Politecnico di Torino

Politecnico di Torino

Master's Thesis in Biomedical Engineering

A.a. 2025/2026

Graduation Session March 2026

Wearable-Oriented Unsupervised Clustering of Heart Rate Variability and Body Position Features for Phenotyping Sleep Disorders

Supervisor:
Gabriella Olmo
Umberto Mosca

Candidate:
Giulia Ciribilli

Abstract

REM Sleep Behaviour Disorder (RBD) is a parasomnia characterized by the loss of physiological muscle atonia during rapid eye movement sleep, leading to dream-enactment. Idiopathic RBD is strongly associated with α-synucleinopathies and frequently precedes the onset of motor or cognitive symptoms by several years, making it a valuable prodromal marker of neurodegenerative diseases, such as Parkinson's disease. Given the substantial healthcare burden of neurodegenerative diseases, early identification of at-risk individuals is essential to enable timely interventions, positioning RBD as a critical target for early detection. However, the current gold standard for RBD diagnosis is video-polysomnography, an expensive and resource-intensive procedure, unsuitable for large-scale screening. Heart Rate Variability (HRV), which reflects autonomic cardiac regulation, differs in individuals with neurodegenerative diseases compared to healthy controls (HC), positioning HRV-derived metrics as non-invasive approach for screening and early disease detection. This study investigates unsupervised clustering using a comprehensive feature set including time-domain, frequency-domain, nonlinear, complexity-based, and Recurrence Quantification Analysis metrics, extracted from overnight electrocardiography and body position signal, subsequently reduced through correlation analysis and Principal Component Analysis. K-means, Gaussian Mixture Models and Fuzzy C-Means (FCM) were applied on the 466-subject dataset and successfully identified three physiologically distinct clusters. Cluster 0 represents a more compromised phenotype, reflected in increased cardiac signal complexity and fractality alongside reduced RMSSD (Root Means Square of Successive Differences), signal entropy, and suppressed VLF (Very Low Frequency) activity. Cluster 1 exhibits an intermediate severity profile driven by sleep disordered breathing, displaying autonomic imbalance with changes in VLF and HF (High Frequency) power. Cluster 2 presents a less compromised group with higher entropy and HRV measures, coupled with lower signal complexity and fractality.

A secondary analysis of 150 RBD patients attempted to distinguish between severe and mild phenotypes. K–means achieved the best results with Accuracy of 0.57 and F1–score of 0.61. Examination of feature value distributions revealed that misclassified patients exhibited feature profiles and membership values closely resembling their algorithmically assigned clusters rather than their clinical labels, suggesting RBD progression follows a continuous spectrum, rather than a dichotomous partition into two discrete severity groups.

A third analysis incorporated both RBD patients and healthy controls (HC: 53, RBD1: 50, RBD2: 57). Clustering was performed with $k = 2$ to separate RBD from HC, and with $k = 3$ to distinguish HC from two RBD severity phenotypes.

Performance was modest for $k = 2$ (K-means: Accuracy = 0.55, F1-score = 0.66) and poor for $k = 3$ (FCM: Accuracy = 0.41, F1-score = 0.40). While algorithms identify geometric structure in the data, emergent clusters do not align with clinical classification of RBD severity subtypes. Visualization in principal component space confirms this discrepancy: predicted clusters appear separated in PC1–PC2 projections, yet true clinical labels exhibit substantial overlap, suggesting that PC1 captures the largest variance without reflecting clinically relevant distinctions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Clinical Background

1.1 Definition of Sleep

Sleep is a fundamental biological process that is crucial for human health and well-being. When normal sleep patterns are altered, whether in terms of sleep duration, sleep continuity, or architecture, this can lead to impairments in cognitive performance and cardiovascular function. These disruptions have also been associated with an increased risk of morbidity and mortality. From an electrophysiological perspective, sleep is a complex process that cyclically alternates between two states: rapid eye movement sleep (REM) and non-rapid eye movement sleep (NREM) [1].

1.2 Hypnogram and Sleep Stages

Sleep architecture refers to the duration and alternation of sleep stages throughout the night, and can be visualized using a hypnogram, a graph that shows the sequence and length of each sleep stage over time (Fig. 1.1). In healthy people, a complete sleep cycle typically lasts around 90 minutes, meaning that during a normal night's sleep, a person may go through up to five of these cycles. The first standardisation of sleep stages was provided by Rechtschaffen and Kales, subsequently updated by the American Academy of Sleep Medicine (AASM), resulting in the union of NREM stages 3 and 4 [1]. Sleep stages are classified according to characteristic patterns of electrical activity in the brain recorded by electroencephalography (EEG). Generally, higher frequencies with lower amplitudes are associated with more alert states, while lower frequencies with higher amplitudes characterize deeper sleep stages.

- **Wake (W)** Wakefulness is characterised by the prevalence of beta waves (13–21 Hz), which represent high-frequency activity. The amplitude of the EEG signal is low, and eye-movements are present.

- **Stage 1 (N1)** Transitional state between relaxed wakefulness to deeper stages, during which alpha waves are gradually replaced by slower, theta waves (4–7 Hz). It accounts for 5% of the total duration of sleep.

- **Stage (N2)** Features a slower heart rate and body temperature. K-complexes and sleep spindles can be found in the EEG. The former are high-amplitude delta waves, in which a negative polarisation is followed by a positive peak, reaching up to 100 µV . These occur around every minute and are frequently followed by sleep spindles. These latter, are rapid bursts of electrical activity, exhibiting frequency components in the range 12–15 Hz, and duration of 0.5 s. These oscillatory patterns are believed to support information processing, memory consolidation, and maintain sleep. It comprises up to 45% of total sleep time.

- **Stage 3 (N3)** Also known as slow-wave sleep (SWS), this stage is dominated by high-amplitude delta activity (0.5–4 Hz), reflecting the most quiescent phase, marked by minimal muscle activity and no eye movements, and it is strongly associated with efficient glymphatic clearance. It accounts for up to 25% of total sleep time.

- **REM stage** Features low amplitude waves, recalling the awake state. It is associated with vivid dreaming accompanied by complete skeletal muscle atonia, with the exception of the eyes and diaphragm. REM sleep can be further divided into two microstates: tonic and phasic. The tonic phase represents the quieter portion with minimal and slow eye movements (amplitude lower than 25 µV in a 4-second range), and sustained muscle atonia. In contrast, the phasic phase is marked by the characteristic rapid eye movements, low-amplitude EEG activity, irregular heart rate, and erratic breathing patterns, leading to a metabolic rate increase of up to 20%. It comprises up to 45% of total sleep time.

1.3 Glymphatic System

Sleep is thought to facilitate the removal of metabolic waste products from the brain through a network of perivascular channels where cerebrospinal fluid (CSF) flows, helping to clear potentially harmful substances. This mechanism, known as glymphatic clearance, operates most efficiently during deep NREM sleep and is significantly diminished when awake. Recent evidence suggests that glymphatic function contributes to the clearance of amyloid-β (Aβ) during slow-wave sleep. If Aβ is not adequately cleared, it can build up and form senile plaques, which are associated with an elevated risk of neurodegenerative disorders. In particular, very

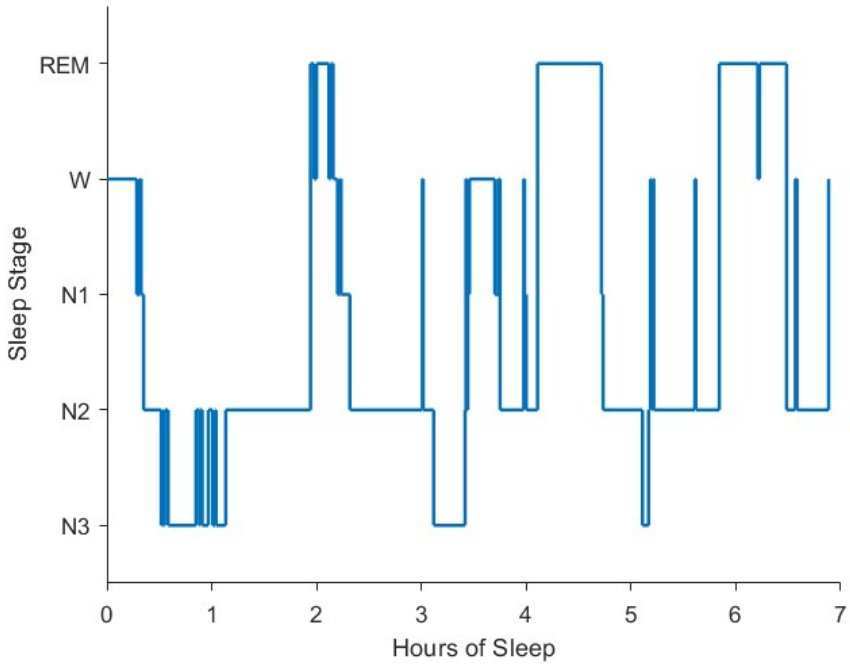


Figure 1.1: Example of an hypnogram of $Pz2$.

slow EEG oscillations (<1 Hz) have been correlated with improved Aβ clearance, while higher-frequency activity has been linked to greater Aβ deposition [1].

1.4 Sleep disorders

Sleep disorders affect quantity and quality of sleep and are often associated with other medical conditions, mental, or mood disorders. They can be classified as follows [2]:

Hypersomnia disorders: Excessive daytime sleepiness, commonly associated with:

- **Obstructive Sleep Apnea Syndrome (OSAS):** A condition marked by snoring and repeated episodes of upper airway blockage during sleep, resulting in decreased oxygen levels in the blood and disrupted sleep quality.

- **Narcolepsy:** A chronic neurological disorder characterized by sudden and uncontrollable sleep attacks.

Insomnia-related and sleep-movement disorders:

- **Insomnia:** A disorder characterized by difficulty falling asleep and/or staying asleep, resulting in fragmented sleep, commonly linked to symptoms of depression.
- **Restless Legs Syndrome (RLS):** A condition involving uncomfortable tingling in the legs that creates an urge to move, that interferes with sleep continuity.
- **Periodic Limb Movements (PLM):** Characterized by repetitive movements of the lower limbs during sleep.

Circadian rhythm sleep-wake disorders: A mismatch between a person's natural sleep–wake schedule and the environmental light–dark cycle.

Parasomnias: Abnormal behaviours occurring during sleep, traditionally categorized as:

- **NREM parasomnias:** Confusional arousals, sleepwalking, and sleep terrors.
- **REM parasomnias:** A primary example is REM sleep behaviour disorder (RBD), characterised by loss of physiological muscle atonia and the presence of abnormal motor activity (kicking, punching, vocalisations), often reflecting dream enactment.

1.5 REM Sleep Behaviour Disorder

REM sleep behaviour disorder is a parasomnia characterized by the loss of the normal skeletal muscle atonia that physiologically occurs during REM sleep. This loss of atonia allows the enactment of dream-related motor behaviours and vocalizations that reflect the ongoing dream content. Dream enactment episodes may include punching, kicking, flailing movements, and potentially injure themselves or their bed partners, accompanied by vocalizations such as shouting, screaming, laughing, or incomprehensible speech. During these episodes, the eyes typically remain closed. RBD affects males, with an approximate male-to-female ratio of 2:1 [3]. Unlike NREM parasomnias, postural control is normally suppressed during REM sleep, therefore individuals with RBD generally do not ambulate during episodes. Dream recall is typically absent unless the patient awakens during an episode. When aroused mid-episode, patients usually demonstrate prompt reorientation and remember their dream content [3].

The current diagnostic framework of RBD relies on a combination of clinical assessment and objective polysomnographic evidence. The clinical history is

frequently obtained from bed partners who witness the nocturnal episodes. The polysomnographic signature of RBD is REM sleep without atonia (RSWA), defined by excessive tonic and/or phasic EMG (electromiography) activity during REM sleep. Definitive diagnosis therefore requires video-polysomnography (vPSG), which includes EEG and EMG recordings. Quantification of RSWA is performed using standardized, amplitude-based scoring systems, with the most widely adopted being the Montréal, SINBAR, and Mayo criteria [3]. Despite its diagnostic value, vPSG carries significant burdens across multiple levels. For patients, spending an entire night in a hospital setting is stressful and frequently associated with the first-night effect, potentially compromising the representativeness of the recorded night. For clinicians, the manual scoring of a full-night is a time-consuming task. For the healthcare system, each study occupies a dedicated room and considerable personnel resources for an entire night.

The neurophysiological basis of REM atonia involves a specialized brainstem network responsible for suppressing spinal motor output during REM sleep. Central to this mechanism are the pontine subcoeruleus/sublaterodorsal nucleus (SLD) and ventral medullary regions, which exert descending inhibitory pathways to suppress motor neurons. RBD is believed to result from dysfunction within these brainstem circuits. More fundamentally, RBD reflects a failure of the specific inhibitory neuronal populations that physiologically suppress motor activity during REM sleep.

A key clinical implication of RBD is its close association with α-synucleinopathies, including Parkinson's disease (PD), dementia with Lewy bodies (DLB), and Amyotrophic Lateral Sclerosis (ALS). When RBD occurs in its idiopathic form (iRBD), when symptoms cannot be explained by medications, substances, or as a comorbidity in other pathologies, it carries a substantial and well-documented risk of conversion to an α-synuclein–related disorder. This supports the interpretation of iRBD as a prodromal manifestation of neurodegeneration, which may precede the onset of motor or cognitive symptoms by several years, up to a decade [3]. Two main hypotheses have been proposed [3]:

- **Hypothesis 1** suggests that RBD may facilitate neurodegenerative processes through disruption of sleep architecture. By reducing sleep efficiency and diminishing slow-wave sleep, RBD could compromise essential restorative mechanisms. Chronic impairment of these clearance pathways may lead to progressive accumulation of toxic protein aggregates, potentially accelerating the neurodegenerative cascade.

- **Hypothesis 2** suggests that RBD does not cause neurodegeneration, but rather represents one of its earliest expressions: α-synuclein pathology may begin in brainstem regions, damaging the regions responsible for REM atonia, producing RBD before the later emergence of the motor and cognitive features.

1.6 Neurodegenerative Diseases

Neurodegenerative diseases (ND) are typically diagnosed only when the first motor symptoms appear, a stage that already corresponds to extensive and largely irreversible neuronal loss. Early diagnosis is therefore of utmost importance: intervening at a presymptomatic stage offers the opportunity to slow disease progression, preserve quality of life, and improve long-term patient outcomes. Beyond the individual benefit, early identification carries significant implications for the healthcare system, as patients with advanced neurodegeneration require continuous and resource-intensive care, both in clinical settings and at home, representing a substantial and growing economic burden, especially in ageing population.

Parkinson's Disease

Parkinson's disease typically manifests after the age of 60 and it is charcterized by the progressive degeneration of dopaminergic neurons in the substantia nigra pars compacta, which disrupts basal ganglia circuits, critical in motor control. Motor symptoms usually become clinically evident only after an estimated loss of 70% of dopamine-producing neurons. A second hallmark is the presence of Lewy pathology, intraneuronal aggregates composed of misfolded α-synuclein protein, known as Lewy bodies. Accordingly to he Braak staging, early stages (I–II) are associated with premotor features including autonomic and olfactory dysfunctions. Stages III–IV coincide with the emergence of classic motor symptoms and involve sleep disturbances. Advanced stages (V–VI) are characterized by extensive neuropsychiatric and cognitive dysfunction. Current diagnostic criteria remain anchored to motor manifestations, particularly the presence of bradykinesia, tremor or rigidity. PD is often classified into subtypes: tremor-dominant, typically associated with slower progression and less disability, and non–tremor-dominant, which tend to be more rapidly progressive and functionally impairing. At present, PD has no curative therapy. Treatment is primarily symptomatic: dopaminergic therapy (levodopa) or, in refractory cases, surgical intervention (lesional surgery or Deep Brain Stimulation) [1].

Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis is a rapidly progressive and ultimately fatal neurodegenerative disease characterized by the degeneration of upper and lower motor neurons. This process leads to progressive muscle weakness, rigidity, sporadic muscle twitches, ultimately leading to speech alterations, and respiratory failure. Current treatment is largely supportive, focusing on maintaining patient autonomy and quality of life [1].

Chapter 2

Material and Methods

2.1 Data Pre-Processing

The dataset comprised 513 patients who underwent a PSG, from which ECG (electrocardiography) and body-position signals were extracted. The selected ECG signal and the body-position signal, encoded as a categorical variable ranging from 0 to 5, where each integer corresponds to a specific sleep position, undergo a standardized pre-processing pipeline.

- **Resampling:** the ECG and position signals were resampled to 256 Hz and 32 Hz, respectively.
- **Trimming:** both signals were trimmed to align with the hypnogram length. This trimming included the removal of the initial and final epochs labelled as not scored (stage "0"). In particular, if the first and/or last 100 epochs contain a majority of not-scored epochs the corresponding portion are discarded up to the first epoch with a score different from 0.

ECG signal pre-processing

The following steps describe the ECG signals pre-processing procedure. Fig. 2.1 illustrates an example comparing the raw and pre-processed signal.

- **Unit Conversion:** ECG amplitudes were converted in µV to ensure comparability across patients.
- **Filtering:** A zero-phase Chebyschev Type 1 bandpass filter between [8-25] Hz was applied. A relatively strong filtering was intentionally adopted, as the objective was to enhance R-peak detectability rather than preserve the full ECG morphology.

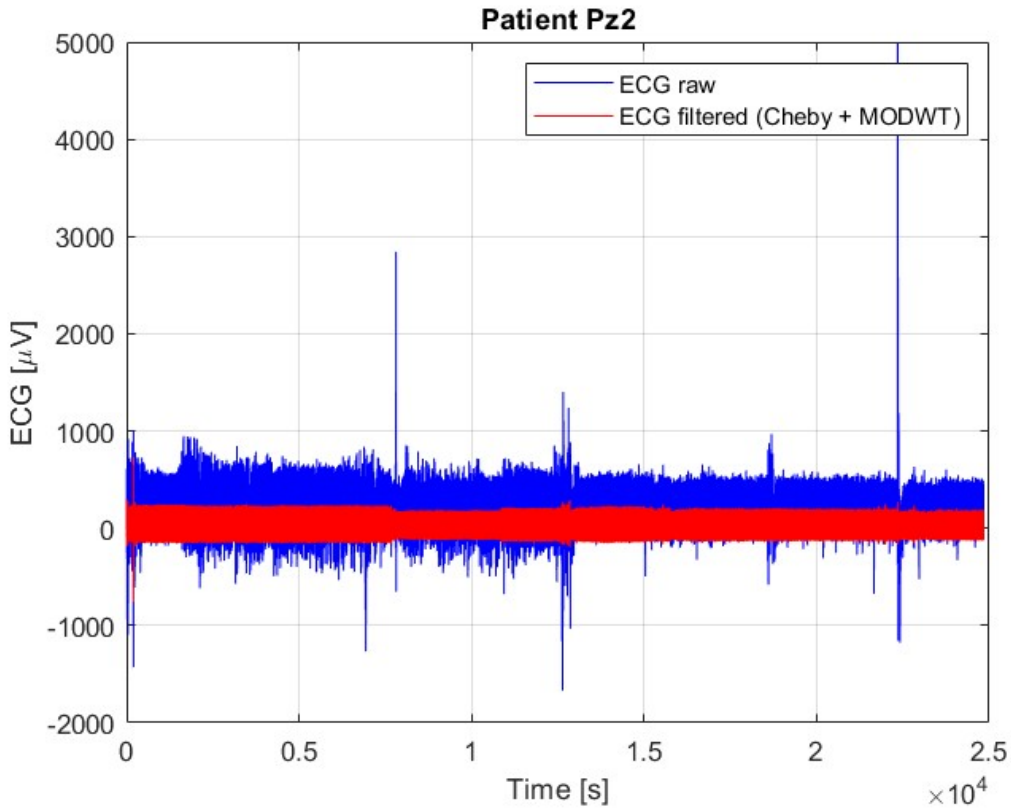


Figure 2.1: Comparison between raw ECG signal and the corrisponding pre-processed signal of $Pz2$.

- **Denoising:** The signal was decomposed into multiple frequency sub-bands using the maximal overlap discrete wavelet transform (MODWT). For each of the decomposition levels, the noise level was estimated and a level-specific soft universal threshold was computed. This threshold was then applied to the wavelet coefficients at that level, shrinking them towards zero rather than discarding the entire level. As a result, all decomposition levels were preserved while noise-dominated coefficients were attenuated [4].

- **Artefact Removal:** Artefact detection was performed on 30-second epochs by comparing a global standard deviation, computed over the entire epoch, with a local standard deviation estimated over a short sliding window. Since artefacts manifest respectively as flat segments in the filtered ECG and saturated segments in the raw ECG, epochs in which the local variability was lower than the global variability were flagged as artefactual and removed (Fig. 2.2).

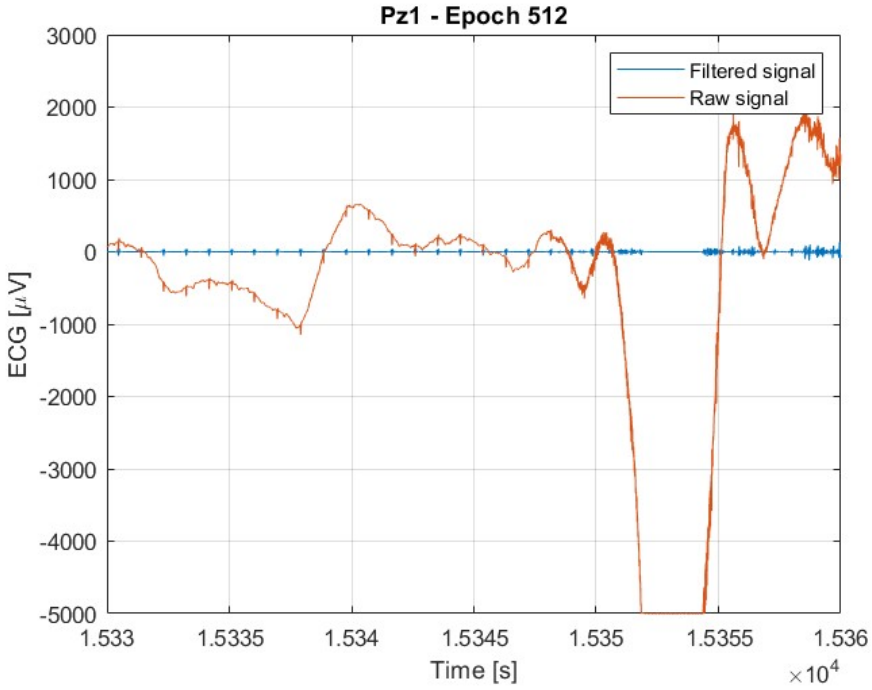


Figure 2.2: Epoch 512 of the ECG signal of subject $Pz1$: ECG affected by an artifact, showing the appearance of the artifact in both the raw and the filtered signal.

2.2 Extraction of RR intervals

Peak Detection:

Hilbert Double Envelope Method (HDEM) was employed for R-peaks detection. The ECG signal was first transformed into its analytic representation using the Hilbert transform, from which upper and lower amplitude envelopes were derived, as illustrated in Fig. 2.3. The intersections of these two envelopes defined discrete intervals along the signal. Within each interval, the candidate R-peak was identified as the point of maximum ECG amplitude [5].

Missed Peak Refinement:

ECG segments, where two consecutive detected peaks were separated by more than 2 s, were flagged as containing potential missed beats and additional candidate peaks were identified using the *findpeaks* function.

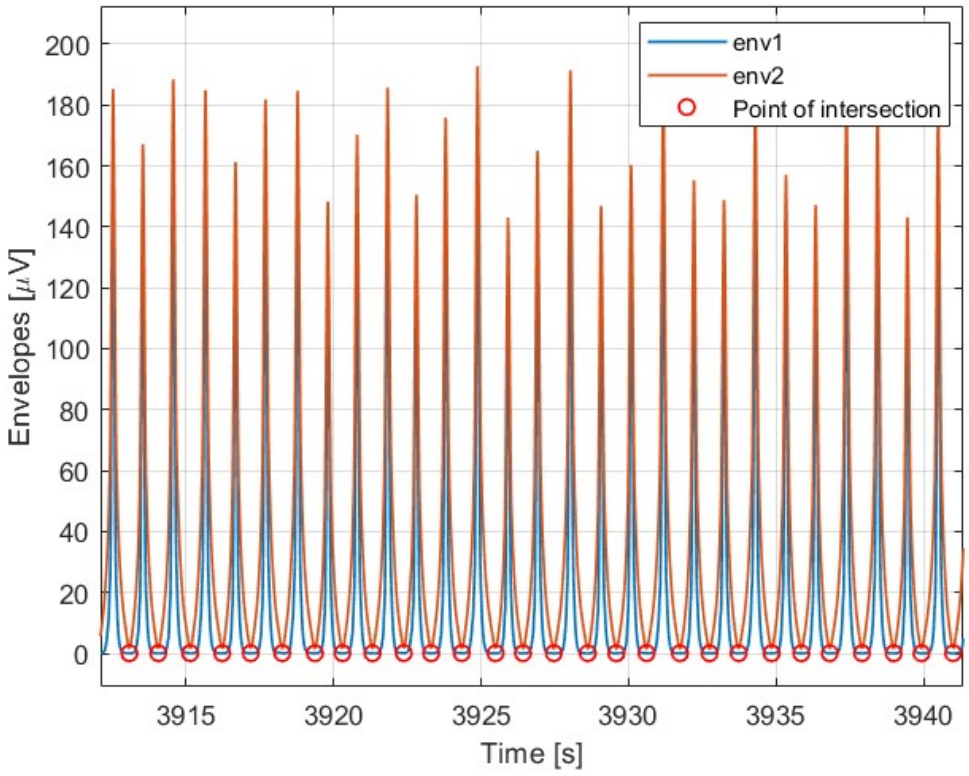


Figure 2.3: Lower, upper envelopes and their intersection points computed on a short segment of the ECG signal from subject $Pz2$.

Refractory Period Peak Pruning:

Candidate peaks occurring within less than 400 ms of each other were considered phisiologically implausible and were therefore pruned to reduce false detections. The peak with the largest amplitude was retained.

RR-interval Post-Processing:

Following RR interval extraction, a post-processing step was applied to improve detection accuracy. Intervals exceeding 2 seconds were flagged as potential missed beats, prompting a secondary search for R-peaks within the corresponding ECG segment using *findpeaks*. Simultaneously, spurious low-amplitude peaks occurring between two larger neighboring peaks were removed, as they are unlikely to represent true R-waves.

2.3 Feature Extraction

Four distinct feature categories were extracted from the processed signals. The first encompassed Heart Rate Variability (HRV) features, spanning time-domain, frequency-domain, and nonlinear metrics derived from the RR interval time series. The second consisted of mixed features, designed to jointly encode information from both the HRV signal and body position data. The third category included Recurrence Quantification Analysis (RQA) and Joint Recurrence Quantification Analysis (JRQA) features, which capture the underlying dynamical structure of the signals through phase-space reconstruction. The fourth drew on position-derived features, extracted directly from the body position signal. The feature engineering process was designed to encompass both established metrics, known to reflect autonomic dysregulation and reduced cardiac variability, well-documented hallmarks of iRBD, and less conventional measures, introduced with the intent of surfacing latent patterns that traditional approaches may fail to capture.

Features HRV

Heart rate variability represents a non-invasive marker of cardiovascular autonomic function. Autonomic impairment and reduced HRV are frequent features in both iRBD and α-synucleinopathies [6, 7, 8, 9, 10]. For these reasons HRV features were extracted, including time-domain indices, which directly quantify the variability of RR intervals, frequency-domain indices, which characterize the cardiovascular autonomic function, by estimating the amount of signal power within specific frequency bands, and nonlinear indices, which quantify the predictability, regularity and uncertainty present in the evolution of the signal [11]. In particular, the cardiac signal was segmented into epochs of different lengths (30 s, 1 min, 5 min, 10 min, and 15 min), and HRV features were computed for each epoch. Table 2.1 shows the feature extracted in each category. Subsequently, descriptive statistical measures (mean, standard deviation, 25^{th} percentile, 75^{th} percentile, kurtosis, skewness) were calculated across all epochs for each feature. As a result, a final feature vector was obtained for each patient, consisting of 105 HRV features, each described by six statistical measures.

Mixed Features

The following features were designed to integrate information derived from the HRV signal with body position data, in particular to investigate the relationship between posture changes and HRV dynamics and evaluate autonomic adaptation to postural transitions. The first analysis focused on how the HRV signal behaves in correspondence with position transitions. This was quantified by measuring the number of R-peaks deviating from the local median HRV value, and their

Category	Features	Description
Time	meanHR, SDNN, RMSSD	Average heart rate, Standard deviation and Root Means Square of Successive Differences of intervals
Frequency	pLF, pHF, LFHFratio, VLF, LF, HF	FFT used to compute the spectral density function of the interpolated RR tachogram
Non-Linear	ApEn, CD, DFA, SD1, SD2, SD1SD2ratio, KC_05, KC_1, KC_2, LZC_05, LZC_1, LZC_2	Approximated Entropy, Correlation Dimension, Detrended Fluctuation Analysis, Standard deviations along the identity line and its perpendicular axis of the Poincaré map and their ratio, Kolmogorov Complexity and Lempel-Ziv calculated on binary sequences where values are set to 1 when the beat-to-beat variation exceeds rispectively 0.5%, 1%, or 2%

Table 2.1: Extracted HRV features grouped by domain.

temporal localization relative to postural changes. The underlying hypothesis, illustrated in Fig. 2.4, is that such deviations tend to cluster around position transitions rather than being uniformly distributed. Out-of-median RR peaks were detected by computing a local median on a sliding window of 11 beats and marking an RR interval as outlier if it deviates by more than ±20%. Table 2.2 reports a description of the features calculated. An higher occurrence of such peaks during posture transitions can be interpreted as noise introduced during signal acquisition due to subject movement. However, R peaks affected by position changes were not removed, as the movements responsible for them may themselves be diagnostically relevant in the context of sleep disorders. The second analysis focuses on capturing autonomic adaptation to postural changes, by selecting two signal windows for each detected body position change: one immediately before and one after the transition. Within each window, standard HRV features were calculated, including meanHR, SDNN, RMSSD, pLF, pHF, LF/HF ratio, VLF, LF, and HF. The features considered for this analysis were the deltas between the HRV metrics computed in the post-change window and those computed in the pre-change window. The underlying hypothesis is that healthy subjects exhibit

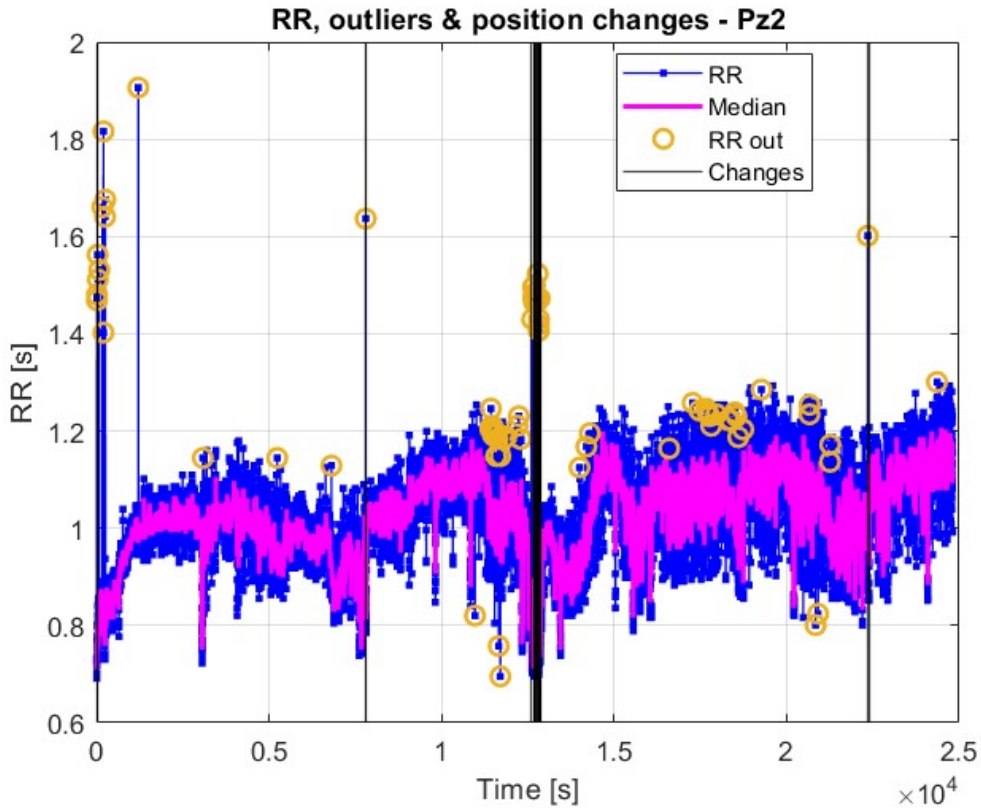


Figure 2.4: Temporal distribution of RR-interval outliers (deviations from the local median) overlaid with body position changes (vertical lines) to evaluate potential alignment.

more effective autonomic adjustment to postural transitions, reflected by larger HRV changes, whereas patients with impaired autonomic function show reduced adaptability and greater difficulty responding to body position changes [9].

Features RQA

Many real-world phenomena are predictable because similar situations tend to evolve in similar ways and because certain situations recur over time. Recurrence, defined as the tendency of a system to return to previously visited states, is therefore a fundamental property that supports memory, experience-based forecasting, and provides a powerful lens for studying dynamical behavior in fields ranging from physics and engineering to finance, medicine, and brain dynamics [12]. In general, a dynamical system is defined by a phase space, a notion of time (continuous or discrete), and an evolution law that governs how the system changes over time.

Feature	Description
`frac_out_global`	Fraction of outlier peaks over the total number of peaks in the signal.
`out_in_change_windows`	Fraction of outlier peaks occurring within windows centered around position changes.
`out_outside_change_windows`	Fraction of outlier peaks occurring outside the windows around position changes.
`frac_bad_around_change`	For each position change, a local window is defined and the fraction of outlier peaks over all peaks is computed.

Table 2.2: Features extracted that measure the number of R peaks that deviate from the median HRV in relation with body position changes information.

At any instant, the system can be represented by a state vector, a vector whose components are the variables that describe the system at that time. The temporal evolution of the system can then be interpreted as a sequence of state vectors that trace a trajectory in the phase space, as time progress [12]. Recurrence Quantification Analysis is a nonlinear framework designed to characterize the dynamics of complex systems by quantifying the recurrence of states in a reconstructed phase space. The phase space reconstruction relies on time-delay embedding, defined by the embedding dimension m, which determines the dimensionality of the reconstructed trajectory, and the time delay τ, which specifies the temporal separation between embedded components. Recurrences are identified by comparing state vectors: if the distance falls below a predefined threshold ϵ, a recurrence point is assigned. The collection of all recurrence points forms the recurrence plot (RP), a binary matrix that visualizes repeating patterns. Diagonal structures indicate deterministic behavior, vertical structures reflect laminar states, and isolated points represent stochastic dynamics [12, 13].

Joint Recurrence Quantification Analysis (JRQA) extends RQA to the analysis of interactions between two distinct dynamical systems. A joint recurrence plot (JRP) must be distinguished from the cross recurrence plot (CRP) that highlights the relationships between two systems by detecting the occurrence of similar states across their trajectories. However, CRP is not appropriate because directly computing distances between state vectors with different phase-space dimensions is not meaningful. JRP helps to overcome this limitation, since it compares two systems separately in their own phase space. As a matter of fact, it is constructed as the element-wise product of the individual recurrence plots of the two systems: joint recurrence point occurs only when both systems exhibit a recurrence at the same

time. JRQA therefore captures coordinated or synchronized dynamics between heterogeneous signals [12, 14]. Kodama et al. [14] demonstrated the applicability of JRQA to mixed data types by studying coordination between continuous body motion signals and categorical language data. In this framework, they transformed categorical signals into numeric sequences and add temporal information (time stamp), then they applied JRQA methods to explore speech-motion coordination coupling during a rap performance.

In this work, RQA was computed on the RR interval series RRi and on the body position signal. For such categorical data temporal information was added at the same sampling rate as RRi, before applying RQA. Then JRQA was obtained by combining the two recurrence plots through a logical AND. From each analysis, RQA descriptors, namely recurrence rate (REC), determinism (DET), laminarity (LAM), trapping time (TT), average diagonal line length (L), divergence (DIV), and Shannon entropy of diagonal line lengths (ENTR), were extracted from the 5 minutes epochs [12]. Table 2.3 reports RQA metrics and their description. Finally, for each feature, summary statistics were computed across epochs.

- **RQA on RRi:** RQA relies on the reconstruction of the system's phase space, so the selection of appropriate embedding parameters, m and τ, is a crucial step. Parameter estimation was performed once per subject on the first 10 minutes of the RRi series, ensuring stable and subject-specific embedding settings that were subsequently applied to all epochs. The time delay τ was determined using the average displacement (AD) method. For a fixed embedding dimension m, different candidate delays τ were tested. For each τ, the average Euclidean distance between the delayed coordinates and the first coordinate in the reconstructed phase space was computed. When τ is too small, the state vectors are very similar, resulting in redundant embedded components; as τ increases, they become more different and provide additional information. The selected delay τ corresponded to the smallest value for which the increase of the average distance becomes small (below 40% of its maximum), indicating that the embedded components were no longer redundant. Once the delay τ was determined, the embedding dimension m was estimated using the gamma test, which evaluates whether adding dimensions improves phase-space reconstruction. For each candidate m, the method computed gamma bar, which represents the residual output variance estimated from nearest-neighbor predictions in the embedded space. The optimal embedding dimension was then chosen as the smallest m for which gamma bar no longer decreased substantially (relative decrease $< 10\%$), indicating that higher dimensions would not meaningfully improve the phase-space reconstruction. Using the estimated parameters, the RRi series was converted into state vectors by time-delay embedding. Then, the Euclidean Distance between state vectors was computed to have the distance matrix D. The recurrence plot RP_{RR} is

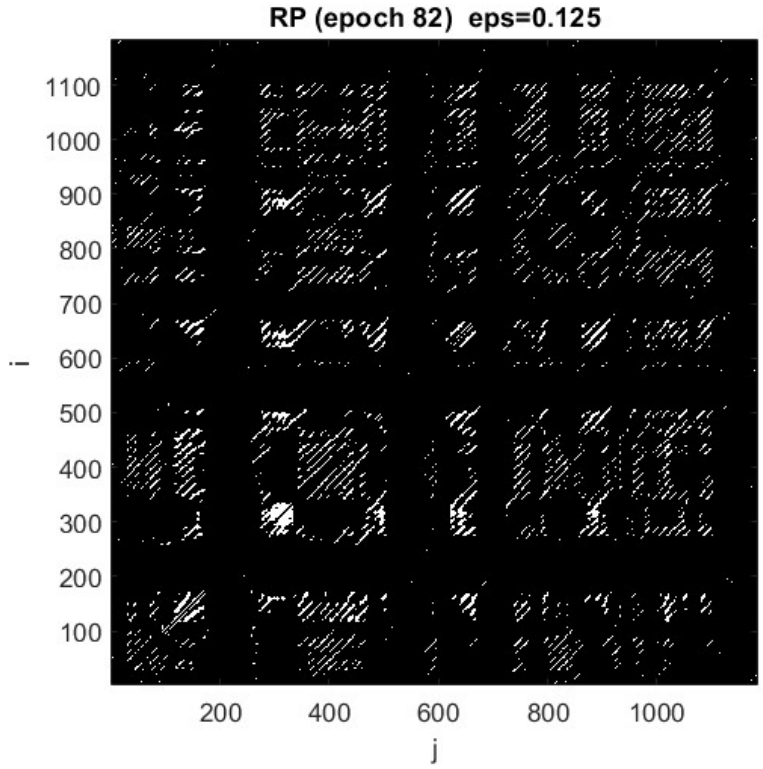


Figure 2.5: Example of a RP calculated on one epoch of the RRi signal from subject $Pz2$.

finally obtained by thresholding the matrix D: a recurrence point is set to 1 when $D_{ij} \leq \epsilon$, and to 0 otherwise. In this work, the threshold ϵ was not fixed a priori, but was selected so that the REC is 10%. From RP_{RR} the set RQA measures reported in Table 2.3 was extracted. In addition permutation entropy (Perm) and the largest Lyapunov exponent (LE) were calculated. Fig. 2.5 reports an example of RP_{RR} obtained from a certain epoch of the RRi signal from subject $Pz2$.

- **RQA on position signal:** RQA was also applied to the body position signal, which is a categorical time series. Following the approach proposed by Kodama et al. [14] for categorical RQA, the categorical sequence was first paired with its temporal information. The recurrence was defined whenever the system revisited the same categorical state.

- **JRQA between RRi and position:** JRP was computed as the element-wise conjunction of the two recurrence plots and the same RQA descriptors were extracted.

Feature	Description
REC	Recurrence rate: percentage of ones in the recurrence plot.
DET	Determinism: percentage of recurrence points forming diagonal lines.
LAM	Laminarity: tendency of the system to remain in laminar states (vertical/horizontal line structures).
TT	Trapping time: average length of vertical lines.
L	Average diagonal line length.
DIV	Divergence: $1/L_{\max}$, where $L_{\max}$ is the maximum diagonal line length.
ENT	Shannon entropy of the diagonal line length distribution.

Table 2.3: RQA descriptors extracted from recurrence plots.

Position Features

Sleep posture and nocturnal mobility have emerged as relevant markers of sleep physiology and neurodegenerative processes. Gnarra et al. [15] demonstrated that PD patients spend a significantly larger proportion of the night in the supine position and exhibit fewer position changes compared with healthy controls, with such alterations already detectable in prediagnostic stages. They associated supine sleep with reduced glymphatic clearance efficiency: lying supine may hinder venous outflow due to gravity, promoting venous congestion and reduced drainage, which could lead to limited clearance. Moreover, time spent in the supine position, especially during specific sleep stages such as REM or slow-wave sleep (N3), has been reported to correlate with the presence of RBD and with disease severity [16]. In this project, body position signal was exploited to quantify nocturnal mobility by computing the number of position changes per hour and the time duration of interval between consecutive position changes.

2.4 Normalization

The extracted features were normalized to ensure that anomalous values do not degrade the resolution. For each feature, a set of descriptive statistical metrics, such as variance, skewness, kurtosis, and outlier prevalence, was calculated to select the most appropriate scaling method for each feature. Features exhibiting approximately symmetric, light-tailed distributions were standardized using z-score normalization, whereas those characterized by a high number of outliers or heavy-tailed distributions were normalized using a robust scaler, which relies on the median and interquartile range rather than the mean and standard deviation. This approach reduces the influence of extreme values. Strongly skewed features were

first transformed using the Yeo–Johnson power transformation and then normalized with robust scaler, while features with near-zero variance across subjects were removed, as they do not contribute discriminative information.

2.5 Dimensionality Reduction

Correlated features were identified using both Pearson and Spearman coefficients. Two features were considered correlated if at least one of the two coefficients exceeded a predefined threshold. For each pair of correlated features, the one exhibiting the higher loading on the first principal component (PC1) was retained, while the other was removed in order to reduce informational redundancy, under the assumption that higher PC1 loadings reflect a stronger contribution to the variance structure of the data. After that, PCA was applied to the entire feature set to analyze the variance explained by the principal components. A scree plot was used to evaluate the contribution of the leading components and to guide the selection of the most informative features. Accordingly, the five features with the highest loadings on the first two principal components were selected. The reduced feature set was then projected onto a two-dimensional PCA space for visualization, allowing an initial inspection of subject distribution and potential clustering tendencies. PCA therefore served a dual purpose: it enabled effective dimensionality reduction of the dataset and provided a two-dimensional embedding space for visualization and interpretation of underlying data patterns.

2.6 Clustering

Clustering is the task of grouping similar objects together with respect to a distance or a similarity measure, by analyzing unlabelled data without any a priori information on group membership. Unsupervised learning has emerged as an effective alternative to supervised approaches, motivated by several key advantages [17]:

- Absence of labelled data in real life scenarios: reliable annotations are often unavailable in monitoring settings or costly to obtain, making supervised learning impractical.
- Inter individual variability and transitional states: strong differences between individuals and the presence of gradual or intermediate states represent a difficult challenge for supervised models, as they rely on rigid, predefined classes.
- Limitations of supervised models: supervised techniques require clearly defined classes and expert annotations, typically available only in controlled laboratory

environments, thus preventing their applicability in autonomous and home monitoring systems.

- Potential for anomaly detection: clustering methods are able identify outliers and transitional patterns.

The objective of this phase is to identify subgroups that may reflect different clinical groups (healthy controls versus patients affected by sleep disorders or other conditions). A wide range of clustering techniques has been proposed in the literature and is commonly grouped into partitioning, hierarchical, grid-based, density-based, model-based, and constraint-based methods [18]. In this project, several algorithms were explored, considering both hard and soft clustering strategies. Hard clustering methods, such as K-means, assign each observation to exactly one cluster. In contrast, soft clustering approaches, such as fuzzy c-means and Gaussian mixture models, allow each data point to belong to multiple clusters simultaneously, with different membership probabilities [19].

- **K-means** partitions the data into k clusters by minimizing the sum of squared Euclidean distances between data points and their assigned cluster centroids. Starting from a random initialization, data points are iteratively reassigned to the nearest centroid, and cluster centroid are updated accordingly. The process continues until the centroids stabilize or a maximum number of iterations is reached [17].

- **Gaussian Mixture Model** assumes that the data are generated by a combination of several gaussian distributions. Each gaussian represents one cluster, characterized by its own mean and variability. Instead of assigning each data point to a single cluster, GMM assigns a probability to each point of belonging to each gaussian cluster.

- **Fuzzy c-means** extends the k-means algorithm by allowing each observation to belong to multiple clusters with different membership probabilities, thereby capturing uncertainty and gradual transitions between groups. The algorithm iteratively updates both the cluster centroid vectors and the membership matrix until convergence of the objective function is achieved or a predefined maximum number of iterations is reached [19].

Global Clustering Analysis – Derivable 1

In this first step, clustering was applied to a subset of subjects, considering only patients with age $> 5^{\text{th}}$ percentile and $< 95^{\text{th}}$ percentile, in order to eliminate too young and too old patients. The result is a total of 466 patients. To ensure comparability across algorithms, the number of clusters was determined using elbow

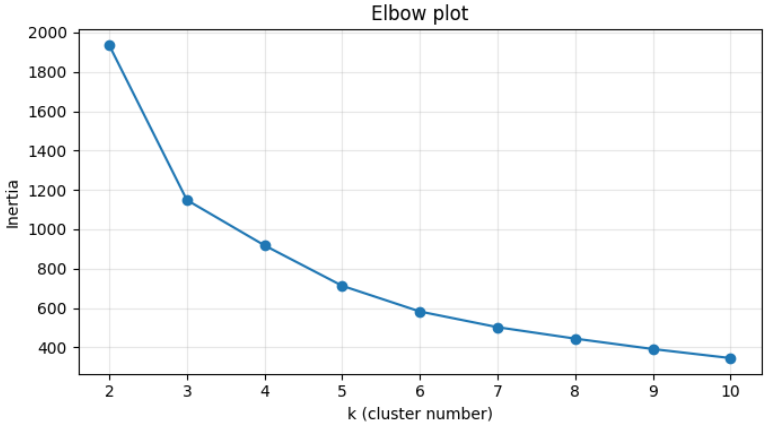


Figure 2.6: Elbow Method: inertia as a function of the number of clusters k.

method, which indicated k=3 as the optimal value, as shown in Fig. 2.6. The primary aims of this step were:

1. Analyze and visualize how patients are partitioned across clusters.
2. Investigate whether patients in the same cluster share common physiological and clinical characteristics and delineate a clinical profile for each cluster.
3. Evaluate the feature values in each cluster in relation to the clinical profile of the patients belonging to that cluster.

Clustering of Isolated RBD – Derivable 2

In this step, K-means and FCM were applied to the subgroup of patients diagnosed with RBD, excluding those RBD patients concurrently affected by OSA and/or PLM, resulting in a group of 150 subjects. The aim was to evaluate whether unsupervised clustering could distinguish between patients with a more severe form of the disorder (RBD1) and those with a milder form (RBD2), characterized by dissociative phenomena limited to specific sleep epochs.

Clustering on RBD and Healthy Controls – Derivable 3

To the previously selected RBD patients, a group of clinically classified healthy controls (HC) was added. These controls were selected based on the absence of RBD, OSA, PLM, and insomnia, and were not receiving any pharmacological treatment. Subsequently, only participants aged between 35 and 75 years were retained from the combined dataset. Three clustering algorithms, K-means, GMM,

and FCM, were applied to this cohort to evaluate their ability to distinguish between diseased and healthy individuals. The number of clusters was set initially to $k = 2$ to differentiate between HC and RBD patients, and then to $k = 3$ to distinguish among HC, RBD subtype 1, and RBD subtype 2. The resulting clusters were visualized in a two-dimensional space defined by the first two principal components. These cluster assignments were then compared with the visualization of the true clinical labels in the same principal component space, in order to evaluate the agreement between unsupervised clustering results and the ground-truth classification.

2.7 Performance Metrics

Global Clustering Analysis – Derivable 1

Due to the absence of ground-truth labels the performance of the clustering approaches was assessed using internal validation metrics, which emphasizes properties such as the separateness and compactness of the clusters. Specifically, the Silhouette Score (SS), the Davies–Bouldin index (DB), and the Calinski–Harabasz index (CH) were employed to evaluate the quality of the resulting partitions in terms of intra-cluster compactness and inter-cluster separation.

- **Silhouette Score** measures how similar each sample is to the elements of its assigned cluster compared to those belonging to the other clusters. The coefficient ranges from -1 to 1, where values close to 1 indicate a high degree of affinity between samples and their corresponding cluster, reflecting a good clustering structure.

- **Davies–Bouldin index** quantifies clusters overlap. Lower values, with a theoretical minimum of zero, correspond to reduced similarity between clusters, indicating greater separation.

- **Calinski–Harabasz index** evaluates clustering quality by considering the ratio between inter-cluster dispersion and intra-cluster compactness. Higher values indicate better clustering.

Clustering of Isolated RBD – Derivable 2

Focused on comparing the clustering results obtained in the subgroup of RBD patients with the true labels, represented by the clinical classification of RBD severity. This assessment was grounded in the definition of confusion matrix, where the rows represent the true classes, while the coloumns the predicted ones. Correct predictions were classified as True Positives (TP) or True Negatives (TN), while misclassifications were classified as False Positives (FP) and False Negatives (FN).

Table 2.4 reports the clinical interpretation of these subgroups. Model performance was then evaluated through four standard metrics, Accuracy, Precision, Recall, and F1 score, defined as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2.2}$$

$$Recall = \frac{TP}{TP + FN} \tag{2.3}$$

$$F_1\text{- score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{2.4}$$

Subgroup	**Description**
TP	RBD2 classified in cluster RBD2
TN	RBD1 classified in cluster RBD1
FP	RBD1 classified in cluster RBD2
FN	RBD2 classified in cluster RBD1

Table 2.4: Confusion matrix interpretation for RBD cluster mapping.

Clustering on RBD and Healthy Controls – Derivable 3

The performance of the clustering algorithms in Deliverable 3 was assessed through both internal and external evaluation metrics. Internal validity was measured using the Silhouette Score, the Davies–Bouldin index, and the Calinski–Harabasz index, while external evaluation relied on Accuracy (Eq. 2.1), Precision (Eq. 2.2), Recall (Eq. 2.3), and F1 score (Eq. 2.4).

Chapter 3

Results

3.1 Dimensionality Reduction

Dimensionality reduction was applied to a matrix of 466×578, where 466 represents the patients remaining after eliminating the tails of the age distribution, and 578 represents the features selected after normalization. This process yielded a decorrelated matrix of dimension 466×155. Subsequently, the explained variance values of the principal components were calculated and represented in a scree plot. The variance explained by the first three components is reported in Table 3.1. The scree plot revealed a pronounced drop after the second principal component, suggesting that PC1 and PC2 capture the majority of the variance (Fig. 3.1). Cumulatively PC1 and PC2 account for 33% of the total variance. Based on this observation, 10 features were selected, specifically the five features that contribute most to each of the first two principal components. The choice of selecting the top-5 features per component was adopted as a practical convention to enhance interpretability and readability, after verifying that the loading profile showed a gradual decreasing trend without evident elbow. Fig. 3.2 shows the contribution of the top 5 features in the first two components.

Principal Component	**Explained Variance**
First Component	0.21
Second Component	0.11
Third Component	0.06

Table 3.1: Explained variance by the first three principal components.

After performing PCA, the 10 selected features were projected onto a two-dimensional PC1–PC2 space to provide a preliminary visualization of potential patient groupings. However, as can be observed in Fig. 3.3, no evident separation

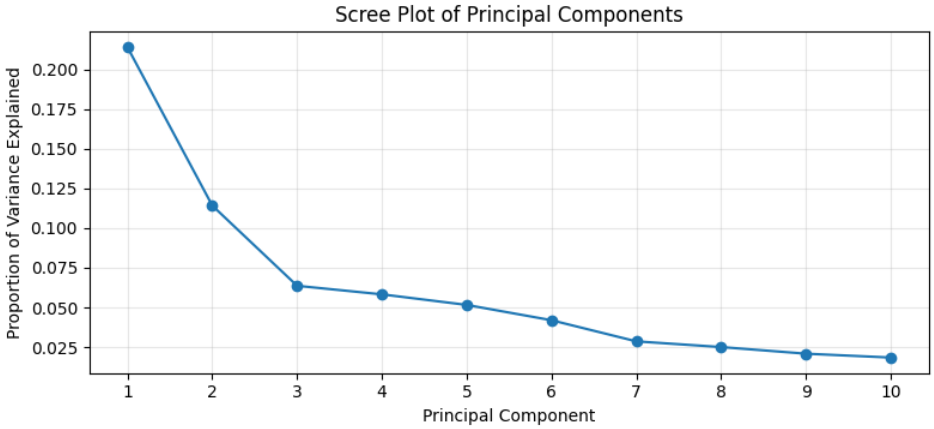


Figure 3.1: Scree plot explained by the principal components.

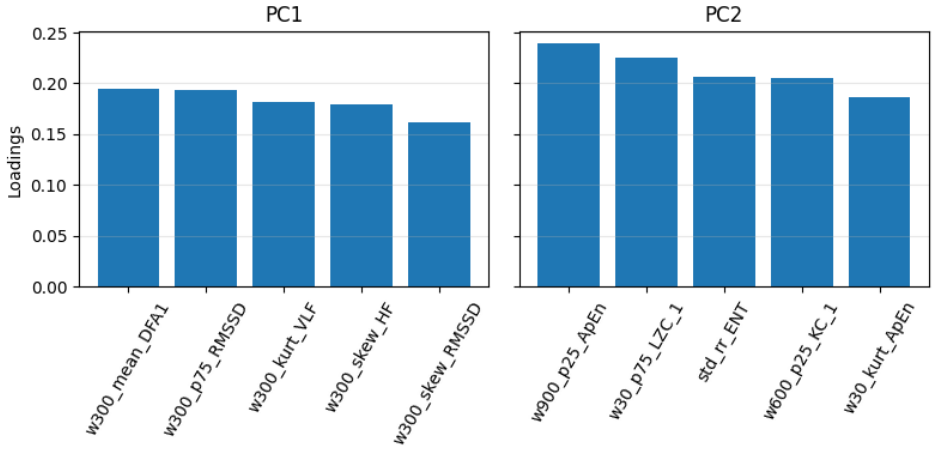


Figure 3.2: Absolute contribution of the top 5 features of the PC1 and PC2.

or distinct clusters were visible. This suggests that the variance captured by the first two principal components alone is insufficient to separate the clinical groups.

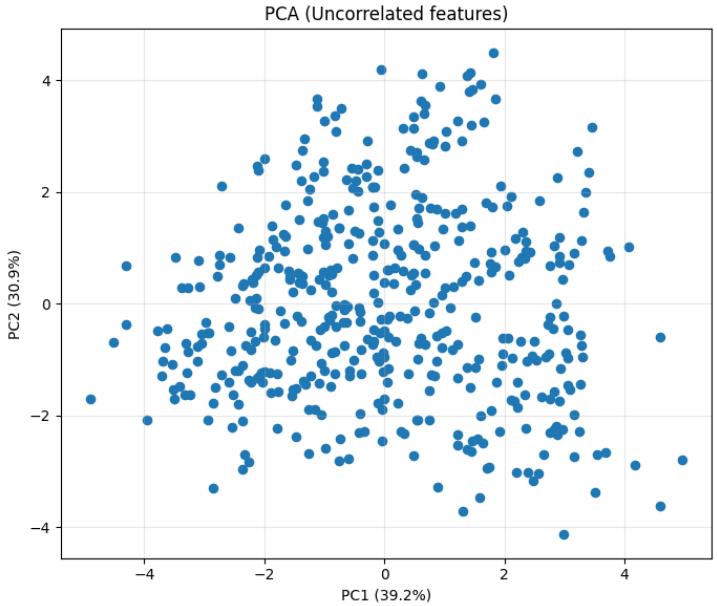


Figure 3.3: Two-dimensional PCA projection of selected features.

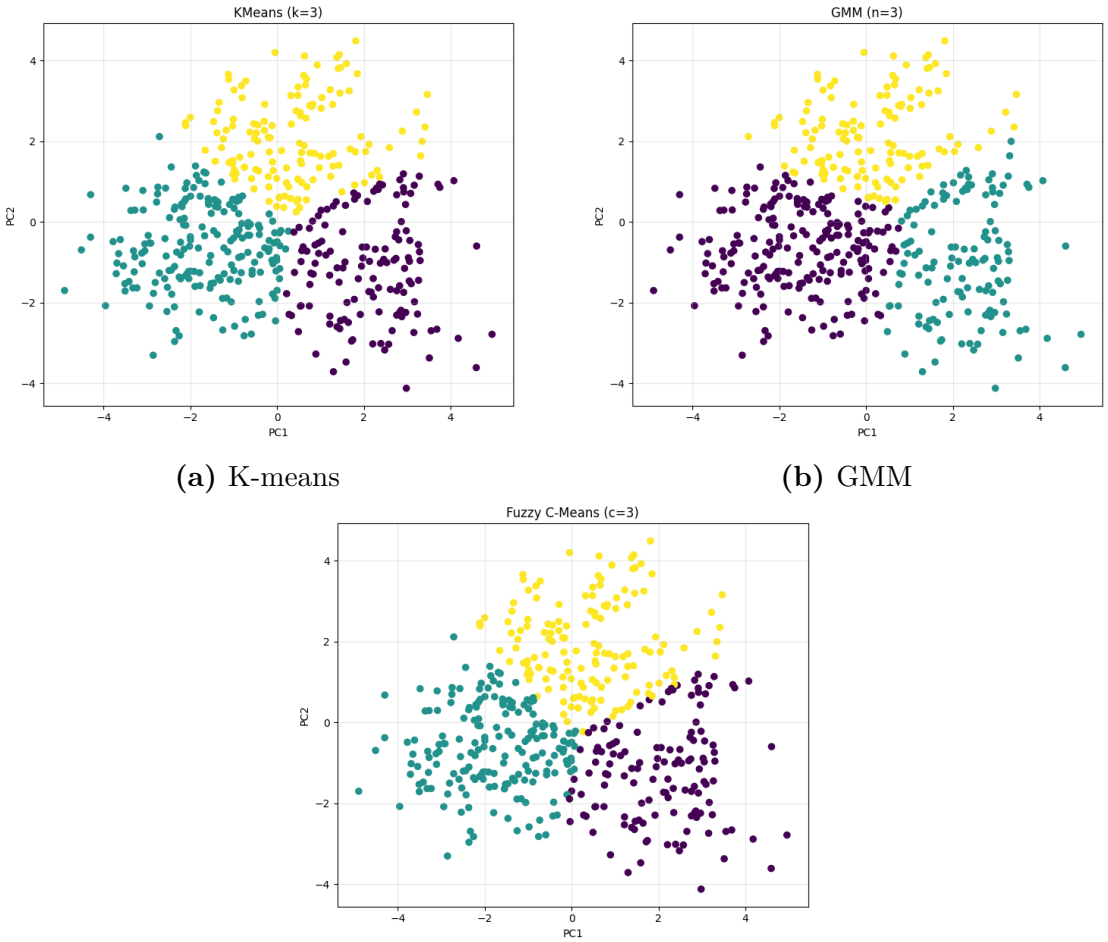
3.2 Global Clustering – Derivable 1

Derivable 1.1 – Analysis and Visualisation of Patients Distribution across Clusters

The 466×10 patient per features matrix, obtained following feature selection and exclusion of patients with extreme age values, was subjected to three clustering algorithms: K–means, Gaussian Mixture Model, and Fuzzy c-means. The number of clusters was set to 3 based on the elbow method analysis. Fig. 3.4 illustrates the cluster solutions obtained from each of the three clustering methods, visualized in the two-dimensional space defined by PC1-PC2. For each clustering method, internal validation metrics were calculated, including the Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index, reported in Table 3.2.

Method	Silhouette	Davies-Bouldin	Calinski-Harabasz
K-means	0.41	0.83	**404.3**
GMM	0.40	0.83	388.9
FCM	**0.63**	**0.40**	401.5

Table 3.2: Perfomance metrics of Global Clustering.



(a) K-means

(b) GMM

(c) FCM

Figure 3.4: Clustering separation in the two-dimensional PC1-PC2 space.

Derivable 1.2 – Definition of a Physiological Profile for each Cluster

Based on the obtained clusters, the mean values of physiological parameters within each cluster were calculated, as well as the percentages of patients classified with specific pathologies, such as RBD, OSA, and PLM. These metrics were used to define a physiological profile for patients within each cluster, summarized in Table 3.3.

- **Cluster 0:** Characterized by higher mean age, mild overweight (elevated BMI), and moderate-to-severe AHI values. This cluster exhibited notable comorbidities, including diabetes and hypertension, with RBD prevalence higher than in any other clusters.
- **Cluster 1:** Features advanced age, mild overweight, and moderate-to-severe AHI. OSA and PLM were highly prevalent, while RBD occurrence was less frequent compared to Cluster 0.
- **Cluster 2:** Distinguished by younger age, normal body weight, and low AHI values. This cluster showed minimal prevalence of OSA and PLM, along with the lowest RBD frequency across all groups.

The considerable within-cluster heterogeneity, evident from the wide standard deviations of the clinical variables reported in Table 3.3 , suggests that the identified clusters were not sharply defined and encompass highly variable clinical profiles. This was confirmed by the statistical analysis. Clinical variables were compared across clusters using the Kruskal–Wallis test and Chi-square, respectively for continous and categorical variables. Age was consistently significant across all clustering methods ($p < 0.041$), as was BMI ($p < 0.046$). PLM reached significance in the GMM solution ($p = 0.0035$), and AHI_NREM in the FCM solution ($p = 0.046$). After post-hoc pairwise comparisons (Dunn's test for continuous, Fisher for categorical, both with Bonferroni-Holm correction) only age survived ($p < 0.04$). PLM differed significantly between clusters 1–2 ($p = 0.0039$) and 0–2 ($p = 0.011$), but not 0–1 ($p = 0.919$). The statistical analysis suggests that the clinical profiles identified do not fully correspond to the clustering structure emerging from the selected features, with age appearing to be the only truly discriminating clinical variable. The fact that the identified clusters do not reflect a clear distinction between healthy and pathological subjects will be further confirmed by the analyses carried out in Deliverable 2 and Deliverable 3.

Description	K–means	GMM	FCM
Higher age, slightly overweight, AHI moderate/severe, co-morbidities, RBD most frequent	N=183 Age=51.85 ± 16.33 BMI=26.01 ± 4.41 AHI_REM=14.02 ± 14.91 AHI_NREM=11.77 ± 12.84 AHI_SONNO=12.35 ± 12.63	N=186 Age=51.76 ± 16.23 BMI=26.09 ± 4.71 AHI_REM=14.22 ± 16.16 AHI_NREM=11.76 ± 13.74 AHI_SONNO=12.36 ± 13.92	N=174 Age=51.96 ± 14.47 BMI=26.01 ± 5.00 AHI_REM=14.14 ± 16.80 AHI_NREM=11.53 ± 13.08 AHI_SONNO=12.17 ± 13.09
High average age, slightly overweight, AHI moderate/severe, high prevalence of OSA and PLM, RBD less frequent	N=116 Age=52.19 ± 17.41 BMI=25.35 ± 5.15 AHI_REM=16.31 ± 18.70 AHI_NREM=9.37 ± 17.03 AHI_SONNO=10.84 ± 17.06	N=108 Age=52.35 ± 17.23 BMI=25.51 ± 5.15 AHI_REM=16.74 ± 17.95 AHI_NREM=9.63 ± 16.65 AHI_SONNO=11.16 ± 16.35	N=126 Age=51.76 ± 17.80 BMI=25.31 ± 5.06 AHI_REM=16.32 ± 18.75 AHI_NREM=9.68 ± 17.52 AHI_SONNO=11.09 ± 17.49
Lower age, normal body weight, low AHI, low prevalence of OSA and PLM,lowest frequency of RBD	N=169 Age=35.75 ± 13.55 BMI=24.26 ± 5.00 AHI_REM=10.60 ± 16.89 AHI_NREM=6.59 ± 13.71 AHI_SONNO=7.20 ± 13.55	N=174 Age=36.22 ± 13.23 BMI=24.10 ± 4.22 AHI_REM=10.33 ± 16.76 AHI_NREM=6.49 ± 11.85 AHI_SONNO=7.12 ± 12.49	N=168 Age=35.85 ± 16.26 BMI=24.31 ± 4.57 AHI_REM=10.32 ± 15.18 AHI_NREM=6.72 ± 13.05 AHI_SONNO=7.32 ± 12.84

Table 3.3: Physiological profiles of the clusters obtained with the three different methods.

Derivable 1.3 – Evaluation of the Features Values in each Cluster in Relation to the Cluster Clinical Profile

To characterize the clinical profile of each cluster, the global mean value of each feature across all patients was computed. For each patient within a cluster, the delta between the individual feature value and the global mean was calculated. These deltas were then averaged within each cluster, and the Standard Error of the Mean (SEM) was calculated. This approach allows the identification of which features are dominant (showed the largest deviations from the global mean) within each cluster, differentiating one patient group from another. Fig. 3.5 shows an example, obtained with GMM, of how the features increase and decrease respect to the global mean in each cluster. As reported in Table 3.4, the three clusters display features patterns, that remain consistent across the three clustering methods:

- **Cluster 0** showed increased fractality and complexity-related features, while RMSSD, the skewness of VLF band power, and the kurtosis of approximate entropy decreased.
- **Cluster 1** was characterized by an increase in the RMSSD, VLF skewness, mean DFA1, and the diagonal line length entropy of the recurrence plot. In contrast, entropy, mean DFA1, and HF band skewness were reduced.
- **Cluster 2** exhibited elevated entropy, RMSSD, and VLF skewness, with

concurrent decreases in complexity features, recurrence plot entropy, and mean DFA1.

Cluster Description	Features increasing	Features diminishing
Cluster 0 – More compromised patients	w300_mean_DFA1, w300_skew_HF, w600_p25_KC1, w300_p75_LZC1	w300_p75_RMSSD, w300_skew_VLF, w300_skew_DFA1, w30_kurt_ApEn
Cluster 1 – Intermediate profile with respiratory impairment	w300_skew_VLF, w300_p75_RMSSD, std_rr_ENT, w300_skew_DFA1	w300_skew_HF, w900_p25_ApEn, w30_kurt_ApEn, w300_mean_DFA1
Cluster 2 – Less compromised patients	w900_p25_ApEn, w30_kurt_ApEn, w300_skew_VLF, w300_p75_RMSSD	w600_p25_KC1, w30_p75_LZC1, std_rr_ENT, w300_mean_DFA1

Table 3.4: Summary of features values across clusters.

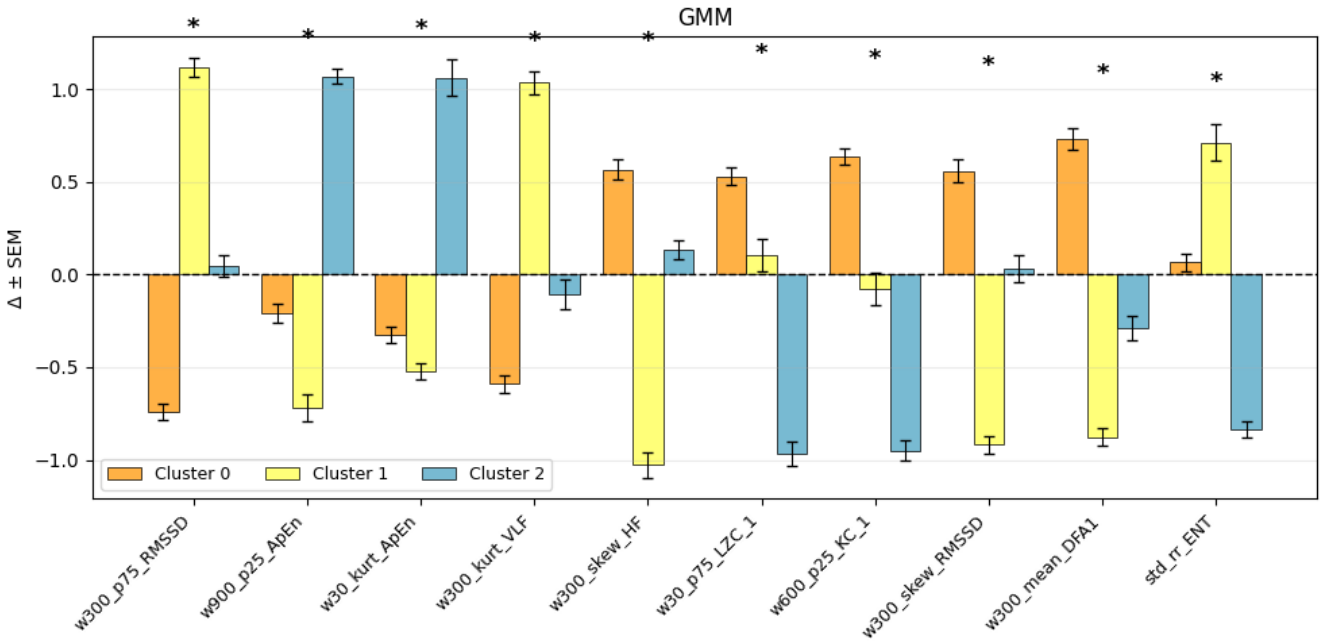


Figure 3.5: GMM cluster-wise feature deviation from the global mean reported with SEM, highlighting the most distinctive descriptors characterizing each group.

To investigate the distribution of the 10 selected features across clusters, violin plots were generated for each feature within every cluster, allowing visual assessment of their distributions and separability. Statistical significance of the observed differences was evaluated using the Kruskal-Wallis test, which revealed that all features exhibited statistically significant differences across clusters. To further identify which specific pairs of clusters differed significantly, a Dunn post-hoc test with Bonferroni correction was applied. The results showed that the majority of features were highly significant across all cluster pairs. Slightly lower p-values, yet still significant, were observed between cluster 0 and cluster 1 for the LZC and ApEn kurtosis features, with p-values of 0.048 and 0.023, respectively. Finally, the effect size was assessed using $Epsilon^2$, a measure of the proportion of variability in each feature attributable to cluster membership. The obtained values ranged from 0.38 to 0.58, indicating that the identified clusters are not only statistically significant, but also substantively meaningful. Fig. 3.6 reports the features distribution trough violin plots, obtained with GMM, showing that the distributions are separated among clusters, though with some overlaps, particularly between cluster 0 and cluster 1 for certain features.

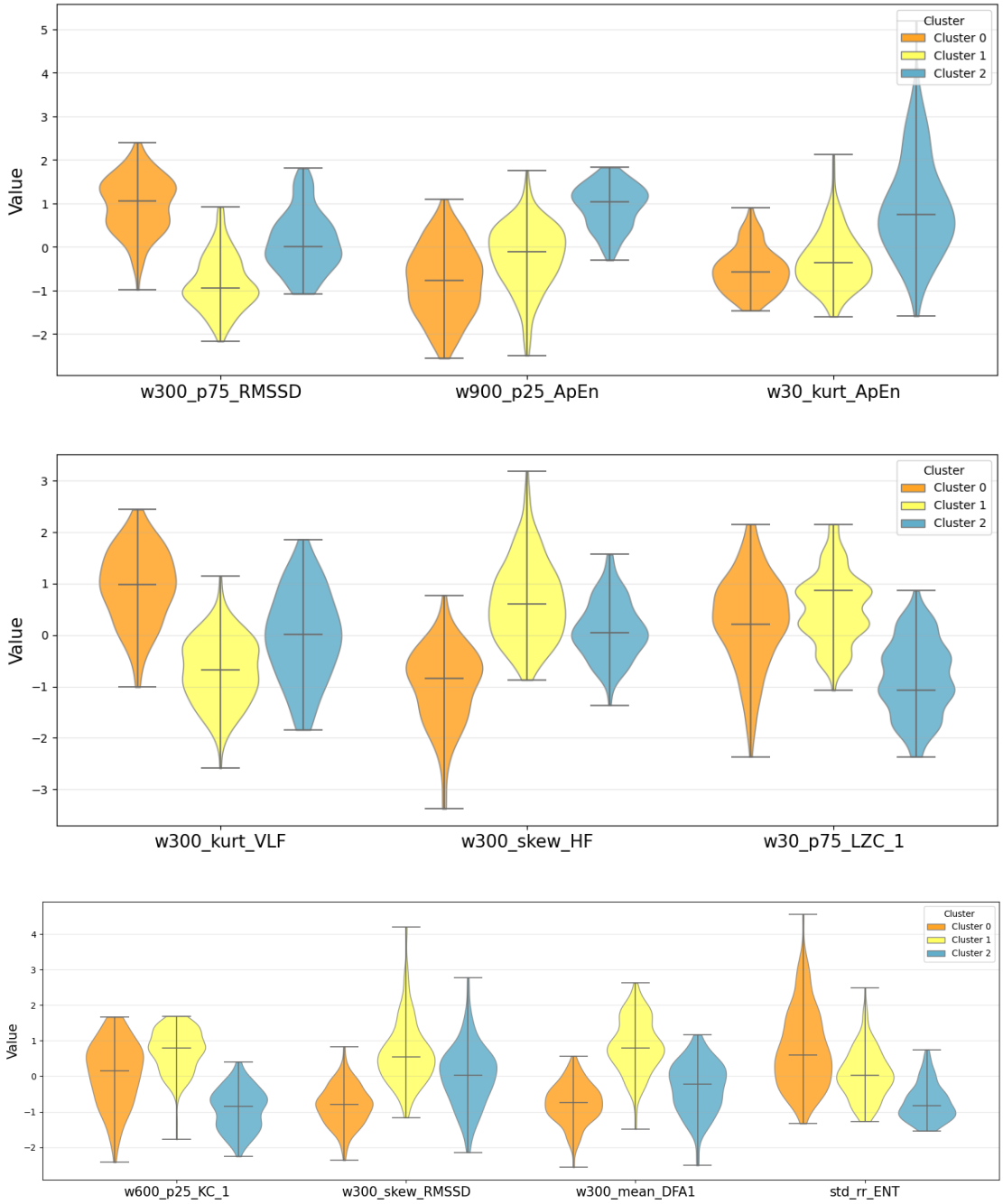


Figure 3.6: Violin plot of the selected 10 features, used in GMM clustering.

Furthermore, the values of some features that were not previously identified as the most discriminative in the initial feature selection, were evaluated in each cluster using violin plots. The features to be analyzed were selected based on their well established relationship with clinical profiles in the literature and on the initial hypothesis that they might be strongly associated with clinical diagnosis. In addition, this analysis is justified by the fact that feature loadings revealed the absence of truly dominant features, suggesting that multiple features contribute to principal components. Fig. 3.2 illustrates the absence of a clear elbow in the contribution curve, with feature loadings gradually decreasing across principal components rather than showing a sharp drop. Consequently, the mean power in HF and LF band, calculated within the 10 minutes window, were evaluated, as these features reflect the autonomic impairment typical of iRBD patients. The analysis of frequency features revealed that HF exhibit a decrease in Cluster 1 compared to an increase in Cluster 0, while LF showed no substantial inter-cluster variation. Fig. 3.7 shows the distribution of these two features in clusters identified by K-means, with findings consistent across all methods.

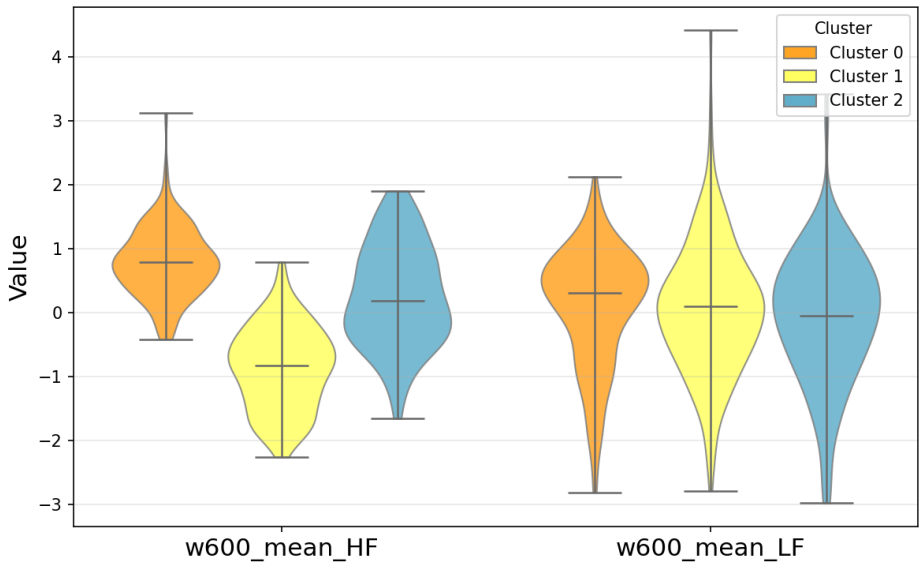


Figure 3.7: Violin plots of mean power in HF and VLF band calculated within 10 minutes windows across the three clusters obtained with K-means.

Secondly, the distribution of four Mixed and Position features was also examined through violin plots, selected on the basis of their hypothesized association with the clinical profiles identified by clustering. The violin plots exhibited substantial overlap, indicating that these features were not highly discriminative in cluster separation (Fig. 3.8). However, some considerations can be drawn regarding their

distribution patterns and the potential relationship between feature values, clusters and associated clinical profile:

- **out_in_changes_windows**: represents the fraction of peaks that exceed the median peak amplitude within a window corresponding to a position change. Cluster 0 exhibited a higher mean value compared to the other two clusters, with a distribution shifted toward higher values.
- **turnInt_mean**: represents the average duration of the time interval between two consecutive position changes. Cluster 2 demonstrated a slightly higher mean value.
- **ChangesPerHour**: frequency of position changes per hour. Cluster 2 displayed lower mean value compared to the others.
- **d_LFHFratio_std**: represents the standard deviation of the difference in the ratio between LF and HF components, calculated in windows preceding and following the position change. Cluster 0 exhibited a greater mean and a distribution shifted toward higher values.

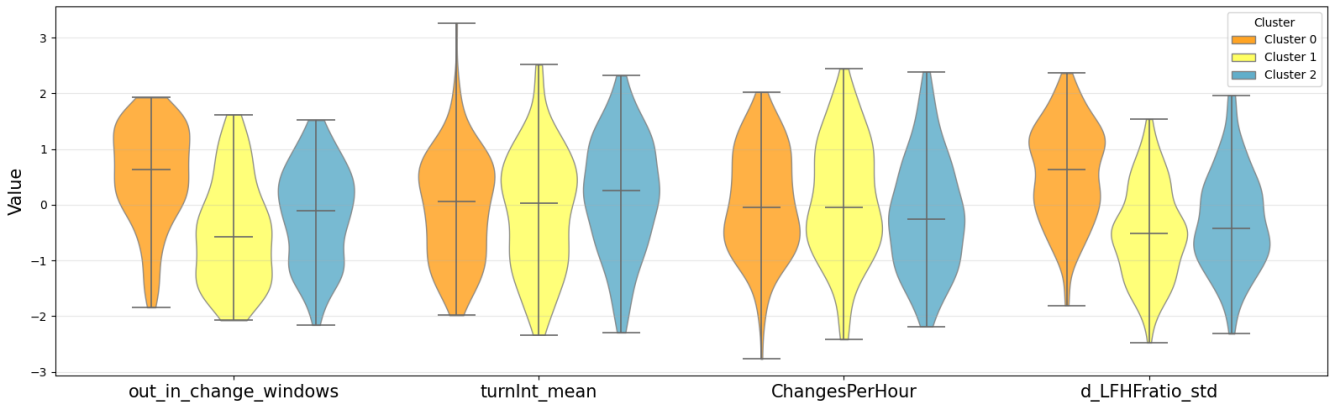


Figure 3.8: Violin plots the four features, selected among Mixed and Position features, within the three cluster obtained with GMM.

3.3 Clustering of Isolated RBD – Derivable 2

All patients identified as having RBD were selected and those concurrently classified with OSA or PLM excluded, yielding a final cohort of 150 patients with isolated RBD. Using the same 10 features selected previously, K-means and FCM clustering were applied to assess whether the methods could discriminate between patients

with severe RBD (RBD1) and those with a milder form of the disorder (RBD2). Classification performance are reported in Table 3.5.

Method	Accuracy	Precision	Recall	F1-score
K-means	0.57	0.64	0.60	0.62
FCM	0.52	0.61	0.49	0.54

Table 3.5: Performance metrics for each clustering method for Derivable 2.

Feature Based Analysis

Given the limited robustness of the obtained metrics, a deeper analysis was conducted by examining the feature profiles that dominated the classification of the four subgroups: True Positives, True Negatives, False Positives, and False Negatives. This approach was motivated by the inherent variability in the clinical severity classification of RBD, which relies on criteria that are not fully standardized and is performed by multiple clinicians, often resulting in inconsistent assessments, with some evaluators regarding this classification as outdated. The feature profile of each subgroup was computed as the mean feature vector across all patients within that group, and pairwise Euclidean distances between profiles were calculated to quantify their similarity. The minimum distance was observed between TN and FN profiles. This finding suggests that FN patients, despite being clinically labelled as mild RBD, exhibit physiological characteristics more closely resembling severe cases (TN), potentially identifying a subgroup at risk of disease progression. This pattern was further confirmed by the box plot distributions of the five most discriminative features across subgroups (Fig. 3.9), where FN patients align closely with TN. From a physiological standpoint, severe RBD was characterized by increased mean DFA1 and HF band skewness, alongside decreased VLF power and RMSSD, consistent with reduced parasympathetic activity and diminished heart rate variability.

Fuzzy Memebership Analysis

In addition, exploiting the fact that FCM, unlike hard clustering approaches, assigns each patient a membership degree to each cluster, rather than discrete categorical assignment, membership distributions were analyzed across TP, TN, FP, and FN subgroups. The aim is to evaluate the sharpness of separation between the two phenotypic groups in the selected feature space. The maximum membership value observed across all subjects was 0.88, indicating that no patient exhibited near-exclusive affiliation to a single cluster (membership values approaching 0.95/0.05).

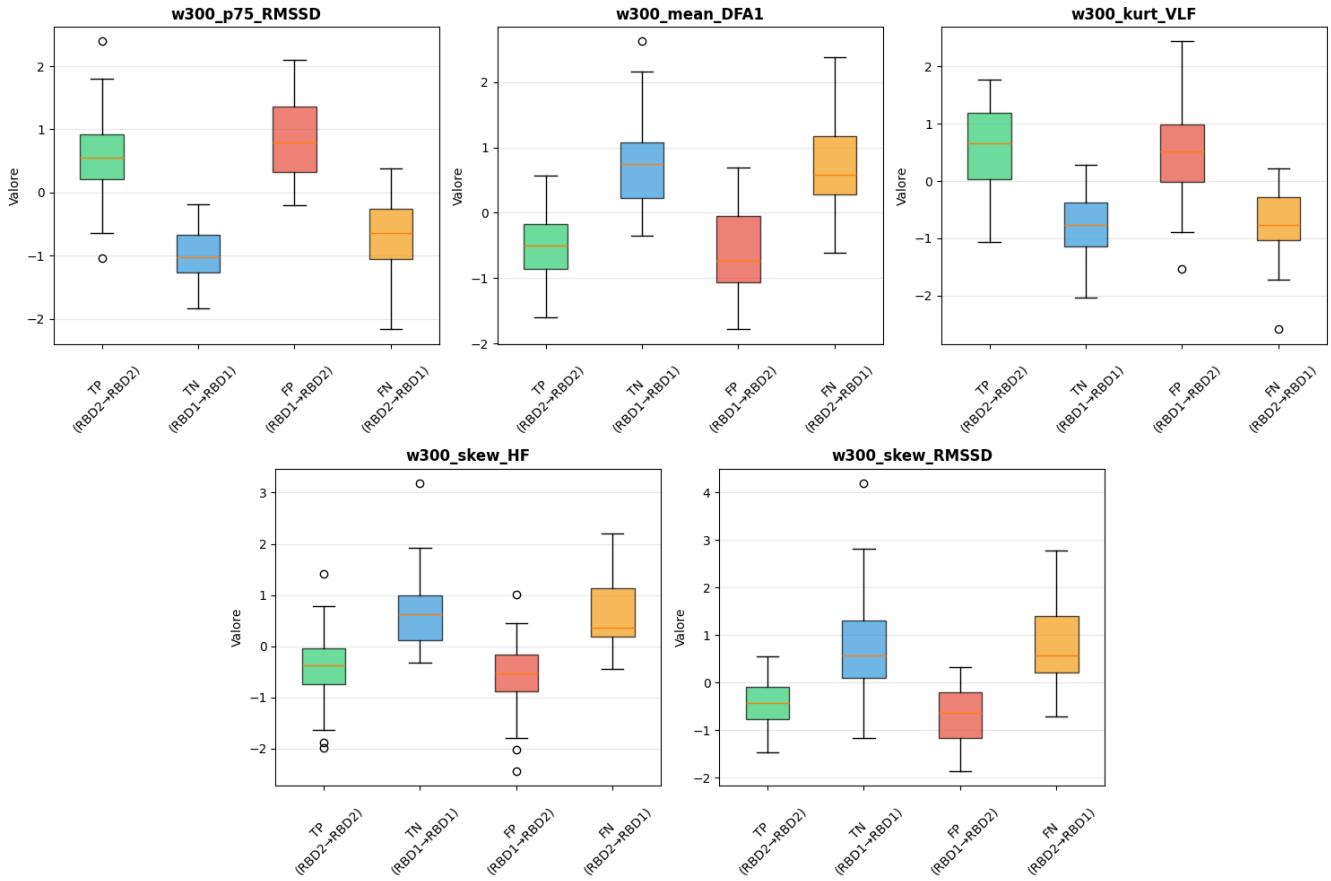


Figure 3.9: Box Plot of 5 features across the four subgroups obtained with K-means.

This suggests moderate separation and substantial overlap between the two phenotypic profiles captured by the selected features. Consistently, only 61 out of 150 patients (40.7%) showed a maximum membership exceeding 0.7. Importantly, the distribution of membership confidence was comparable between correctly and incorrectly classified subjects. Among correctly classified patients (TP + TN, n = 78), 30 (38.5%) had max(membership) > 0.7, whereas among misclassified patients (FP + FN, n = 72), 31 (43.1%) exceeded the same threshold. This pattern indicates that classification errors are not primarily attributable to uncertain or borderline cluster assignments. The violin plot of the FCM membership degrees, in Fig. 3.10, corroborates the previous analysis, indicating a moderate separation between the two clusters, with a substantial overlap across distributions. No subject reaches near-complete affiliation to a single cluster: membership values do not approach 0.9, while a considerable portion of patients exhibits values close to 0.6, consistent

with an intermediate position between the two profiles. A further aspect concerns the pattern of disagreement with clinical labels. If misclassifications were arbitrary, FP and FN subjects would cluster around 0.5, indicating ambiguous membership. Instead, a systematic pattern emerged: FN patients (clinically RBD2, assigned to RBD1 cluster) exhibited membership distributions toward the RBD1 cluster, overlapping with TN, while FP patients (clinically RBD1, assigned to RBD2 cluster) showed membership toward the RBD2 cluster comparable to TP.

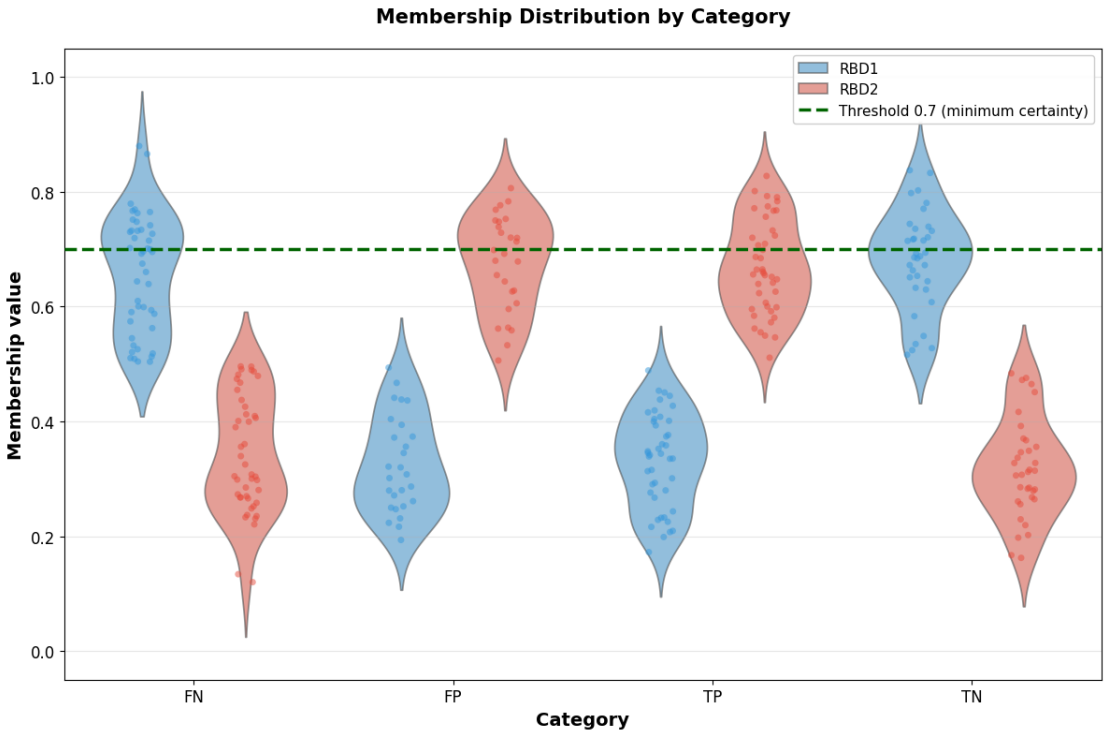


Figure 3.10: Violin plot showing the membership values of FCM for each category.

3.4 Clustering on RBD and Healthy Controls – Derivable 3

To the previously selected RBD patients, a group of clinically classified healthy controls was added. These controls were selected based on the absence of RBD, OSA, PLM, and insomnia, and were not receiving any pharmacological treatment. After age-based filtering (35-75 years), the combined dataset comprised 107 RBD patients (RBD1: 50 patients, RBD2: 57 patients) and 53 healthy controls. Three clustering algorithms (K-Means, GMM, and FCM) were applied with two different

configurations: $k = 2$ to distinguish between HC and RBD patients, and $k = 3$ to differentiate among HC and two RBD subtypes (RBD1: 50 patients, RBD2: 57 patients). For each clustering method the cluster solutions were visualized in the two-dimensional space defined by PC1-PC2 and compared against the true clinical labels in the same space. Additionally, the same comparison has been done in the PC2-PC3 plane. For each clustering method and configuration, performance was evaluated using both internal cluster validity indices (SS, DB, CH) and external validation metrics computed against clinical labels (Accuracy, Precision, Recall, F1-score). Results for $k = 2$ and $k = 3$ configurations are reported in Tables 3.6, 3.7, 3.8, and 3.9, respectively. With $k = 2$ K-means achieved the best internal metrics (SC = 0.26, DB = 1.41, CH = 62.15, Table 3.6), though these values reflect low cluster separation. External validation against clinical labels (HC vs. RBD) yielded modest performance, with K-means reaching Accuracy = 0.55 and F1-score = 0.66 (Table 3.7). With $k = 3$ internal metrics remained comparable (K-means: SC = 0.24, DB = 1.33, CH = 60.95, Table 3.8), while external performance further deteriorated, with FCM achieving the best results at Accuracy = 0.41 and F1-score = 0.40, (Table 3.9). Across both configurations, the overall low performance suggests that while the clustering algorithms succeed in identifying latent structure in the data, the emergent partitions do not align well with the clinical classification, particularly in distinguishing between RBD severity subtypes.

Method	**SS**	**DB**	**CH**
K-means	0.26	1.41	62.15
GMM	0.22	1.60	52.34
FCM	0.24	1.55	58.24

Table 3.6: Internal validation metrics for each clustering method with $k = 2$.

Acc.	**Prec.**	**Recall**	**F1-score**
0.55	0.66	0.66	0.66
0.54	0.67	0.62	0.64
0.54	0.68	0.59	0.63

Table 3.7: External validation metrics for each clustering method with $k = 2$.

Method	**SS**	**DB**	**CH**
K-means	0.24	1.33	60.95
GMM	0.18	1.62	46.80
FCM	0.23	1.42	60.08

Table 3.8: Internal validation metrics for each clustering method with $k = 3$.

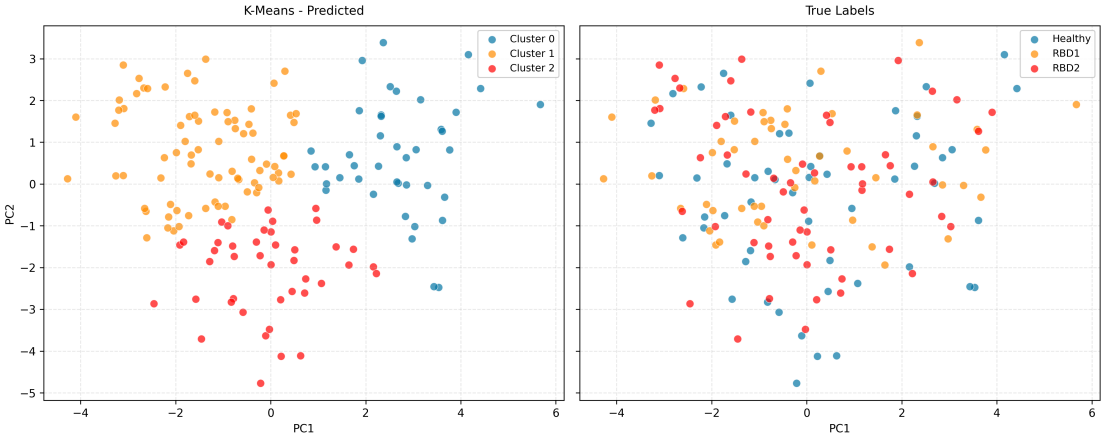
Acc.	**Prec.**	**Recall**	**F1-score**
0.39	0.40	0.39	0.38
0.38	0.38	0.38	0.37
0.41	0.40	0.41	0.40

Table 3.9: External validation metrics for each clustering method with $k = 3$.

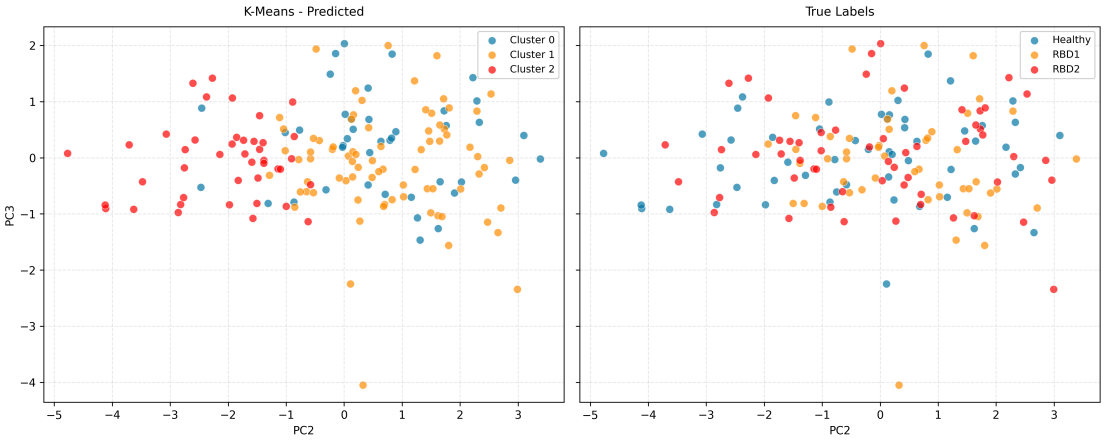
The visualization in the principal components space allowed the identification of clusters that appeared geometrically well-separated. However, when observing the

distribution of true clinical labels, it becomes evident that the separations do not correspond to the clinically relevant distinctions, neither between RBD patients and HC nor among RBD1, RBD2, and HC. The separation observed in PC1–PC2 projections was primarily driven by the first principal component, which captured the largest proportion of variance but appeared uninformative with respect to clinical diagnosis. This suggests that the dominant source of variability in the data does not align with the biological or pathological distinctions of clinical interest. PC2-PC3 visualizations revealed a more overlapping distribution across clinical classes, suggesting that clinically relevant variability is not concentrated along a single dominant direction but is instead weaker and distributed across higher order components. To assess whether a clearer separation could be observed in a higher-dimensional space, a three-dimensional visualization of the data projected onto the PC1-PC2-PC3 space was performed, however, no distinct clustering structure has been found.

Fig. 3.11 illustrates the K-means clustering solution as a representative example for the $k = 3$ configuration, visualized in both PC1-PC2 and PC2–PC3 planes. In the PC1–PC2 projection (top panels), the predicted clusters appear geometrically well-separated, suggesting distinct subgroups in the data. However, when true clinical labels (HC, RBD1, RBD2) are overlaid in the same space, they exhibit no clear correspondence to the algorithmic partition. The PC2–PC3 projections (bottom panels) shows that both predicted clusters and true labels appear highly overlapping. Fig. 3.12 reports the cluster obtained with K-means with $k = 3$ projected onto the three-dimensional PCA space, confirming that no clear separation could be drawn, even when considering an higher-dimensional representation of the data.



(a) Clusters visualization on PC1-PC2



(b) Clusters visualization on PC2-PC3

Figure 3.11: Comparison between K-means clustering solution with $k = 3$ (right) and true clinical labels (left).

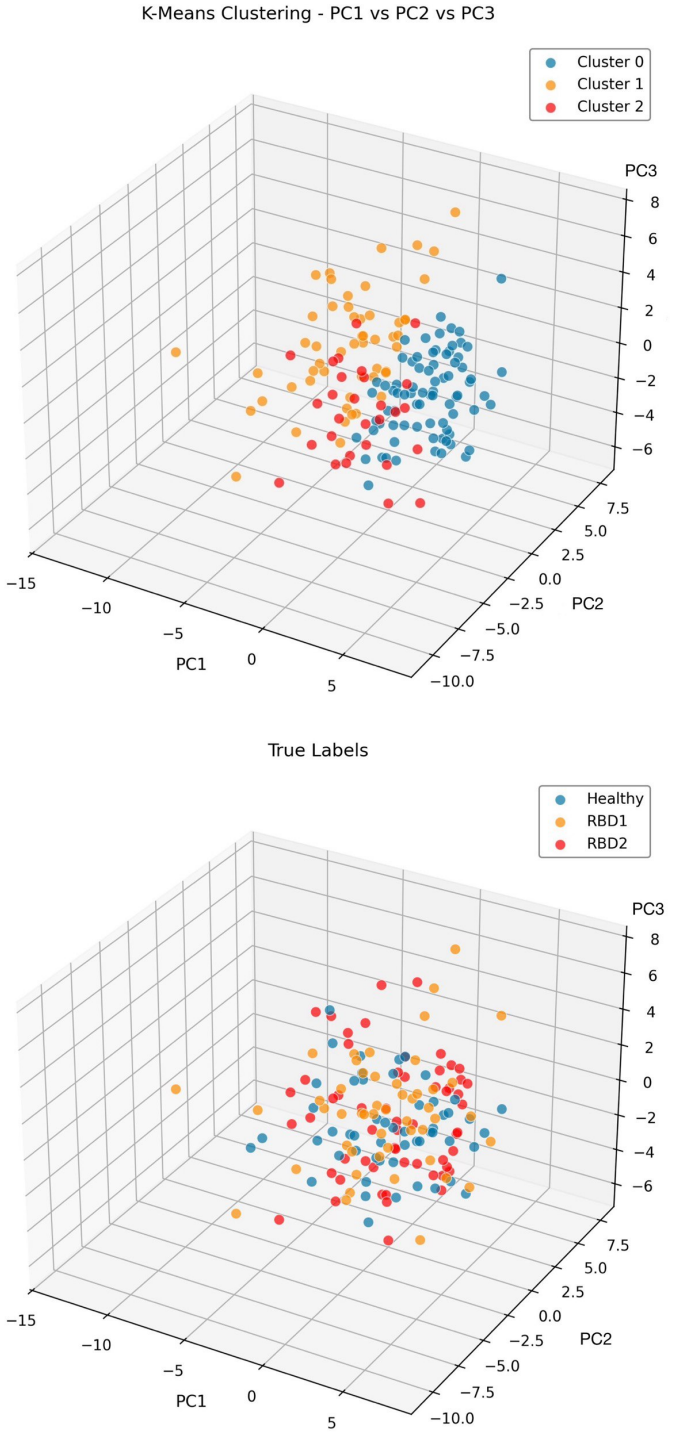


Figure 3.12: Comparison between K-means clustering solution with $k = 3$ and the true clinical labels, visualized in a the three-dimensional PC1-PC2-PC3 space.

Chapter 4

Discussion

4.1 Dimensionality Reduction

PCA was employed as a dimensionality reduction method, enabling the selection of 10 features with the highest loadings across the first two principal components. The first two components together explained only 33% of the total variance, a relatively modest proportion, suggesting that the variance is broadly distributed across a large number of components rather than being concentrated in a few dominant directions. This may reflect the intrinsic complexity and multidimensionality of the data, where no single feature or small subset of features captures the majority of the underlying variability. Furthermore, the observation that many features contribute in a comparably modest manner to the principal components, rather than a few features dominating, carries an important implication: the dataset may lack a defined structure, or the relevant information is distributed across many dimensions.

4.2 Global Clustering – Derivable 1

Unsupervised cluster analysis applied to the entire patient cohort yielded acceptable separation quality and intra-cluster cohesion. Internal validation indices were consistent across all three algorithms, reflecting satisfactory performance in terms of both cluster compactness and inter-cluster separation. In particular, FCM achieved the best results reaching: SS= 0.63, DB= 0.40, CH= 401.5. The analysis consistently identified three distinct patient groups across all three clustering methods (Table 3.3): Cluster 0 comprised patients with greater clinical severity; Cluster 1 patients characterized by an intermediate profile, marked presence of OSA and respiratory impairment; and Cluster 2 consisting of less compromised patients. However, the clinical boundaries between clusters were not significant, suggesting

that the features driving the clustering structure do not capture clinical categories. The association of a specific clinical profile with the identified clusters is therefore complex and not straightforward. In the following paragraphs, the features that emerged as most significant in driving the clustering structure are examined in light of the existing literature, in order to explore what clinical conditions they have been previously associated with. Addressing this problem is where the key advantage of employing Machine Learning approaches becomes particularly relevant, as it lies in the ability to evaluate and interpret feature values within the identified clusters. Existing literature showed that sleep is tightly related to autonomic nervous system (ANS) activity, which governs essential involuntary functions including respiration, heart rate, digestion, and blood pressure [6]. Several studies have consistently reported a global reduction in heart rate variability and impaired cardiac vagal modulation in patients with iRBD compared with healthy controls [6, 7, 8]. In addition, autonomic instability, commonly manifested as alterations in HRV, is a hallmark of both iRBD and α-synucleinopathies [6], indicating dysfunctions involving both the sympathetic and parasympathetic components of the autonomic nervous system [7, 9]. However, autonomic regulation of heart rate is the result from a complex and dynamic interaction between the parasympathetic (PNS) and sympathetic (SNS) nervous systems. In healthy individuals at rest, parasympathetic influence predominates, ensuring a low heart rate. Importantly, sympathetic and parasympathetic activities do not operate as simple opposites: sympathetic activation may inhibit parasympathetic tone but can also facilitate its responsiveness. Therefore, autonomic control cannot be described as a zero-sum process. This interplay has been compared to the "left-foot braking" strategy in Formula 1® driving, where acceleration and braking are simultaneously modulated to achieve optimal control [20]. It is also important to notice that HRV is influenced by both demographic and methodological factors. Time-domain HRV measures are known to decline with age. For what concerns sex-related differences, women reported higher mean heart rate compared to men, while in the frequency domain, they showed lower total, VLF, and LF power but higher HF power, suggesting a relative vagal predominance, whereas men tend to exhibit relative sympathetic dominance despite a lower mean heart rate [20, 21]. From a methodological point of view, the recording length is crucial: HRV measurements are commonly classified as ultra-short-term (<5 min) or short-term (5 min). Ultra-short-term recordings are attractive for their efficiency; however, they may fail to capture a sufficient proportion of the underlying heart rate variability, particularly those reflecting low-frequency and non-linear dynamics. In conclusion, their reliability strongly depends on the specific HRV metric and the duration of the segment [20].

Time Domain Measures

RMSSD reflects beat-to-beat heart rate variance and is the primary time-domain measure used to estimate vagally mediated changes in HRV. A healthy heart is not a metronome: its beat-to-beat fluctuations exhibit complex, chaotic patterns, best described by mathematical chaos, that allow the cardiovascular system to rapidly adjust to sudden changes [20]. Lower RMSSD values, as observed in Cluster 0, indicate a reduced capacity to generate complex and adaptive responses, typical of an healthy cardiovascular function and indicates autonomic dysfunction, consistent with greater clinical severity. Lower RMSSD has been broadly associated with older age, physical inactivity, depression, heart failure, and increased cardiac morbidity, as well as with elevated risk of all-cause mortality [22]. Beyond these established associations, reduced vagal tone as reflected by RMSSD has also been linked to post-traumatic stress disorder [23], and to a spectrum of medically unexplained physical symptoms (MUPS) and related syndromes, including chronic fatigue syndrome, irritable bowel syndrome, and fibromyalgia [24]. Taken together, these findings underscore the broad clinical relevance of RMSSD as a marker of autonomic dysregulation, extending well beyond the cardiovascular domain.

Frequency Domain Measures

Frequency-domain HRV analysis estimates the distribution of absolute or relative signal power across specific frequency bands: Ultra Low Frequency (ULF, < 0.003 Hz), Very Low Frequency (VLF, 0.0033–0.04 Hz), Low Frequency (LF, 0.04–0.15 Hz), and High Frequency (HF, 0.15–0.40 Hz). The LF component reflects a combination of sympathetic and parasympathetic influences, whereas the HF band mainly represents parasympathetic activity. Based on these assumptions, the LF/HF ratio is commonly interpreted as an index of sympathovagal balance: low values indicate parasympathetic dominance, while high values suggest sympathetic dominance or parasympathetic withdrawal [20].

- The VLF band, which requires longer monitoring periods, has been shown to be strongly associated with all-cause mortality than LF or HF power and may represent a fundamental rhythm of cardiovascular regulation. Reduced VLF power has been linked to arrhythmic death, post-traumatic stress disorder, and inflammatory states [20]. In OSA, alterations in VLF power have been proposed to reflect long-term autonomic instability related to chronic intermittent hypoxia and systemic inflammation [25]. Beyond autonomic modulation, VLF oscillations are also influenced by thermoregulation and humoral systems, particularly the renin-angiotensin: adenosine has been shown to modulate heart rate in a dose-dependent manner, representing a neuromodulatory mechanism that may further contribute to VLF variability [26]. Cluster 0

shows a reduction in the VLF power band, reflecting impaired parasympathetic nervous system (Table 3.4). Meanwhile, Cluster 1 shows an increase in the VLF band (Table 3.4).

- The HF band, often called respiratory band because it corresponds to heart rate variations related to the respiratory cycle, reflects parasympathetic activity and corresponds to respiratory sinus arrhythmia, with heart rate accelerating during inspiration and decelerating during expiration [20]. Beyond the central coupling of respiratory oscillators with autonomic brainstem centers, a mechanical cardiopulmonary mechanism also contributes to HF variability: the Bainbridge reflex, initiated by atrial mechanoreceptors in response to respiratory changes in central blood volume, generates corresponding fluctuations in cardiac autonomic activity [26]. During inspiration, intrathoracic pressure decreases, facilitating venous return to the right atrium and increasing atrial filling volume. Atrial mechanoreceptors detect this volume increase and trigger a reflex tachycardia via autonomic efferent pathways; during expiration, the reverse occurs, producing bradycardia. Since only the parasympathetic limb of this reflex operates at respiratory frequencies, HF power integrates both centrally-driven vagal modulation and peripherally-driven mechanical afferences, a distinction particularly relevant in patients with OSA or respiratory compromise, where mechanical cardiopulmonary coupling may be significantly altered. Indeed, in OSA patients, HF power has been reported to be elevated, reflecting increased parasympathetic activity triggered by recurrent obstructive events: each apneic episode activates vagal reflexes in response to hypoxia and mechanical airway obstruction [25]. In this study, Cluster 0 reported an increase in the skewness HF power, while Cluster 1 a decrease (Table 3.4).

Complexity Measures

Cluster 0, with the greatest clinical severity, exhibited an increase in complexity measures, such as Kolmogorov Complexity (KC) and Lempel-Ziv Complexity (LZC)(Table 3.4). Cluster 1 shows a smaller increase in complexity values, while Cluster 2 exhibits the lowest complexity measures. This gradient suggests that in more compromised patients, complexity is increased. Some studies have correlated increased HRV complexity with depression and anxiety disorders [27]. Notably, disease-specific patterns have also been identified: heart failure patients tend to exhibit reduced complexity and entropy, reflecting a loss of adaptive variability consistent with severe autonomic impairment, whereas apnea patients show an opposite pattern, characterized by increased irregularity [28]. This divergence underscores the importance of interpreting complexity measures: reduced complexity may signal a rigid system, while elevated complexity may reflect pathological disorganization.

Fractality Measures

Cluster 0, with the greatest clinical severity, exhibited higher DFA α_1 values (Table 3.4). Detrended Fluctuation Analysis (DFA) quantifies self-similar correlations between successive RR intervals across different time scales, yielding two scaling exponents: α_1, capturing short-term fluctuations linked to the baroreceptor reflex, and α_2, reflecting long-term regulatory dynamics. Values of α_1 greater than 1 indicate greater system rigidity and a reduced capacity for rapid autonomic adaptation [20]. From a prognostic standpoint, DFA α_1 has emerged as one of the strongest independent predictors of both total and cardiovascular mortality, outperforming several conventional HRV measures [29].

Entropy Measures

Approximate Entropy (ApEn) quantifies the regularity and complexity of the RR interval time series: large ApEn values indicate low predictability of successive RR interval fluctuations, while small values indicate a regular and predictable signal [20]. Cluster 0 and Cluster 1 showed a decrease in ApEn, while Cluster 2 showed an increase, suggesting that less compromised patients have higher entropy values, that reflect the degree of unpredictability necessary to enable the variability typical of a healthy system. Nevertheless, elevated ApEn has also been reported in congestive heart failure patients, where it reflects not healthy variability but rather a more erratic and disorganized heart rate dynamics [30].

4.3 Clustering of Isolated RBD – Derivable 2

Unsupervised K-means and GMM clustering were applied to 150 RBD patients using 10 selected HRV features, with the aim of distinguishing severe (RBD1) from mild (RBD2) cases. The classification performance metrics, reported in Table 3.5, indicated only moderate accuracy. These results are consistent with what emerged in the first analysis, suggesting that autonomic signatures captured by these features do not straightforwardly reflect sleep disorder-related clinical profiles. Furthermore, the inability to discriminate between RBD severity levels is not entirely surprising: many clinicians consider the current binary severity classification obsolete, viewing RBD progression as a continuous spectrum rather than a set of discrete categories. To further investigate the nature of misclassifications, the feature profiles across the four subgroups (TP, TN, FP, FN) were examined. A systematic similarity emerged between FN and TN profiles, indicating that misclassified patients tend to exhibit physiological characteristics more consistent with their assigned cluster than with their clinical label. This pattern may suggest the presence of patients whose biological features anticipate or partially reflect a more severe disease state,

potentially identifying individuals at risk of clinical progression. Consistent with this interpretation, the analysis of FCM membership confidence scores revealed no clear discriminative boundary between RBD1 and RBD2 patients. The overlap in membership distributions and the absence of extreme membership values across both correctly and incorrectly classified subjects point to a gradual transition between phenotypes rather than a sharp dichotomy, further supporting the view of RBD severity as a spectrum rather than a binary clinical construct.

4.4 Clustering on RBD and Healthy Controls – Derivable 3

Unsupervised clustering was applied to a combined cohort of RBD patients and healthy controls. Three clustering algorithms were evaluated under two configurations ($k = 2$ and $k = 3$), and the resulting clusters were interpreted with respect to clinical labels and their representation in principal component space. Overall performance was low: for $k = 2$, GMM achieved an Accuracy of 0.54 and an F1-score of 0.64 (Table 3.7), whereas for $k = 3$ Accuracy and F1-score decreased to 0.38 and 0.37, respectively (Table 3.9). This confirms that the clinical separation between HC, RBD1 and RBD2 is not supported by a corresponding differentiation in the features used for clustering. Further insight is provided by the analysis of the principal component projections. In the PC1–PC2 space, clustering solutions exhibit an apparently well-defined geometric separation. This separation is largely driven by PC1, which accounts for the largest proportion of variance, but does not align with the clinically relevant distinction. The lack of correspondence between predicted clusters and true clinical labels in this space suggests that PC1 primarily captures sources of variability unrelated to disease status or severity, such as demographic effects, inter-individual variability, or technical characteristics of the extracted features. In the PC2–PC3 space, the predicted representation shows weaker geometric separation.

Chapter 5

Conclusion

The objective of this study was to investigate the possible presence of distinct subgroups within the dataset by applying unsupervised clustering methods, to features extracted from ECG and body position signals. Additionally, by interpreting the differences among the identified clusters, the analysis aimed to explore whether such groupings could be associated with the presence of RBD and its severity. The initial clustering analysis consistently revealed three well-separated clusters across multiple methods, suggesting a degree of robustness in the underlying structure of the data. However, the clinical profiles associated with these three clusters, while coherent across methods, did not reach statistical significance, and the high variability in terms of standard deviation within each group limits the interpretability of the results. The clustering structure did not correspond to the clinical classification of subjects as healthy or affected by RBD. This observation was confirmed by additional clustering analyses performed on subsets of the original dataset, whose results were validated against the ground-truth labels indicating the presence or absence of RBD assigned by clinicians. Further insight was gained by visualizing the true clinical labels in the PC1–PC2 space: no clear groupings emerged with no discernible boundary separating the HC from RBD. This suggests that the feature space, as currently defined, does not encode the information necessary to distinguish RBD from healthy controls in a geometrically separable way. This outcome points to the inherent complexity of the problem: PCA revealed that variance is broadly distributed across many components, with a large number of features each contributing marginally to the principal components. Explaining a greater proportion of variance would require retaining more features per component, which in turn would demand a substantially larger dataset to avoid overfitting and ensure generalizability. What can be stated, however, is that the data consistently supports a subdivision of patients into three distinct groups. While this partition does not align with the clinical dichotomy of healthy versus RBD, it should not be

dismissed, rather, it calls for a different clinical interpretation. The extracted features, by their nature, are not exclusive to RBD, and previous literature associates them with a broad spectrum of conditions, including depression, bipolar disorder, anxiety, heart failure and sudden cardiac death.

Bibliography

[1] Irene Rechichi. «Sleep in Neurodegenerative Diseases: An Integrated Approach to Diagnosis and Monitoring». PhD thesis. Turin, Italy: Politecnico di Torino, 2024.

[2] Michael J Sateia. «International classification of sleep disorders». In: *Chest* 146.5 (2014), pp. 1387–1394.

[3] Michele T. Hu. «REM sleep behavior disorder (RBD)». In: *Neurobiology of Disease* 143, 104996 (2020).

[4] Shanti Chandra, Ambalika Sharma, and Girish Kumar Singh. «Feature extraction of ECG signal». In: *Journal of Medical Engineering* 42.4 (2018), pp. 306–316.

[5] Filipa Esgalhado, Arnaldo Batista, Valentina Vassilenko, Sara Russo, and Manuel Ortigueira. «Peak detection and HRV feature evaluation on ECG and PPG signals». In: *Symmetry* 14.6, 1139 (2022).

[6] Jeong Hun Yang, Sang Ho Choi, Mi Hyun Lee, Seong Min Oh, Jae-Won Choi, Jee Eun Park, Kwang Suk Park, and Yu Jin Lee. «Association of heart rate variability with REM sleep without atonia in idiopathic REM sleep behavior disorder». In: *Journal of Clinical Sleep Medicine* 17.3 (2021), pp. 461–469.

[7] Clara Gino, Laura Adelaide Dalla Vecchia, Monica Parati, Davide Sattin, Serena Covino, Graziella Cappelletti, Alessandro Pincherle, and Beatrice De Maria. «Role of heart rate variability analysis in differentiating patients with idiopathic REM sleep behaviour disorders from healthy subjects. A systematic review with meta-analysis». In: *Sleep Medicine Reviews* (2025), p. 102160.

[8] Masashi Suzuki et al. «Wearable sensor device-based detection of decreased heart rate variability in Parkinsons disease». In: *Journal of Neural Transmission* 129.10 (2022), pp. 1299–1306.

[9] Adeel A Memon, Ethan B George, Talha Nazir, Yatharth Sunkara, Corina Catiul, and Amy W Amara. «Heart rate variability during sleep in synucleinopathies: a review». In: *Frontiers in Neurology* 14 (2024), p. 1323454.

[10] Umberto Mosca, Irene Rechichi, Alessandro Cicolin, and Gabriella Olmo. «Exploring Heart Rate Variability Feature Importance for REM Sleep Behavior Disorder Classification: A Comprehensive Multi Dataset Study». In: (2025).

[11] Tam Pham, Carly J Johnco, Zen J Lau, Dominique Makowski, and Miriam K Forbes. «Which Heart Rate Variability (HRV) Indices Should I Use? A Data-driven Approach to Identifying Clusters of HRV Indices». In: (2025).

[12] Norbert Marwan, M Carmen Romano, Marco Thiel, and Jürgen Kurths. «Recurrence plots for the analysis of complex systems». In: *Physics reports* 438.5-6 (2007), pp. 237–329.

[13] Jerome Rolink, Martin Kutz, Pedro Fonseca, Xi Long, Berno Misgeld, and Steffen Leonhardt. «Recurrence quantification analysis across sleep stages». In: *Biomedical Signal Processing and Control* 20 (2015), pp. 107–116.

[14] Kentaro Kodama, Daichi Shimizu, Rick Dale, and Kazuki Sekine. «An approach to aligning categorical and continuous time series for studying the dynamics of complex human behavior». In: *Frontiers in Psychology* 12 (2021), p. 614431.

[15] Oriella Gnarra et al. «Exploring the association linking head position and sleep architecture to motor impairment in Parkinson's disease: an exploratory study». In: *Journal of Personalized Medicine* 13.11 (2023), p. 1591.

[16] Femke Dijkstra, Nathan Reyn, Barbara de Bruyn, Karlien Van den Bossche, Ilse de Volder, Marc Willemen, Mineke Viaene, Patrick Cras, David Crosiers, et al. «REM sleep without atonia and nocturnal body position in prediagnostic Parkinson's disease». In: *Sleep Medicine* 84 (2021), pp. 308–316.

[17] Galya Georgieva-Tsaneva, Krasimir Cheshmedzhiev, Yoan-Aleksandar Tsanev, and Miroslav Dechev. «Physiological State Recognition via HRV and Fractal Analysis Using AI and Unsupervised Clustering». In: *Information* 16.9 (2025), p. 718.

[18] Kamran Khan, Saif Ur Rehman, Kamran Aziz, Simon Fong, and Sababady Sarasvady. «DBSCAN: Past, present and future». In: *The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014)*. IEEE. 2014, pp. 232–238.

[19] Mohammadreza Ghorvei et al. «A comparative analysis of unsupervised machine-learning methods in PSG-related phenotyping». In: *Journal of Sleep Research* 34.3 (2025), e14349.

[20] Fred Shaffer and Jay P Ginsberg. «An overview of heart rate variability metrics and norms». In: *Frontiers in public health* 5 (2017), p. 258.

[21] Glaucylara Reis Geovanini, Enio Rodrigues Vasques, Rafael de Oliveira Alvim, José Geraldo Mill, Rodrigo Varejão Andreão, Bruna Kim Vasques, Alexandre Costa Pereira, and Jose Eduardo Krieger. «Age and sex differences in heart rate variability and vagal specific patterns–Baependi heart study». In: *Global heart* 15.1 (2020), p. 71.

[22] Robert M. Carney, Kenneth E. Freedland, Phyllis K. Stein, Judith A. Skala, Patricia Hoffman, and Allan S. Jaffe. «Change in Heart Rate and Heart Rate Variability During Treatment for Depression in Patients With Coronary Heart Disease». In: *Psychosomatic Medicine* 62.5 (2000), pp. 639–647.

[23] Fei Ge, Meng Yuan, Yan Li, and Wei Zhang. «Posttraumatic Stress Disorder and Alterations in Resting Heart Rate Variability: A Systematic Review and Meta-Analysis». In: *Psychiatry Investigation* 17.1 (2020), pp. 9–20.

[24] Sarah R. Vreijling, Yael Troudart, and Jos F. Brosschot. «Reduced Heart Rate Variability in Patients With Medically Unexplained Physical Symptoms: A Meta-Analysis of HF-HRV and RMSSD». In: *Psychosomatic Medicine* 83.1 (2021), pp. 2–15.

[25] Doo-Heum Park, Chul-Jin Shin, Seok-Chan Hong, Jaehak Yu, Seung-Ho Ryu, Eui-Joong Kim, Hong-Beom Shin, and Byoung-Hak Shin. «Correlation between the severity of obstructive sleep apnea and heart rate variability indices». In: *Journal of Korean medical science* 23.2 (2008), p. 226.

[26] Harald M. Stauss. «Heart Rate Variability». In: *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 285.5 (2003), R927–R931.

[27] Matteo Migliorini, Martin O. Mendez, and Anna M. Bianchi. «Study of Heart Rate Variability in Bipolar Disorder: Linear and Non-Linear Parameters during Sleep». In: *Frontiers in Neuroengineering* 4 (2011), p. 22.

[28] AM Bianchi, MO Mendez, Manuela Ferrario, L Ferini-Strambi, and Sergio Cerutti. «Long-term correlations and complexity analysis of the heart rate variability signal during sleep». In: *Methods of information in medicine* 49.05 (2010), pp. 479–483.

[29] Elisa Karhumaa, Antti Vuoti, Antti M Kiviniemi, M Juhani Junttila, Mikko P Tulppo, Heikki V Huikuri, Olavi H Ukkola, and Juha S Perkiömäki. «Changes and prognostic significance of autonomic cardiac regulation during ageing». In: *Autonomic Neuroscience* 258 (2025), p. 103255.

[30] Frank Beckers, Dirk Ramaekers, and André E Aubert. «Approximate entropy of heart rate variability: validation of methods and application in heart failure». In: *Cardiovascular Engineering: An International Journal* 1.4 (2001), pp. 177–182.