that sub-window. The envelopes obtained were then subjected to standard deviation computation over the 10-second window.

$$\sigma(x) = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \tag{3.2}$$

A signal that is flat, corrupted , or physiological is prone to show variability marked by distinct minimum and maximum values of the envelope as compared to a healthy dynamic signal. On the other hand, variability that is very high may be indicative of noise or artifacts. This metric proved crucial in the identification of noisy peaks and was previously used for assessing the quality of pulse oximetry signals. [40]

Skewness: Skewness gives more detail on the distribution of the signal. Practically, near-zero skewness suggests symmetric, artifact-free signal, whereas large positive or negative values often indicate motion artifacts, baseline shifts, clipping, or electrode issues. It was used as a metric of quality in different types of signals such as pulse oximetry [41], ECG [42], and EEG [43].

$$\text{Skewness}(x) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma}\right)^3 \tag{3.3}$$

Kurtosis: Kurtosis shows how much the signal has outliers or "spiky" activity, meaning whether a given signal contains extreme values or not. For ECG signals, the appropriate kurtosis value corresponds to the QRS spikes, while in EEG signals, a high kurtosis value shows the presence of episodic noise spikes. Unlike skewness kurtosis was used for pulse ECG [42], and EEG [43] and oximetry [44].

$$\text{Kurtosis}(x) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma}\right)^4 \tag{3.4}$$

Spectral entropy: The unpredictability of the signal in the frequency domain was also analyzed using spectral entropy. [43] Low entropy is exhibited by signals with narrow frequency bands of energy (e.g., rhythmic breathing), while high entropy is exhibited by noisy signals with wide frequency bands. It has been used for measuring signal quality of ECG [45], and EEG [43] signals and pulse oximetry [44], In this thesis, normalized Shannon (spectral) entropy is used:

$$\text{Entropy}(x) = -\frac{1}{\log_2 M}\sum_{i=1}^{M} p(x)_i \log_2\left(p(x)_i\right) \tag{3.5}$$

where,

- M: Number of discrete bins (e.g., frequency bins in a power spectrum).
- $p(x)_i$: Normalized probability or normalized spectral power in the i^{th} bin.
- The division by $\log_2 M$ ensures that the entropy value ranges from 0 (perfectly ordered) to 1 (maximally random).

Zero-Crossing rate: The technique is based on the idea that it is possible to simultaneously separate signal and noise if one knows that signal is present only in a certain limited low-frequency range – the case for EEG is 0–30 Hz [46] and for ECG 0.05–100 Hz [47], but with the major part of the signal below 30 Hz [48]. On the contrary, noise is present everywhere in the frequency domain and generally has a higher power at both ends of the frequency range [47]. Several biosignals have made use of this method before, such as electrocardiography (ECG) [4] and electroencephalography (EEG) [49]. The zero-crossing rate is defined in the following way:

$$\text{Zero-Crossing Rate}(x) = \frac{1}{N-1}\sum_{i=1}^{N}\mathbb{I}(x_i\, x_{i-1} < 0), \tag{3.6}$$

where

$$\mathbb{I} = \begin{cases} 1, & \text{if } x_i\, x_{i-1} < 0, \\ 0, & \text{otherwise.} \end{cases}$$

Physiological Plausibility Checks:

The process for the third layer of SQIs was a bit different since it was applied to the ECG signals alone with a set of rules relating to physiological plausibility rules. First the R-peaks were detected with multiple QRS detectors (namely, Pan-Tompkins, Hamilton, and Christov). Each of the 10-second windows [38] was then tested with the following conditions proposed by [50], which is especially suitable for surgery data with many motion artifacts [6]:

- Heart rate between 40 and 180 beats per minute. Heart rates outside of this range are nearly impossible and typically indicate misdetected peaks or artifacts.
- Maximum RR interval 3 seconds, corresponds to a minimum heart rate of 20 bpm. It ensures that more than one beat can not be missed. Extremely long pauses in detected heartbeats are either artifacts or pathological events not expected in the normative SIESTA recordings.
- Max/min RR ratio 2.2. This is to ensure consistency in beat-to-beat intervals and to lower the effect of spurious detections.

We used a majority voting method to combine each detector's decision. This means that if two of the three detectors agreed the window met the physiological guidelines, the window was regarded as suitable. This method lessens the influence of any single detector and enables a more reliable assessment of the quality of the ECG.

Macro and Micro-Averaging of Signal Quality

To measure overall signal quality of ECG across the recordings, two complementary averaging methods were used: *macro-averaging* and *micro-averaging*. In the macro-averaged metric, the first step was to calculate the percentage of segments with good-quality signal for each individual recording, and then the average of these percentages was calculated across all recording samples. This approach assumes that every recording is of equal importance irrespective of its length. On the other hand, the micro-averaged metric was computed by combining all segmented windows from every recording and calculating the ratio of good-quality segments in this combined set. This approach gives weight to the recordings in proportion to their length, as the longer the recording, the more windows will be contributed.

3.5 Multimodal strategy

The multimodal approach is developed to evaluate biosignal quality sequentially by merging segmentation, signal quality indicators (SQIs), and conservative combination rules. The multimodal framework differs from unimodal SQIs because it does not analyze one channel at a time but makes sure that every segment shows physiologically valid activity across EEG, ECG, and respiration. This part explains the fundamental elements of the approach: segmentation, unimodal SQIs, and multimodal integration. [7, 8]

Multimodal Integration

After computing the SQIs for the channels, the segments are retained only if every modality passes the quality checks. This is a conservative approach that helps in preserving the quality of data for all further analyses. Cardiorespiratory coupling studies need both ECG and respiration data to be noise-free, while studies involving EEG and ECG data need to have neural and cardiac data clean and artifact-free. Reducing false positives and improving robustness to noise in a single channel is achieved through the enforcement of agreement across all modalities, which they call cross-modal agreement. [51, 7]

Application in SIESTA

Adopting this model in SIESTA guarantees that the sleep studies are based on high-quality multimodal datasets. Although ICUs and surgical wards have compromised conditions, laboratories for sleep studies also suffer from artifacts. These are caused by the detachment of electrodes, shifting of the body, or drift in the airflow sensors. Earlier research has also established that in the absence of cautious screening, nearly 15 to 20 percent of the sleep studies contain segments that are unusable. [2, 52, 53].Integrating this multimodal approach, the present thesis ensures against such artifacts and improves the reliability of subsequent analyses of sleep architecture and physiological interactions. [54, 6, 55, 56, 53]

3.5.1 Modalities Used

The multimodal approach in this thesis considers the three key physiological parameters from the SIESTA dataset – the electroencephalogram (EEG), electrocardiogram (ECG), and respiration (airflow signal). The exploration into the benefits of merging physiological signals came about by analyzing both pairwise and triplet combinations of the different modalities. The primary reason for this was to discover if the

multisignal integration actually made the data segments' reliability better compared to when the individual signals were assessed.

For each of the three modalities (EEG_i, ECG, and respiration), a 10-second window was designated and the corresponding Signal Quality Index (SQI) values were computed. In the case of a multimodal evaluation, the SQIs were synchronized and overlaid on the time axis of the different modalities so that when the signals were time-sequenced, they all referred to the same segments.

Pairwise Combinations. The binary quality classification for the two modalities selected for a particular window i.e. *good* or *bad* determined their comparison. If no physiological inconsistency existed between them (e.g., unrealistic heart–respiration timing), a segment was characterized as *usable* if at least one modality came out with good quality. The pairwise quality score (Q_{pair}) was calculated as the proportion of usable windows to the total number of analyzed windows:

$$Q_{pair} = \frac{N_{\text{usable}}}{N_{\text{total}}} \times 100\% \tag{3.7}$$

where
$\mathrm{Q_{pair}}$ = pairwise quality score,
N_{usable} = number of usable windows
N_{total} = total number of analyzed windows.

The whole process was executed for all the possible pairs of signals (e.g., ECG–Resp, EEG_3–ECG, EEG_4–Resp), finally giving rise to the pairwise scores mentioned in table 4.6.

Triplet Combinations. To examine three-signal combination (e.g., EEG, ECG, and respiration), the same concept was applied using the logical AND/OR rule. A window received a *good* label if quality thresholds for two out of the three modalities were at least met which, of course, provided redundancy and was strong against single-channel dropouts:

$$Q_{tri} = \frac{N_{(\text{ 2 good})}}{N_{\text{total}}} \times 100\% \tag{3.8}$$

where
$\mathrm{Q_{tri}}$ = three-signal quality score,

$N_{(\text{ 2 good})}$ = number of windows where at least two of the three modalities are of good quality,
N_{total} = total number of analyzed windows.

Table 4.7 therefore, represents the proportion of time windows where multimodal agreement confirmed data suitability. These analyses quantify how integrating correlated physiological signals can improve the overall reliability of multimodal biosignal datasets.

Looking at EEG, ECG, and respiration signals together, the multimodal approach takes advantage of both overlap and differences in the information. Sleep staging would not be possible without EEG as it provides essential neural information, and the ECG gives strong cardiac dynamics, while the respiration is used to check the breathing rhythm and also to help in the identification of sleep-disordered breathing. The use of these different signals at the same time guarantees reliable and quality data that is physiologically consistent. In SIESTA, this is a great justification for the use of multimodal analysis.

3.5.2 Synchronization and Alignment of Modalities

A crucial part of analysing multimodal signals is the verification of the horizontal and temporal alignment of the different biosignals that have been collected from the different physiological systems. The absence of alignment might hide or even lead to unnecessary distortion of the interactions between different modes, for example, cardiorespiratory and EEG-ECG interactions. As in the case of SIESTA, the use of simultaneous polysomnography (PSG) recordings goes a long way in solving many of the synchronization problems; there is however, still the need to solve some, approach missing data, sensor delays, and differences in sampling rates.

Sampling Rate Differences

SIESTA modalities vary in sampling rate: EEG and ECG are sampled at 100 Hz, while respiration is sampled at lower rates ($\approx 16\,\text{Hz}$) . For comparability, respiration signals are resampled to 100 Hz to coincide with the higher-frequency signals before windowing. Resampling techniques have also been widely applied in multimodal physiological recordings, e.g., PhysioNet and fetal ECG research, where maternal–fetal signals must be properly synchronized. [57]

Handling Missing or Corrupted Data

Often in multimodal synchronization, there tends to be a little bit of data loss, which

can arise from something as simple as a short disconnection of a device or a bug in the recording. Instead of dropping entire multimodal windows due to damage in one channel, the data gets padded (for example, with zeros) so that the alignment remains temporal. Since SQI's function in classification, such windows can be flagged as unsatisfactory. Because data from such windows is marked as unsatisfactory, it is retained for comparison, removing any bias. Both clinical monitoring and affective computing studies identify that incorrectly aligned or missing data is the primary cause for the decline of multimodal systems' performance. [58]

Windowing and Metadata Alignment

Once finished re-sampling and handling missing data, all signals are divided into 10-second segments with synchronized data and 5-second overlaps. All modalities share the same start and end time. Each segment is labeled with the following metadata: subject id, channel type, and time indices. This allows for easier verification that successive SQIs of suitability evaluations are performed on the same data segments. Other multimodal datasets such as PhysioNet and VitalDB have also indicated that retained synchronization of metadata is crucial in supporting reproducibility of biomedical signal research. [54, 20]

3.5.3 Ethical Considerations

Patient data are collected in biomedical research with significant ethical responsibilities attached. In this thesis, the SIESTA dataset was utilized to assess multimodal signal quality. The SIESTA is a large sleep database containing recordings of EEG, ECG, and respiration. Biosignals can reveal sensitive information regarding a person's state of health and physiology; thus, issues related to consent, privacy, and the responsible secondary use of data needed to be considered. The legislation concerning secondary usage of SIESTA data serves as an ethical safeguard: the recordings were obtained through informed consent, in which the participants voluntarily accepted that their data would be anonymized and used for future research, in line with the Belmont Report and the Declaration of Helsinki's stress on respect for autonomy [59]. The SIESTA database is made available without direct identifiers and with the retention of only non-identifiable metadata (e.g., age, sex, and sleep-related clinical parameters). This is in accordance with the EU General Data Protection Regulation (GDPR) principles of data minimization and pseudonymization in scientific use [60]. Even anonymized biosignals require data security and stewardship, which are crucial in this case; in the course of this thesis, all data were stored on password-protected institutional servers and processed according to the university's data management policies. This was done with a special focus on avoiding "function

creep," that is, employing health data for purposes other than those originally justified. [24] Finally, since the multimodal biosignals could potentially be indicative of poor sleep quality or health risks, the researchers need to be careful not to extrapolate from the results or make diagnoses without clinical support; the analyses here were intentionally restricted to the methodological signal-quality evaluation, which was along the lines of the wider ethical inference of responsible data interpretation. [61]

Lastly, this thesis was conducted with considerations of justice, transparency, reproduction, and data usage. Since SIESTA is an open database, this thesis contributes to open science, in which the emphasis is placed on being transparent in methods and results and on having maximum collective benefit. The FAIR principles (Findability, Accessibility, Interoperability, and Reusability) are increasingly being considered as ethical facial ideals in biomedical informatics to ensure that data usage advances knowledge and does not exploit participants. [62]

4. RESULTS

Signal quality is assessed by three complementary categories of methods: rule-based, statistical, and physiological. Each approach gave altogether different information so that a signal segment could be considered suitable for further analysis.

4.1 Data Integrity Analysis

The signal continuity and out-of-range analyses showed (Table 4.1) that the data quality varies for different biosignal modalities. The ECG had the highest signal integrity with just 0.06% signal discontinuity and 2.88% out of range, implying stable recording. The EEG channels EEG_6 and EEG_3 were better performers with discontinuities of 0.18% and 0.2% and out-of-range values of 11.06% and 10.62%, respectively. On the other hand, EEG_4 appeared to be of the least quality, with 0.31% discontinuity and 27.33% out-of-range values; in all probability, this pointed to problems with electrode placement or contact. The airflow signals recorded little discontinuity (1.61%) but as it was dimensionless in SIESTA, no acceptable range was possible to establish for out-of-range analysis. The average signal discontinuity across all channels was 0.41%, whereas the average out-of-range values amounted to 15.71%.

Signal	Discontinuity (%)	Out of Range (%)
Airflow	1.61	–
ECG	0.06	2.88
EEG_1	0.46	25.53
EEG_2	0.20	16.33
EEG_3	0.20	10.62
EEG_4	0.31	27.33
EEG_5	0.25	16.27
EEG_6	0.18	11.06
Mean	**0.41**	**15.71**

Table 4.1 *Signal continuity and out-of-range percentages across modalities*

4.2 Statistical Feature Analysis

Envelope Features: The standard deviations of the upper and lower envelopes showed remarkable consistency across EEG channels (from 0.2±0.03 to 0.26±0.05 for upper, and 0.2±0.03 to 0.24±0.04 for lower), suggesting stable signal amplitude characteristics. Airflow signals tend to show slightly lower values (0.14±0.07 and 0.15±0.08), whereas ECG envelopes with their characteristic sharp, periodic waveform display the lowest values (0.05±0.03 for both).

Channel	Upper mean ± SD	Lower mean ± SD
Airflow	0.14 ± 0.07	0.15 ± 0.08
ECG	0.05 ± 0.03	0.05 ± 0.03
EEG_1	0.26 ± 0.05	0.24 ± 0.04
EEG_2	0.22 ± 0.03	0.21 ± 0.03
EEG_3	0.20 ± 0.03	0.21 ± 0.03
EEG_4	0.25 ± 0.04	0.23 ± 0.04
EEG_5	0.22 ± 0.03	0.21 ± 0.03
EEG_6	0.20 ± 0.03	0.20 ± 0.03

Table 4.2 *Upper and lower envelope mean ± standard deviation values across channels.*

Distribution Characteristics: Most channels were characterized by very low skewness values (ranging between -0.038 and 0.14 for EEG, and -0.011 for airflow), which means that their distributions can be taken as approximately symmetric. On the other hand, the ECG was highly skewed (0.965±2.534), with a huge variability, which fits well with the pointedness of the QRS complexes in the ECG signal. The kurtosis analysis showed that the ECG obtained an extremely high kurtosis value (22.58±7.851), suggesting a highly leptokurtic distribution with very sharp peaks in the signal. The EEG channels had moderate values of kurtosis (6.95-7.49), and airflow had the lowest (5.504±1.053).

Channel	Skewness (mean ± SD)	Kurtosis (mean ± SD)
Airflow	-0.011 ± 0.309	5.504 ± 1.053
ECG	0.965 ± 2.534	22.58 ± 7.851
EEG_1	0.126 ± 0.183	7.491 ± 1.250
EEG_2	0.042 ± 0.149	7.164 ± 1.000
EEG_3	-0.038 ± 0.212	6.953 ± 0.990
EEG_4	0.140 ± 0.175	7.463 ± 0.766
EEG_5	0.057 ± 0.155	7.176 ± 1.009
EEG_6	-0.015 ± 0.198	6.989 ± 1.088

Table 4.3 *Skewness and kurtosis mean ± standard deviation values across channels.*

Entropy and Zero-Crossing Rate: The entropy and zero-crossing rate values for ECG (0.664±0.083 and 0.089±0.061, respectively) were the highest ones, which

shows the complexity and variability of its morphology. The EEG channels presented average values of entropy (from 0.478 to 0.568), and zero-crossing rate (from 0.114 to 0.132), and EEG_3 had the highest entropy (0.568±0.079). On the contrary, airflow signals exhibited the smallest values of entropy (0.201±0.056) and zero-crossing rate (0.023±0.018), which reflect their slow and smooth respiratory pattern.

Channel	Entropy (mean ± SD)	ZCR (mean ± SD)
Airflow	0.201 ± 0.056	0.023 ± 0.018
ECG	0.664 ± 0.083	0.089 ± 0.061
EEG_1	0.486 ± 0.086	0.114 ± 0.058
EEG_2	0.521 ± 0.127	0.120 ± 0.049
EEG_3	0.568 ± 0.079	0.132 ± 0.043
EEG_4	0.478 ± 0.103	0.115 ± 0.070
EEG_5	0.564 ± 0.075	0.129 ± 0.047
EEG_6	0.523 ± 0.130	0.125 ± 0.053

Table 4.4 *Entropy and zero-crossing rate mean ± standard deviation values across channels.*

4.3 Histogram Analysis of SQIs

To complement the descriptive statistics, normalized histograms were generated for each SQI, illustrating the distribution across modalities. For better comparison with the referenced paper [3], all the histograms were plotted in a mirror position.

Zero-Crossing Rate (ZCR) The histograms (Figure 4.1) provide ZCR distributions for EEG, ECG, and airflow channels. It was observed that ZCR varies more widely for EEG channels than for ECG and airflow, reflecting the greater oscillatory content of brain signals. ECG ZCR, on the other hand, clustered narrowly due to its periodic nature, while airflow showed sparse zero-crossing events, owing to the low-frequency nature of respiration cycles.

Entropy: Entropy values reflected the inherent complexity and variability of each biosignal type. ECG displayed the highest entropy (0.664 ± 0.083), indicating a greater degree of irregularity due to its rapid morphological changes between cardiac cycles. EEG signals showed moderate entropy levels (0.478–0.568), corresponding to their oscillatory but structured neural activity during sleep. EEG_3 recorded the highest entropy, implying slightly broader fluctuations in that channel's electrical activity. On the other hand, airflow showed the lowest entropy (0.201 ± 0.056), consistent with its smooth, periodic respiratory pattern and reduced signal unpredictability. Overall, the entropy distributions confirm that ECG signals contain the

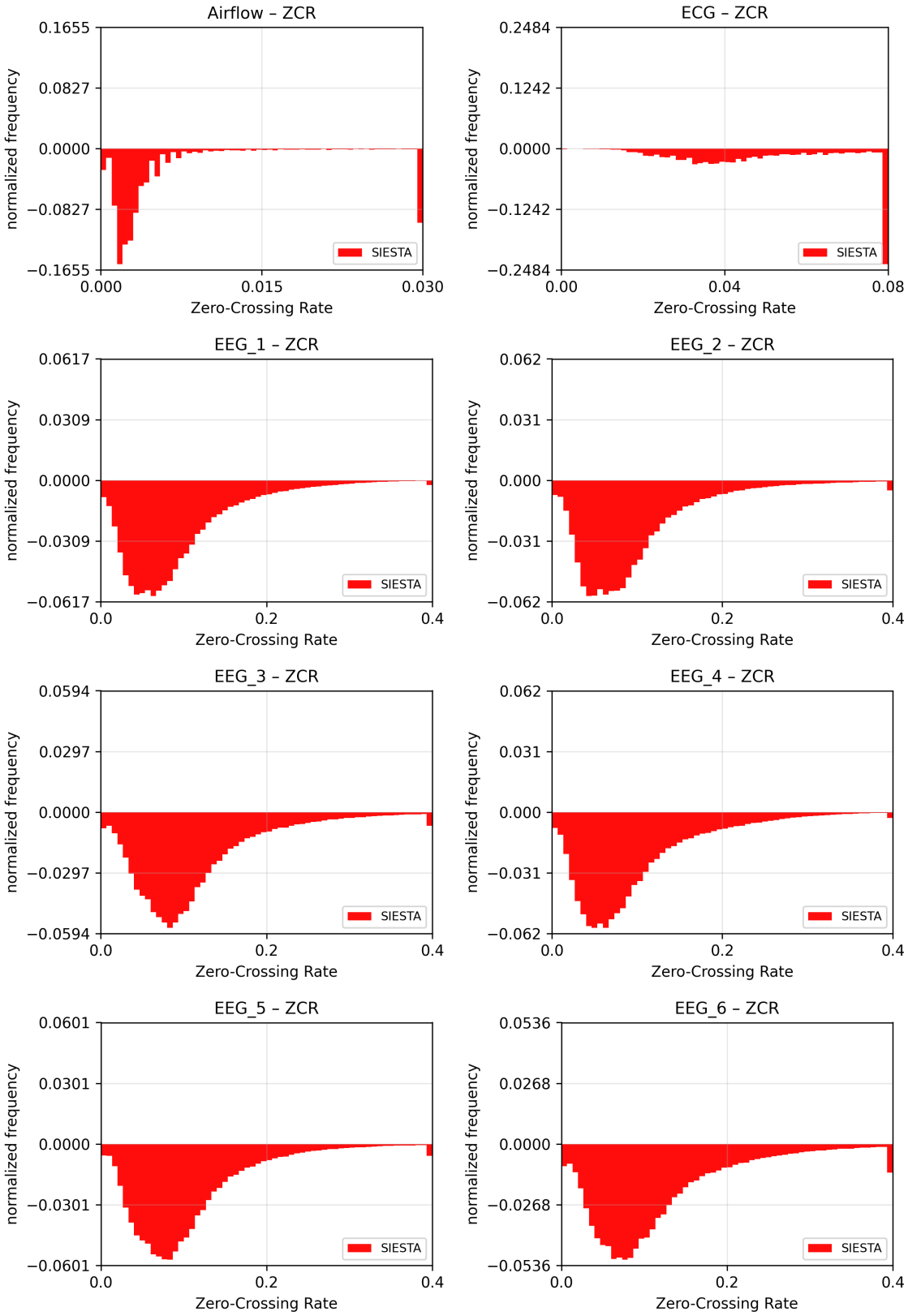

Figure 4.1 *Zero-crossing rate distributions across all channels*

most dynamic information, while respiratory signals remain the most regular, thus reflecting different physiological behaviors over the modalities.

Skewness Figure 4.3 illustrates the skewness histograms, in which EEG signals were centered around zero skewness, reflecting relatively symmetric oscillatory activity. The airflow likewise hovered around zero skewness, which corresponds with consistent breathing. However, the ECG skewness deviated noticeably, thus implying a waveform asymmetry either induced by artifacts or due to morphological variability of the QRS complexes.

Kurtosis Kurtosis results (Figure 4.4) indicated heavy-tailed distribution for ECG cases with values surpassing 20 for many instances, implying the presence of sharp peaks or transient noise bursts. EEG kurtosis was moderate, showing a more stable oscillatory structure. Airflow kurtosis values were mostly low, consistent with smooth respiratory cycles.

Envelope Standard Deviation (Upper and Lower) Figures 4.5 and 4.6 represent histograms for the standard deviation of upper and lower Hilbert envelopes. SQIs capture channel amplitude variability. EEG channels displayed a spread of values between 0 and 0.5, reflecting modulation of oscillatory rhythms. ECG and airflow envelopes were more concentrated, showing less variability. This distinction highlights modality-specific dynamics in amplitude fluctuations.

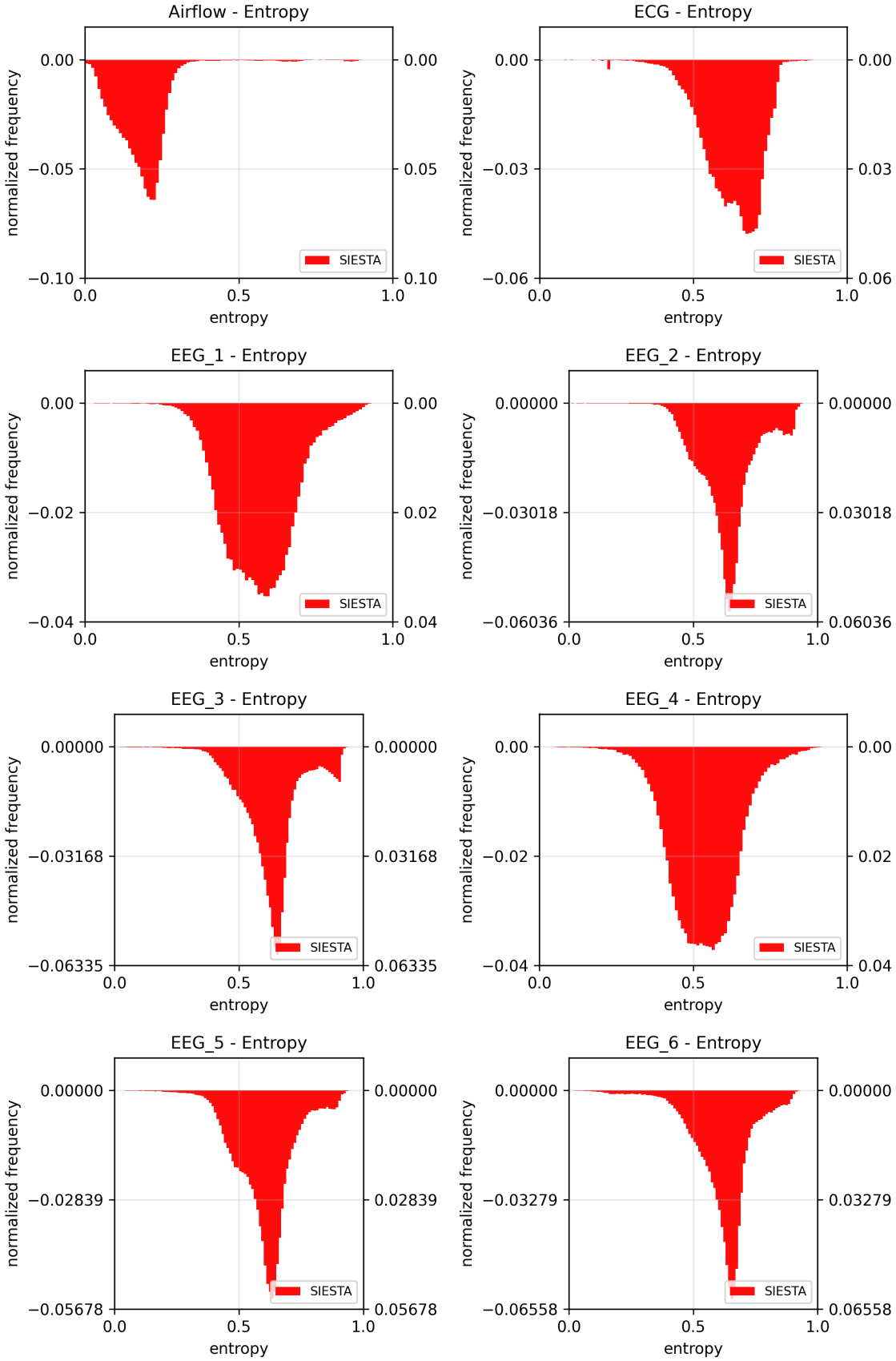

Figure 4.2 *Entropy distributions across all channels*

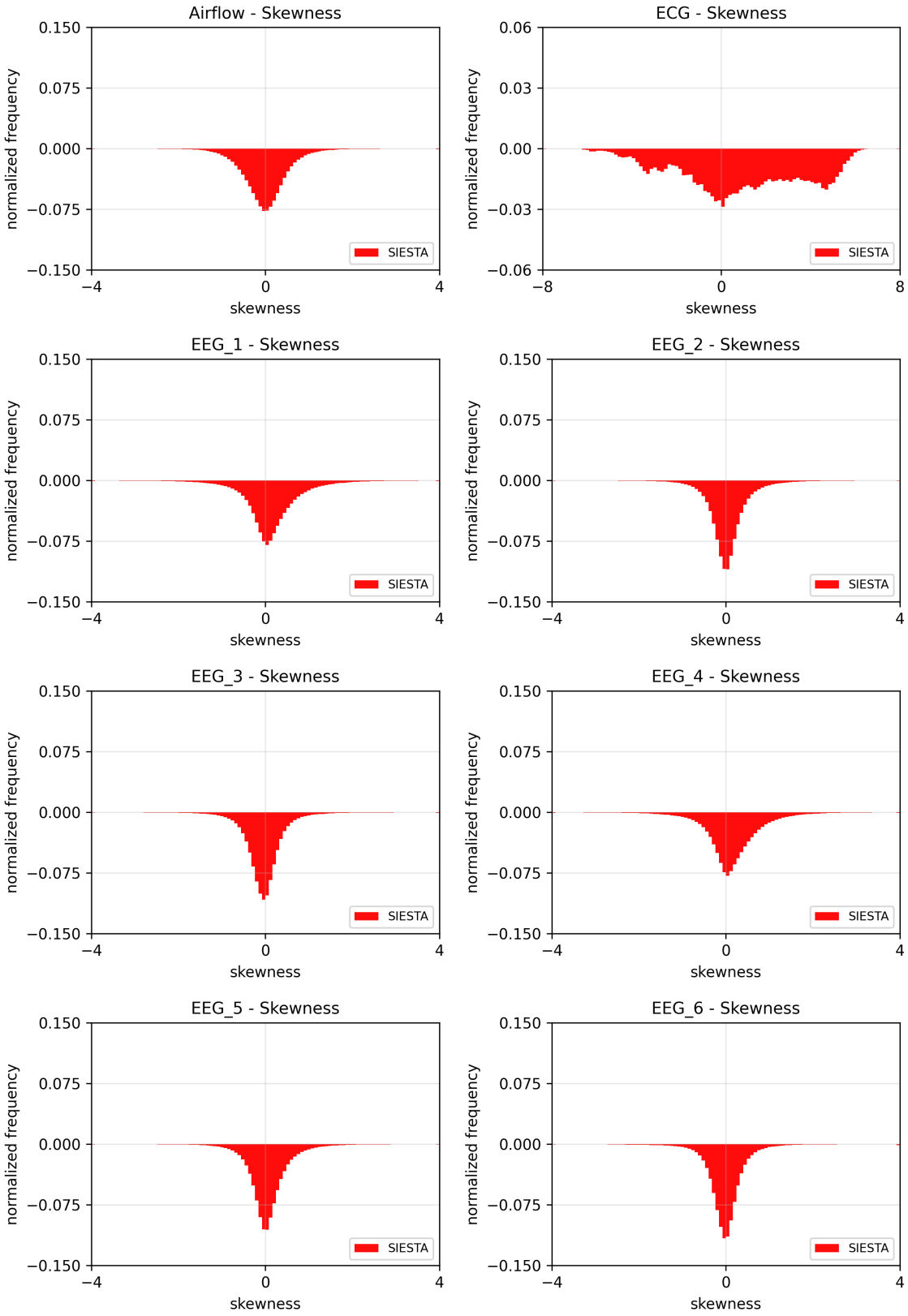

Figure 4.3 *Skewness distributions showing waveform symmetry across all modalities.*

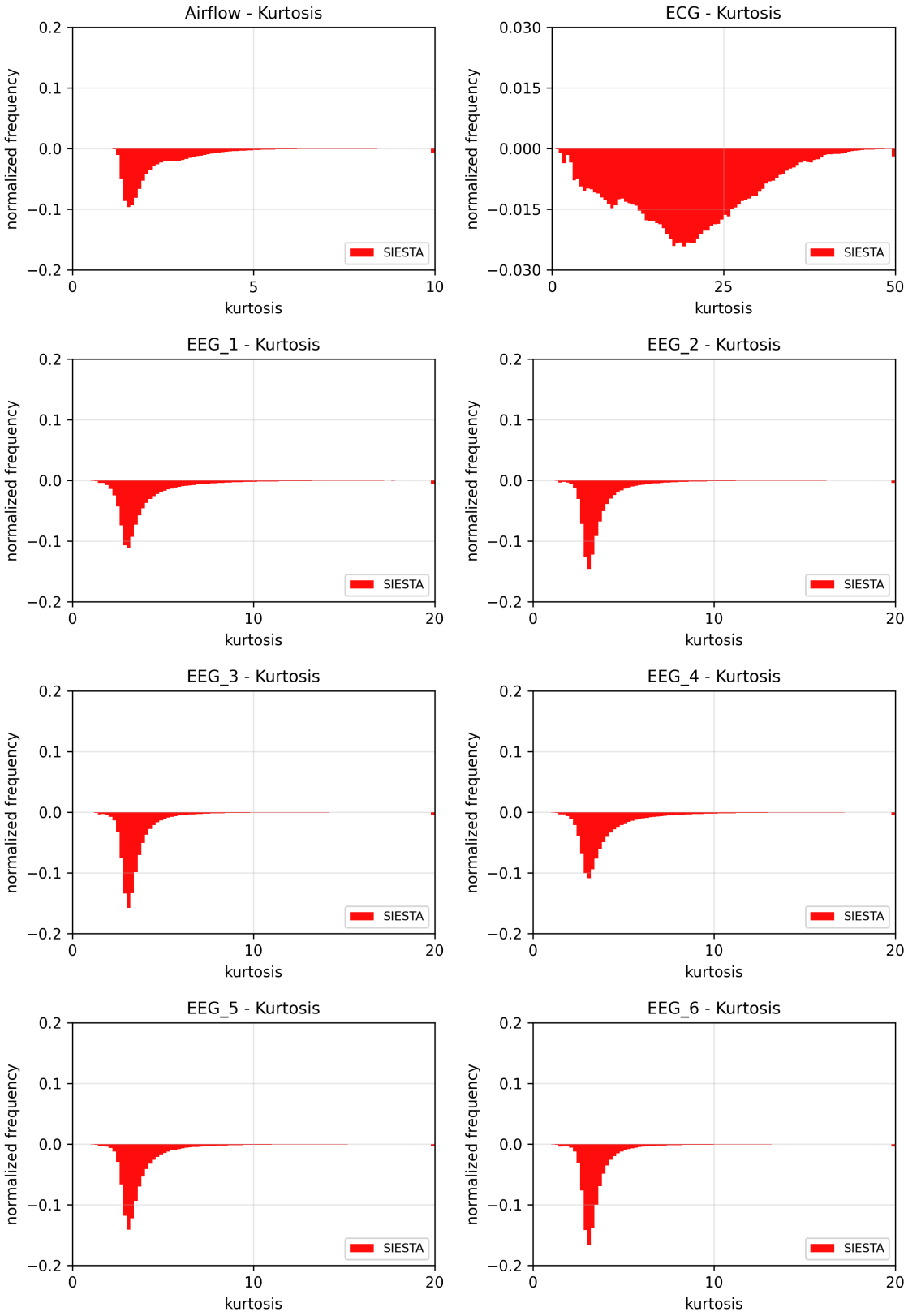

Figure 4.4 *Kurtosis distributions highlighting peakedness variations across all channels*

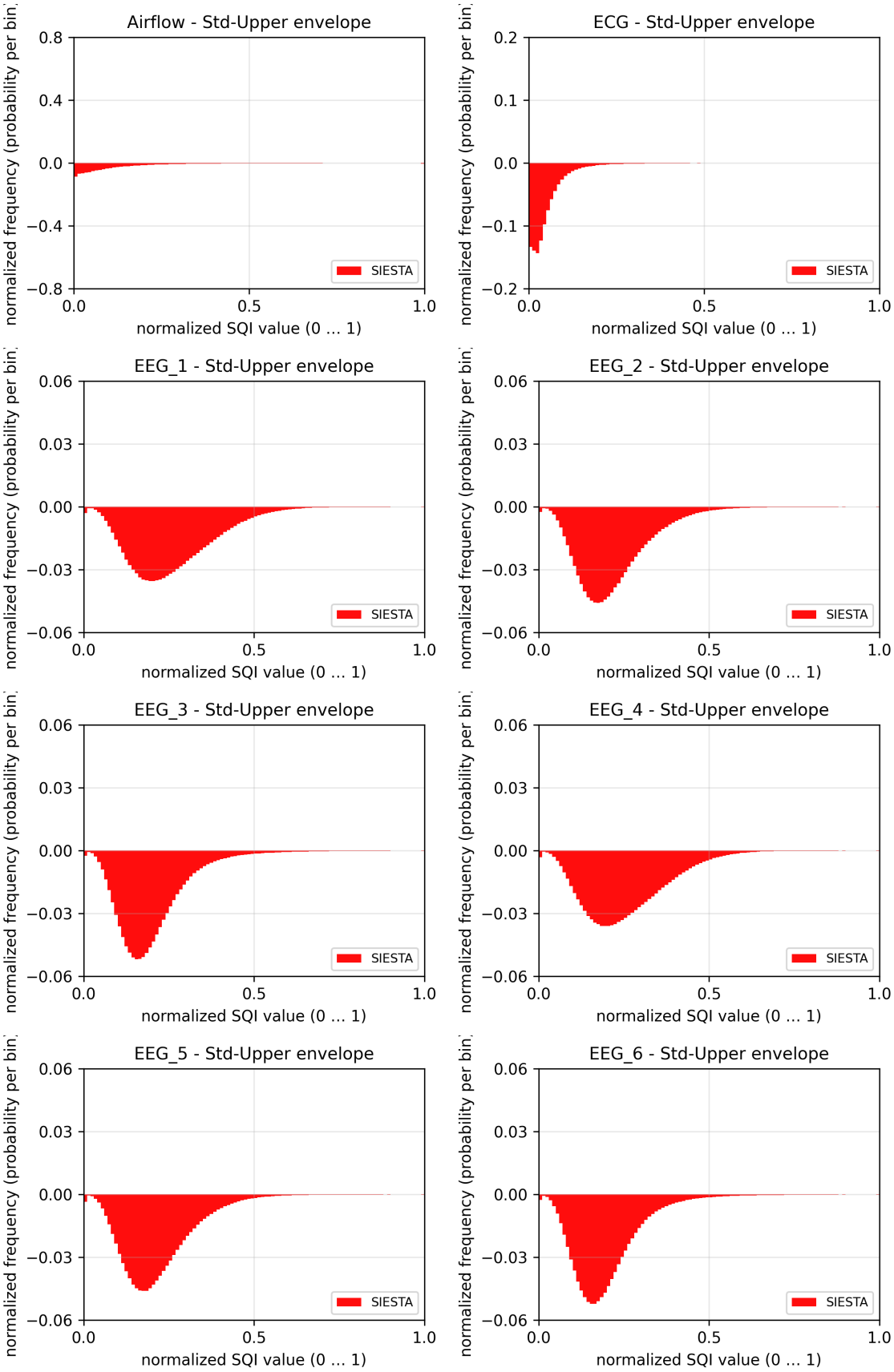

Figure 4.5 *Upper envelope standard deviation distributions across all channels.*

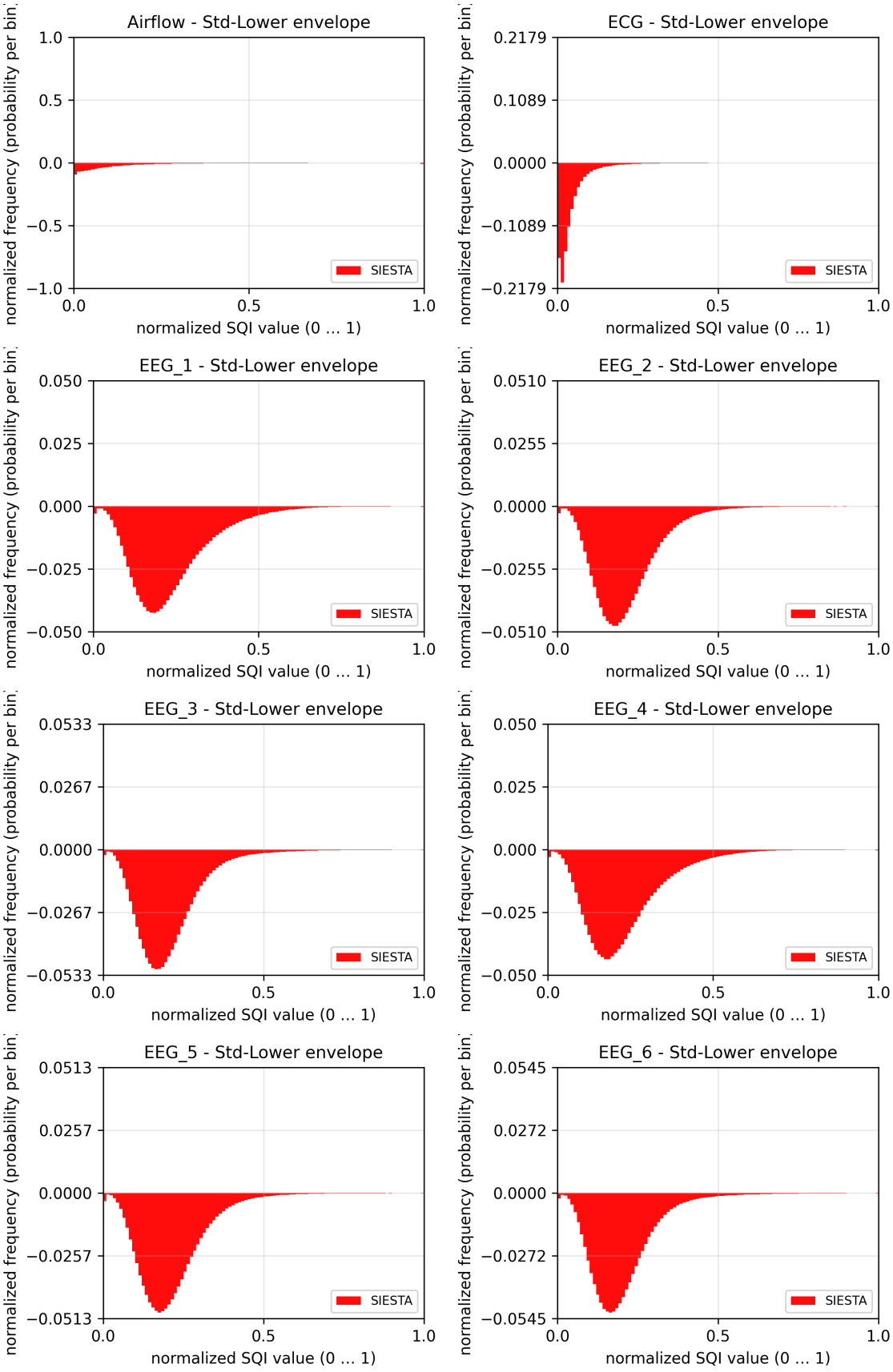

Figure 4.6 *Lower envelope standard deviation distributions across all channels.*

4.4 Overall Signal Quality Assessment

The physiological indicators were implemented in ECGs to ensure that the data retained biological plausibility. These indicators were based on average heart rate, maximum allowable R-R interval, and allowable beat-to-beat variability.

Findings in Table 4.5 proved that the vast majority of ECG windows in SIESTA did satisfy these physiological criteria. A minority of segments were rejected, almost always due to missed QRS detections in noisy intervals or brief electrode disturbances. This justifiably reflects the control sleep laboratory's usually high setting of ECG quality, with macro-averaged good signal percentages at 89.15% (mean of per-record means) and micro-averaged at 89.22% (pooled windows). The close agreement between macro and micro metrics indicates that the overall signal quality is good across all recordings and not affected by outliers or very fluctuating sessions.

Metric	Macro-Average (%)	Micro-Average (%)
Good-Quality Segments	89.15	89.22
Poor-Quality Segments	10.85	10.78

Table 4.5 *Performance metrics for ECG signal quality assessment (macro- and micro-averaged percentages).*

Although physiological methods were limited to ECG in this study, they provided an important clinical validation layer by confirming that signals classified as "good" were not just noise-free but also physiologically meaningful.

4.5 Signal Combination Performance

Pairwise Signal Combinations: There were varying degrees of signal quality in each pair of channels. Table 4.6 shows that the ECG-Respiratory combination had the highest quality score (95.14%). In practical terms, it means that more than 95% of the recorded segments were of sufficient quality for reliable physiological interpretation. This outcome is a clear indication of the benefit of the integration of modalities whereby two related physiological signals reinforce one another to give good reliability in data. Respiratory signals here, for example, helped to cover ECG short-term flaws while ECG was giving a reference in situations when respiratory signals were weak or bad. The pair's great performance is a reflection of their natural physiological link and at the same time emphasizes the strength of multimodal quality control over unimodal, therefore deemed the best complement. Similarly, of the EEG signals paired with the ECG, EEG_3 performed best (89.11%), followed

by EEG_6 (88.67%), with EEG_4 always giving the poorest result (72.47%). Good scoring also went to the EEG_3-Respiratory and EEG_6-Respiratory, which had 88.1% and 87.7% scores, respectively. EEG_4 tended to be the worst across all pairs (67.88% to 72.47%), confirming its inferior quality.

Signal Pair	Quality (%)	Interpretation
EEG_1 – ECG	74.25	Moderate pairing performance
EEG_2 – ECG	83.41	Good consistency and reliability
EEG_3 – ECG	89.11	High-quality pairing
EEG_4 – ECG	72.47	Weak quality due to EEG_4 noise
EEG_5 – ECG	83.47	Stable pair performance
EEG_6 – ECG	88.67	Strong reliability
EEG_1 – Respiratory	73.41	Moderate cross-modal consistency
EEG_2 – Respiratory	82.53	Good reliability
EEG_3 – Respiratory	88.10	Strong cross-modal complementarity
EEG_4 – Respiratory	71.53	Low reliability from EEG_4
EEG_5 – Respiratory	82.58	Good redundancy
EEG_6 – Respiratory	87.70	Stable overall performance
Mean ± SD	**82.49 ± 7.49**	

Table 4.6 *Pairwise signal-combination performance metrics showing the percentage of good-quality data for each channel pairing.*

Multimodal Combinations: Triple-signal combinations in Table 4.7 for (EEG, ECG, Respiratory) were slightly down in performance relative to the two-signal combinations. By user base, EEG_3-based combinations scored best (87.84% triplets, 89.11% EEG_3-ECG, 88.1% EEG_3-Respiratory), while EEG_4 combinations remained poor (71.34%-72.47%). The mean performance of all combinations tested was 81.20±6.63%, while for all available channel combinations it was 82.49±7.49%.

Signal Triple	Quality (%)	Interpretation
EEG_1 – ECG – Respiratory	73.21	Moderate triple-modal performance
EEG_2 – ECG – Respiratory	82.27	Good multimodal consistency
EEG_3 – ECG – Respiratory	87.84	Strongest overall performance
EEG_4 – ECG – Respiratory	71.34	Low reliability due to EEG_4 noise
EEG_5 – ECG – Respiratory	82.33	Stable multimodal performance
EEG_6 – ECG – Respiratory	87.44	High cross-modal complementarity
Mean ± SD	**81.20 ± 6.63**	

Table 4.7 *Triple-channel (EEG–ECG–Respiratory) performance metrics showing the percentage of good-quality data for each EEG channel.*

5. DISCUSSION

5.1 Summary of Findings

The thesis investigates EEG/ECG/respiratory signal quality in the SIESTA dataset using rule-based, statistical, and entropy-based SQIs, with an emphasis placed on multimodal analysis.

At the unimodal level, continuity failures were negligible in both EEG and ECG but were more prevalent in respiration due to the inherent technical flaws of airflow sensors. Out-of-range violations in EEG were rare but common in ECG, corresponding with motion artifacts or baseline drift. Table 4.1 summarizes these results across all channels, thus setting a quantitative baseline for data quality.

Statistical SQIs, such as skewness, kurtosis, and zero-crossing rate, provided further insight into the quality landscape of the signal. EEG distributions remained close to their expected ranges, while respiration showed wider tails due to its increased instability. Zero-crossing rates parted EEG from ECG and respiration in a physiologically meaningful manner.

Meanwhile, it was only during the multimodal analyses (Tables 4.4 and 4.5) that a percentage of data suitable for any pair and tri-modal combination of signals got measured and directly addressed the central research question by demonstrating that integrating different biosignals enhances robustness against noise and sensor failure. Pairwise combinations such as ECG–Respiration (95.14%) achieved the highest reliability, confirming that cardiac and respiratory rhythms mutually validate each other. EEG combinations with ECG or respiration also performed well (around 88–89%), though performance declined when including noisier EEG channels (notably EEG_4).

This confirms that a multimodal approach can “fill in” missing or corrupted information from one channel by utilizing reliable clues from another—something unimodal SQIs are not capable of achieving. The multimodal method successfully reduced false rejections (useful data discarded as bad) and false inclusions (noisy data accepted as good), providing a more balanced, physiologically consistent signal quality

evaluation. Thus, the research question was answered:

By applying synchronized, rule-based SQIs across multiple physiological modalities, it is possible to reliably assess biosignal usability in real-world, noisy sleep data.

5.2 Limitations

Several limitations must be acknowledged.

Scope of SQIs: Only continuity, amplitude range, and a few statistical SQIs were used. Other SQIs possibly would be considered: spectral measures or electrode impedance [63], for example, and were not addressed in this study due to the non-availability of such data.

Entropy thresholds: Entropy-based SQIs have been shown to be sensitive to parameter choices and dataset characteristics [64], so heuristically set at the minima of the histograms may not be valid for all datasets or populations. Automatic thresholding is still an open challenge and may require some kind of adaptive or machine learning-based approach.

Multimodal dependencies: Multimodal analysis hypothesized independence of signals. At the same time, artifacts could co-occur (e.g., a subject moving so that both the ECG and Respiration are affected), thereby decreasing the actual suitability percentages for any modality.

Computational constraints: The size of the dataset forced the implementation of windowed analysis and simplified implementations of detectors. This may render it less sensitive to very brief and subtle artifacts.

Limitations of ML Approaches to SQIs: The assessment process in the present study employs statistics- and rules-based approaches, where each method provides clear and interpretable measures of signal quality. Thus, machine learning methods were not brought forth, even though they are capable of capturing subtle multimodal patterns, pose challenges for interpretability and reproducibility in clinical use [65]. This choice reflects the study's focus on transparency and reproducibility, while leaving the integration of advanced ML techniques as a direction for future research.

5.3 Practical Implications

Despite these limitations, the findings are immensely relevant for practical application. Rule-based SQIs, such as continuity and amplitude range, are cheap enough to

compute so that they can be incorporated into real-time monitoring systems. High precision in quality detection of ECGs ensures that other fully automated down-line tasks, such as arrhythmia detection or heart rate variability analysis, do not get adulterated with spurious signals. Entropy and statistical SQIs act as secondary indicators by flagging suspicious segments to be considered for manual review. The demonstrated multimodal redundancy indicates that modern monitoring systems should promote the collection of complementary sets of signals because when one modality fails, it is quite often that others can make up for it. This principle may therefore steer research datasets and the clinical devices, where robust multimodal monitoring is increasingly valued.

5.4 Future Directions

Future research should expand analyses to examine other databases including, for example, VitalDB, to allow cross-setting validation between sleep studies and in perioperative recordings. Machine learning SQIs could well be used to detect very complex, multimodal artifact patterns that are missed by rule-based thresholds. Adaptive thresholding mechanisms should be explored that allow SQIs to interactively adjust to subject-specific physiology or noise situation. Multimodal fusion of SQIs across signals toward a single quality score might increase its interpretability and implementation in clinical practice. Finally, the downstream evaluation should relate how filtering based on SQIs influences various downstream measures, for example, sleep stage scoring accuracy, arrhythmia detection reliability, or respiratory event classification. Directly linking signal quality assessment to the ultimate clinical and scientific applications will be essential. Nonetheless, ML might be applied.

6. CONCLUSION

This thesis focuses on the problems and methods of evaluating biomedical signal quality within the multimodality scheme. The study proceeded from the SIESTA dataset in subjecting EEG, ECG, and respiratory signals to SQI assessments that were rule-based, statistical, or physiological, and then examined their efficacy from a unimodal or multimodal perspective.

The findings showed that none of the SQI methods could stand on its own. Rule-based checks identify obvious artifacts but risk discarding clean data. Statistical measures described signal distortion in greater detail and could be used effectively, but they do not offer direct physiological interpretation. Physiological tests validate in the clinical domain but are extremely modality-biased.

The multimodal setup that brought together SQIs from different sources was able to significantly reduce the number of false rejections and false inclusions, hence it led to higher quality decisions concerning the physiological coherence of data. The combination of signals in pairs—most notably the ECG and breathing signals—scored the highest performance and indicated that inter-system redundancy is a contributor to reliability. Thus, it was demonstrated that the synchronized, rule-based multimodal evaluation could satisfactorily determine the usability of biosignals even in the presence of noise, as in the case of sleep studies.

Overall, the findings verified that the quality assessment of biosignals done through multimodal evaluation was truly robust and much more reliable. ECG was at the top in terms of its overall reliability, having no discontinuity and very stable morphological features, while different EEG channels had variable performance based upon electrode positioning and quality of contact. However, the respiratory signals, with their smoother and less complex nature, were still very much informative in detecting and validating artifacts, hence were seen as a valuable modality. On the other hand, the statistical metrics—entropy, kurtosis, and zero-crossing rate—were the ones that showed the complexity and the stability of each of the modalities effectively, whereas the envelope features were the ones that picked up the amplitude consistency across the timeline.

The method, however, in spite of its advantages, is still limited because of the dependencies on arbitrary thresholds and features that are manually designed. It is not yet able to take advantage of adaptive machine-learning-based models that could learn the complex, nonlinear patterns of signal degradation. What remains to be done in the future is to apply data-driven methods to the existing framework and to validate them on datasets that are larger and more varied, including real-world, noise-prone data such as sleep studies.

To sum up, the research indicates that multimodal processing provides a powerful and interpretable framework for biomedical signal quality evaluation. Leveraging synchronized information from multiple physiological systems makes it possible to monitor more reliably, to give more assurance of the correctness of the data for analysis later on, and to set up intelligent, real-time quality control in clinical and wearable health applications.

BIBLIOGRAPHY

[1] G. D. Clifford and G. B. Moody, "Signal quality in cardiorespiratory monitoring," *Physiological Measurement*, vol. 33, no. 9, pp. E1–E10, 2012.

[2] T. Penzel, B. Kemp, G. Klosch, A. Schlogl, J. Hasan, A. Varri, and I. Korhonen, "Acquisition of biomedical signals databases," *IEEE Eng Med Biol Mag*, vol. 20, no. 3, pp. 25–32, 2001.

[3] E. Idrobo-Ávila, G. Bognár, D. Krefting, T. Penzel, P. Kovács, and N. Spicher, "Quantifying the suitability of biosignals acquired during surgery for multimodal analysis," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 5, pp. 250–260, 2024.

[4] U. Satija, B. Ramkumar, and M. S. Manikandan, "Review of signal processing techniques for evaluation of electrocardiogram signal quality," *IEEE Reviews in Biomedical Engineering*, vol. 11, pp. 36–52, 2018.

[5] J. G. Webster, ed., *Medical instrumentation: Application and design*. Wiley, 4th ed., 2009.

[6] G. Takla, J. H. Petre, D. J. Doyle, M. Horibe, and B. Gopakumaran, "The problem of artifacts in patient monitor data during surgery: A clinical and methodological review," *Anesthesia & Analgesia*, vol. 103, no. 5, pp. 1196–1204, 2006.

[7] R. P. Bartsch, K. K. L. Liu, A. Bashan, and P. C. Ivanov, "Network physiology: How organ systems dynamically interact," *PLoS One*, vol. 10, no. 11, p. e0142143, 2015.

[8] P. C. Ivanov, "The new field of network physiology: Building the human physiolome," *Frontiers in Network Physiology*, vol. 1, p. 711778, 2021.

[9] S. Schulz, F. C. Adochiei, I. R. Edu, R. Schroeder, H. Costin, K. J. Bär, and A. Voss, "Cardiovascular and cardiorespiratory coupling analyses: A review," *Philosophical Transactions of the Royal Society A*, vol. 371, no. 1997, p. 20120191, 2013.

[10] R. Pillalamarri and U. Shanmugam, "A review on EEG-based multimodal learning for emotion recognition," *Artificial Intelligence Review*, vol. 58, no. 5, p. 131, 2025.

[11] E. Lopez, A. Uncini, and D. Comminiello, "Phemonet: A multimodal network for physiological signals," in *2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI)*, pp. 260–264, IEEE, 2024.

[12] V.-R. Xefteris, M. Dominguez, J. Grivolla, A. Tsanousa, F. Zaffanela, M. Monego, S. Symeonidis, S. Diplaris, L. Wanner, S. Vrochidis, *et al.*, "A multimodal late fusion framework for physiological sensor and audio-signal-based stress detection: An experimental study and public dataset," *Electronics*, vol. 12, no. 23, p. 4871, 2023.

[13] D. Park, K. So, S. K. Prabhakar, C. Kim, J. J. Lee, J.-H. Sohn, J.-H. Kim, S.-H. Lee, and D.-O. Won, "Early warning score and feasible complementary approach using artificial intelligence-based bio-signal monitoring system: a review," *Biomedical Engineering Letters*, pp. 1–18, 2025.

[14] I. Jeong, W. G. Chung, E. Kim, W. Park, H. Song, J. Lee, M. Oh, E. Kim, J. Paek, T. Lee, *et al.*, "Machine learning in biosignal analysis from wearable devices," *Materials Horizons*, 2025.

[15] W.-B. Jiang, X. Fu, Y. Ding, and C. Guan, "Towards robust multimodal physiological foundation models: Handling arbitrary missing modalities," *arXiv preprint arXiv:2504.19596*, 2025.

[16] S. Frey, G. Spacone, A. Cossettini, M. Guermandi, P. Schilk, L. Benini, and V. Kartsch, "BioGAP-Ultra: A modular edge-AI platform for wearable multimodal biosignal acquisition and processing," *arXiv preprint arXiv:2508.13728*, 2025.

[17] M. Elgendi, R. Fletcher, Y. Liang, N. Howard, and N. H. Lovell, "The use of photoplethysmography for assessing hypertension," *NPJ Digital Medicine*, vol. 2, no. 1, p. 60, 2019.

[18] H. Kim, S. M. Lee, and S. Choi, "Automatic sleep stages classification using multi-level fusion," *Biomedical Engineering Letters*, vol. 12, no. 4, pp. 413–420, 2022.

[19] A. Bashan, R. P. Bartsch, J. W. Kantelhardt, S. Havlin, and P. C. Ivanov, "Network physiology reveals relations between network topology and physiological function," *Nature communications*, vol. 3, no. 1, p. 702, 2012.

[20] H.-C. Lee, Y. Park, S. B. Yoon, S. M. Yang, D. Park, and C.-W. Jung, "Vitaldb, a high-fidelity multi-parameter vital signs database in surgical patients," *Scientific Data*, vol. 9, no. 1, p. 279, 2022.

[21] R. Alam, D. B. Peden, and J. C. Lach, "Wearable respiration monitoring: interpretable inference with context and sensor biomarkers," *IEEE journal of biomedical and health informatics*, vol. 25, no. 6, pp. 1938–1948, 2020.

[22] European Parliament and Council, "Regulation (eu) 2016/679 of the european parliament and of the council ... (general data protection regulation)." `https://eur-lex.europa.eu/eli/reg/2016/679/oj`, 2016. Accessed: 2025-11-20.

[23] A. E. W. Johnson, T. J. Pollard, L. Shen, L.-W. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "MIMIC-III, a freely accessible critical care database," *Scientific Data*, vol. 3, p. 160035, 2016.

[24] M. Shabani and P. Borry, "Rules for processing genetic data for research purposes in view of the new EU general data protection regulation," *European Journal of Human Genetics*, vol. 26, no. 2, pp. 149–156, 2018.

[25] R. P. Bartsch, K. K. Liu, A. Bashan, and P. C. Ivanov, "Network physiology: how organ systems dynamically interact," *PloS one*, vol. 10, no. 11, p. e0142143, 2015.

[26] A. John, B. Cardiff, and D. John, "A review on multisensor data fusion for wearable health monitoring," *arXiv preprint arXiv:2412.05895*, 2024.

[27] R. Adarshi Rammohan, M. Meier, D. Küster, and T. Schultz, "EASELAN: An open-source framework for multimodal biosignal annotation and data management," *arXiv e-prints*, pp. arXiv–2510, 2025.

[28] D. Schouten, G. Nicoletti, B. Dille, C. Chia, P. Vendittelli, M. Schuurmans, G. Litjens, and N. Khalili, "Navigating the landscape of multimodal AI in medicine: a scoping review on technical challenges and clinical applications," *Medical Image Analysis*, p. 103621, 2025.

[29] G. Klösch, B. Kemp, T. Penzel, C. Becker, P. Rappelsberger, E. Trenker, G. Gruber, J. Zeitlhofer, B. Saletu, S. L. Himanen, D. Kunz, A. Varri, and G. Dorffner, "The SIESTA project polygraphic and clinical database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 51–57, 2001.

[30] B. Kemp, A. Värri, A. C. Rosa, K. D. Nielsen, and J. Gade, "A simple format for exchange of digitized polygraphic recordings," *Electroencephalography and Clinical Neurophysiology*, vol. 82, no. 5, pp. 391–393, 2000.

[31] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen, "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, pp. 446–460, 2014.

[32] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S.-A. A. Chen, "Neurokit2: A python toolbox for neurophysiological signal processing," *Behavior Research Methods*, vol. 53, no. 4, pp. 1689–1696, 2021.

[33] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.

[34] P. Hamilton, "Open source ECG analysis," in *Computers in cardiology*, pp. 101–104, IEEE, 2002.

[35] I. I. Christov, "Real time electrocardiogram QRS detection using combined adaptive threshold," *Biomedical engineering online*, vol. 3, no. 1, p. 28, 2004.

[36] R. Rosipal, "Clustering probabilistic sleep microstate curves: a functional data analysis approach," in *The 9th International Conference MEASUREMENT 2013 was organized by the Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia.*, p. 101, 2013.

[37] L. Kramer, C. Menon, and M. Elgendi, "ECGAssess: A python-based toolbox to assess ECG lead signal quality," *Frontiers in Digital Health*, vol. 4, p. 847555, 2022.

[38] C. Liu, X. Zhang, L. Zhao, F. Liu, X. Chen, Y. Yao, and J. Li, "Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1363–1374, 2018.

[39] S. Vaid, P. Singh, and C. Kaur, "EEG signal analysis for bci interface: A review," in *2015 fifth international conference on advanced computing & communication technologies*, pp. 143–147, IEEE, 2015.

[40] I. García-López, R. X. A. Pramono, and E. Rodriguez-Villegas, "Artifacts classification and apnea events detection in neck photoplethysmography signals," *Medical & Biological Engineering & Computing*, vol. 60, no. 12, pp. 3539–3554, 2022.

[41] R. Krishnan, B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE transactions on biomedical engineering*, vol. 57, no. 8, pp. 1867–1876, 2010.

[42] S. Rahman, C. Karmakar, I. Natgunanathan, J. Yearwood, and M. Palaniswami, "Robustness of electrocardiogram signal quality indices," *Journal of the Royal Society Interface*, vol. 19, no. 189, p. 20220012, 2022.

[43] J. Witteveen, P. Pradhapan, and V. Mihajlović, "Comparison of a pragmatic and regression approach for wearable EEG signal quality assessment," *IEEE journal of biomedical and health informatics*, vol. 24, no. 3, pp. 735–746, 2019.

[44] J. Li, K. Vatanparvar, L. Zhu, J. Kuang, and A. Gao, "Enhancement of remote PPG and heart rate estimation with optimal signal quality index," in *2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1–4, IEEE, 2022.

[45] X. Shi, K. Yamamoto, T. Ohtsuki, Y. Matsui, and K. Owada, "Non-invasive fetal ECG signal quality assessment based on unsupervised learning approach," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 1296–1299, IEEE, 2022.

[46] C. Dai, J. Wang, J. Xie, W. Li, Y. Gong, and Y. Li, "Removal of ECG artifacts from EEG using an effective recursive least square notch filter," *IEEE Access*, vol. 7, pp. 158872–158880, 2019.

[47] F. Liu, Y. Xu, and Y. Yao, "Highly efficient low noise solutions in ECG signals," in *Journal of Physics: Conference Series*, vol. 2246, p. 012030, IOP Publishing, 2022.

[48] M. H. Chowdhury and R. C. Cheung, "Reconfigurable architecture for multi-lead ECG signal compression with high-frequency noise reduction," *Scientific reports*, vol. 9, no. 1, p. 17233, 2019.

[49] F. Grosselin, X. Navarro-Sune, A. Vozzi, K. Pandremmenou, F. de Vico Fallani, Y. Attal, and M. Chavez, "Quality assessment of single-channel EEG for wearable devices," *Sensors*, vol. 19, no. 3, p. 601, 2019.

[50] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, "Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring," *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 832–838, 2014.

[51] S. Schulz, F.-C. Adochiei, I.-R. Edu, R. Schroeder, H. Costin, K.-J. Bär, and A. Voss, "Cardiovascular and cardiorespiratory coupling analyses: a review," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1997, p. 20120191, 2013.

[52] C. W. Whitney, D. J. Gottlieb, S. Redline, R. G. Norman, R. R. Dodge, E. Shahar, S. Surovec, and F. J. Nieto, "Reliability of scoring respiratory disturbance indices and sleep staging," *Sleep*, vol. 21, no. 7, pp. 749–757, 1998.

[53] S. Redline, R. Budhiraja, V. Kapur, C. L. Marcus, J. H. Mateika, R. Mehra, S. Parthasarthy, V. K. Somers, K. P. Strohl, D. Gozal, *et al.*, "The scoring of respiratory events in sleep: reliability and validity," *Journal of Clinical Sleep Medicine*, vol. 3, no. 02, pp. 169–200, 2007.

[54] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, Physiotoolkit, and Physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[55] G. Inuso, F. La Foresta, N. Mammone, and F. C. Morabito, "Brain activity investigation by EEG processing: wavelet analysis, kurtosis and renyi's entropy for artifact detection," in *2007 International Conference on Information Acquisition*, pp. 195–200, IEEE, 2007.

[56] L. Sörnmo and P. Laguna, *Bioelectrical signal processing in cardiac and neurological applications*. Academic press, 2005.

[57] R. Sameni and G. D. Clifford, "A review of fetal ECG signal processing; issues and promising directions," *The open pacing, electrophysiology & therapy journal*, vol. 3, p. 4, 2010.

[58] G. Zhao, Y. Zhang, and Y. Ge, "Frontal EEG asymmetry and middle line power difference in discrete emotions," *Frontiers in behavioral neuroscience*, vol. 12, p. 225, 2018.

[59] World Medical Association, "World medical association declaration of helsinki: ethical principles for medical research involving human subjects," *Jama*, vol. 310, no. 20, pp. 2191–2194, 2013.

[60] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A practical guide, 1st ed., Cham: Springer International Publishing*, vol. 10, no. 3152676, pp. 10–5555, 2017.

[61] L. Floridi and M. Taddeo, "What is data ethics?," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2083, p. 20160360, 2016.

[62] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, *et al.*, "The FAIR guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[63] E. S. Kappenman and S. J. Luck, "The effects of electrode impedance on data quality and statistical significance in erp recordings," *Clinical Neurophysiology*, vol. 121, no. 7, pp. 1234–1245, 2010.

[64] Y. Zhang, S. Wei, Y. Long, and C. Liu, "Performance analysis of multiscale entropy for the assessment of ecg signal quality," *Journal of Electrical and Computer Engineering*, vol. 2015, no. 1, p. 563915, 2015.

[65] A. Jovic, N. Frid, K. Brkic, and M. Cifrek, "Interpretability and accuracy of machine learning algorithms for biomedical time series analysis–a scoping review," *Biomedical Signal Processing and Control*, vol. 110, p. 108153, 2025.