



Uniwersytet Zielonogórski
Szkoła Doktorska Nauk Ścisłych i Technicznych
Instytut Fizyki

Symetryczne i niesymetryczne parametry złożoności w wykrywaniu i różnicowaniu stanów fizjologicznych i patologicznych

*Symmetric and asymmetric complexity parameters in the
detection and differentiation of physiological and pathological
states*

Bartosz Biczuk

Rozprawa doktorska napisana
pod kierunkiem: dra hab. Jarosława Piskorskiego

promotor pomocniczy: dr Sebastian Żurek

Dziedzina: NAUK ŚCISŁYCH I PRZYRODNICZYCH
Dyscyplina: NAUKI FIZYCZNE

Akceptacja promotora:

Zielona Góra, 2025

Dedykacja
Dla mojej mamy Wioletty, w podziękowaniu za nieskończone wsparcie,
cierpliwość i wiarę we mnie.

Streszczenie

Niniejsza rozprawa doktorska skupia się na analizie własności rytmów serca (na podstawie sygnału EKG, a w szczególności odstępów RR) z wykorzystaniem serii monotonicznych oraz miar entropijnych. Rozwijana metodologia służy zarówno pogłębieniu teoretycznego zrozumienia zjawisk związanych ze zmiennością rytmu serca, jak i ocenie przydatności tych metod w praktycznych zastosowaniach diagnostycznych oraz klasyfikacyjnych.

W pierwszej kolejności przedstawiono studium literaturowe, w którym dokonano przeglądu kluczowych metod analizy zmienności rytmu serca (HRV), metod złożonościowych oraz wskaźników asymetrii rytmu serca (HRA), szczególnie w kontekście badań nad snem. Następnie wprowadzono teoretyczny model o charakterze ciągłym oparty na funkcji Weierstrassa z asymetrycznym szumem, stanowiący punkt wyjścia do badania wpływu asymetrii na parametry entropijne serii monotonicznych oraz porównania z innymi metodami, takimi jak HRA czy ADFA.

Kolejnym etapem było wykorzystanie tych metod do analizy rzeczywistych zapisów sygnałów EKG, w szczególności w kontekście klasyfikacji faz snu oraz detekcji migotania przedsionków. W pracy sprawdzono, czy entropia serii monotonicznych, w połączeniu z innymi miarami zmienności i złożoności rytmu serca, umożliwia skuteczne rozróżnianie stanów fizjologicznych i patologicznych.

Ostatnią częścią pracy jest opracowanie własnych narzędzi programistycznych, umożliwiających automatyczną analizę sygnałów EKG (m.in. w celu detekcji faz snu czy migotania przedsionków), a także przetwarzanie komercyjnych zapisów PSG tak, by nadawały się do dalszych badań naukowych. Dodatkowo udoskonalono algorytmy niezbędne do obliczania entropii z próby, zapewniając większą wydajność obliczeniową oraz transparentność ich działania.

W rezultacie praca dostarcza zarówno nowych wniosków teoretycznych dotyczących roli i struktur serii monotonicznych w sygnałach rytmu serca, jak i praktycznych narzędzi, które mogą znaleźć zastosowanie w badaniach naukowych i praktyce klinicznej, w szczególności w diagnostyce zaburzeń rytmu serca oraz analizie snu.

Abstract

This dissertation focuses on the analysis of heart rhythm properties (based on EKG signals, specifically RR intervals) using monotonic series and entropic measures. The developed methodology aims to deepen the theoretical understanding of phenomena related to heart rate variability (HRV) and assess the applicability of these methods for diagnostic and classification purposes.

First, a literature review is presented, surveying key HRV analysis techniques, complexity-based methods, and heart rhythm asymmetry (HRA) indicators, especially in the context of sleep research. Subsequently, a continuous theoretical model based on the Weierstrass function with asymmetric noise is introduced. This model serves as a foundation for examining how asymmetry influences entropic parameters of monotonic series and for comparing these parameters with other approaches such as HRA and ADFA.

Next, these methods are applied to real EKG recordings, particularly with respect to sleep stage classification and atrial fibrillation detection. The dissertation investigates whether the entropy of monotonic series, combined with other measures of heart rate variability and complexity, allows effective differentiation between physiological and pathological states.

The final part of the work concerns the development of custom software tools that enable automated EKG signal analysis (e.g., for sleep stage detection or atrial fibrillation). This includes the processing of commercial PSG recordings to make them suitable for further scientific studies. In addition, the algorithms required for calculating sample entropy were improved to offer greater computational efficiency and clarity.

As a result, the dissertation provides novel theoretical insights into the role and structure of monotonic series in heart rhythm signals and practical tools that could be applied in both scientific research and clinical practice, particularly for diagnosing heart rhythm disorders and analyzing sleep.

Spis treści

Spis treści	5
1 Wprowadzenie w problematykę badawczą	14
1.1 Podstawy elektrofizjologii serca	14
1.1.1 Szereg odstępów RR	16
1.1.2 Migotanie przedsionków	18
1.2 Sen i polisomnografia	19
1.2.1 Architektura snu	19
1.2.2 Polisomnografia	21
2 Opis teoretyczny	23
2.1 Zmienność rytmu serca	23
2.1.1 Domena czasowa	23
2.1.2 Domena częstotliwościowa	26
2.1.2.1 Periodogram Lomba-Scargle	29
2.2 Asymetryczna analiza fluktuacji z wytraceniem trendu . . .	31
2.3 Asymetria rytmu serca	32
2.4 Serie monotoniczne szeregu RR	37
2.4.1 Entropia informacyjna	40
2.4.2 Teoretyczny rozkład przetasowanego szeregu serii monotonicznych	41
2.5 Parametry entropijne	45
2.5.1 Entropia przybliżona	45
2.5.2 Entropia z próby	47
2.6 Metody uczenia maszynowego	48
2.6.1 Regresja logistyczna	49
2.6.2 Uogólnione Równania Estymujące (GEE)	50
2.6.3 XGBoost	52
2.7 Studium literaturowe związków HRV, złożonością i HRA z badaniami nad snem	53
2.7.1 Związek między snem a HRV	54

2.7.2	Związek między snem a złożonością	56
2.7.3	Związek między snem a HRA	57
2.7.4	Ograniczenia badań nad klasyfikacją faz snu opartych wyłącznie na szeregu odstępów RR	59
3	Materiały badawcze	61
3.1	Nagrania holterowskie	61
3.2	Nagrania 30 minutowe	62
3.3	Dane polisomnograficzne	62
4	Cele	69
5	Wyniki	72
5.1	Matematyczne własności entropii serii monotonicznych . . .	73
5.1.1	Wariancja sumy - dwa elementy	74
5.1.2	Wariancja sumy - n elementów	74
5.1.3	Wariancja funkcji z asymetrycznym szumem	75
5.1.4	Wyprowadzenie wariancji procesu opartego na funkcji Weierstrassa	76
5.1.5	Wyniki badania nad wykrywaniem asymetrii w sygna- łach Weierstrassa	78
5.1.5.1	Sygnał Weierstrassa bez asymetrii	79
5.1.5.2	Sygnał Weierstrassa z asymetrycznym szumem	81
5.1.6	Wnioski	83
5.2	Związek między deskryptorami serii monotonicznych a ADFA	83
5.2.1	Związki między obecnością asymetrii w deskrypto- rach serii a asymetrią w lokalnych asymetrycznych wykładnikach skalowania ADFA	83
5.2.2	Związki między asymetrią w deskryptorach entropij- nych HRA a ADFA	84
5.2.3	Dyskusja	84
5.3	Klasyfikacja faz snu na podstawie asymetrii rytmu serca . .	85
5.3.1	Statystyki opisowe – Rozkład faz snu	86
5.3.2	Analiza statystyczna	88
5.3.2.1	Faza N1	88
5.3.2.2	Faza N2	89
5.3.2.3	Przejście N1–N2	90
5.3.2.4	Przejście W–N1	90
5.3.2.5	Przejście N2–N3	91
5.3.2.6	Faza N3	92
5.3.2.7	Przejście N3–N2	92

5.3.2.8	Przejście N2-R	93
5.3.2.9	Przejście R-N2	93
5.3.2.10	Faza R	94
5.3.2.11	Przejście N3-R	94
5.3.2.12	Faza W (czuwanie)	95
5.3.3	Wyniki modelowania predykcyjnego	95
5.3.3.1	Faza N1	96
5.3.3.2	Faza N2	97
5.3.3.3	Faza N3	98
5.3.3.4	Faza R	99
5.3.3.5	Faza W	100
5.3.3.6	Przejście N2-N3	101
5.3.3.7	Przejście N2-R	102
5.3.3.8	Przejście N3-N2	103
5.3.3.9	Przejście N3-R	104
5.3.3.10	Przejście R-N2	105
5.3.3.11	Przejście W-N1	106
5.3.4	Dyskusja	106
5.3.5	Ograniczenia badania	110
5.4	Detekcja migotania przedsionków w oparciu o entropie serii .	112
5.4.1	Opis danych	113
5.4.2	Wartości diagnostyczne modeli tworzonych z pRR30, pRR3.25% oraz miarami entropii asymetrycznej . . .	116
5.4.3	Dyskusja	119
5.5	Refaktoryzacja algorytmu NCM	122
5.5.1	Algorytm NCM	122
5.5.2	Implementacja NCM: wersja <i>wprost</i>	127
5.5.3	Implementacja NCM: wersja <i>wzmocniona</i>	130
5.5.4	Analiza wydajności różnych implementacji algorytmów do obliczeń sum korelacyjnych	135
6	Wnioski	142
7	Podsumowanie	149
	Spis rysunków	152
	Spis tabel	154
	Bibliografia	157

Wstęp

Historia badań nad zmiennością rytmu serca (HRV, z ang. heart rate variability) sięga początków XX wieku, kiedy naukowcy po raz pierwszy zaobserwowali, że rytm serca podlega zmienności zależnej od czynników wewnętrznych i zewnętrznych [1]. Pionierzy badań nad HRV, jak William Einthoven – wynalazca elektrokardiografii (EKG) [2] – i późniejsi badacze, odkryli, że nie tylko rytm serca, ale także jego zmienność odgrywa istotną rolę w zdrowiu i fizjologii człowieka. Na przestrzeni dekad HRV stała się obiektem intensywnych badań, szczególnie ze względu na jej związek z aktywnością układu nerwowego autonomicznego (ANS, z ang. autonomic nervous system), a także jej wartość prognostyczną w modelach predykcyjnych dotyczących zdrowia i choroby [3].

Zmienność rytmu serca odzwierciedla reakcje organizmu na zmieniające się warunki zewnętrzne, takie jak stres, wysiłek fizyczny, czy stan emocjonalny, co czyni ją istotnym biomarkerem. Zjawisko to jest wypadkową dwóch podstawowych komponentów autonomicznego układu nerwowego: sympatycznego, odpowiedzialnego za reakcje „walki lub ucieczki,” oraz przywspółczulnego, aktywowanego w stanach odpoczynku i regeneracji. Liczne badania wykazały, że zróżnicowane wartości HRV mogą wskazywać na podatność na różne schorzenia, takie jak choroby sercowo-naczyniowe, zaburzenia psychiczne czy przewlekły stres [4]. W konsekwencji HRV znalazło szerokie zastosowanie w budowie modeli predykcyjnych, mających na celu prognozowanie zdrowotnych konsekwencji fizjologicznych oraz monitorowanie stanu zdrowia w czasie rzeczywistym.

Z biegiem lat rozwój technik przetwarzania danych i zaawansowanych metod analizy, takich jak algorytmy uczenia maszynowego, umożliwił dokładniejsze modele predykcyjne, w których HRV jest kluczowym parametrem wejściowym. Przykładowo, modele oparte na HRV są stosowane w predykcji nagłych incydentów sercowych, monitorowaniu stanu zdrowia pacjentów z chorobami przewlekłymi, a także w optymalizacji wydolności fizycznej u sportowców [5].

Interesującym kierunkiem badań nad analizą HRV są metody oparte

na mierzeniu złożoności sygnału, które dostarczają istotnych informacji o strukturze czasowej i dynamice układu autonomicznego. Jednym z przykładów takich podejść jest wykorzystanie entropii z próby (ang. *sample entropy*, SampEn), która ocenia nieregularność sygnału. SampEn znalazła zastosowanie w analizie HRV jako narzędzie do oceny autonomicznej regulacji serca, szczególnie w kontekście detekcji patologicznych zmian w systemie sercowo-naczyniowym. Na przykład Lake i współpracownicy wykorzystali entropię z próby do badania dynamiki rytmu serca u pacjentów z niewydolnością serca, wykazując jej zdolność do różnicowania stanów zdrowotnych i przewidywania ryzyka zdarzeń sercowych [6].

W trakcie badań nad HRV wyłoniło się także zjawisko asymetrii rytmu serca, które dopiero od niedawna wzbudza zainteresowanie naukowców. Tradycyjna analiza HRV opiera się na założeniu symetrycznych wahań czasowych między kolejnymi uderzeniami serca, jednakże nowe badania wykazały, że HRV może mieć wyraźnie asymetryczny charakter.

HRA oparta na wariancji kwantyfikuje różne wkłady zwolnień i przyspieszeń rytmu serca w zmienność odstępów RR, podkreślając, że zwolnienia mają większy wpływ na zmienność krótkoterminową, podczas gdy przyspieszenia bardziej wpływają na zmienność długoterminową i całkowitą [7]. HRA oparta na sekwencjach analizuje monotoniczne ciągi zwolnień i przyspieszeń, ujawniając, że przyspieszenia tworzą dłuższe i liczniejsze sekwencje niż zwolnienia [8].

Asymetria rytmu serca stanowi obiecujący, lecz wciąż niedostatecznie zbadany aspekt w budowie modeli predykcyjnych. Może dostarczyć nowych informacji o zdrowiu pacjenta, które są pomijane przy standardowej analizie HRV. Wprowadzenie deskryptorów asymetrii jako zmiennej predykcyjnej do modeli prognozujących zdrowie człowieka może potencjalnie zwiększyć dokładność prognozowania ryzyka wystąpienia zdarzeń sercowo-naczyniowych, a także umożliwić bardziej spersonalizowane podejście do monitorowania i oceny zdrowia. Dalsze badania nad asymetrią rytmu serca mogą przyczynić się do rozwinięcia nowych narzędzi analitycznych oraz wzbogacić obecne modele o dodatkowy wymiar, poprawiając ich efektywność w różnych kontekstach fizjologicznych.

Badania nad arytmiami serca, zwłaszcza migotaniem przedsionków (AF, z ang. atrial fibrillation), mają istotne znaczenie dla współczesnej medycyny ze względu na ich szerokie implikacje zdrowotne i społeczne. Migotanie przedsionków to jedna z najczęstszych form zaburzeń rytmu serca, która występuje szczególnie u osób starszych, a jej częstość rośnie wraz z wiekiem. To zaburzenie charakteryzuje się nieregularnym i często przyspieszonym rytmem serca, co prowadzi do pogorszenia efektywności pracy serca i może skutkować groźnymi powikłaniami, takimi jak udar mózgu, niewydolność

serca czy zatorowość płucna [9]. Z tego względu wczesne wykrywanie i monitorowanie AF jest kluczowe w profilaktyce oraz w optymalizacji opieki nad pacjentami zagrożonymi tym schorzeniem.

Badania nad migotaniem przedsionków koncentrują się na identyfikacji czynników ryzyka, takich jak nadciśnienie tętnicze, cukrzyca, otyłość oraz przewlekły stres, a także na opracowywaniu metod predykcyjnych, które pozwoliłyby na skuteczne monitorowanie osób z grup ryzyka [10]. W tym kontekście analiza zmienności rytmu serca, jego złożoności oraz nowych parametrów, takich jak asymetria rytmu serca, może dostarczyć dodatkowych informacji o funkcjonowaniu układu sercowo-naczyniowego. Algorytmy oparte na uczeniu maszynowym stają się coraz częściej stosowane w detekcji arytmii na podstawie sygnałów EKG oraz HRV, co może umożliwić precyzyjniejszą identyfikację osób zagrożonych migotaniem przedsionków [11].

Równoległe z badaniami nad arytmiami, dynamicznie rozwija się dziedzina badań nad snem oraz jego higieną, szczególnie dzięki postępom w technologii polisomnograficznej. Polisomnografia (PSG) to kompleksowa metoda badań, która obejmuje monitorowanie różnych parametrów fizjologicznych podczas snu, takich jak EEG (elektroencefalografia), EMG (elektromiografia), EOG (elektrookulografia), a także HRV. Badania polisomnograficzne są podstawowym narzędziem diagnozowania zaburzeń snu, takich jak bezdech senny, narkolepsja, czy zaburzenia ruchowe w czasie snu, które wpływają na zdrowie i jakość życia pacjentów [12].

Znaczenie higieny snu i jej wpływu na zdrowie człowieka zostało wielokrotnie udokumentowane. Sen pełni kluczową rolę w regeneracji organizmu, regulacji funkcji układu odpornościowego, procesach pamięciowych oraz w zachowaniu równowagi hormonalnej. Niedostateczna ilość snu, jego niska jakość lub zaburzenia snu są związane z wyższym ryzykiem chorób sercowo-naczyniowych, zaburzeń metabolicznych, a nawet zaburzeń psychicznych, takich jak depresja i lęk [13]. W tym kontekście HRV oraz inne parametry monitorowane podczas snu mogą dostarczyć wartościowych informacji o zdrowiu autonomicznego układu nerwowego i ogólnym stanie zdrowia pacjenta.

Polisomnografia, wsparta nowoczesnymi metodami analitycznymi, pozwala nie tylko na diagnozowanie, ale także na monitorowanie wpływu zmian w higienie snu na zdrowie pacjentów. Stosowanie odpowiednich technik analizy, takich jak analiza HRV w czasie snu, może pomóc w identyfikacji okresów stresu oraz regeneracji organizmu. Badania pokazują, że osoby z niską HRV podczas snu mogą być bardziej narażone na rozwój schorzeń, co czyni HRV, złożoność i HRA [14] potencjalnie użytecznym wskaźnikiem zdrowotnym w profilaktyce i monitorowaniu stanu zdrowia [15].

Kontynuacja badań nad snem, jego higieną oraz parametrami fizjologicznymi monitorowanymi podczas snu, jak HRV, umożliwia rozwijanie zaawansowanych modeli predykcyjnych, które mogą służyć do prognozowania zdrowia pacjentów, wczesnej diagnozy zaburzeń snu oraz oceny efektywności interwencji związanych z higieną snu. Dzięki dynamicznemu rozwojowi algorytmów sztucznej inteligencji oraz analizie dużych zbiorów danych, modele te stają się coraz dokładniejsze, oferując nowe możliwości diagnostyczne i terapeutyczne w obszarze medycyny snu oraz kardiologii.

Struktura pracy

Praca składa się ze wstępu i siedmiu rozdziałów. Pierwszy rozdział pełni funkcję wprowadzającą do tematyki badawczej. Przedstawiono w nim podstawy elektrokardiografii oraz zjawisko migotania przedsionków (AF). Opisano także podstawowe zasady polisomnografii, która stanowi standardową metodę oceny jakości snu. Omówiono charakterystykę poszczególnych faz snu, które stanowią istotny kontekst dla analizy fizjologicznych parametrów związanych z pracą serca. Drugi rozdział zawiera szczegółowy opis teoretyczny związany z analizą zmienności rytmu serca (HRV) oraz asymetrii rytmu serca (HRA). W tej części pracy scharakteryzowano najważniejsze parametry HRV oraz HRA, które są używane do ilościowej oceny regularności i dynamiki rytmu serca. Rozdział obejmuje również opis metod analizy serii monotonicznych szeregów odstępów RR oraz omówienie metod entropijnych, takich jak entropia Shannona oraz entropia z próby, które są stosowane jako miary złożoności sygnałów czasowych. Zawiera też szczegółowe studium literaturowe badań nad fazami snu przy użyciu metod opartych o zmienność rytmu serca, jego złożoność oraz asymetrię. Studium literaturowe stanowi istotny wkład badawczy w niniejszą rozprawę doktorską. Rozdział trzeci jest poświęcony celom pracy. Sformułowane zostały pytania badawcze oraz przedstawiono cele, które mają zostać osiągnięte w toku przeprowadzonych analiz. Cele te koncentrują się na analitycznej i numerycznej analizie własności parametrów HRV, wykorzystaniu parametrów HRV oraz HRA do budowy modeli predykcyjnych, umożliwiających m.in. detekcję migotania przedsionków oraz ocenę faz snu oraz rozwoju oprogramowania do obliczeń parametrów złożonościowych. Czwarty rozdział zawiera opis materiałów zastosowanych w niniejszej pracy. Zaprezentowano narzędzia pomiarowe oraz techniki analizy danych wykorzystane do oceny parametrów HRV i HRA oraz szczegółowo omówiono procedury związane z przetwarzaniem danych EKG i polisomnograficznych. Piąty rozdział przedstawia wyniki przeprowadzonych badań. W pierwszej części rozdziału opisano wyniki badań

teoretycznego rozkładu serii monotonicznych oraz zastosowanie tej metody na szeregach czasowych wygenerowanych przez funkcję Weierstrassa oraz realizacji tej funkcji z asymetrycznym szumem. Kolejna część rozdziału przedstawia wyniki badań nad budową modelu do klasyfikacji faz snu oraz przejść między fazami z wykorzystaniem parametrów HRA oraz rezultaty dotyczące detekcji migotania przedsionków, również oparte na metodach HRA. To właśnie w tej części są analizowane zebrane dane. Rozdział ten stanowi zasadniczą część pracy i przygotowanie wyników opisanych w tym rozdziale stanowiło większość wysiłku badawczego niniejszego doktoratu. Szósty rozdział poświęcony jest wnioskowi wyciągnięty z realizacji zaplanowanych zadań badawczych. Siódmy, ostatni rozdział stanowi podsumowanie pracy. Zawiera on podsumowanie uzyskanych wyników oraz ich implikacje w kontekście przyszłych badań nad zastosowaniem parametrów HRV i HRA w modelach predykcyjnych dla oceny zaburzeń rytmu serca oraz jakości snu.

Źródła naukowe, materiały badawcze oraz współpraca naukowa

Do materiałów źródłowych oraz materiałów badawczych (danych pomiarowych), wykorzystanych w przygotowaniu tej pracy doktorskiej, zaliczają się:

- krajowa i zagraniczna literatura naukowa związana z badaniami rytmu serca, metodami analizy szeregów czasowych, uczeniem maszynowym, statystyczną analizą danych pomiarowych,
- zapisy sygnałów polisomnograficznych, zarejestrowanych w *klinicznym oddziale neurologii* szpitala uniwersyteckiego w Zielonej Górze ; szczegółowy opis danych jak i procedur związanych z ich pozyskiwaniem zamieszczono w rozdziale piątym tej pracy.
- zapisy nagrań EKG uzyskane dzięki współpracy z pracownikami Katedry i Kliniki Intensywnej Terapii Kardiologicznej i Chorób Wewnętrznych Uniwersytetu Medycznego w Poznaniu.
- zapisy nagrań EKG z Long Term AF Database [16], szczegółowy opis znajduje się w rozdziale piątym.

Badania przedstawione w rozprawie zostały wykonane w grupie badawczej, utworzonej z pracowników *Instytutu Fizyki* Uniwersytetu Zielonogórskiego oraz pracowników *klinicznego oddziału neurologicznego* szpitala uniwersyteckiego w Zielonej Górze oraz kliniki Intensywnej Terapii Kardiologicznej i

Chorób Wewnętrznych Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu, Oddziału Klinicznego Intensywnej Terapii Kardiologicznej i Chorób Wewnętrznych Szpitala Klinicznego im. Heliodora Święcickiego w Poznaniu.

Podziękowania

Wyniki badań opisanych w kolejnych rozdziałach dysertacji uzyskano dzięki współpracy i zaangażowaniu grupy naukowej składającej się z lekarzy klinicznych, techników medycznych oraz fizyków.

Panu **dr hab. Jarosławowi Piskorskiemu – prof. UZ** – dziękuję za wszechstronne wsparcie, cierpliwość i życzliwość, a przede wszystkim za nieocenioną opiekę naukową, okazywaną mi przez cały okres przygotowywania niniejszej rozprawy. Jestem wdzięczny za liczne, cenne rady naukowe i życiowe, które otrzymywałem w kluczowych momentach pracy.

Panu **dr Sebastianowi Żurkowi** – dziękuję za ogromny wkład czasu i pełną zaangażowania współpracę. Pragnę wyrazić wdzięczność za wspólny trud włożony w badania, za liczne wyjazdy konferencyjne oraz za otwieranie drzwi, których sam nie byłbym w stanie przekroczyć.

Panu **dr Szymonowi Jurdze** – dziękuję za możliwość wykorzystania aparatury badawczej i zapewnienie dostępu do pracy z rzeczywistymi zapisami elektrokardiograficznymi, pochodzącymi z eksperymentu poświęconego badaniom nad zmiennością i złożonością rytmu serca oraz za merytoryczną pomoc, opartą na jego wiedzy medycznej. Wyrażam wdzięczność wszystkim osobom współuczestniczącym w skomplikowanym procesie pozyskiwania danych.

Panu **Marcinowi Wierzejewskiemu** w sposób szczególny pragnę podziękować, za wyjątkową życzliwość, bezinteresowną chęć niesienia pomocy i dzielenie się swoją obszerną wiedzą.

Panu **dr hab. n. med. Przemysławowi Guzikowi** dziękuję za udostępnione dane, za merytoryczne dyskusje i konstruktywną krytykę oraz za pozytywną mobilizację, która okazała się nieoceniona w wyjaśnianiu złożonych kwestii związanych z rytmem serca. Dzięki Jego wsparciu możliwe było przeprowadzenie kompleksowych analiz i weryfikacja wyników na podstawie rzeczywistych zapisów elektrokardiograficznych.

Rozdział 1

Wprowadzenie w problematykę badawczą

Rozdział wstępny poświęcony jest przedstawieniu podstaw elektrofizjologii oraz elektrokardiografii, w ramach których opisujemy pojedynczy cykl pracy serca. W tej części rozdziału omówimy mechanizmy generowania sygnałów elektrycznych przez serce, które stanowią podstawę dla interpretacji zapisu EKG, jak również wskażemy kluczowe parametry diagnostyczne, takie jak interwał RR, pozwalające na szczegółową analizę rytmu serca. Następnie, w drugiej części tego rozdziału, skupimy się na podstawach polisomnografii, w tym na opisie oraz charakterystyce poszczególnych faz snu. Omówimy znaczenie faz NREM i REM, ich kolejność, wraz z charakterystyką związaną z aktywnością mózgową i innymi parametrami fizjologicznymi.

1.1 Podstawy elektrofizjologii serca

Elektrokardiografia (EKG) jest nieinwazyjną metodą diagnostyczną, która pozwala na rejestrowanie i analizę elektrycznej aktywności serca. Od momentu jej wynalezienia przez Willema Einthovena w 1903 roku, EKG stało się jednym z najważniejszych narzędzi w kardiologii, umożliwiając ocenę rytmu serca, przewodzenia impulsów oraz wykrywanie różnych patologii układu sercowo-naczyniowego [17].

Serce jest organem mięśniowym, którego skoordynowana praca jest kontrolowana przez układ przewodzący, złożony z wyspecjalizowanych komórek zdolnych do generowania i przewodzenia impulsów elektrycznych. Głównymi elementami tego układu są:

- **Węzeł zatokowo-przedsionkowy (SA)** – naturalny rozrusznik serca, generujący impulsy elektryczne inicjujące każdy cykl serca,

- **Węzeł przedsionkowo-komorowy (AV)** – struktura opóźniająca przekazywanie impulsów z przedsionków do komór, co pozwala na ich skuteczne wypełnienie krwią przed skurczem komór,
- **Pęczek Hisa i włókna Purkinjego** – drogi przewodzenia impulsów do mięśnia komór, zapewniające synchroniczny skurcz serca.

Elektryczna aktywność serca wynika z cyklicznych zmian potencjałów błonowych komórek mięśniowych serca, które przechodzą przez fazy depolaryzacji i repolaryzacji. Te zmiany potencjałów generują pola elektryczne, które mogą być rejestrowane na powierzchni ciała za pomocą elektrod EKG.

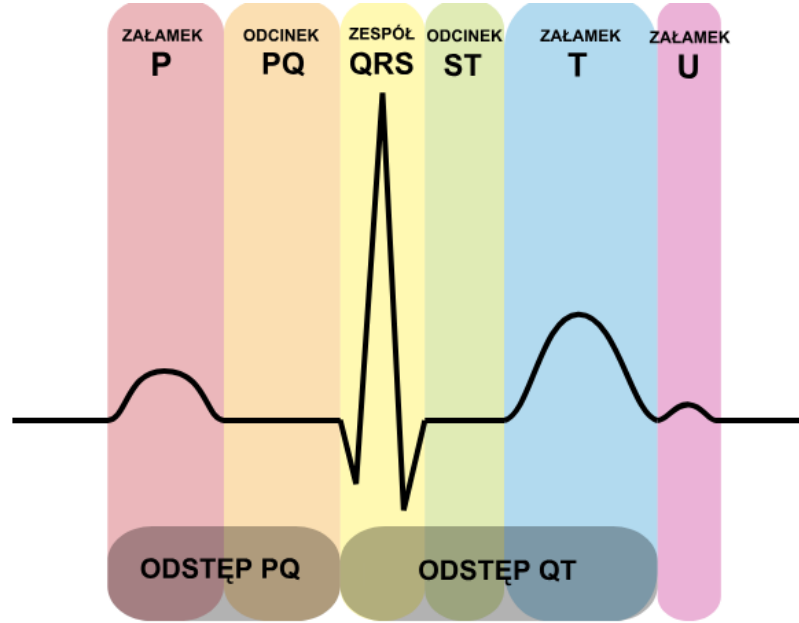
Elektrokardiograf mierzy różnice potencjałów elektrycznych pomiędzy dwiema lub więcej elektrodami umieszczonymi na powierzchni ciała pacjenta. Tradycyjny 12-odprowadzeniowy EKG obejmuje trzy odprowadzenia kończynowe dwubiegunowe (I, II, III), trzy odprowadzenia kończynowe jednobiegunowe wzmocnione (aVR, aVL, aVF) oraz sześć odprowadzeń przedsercowych (V1–V6). Każde z tych odprowadzeń rejestruje projekcję wektora elektrycznego serca w określonej płaszczyźnie, co pozwala na trójwymiarową ocenę jego aktywności [18].

Typowy zapis EKG składa się z charakterystycznych fal i odstępów, które odpowiadają poszczególnym fazom cyklu serca:

- **Załamek P** – reprezentuje depolaryzację przedsionków.
- **Odcinek PQ** - jest wyrazem przewodnictwa pobudzenia węzła przedsionkowo-komorowego,
- **Odstęp PR** – czas przewodzenia impulsu od węzła SA przez węzeł AV do komór,
- **Zespół QRS** – odpowiada depolaryzacji komór,
- **Odcinek ST** – okres pomiędzy depolaryzacją a repolaryzacją komór,
- **Załamek T** – reprezentuje repolaryzację komór.

Analiza tych elementów pozwala na ocenę rytmu i osi serca, wykrywanie zaburzeń przewodzenia, arytmii, niedokrwienia mięśnia sercowego oraz innych patologii [19].

Automatyczna analiza surowego nagrania EKG i wyszukiwanie w nim nieprawidłowości jest bardzo trudnym procesem co zaprezentowano w [20]



Rysunek 1.1: Przykład schemat pojedynczej ewolucji serca na EKG; źródło [21]

1.1.1 Szereg odstępów RR

Odstępy RR są definiowane jako różnice czasowe pomiędzy kolejnymi szczytami załamków R w sygnale EKG, które reprezentują depolaryzację komórek serca [22]. Matematycznie, odstęp RR można wyrazić jako:

$$RR_i = t_{R_{i+1}} - t_{R_i}, \quad (1.1)$$

gdzie t_{R_i} to czas wystąpienia i -tego załamka R. Odstępy RR są miarą tętna, ponieważ odwrotność odstępów RR odpowiada częstości rytmu serca:

$$HR_i = \frac{60}{RR_i}, \quad (1.2)$$

gdzie HR_i jest tętnem w uderzeniach na minutę.

Szereg odstępów RR , czyli ciąg kolejnych odstępów RR , jest kluczowy w badaniach kardiologicznych, ponieważ pozwala na analizę zmienności rytmu serca (HRV), która jest wskaźnikiem aktywności autonomicznego układu nerwowego [23].

Z matematycznego punktu widzenia, szereg odstępów RR jest procesem niestacjonarnym zarówno w sensie słabym, jak i mocnym. Niestacjonarność w sensie słabym oznacza, że średnia i autokowariancja procesu zmieniają

się w czasie [24]. W przypadku szeregu odstępów RR , średnia wartość odstępów może ulegać zmianom z powodu wpływu czynników fizjologicznych i patologicznych, takich jak aktywność fizyczna, stres czy arytmie. Ponadto, autokowariancja odstępów RR również nie jest stała w czasie, co jest konsekwencją dynamicznych zmian w regulacji rytmu serca.

Niestacjonarność w sensie mocnym oznacza, że rozkład prawdopodobieństwa procesu zmienia się w czasie. W szeregu odstępów RR , rozkład wartości odstępów może ulegać zmianom z powodu nagłych zdarzeń sercowych lub zmian w stanie fizjologicznym pacjenta [25].

Dla celów naszej pracy, analizowany sygnał powinien być pochodzenia zatokowego, co oznacza, że pochodzi z prawidłowego węzła zatokowego. Konieczne jest usunięcie artefaktów, takich jak załamki R wynikające z pobudzeń ektopowych (np. pobudzenia komorowe) oraz artefakty, ponieważ mogą one zafałszować analizę HRV [26].

Proces filtracji obejmuje identyfikację i usunięcie nieprawidłowych załamków R. Następnie należy upewnić się, że powstały szereg odstępów RR nie zawiera sztucznych odstępów, które nie odpowiadają rzeczywistym fizjologicznym interwałom.

Rozważmy przykład sygnału czasów załamków R:

$$[t_{R_1}, t_{R_2}, t_{R_3}, t_{R_4}], \quad (1.3)$$

gdzie t_{R_3} jest czasem wystąpienia artefaktycznego załamka R, który musi zostać usunięty. Po jego usunięciu nie można skonstruować odstępów RR jako:

$$RR_1 = t_{R_2} - t_{R_1}, \quad RR_2 = t_{R_4} - t_{R_2}, \quad (1.4)$$

ponieważ odstęp $t_{R_4} - t_{R_2}$ nie odzwierciedla rzeczywistego fizjologicznego interwału między sąsiednimi załawkami R w oryginalnym nagraniu. W rzeczywistości, odstęp ten obejmowałby dodatkowy cykl serca, który został pominięty z powodu usunięcia artefaktu. Dlatego jedynym prawidłowym odstępem RR w tym przypadku jest:

$$RR_1 = t_{R_2} - t_{R_1}. \quad (1.5)$$

Pominięcie tego faktu może prowadzić do błędnych wniosków w analizie HRV, dlatego tak ważne jest precyzyjne przetwarzanie sygnału EKG i prawidłowa konstrukcja szeregu odstępów RR [27].

1.1.2 Migotanie przedsionków

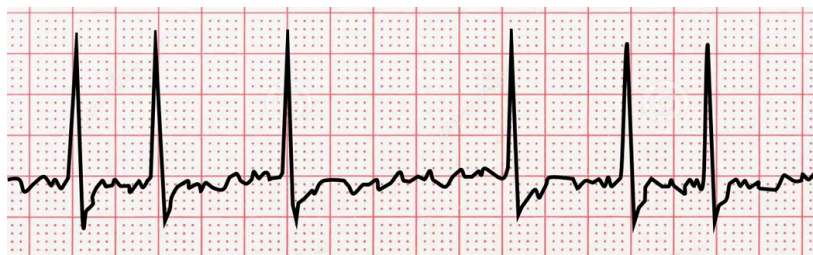
Migotanie przedsionków (ang. *Atrial Fibrillation*, AF) jest najczęstszą utrwaloną arytmia nadkomorową u dorosłych, charakteryzującą się nieskoordynowaną aktywnością elektryczną przedsionków, prowadzącą do ich nieskutecznej mechanicznej pracy [28]. Szacuje się, że dotyka około 1–2% populacji ogólnej, a częstość występowania wzrasta wraz z wiekiem [29].

W migotaniu przedsionków dochodzi do chaotycznej aktywności elektrycznej w przedsionkach, co skutkuje utratą ich efektywnego skurczu. Przyczyną są zwykle ogniska ektopowe generujące szybkie impulsy lub zjawisko reentry. To prowadzi do nieefektywnego przepływu krwi przez przedsionki i zwiększa ryzyko powstawania skrzeplin [30].

Pacjenci z migotaniem przedsionków mogą doświadczać objawów takich jak kołatanie serca, duszność, zawroty głowy czy zmęczenie. Jednak u niektórych osób arytmia może przebiegać bezobjawowo, co utrudnia jej wykrycie [28].

Migotanie przedsionków można rozpoznać na podstawie elektrokardiogramu (EKG). Charakterystyczne cechy migotania przedsionków na EKG to [31]:

- **Brak załamków P** – zamiast nich występują nieregularne, drobnofaliste fale migotania (f),
- **Nieregularne odstępy RR** – rytm komór jest całkowicie nieregularny (tzw. *arytmia absolutna*),
- **Wąskie zespoły QRS** – wskazujące na nadkomorowe pochodzenie rytmu.



Rysunek 1.2: Przykład EKG z migotaniem przedsionków; źródło [32]

Rozpoznanie AF wymaga rejestracji EKG trwającej co najmniej 30 sekund, aby potwierdzić obecność arytmii [33]. Wczesne wykrycie migotania przedsionków jest kluczowe ze względu na poważne powikłania związane z tą arytmia. Należą do nich [30]:

- **Udar mózgu** – AF zwiększa ryzyko udaru pięciokrotnie, co wynika z powstawania skrzeplin w lewym przedsionku i ich ewentualnej embolizacji do naczyń mózgowych,
- **Niewydolność serca** – przewlekłe AF może prowadzić do osłabienia funkcji serca z powodu utraty efektywnego skurczu przedsionków i niekontrolowanego rytmu komór,
- **Zwiększona śmiertelność** – AF jest związane z wyższym ryzykiem zgonu z przyczyn sercowo-naczyniowych.

Wczesna diagnoza pozwala na wdrożenie odpowiedniego leczenia, takiego jak antykoagulacja w celu zmniejszenia ryzyka udaru, kontrola rytmu lub częstości serca, oraz interwencje inwazyjne jak ablacja [33].

1.2 Sen i polisomnografia

1.2.1 Architektura snu

Sen jest naturalnym, cyklicznie powtarzającym się stanem organizmu, charakteryzującym się zmniejszoną świadomością, obniżoną reakcją na bodźce zewnętrzne oraz specyficznymi wzorcami aktywności mózgu [34]. Pełni kluczową rolę w procesach regeneracji fizycznej i psychicznej, konsolidacji pamięci oraz regulacji funkcji metabolicznych i hormonalnych [13].

Regulacja snu jest wynikiem działania dwóch głównych procesów:

- **Proces homeostatyczny** (*Proces S*) – nasilenie potrzeby snu w zależności od czasu czuwania,
- **Proces okołodobowy** (*Proces C*) – rytm biologiczny kontrolowany przez zegar biologiczny w jądrze nadskrzyżowaniowym podwzgórza, synchronizowany przez bodźce zewnętrzne, takie jak światło [35].

Sen nie jest stanem jednorodnym; składa się z cyklicznie następujących po sobie faz, tworzących tzw. architekturę snu. Wyróżnia się dwie główne fazy:

- **Sen NREM** (*Non-Rapid Eye Movement*) – sen wolnofalowy, dzielony na trzy stadia (N1, N2, N3),
- **Sen REM** (*Rapid Eye Movement*) – faza snu z szybkimi ruchami gałek ocznych.

Stadium N1 Jest to najlżejsze stadium snu, stan przejściowy między czuwaniem a snem. Charakteryzuje się spowolnieniem aktywności fal mózgowych (pojawienie się fal theta o częstotliwości 4–7 Hz) w EEG, zmniejszeniem napięcia mięśniowego oraz sporadycznymi ruchami gałek ocznych [36].

Stadium N2 Stanowi około 50% całkowitego czasu snu. W tym stadium występują wrzeciona snu (fale o częstotliwości 12–14 Hz) w EEG oraz zespoły K (krótkotrwałe, wysokowoltazowe fale), które są charakterystyczne dla tego etapu [37].

Stadium N3 Znane jako sen głęboki lub sen wolnofalowy. Charakteryzuje się dominacją fal delta (o częstotliwości 0,5–2 Hz) w EEG. Jest to stadium najtrudniejsze do przerywania, pełniące kluczową rolę w procesach regeneracyjnych organizmu [38].

Stadium REM Faza REM występuje cyklicznie co około 90 minut i stanowi około 20–25% całkowitego czasu snu. Charakteryzuje się:

- **Szybkimi ruchami gałek ocznych** – stąd nazwa fazy.
- **Desynchronizacją EEG** – aktywność mózgu podobna do stanu czuwania (fale beta o częstotliwości powyżej 14 Hz).
- **Atonią mięśniową** – zahamowanie napięcia mięśni szkieletowych, co zapobiega realizacji ruchów podczas snu.

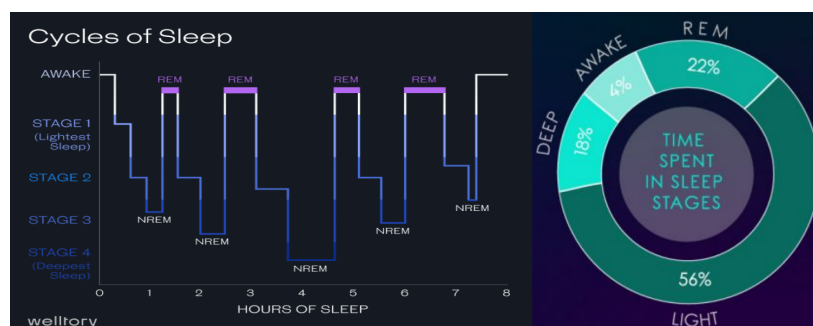
Faza REM jest uważana za kluczową dla procesów konsolidacji pamięci oraz przetwarzania emocjonalnego [39]. Przejścia między tymi etapami odgrywają kluczową rolę w zrozumieniu architektury snu. Przejście z fazy N3 do N2 oznacza spłycenie snu, które często występuje tuż przed wejściem w fazę REM lub krótkim wybudzeniem [40]. Bezpośrednie przejścia z N3 do REM są rzadkie i zazwyczaj wymagają etapu pośredniego; kiedy jednak się pojawiają, wskazują na znaczącą zmianę aktywności mózgu z procesów regeneracyjnych na funkcje poznawcze [34]. Z kolei przejście z fazy REM do N2 jest powszechne i świadczy o zakończeniu epizodu REM, po czym zwykle następuje powrót do N2 przed przejściem w inne fazy snu [41].

Podczas typowej nocy zdrowy dorosły przechodzi przez 4–6 cykli snu, z których każdy trwa około 90–110 minut. W miarę postępu nocy udział fazy REM zwiększa się kosztem snu N3 [36].

Zrozumienie tych przejść oraz dynamiki poszczególnych faz snu ma kluczowe znaczenie dla dokładnej detekcji i analizy wzorców snu, co może

mieć istotne konsekwencje zarówno w praktyce klinicznej, jak i w badaniach naukowych. Klasyfikacją faz snu i przejść między nimi najczęściej zajmuje się wykwalifikowany personel na podstawie zapisu EEG, ale w tym zadaniu pomocne są również pozostałe sygnały dostępne w trakcie badania polisomnograficznego [34].

Spośród tych sygnałów, szereg odstępów RR — będąca pochodną zapisu EKG — wydaje się najbardziej przystępna, praktyczna i najtańsza w analizie. Wynika to z dobrze zdefiniowanego charakteru EKG i stosunkowo łatwej identyfikacji załamka R oraz pomiaru odległości między kolejnymi pobudzeniami. Prawdopodobnie właśnie dlatego tak wiele urządzeń typu wearable czyli urządzenia mogące być noszone jako akcesoria odzieżowe pozwalające między innymi na monitoring procesów fizjologicznych bazuje na tym sygnale, wykorzystując go w różnych analizach [42].



Rysunek 1.3: Przykład hipnogramu-wykresu przedstawiającego zmiany faz snu oraz diagram kołowy procentowego udziału czasowego danej fazy snu do całości snu; źródło [43]

1.2.2 Polisomnografia

Polisomnografia (PSG) jest kluczowym narzędziem w diagnostyce i leczeniu zaburzeń snu, które mają istotny wpływ na zdrowie i jakość życia pacjentów. Pozwala na dokładne zrozumienie mechanizmów leżących u podstaw tych zaburzeń oraz monitorowanie skuteczności terapii [44].

PSG jest wieloparametrowym, nieinwazyjnym badaniem diagnostycznym służącym do oceny czynności fizjologicznych podczas snu. Pozwala na jednoczesną rejestrację i analizę wielu parametrów, takich jak aktywność mózgu, ruchy gałek ocznych, napięcie mięśniowe, czynność serca, oddychanie oraz ruchy kończyn [40].

- **Elektroencefalografia (EEG)** – rejestracja aktywności elektrycznej mózgu.

- **Elektrookulografia (EOG)** – rejestracja ruchów gałek ocznych.
- **Elektromiografia (EMG)** – rejestracja napięcia mięśniowego, zwykle mięśni bródki.
- **Elektrokardiografia (EKG)** – monitorowanie czynności serca.
- **Badanie parametrów oddechowych** – przepływ powietrza przez nos i usta, ruchy klatki piersiowej i brzucha, saturacja krwi tlenem (SpO_2).
- **Aktywność ruchowa** – rejestracja ruchów kończyn, szczególnie nóg.

Polisomnografia jest złotym standardem w diagnostyce zaburzeń snu, takich jak [45]:

- **Obturacyjny bezdech senny** – powtarzające się epizody zamknięcia górnych dróg oddechowych podczas snu.
- **Narkolepsja** – nadmierna senność dzienna z nagłymi napadami snu.
- **Zaburzenia ruchowe podczas snu** – np. zespół niespokojnych nóg.
- **Parasomnie** – nieprawidłowe zachowania podczas snu, takie jak somnambulizm czy lęki nocne.

Badanie polisomnograficzne jest zwykle przeprowadzane w specjalistycznym laboratorium snu. Pacjent spędza noc w komfortowych warunkach, a do jego ciała podłączane są czujniki rejestrujące wymienione wcześniej parametry. Dane są zapisywane i analizowane przez wykwalifikowany personel medyczny [40].

Po zakończeniu badania dane są poddawane szczegółowej analizie. Na podstawie zapisu EEG dokonuje się podziału snu na poszczególne fazy i stadia. Analiza parametrów oddechowych pozwala na wykrycie zaburzeń oddychania, takich jak bezdechy i spłylenia oddechu. Wyniki polisomnografii są interpretowane w kontekście klinicznym pacjenta, co umożliwia postawienie diagnozy i zaplanowanie odpowiedniego leczenia [46].

W ostatnich latach rozwijane są nieinwazyjne, przenośne urządzenia umożliwiające przeprowadzanie uproszczonych badań polisomnograficznych w warunkach domowych. Ponadto, zastosowanie sztucznej inteligencji i algorytmów uczenia maszynowego pozwala na automatyczną analizę i interpretację danych, co zwiększa dostępność i efektywność diagnostyki zaburzeń snu [47].

Rozdział 2

Opis teoretyczny

W niniejszym rozdziale przedstawiono metody wykorzystane w pracy, ze szczególnym uwzględnieniem ich podstaw teoretycznych. Kluczowe aspekty obejmują metody służące do opisu zmienności rytmu serca (HRV, jego asymetrii oraz metody entropijne, ze szczególnym naciskiem na entropię z próby (*Sample Entropy*), wykorzystaną w tej pracy jako miara złożoności sygnałów RR. Wymienione metody są podstawą mojej pracy. Kolejnym istotnym elementem są metody uczenia maszynowego zastosowane do budowy modeli predykcyjnych oraz walidacji uzyskanych wyników. W ramach tego omówienia przedstawione zostaną ich matematyczne podstawy oraz metody oceny ich efektywności i wiarygodności.

2.1 Zmienność rytmu serca

Zmienność rytmu serca można mierzyć przy użyciu wielu różnych metod. Szereg odstępów RR, podobnie jak wiele realistycznych szeregów czasowych, posiada złożoną strukturę, której nie jest w stanie w pełni opisać jedna znana metoda analityczna. Do tej pory nie został również opracowany model matematyczny zdolny do dokładnego odwzorowania struktury zmienności rytmu serca i generowania realistycznych odstępów RR. W poniższej sekcji zostaną omówione metody wykorzystane dalszej części tej pracy, podzielone na domeny czasową i częstotliwościową.

2.1.1 Domena czasowa

Parametr *SDNN* (Standard Deviation of Normal to Normal intervals) jest miarą całkowitej zmienności rytmu serca i oblicza się go na podstawie odstępów między kolejnymi uderzeniami serca (*RR*). Definiuje się go wzorem:

$$SDNN = \sqrt{\frac{1}{n} \sum_{i=1}^n (RR_i - \bar{RR})^2}, \quad (2.1)$$

gdzie n to liczba odstępów RR w analizowanym przedziale, a $\bar{RR}$ to średnia wartość tych odstępów:

$$\bar{RR} = \frac{1}{n} \sum_{i=1}^n RR_i. \quad (2.2)$$

Ponieważ szereg RR jest na ogół niestacjonarny, średnia i odchylenie standardowe mogą zmieniać się w czasie w zależności od wybranego przedziału analizy [48]. Oznacza to, że $SDNN$ nie ma stałych właściwości statystycznych i wymaga standaryzacji długości nagrania w badaniach. Najczęściej analizuje się zapisy o długości 5, 30 minut lub 24 godzin [23].

Wykres Poincaré jest narzędziem graficznym służącym do analizy zmienności rytmu serca poprzez przedstawienie zależności między kolejnymi odstępami RR . Każdy punkt na wykresie reprezentuje parę kolejnych odstępów (RR_i, RR_{i+1}) [49, 50].

Deskryptory $SD1$ i $SD2$ służą do ilościowego opisu kształtu chmury punktów na wykresie Poincaré. Parametr $SD1$ odnosi się do zmienności krótkoterminowej i jest związany z rozproszeniem punktów w kierunku prostopadłym do linii identyczności (prosta $y = x$). Natomiast $SD2$ opisuje zmienność długoterminową i mierzy rozproszenie wzdłuż linii identyczności [51, 52].

Parametry $SD1$ i $SD2$ można wyrazić matematycznie jako:

$$SD1^2 = \text{Var} \left(\frac{RR_{i+1} - RR_i}{\sqrt{2}} \right), \quad (2.3)$$

$$SD2^2 = \text{Var} \left(\frac{RR_{i+1} + RR_i}{\sqrt{2}} \right). \quad (2.4)$$

W tych wzorach Var oznacza wariancję odpowiednich kombinacji odstępów RR .

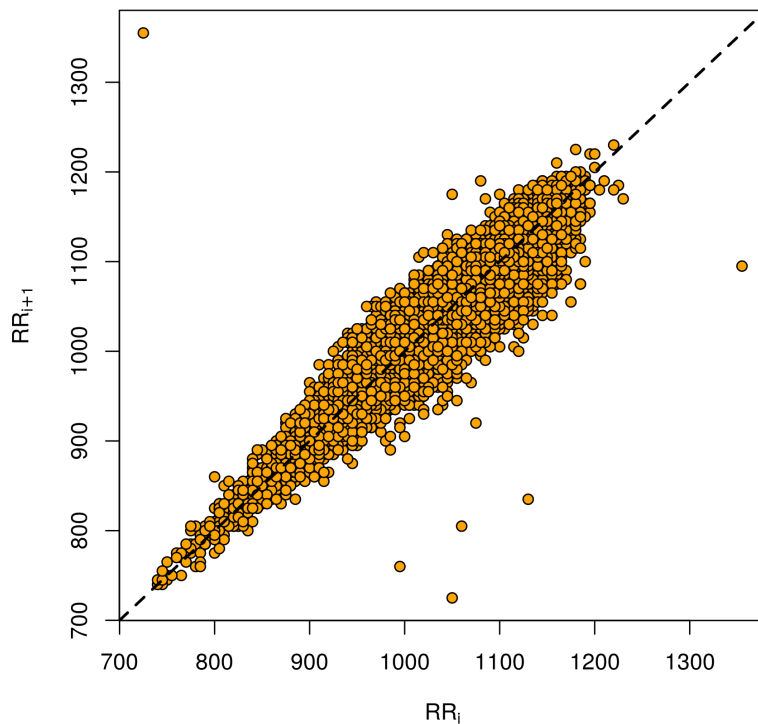
Parametr $SD1$ odzwierciedla krótkoterminową zmienność rytmu serca i jest wrażliwy na wpływ układu przywspółczulnego [23]. Parametr $SD2$ reprezentuje długoterminowe zmiany rytmu serca, związane z aktywnością zarówno układu przywspółczulnego, jak i współczulnego [9]. Istnieje zależność łącząca parametry $SDNN$, $SD1$ i $SD2$:

$$SDNN^2 = \frac{1}{2}(SD1^2 + SD2^2). \quad (2.5)$$

Ta relacja wynika z własności wariancji i sposobu definiowania parametrów $SD1$ i $SD2$. Aby skonstruować wykres Poincaré:

1. uzyskujemy szereg odstępów RR z zapisu EKG,
2. na osi poziomej odkładamy RR_i , a na osi pionowej RR_{i+1} ,
3. otrzymujemy chmurę punktów, której kształt analizujemy za pomocą parametrów $SD1$ i $SD2$.

Na Rysunku 2.1 przedstawiono przykładowy wykres Poincaré wraz z oznaczeniem linii identyczności oraz kierunków odpowiadających parametrom $SD1$ i $SD2$. Ważnym statystycznym wskaźnikiem procentowym, który określa



Rysunek 2.1: Przykład wykresu Poincaré odstępów RR uzyskany z całonocnego zapisu EKG. Grafika wyeksportowana z oprogramowania *HRAexplorer* [53].

odsetek par kolejnych odstępów NN różniących się o więcej niż 50 milisekund

jest paramter pNN50 (*percentage of NN50*) [23]. Matematycznie wyraża się to jako:

$$\text{pNN50} = \frac{\text{NN50}}{N - 1} \times 100\%, \quad (2.6)$$

gdzie:

- NN50 — liczba par kolejnych odstępów RR spełniających warunek $|RR_{i+1} - RR_i| > 50 \text{ ms}$,
- N — całkowita liczba odstępów RR.

pNN50 jest miarą aktywności przywspółczulnej i odzwierciedla zmienność rytmu serca na krótkich skalach czasowych.

Pochodne tego wskaźnika, takie jak pNN x i pNN $x\%$, generalizują tę miarę poprzez zastosowanie innego progu x niż 50 ms:

$$\text{pNN}x = \frac{\text{NN}x}{N - 1} \times 100\%, \quad (2.7)$$

gdzie NN x to liczba par odstępów RR różniących się o więcej niż x ms [54]. Wybór wartości x pozwala na dostosowanie analizy do specyfiki badanego sygnału.

2.1.2 Domena częstotliwościowa

Analiza spektralna jest jedną z kluczowych metod badania szeregu odstępów RR w zmienności rytmu serca (HRV, ang. *Heart Rate Variability*). Większość wniosków dotyczących dynamiki rytmu serca, jego źródeł oraz czynników wpływających została sformułowana na podstawie badań w dziedzinie częstotliwości [23, 22, 55].

Istotą analizy spektralnej jest określenie, jak wariancja sygnału (w tym kontekście nazywana mocą spektralną) jest rozłożona w funkcji częstotliwości w analizowanym zakresie. Standardowo oblicza się funkcję gęstości widmowej mocy (ang. *Power Spectral Density*, PSD), a ilość wariancji przypadającej na określone pasmo częstotliwości otrzymuje się poprzez całkowanie PSD w granicach tego pasma [56]. Funkcję PSD można obliczyć bezpośrednio, stosując metody nieparametryczne, lub estymować ją w metodach parametrycznych za pomocą odpowiedniego modelu – rezultatem tego procesu jest periodogram [57]. Najpopularniejszą i najbardziej dostępną w narzędziach analitycznych metodą nieparametryczną jest szybka transformata Fouriera (FFT, ang. *Fast Fourier Transform*). Metoda ta wykorzystuje pełne dane zawarte w szeregu czasowym, jest szybka, a jej algorytm jest stosunkowo prosty do zrozumienia [58]. Metody parametryczne, takie jak modelowanie

autoregresyjne (AR), estymują spektrum za pomocą modelu matematycznego, co oznacza, że dane nie są w pełni reprezentowane w periodogramie. Jednakże metoda ta dostarcza oszacowań spektrum wraz z niepewnościami, co umożliwia prowadzenie statystycznych wnioskowań [59].

Obie powyższe metody wymagają resamplingu, czyli interpolacji tachogramu (szeregu odstępów RR) do próbek w równych odstępach czasu. Zazwyczaj resampłowanie przeprowadza się z częstotliwością 4 Hz. Dodatkowo konieczne jest przetwarzanie uderzeń ektopowych oraz artefaktów poprzez ich interpolację lub usunięcie [26]. Dla dyskretnego sygnału czasowego $x(n)$ o N próbkach, dyskretna transformata Fouriera (DFT) jest zdefiniowana jako:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N}, \quad k = 0, 1, \dots, N-1, \quad (2.8)$$

gdzie $X(k)$ to współczynniki Fouriera, a j jest jednostką urojoną. Funkcja gęstości widmowej mocy (PSD) jest następnie obliczana jako moduł kwadratowy współczynników Fouriera:

$$P(k) = \frac{1}{N} |X(k)|^2. \quad (2.9)$$

Obliczenie PSD pozwala na analizę, jak moc sygnału jest rozłożona w funkcji częstotliwości. Szybka transformata Fouriera (FFT) jest algorytmem efektywnego obliczania dyskretnego transformaty Fouriera (DFT). Zamiast bezpośredniego obliczania sumy w DFT, które wymaga $O(N^2)$ operacji, FFT wykorzystuje faktoryzację liczby próbek N do podziału obliczeń na mniejsze części, co redukuje złożoność obliczeniową do $O(N \log N)$ [58].

Algorytm FFT polega w uproszczeniu na:

1. podzieleniu sygnału na parzyste i nieparzyste próbki,
2. rekursywnym obliczeniu DFT dla tych podzbiorów,
3. połączeniu wyników z wykorzystaniem symetrii i okresowości funkcji wykładniczych.

Dzięki temu FFT jest znacznie szybsza niż bezpośrednio obliczanie DFT, szczególnie dla dużych N . Analiza widmowa szeregu odstępów RR zazwyczaj skupia się na obliczaniu mocy sygnału w określonych przedziałach częstotliwości. Dla krótkich zapisów (2–5 minut) wyróżnia się następujące pasma [23]:

- **wysoka częstotliwość (HF):** 0,15–0,4 Hz,

- **wiska częstotliwość (LF)**: 0,04–0,15 Hz,
- **bardzo niska częstotliwość (VLF)**: poniżej 0,04 Hz.

W bardzo długich nagraniach dodatkowo wyodrębnia się pasmo ultraniskich częstotliwości (ULF): poniżej 0,003 Hz.

Moc w paśmie HF jest generalnie uważana za odzwierciedlenie aktywności układu przywspółczulnego oraz wpływu oddychania na rytm serca [55]. Pasmo LF jest związane zarówno z aktywnością układu współczulnego, jak i przywspółczulnego. Natomiast moce w paśmie VLF i ULF są często trudne do interpretacji i uważa się, że są silnie zależne od warunków rejestracji oraz zastosowanych metod przetwarzania sygnału [60].

Warto zauważyć, że powyższe podziały pasm częstotliwości są umowne i nie zawsze mają ścisłe podstawy fizjologiczne. Na przykład częstość oddychania może w niektórych przypadkach wpływać na pasmo LF zamiast HF.

Jednym z głównych problemów w analizie szeregu odstępów RR za pomocą metod spektralnych jest niestacjonarność sygnału. Chociaż technicznie każdy szereg czasowy można poddać transformacji Fouriera, to jednak skład widmowy wynikający z tej analizy może nie odzwierciedlać rzeczywistych oscylacji w układzie sercowo-naczyniowym. Szczególnie wolne trendy oraz efekty brzegowe mogą znacząco wpływać na widmo, zwłaszcza w niskich częstotliwościach [61]. Istnieją metody mające na celu usuwanie niestacjonarności i minimalizowanie wpływu gwałtownych zmian fazy, jednak często ingerują one w strukturę sygnału, co może być niepożądane. Dlatego w niektórych przypadkach korzystne jest zastosowanie innych metod analizy dostosowanych do specyfiki sygnału HRV. Podczas standardowej analizy spektralnej wykorzystuje się jedynie moduł transformaty Fouriera (widmo amplitudowe), pomijając informację o fazie sygnału. Odrzucenie komponentu fazowego uniemożliwia pełną rekonstrukcję sygnału w dziedzinie czasu na podstawie samego widma amplitudowego. Ma to znaczenie w kontekście analizy wzorców przyspieszeń i zwolnień rytmu serca, które są zależne od kształtu szeregu RR. Resampling sygnału jest konieczny w metodach opartych na transformacie Fouriera, ponieważ wymaga ona równomiernego próbkowania sygnału. Tachogramy są naturalnie nierównomiernie próbkowane ze względu na zmienność rytmu serca. Resampling polega na interpolacji sygnału w równych odstępach czasu, co pozwala na zastosowanie FFT [61]. Artefakty i pobudzenia ektopowe również muszą być przetworzone przed analizą spektralną, zazwyczaj poprzez ich usunięcie lub interpolację, aby nie zaburzały wyników analizy.

2.1.2.1 Periodogram Lomb-Scargle

Metoda Lomb-Scargle stanowi podejście do estymacji widma mocy sygnału z próbek rozmieszczonych w niejednakowych odstępach czasu [62, 63]. W przeciwieństwie do klasycznej transformacji Fouriera, która wymaga jednostajnego próbkowania w dziedzinie czasu, metoda Lomb-Scargle pozwala na bezpośrednią analizę szeregów o nieregularnym rozmieszczeniu próbek. Jest ona szczególnie użyteczna w przypadku szeregów czasowych o nierównej częstości próbkowania, w tym szeregów z długimi przerwami pomiędzy próbkami. Cecha ta sprawia, że po pierwsze nie trzeba przeprowadzać resamplingu szeregu RR, po drugie, w przypadku pobudzeń ektopowych czy artefaktów technicznych, nie ma potrzeby interpolowania wartości pomiędzy pobudzeniami zatokowymi.

Metoda Lomb-Scargle opiera się na analizie wariancji sygnału w dziedzinie częstości (lub pulsacji ω) przy uwzględnieniu rzeczywistych momentów próbkowania t_i . Rozważmy szereg odstępów RR obserwowanych w chwilach czasu t_i (gdzie $i = 1, 2, \dots, N$). Periodogram Lomb-Scargle można zdefiniować w następujący sposób [62, 63]:

$$P_X(\omega) = \frac{1}{2SDNN^2} \left[\frac{\left(\sum_{i=1}^N [RR(t_i) - \overline{RR}] \cos[\omega(t_i - \tau)] \right)^2}{\sum_{i=1}^N \cos^2[\omega(t_i - \tau)]} + \frac{\left(\sum_{i=1}^N [RR(t_i) - \overline{RR}] \sin[\omega(t_i - \tau)] \right)^2}{\sum_{i=1}^N \sin^2[\omega(t_i - \tau)]} \right], \quad (2.10)$$

gdzie $\overline{RR}$ jest średnią z próbek $RR(t_i)$, $SDNN^2$ to ich wariancja, natomiast τ jest przesunięciem fazowym definiowanym przez:

$$\tan[2\omega\tau] = \frac{\sum_{i=1}^N \sin(2\omega t_i)}{\sum_{i=1}^N \cos(2\omega t_i)}. \quad (2.11)$$

t_i tutaj jest zdefiniowane jako suma kumulatywna szeregu RR, tj. zbioru $\{t_0 = RR_0, t_1 = RR_0 + RR_1, t_2 = RR_0 + RR_1 + RR_2, \dots, \sum_{i=0}^n RR_i\}$, odpowiadające poszczególnym odstępom RR w następujący sposób:

$$RR(t_i) \equiv RR_i \equiv t_i - t_{i-1}, \quad (2.12)$$

W publikacji [64], autorzy przedstawili skuteczność metod bazujących na Lomb-Scargle (LSP) przy wyznaczaniu widma szeregu RR w obecności luk pomiarowych, wynikających choćby z usuwania pobudzeń ektopowych

czy braku równomiernych odstępów czasowych między kolejnymi uderzeniami serca. Z kolei w innej pracy [65], pojawia się koncepcja uśredniania uzyskanych widm (LSP) po pulach okien czasowych o identycznej szerokości, wybranych z pojedynczego szeregu.

Metoda uśrednionego periodogramu Lomb-Scargle’a (ang. *averaged Lomb-Scargle periodogram*, ALSP), została zaproponowana w [66] i bazuje na schemacie przedstawionym w [65]. Podobnie jak metoda opisana w [65], opiera się na sumowaniu wielu widm w jedną reprezentatywną krzywą. Jednocześnie różni się w sposobie interpretowania informacji oraz w pewnych własnościach matematycznych, gdyż w ALSP widma powstają z odmiennych szeregów (pochodzących od różnych osób), a nie z okien wycinkowych tego samego zapisu.

Pierwsza kluczowa rozbieżność w stosunku do konwencjonalnej definicji (2.10) dotyczy denormalizacji otrzymanego periodogramu Lomb-Scargle’a. Zamiast posługiwać się całościową wariancją szeregu SDN_N^2 , wykorzystujemy informację jedynie z przedziału częstotliwości 0,04–0,4 Hz. W związku z tym, w równaniu (2.10), wartość SDN_N^2 zostaje zastąpiona wielkością σ , przy czym

$$\sigma = \sum_{f \in [0.04, 0.4 \text{ Hz}]} A_f^2, \quad (2.13)$$

gdzie A_f^2 odnosi się do kwadratu amplitudy sinusoidy o częstotliwości f . Tak skorygowany periodogram, będący już częściowo pozbawiony wpływu mniej istotnych komponentów, podlega następnie procesowi uśredniania:

$$P_X(f) = \frac{1}{M} \sum_{k=1}^M P_X^k(f), \quad (2.14)$$

gdzie M oznacza liczbę niezależnych szeregów RR, a $P_X^k(f)$ jest standardowym periodogramem LSP dla k -tego sygnału.

Uśredniony periodogram Lomb-Scargle’a zapewnia wgląd w statystycznie najbardziej charakterystyczne częstości, obejmujące dynamikę wielu różnych przebiegów RR. Jeżeli dane częstości w całej populacji szeregów wносиły podobny udział w całkowitej wariancji (tzn. wszędzie występuje zbliżony wysoki lub niski poziom mocy), wówczas takie częstości zostaną wzmocnione w wyniku sumowania (2.21). Równocześnie wartości częstości, których wkład w całkowitą moc znacząco się różni między poszczególnymi badaniami, zostają w tym procesie osłabione lub nawet zniwelowane. Dzięki temu otrzymuje się przejrzysty obraz wspólnej charakterystyki spektralnej w wybranym zakresie 0,04–0,4 Hz [66].

2.2 Asymetryczna analiza fluktuacji z wytraceniem trendu

Analiza fluktuacji z wytraceniem trendu (Detrended Fluctuation Analysis, DFA) jest jedną z najczęściej stosowanych metod analizy szeregów czasowych odstępów RR . Główną informacją dostarczaną przez tę metodę są własności skalowania szumu pozostałego po wytraceniu trendu z szeregów czasowych. Szczegóły metody można znaleźć w [67, 68, 69], a poniżej zostaną one szybko przedstawione w celu ustalenia notacji.

Zdefiniujmy szeregi czasowe odstępów RR jako odległość między kolejnymi załamkami R w elektrokardiogramie [50, 70] w następujący sposób:

$$\mathbf{RR} = (RR_1, RR_2, \dots, RR_n). \quad (2.15)$$

Zdefiniujmy także pochodną, sumowaną i średnio-wyzerowaną serię czasową:

$$y(k) = \sum_{i=1}^k RR_i - \overline{RR}, \quad (2.16)$$

gdzie $\overline{RR}$ oznacza średnią z całego szeregu czasowego, a y_k definiuje tzw. "okna" o długości n , w których trend jest wyznaczany i odejmowany. Ponieważ naszym celem jest identyfikacja trendów rosnących i malejących, używamy jedynie wielomianów pierwszego stopnia (liniowych):

$$y(k) = y_n(k) + \epsilon_k, \quad (2.17)$$

gdzie $y_n(k)$ to linia dopasowana do konkretnego okna, a ϵ_k to składnik błędu, który jest przedmiotem naszych badań. Funkcja średniokwadratowa jest zdefiniowana jako:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(k) - y_n(k))^2}. \quad (2.18)$$

i obliczana dla wszystkich skal dostępnych w analizowanym szeregu czasowym - w praktyce n zmienia się od 4 do $N/4$, gdzie N to długość szeregu czasowego. Wartości $\log_{10}(F(n))$ są wykreślane względem $\log_{10}(n)$, a jeśli uzyskany wykres tworzy linię prostą, wnioskuje się o istnieniu prawa potęgowego w skalowaniu funkcji średniokwadratowej, tj.

$$F(n) \sim n^\alpha, \quad (2.19)$$

gdzie α to wykładnik skalowania. Wartości tego wykładnika interpretuje się jako wskazujące na obecność negatywnych długozasięgowych korelacji

($0 < \alpha < 0.5$), szumu białego ($\alpha = 0.5$), pozytywnych długozasięgowych korelacji ($0.5 < \alpha < 1$), szumu $1/f$ ($\alpha = 1$), długozasięgowych korelacji nienależących do prawa potęgowego ($1 < \alpha < 1.5$) lub zgodności szeregów czasowych z przypadkowym spacerem ($\alpha = 1.5$) [67, 68, 69].

Fakt, że lokalne trendy w oknach mogą być liniowe, umożliwia zdefiniowanie dwóch różnych funkcji średniokwadratowych w zależności od znaku tangensa dopasowanej linii.

Definiując M jako całkowitą liczbę segmentów, a M_n^+ jako liczbę segmentów, w których trend jest rosnący (lub, używając równania (2.17), $a_n > 0$), definiujemy:

$$F^+(n) = \sqrt{\frac{1}{M_n^+ n} \sum_{j=1}^M \delta^+(j) \sum_{k=1}^n (y((j-1) \cdot n + k) - y_n((j-1) \cdot n + k))^2}, \quad (2.20)$$

a analogicznie dla trendów malejących:

$$F^-(n) = \sqrt{\frac{1}{M_n^- n} \sum_{j=1}^M \delta^-(j) \sum_{k=1}^n (y((j-1) \cdot n + k) - y_n((j-1) \cdot n + k))^2}, \quad (2.21)$$

gdzie zewnętrzne sumowanie (po j) obejmuje wszystkie okna, $\delta(\pm)$ wybiera okna z rosnącymi (równanie (2.20)) lub malejącymi trendami (równanie (2.21)) wyłącznie [71, 72].

Jeśli powyższe funkcje zostaną przedstawione na podwójnie logarytmicznej skali względem n , mogą być zdefiniowane dwa wykładniki skalowania:

$$F^\pm(n) \sim n^{\alpha^\pm}, \quad (2.22)$$

o ile zależność ta jest liniowa.

W pracy [73] rozwinięto lokalną wersję ADFA, w której wykładniki skalowania $\alpha^\pm$ są obliczane w oknie przesuwającym się wzdłuż analizowanego szeregu czasowego. Alternatywą dla ADFA w badaniach zmienności rytmu serca może być asymetryczna wersja wieloskalowej, wielofrakalnej analizy (ang. *Asymmetric multiscale multifractal analysis*, AMMA) opierającej się na analizie powierzchni Hursta [74].

2.3 Asymetria rytmu serca

Asymetria rytmu serca (ang. *heart rate asymmetry*, HRA) jest niedawno odkrytym zjawiskiem fizjologicznym, polegającym na tym, że przyspieszenia

rytmu serca mają inny wpływ na jego zmienność niż zwolnienia. Konkretnie, wkład zwolnień do krótkoterminowej zmienności rytmu serca jest mniejszy niż wkład przyspieszeń, natomiast w zmienności długoterminowej i całkowitej przyspieszenia dominują nad zwolnieniami [7].

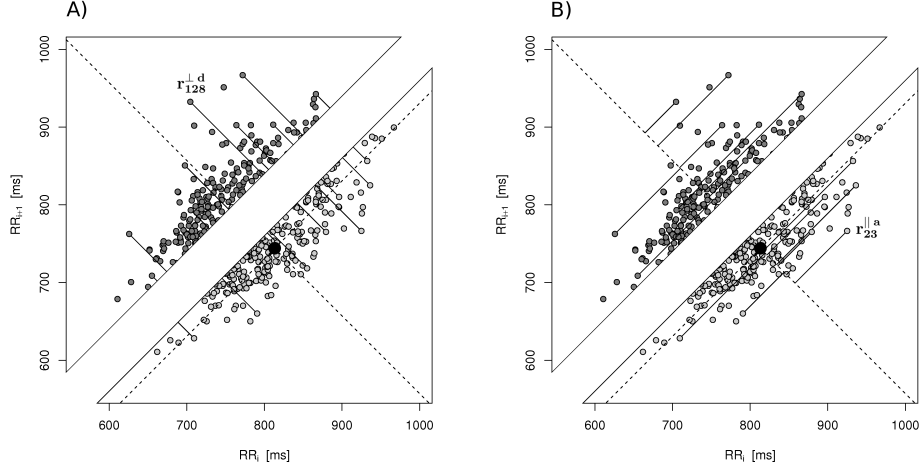
Wykres Poincaré (Poincaré) jest narzędziem graficznym służącym do analizy zmienności rytmu serca. Na tym wykresie linia identyczności odgrywa kluczową rolę, dzieląc go na dwa obszary: punkty powyżej linii identyczności reprezentują zwolnienia (następny odstęp RR jest dłuższy od poprzedniego), a punkty poniżej odpowiadają przyspieszeniom. Punkty leżące dokładnie na linii identyczności wskazują brak zmiany długości odstępu RR .

Matematycznie, parametry wykresu Poincaré, takie jak $SD1^2$, $SD2^2$ oraz $SDNN^2$, można podzielić na części zależne wyłącznie od zwolnień lub przyspieszeń. Dzięki temu uzyskujemy następujące składowe:

- $SD1^2d$ – część $SD1^2$ zależna tylko od zwolnień,
- $SD1^2a$ – część $SD1^2$ zależna tylko od przyspieszeń,
- $SD2^2d$ – część $SD2^2$ zależna tylko od zwolnień,
- $SD2^2a$ – część $SD2^2$ zależna tylko od przyspieszeń,
- $SDNN^2d$ – część $SDNN^2$ zależna tylko od zwolnień,
- $SDNN^2a$ – część $SDNN^2$ zależna tylko od przyspieszeń,
- $C1_d$ – udział zwolnień w zmienności krótkoterminowej,
- $C1_a$ – udział przyspieszeń w zmienności krótkoterminowej,
- $C2_d$ – udział zwolnień w zmienności długoterminowej,
- $C2_a$ – udział przyspieszeń w zmienności długoterminowej,
- C_d – udział zwolnień w zmienności całkowitej,
- C_a – udział przyspieszeń w zmienności całkowitej.

Warto zauważyć, że użycie kwadratów poszczególnych deskryptorów wynika z właściwości matematycznych wariancji – w nieparametrycznych testach statystycznych można stosować zarówno kwadraty, jak i wartości przed podniesieniem do potęgi drugiej.

Proces tworzenia deskryptorów asymetrii krótkoterminowej przedstawiono na rysunku 2.2 w panelu A), natomiast konstrukcję deskryptorów asymetrii długoterminowej zaprezentowano w panelu B).



Rysunek 2.2: Grafika A) przedstawia sposób konstrukcji deskryptorów asymetrii krótkoterminowej, natomiast obrazek B) pokazuje proces budowy deskryptorów asymetrii długoterminowej. W obu przypadkach kluczową rolę odgrywa linia identyczności, która określa, czy dany punkt przyczynia się do części wariancji ($SD1^2$, $SD2^2$ lub $SDNN^2$) związanej ze zwolnieniami lub przyspieszeniami. $r_{128}^{\perp;d}$ oznacza prostopadłą odległość punktu odpowiadającego zwolnieniu o numerze 128 na wykresie Poincaré od linii identyczności, natomiast $r_{23}^{\parallel;a}$ reprezentuje prostopadłą odległość punktu przyspieszenia o numerze 23 od linii l_2 [75]. Grafika na licencji Creative Commons.

Istnienie asymetrii definiuje się jako sytuację, w której w grupie N badanych osób zachodzi:

$$C1_d > C1_a, \quad (2.23)$$

czyli udział zwolnień w zmienności krótkoterminowej jest statystycznie istotnie większy niż udział przyspieszeń. Asymetria długoterminowa występuje, gdy:

$$C2_d < C2_a, \quad (2.24)$$

co oznacza, że udział przyspieszeń w zmienności długoterminowej jest statystycznie istotnie większy niż udział zwolnień. Asymetrię całkowitą określa się, gdy:

$$C_d < C_a, \quad (2.25)$$

również w stopniu istotnym statystycznie.

Badania wykazały, że asymetria krótkoterminowa jest obecna w szeregach odstępów RR zarejestrowanych u zdrowych osób w spoczynku, a jej obecność jest wysoce istotna statystycznie. Podobnie, asymetria długoterminowa została potwierdzona w takich samych warunkach. U zdrowych osób w spoczynku stwierdzono również asymetrię rytmu serca względem podziału $SDNN^2$, co jest statystycznie istotne. Kliniczne zastosowania wariacyjnych parametrów asymetrii rytmu serca zostały opisane w literaturze.

Podobne analizy można przeprowadzić dla ciśnienia tętniczego – wykazano, że sygnał ciśnienia również wykazuje asymetrię krótkoterminową, niezależną od asymetrii rytmu serca.

Wariancję odpowiadającą zmienności krótkoterminowej $SD1^2$ można rozłożyć na części odpowiadające zwolnieniom i przyspieszeniom. Wprowadzamy następujące oznaczenia:

1. $r^{\perp;d}_i$ – prostopadła odległość zwalniającego punktu wykresu Poincaré o indeksie i od linii identyczności,
2. $r^{\perp;a}_j$ – prostopadła odległość przyspieszającego punktu wykresu Poincaré o indeksie j od linii identyczności.

Zmienność krótkoterminową $SD1^2$ możemy podzielić na części odpowiadające zwolnieniom i przyspieszeniom:

$$SD1^2 = \frac{1}{n} \left(\sum_{i=1}^{n_d} [r^{\perp;d}_i]^2 + \sum_{j=1}^{n_a} [r^{\perp;a}_j]^2 \right), \quad (2.26)$$

gdzie n_d to liczba zwolnień (punktów powyżej linii identyczności), a n_a to liczba przyspieszeń (punktów poniżej linii identyczności). Punkty na linii identyczności nie wnoszą wkładu do zmienności krótkoterminowej, więc:

$$n = n_d + n_a + n_{on}, \quad (2.27)$$

gdzie n_{on} to liczba punktów na linii identyczności.

Podział ten prowadzi do definicji deskryptorów odpowiadających wkładowi zwolnień i przyspieszeń do $SD1^2$:

$$SD1^2 = SD1_d^2 + SD1_a^2, \quad (2.28)$$

gdzie:

$$SD1_d^2 = \frac{1}{n} \sum_{i=1}^{n_d} [r^{\perp;d}_i]^2, \quad SD1_a^2 = \frac{1}{n} \sum_{j=1}^{n_a} [r^{\perp;a}_j]^2. \quad (2.29)$$

Aby umożliwić porównanie pomiarów o różnej wariancji, definiujemy znormalizowane wkłady zwolnień i przyspieszeń do $SD1^2$:

$$C1_d = \frac{SD1_d^2}{SD1^2}, \quad C1_a = \frac{SD1_a^2}{SD1^2}, \quad (2.30)$$

przy czym:

$$C1_d + C1_a = 1. \quad (2.31)$$

W przypadku podziału $SD2^2$, również używamy linii identyczności jako kryterium podziału, ze względu na jej jasną interpretację fizjologiczną. Definiujemy:

1. $r^{\parallel;d}_i$ – odległość wzdłuż linii identyczności zwalniającego punktu wykresu Poincaré o indeksie i od linii $l2$,
2. $r^{\parallel;a}_j$ – odległość wzdłuż linii identyczności przyspieszającego punktu o indeksie j od linii $l2$.

Wariancję długoterminową $SD2^2$ wyrażamy jako:

$$SD2^2 = \frac{1}{n} \left(\sum_{i=1}^{n_d} [r^{\parallel;d}_i]^2 + \sum_{j=1}^{n_a} [r^{\parallel;a}_j]^2 + \sum_{k=1}^{n_{on}} [r_k^{\parallel;on}]^2 \right). \quad (2.32)$$

Aby uzyskać pełny podział $SD2^2$, wkład punktów na linii identyczności dzielimy po równo między zwolnienia i przyspieszenia:

$$SD2^2 = SD2_d^2 + SD2_a^2, \quad (2.33)$$

gdzie:

$$SD2_d^2 = \frac{1}{n} \left(\sum_{i=1}^{n_d} [r^{\parallel;d}_i]^2 + \frac{1}{2} \sum_{k=1}^{n_{on}} [r_k^{\parallel;on}]^2 \right), \quad (2.34)$$

$$SD2_a^2 = \frac{1}{n} \left(\sum_{j=1}^{n_a} [r^{\parallel;a}_j]^2 + \frac{1}{2} \sum_{k=1}^{n_{on}} [r_k^{\parallel;on}]^2 \right). \quad (2.35)$$

Znormalizowane wkłady zwolnień i przyspieszeń do $SD2^2$ definiujemy jako:

$$C2_d = \frac{SD2_d^2}{SD2^2}, \quad C2_a = \frac{SD2_a^2}{SD2^2}, \quad (2.36)$$

z warunkiem:

$$C2_d + C2_a = 1. \quad (2.37)$$

Pełną zmienność $SDNN^2$ możemy podzielić na części związane ze zwolnieniami i przyspieszeniami, korzystając z relacji:

$$SDNN^2 = \frac{1}{2} (SD1^2 + SD2^2). \quad (2.38)$$

Definiujemy:

$$SDNN^2 = SDNN_d^2 + SDNN_a^2, \quad (2.39)$$

gdzie:

$$SDNN_d^2 = \frac{1}{2} (SD1_d^2 + SD2_d^2), \quad SDNN_a^2 = \frac{1}{2} (SD1_a^2 + SD2_a^2). \quad (2.40)$$

Znormalizowane wkłady zwolnień i przyspieszeń do $SDNN^2$ definiujemy jako:

$$C_d = \frac{SDNN_d^2}{SDNN^2}, \quad C_a = \frac{SDNN_a^2}{SDNN^2}, \quad (2.41)$$

przy czym:

$$C_d + C_a = 1. \quad (2.42)$$

Podsumowując, asymetria rytmu serca odzwierciedla nierówny wkład zwolnień i przyspieszeń w zmienność rytmu serca. Dzięki matematycznemu podziałowi wariancji na części związane z tymi dwoma zjawiskami, możliwa jest precyzyjna analiza ich wpływu na rytm serca oraz ich znaczenie kliniczne.

2.4 Serie monotoniczne szeregu RR

W celu bardziej szczegółowego opisu asymetrii rytmu serca w niniejszej pracy wykorzystamy analizę serii monotonicznych szeregu RR. Analiza sekwencji tych odstępów pozwala na ocenę funkcjonowania układu autonomicznego oraz wykrywanie potencjalnych zaburzeń rytmu serca. W szczególności, badanie monotonicznych serii odstępów RR umożliwia zrozumienie dynamiki przyspieszeń i zwolnień rytmu serca.

Definiujemy szereg odstępów RR jako:

$$\mathbf{RR}_n \equiv (RR_1, RR_2, \dots, RR_n). \quad (2.43)$$

Aby zbadać zmiany między kolejnymi odstępami, przekształcamy ten szereg na szereg różnic pierwszych:

$$\Delta \equiv (\delta_1, \delta_2, \dots, \delta_{n-1}), \quad (2.44)$$

gdzie:

$$\delta_i = RR_{i+1} - RR_i. \quad (2.45)$$

Różnice δ_i mogą być dodatnie (przyspieszenie rytmu serca), ujemne (zwolnienie rytmu serca) lub równe zero (brak zmiany). Wprowadzamy trójelementowe odwzorowanie symboliczne tych różnic:

$$\text{sgn}(\delta_i) = \begin{cases} + \\ - \\ 0 \end{cases}. \quad (2.46)$$

Przy użyciu powyższego zapisu możemy zdefiniować *monotoniczne serie zwolnień i przyspieszeń*.

Definicja 1: *Monotoniczna seria zwolnień* o długości i (DRi) to segment symbolicznego szeregu (2.46), który rozpoczyna się i kończy jednym ze znaków „-” lub „0” i składa się z i kolejnych znaków „+”.

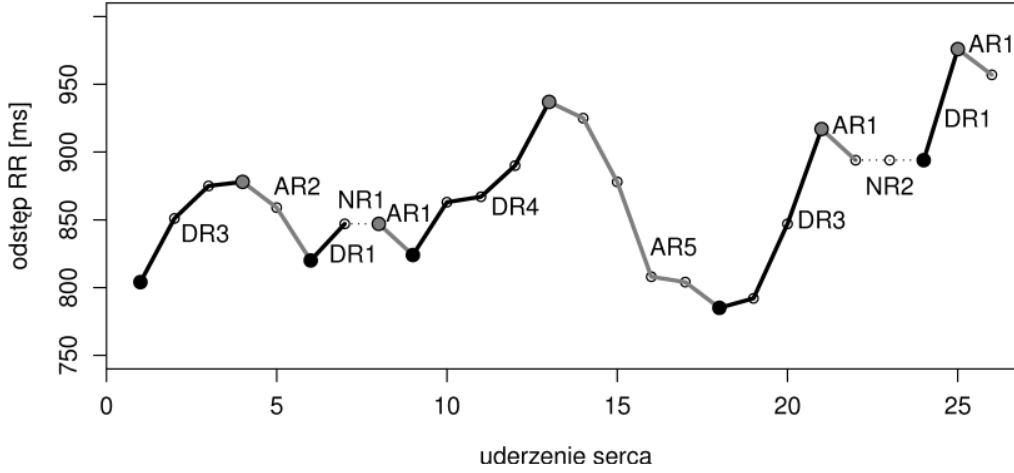
Definicja 2: *Monotoniczna seria przyspieszeń* o długości i (ARi) to segment symbolicznego szeregu (2.46), który rozpoczyna się i kończy jednym ze znaków „+” lub „0” i składa się z i kolejnych znaków „-”.

Skróty DR oraz AR pochodzą z nazw angielskich *Deceleration Run* oraz *Acceleration Run*. Poniższa tabela zawiera kilka przykładów monotonicznych serii zwolnień i przyspieszeń wraz z oznaczeniami, opisanymi przy pomocy symboli zdefiniowanych w (2.46).

Symbol	Seria
AR1	0 - +
AR4	+ - - - - +
DR3	- + + + -
DR7	- + + + + + + + 0

Tabela 2.1: Przykłady monotonicznych serii.

Procedura powyższa dzieli szereg odstępów $\mathbf{RR}_n$, zrzucony na szereg symboliczny (2.46), na rozłączne serie zwolnień i przyspieszeń oraz serie niezmiennych odstępów. Przykład takiego podziału pokazany jest na rysunku 2.3.



Rysunek 2.3: Podział fragmentu tachogramu według serii monotonicznych, które oznaczone są jako DR_i (seria zwolnień długości i) oraz AR_i (seria przyspieszeń długości i). Symbole N_i oznaczają serie neutralne, w których kolejne odstępy RR są równe i które mogą przerywać serie zwolnień i przyspieszeń. Pełne szare punkty oznaczają początki serii zwolnień, a pełne czarne punkty rozpoczynają serie przyspieszeń – są to punkty odniesienia dla tych serii [75]. Grafika na licencji Creative Commons.

Serie monotoniczne zwolnień i przyspieszeń są zliczane, więc najwygodniejszymi parametrami podsumowującymi pełen rozkład są parametry zliczające. Podstawowym takim parametrem jest entropia Shannona, która podsumowuje rozkład prawdopodobieństwa przy pomocy jednej liczby. Dla entropii Shannona można wydzielić składowe odpowiadające poszczególnym elementom – w naszym przypadku będzie to funkcja częstości występowania poszczególnych serii.

Entropię Shannona dla rozkładu prawdopodobieństwa, że dany RR będzie częścią pewnej serii:

$$\begin{aligned}
 H_{Runs} = & - \sum_{i=1}^{\max(i)_{Dec}} p_{i,Dec} \cdot \log i_{i,Dec} \\
 & - \sum_{i=1}^{\max(i)_{Acc}} p_{i,Acc} \cdot \log i_{i,Acc}.
 \end{aligned} \tag{2.47}$$

Wyrażenie to ma dwie części – pierwsza z nich zależy jedynie od zwolnień, druga jedynie od przyspieszeń. W związku z tym możemy zdefiniować

$$H_{rD} = \sum_{i=1}^{\max(i)_{Dec}} p_{i,Dec} \cdot \log p_{i,Dec}, \quad (2.48)$$

oraz

$$H_{rA} = \sum_{i=1}^{\max(i)_{Acc}} p_{i,Acc} \cdot \log p_{i,Acc}, \quad (2.49)$$

gdzie $p_{i,Dec}$ i $p_{i,Acc}$ są prawdopodobieństwami wystąpienia odpowiednio serii zwolnień i przyspieszeń o długości i .

2.4.1 Entropia informacyjna

Entropia jest fundamentalnym pojęciem w teorii informacji, wprowadzonym przez Claude'a E. Shannona w 1948 roku [76]. Służy jako miara niepewności lub złożoności związanej z losowym źródłem informacji. Entropia pozwala na ilościowe ujęcie informacji, jaką można uzyskać z danego źródła, oraz na ocenę efektywności systemów kodowania i przesyłania danych. Entropia Shannona $H(X)$ zdefiniowana jest następująco: niech X będzie dyskretną zmienną losową $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ przyjmującą wartości z prawdopodobieństwami $P(X = x_i) = p_i$, gdzie $i = 1, 2, \dots, n$. Entropia Shannona $H(X)$ jest zdefiniowana jako:

$$H(X) = - \sum_{i=1}^n p_i \ln p_i. \quad (2.50)$$

Entropia wyraża średnią ilość informacji (w *natach*) uzyskiwaną na sygnał wygenerowany przez źródło X . Jednostka *nat* jest używana, gdy stosujemy logarytm naturalny. Jeśli stosujemy logarytm o podstawie 2, jednostką entropii jest *bit*. Różnica między mierzeniem entropii w bitach i natakach polega na podstawie logarytmu użytej w definicji. Konwersja między nimi jest możliwa poprzez mnożenie przez stałą:

$$1 \text{ nat} = \frac{1}{\ln 2} \text{ bit} \approx 1.4427 \text{ bitów}. \quad (2.51)$$

W systemie jednostek SI entropia jest traktowana jako wielkość bezwymiarowa [77]. Dla celów niniejszej pracy będziemy stosować logarytm naturalny do obliczeń entropii, a więc mierzona entropia będzie bezwymiarowa. Kluczowe cechy entropii Shannona to:

- $H(X) \geq 0$; entropia jest zawsze nieujemna,

- $H(X) = 0$ wtedy i tylko wtedy, gdy X jest zmienną deterministyczną (tj. $p_i = 1$ dla pewnego i),
- entropia osiąga maksimum, gdy rozkład prawdopodobieństwa jest jednorodny, tj. $p_i = \frac{1}{n}$ dla wszystkich i ,
- jest addytywna dla zmiennych niezależnych: $H(X, Y) = H(X) + H(Y)$, jeśli X i Y są niezależne.

Entropia jest wykorzystywana jako miara złożoności sygnału. Wyższa entropia wskazuje na większą nieprzewidywalność i złożoność sygnału, co jest istotne w analizie danych biologicznych, takich jak sygnały EEG czy EKG [78]. Może służyć do oceny stosunku sygnału do szumu w systemach komunikacyjnych. Poprzez analizę entropii sygnału odbieranego można ocenić, na ile szum wpływa na jakość przesyłanych danych [77].

Chociaż termin "entropia" jest używany zarówno w teorii informacji, jak i w termodynamice, reprezentuje on różne koncepcje w tych dziedzinach. Entropia termodynamiczna jest miarą nieuporządkowania układu fizycznego i jest związana z ilością energii niedostępnej do wykonania pracy. Natomiast entropia informacyjna mierzy niepewność związaną z losowym źródłem informacji.

Entropia termodynamiczna jest miarą liczby mikroskopowych konfiguracji odpowiadających stanowi makroskopowemu, odzwierciedlająca nieuporządkowanie lub losowość w układzie fizycznym. Kluczową jej cechą jest to że zgodnie z drugą zasadą termodynamiki entropia układu izolowanego powinna wzrastać. Entropia informacyjna jest miarą złożoności szeregu czasowego. Dla zobrazowania tych metryk weźmiemy jakiś stacjonarny szereg czasowy, na przykład wygenerowany przez funkcję piły zębatej *Sawtooth function*. Jeśli wzorzec tej funkcji jest ściśle regularny, entropia informacyjna pozostaje stała w czasie, co jest zgodne z charakterystyką stacjonarności. Entropia nie wzrasta: W przypadku prostej, deterministycznej i stacjonarnej fali zębatej nie istnieje żaden wewnętrzny mechanizm prowadzący do wzrostu entropii. Układ utrzymuje swój porządek i przewidywalność w sposób nieograniczony, co kontrastuje z układami termodynamicznymi, w których entropia ma tendencję do wzrostu.

2.4.2 Teoretyczny rozkład przetasowanego szeregu serii monotonicznych

Seria w szeregu RR jest jednoznacznie związana z funkcjonowaniem układów fizjologicznych odpowiedzialnych za jej powstawanie, takich jak układ sercowo-naczyniowy i układ nerwowy. W pracy [79] zastosowano

metodę tasowania szeregu RR , aby zweryfikować, czy HRA (asymetria rytmu serca) jest rzeczywistym zjawiskiem fizjologicznym, czy jedynie efektem użytego algorytmu. Innymi słowy, sprawdzono, czy wyniki nie są artefaktem metody. Analiza wykazała, że po przetasowaniu danych klasyczna asymetria zanika.

W niniejszej pracy metoda tasowania szeregu zostanie wykorzystana z dwóch powodów. Przetasowane szeregi służą do wyznaczenia wartości referencyjnych, które umożliwią ocenę, czy i w jaki sposób rozkład monotonicznych serii w rzeczywistym szeregu RR różni się od przetasowanego. Tego rodzaju analiza dostarczy wartościowych informacji z perspektywy fizjologicznej.

Dla przetasowanych szeregów, w których elementy nie powtarzają się, można również wyprowadzić teoretyczny wzór na oczekiwaną liczbę serii o określonej długości. W takim przypadku rozkłady serii zwalniających i przyspieszających okazują się identyczne. Podstawy teoretyczne tego zagadnienia zostały opisane przez Wolfowitza [80] a poniżej zostały dostosowane do problemu, którym zajmuję się w niniejszej pracy.

Jeśli znak (+ albo $-$) różnicy $(h_{i+1} - h_i)$ stanowi początkowy znak serii (zdefiniowanego jak wyżej), wówczas nazywamy h_i początkowym punktem zwrotnym (tzw. i.t.p.) tej serii. W związku z tym h_1 zawsze jest i.t.p., natomiast przyjmujemy konwencję, że h_n nigdy nie jest i.t.p. Wprowadzamy następujące zmienne losowe:

$$x_i = \begin{cases} 1, & \text{jeśli } h_i \text{ jest i.t.p.,} \\ 0, & \text{w przeciwnym razie,} \end{cases} \quad (2.1)$$

$$x_{pi} = \begin{cases} 1, & \text{jeśli } h_i \text{ jest i.t.p. serii o długości } p, \\ 0, & \text{w przeciwnym razie,} \end{cases} \quad (2.2)$$

$$w_{pi} = \begin{cases} 1, & \text{jeśli } h_i \text{ jest i.t.p. serii o długości } p \text{ lub większej,} \\ 0, & \text{w przeciwnym razie.} \end{cases} \quad (2.3)$$

dla $i = 1, 2, \dots, n$. Wprowadzamy również oznaczenia:

r = liczba serii w zbiorze R ,

r_p = liczba serii o długości p w zbiorze R ,

r'_p = liczba serii o długości p lub więcej w zbiorze R .

Oczywiście $r = \sum_{i=1}^n x_i$, $r_p = \sum_{i=1}^n x_{pi}$ oraz $r'_p = \sum_{i=1}^n w_{pi}$.

Jeśli X i Y są zmiennymi losowymi, wówczas $E(X)$ oznacza ich wartość oczekiwaną, $\sigma(XY)$ to kowariancja X i Y , a $\sigma^2(X)$ to wariancja X , a ile

istnieją. Przez dystrybuantę X rozumiemy funkcję $f(x)$ taką, że $P\{X < x\} \equiv f(x)$, gdzie $P\{\cdot\}$ jest prawdopodobieństwem zdarzenia opisanego w nawiasach.

Niech Y' będzie zmienną losową z dowolną ciągłą funkcją rozkładu $f(y)$. Niech $Y = (y_1, y_2, \dots, y_n)$ będzie ciągiem n niezależnych obserwacji zmiennej Y' . Ponieważ $P\{y_i = y_j\} = 0$ (dla $i \neq j; i, j = 1, 2, \dots, n$), rozkład serii rosnących i malejących w Y jest taki sam jak w S . Wybieramy funkcję rozkładu $f(y)$ w następujący sposób:

$$\begin{aligned} f(y) &= 0, & (y \leq 0), \\ f(y) &= y, & (0 \leq y \leq 1), \\ f(y) &= 1, & (y \geq 1). \end{aligned} \quad (2.52)$$

Wtedy:

$$P\{y_{i-1} < y_i > y_{i+1}\} = \int_0^1 \left[\int_{y_{i+1}}^1 \left(\int_0^{y_i} dy_{i-1} \right) dy_i \right] dy_{i+1} = \frac{1}{3}. \quad (2.53)$$

Ze względu na symetrię,

$$E(x_i) = P\{y_{i-1} < y_i > y_{i+1}\} + P\{y_{i-1} > y_i < y_{i+1}\} = \frac{2}{3}, \quad (i = 2, 3, \dots, n-1), \quad (2.54)$$

$$E(x_1) = 1, \quad E(x_n) = 0. \quad (2.55)$$

Będziemy potrzebować wzorów typu

$$U = \int_0^{y_{p+1}} \dots \int_0^{y_2} \frac{(y_1)^k}{k!} dy_1 \dots dy_p = \frac{(y_{p+1})^{k+p}}{(k+p)!}, \quad (2.56)$$

oraz

$$V = \int_{y_{p+1}}^1 \dots \int_{y_2}^1 \frac{(y_1)^k}{k!} dy_1 \dots dy_p. \quad (2.57)$$

Wynika to z faktu, że

$$\int_{y_{p+1}}^1 \cdots \int_{y_2}^1 dy_1 \cdots dy_p = \int_0^{v_{p+1}} \cdots \int_0^{r_2} dv_1 \cdots dv_p = \frac{(v_{p+1})^p}{p!}, \quad (2.58)$$

gdzie $v_j = (1 - y_j)$ dla $j = 1, \dots, p+1$. Łatwo wtedy wykazać, że

$$V = \sum_{s=1}^p (-1)^{s+1} \frac{(v_{p+1})^{p-s}}{(k+s)!(p-s)!} + (-1)^p \frac{(y_{p+1})^{k+p}}{(k+p)!} = \sum_{r=0}^k (-1)^{k+r} \frac{(v_{p+1})^{p+k-r}}{(p+k-r)! r!}, \quad (3.3)$$

a także będziemy potrzebować wartości $\int_0^1 V dy_{p+1}$ oraz $\int_0^1 \int_0^{y_{p+2}} V dy_{p+1} dy_{p+2}$.

Zauważmy, że

$$\int_0^1 V dy_{p+1} = \sum_{r=0}^k (-1)^{k+r} \frac{1}{(p+k-r+1)! r!}. \quad (2.60)$$

Posługując się tożsamością:

$$\sum_{r=0}^t (-1)^r \frac{1}{(n-r)! r!} = (-1)^t \frac{1}{n(n-t-1)! t!} \quad (t < n), \quad (2.61)$$

otrzymujemy:

$$\int_0^1 V dy_{p+1} = \frac{1}{(p+k+1)p! k!}. \quad (2.62)$$

Podobnie:

$$\begin{aligned} \int_0^1 \int_0^{y_{p+2}} V dy_{p+1} dy_{p+2} &= \frac{1}{(p+k+1)p! k!} \\ &- \frac{1}{(p+k+2)(p+1)! k!}. \end{aligned} \quad (2.63)$$

4. Kowariancje biegów rosnących i malejących.

Najpierw obliczamy $E(r_p)$ oraz $E(r'_p)$. Wprowadzamy symbol

$$P\{-, +^p, -\} = P\{y_{i-1} > y_i < y_{i+1} < \cdots < y_{i+p} > y_{i+p+1}\}. \quad (2.64)$$

Wartość po prawej stronie jest niezależna od i (gdy jest ono zdefiniowane, tj. $i - 1 \geq 1$, $i + p + 1 \leq n$). Zatem

$$E(x_{pi}) = P\{-, +^p, -\} + P\{+, -^p, +\} = 2 P\{-, +^p, -\} \quad (2.65)$$

$$= 2 \int_0^1 \int_{y_{i+p+1}}^1 \int_0^{y_{i+p}} \cdots \int_0^{y_{i+1}} \int_{y_i}^1 dy_{i-1} \cdots dy_{i+p+1} = 2 \frac{p^2+3p+1}{(p+3)!} \quad (2.66)$$

$$(i = 2, 3, \dots, n - p - 1). \quad (2.67)$$

Dla symetrii $E(x_{p1}) = E(x_{p,n-p})$, a obie te wartości wynoszą $2 \frac{p+1}{(p+2)!}$.

Ponadto $E(x_{pi}) = 0$ przy $i > n - p$.

Stąd

$$\begin{aligned} E(r_p) &= E\left(\sum_{i=1}^n x_{pi}\right) = 2 E(x_{p1}) + (n - p - 2) E(x_{pi}) \\ &= 2n \frac{p^2 + 3p + 1}{(p+3)!} - 2 \frac{p^3 + 3p^2 - p - 4}{(p+3)!}, \quad (p \leq n - 2). \end{aligned} \quad (2.68)$$

Finalnie, Wolfowitz [80] sformułował wzór pozwalający oszacować przewidywaną liczbę wystąpień poszczególnych monotonicznych sekwencji w danym wektorze. Model tasowania, który tu przyjmujemy, zakłada, po pierwsze, że obserwacje są niezależne, a po drugie, że pochodzą z ciągłego procesu losowego. W świetle tej drugiej przesłanki prawdopodobieństwo uzyskania dwóch identycznych wartości zmiennej jest równe zero [80, 81]. W przypadku badania szeregów odstępów RR, rozdzielczość urządzenia pomiarowego sprawia, że prawdopodobieństwo uzyskania dwóch lub więcej identycznych wartości jest niezerowe - stąd wprowadzamy serie neutralne, w których zliczamy występujące po sobie identyczne wartości. Warto jednak zwrócić uwagę, że jest to artefakt urządzenia pomiarowego. W wielu pracach serie neutralne są usuwane poprzez dodawanie szumu o bardzo niskiej wariancji [7]. Używając rozkładu prawdopodobieństwa wyprowadzonego powyżej, można podać referencyjne wartości zarówno dla serii monotonicznych AR i DR, jak i entropii danych wzorami 29-41 [75].

2.5 Parametry entropijne

2.5.1 Entropia przybliżona

Entropia przybliżona (*Approximate Entropy*, *ApEn*) została zaproponowana przez Stephena Pincusa w serii prac naukowych z lat 90. XX wieku

[82, 83, 84]. Celem wprowadzenia tej miary było umożliwienie klasyfikacji zarejestrowanych eksperymentalnie sygnałów, pochodzących z układów o potencjalnie chaotycznym charakterze, umożliwiające klasyfikacji entropowych.

Podstawowa definicja $ApEn$ zakłada, że dla ustalonego zbioru parametrów (m, r) oraz przy założeniu bardzo dużej liczby próbek N , można wprowadzić funkcję:

$$ApEn(m, r) = \lim_{N \rightarrow \infty} [\Phi^m(r) - \Phi^{m+1}(r)], \quad (2.69)$$

gdzie $\Phi^m(r)$ jest pomocniczą wielkością opisującą tzw. *strukturę podobieństw* w przestrzeni m -wymiarowych wektorów (tzw. wzorców), zdefiniowaną jako:

$$\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log B_i^m(r), \quad (2.70)$$

gdzie:

$$B_m^i(r) = \begin{cases} \frac{1}{L_m} \sum_{j=1}^{L_m} \Theta(r - \|\vec{v}_m(i) - \vec{v}_m(j)\|) & \text{jeśli } i \leq L_m, \\ 0 & \text{jeśli } i > L_m \end{cases}, \quad (2.71)$$

Θ jest funkcją Heaviside'a przyjmującą wartość 1 dla argumentu ≥ 0 i 0 w przeciwnym przypadku. Wektory v_m określone są na podstawie szeregu czasowego składającego się z N pomiarów:

$$U_i = \{u(1), u(2), \dots, u(N)\}, \quad (2.72)$$

dla którego definiuje się tzw. *przestrzeń zanurzeniową* (ang. *embedding space*), w której rozpatruje się wektory m -wymiarowe postaci (wektory zanurzeniowe):

$$\vec{v}_{m,\tau}(i) = [u(i), u(i + \tau), u(i + 2\tau), \dots, u(i + (m - 1)\tau)]. \quad (2.73)$$

Liczba takich wektorów (tzw. wzorców) wynosi $L_m = N - (m - 1)\tau$. Parametr τ nazywa się opóźnieniem czasowym i odpowiada liczbie próbek, o którą przesuwane są kolejne składowe wektora. Wielkość r pełni rolę *promienia porównania* (ang. *tolerance radius*), a odległość między wektorami $\vec{v}_{m,\tau}(i)$ i $\vec{v}_{m,\tau}(j)$ jest definiowana jako

$$\|\vec{v}_{m,\tau}(i) - \vec{v}_{m,\tau}(j)\|_\infty = \max_{1 \leq k \leq m} |u(i + (k - 1)\tau) - u(j + (k - 1)\tau)|. \quad (2.74)$$

Ostatecznie, używając powyższych definicji, dla zbioru danych o ograniczonej liczbie punktów N wprowadza się następnie statystykę:

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r). \quad (2.75)$$

Im większa wartość ApEn, tym większa „nieuporządkowana dynamika” lub „nieprzewidywalność” badanego szeregu czasowego. *ApEn* zawsze oblicza się dla konkretnych wartości parametrów m oraz r . Miarę tę można stosować do zestawu danych o praktycznie dowolnej wielkości - literatura sugeruje $N > 300$ [82].

2.5.2 Entropia z próby

Entropia z próby (*Sample Entropy*, *SampEn*) jest miarą złożoności sygnału, która została wprowadzona jako poprawiona wersja entropii przybliżonej [85]. Także tutaj definicja opiera się na analizie prawdopodobieństwa, że dwa fragmenty sygnału o długości m są podobne w zakresie tolerancji r , a ich przedłużenia do długości $m + 1$ pozostają podobne - podstawowa różnica polega na usunięciu sumowania samopodobieństw wektorów zanurzeniowych oraz położeniu funkcji logarytmującej (w przypadku *SampEn* jest ona zewnętrzna w stosunku do sumy podobieństw). W przestrzeni zanurzeniowej wektorów v_m można wprowadzić tzw. *sumę korelacyjną*:

$$C^m(r) = \frac{1}{L_m} \sum_{i=1}^{L_m} C_i^m(r), \quad (2.76)$$

gdzie

$$C_i^m(r) = \frac{1}{L_m - 1} \sum_{\substack{j=1 \\ j \neq i}}^{L_m} \Theta\left(r - \|\vec{v}_{m,\tau}(i) - \vec{v}_{m,\tau}(j)\|_\infty\right). \quad (2.77)$$

Podsumowując powyższe Richman i Moorman proponują praktyczne rozważanie stosunku sum korelacyjnych i definiują entropię z próby jako pewną formę analogi do entropii przybliżonej oryginalnie wprowadzonej Pincusa [86]:

$$SampEn(m, r, N) = -\ln \left[\frac{C^{m+1}(r)}{C^m(r)} \right]. \quad (2.78)$$

Entropia z próby podobnie jak entropia przybliżona jest zatem wariantem estymatora wywiedzionego z tzw. entropii Eckmanna i Ruelle (nie jest jednak jej estymatorem) [82], zdefiniowanej jako:

$$ER = \lim_{r \rightarrow 0} \lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} [\Phi_m(r) - \Phi_{m+1}(r)], \quad (2.79)$$

dostosowanym do analizy danych eksperymentalnych, gdzie jak już wspomniano kluczowa idea polega na zbadaniu, z jakim prawdopodobieństwem

wektory, które były podobne w m -wymiarze, pozostają podobne po przejściu do wymiaru $m + 1$.

Dla bardzo długich szeregów czasowych (teoretycznie $N \rightarrow \infty$) definicja ta zbliża się do formalnej entropii Ruelle, jednak w przypadku skończonych danych stanowi praktyczny, obliczalny wskaźnik. W SampEn wyłączone są tzw. *samodopasowania*, czyli porównania wektora z nim samym, co pozwala uniknąć pewnych niekorzystnych efektów obecnych w ApEn.

Interesującym z punktu widzenia niniejszej pracy doktorskiej zastosowaniem SampEn w naukach biomedycznych jest analiza sygnałów *RR* (odstępów między kolejnymi załamkami *R* w sygnale EKG), która to analiza została włączona do obowiązujących standardów opisu HRV [87, 88]. Inne interesujące zastosowania *SampEn* w kontekście analizy szeregów odstępów *RR* to wykorzystanie tego parametru w testach pochyleniowych czy badaniach nad omdleniami wazowagalnymi [89, 90, 91]. Zaletą SampEn jest relatywnie prosta interpretacja związana z miarą nieuporządkowania w sygnale — im wyższa wartość entropii z próby, tym większa „nieprzewidywalność” dynamiki sygnału. *SampEn* powstało jako modyfikacja *ApEn* a samo z kolei jest podstawą do takich miar jak Fuzzy Entropy (FuzzyEn), czy Permutation Entropy (PE) [92]

W tej pracy do obliczenia entropii z próby wykorzystany zostanie algorytm NCM (*Norm Component Matrix*) przedstawiony w pracy Żurek i in. [86]. Algorytm ten pozwala na szybkie i efektywne obliczenie sum korelacyjnych dla danego sygnału, na podstawie których jesteśmy w stanie obliczyć entropię przybliżoną i entropię z próby. Jest to szczególnie przydatne przy analizie dużych zbiorów danych. Na potrzeby analizy wielogodzinnych polisomnograficznych szeregów czasowych opisywanych w tej pracy oraz celem przygotowania szeroko dostępnego i łatwego w instalacji oprogramowania NCM w ramach prac badawczych realizowanych w niniejszym doktoracie przeprowadzono reimplementację wspomnianego algorytmu wraz z jego szczegółową dokumentacją oraz analizą wyników. Jest to jeden z oryginalnych wkładów badawczych tej rozprawy doktorskiej, szczegółowo opisany w oczekującym na recenzję artykule [93].

2.6 Metody uczenia maszynowego

W tym podrozdziale omówiono zastosowanie metod uczenia maszynowego w budowie i ewaluacji modeli predykcyjnych, opartych na metodach przedstawionych w poprzednich podrozdziałach, takich jak serie monotoniczne szeregów *RR*. Opisane wcześniej miary i wzorce stanowią kluczowy wkład do konstrukcji cech wejściowych dla modeli. Do treningu i oceny

tych modeli wykorzystano algorytmy uczenia maszynowego, w tym regresję logistyczną oraz uogólnione równania estymacyjne (GEE) a także XGBoost. Przedstawione podejście umożliwia analizę złożonych relacji między danymi i predykcję określonych stanów.

2.6.1 Regresja logistyczna

Regresja logistyczna jest jedną z najważniejszych metod statystycznych używanych do analizy danych binarnych, gdzie zmienna zależna przyjmuje tylko dwie wartości, często kodowane jako 0 i 1 [94]. Jest szeroko stosowana w różnych dziedzinach, w tym w uczeniu maszynowym i medycynie, do modelowania prawdopodobieństwa wystąpienia zdarzenia w zależności od zestawu zmiennych niezależnych.

Początki regresji logistycznej sięgają XIX wieku, kiedy Pierre François Verhulst wprowadził funkcję logistyczną do modelowania wzrostu populacji [95]. Jednak dopiero w XX wieku metoda ta została zastosowana w analizie danych binarnych przez naukowców takich jak Joseph Berkson [96], który wprowadził termin "logit" do opisu logarytmu ilorazu szans.

Regresja logistyczna modeluje logarytm ilorazu szans (*logit*) jako liniową kombinację zmiennych niezależnych:

$$\log \left(\frac{P(Y = 1|X)}{1 - P(Y = 1|X)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2.80)$$

Gdzie:

- $P(Y = 1|X)$ jest prawdopodobieństwem, że zmienna zależna Y przyjmuje wartość 1 przy danych wartościach zmiennych niezależnych $X = (X_1, X_2, \dots, X_n)$.
- β_0 jest wyrazem wolnym (intercept).
- $\beta_1, \beta_2, \dots, \beta_n$ są współczynnikami regresji odpowiadającymi zmiennym niezależnym.

Funkcja logistyczna przekształca liniową kombinację zmiennych w przedział $(0,1)$:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}} \quad (2.81)$$

Iloraz szans (*odds*) jest definiowany jako stosunek prawdopodobieństwa wystąpienia zdarzenia do prawdopodobieństwa jego nie wystąpienia:

$$\text{odds} = \frac{P(Y = 1|X)}{1 - P(Y = 1|X)} \quad (2.82)$$

Logarytm ilorazu szans (*logit*) liniowo zależy od zmiennych niezależnych, co umożliwia interpretację współczynników β_i jako wpływu jednostkowej zmiany zmiennej X_i na logarytm ilorazu szans. Parametry β w modelu regresji logistycznej są estymowane metodą maksymalnej wiarygodności (*maximum likelihood estimation, MLE*) [97]. Funkcja wiarygodności jest konstruowana na podstawie obserwowanych danych, a następnie maksymalizowana w celu znalezienia najlepszych wartości parametrów. W kontekście uczenia maszynowego, regresja logistyczna jest stosowana jako algorytm klasyfikacji binarnej [98]. Jej zalety to prostota implementacji, efektywność obliczeniowa oraz interpretowalność wyników.

Celem jest przypisanie obserwacji do jednej z dwóch klas na podstawie wartości zmiennych niezależnych. Decyzja jest podejmowana na podstawie progowego prawdopodobieństwa, zwykle 0.5:

$$\hat{Y} = \begin{cases} 1 & \text{jeśli } P(Y = 1|X) \geq 0.5 \\ 0 & \text{w przeciwnym razie} \end{cases} \quad (2.83)$$

Aby zapobiec przeuczeniu (*overfitting*), stosuje się techniki regularizacji, takie jak L1 (Lasso) i L2 (Ridge) [99]. Regularizacja dodaje do funkcji kosztu karę za wielkość współczynników, co prowadzi do prostszych i bardziej ogólnych modeli. Regularizacja (zwłaszcza L1) pośrednio prowadzi do zmniejszenia liczby niezerowych współczynników.

W medycynie regresja logistyczna jest narzędziem do analizy danych klinicznych i epidemiologicznych [100, 101].

2.6.2 Uogólnione Równania Estymujące (GEE)

Uogólnione Równania Estymujące (GEE) są metodą statystyczną używaną do analizy danych zależnych lub powtarzanych pomiarów, szczególnie gdy dane nie spełniają założeń tradycyjnych modeli regresji [102]. GEE rozszerza modele uogólnionej regresji liniowej, umożliwiając modelowanie średniej odpowiedzi w funkcji czasu lub innych czynników, uwzględniając korelacje między obserwacjami.

Metoda GEE została wprowadzona przez Lianga i Zegera w 1986 roku jako narzędzie do analizy danych podłużnych [102]. GEE są uogólnieniem modeli liniowych, które pozwalają na uwzględnienie struktury korelacji

między powtarzanymi obserwacjami tego samego obiektu, co jest istotne w badaniach medycznych i biologicznych.

W modelu GEE zakłada się, że obserwacje Y_i są powiązane ze zmiennymi niezależnymi X_i poprzez funkcję łączącą $g(\cdot)$:

$$g(\mu_i) = X_i\beta \quad (2.84)$$

gdzie:

- $\mu_i = E(Y_i)$ jest wartością oczekiwaną zmiennej zależnej Y_i ,
- X_i to macierz zmiennych niezależnych dla i -tego obiektu,
- β to wektor parametrów modelu.

Estymacja parametrów β odbywa się poprzez rozwiązanie uogólnionych równań estymujących:

$$\sum_{i=1}^N D_i^T V_i^{-1} (Y_i - \mu_i) = 0 \quad (2.85)$$

gdzie:

- $D_i = \frac{\partial \mu_i}{\partial \beta}$ jest macierzą pochodnych,
- V_i jest macierzą wariancji i kowariancji dla Y_i ,
- N to liczba jednostek (np. pacjentów) w badaniu.

Macierz V_i jest modelowana za pomocą tzw. *roboczej macierzy korelacji* $R(\alpha)$, która opisuje strukturę korelacji między powtarzanymi obserwacjami [103]. W praktyce często stosuje się proste struktury korelacji, takie jak:

- **Niezależność:** brak korelacji między obserwacjami.
- **Korelacja wymienna (tożsamościowa):** wszystkie pary obserwacji mają tę samą korelację.
- **Korelacja autoregresyjna:** korelacja maleje wraz ze wzrostem odległości między obserwacjami w czasie.

W uczeniu maszynowym GEE są stosowane do analizy danych sekwencyjnych i czasowych, gdzie istnieje zależność między kolejnymi obserwacjami [104]. W przeciwieństwie do tradycyjnych metod, takich jak modele mieszane, GEE skupia się na modelowaniu średniej populacyjnej, co jest korzystne w kontekście dużych zbiorów danych. GEE są wykorzystywane do tworzenia

modeli predykcyjnych w sytuacjach, gdy dane są skorelowane lub występują powtarzane pomiary [105]. W medycynie GEE są szeroko stosowane do analizy danych podłużnych i powtarzanych pomiarów [106].

W kontekście GEE, **klaster** to grupa obserwacji, które są ze sobą powiązane, zwykle poprzez wspólne cechy lub jednostki badane, takie jak powtarzane pomiary tego samego pacjenta, szkoły w badaniach edukacyjnych czy rodziny w badaniach genetycznych [104]. Kluczowym aspektem GEE jest modelowanie korelacji między obserwacjami wewnątrz klastrów, co pozwala na uzyskanie poprawnych estymatorów parametrów modelu i ich wariancji.

GEE wprowadza tzw. *roboczą strukturę korelacji* $R(\alpha)$, która opisuje korelacje między obserwacjami w ramach tego samego klastra [102].

W tradycyjnym uczeniu maszynowym, termin *klaster* odnosi się zazwyczaj do grup obserwacji zgrupowanych na podstawie podobieństwa cech, identyfikowanych za pomocą algorytmów klasteryzacji, takich jak *k-means* czy *hierarchiczna klasteryzacja* [107]. Jest to podejście **uczenia bez nadzoru**, gdzie celem jest odkrycie naturalnych grup w danych bez wcześniejszej wiedzy o strukturze.

Natomiast w GEE klastry są **zdefiniowane a priori** na podstawie struktury danych lub projektu badania. Przykładem mogą być pomiary różnych parametrów u tych samych pacjentów w różnych punktach czasowych. W tym przypadku klaster to zestaw obserwacji dla pojedynczego pacjenta [105].

W GEE kluczowe jest modelowanie korelacji wewnątrz klastrów w celu uzyskania poprawnych estymatorów parametrów modelu. Korelacja między obserwacjami w klastrze jest traktowana jako znana lub szacowana na podstawie danych, ale struktura klastrów jest ustalona z góry [103].

W tradycyjnym uczeniu maszynowym, algorytmy klasteryzacji mają na celu odkrycie nieznanej struktury klastrów w danych, gdzie korelacje czy podobieństwa między obserwacjami są wykorzystywane do grupowania danych [108]. Nie ma tu założonej struktury korelacji ani zdefiniowanych wcześniej klastrów.

2.6.3 XGBoost

XGBoost (Extreme Gradient Boosting) to zaawansowana implementacja metody gradient boosting, szeroko stosowana w zadaniach klasyfikacji i regresji [109]. Model ten iteracyjnie dopasowuje drzewa decyzyjne, minimalizując funkcję straty poprzez optymalizację gradientową. Dzięki wysokiej wydajności oraz zdolności do modelowania złożonych, nieliniowych relacji, XGBoost jest jednym z najczęściej używanych algorytmów w uczeniu maszynowym [109].

Matematycznie, funkcja przewidywania w XGBoost wyrażana jest jako suma K drzew decyzyjnych:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F}, \quad (2.86)$$

gdzie f_k to pojedyncze drzewo decyzyjne z przestrzeni $\mathcal{F}$, a x_i to wektor cech dla i -tej obserwacji. Funkcja celu składa się z funkcji straty l i regularyzacji Ω :

$$\mathcal{L}(\Theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k), \quad (2.87)$$

gdzie $\Omega(f_k)$ kontroluje złożoność drzewa:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2, \quad (2.88)$$

z γ i λ jako hiperparametrami regularyzacji, a w reprezentującym wektor wag liści drzewa. Gradient boosting iteracyjnie aktualizuje model, dodając drzewa dopasowane do gradientów i hesjanów funkcji straty:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t), \quad (2.89)$$

gdzie $g_i = \partial_{\hat{y}_i} l(y_i, \hat{y}_i)$ oraz $h_i = \partial_{\hat{y}_i}^2 l(y_i, \hat{y}_i)$.

W porównaniu z regresją logistyczną, która przewiduje prawdopodobieństwo przynależności do klasy na podstawie kombinacji liniowej zmiennych:

$$P(y = 1|x) = \frac{1}{1 + \exp(-w^T x)}, \quad (2.90)$$

XGBoost nie zakłada liniowej relacji między zmiennymi wejściowymi a odpowiedzią, dzięki czemu lepiej radzi sobie z danymi o charakterze nieliniowym i wielowymiarowym [109].

2.7 Studium literaturowe związków HRV, złożonością i HRA z badaniami nad snem

Odstępy RR odzwierciedlają zmienność rytmu serca, która jest kształtowana przez aktywność autonomicznego układu nerwowego (AUN) i ściśle wiąże się z architekturą snu [110, 111]. Monitorowanie tych odstępów może dostarczyć cennych informacji na temat przejść między różnymi fazami snu, w tym zarówno fazą szybkich ruchów gałek ocznych, jak i fazami NREM [111, 112].

2.7.1 Związek między snem a HRV

Zależność między stadami snu a wartością *SDNN* (odchylenie standardowe odstępów RR) była przedmiotem wielu badań, wskazując na powiązanie między zmiennością rytmu serca, aktywnością układu autonomicznego i różnymi fazami snu. *SDNN*, jako ogólna miara HRV, wyraża odchylenie standardowe wszystkich odstępów RR, a jego wartość można mierzyć zarówno w trakcie całego zapisu, jak i w poszczególnych fazach snu. Liczne prace sugerują, że *SDNN* może pokazywać ogólną aktywność AUN podczas snu oraz stanowić potencjalny wskaźnik jakości snu. Na przykład zaobserwowano, że wartości *SDNN* zmieniają się w zależności od fazy snu — spadają w głębokim śnie (N3) i rosną w fazie REM, co może świadczyć o zmianach w równowadze autonomicznej [111].

Wykazano również, że w fazie REM aktywność autonomiczna jest bardziej zmienna, co uwidacznia się wyższymi wartościami *SDNN* odzwierciedlającymi zwiększoną zmienność rytmu serca w tej fazie [113].

W innym badaniu przeanalizowano dokładność klasyfikacji stadiów snu na podstawie parametrów HRV, wskazując, że *SDNN* i inne miary oparte na HRV mogą zapewniać stosunkowo wysoką skuteczność w różnicowaniu poszczególnych faz [114].

Klasyfikacja faz snu z wykorzystaniem wskaźników HRV, w tym wartości *SDNN*, była również podejmowana w pracach nad modelami uczenia maszynowego. Przykładowo, modele wykorzystujące parametry HRV pochodzące z sygnału EKG wykazały, że *SDNN* należy do kluczowych wskaźników umożliwiających rozróżnienie stanu czuwania, snu REM i snu NREM. Rezultaty te wskazują na możliwość włączenia *SDNN* do automatycznych systemów klasyfikacji snu, co może sprzyjać rozwojowi nieinwazyjnych i wygodnych metod monitorowania snu [115].

Przyjmuje się często, że HRV odzwierciedla wzajemne oddziaływanie między gałęzią współczulną i przywspółczulną AUN. W różnych stadiach snu można zaobserwować charakterystyczne wzorce HRV związane z aktywnością obu tych gałęzi. Często zakłada się (choć jest to pewne uproszczenie), że składowa niskoczęstotliwościowa (LF) w widmie HRV odnosi się do aktywności przywspółczulnej, natomiast składowa wysokoczęstotliwościowa (HF) do aktywności współczulnej [110]. Jeśli przyjmiemy takie założenia, w fazie NREM wyższy udział składowej HF i niższy udział LF wskazują na przewagę aktywności przywspółczulnej, natomiast w fazie REM obserwuje się wzrost stosunku LF/HF [111].

Wyższa aktywność przywspółczulna i niższa aktywność współczulna, widoczne w specyficznych wzorcach HRV, są zazwyczaj łączone z lepszą jakością snu, zwłaszcza w głębokim śnie (N3). Wskazuje to na istotną rolę

HRV w ocenie jakości snu i w zrozumieniu fizjologicznych mechanizmów związanych ze snem [112].

Ichimaru i wsp. zbadali wpływ różnych stadiów snu na relację między oddychaniem a HRV, wykazując, że fazy snu w odmienny sposób modulują zmienność rytmu serca. Odkryli oni, że aktywność przywspółczulna ulega obniżeniu w fazie REM w porównaniu z fazą NREM [116].

W pracy [117] przeanalizowano powtarzalność parametrów HRV w różnych stadiach snu i stwierdzono, że w fazie snu wolnofalowego (SWS, ang. *slow-wave sleep*) częstość rytmu serca (HR) oraz moc w zakresie HF pozostają względnie stabilne, natomiast zmienne takie jak moc LF czy stosunek LF/HF podlegają istotnym zmianom między poszczególnymi fazami snu.

Badania Tsunody i wsp. pokazały wpływ natężenia światła na HRV, wykazując, że stosunek LF/HF rośnie przy intensywnym oświetleniu lub w skrajnej ciemności w porównaniu z oświetleniem przytłumionym. Autorzy zaobserwowali również, że moc LF jest wyższa w fazie REM, podczas gdy moc HF wzrasta w różnych stadiach snu, prowadząc do spadku stosunku LF/HF w fazie wolnofalowej [118].

Badania wskazują również na znaczący wpływ wieku na HRV podczas snu. Starsze osoby mają zazwyczaj niższe składowe HF w fazach NREM w porównaniu z osobami młodszymi, co interpretowane jest jako oznaka słabszej aktywności przywspółczulnej w procesie starzenia. Dlatego też przy ocenie snu i jakości snu za pomocą HRV należy uwzględniać zmiany związane z wiekiem [111].

HRV jest też istotnym narzędziem w badaniach nad różnymi zaburzeniami snu. Na przykład u pacjentów z obturacyjnym bezdechem sennym (OSA) widoczne są specyficzne wzorce HRV, takie jak zwiększona zmienność rytmu serca podczas epizodów bezdechu, wynikająca z intensyfikacji aktywności współczulnej [119].

Ponadto liczne prace ukazują, że na HRV podczas snu wpływ mogą mieć inne czynniki, np. zaburzenia oddychania, ruchy kończyn czy bezsenność. Obecność tych zaburzeń może istotnie modyfikować normalne wzorce regulacji autonomicznej w czasie snu [111].

W ostatnich latach coraz częściej łączy się analizę HRV z innymi sygnałami fizjologicznymi, takimi jak EKG, EEG czy pomiary wysiłku oddechowego, co zwiększa precyzję klasyfikacji stadiów snu. Rozwój algorytmów uczenia maszynowego i zaawansowanych metod przetwarzania sygnałów umożliwia lepsze wykrywanie poszczególnych faz i przejść między nimi [120].

Mimo że wykorzystanie samej analizy HRV do klasyfikacji stadiów snu jest trudniejsze (ze względu na mniejszą ilość danych niż w polisomnografii), ta metoda cieszy się dużym zainteresowaniem, ponieważ jest znacznie bardziej dostępna i tańsza w realizacji. Może być potencjalnie implementowana

w niewielkich, tanich i przenośnych urządzeniach lub nawet w telefonach komórkowych.

Analiza danych HRV pozwoliła na rozwój modeli różnicujących stan czuwania, fazę REM i poszczególne stadia snu NREM. Przykładowo, podejścia wykorzystujące cechy wyodrębnione z EKG i parametrów oddechowych osiągały satysfakcjonującą dokładność w klasyfikacji stadiów snu, podkreślając rosnące możliwości nieinwazyjnego monitorowania snu [120].

Dalszy rozwój technologii ubieralnych pozwala na ciągłe monitorowanie HRV w trakcie całej nocy, co otwiera nowe możliwości prowadzenia długoterminowych badań nad snem i zarządzania zdrowiem w sposób spersonalizowany. Urządzenia takie jak pierścionek Oura wykazują wysoką dokładność pomiaru HRV w porównaniu z urządzeniami klinicznymi klasy medycznej (ECG), co potwierdza ich użyteczność w codziennym monitoringu snu i stylu życia [42].

2.7.2 Związek między snem a złożonością

Kolejny ważny obszar badań nad snem i jego fazami w kontekście sygnału z odstępów RR dotyczy *złożoności* (ang. *complexity*). Chociaż analiza złożoności często jest postrzegana jako część oceny HRV, można też traktować ją jako odrębny aspekt badawczy, bazujący na tych samych szeregach czasowych odstępów RR. Najczęściej stosowaną miarą złożoności sygnału EKG (oraz sygnału EEG) jest *SampEn*. Wiele prac poświęcono związkowi między fazami snu a wartościami

SampEn obliczanymi na podstawie sygnałów EKG i EEG. Najnowsze badania pokazały, że zastosowanie metody „fizycznego progu” (ang. *physical threshold-based sample entropy*) pozwala skutecznie uchwycić przejścia fizjologiczne między różnymi fazami snu, w szczególności u pacjentów z bezdechem sennym. W analizie segmentów odstępów RR u osób zmagających się z bezdechem sennym zaobserwowano, że przejścia z fazy snu do czuwania charakteryzują się znacznie większym spadkiem *SampEn* niż przejścia w obrębie samego snu. Ten spadek *SampEn* był bardziej wyraźny niż zmiany w standardowych miarach HRV, wskazując na wyższą wiarygodność *SampEn* przy monitorowaniu zmian fizjologicznych związanych ze snem [121, 122].

Zarówno HRV, jak i *SampEn* korelują z różnorodnymi zjawiskami fizjologicznymi i patologicznymi, odzwierciedlając aktywność współczulną i mechanizmy kontrolne układu oddechowego. Badania wykazały zmniejszoną aktywność przywspółczulną i obniżone wartości *SampEn* u osób z OSA w porównaniu z grupą kontrolną, co wskazuje na zachwianą równowagę pomiędzy aktywnością współczulną i przywspółczulną [123].

Zastosowanie *SampEn* do analizy odstępów RR u pacjentów z obturacyjnym bezdechem sennym (OSAS) pozwoliło na określenie wyraźnych zależności między złożonością sygnału a aktywnością układu autonomicznego, rzucając nowe światło na ciężkość choroby i jej wpływ na organizm [124]. W przywołanej pracy stwierdzono, że wzorzec *RR* u pacjentów z OSAS jest bardziej regularny (mniej złożony) niż u osób zdrowych, co wiąże się z cyklicznymi zmianami częstości akcji serca spowodowanymi epizodami bezdechu i następującymi po nich wybudzeniami, odzwierciedlającymi zaburzenia w funkcji autonomicznej [124].

Ponadto, inne badania w dziedzinie neurofizjologii sugerują, że *SampEn* może służyć jako kryterium rozróżniania stadiów snu, z niższymi wartościami entropii w głębszych fazach snu [125].

Kolejne prace pokazują, że uzupełnienie klasycznych miar HRV o parametry entropijne poprawia dokładność klasyfikacji stadiów snu. Na przykład w [121, 122] podkreślono, że analiza indeksów nieliniowych (w tym entropii) znacząco wzmacnia precyzję klasyfikacji snu dzięki uchwyceniu dodatkowych wzorców czasowych. Również wykorzystanie sztucznych sieci neuronowych do rozróżnienia czuwania, NREM i REM z uwzględnieniem miar entropii prowadzi do poprawy skuteczności klasyfikacji [126].

Zebrane wyniki wskazują, że obliczana z sygnału EKG wartość *SampEn* stanowi solidny miernik do odróżniania stadiów snu. Konsekwentne zmniejszanie się *SampEn* w głębszych stadiach snu, obserwowane w różnych badaniach, sugeruje, że jest to wiarygodny wskaźnik głębokości snu oraz jego dynamiki.

2.7.3 Związek między snem a HRA

Relatywnie nowym podejściem do badania snu i jego faz jest asymetria rytmu serca. Uwzględnia ona zarówno miary oparte na wariancji, jak i na analizie serii monotonicznych. Asymetria oparta na wariancji odzwierciedla różnicę w wkładzie zwolnień i przyspieszeń rytmu serca do zmienności odstępów RR — zwolnienia w większym stopniu wpływają na zmienność krótkoterminową, natomiast przyspieszenia mają większy wpływ na zmienność długoterminową i całkowitą [7]. Natomiast analiza sekwencji monotonicznych koncentruje się na strukturze przyspieszeń i zwolnień w kolejnych odstępach RR, wykazując, że przyspieszenia częściej tworzą dłuższe i liczniejsze ciągi (runs) niż zwolnienia [8].

Związek między HRA a poszczególnymi fazami snu, zwłaszcza w kontekście tzw. serii, był przedmiotem kilku badań.

W pracy Guzika i wsp. [127] przeanalizowano mikrostrukturę HRA u pacjentów z obturacyjnym bezdechem sennym podczas snu. Stwierdzono,

że przyspieszenia i zwolnienia rytmu serca w różny sposób wpływają na zmienność rytmu serca (HRV), a sama asymetria jest zależna od nasilenia OSA. Dane EKG zostały poddane analizie w celu ilościowego opisu mikrostruktury HRA poprzez zliczanie monotonicznych sekwencji zwolnień oraz przyspieszeń rytmu serca.

Wyniki pokazały, że u pacjentów z ciężką postacią OSA liczba sekwencji zwolnień i przyspieszeń o długości 1 była istotnie mniejsza w porównaniu z pacjentami z umiarkowanym lub łagodnym brakiem OSA. Z kolei w grupie z ciężką OSA zaobserwowano znacznie większą liczbę dłuższych sekwencji (przyspieszeń w zakresie 5–10 oraz zwolnień w zakresie 5–8). Ponadto najdłuższe sekwencje przyspieszeń były wyraźnie dłuższe u pacjentów z ciężką OSA niż u osób bez tej choroby lub z jej łagodną postacią.

Autorzy konkludują, że mikrostruktura HRA jest powiązana z ciężkością OSA, a w przypadku głębokiej postaci choroby częściej występują dłuższe sekwencje zwolnień i przyspieszeń. Analiza ta wskazuje, że badanie mikrostruktury HRA może stanowić cenne źródło informacji o zmianach fizjologicznych w rytmie serca związanych z nasiloną formą OSA.

Z kolei Jiang i wsp. [128] przeanalizowali wzorce przyspieszeń (ARs, ang. *acceleration runs*) i zwolnień (DRs, ang. *deceleration runs*) u pacjentów z OSA, wykorzystując metodę sekwencji monotonicznych, która polega na zliczaniu kolejnych odstępów RR zmieniających się w tym samym kierunku. Wyniki wskazują, że OSA znacząco wpływa na te wzorce, sugerując asymetryczną reakcję częstości rytmu serca na epizody bezdechu [128].

Celem cytowanego badania było dalsze zgłębienie relacji między HRA a poszczególnymi fazami snu — obszaru stosunkowo słabo poznanego, głównie ze względu na nowość koncepcji HRA w badaniach nad snem. Wykorzystanie tej metody oferuje świeże spojrzenie na analizę snu, a praca ta dąży do wypełnienia luki w wiedzy, oceniając, w jaki sposób asymetryczne opisy rytmu serca korelują z różnymi fazami snu.

Aby przeprowadzić kompleksową analizę, porównuje się także otrzymane wyniki z innymi, już dobrze znanymi miernikami, które zwykle wykorzystuje się w celu badania relacji między odstępami RR a stadiów snu. Wskaźniki te obejmują m.in. *SDNN*, stosunek składowej wysokiej do niskiej częstotliwości (HF/LF) oraz entropię próbkową (SampEn). Dzięki temu możliwe jest ukazanie zarówno zalet, jak i ograniczeń podejścia opartego na HRA w kontekście tradycyjnych miar, a także zademonstrowanie potencjalnych korzyści płynących z nowej metody dla badań nad snem.

2.7.4 Ograniczenia badań nad klasyfikacją faz snu opartych wyłącznie na szeregu odstępów RR

Szereg czasowy odstępów RR oraz jego pochodne miary, takie jak HRV, miary złożoności czy HRA, są zdecydowanie bardziej dostępne oraz tańsze w pozyskaniu niż pełne zapisy polisomnograficzne. Z tego powodu wzbudzają duże zainteresowanie, ponieważ można je potencjalnie zaimplementować w niewielkich, tanich i przenośnych urządzeniach, a nawet w telefonach komórkowych. Ostatnie postępy w rozwoju technologii ubieralnych sprawiły, że możliwe stało się ciągle monitorowanie HRV w trakcie całej nocy, co stwarza nowe możliwości w zakresie długoterminowych badań nad snem oraz spersonalizowanego zarządzania zdrowiem snu. Urządzenia takie jak pierścionek Oura cechują się wysoką dokładnością pomiaru HRV w porównaniu z urządzeniami klasy medycznej (ECG), co potwierdza ich użyteczność w codziennym monitorowaniu snu i zarządzaniu stylem życia [42].

Wiele prac poświęcono badaniom nad wykorzystaniem samego HRV do klasyfikacji stadiów snu. Jest to jednak zadanie trudniejsze, gdyż dysponuje się mniejszą liczbą danych. Można przypuszczać, że podobne ograniczenia będą dotyczyły metod opartych na złożoności czy HRA.

Klasyfikacja stadiów snu wyłącznie na podstawie zmienności rytmu serca, miar złożoności czy asymetrii bywa utrudniona ze względu na złożoną interakcję między ośrodkowym (CNS, ang. *central nervous system*) a autonomicznym (ANS, ang. *autonomic nervous system*) układem nerwowym podczas snu. Jak podkreślono w [129], HRV nie w pełni oddaje subtelne sprzężenia pomiędzy tymi systemami, a stadia snu zdefiniowane na podstawie EEG nie zawsze mają bezpośrednie odwzorowanie w parametrach HRV. Dlatego, mimo że HRV dostarcza informacji o regulacji autonomicznej, niekoniecznie w sposób niezawodny przekłada się na klasyfikację snu w ujęciu EEG. Potwierdzają to także wyniki Stein i wsp. [111], którzy wykazali, że choć ogólne wzorce HRV, takie jak obniżona częstość rytmu serca i zmienność w stadiach snu 1–3 oraz wyższe wartości w fazie REM, można zidentyfikować, to zależą one w dużym stopniu od wieku i stanu klinicznego. Na przykład starsi uczestnicy badania *Sleep Heart Health Study* nie wykazywali typowego wzrostu stosunku LF/HF w fazie REM, a u pacjentów po zawale mięśnia sercowego (MI) zaobserwowano wzrost zamiast spodziewanego spadku LF/HF przy przejściu ze stanu czuwania do snu NREM.

Należy również zauważyć, że zapis EKG, choć odzwierciedla zmiany w układzie autonomicznym podczas snu, nie umożliwia samodzielnie precyzyjnego wyodrębnienia wszystkich stadiów snu, nawet jeśli jest uzupełniony o sygnały oddechowe. Jak zauważono w [130], sygnały fizjologiczne takie jak EKG czy oddychanie są wprawdzie modulowane przez stadia snu, jednak

ich dokładna identyfikacja wymaga dodatkowych danych, np. nagrań EEG.

Celem niniejszej pracy jest zbadanie relacji między asymetrią rytmu serca (HRA) a poszczególnymi fazami snu — zagadnienia, które nie zostało dotąd szczegółowo przeanalizowane ze względu na nowość koncepcji HRA w badaniach nad snem. Aby przeprowadzić kompleksową analizę, oprócz HRA wykorzystuje się także inne, dobrze znane miary, tradycyjnie stosowane w opisie relacji pomiędzy odstępami RR a stadiów snu: *SDNN*, stosunek niskiej do wysokiej częstotliwości (LF/HF) oraz entropię próbkową (SampEn). Porównanie asymetrycznych deskryptorów rytmu serca z tymi konwencjonalnymi wskaźnikami ma na celu podkreślenie unikalnych informacji płynących z HRA oraz wskazanie jej potencjalnego zastosowania w badaniach nad snem.

Uzasadnieniem dla próby wykrywania stadiów snu wyłącznie na podstawie odstępów RR jest fakt, że — jak wspomniano we wstępie — istnieje wiele urządzeń, które mogą rejestrować sygnał EKG i wyznaczać odstęp RR w sposób znacznie tańszy niż pełna polisomnografia. Metoda pozwalająca na detekcję faz snu bez konieczności wykorzystania EEG mogłaby stanowić rozwiązanie mniej inwazyjne, bardziej dostępne i potencjalnie tańsze, co umożliwiłoby szersze zastosowania w warunkach klinicznych i domowych.

Rozdział 3

Materiały badawcze

W niniejszej pracy wykorzystano trzy rodzaje danych: dane holterowskie pochodzące z otwartych źródeł, 30 minutowe nagrania EKG pochodzące z Katedry i Kliniki Intensywnej Terapii Kardiologicznej i Chorób Wewnętrznych Uniwersytetu Medycznego w Poznaniu oraz dane polisomnograficzne zebrane w Oddziale Neurologii Szpitala w Zielonej Górze.

Wszystkie dane zostały zanonimizowane zgodnie z obowiązującymi przepisami dotyczącymi ochrony danych osobowych i etyki badań naukowych. Uzyskano odpowiednie zgody na wykorzystanie danych klinicznych do celów badawczych, co zapewnia zgodność z normami etycznymi i prawnymi.

3.1 Nagrania holterowskie

W tym badaniu wykorzystano zanonimizowane dane z Bazy Danych długotrwałych zapisów migotania przedsionków (LTAfDB) [16]. LTAfDB składa się z 84 rozszerzonych, 24-godzinnych zapisów elektrokardiograficznych (EKG) metodą Holtera, próbkowanych z częstotliwością 128 Hz. Baza dostarcza informacje o lokalizacjach załamków R oraz odpowiadających im rytmach serca (normalny, nadkomorowy, komorowy, przedsionkowy i artefakty techniczne) u pacjentów z napadowym migotaniem przedsionków i innymi arytmiami. Do analizy wybrano wyłącznie ciągłe fragmenty EKG z migotaniem przedsionków (AF) lub rytmem zatokowym (SR) trwające co najmniej 60 sekund, pomijając segmenty oznaczone jako inne rytmy.

Procedura wstępnego przetwarzania danych została zilustrowana na Rysunku 3.1. Szeregi czasowe odstępów RR podzielono na przyległe segmenty trwające 60 sekund. Aby cały segment został zaklasyfikowany jako SR, każdy odstęp RR musiał pochodzić z rytmu zatokowego; w przeciwnym razie segment był wykluczany z analizy. Analogicznie, dla segmentów AF, każde

pobudzenie musiało być AF, a segmenty zawierające pobudzenia komorowe również były usuwane. W celu zminimalizowania liczby potencjalnie niezidentyfikowanych artefaktów technicznych, usunięto odstępy RR krótsze niż 240 ms lub dłuższe niż 3000 ms zarówno dla SR, jak i AF. Segmenty, w których całkowita długość usuniętych odstępów RR przekraczała 3% (1,8 s) długości segmentu, były wykluczane z analizy.

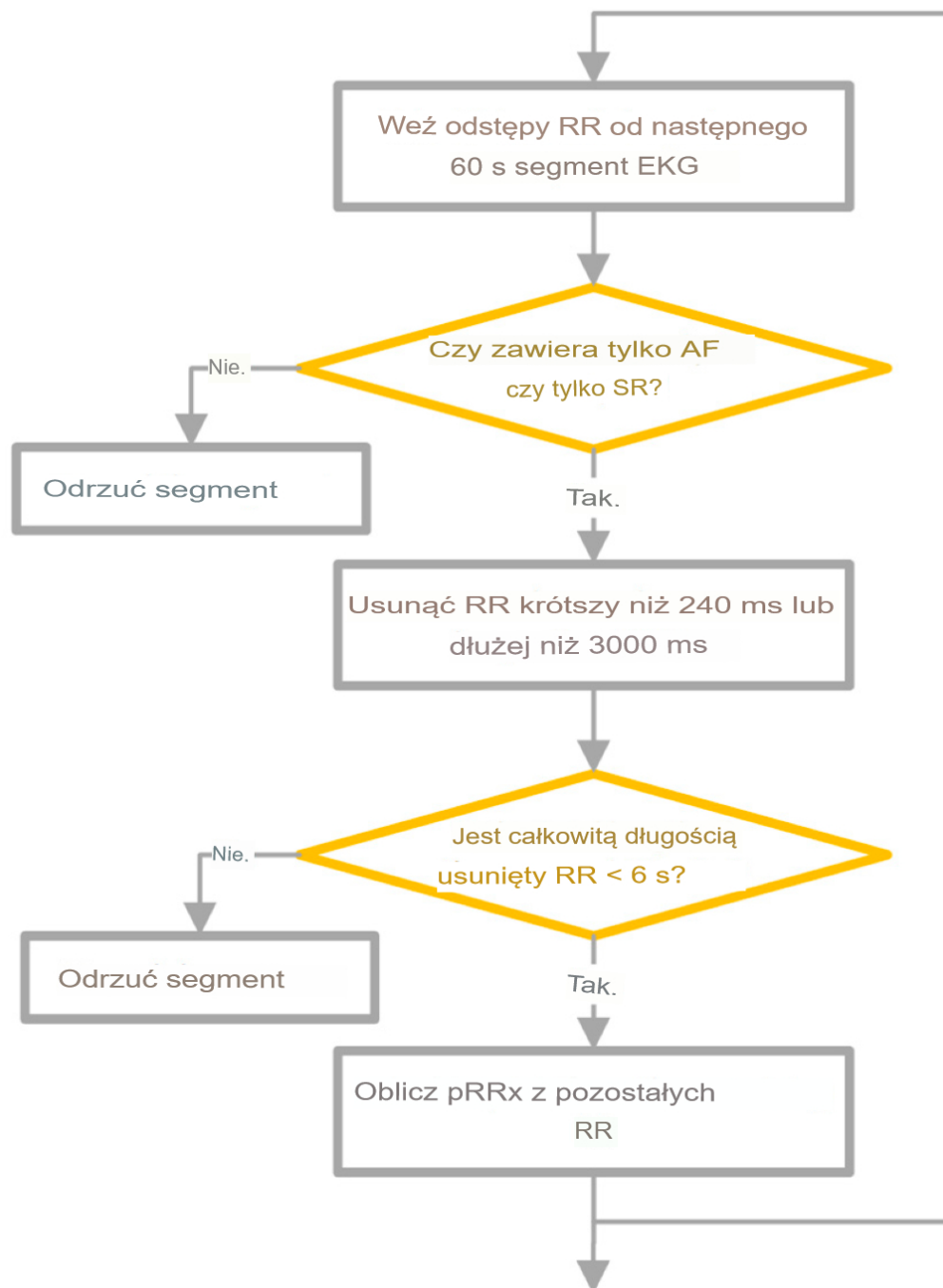
Wybór progu 3% dla wykluczania segmentów na podstawie całkowitej długości usuniętych odstępów RR był oparty na ustalonych wytycznych w tej dziedzinie. W szczególności, [9] zalecają wykluczanie segmentów, w których ponad 10% odstępów RR nie pochodzi z rytmu zatokowego lub są artefaktami. Nasza decyzja o zmniejszeniu tego progu do 3% wynikała głównie ze specyfiki naszego badania, które obejmowało bardzo krótkie, jednodominutowe segmenty. Po wstępnym przetworzeniu uzyskano łącznie 63 636 serii RR trwających 60 sekund (31 919 SR, 31 717 AF). Segmenty EKG, w których całkowita długość odstępów RR po filtrowaniu była krótsza niż 58 sekund, również zostały odrzucone.

3.2 Nagrania 30 minutowe

W badaniu wykorzystano 388 stacjonarnych 30-minutowych zapisów EKG pochodzących od zdrowych młodych osób w wieku od 20 do 40 lat, w tym 233 kobiet. Badanie przeprowadzono w spoczynku w pozycji leżącej na plecach, w neutralnych warunkach środowiskowych. Podczas całego badania uczestnicy oddychali spontanicznie. Zapis 30-minutowy został wykonany po uprzednim 15-minutowym okresie przeznaczonym na adaptację układu sercowo-naczyniowego. Krzywe EKG próbkowano z częstotliwością 1600 Hz przy użyciu przetwornika analogowo-cyfrowego (Porti 5, TMSI, Holandia). Do przetwarzania danych oraz automatycznej klasyfikacji uderzeń serca jako zatokowych, komorowych, nadkomorowych oraz artefaktów zastosowano oprogramowanie libRASCH/RASCHlab (v. 0.6.1, www.librash.org, Raphael Schneider, Niemcy). Klasyfikacja automatyczna została następnie zweryfikowana przez wykwalifikowanego technika, który skorygował błędne klasyfikacje uderzeń.

3.3 Dane polisomnograficzne

W badaniu wzięło udział 31 osób, w tym 11 kobiet, z różnymi schorzeniami, głównie skupiającymi się na zaburzeniach snu. Wiek uczestników wahał się od 27 do 84 lat, przy medianie wieku wynoszącej 58 lat.



Rysunek 3.1: Diagram przedstawiający algorytm wstępnego przetwarzania danych holterowskich.

Uczestnicy prezentowali różnorodne choroby współistniejące. Do najczęściej występujących należały:

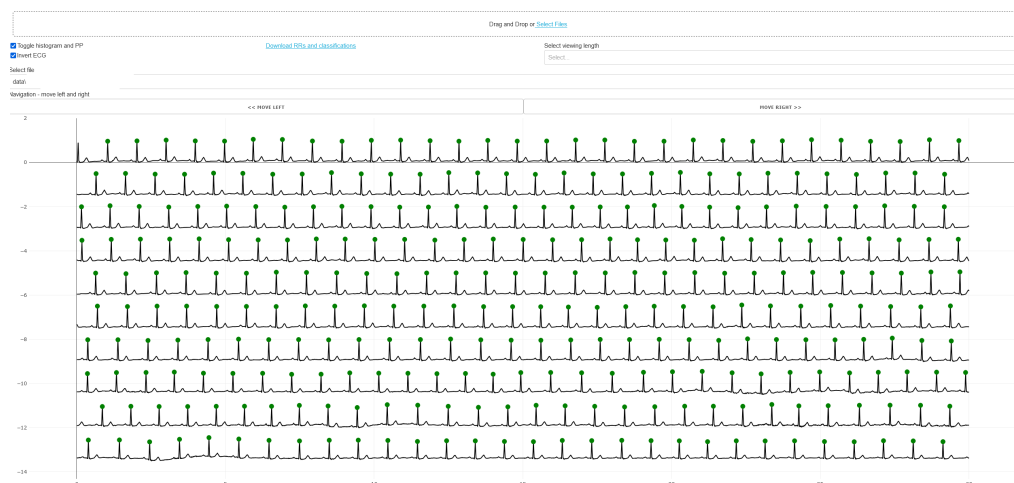
- nadciśnienie tętnicze — 8 pacjentów,
- obturacyjny bezdech senny (OBS) — 5 pacjentów,
- bezdech senny — 12 pacjentów,
- zaburzenia snu — 11 pacjentów,
- cukrzyca — 2 pacjentów,
- udar niedokrwienny mózgu — 1 pacjent,
- przemijający atak niedokrwienny (TIA) — 2 pacjentów,
- otyłość — 1 pacjent,
- astma — 1 pacjent.

W niniejszym badaniu wykorzystano dane zarejestrowane podczas polisomnografii, szczególnie z kanału EKG. Użyto urządzenia Comet PSG Plus XL (Grass Technologies). Dane zostały wyeksportowane do plików EDF, a następnie z każdego zapisu wyodrębniono sygnał EKG. Kompleksy QRS zostały początkowo zaanotowane automatycznie, następnie po przeszkoleniu w tym zakresie osobiście zweryfikowałem poprawność automatycznej anotacji i w razie potrzeby ją skorygowałem. Do obu tych zadań wykorzystano autorskie oprogramowanie Signalweaver (<https://github.com/jaropis/signalweaver>).

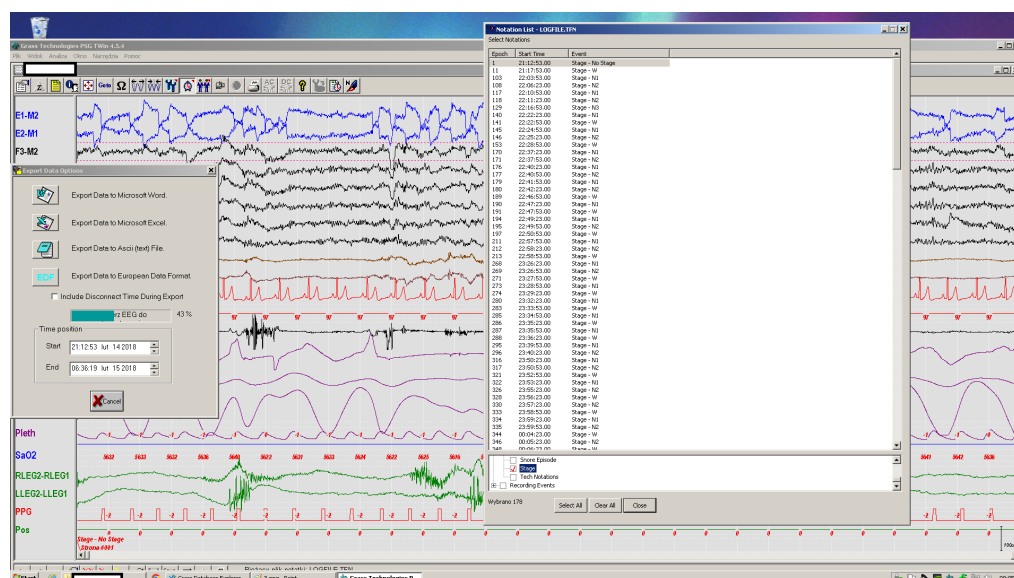
Oryginalne zapisy zostały zaanotowane przez eksperta od snu, który oznaczył poszczególne stadia snu. Niestety pojawił się problem techniczny. Urządzenie Comet PSG Plus XL pozwala na zapisywanie anotacji faz snu wyłącznie do binarnego formatu .TNF, który można otworzyć wyłącznie na urządzeniu z zainstalowanym oprogramowaniem Grass technologies Twin. Jedynym rozwiązaniem tego problemu okazało się utworzenie zrzutów ekranów anotacji, a następnie w warunkach domowych ręczne ich przepisywanie. Był to jeden z głównych problemów w trakcie prac nad doktoratem.

W celu ułatwienia bardziej szczegółowej analizy, zapisy podzielono na segmenty 5-minutowe. Następnie do tych segmentów przypisano stadia snu na podstawie zaanotowanych danych i określonych poniżej reguł.

Fazy snu i przejścia między nimi zostały określone za pomocą własnego skryptu w Pythonie, który wykorzystuje specyficzne reguły do klasyfikacji i wykrywania przejść:



Rysunek 3.2: Zrzut ekranu przedstawiający aplikację signalweaver do automatycznej detekcji odstępów RR wraz z możliwością ich ręcznej korekty. Signalweaver stanowi jeden z oryginalnych wkładów mojej pracy, pozwolił na skrócenie wstępnej analizy zapisów polisomnograficznych z około 4 godzin do 2 minut.



Rysunek 3.3: Zrzut ekranu przedstawiający aplikację Grass technologies Twin z otwartym oknem do anotowania faz snu. Te dane były potem ręcznie przepisywane ze względu na brak możliwości ich eksportu do formatu tekstowego.

1. **Wykrywanie określonych faz:** Pewne fazy (np. N1, N2) zostały zidentyfikowane poprzez dodanie nowych kolumn i oznaczenie segmentów, w których wartość stadium była większa niż ustalony próg (0,6).
2. **Inicjalizacja kolumn przejść:** Dla każdej zmiany fazy snu (np. z N1 do N2, z W do N1) w zbiorze danych zainicjalizowano nową kolumnę do oznaczania wystąpienia przejścia.
3. **Wykrywanie przejść:** Przejścia między fazami snu w każdym zapisie zostały zidentyfikowane na podstawie następujących kryteriów:
 - Przejście jest oznaczone, jeśli w analizowanym segmencie zarówno kolumny fazy początkowej, jak i końcowego są niezerowe, wartość fazy końcowej jest mniejsza lub równa progowi (0,6), a wartość fazy końcowej nie jest mniejsza od poprzedniej wartości.
 - Przejście jest również oznaczone, jeśli kolumna fazy początkowej jest niezerowa, a odpowiadający jej element w kolumnie fazy końcowej wynosi zero, ale kolejny element w kolumnie fazy końcowej jest niezerowy i mniejszy lub równy progowi.

Te metody zastosowano do zbioru danych w celu utworzenia nowych kolumn wskazujących określone fazy i przejścia, co ułatwiło klasyfikację faz snu.



Rysunek 3.4: Poglądowe zdjęcie przedstawiające badanie polisomnograficzne. W trakcie badania wykwalifikowany technik ręcznie anotuje fazy snu pacjenta na podstawie mierzonych sygnałów oraz reakcji pacjenta. (źródło: [131])



Rysunek 3.5: Zdjęcie przedstawiające polisomnograf Comet PSG Plus XL (Grass Technologies) należący do oddziału Neurologicznego Szpitala Zielonogórskiego w Zielonej Górze.

Rozdział 4

Cele

W niniejszym rozdziale sformułowane zostaną kluczowe cele pracy, stanowiące naturalne rozwinięcie zagadnień przedstawionych we wcześniejszych częściach. Dotychczasowe analizy wskazały na istotną rolę serii monotonicznych uzyskanych z szeregu odstępów RR oraz na możliwość ich teoretycznego opisu z uwzględnieniem miar entropijnych wykorzystanych do ich oceny ilościowej. Otwiera to perspektywy nie tylko dla pogłębienia rozumienia struktury i złożoności rytmów serca, ale także dla praktycznych zastosowań w klasyfikacji, diagnostyce oraz charakteryzowaniu stanów fizjologicznych.

Pierwszym obszarem zainteresowań będzie zbadanie, czy parametry entropiczne serii monotonicznych są w stanie skutecznie wykrywać asymetrie w szeregach czasowych oraz czy wyniki uzyskane przy ich użyciu są spójne z innymi metodami detekcji asymetrii, takimi jak parametry HRA (np. $SD1a/d$).

W tym celu opracowano teoretyczny model oparty na funkcji Weierstrassa z asymetrycznym szumem. W ramach tej analizy przedstawiono analityczny opis wkładu asymetrycznego szumu do wariancji, umożliwiając teoretyczne zrozumienie wpływu asymetrii na badane parametry. Następnie przeprowadzono numeryczne obliczenia, które pozwoliły porównać różnice między wartościami parametrów entropicznych a parametrami HRA. Wyniki tych analiz ilościowych dostarczą istotnych informacji na temat przydatności różnych metod wykrywania asymetrii oraz ich wzajemnej zgodności w kontekście analiz szeregów czasowych.

Jako pewnego rodzaju kontynuacja badań nad zdolnością parametrów serii monotonicznych do wykrywania asymetrii w szeregach czasowych, będzie porównanie ich ze skalami wykładniczymi z metody ADFA. Sprawdzenie zbieżności wyników między tymi dwiema metodami pozwoli nam ocenić czy są one zdolne do wykrywania tych samych zjawisk fizjologicznych pomimo różnych metod matematycznych.

Kolejną płaszczyzną badań będzie ocena zdolności predykcyjnych miar entropijnych wyznaczonych z serii monotonicznych szeregu RR, wspomagana innymi miarami zmienności rytmu serca. W szczególności istotne jest sprawdzenie, czy powyższe miary pozwalają na skuteczną klasyfikację faz snu umożliwiając w ten sposób głębsze zrozumienie neurofizjologicznych korelatów zmian rytmów serca w różnych stanach czuwania i wypoczynku. Analogiczne analizy zostaną przeprowadzone dla detekcji migotania przedsionków - zaburzeniu o istotnym znaczeniu klinicznym.

Ostatnim obszarem badawczym tej pracy jest rozwój oprogramowania komputerowego wykorzystywanego w tej pracy. Na potrzeby niniejszych badań zostało przygotowane oprogramowanie komputerowe napisane w języku Python do obliczeń serii monotonicznych dla dowolnych szeregów czasowych wraz z obliczaniem parametrów entropijnych, co zostało upublicznione w postaci oprogramowania open-source. Kolejnym oprogramowaniem, które zostało opracowane i upublicznione w ramach tej pracy doktorskiej zostało narzędzie do przetwarzania zapisów PSG pochodzących z komercyjnych urządzeń i przygotowania ich do pracy badawczej, zarówno w trybie aplikacji z graficznym interfejsem nad pojedynczym nagraniem jak i do automatycznego przygotowania wielu nagrań [132]. W badaniach nad klasyfikacją faz snu, jeden z parametrów - entropia z próby został poddany wnikliwej analizie od strony algorytmicznej. W efekcie czego zrefaktorowano oprogramowanie obliczeniowe sum korelacyjnych, niezbędnych do obliczenia entropii z próby, oraz przygotowano analizę wydajnościową wykorzystywanego algorytmu. Wyniki tych prac są przygotowane do opublikowania w piśmie naukowym [93].

Cele pracy można zdefiniować następująco:

1. studium literaturowe wykorzystania HRV, metod złożonościowych oraz HRA w badaniach nad snem,
2. opisanie oraz przetestowanie, przy użyciu danych syntetycznych i eksperymentalnych, teoretycznego rozkładu serii monotonicznych wraz z uwzględnieniem szumu asymetrycznego,
3. zbadanie relacji między ADFA i deskryptorami serii monotonicznych,
4. zbadanie predykcyjnych zdolności miar entropijnych serii monotonicznych szeregu RR, we współpracy z pozostałymi miarami zmienności i złożoności rytmu serca, w szczególności *SampEn*, do detekcji poszczególnych faz snu,

5. zbadanie predykcyjnych zdolności miar entropijnych serii monotonicznych szeregu RR, w połączeniu z innymi miarami zmienności rytmu serca, do detekcji migotania przedsionków,
6. zbadanie algorytmu do obliczeń sum korelacyjnych oraz jego reimplementacja,
7. opracowanie metodologii badawczej służącej przystosowaniu sygnałów dostępnych w technologii komercyjnej do badań naukowych.

Rozdział 5

Wyniki

W niniejszym rozdziale zaprezentowane zostaną wszystkie wyniki uzyskane w toku moich badań naukowych, zarówno te, które doczekały się już publikacji w recenzowanych czasopismach naukowych, jak i te, które obecnie są przygotowywane do przesłania. Dla większej przejrzystości rozdział podzielony został na pięć sekcji, odpowiadające odmiennym, lecz komplementarnym obszarom badawczym.

W pierwszej sekcji zostaną omówione prace dotyczące zbadania własności serii monotonicznych oraz entropii tych serii w kontekście wykrywania asymetrii w szeregach czasowych. Podstawową motywacją do tych badań jest fakt, że metody HRA . Aby zilustrować tę kwestię, przeprowadzono konstrukcję procesów opartych na modyfikacjach funkcji Weierstrassa [133] z dodanym asymetrycznym szumem [134, 135]. Dla uzyskanych procesów obliczyłem wariancje całkowite oraz określiłem wkłady poszczególnych komponentów (tj. szumów oraz dodanych serii) w ich wartość. Następnie, na podstawie wygenerowanych danych, policzyłem deskryptory HRA ($SD1a$ i $SD1d$) i entropie serii monotonicznych w celu oceny, w jakim stopniu uwzględniają one asymetrię wzrostów i spadków. Otrzymane wyniki zostały poddane analizie z zastosowaniem testu dwumianowego, który pozwala zweryfikować istotność statystyczną zdolności poszczególnych miar do wykrywania asymetrii.

W drugiej sekcji zaprezentowane zostaną wyniki porównujące zdolność wykrywania asymetrii za pomocą metody serii monotonicznych oraz metody ADFA. Metoda ADFA jest już dobrze przebadana i wiemy, że skutecznie wykrywa asymetrię zarówno w danych syntetycznych, jak i w szeregach odstępów RR. W niniejszym badaniu szeregi czasowe parametrów α^+ i α^- , uzyskane dla każdego zapisu, zostały podsumowane za pomocą median w celu przeprowadzenia porównań. Ponieważ wartości średnie α^+ i α^- nie miały rozkładu normalnego (test Shapiro-Wilka), do porównań zastosowano nieparametryczny test sparowany Wilcoxa.

Analizy zależności pomiędzy asymetrycznymi wykładnikami ADFA, a deskryptorami zmienności rytmu serca (HRA) przeprowadzono przy użyciu nieparametrycznego testu korelacji Spearmana. Aby sprawdzić czy asymetria HRA wiąże się z relacją $\alpha^+ < \alpha^-$, skonstruowano jednowymiarowe modele regresji logistycznej dla tablic kontyngencji 2×2 , które zestawiały istnienie HRA z warunkiem $\alpha^+ < \alpha^-$.

Badanie to pozwoliło na dokładne porównanie metod oraz lepsze zrozumienie relacji między asymetrią wykrytą metodą ADFA a zmiennością rytmu serca (HRA).

W trzeciej sekcji przedstawione zostaną rezultaty dotyczące wykorzystania wspomnianych wyżej deskryptorów entropicznych oraz wariacyjnych wyznaczanych z szeregu RR w kontekście klasyfikacji faz snu i detekcji przejść między tymi fazami. Zaprezentuję zarówno metodykę analiz, jak i szczegółową dyskusję uzyskanych wyników, wskazując na korzyści oraz ograniczenia proponowanego podejścia.

W czwartej sekcji skoncentruję się na wynikach związanych z zastosowaniem entropii serii monotonicznych w analizie krótkich, jednominutowych segmentów interwałów RR w celu detekcji migotania przedsionków. Przedstawię zarówno kluczowe aspekty metodologii, jak i uzyskane wskaźniki skuteczności tej detekcji.

Piąta, ostatnia sekcja zawiera szczegółowy opis reimplementacji algorytmu NCM w dwóch wersjach, służącego do obliczeń sum korelacyjnych, wraz z dogłębną analizą wydajnościową.

Taki układ rozdziału ma na celu ukazanie wielopłaszczyznowego zastosowania zarówno deskryptorów HRA, jak i miar opartych na entropii serii monotonicznych. Otwiera to możliwość porównania oraz zidentyfikowania, w jakich warunkach i kontekstach dane metody analityczne ujawniają swoją największą przydatność, a także wskazuje kierunki dalszych badań nad analizą sygnałów fizjologicznych.

5.1 Matematyczne własności entropii serii monotonicznych

W tej sekcji przedstawione zostanie teoretyczna analiza ogólnego procesu z asymetrycznym szumem oraz wyprowadzenie teoretycznej postaci wariancji realizacji funkcji Weierstrassa, która jest najczęściej używana w obliczeniach numerycznych [136]. W trakcie kwerendy nad tym problemem nie udało się znaleźć nigdzie spójnej postaci tej wariancji, więc została ona ręcznie wyprowadzona, po to, żeby sprawdzić, czy można stosować poniższe rozważania

po dodaniu asymetrycznego szumu. Dla ustalenia notacji najpierw przedstawione zostaną powszechnie znane rozważania dotyczące wariancji sumy procesów, a następnie wyprowadzony zostanie wzór na wkłady do wariancji ogólnego procesu po dodaniu asymetrycznego szumu. Jest to oryginalny wkład doktoranta. Obliczenia teoretyczne zostały potwierdzone symulacjami numerycznymi.

5.1.1 Wariancja sumy - dwa elementy

Rozważmy zmienną losową $f(x)$ o wariancji A . W przypadku niezależnych zmiennych losowych, wariancje się sumują. W szczególności, jeśli X i Y są niezależne:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y). \quad (5.1)$$

Ta własność dotyczy również sum więcej niż dwóch niezależnych zmiennych losowych. Załóżmy:

$$f(x) \text{ ma wariancję } A, \quad (5.2)$$

$$X_1 \sim N(0, 1), \quad X_2 \sim N(0, 1/2), \quad X_3 \sim N(0, 1/2), \quad (5.3)$$

oraz wszystkie zmienne są niezależne od siebie i od $f(x)$.

Pierwszy scenariusz:

$$\text{Var}(f(x) + X_1) = \text{Var}(f(x)) + \text{Var}(X_1) = A + 1. \quad (5.4)$$

Drugi scenariusz:

$$\text{Var}(f(x) + X_2 + X_3) = \text{Var}(f(x)) + \text{Var}(X_2) + \text{Var}(X_3). \quad (5.5)$$

Ponieważ $\text{Var}(X_2) = \text{Var}(X_3) = \frac{1}{2}$:

$$\text{Var}(f(x) + X_2 + X_3) = A + \frac{1}{2} + \frac{1}{2} = A + 1. \quad (5.6)$$

W związku z tym oba scenariusze dają tę samą całkowitą wariancję $A + 1$.

5.1.2 Wariancja sumy - n elementów

Zacznijmy od znanej wariancji z pierwszego scenariusza:

$$\text{Var}(f(x) + N(0, 1)) = \text{Var}(f(x)) + \text{Var}(N(0, 1)) = A + 1. \quad (5.7)$$

Teraz chcemy, aby w drugim scenariuszu, w którym uzyskujemy losową liczbę n z zakresu od 1 do 10, całkowita wariancja również wynosiła $A + 1$.

W tym drugim przypadku rozważamy:

$$f(x) + \sum_{i=1}^n N(0, S_n). \quad (5.8)$$

Warunkując na konkretną wartość n , wariancja wynosi:

$$\text{Var}(f(x) + \sum_{i=1}^n N(0, S_n)) = \text{Var}(f(x)) + n \cdot S_n = A + nS_n. \quad (5.9)$$

Aby ta wariancja dla dowolnego zrealizowanego n była równa $A + 1$, musi zachodzić:

$$nS_n = 1 \quad \text{dla każdego } n. \quad (5.10)$$

Rozwiązując to równanie otrzymujemy:

$$S_n = \frac{1}{n}. \quad (5.11)$$

Przy tym wyborze, niezależnie od wylosowanej liczby n , suma n niezależnych zmiennych normalnych, każda o wariancji $\frac{1}{n}$, daje łączną wariancję równą 1. W rezultacie całkowita wariancja pozostaje $A + 1$, czyli taka sama jak w przypadku dodania tylko jednej zmiennej $N(0, 1)$.

5.1.3 Wariancja funkcji z asymetrycznym szumem

Definicja problemu:

Mamy dany proces stacjonarny $f(t_i)$ o średniej 0 i wariancji A . Definiujemy nowy proces:

$$g(t_i) = f(t_i) + \eta_i, \quad (5.12)$$

gdzie zmienna losowa η_i zależy od kierunku zmiany $f(t_i)$:

$$\eta_i = \begin{cases} N(0, 1 - 1/k), & \text{jeśli } f(t_i) > f(t_{i-1}), \\ N(0, 1 + 1/k), & \text{jeśli } f(t_i) < f(t_{i-1}). \end{cases} \quad (5.13)$$

Założmy, że $f(t_i)$ jest symetryczny i stacjonarny, co oznacza w szczególności, że prawdopodobieństwo wzrostu jest równe prawdopodobieństwu spadku i wynosi $1/2$.

Wariancja $g(t_i)$:

Ponieważ $g(t_i) = f(t_i) + \eta_i$ oraz $f(t_i)$ jest niezależne od η_i , dostajemy:

$$\text{Var}(g(t_i)) = \text{Var}(f(t_i)) + \text{Var}(\eta_i) = A + \text{Var}(\eta_i). \quad (5.14)$$

Obliczmy $\text{Var}(\eta_i)$:

$$\text{Var}(\eta_i) = \frac{1}{2}(1 - \frac{1}{k}) + \frac{1}{2}(1 + \frac{1}{k}) = \frac{1}{2}(1 - \frac{1}{k} + 1 + \frac{1}{k}) = \frac{1}{2}(2) = 1. \quad (5.15)$$

Zatem

$$\text{Var}(g(t_i)) = A + 1. \quad (5.16)$$

Wkład wzrostów i spadków do wariancji:

Ponieważ prawdopodobieństwo wzrostu i spadku jest równe $1/2$, możemy założyć, że połowa wariancji $f(t_i)$, czyli $\frac{A}{2}$, przypada na wzrosty, a połowa, czyli $\frac{A}{2}$, na spadki. W przypadku wzrostu:

$$\text{Wkład} = \frac{1}{2}A + \frac{1}{2}(1 - \frac{1}{k}). \quad (5.17)$$

W przypadku spadku wyraźnie widać, że wkłady spadków i wzrostów są różne przy takiej definicji procesu z asymetrycznym szumem. Jest to analogiczne do szeregu odstępów RR.:

$$\text{Wkład} = \frac{1}{2}A + \frac{1}{2}(1 + \frac{1}{k}). \quad (5.18)$$

Po zsumowaniu wkładów z obu przypadków otrzymujemy całkowitą wariancję $A + 1$:

$$\left(\frac{1}{2}A + \frac{1}{2}(1 - \frac{1}{k})\right) + \left(\frac{1}{2}A + \frac{1}{2}(1 + \frac{1}{k})\right) = A + 1. \quad (5.19)$$

W ten sposób pokazaliśmy, że przy równych szansach na wzrost i spadek procesu $f(t_i)$, wzrosty wnoszą do całkowitej wariancji $g(t_i)$ wkład równy $\frac{1}{2}A + \frac{1}{2}(1 - \frac{1}{k})$.

5.1.4 Wyprowadzenie wariancji procesu opartego na funkcji Weierstrassa

Rozważamy proces losowy zdefiniowany przez:

$$W(t) = \sum_{n=0}^{\infty} \frac{\cos(2^n t)}{2^{(2-D)n}}, \quad (5.20)$$

gdzie $1 < D < 2$. Aby zdefiniować wariancję, traktujemy t jako zmienną losową o równomiernym rozkładzie na przedziale $[0, 2\pi]$. Oznacza to, że wartość oczekiwana $E[\cdot]$ jest rozumiana jako:

$$E[f(t)] = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt. \quad (5.21)$$

Obliczanie wartości oczekiwanej

Najpierw obliczmy wartość oczekiwaną $E[W(t)]$:

$$E[W(t)] = E \left[\sum_{n=0}^{\infty} \frac{\cos(2^n t)}{2^{(2-D)n}} \right]. \quad (5.22)$$

Zakładając zbieżność i możliwość zamiany sumy z całką, mamy:

$$E[W(t)] = \sum_{n=0}^{\infty} \frac{1}{2^{(2-D)n}} E[\cos(2^n t)]. \quad (5.23)$$

Ponieważ t jest równomiernie rozłożone na $[0, 2\pi]$, obliczamy:

$$E[\cos(2^n t)] = \frac{1}{2\pi} \int_0^{2\pi} \cos(2^n t) dt. \quad (5.24)$$

Całka z $\cos(kt)$ po jednym pełnym okresie (lub jego wielokrotności) jest równa zero, o ile $k \neq 0$. Ponieważ $2^n > 0$, otrzymujemy:

$$E[\cos(2^n t)] = 0. \quad (5.25)$$

Stąd:

$$E[W(t)] = \sum_{n=0}^{\infty} \frac{0}{2^{(2-D)n}} = 0. \quad (5.26)$$

Obliczanie drugiego momentu i wariancji

Następnie obliczmy drugi moment $E[W(t)^2]$:

$$W(t)^2 = \left(\sum_{n=0}^{\infty} \frac{\cos(2^n t)}{2^{(2-D)n}} \right) \left(\sum_{m=0}^{\infty} \frac{\cos(2^m t)}{2^{(2-D)m}} \right). \quad (5.27)$$

Możemy to zapisać jako podwójną sumę:

$$W(t)^2 = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\cos(2^n t) \cos(2^m t)}{2^{(2-D)(n+m)}}. \quad (5.28)$$

Bierzemy wartość oczekiwaną:

$$E[W(t)^2] = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{E[\cos(2^n t) \cos(2^m t)]}{2^{(2-D)(n+m)}}. \quad (5.29)$$

Wykorzystujemy tożsamość:

$$\cos(\alpha) \cos(\beta) = \frac{\cos(\alpha - \beta) + \cos(\alpha + \beta)}{2}. \quad (5.30)$$

Rozważmy wartość oczekiwaną:

$$E[\cos(2^n t) \cos(2^m t)] = \frac{1}{2\pi} \int_0^{2\pi} \cos(2^n t) \cos(2^m t) dt. \quad (5.31)$$

Jeśli $n \neq m$, to po zastosowaniu wzoru na sumę i różnicę argumentów otrzymujemy całki z $\cos(kt)$ dla $k \neq 0$, które wynoszą zero. Stąd wszystkie wyrazy o $n \neq m$ znikają. Jeśli $n = m$:

$$E[\cos^2(2^n t)] = \frac{1}{2\pi} \int_0^{2\pi} \cos^2(2^n t) dt. \quad (5.32)$$

Ponieważ $\cos^2(x)$ uśrednia się do $\frac{1}{2}$ po pełnym okresie:

$$E[\cos^2(2^n t)] = \frac{1}{2}. \quad (5.33)$$

Zatem:

$$E[W(t)^2] = \sum_{n=0}^{\infty} \frac{1}{2^{2(2-D)n}} \cdot \frac{1}{2} = \frac{1}{2} \sum_{n=0}^{\infty} \left(2^{-2(2-D)}\right)^n. \quad (5.34)$$

Oznaczmy $\alpha = 2 - D$. Ponieważ $1 < D < 2$, mamy $0 < \alpha < 1$. Wtedy:

$$2^{-2(2-D)} = 2^{-2\alpha}. \quad (5.35)$$

Suma geometryczna:

$$\sum_{n=0}^{\infty} (2^{-2\alpha})^n = \frac{1}{1 - 2^{-2\alpha}}. \quad (5.36)$$

Stąd:

$$E[W(t)^2] = \frac{1}{2} \cdot \frac{1}{1 - 2^{-2(2-D)}}. \quad (5.37)$$

Ponieważ $E[W(t)] = 0$, wariancja to:

$$\text{Var}(W(t)) = E[W(t)^2] - (E[W(t)])^2 = E[W(t)^2] = \frac{1}{2(1 - 2^{-2(2-D)})}. \quad (5.38)$$

5.1.5 Wyniki badania nad wykrywaniem asymetrii w sygnałach Weierstrassa

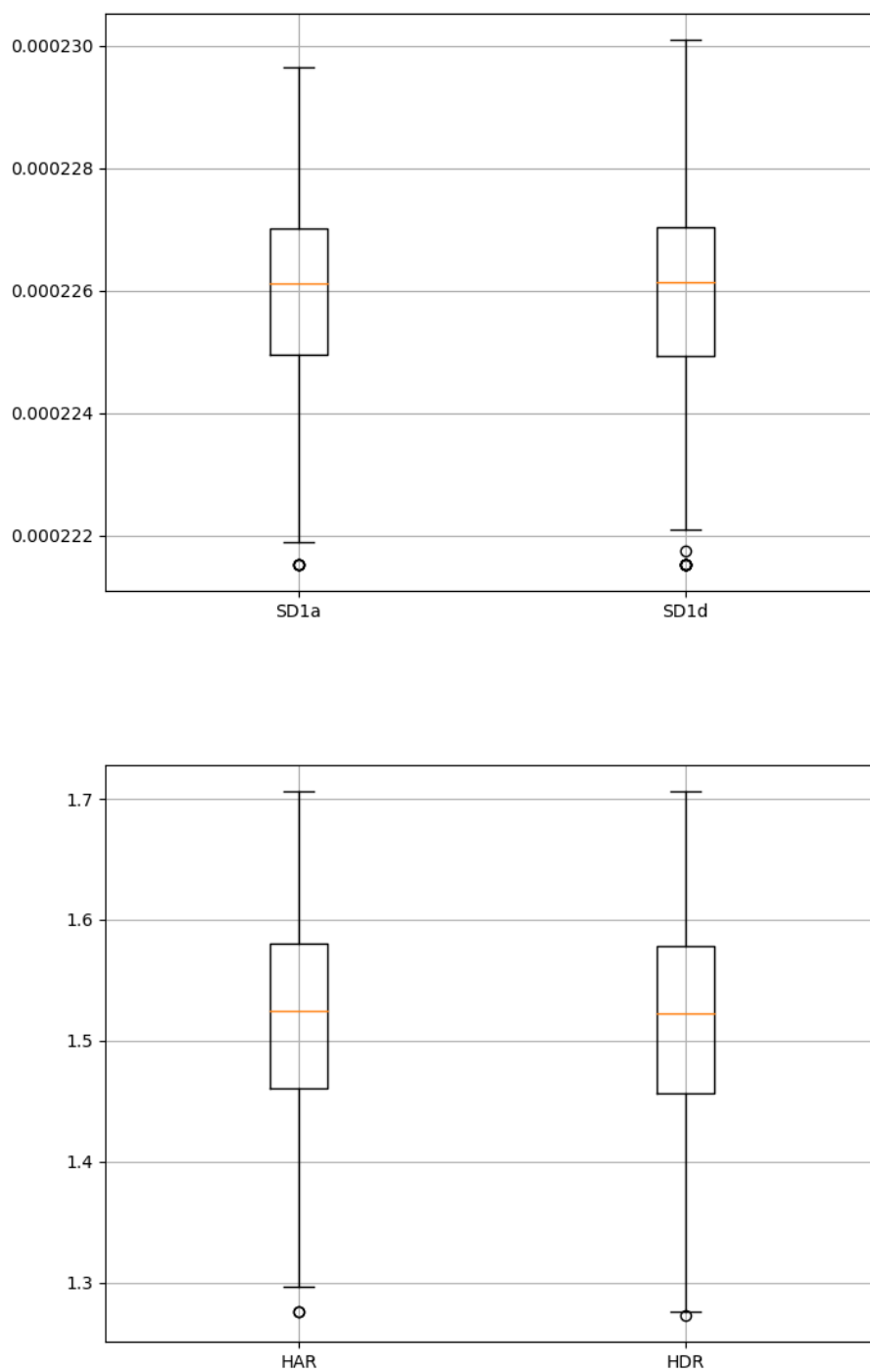
Następna część zajmuje się opisem analizy numerycznej poświęconej zdolności deskryptorów wariancyjnych ($SD1_a$, $SD1_d$) jak i entropijnych opartych na seriach (HAR oraz HDR) do detekcji asymetrii dla sygnałów opartych na funkcji Weierstrassa oraz realizacji funkcji Weierstrassa z dodanym szumem asymetrycznym.

5.1.5.1 Sygnał Weierstrassa bez asymetrii

W pierwszej części badania wygenerowano sygnał funkcji Weierstrassa o długości 1 000 000 elementów. Następnie sygnał został podzielony na 10 000 segmentów, z których losowo wybrano 2 000. Dla wybranych segmentów obliczono parametry $SD1_a$, $SD1_d$, HAR oraz HDR . Dodatkowo przeanalizowano dla ilu segmentów zachodzi relacja $SD1_a > SD1_d$ oraz $HAR > HDR$. Liczby te zostały następnie wykorzystane do przeprowadzenia testu dwumianowego. Wyniki testu dwumianowego wykazały, że zarówno parametry wariancyjne ($SD1_a$ i $SD1_d$), jak i parametry entropiczne (HAR i HDR) nie wykrywały asymetrii w badanym sygnale Weierstrassa. Obserwacja ta jest zgodna z teoretycznymi założeniami, które wskazują na brak asymetrii w takim deterministycznym procesie.

Tabela 5.1: Tabela wyników testu binomialnego wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa.

Porównanie	p-value	Liczba przypadków
$SD1_a > SD1_d$	0.4339	982
$HAR > HDR$	0.4882	984



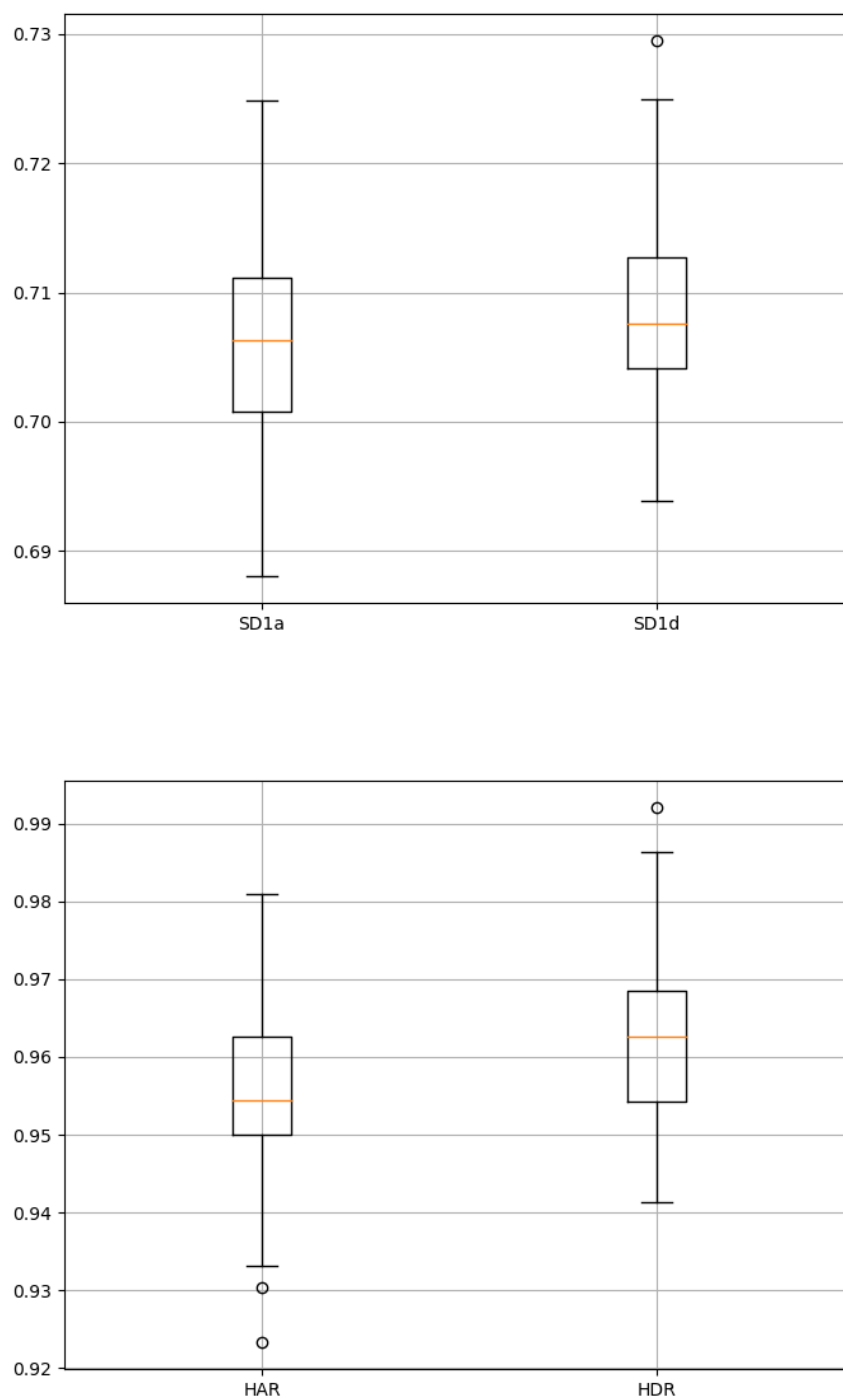
Rysunek 5.1: Wykresy pudełkowe przedstawiające różnice w rozkładach wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa.

5.1.5.2 Sygnał Weierstrassa z asymetrycznym szumem

W kolejnej części badania przeprowadzono podobny eksperyment, lecz dla sygnału funkcji Weierstrassa z asymetrycznym szumem. W tym przypadku, z uwagi na obecność asymetrycznego szumu, proces nie był deterministyczny. Dlatego też nie było potrzeby dzielenia sygnału na segmenty ani losowania ich próbek. Zamiast tego wygenerowano 100 sygnałów o tej samej długości, korzystając z tego samego procesu generowania asymetrycznego szumu. Tylko 100 elementów sprawia, że trudniej jest wykryć różnicę w teście binomialnym. Gdybyśmy zastosowali tyle elementów ile w przypadku procesu bez szumu, to otrzymalibyśmy różnicę istotną statystycznie z powodu samej liczby analizowanych elementów. Zastosowano więc podejście konserwatywne, żeby wyeliminować efekt skali [137]. Ogromna liczba elementów w przypadku deterministycznego procesu Weierstrassa podkreśla jego perfekcyjną symetrię. Podobnie jak wcześniej, dla każdego sygnału obliczono parametry $SD1_a > SD1_d$ oraz $HAR > HDR$, a następnie wyniki wykorzystano do testu dwumianowego. W tym przypadku oba rodzaje deskryptorów skutecznie wykrywały asymetrię w sygnale.

Tabela 5.2: Tabela wyników testu binomialnego wartości $SD1_a$ i $SD1_d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa z asymetrycznym szumem.

Porównanie	p-value	Liczba przypadków
$SD1_a$ większe	3.43×10^{-250}	26
HAR większe	4.50×10^{-231}	39



Rysunek 5.2: Wykresy pudełkowe przedstawiające różnice w rozkładach wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa z asymetrycznym szumem.

5.1.6 Wnioski

Wyniki badania pokazały, że:

1. W przypadku deterministycznego sygnału Weierstrassa brak było wykrywalnej asymetrii zarówno przez parametry wariancyjne HRA ($SD1_a$, $SD1_d$), jak i przez parametry entropiczne serii monotonicznych (HAR , HDR).
2. Dla sygnału funkcji Weierstrassa z asymetrycznym szumem oba rodzaje deskryptorów skutecznie wykrywały asymetrię.

Obserwacje te potwierdzają, że zarówno parametry wariancyjne, jak i entropiczne są spójnymi miarami zdolnymi do wykrywania asymetrii w sygnałach, zgodnie z teoretycznymi przewidywaniami. Eksperymenty te podkreślają potencjał obu rodzajów deskryptorów w analizie dynamicznych systemów z obecnością asymetrii.

5.2 Związek między deskryptorami serii monotonicznych a ADFA

Przeprowadzono analizy, których celem było ustalenie związku między seriami monotonicznymi spowolnień (DR) i przyspieszeń (AR) a asymetrycznymi wykładnikami skalowania ADFA (α^+ i α^-) [135]. W przypadku spowolnień (DR) stwierdzono, że krótkie serie (o długości ≤ 3) nie wykazują silnych korelacji, natomiast serie dłuższe (> 3) cechuje istotna statystycznie współzależność z obydwoma wykładnikami α^+ oraz α^- . Analogiczne wyniki uzyskano dla serii przyspieszeń (AR) — obserwowano wyraźne, istotne korelacje z ADFA przede wszystkim w przypadku dłuższych serii. Sugeruje to, że w szeregach odstępów RR istotną rolę odgrywa zarówno asymetria w rytmie serca (ang. Heart Rate Asymmetry, HRA), jak i obecność wydłużających się ciągów monotonicznych.

5.2.1 Związki między obecnością asymetrii w deskryptorach serii a asymetrią w lokalnych asymetrycznych wykładnikach skalowania ADFA

Dalsza część analizy skupiła się na dłuższych seriach (> 3) oraz sprawdzeniu, w jakim stopniu stwierdzana tam asymetria rytmu serca współwystępuje

z asymetrią w ADFA. Wykazano, że jeśli w danym zapisie EKG dominowały określone wzorce w seriach spowolnień i przyspieszeń, to prawdopodobieństwo występowania określonego kierunku asymetrii ADFA ($\alpha^+ > \alpha^-$ lub $\alpha^+ < \alpha^-$) było znacząco wyższe niż w przypadku zapisu bez tak wyraźnych serii monotonicznych.

Zastosowane modele regresji logistycznej potwierdziły, że stwierdzenie asymetrii w seriach monotonicznych może służyć jako predyktor asymetrii w metodzie ADFA. Współczynniki modeli były istotne statystycznie i wskazywały, że obecność bądź siła asymetrii w ciągach spowolnień i przyspieszeń istotnie wpływa na wartość i kierunek asymetrycznych wykładników skalowania α^+ oraz α^- .

5.2.2 Związki między asymetrią w deskryptorach entropijnych HRA a ADFA

W analizie rozszerzającej przedstawione wyniki uwzględniono także parametry entropijne związane z seriami spowolnień (H_{DR}) i przyspieszeń (H_{AR}). Zaobserwowano istotne i często wysokie korelacje tych parametrów z α^+ i α^- . Przykładowo, im wyższe wartości entropijne H_{DR} , tym większa tendencja do wyższych wartości jednego z asymetrycznych wykładników ADFA, co wspiera hipotezę, że oba podejścia (seria monotoniczna i analiza skalowania) opisują to samo zjawisko fizjologiczne, wykorzystując jednak inne narzędzia i założenia [135].

5.2.3 Dyskusja

Przedstawione wyniki wskazują, że asymetria rytmu serca (HRA), określana na bazie deskryptorów serii monotonicznych, jest silnie powiązana z asymetrią w analizie ADFA. Co istotne, największe znaczenie mają dłuższe serie spowolnień i przyspieszeń (o długości > 3), podczas gdy krótsze serie (o długości ≤ 3) nie wykazują aż tak wyraźnego związku z α^+ i α^- . Silna współzależność tych dwóch podejść (serii monotonicznych oraz lokalnie definiowanych asymetrycznych wykładników skalowania) została dodatkowo wzmocniona analizami regresji logistycznej, które wykazały, że obecność asymetrii w deskryptorach opartych na seriach monotonicznych pozwala przewidywać kierunek i siłę asymetrii w ADFA.

Wyniki potwierdzają także, że parametry entropijne, takie jak H_{DR} i H_{AR} , istotnie korelują z asymetrycznymi wykładnikami skalowania. Jest to kolejne potwierdzenie, że długotrwała struktura i organizacja serii spowolnień oraz przyspieszeń w zapisie EKG ma znaczący wpływ na sposób, w jaki

sygnał rytmu serca wykazuje (lub nie wykazuje) asymetrię w skalowaniu długozasięgowym [135].

Podsumowując, przedstawione badania dowodzą, że:

1. dłuższe serie monotoniczne (zarówno spowolnień, jak i przyspieszeń) wykazują istotne korelacje z asymetrycznymi wykładnikami skalowania ADFA,
2. asymetria serii monotonicznych może być traktowana jako predyktor kierunku i siły asymetrii ADFA, co potwierdzają modele regresji logistycznej,
3. parametry entropijne (np. H_{DR} , H_{AR}) silnie korelują z α^+ i α^- , co wzmacnia wnioski, że oba te podejścia opisują zbliżone aspekty zjawisk fizjologicznych związanych z regulacją rytmu serca,
4. zbieżność dwóch metod — tj. deskryptorów HRA opartych na seriach monotonicznych i wskaźników ADFA — pokazuje, że stanowią one komplementarne sposoby opisu asymetrii rytmu serca.

W efekcie zarówno analiza serii monotonicznych, jak i metoda ADFA (z jej asymetrycznymi wykładnikami α^+ i α^-) można uznać za przydatne i wzajemnie uzupełniające się narzędzia w ocenie zjawisk dynamiki rytmu serca. Pozostaje zachęcające, że uzyskane wyniki w grupie zdrowych młodych osób (rejestrowanych w warunkach stacjonarnego 30-minutowego zapisu EKG) wskazują na spójność obu metod, co wyjaśnia i potwierdza wcześniejsze doniesienia literaturowe [135].

5.3 Klasyfikacja faz snu na podstawie asymetrii rytmu serca

Niniejsza sekcja przedstawia szczegółową analizę rozkładu faz snu oraz rozkładu wybranych parametrów w różnych fazach snu i podczas przejść między nimi. Na początku omawiamy rozkład poszczególnych faz snu oraz przejść między nimi, ilustrując częstotliwość i procentowy udział każdej fazy i przejścia we wszystkich nagraniach, a także w odniesieniu do poszczególnych badanych osób. W dalszej części analizy poddano rozkład symetrycznych i asymetrycznych parametrów w różnych fazach snu. Dane zaprezentowano zarówno w formie graficznej, jak i tabelarycznej, pokazując rozpiętość i tendencje centralne tych parametrów.

Dane są ograniczone, dlatego konieczne jest dokonanie wyboru zmiennych (cech) do analizy. Do dalszych analiz wybrano następujące zmienne:

Zmiennie symetryczne: $SDNN$, $SD1$, SampEn, LF/HF . Te zmienne odzwierciedlają miary zmienności i złożoności sygnałów fizjologicznych i są uznawane za istotne dla zrozumienia dynamiki faz snu. Kolejną symetryczną zmienną używaną w analizie jest $pnn30$, która jest blisko powiązana z powszechnie stosowaną metryką $pnn50$ [110, 138].

Zmiennie asymetryczne: $SD1_a$, $SD1_d$, $pnn30_{dec}$, $pnn30_{acc}$, $AR4$, $DR4$, AR_{MAX} , DR_{MAX} , H_{DR} , oraz H_{AR} .

Wszystkie wybrane metryki zostały obliczone za pomocą darmowego oprogramowania HraExplorer (<https://hraexplorer.com/>) w oknach czasowych trwających 5 minut, odpowiadających segmentacji danych dotyczących faz snu. Kod źródłowy tego oprogramowania jest dostępny w [53].

Następnie przedstawiono wyniki analiz regresyjnych typu GEE (Generalized Estimating Equations) dla faz snu i przejść między nimi, wskazując istotne statystycznie predyktory oraz ich wpływ na prawdopodobieństwo wystąpienia konkretnego przejścia. Kluczowe zmienne, tj. te wymienione w punkcie 5.3, zostały zbadane pod kątem istotności statystycznej oraz wartości prognostycznej.

W ostatniej części omówiono skuteczność modeli predykcyjnych służących do identyfikacji różnych faz snu i przejść pomiędzy nimi. Zaprezentowano macierze konfuzji, miary oceny wydajności (dokładność, precyzja, czułość, F1) oraz krzywe ROC dla każdej fazy snu i przejścia. Umożliwia to jasne przedstawienie efektywności modeli i wskazanie obszarów wymagających dalszej poprawy.

5.3.1 Statystyki opisowe – Rozkład faz snu

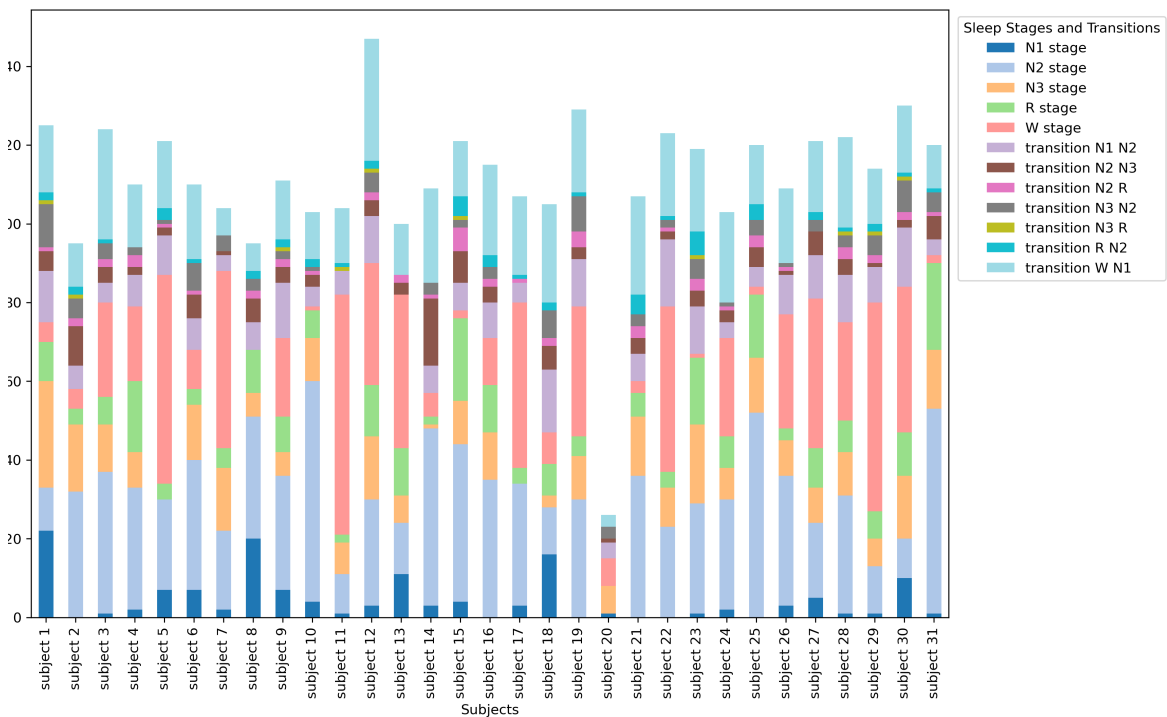
W tej podsekcji przedstawiamy rozkład faz snu i przejść między nimi. Prezentujemy częstość i procentowy udział każdej fazy oraz każdego przejścia, zarówno w odniesieniu do wszystkich zapisów łącznie, jak i w ramach poszczególnych badanych osób. Takie zestawienie pozwala zrozumieć ogólną strukturę i zmienność danych.

Rozkład faz snu oraz przejść między nimi ilustruje Tabela 5.3:

Rozkład faz snu i przejść między nimi w zapisach poszczególnych uczestników badania przedstawiono na Rysunku

Tabela 5.3: Całkowity rozkład faz snu i przejść między nimi w analizowanych zapisach.

	Liczba faz we wszystkich zapisach	Procent wszystkich zapisów
N1 faza	138	4.0%
N2 faza	868	25.17%
N3 faza	328	9.51%
R faza	270	7.83%
W faza	680	19.72%
przejście N1-N2	264	7.65%
przejście N2-N3	127	3.68%
przejście N2-R	55	1.59%
przejście N3-N2	112	3.25%
przejście N3-R	10	0.29%
przejście R-N2	53	1.54%
przejście W-N1	544	15.77%



Rysunek 5.3: Rozkład faz snu i przejść w zapisach poszczególnych osób.

5.3.2 Analiza statystyczna

W tej podsekcji opisano wyniki regresji z wykorzystaniem uogólnionych równań estymujących (GEE, ang. Generalized Estimating Equations) dla różnych przejść między fazami snu. Najlepsze modele wybrano za pomocą selekcji wstecznej, przyjmując kryterium AIC. Zidentyfikowano istotne predyktory oraz oceniono ich wpływ na prawdopodobieństwo konkretnych przejść, co pozwala poznać istotne związki statystyczne między parametrami snu a przejściami między fazami.

Zaprezentowano wyniki dla poszczególnych zmiennych, ale nie podano wartości p dla całych modeli, ponieważ w przypadku GEE zwykle wykorzystuje się QIC lub AIC do porównań modeli [102]. Nie przedstawiono również miar takich jak pseudo- R^2 , gdyż w tym badaniu nie uznano ich za użyteczne.

W dalszej części, omawiając p -wartość lub współczynniki dla danej zmiennej, należy rozumieć, że pozostałe zmienne pozostają niezmiennione. Uznaliśmy, że ciągle powtarzanie tego zastrzeżenia przy każdej wzmiance o zmiennej utrudniłoby klarowność przekazu.

5.3.2.1 Faza N1

Tabela 5.4: Wyniki regresji GEE, faza N1.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-3.5554	0.709	-5.015	0.000	[-4.945, -2.166]
<i>SDNN</i>	-0.0236	0.009	-2.693	0.007	[-0.041, -0.006]
<i>SD1_a</i>	0.0322	0.012	2.682	0.007	[0.009, 0.056]
<i>pnn30_{acc}</i>	0.0246	0.019	1.271	0.204	[-0.013, 0.063]
<i>DR_{MAX}</i>	0.0863	0.053	1.632	0.103	[-0.017, 0.190]
<i>H_{DR}</i>	0.9044	0.730	1.239	0.215	[-0.527, 2.335]
SampEn	-0.4135	0.160	-2.581	0.010	[-0.728, -0.099]

Tabela 5.4 prezentuje wyniki analizy GEE dla zmiennej zależnej binarnej opisującej obecność fazy N1. Jest to najlepszy model wybrany metodą selekcji wstecznej przy użyciu AIC.

Zmienne symetryczne

Analiza zmiennych symetrycznych dla fazy N1 pokazuje, że *SDNN* ma ujemny współczynnik (-0.0236), co oznacza, iż wzrost *SDNN* obniża prawdopodobieństwo zaklasyfikowania danego fragmentu jako faza N1. Efekt ten jest istotny statystycznie ($p = 0.007$). Z kolei SampEn ma także współczynnik ujemny (-0.4135), sugerując, że wyższe wartości SampEn zmniejszają szansę identyfikacji fazy N1; efekt ten również jest istotny ($p = 0.010$).

Zmienne asymetryczne

Wśród zmiennych asymetrycznych, $SD1_a$ ma dodatni współczynnik (0.0322), co oznacza, że wzrost $SD1_a$ zwiększa prawdopodobieństwo wystąpienia fazy N1, przy istotności statystycznej ($p = 0.007$). $pnn30_{acc}$ również ma dodatni współczynnik (0.0246), ale nie jest istotny statystycznie ($p = 0.204$). DR_{MAX} (0.0863) sugeruje wyższe prawdopodobieństwo fazy N1 przy jego wzroście, choć przy $p = 0.103$ nie jest to istotne. Podobnie H_{DR} (0.9044) wskazuje możliwy wzrost prawdopodobieństwa fazy N1 przy wyższych wartościach tego parametru, lecz również nie osiąga istotności ($p = 0.215$).

5.3.2.2 Faza N2

Tabela 5.5: Wyniki regresji GEE, faza N2.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-1.8036	0.579	-3.116	0.002	[-2.938, -0.669]
$SDNN$	-0.0054	0.003	-1.581	0.114	[-0.012, 0.001]
$SD1$	-0.0317	0.021	-1.511	0.131	[-0.073, 0.009]
$SD1_d$	0.0436	0.025	1.719	0.086	[-0.006, 0.093]
$pnn30$	0.0102	0.008	1.317	0.188	[-0.005, 0.025]
$DR4$	-0.1330	0.045	-2.985	0.003	[-0.220, -0.046]
AR_{MAX}	-0.2090	0.061	-3.419	0.001	[-0.329, -0.089]
H_{DR}	1.2553	0.607	2.069	0.039	[0.066, 2.445]
H_{AR}	0.4440	0.161	2.754	0.006	[0.128, 0.760]
SampEn	0.2418	0.137	1.770	0.077	[-0.026, 0.509]
LF/HF	0.0835	0.015	5.464	0.000	[0.054, 0.113]

Tabela 5.5 przedstawia wyniki regresji GEE dla fazy N2 (zmienna binarna) w najlepszym modelu.

Zmienne symetryczne

W przypadku zmiennych symetrycznych, $SDNN$ ma ujemny współczynnik (-0.0054), co oznacza nieznaczny spadek prawdopodobieństwa fazy N2 przy rosnącym $SDNN$, ale nie jest to efekt istotny ($p = 0.114$). Podobnie $SD1$ ma wartość ujemną (-0.0317), także nieistotną ($p = 0.131$). Z kolei $pnn30$ (0.0102) i SampEn (0.2418) mają współczynniki dodatnie, ale oba efekty nie są istotne ($p = 0.188$ i $p = 0.077$). Natomiast LF/HF (0.0835) jest wysoce dodatni i istotny ($p < 0.001$), co wskazuje, że wyższe wartości LF/HF silnie zwiększają prawdopodobieństwo fazy N2.

Zmienne asymetryczne

Wśród zmiennych asymetrycznych, $SD1_d$ (0.0436) ma dodatni wpływ, lecz przy $p = 0.086$ nie jest istotny. Natomiast $DR4$ (-0.1330) i AR_{MAX} (-0.2090) mają ujemne współczynniki i oba są istotne ($p = 0.003$ i $p = 0.001$), co oznacza, że wzrost tych parametrów zmniejsza prawdopodobieństwo fazy N2. Z kolei H_{DR} (1.2553) i H_{AR} (0.4440) działają dodatnio i w obu przypadkach efekt jest istotny ($p = 0.039$ i $p = 0.006$).

5.3.2.3 Przejście N1–N2

Tabela 5.6: Wyniki regresji GEE, przejście N1-N2.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-3.1716	0.218	-14.542	0.000	[-3.599, -2.744]
$SDNN$	0.0088	0.004	2.429	0.015	[0.002, 0.016]
$SD1_d$	-0.0067	0.004	-1.651	0.099	[-0.015, 0.001]
H_{AR}	0.4436	0.197	2.257	0.024	[0.058, 0.829]

Tabela 5.6 pokazuje wyniki regresji GEE dla odpowiedzi binarnej opisującej przejście z fazy N1 do N2.

Zmienne symetryczne

Dla zmiennych symetrycznych, $SDNN$ (0.0088) ma dodatni i istotny statystycznie wpływ ($p = 0.015$), co oznacza, że wyższe $SDNN$ sprzyja przejściu z N1 do N2.

Zmienne asymetryczne

W przypadku zmiennych asymetrycznych, $SD1_d$ (-0.0067) ma współczynnik ujemny, sugerując, że rosnąca wartość $SD1_d$ zmniejsza prawdopodobieństwo przejścia N1-N2, choć efekt nie jest istotny ($p = 0.099$). Natomiast H_{AR} (0.4436) wykazuje dodatni i istotny wpływ ($p = 0.024$), wskazując na wyższe prawdopodobieństwo przejścia N1-N2 przy wzroście H_{AR} .

5.3.2.4 Przejście W–N1

Tabela 5.7 przedstawia wyniki regresji GEE dla zmiennej binarnej opisującej przejście z czuwania (W) do fazy N1.

Zmienne symetryczne

Wśród zmiennych symetrycznych, $SDNN$ (0.0135) wpływa dodatnio i istotnie ($p < 0.001$), co oznacza, że wyższe wartości $SDNN$ zwiększają prawdopodobieństwo przejścia W-N1. $SD1$ (-0.4619) ma współczynnik ujemny, lecz nieistotny ($p = 0.087$).

Zmienne asymetryczne

Tabela 5.7: Wyniki regresji GEE, przejście W-N1.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-2.7905	0.331	-8.433	0.000	[-3.439, -2.142]
<i>SDNN</i>	0.0135	0.003	4.097	0.000	[0.007, 0.020]
<i>SD1</i>	-0.4619	0.270	-1.710	0.087	[-0.991, 0.067]
<i>SD1_a</i>	0.3156	0.182	1.733	0.083	[-0.041, 0.672]
<i>SD1_d</i>	0.3268	0.200	1.637	0.102	[-0.065, 0.718]
<i>AR_{MAX}</i>	0.0632	0.032	1.951	0.051	[-0.000, 0.127]
<i>H_{AR}</i>	0.3915	0.246	1.592	0.111	[-0.090, 0.873]

Dla zmiennych asymetrycznych, *SD1_a* (0.3156) i *SD1_d* (0.3268) mają dodatnie współczynniki, wskazując na wzrost prawdopodobieństwa przejścia W-N1, jednak oba efekty nie są istotne ($p = 0.083$ i $p = 0.102$). *AR_{MAX}* (0.0632) także działa dodatnio i jest bliskie istotności ($p = 0.051$). *H_{AR}* (0.3915) jest dodatnie, lecz nieistotne ($p = 0.111$).

5.3.2.5 Przejście N2-N3

Tabela 5.8: Wyniki regresji GEE, przejście N2-N3.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-3.5729	0.476	-7.502	0.000	[-4.506, -2.639]
<i>SDNN</i>	-0.0150	0.005	-3.103	0.002	[-0.025, -0.006]
<i>SD1</i>	0.0164	0.004	3.897	0.000	[0.008, 0.025]
<i>AR4</i>	-0.1293	0.069	-1.877	0.061	[-0.264, 0.006]
<i>DR4</i>	-0.0966	0.062	-1.568	0.117	[-0.217, 0.024]
SampEn	0.4543	0.205	2.216	0.027	[0.053, 0.856]
<i>LF/HF</i>	0.1009	0.008	12.789	0.000	[0.085, 0.116]

Zmienne symetryczne

Analiza zmiennych symetrycznych dla przejścia N2-N3 wskazuje, że *SDNN* (-0.0150) istotnie zmniejsza prawdopodobieństwo przejścia ($p=0.002$). *SD1* (0.0164) działa dodatnio i istotnie ($p < 0.001$). SampEn (0.4543) też jest dodatnie i istotne ($p = 0.027$). *LF/HF* (0.1009) silnie zwiększa prawdopodobieństwo przejścia N2-N3 i jest wysoce istotne ($p < 0.001$).

Zmienne asymetryczne

W przypadku zmiennych asymetrycznych, $AR4$ (-0.1293) i $DR4$ (-0.0966) mają ujemne współczynniki, co sugeruje spadek prawdopodobieństwa przejścia N2-N3 przy ich wzroście, jednak nie są istotne ($p = 0.061$ i $p = 0.117$).

5.3.2.6 Faza N3

Tabela 5.9: Wyniki regresji GEE, faza N3.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-1.0665	0.764	-1.396	0.163	[-2.564, 0.431]
$SDNN$	-0.0092	0.008	-1.182	0.237	[-0.025, 0.006]
$SD1$	0.0060	0.006	0.962	0.336	[-0.006, 0.018]
$DR4$	-0.1895	0.088	-2.146	0.032	[-0.363, -0.016]
AR_{MAX}	-0.1666	0.082	-2.023	0.043	[-0.328, -0.005]
H_{AR}	-0.4781	0.399	-1.197	0.231	[-1.261, 0.305]
SampEn	0.4975	0.185	2.692	0.007	[0.135, 0.860]

Zmienne symetryczne

Dla fazy N3, $SDNN$ (-0.0092) sugeruje nieistotny spadek prawdopodobieństwa ($p=0.237$). $SD1$ (0.0060) wskazuje nieistotny wzrost ($p=0.336$). Natomiast SampEn (0.4975) jest dodatnie i istotne ($p=0.007$).

Zmienne asymetryczne

Wśród zmiennych asymetrycznych, $DR4$ (-0.1895) i AR_{MAX} (-0.1666) zmniejszają prawdopodobieństwo fazy N3, przy czym oba są istotne ($p = 0.032$ i $p = 0.043$). H_{AR} (-0.4781) również jest ujemne, lecz nieistotne ($p = 0.231$).

5.3.2.7 Przejście N3-N2

Tabela 5.10: Wyniki regresji GEE, przejście N3-N2.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-3.4785	0.209	-16.609	0.000	[-3.889, -3.068]
$SD1$	-0.0240	0.012	-2.050	0.040	[-0.047, -0.001]
$SD1_a$	0.0348	0.016	2.130	0.033	[0.003, 0.067]
$pnn30_{acc}$	0.0184	0.012	1.487	0.137	[-0.006, 0.043]

Zmienne symetryczne

Analiza dla przejścia N3-N2 pokazuje, że $SD1$ (-0.0240) zmniejsza prawdopodobieństwo przejścia i jest istotny ($p = 0.040$).

Zmienne asymetryczne

$SD1_a$ (0.0348) ma dodatni wpływ na przejście N3-N2 i jest istotny ($p = 0.033$). $pnn30_{acc}$ (0.0184) wskazuje dodatni, lecz nieistotny efekt ($p = 0.137$).

5.3.2.8 Przejście N2-R

Tabela 5.11: Wyniki regresji GEE, przejście N2-R.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-5.3598	0.846	-6.333	0.000	[-7.019, -3.701]
AR_{MAX}	-0.2430	0.153	-1.584	0.113	[-0.544, 0.058]
DR_{MAX}	0.2123	0.136	1.564	0.118	[-0.054, 0.478]
H_{DR}	2.4855	1.083	2.295	0.022	[0.363, 4.608]
H_{AR}	-1.5457	0.860	-1.798	0.072	[-3.231, 0.139]
SampEn	0.3578	0.186	1.923	0.055	[-0.007, 0.723]

Zmienne symetryczne

SampEn (0.3578) ma dodatni wpływ na przejście z N2 do fazy R, chociaż graniczny ($p=0.055$).

Zmienne asymetryczne

AR_{MAX} (-0.2430) zmniejsza prawdopodobieństwo N2-R ($p=0.113$), DR_{MAX} (0.2123) wskazuje wyższe prawdopodobieństwo, lecz nieistotne ($p = 0.118$). Natomiast H_{DR} (2.4855) jest wyraźnie dodatnie i istotne ($p = 0.022$). H_{AR} (-1.5457) jest ujemne i bliskie istotności ($p = 0.072$).

5.3.2.9 Przejście R-N2

Tabela 5.12: Wyniki regresji GEE, przejście R-N2.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-5.8863	0.458	-12.862	0.000	[-6.783, -4.989]
$SD1$	-0.0584	0.046	-1.276	0.202	[-0.148, 0.031]
$SD1_d$	0.0775	0.058	1.332	0.183	[-0.037, 0.191]
AR_{MAX}	-0.2505	0.110	-2.277	0.023	[-0.466, -0.035]
DR_{MAX}	0.1279	0.052	2.450	0.014	[0.026, 0.230]
H_{DR}	2.4818	0.595	4.169	0.000	[1.315, 3.649]

Zmienne symetryczne

Dla przejścia z fazy R do N2, $SD1$ (-0.0584) jest nieistotnie ujemne ($p=0.202$).

Zmienne asymetryczne

AR_{MAX} (-0.2505) obniża prawdopodobieństwo przejścia ($p=0.023$), natomiast DR_{MAX} (0.1279) je podnosi ($p=0.014$). H_{DR} (2.4818) ma silny dodatni i wysoce istotny wpływ ($p<0.001$). $SD1_d$ (0.0775) jest dodatnie, lecz nieistotne ($p=0.183$).

5.3.2.10 Faza R

Tabela 5.13: Wyniki regresji GEE, faza R.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-2.9060	0.349	-8.337	0.000	[-3.589, -2.223]
$SDNN$	0.0165	0.004	4.694	0.000	[0.010, 0.023]
$SD1$	-0.0785	0.043	-1.827	0.068	[-0.163, 0.006]
$SD1_d$	0.0887	0.054	1.649	0.099	[-0.017, 0.194]
$pnn30$	-0.0119	0.012	-0.964	0.335	[-0.036, 0.012]
SampEn	0.1863	0.148	1.256	0.209	[-0.104, 0.477]

Zmienne symetryczne

W fazie R, $SDNN$ (0.0165) jest dodatnie i istotne ($p<0.001$). $SD1$ (-0.0785) nie osiąga istotności ($p=0.068$). SampEn (0.1863) i $pnn30$ (-0.0119) nie są istotne.

Zmienne asymetryczne

$SD1_d$ (0.0887) ma dodatni współczynnik, ale nie jest istotny ($p=0.099$).

5.3.2.11 Przejście N3-R

Tabela 5.14: Wyniki regresji GEE, przejście N3-R.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	-6.9436	1.091	-6.364	0.000	[-9.082, -4.805]
AR_{MAX}	-0.4793	0.321	-1.493	0.136	[-1.109, 0.150]
H_{DR}	3.7051	2.091	1.772	0.076	[-0.394, 7.804]

Zmienne asymetryczne

Dla przejścia N3-R, AR_{MAX} (-0.4793) ma ujemny współczynnik, ale nie jest istotny ($p=0.136$). H_{DR} (3.7051) sugeruje dodatni wpływ, bliski istotności ($p=0.076$).

5.3.2.12 Faza W (czuwanie)

Tabela 5.15: Wyniki regresji GEE, faza W.

	coef	std err	z	$P > z $	[0.025, 0.975]
Intercept	0.2604	0.823	0.316	0.752	[-1.353, 1.874]
$SDNN$	0.0040	0.005	0.792	0.429	[-0.006, 0.014]
$SD1$	-0.0352	0.020	-1.753	0.080	[-0.075, 0.004]
$SD1_a$	0.0506	0.028	1.804	0.071	[-0.004, 0.106]
$pnn30_{dec}$	-0.0351	0.021	-1.679	0.093	[-0.076, 0.006]
$DR4$	0.1820	0.054	3.373	0.001	[0.076, 0.288]
AR_{MAX}	0.2974	0.082	3.639	0.000	[0.137, 0.458]
H_{DR}	-1.9586	0.663	-2.955	0.003	[-3.258, -0.660]
H_{AR}	-0.4702	0.266	-1.768	0.077	[-0.991, 0.051]
SampEn	-0.2381	0.191	-1.248	0.212	[-0.612, 0.136]
LF/HF	-0.1537	0.047	-3.291	0.001	[-0.245, -0.062]

Zmienne symetryczne

Dla fazy W, $SDNN$ (0.0040) jest dodatnie, lecz nieistotne ($p=0.429$). $SD1$ (-0.0352) sugeruje spadek prawdopodobieństwa ($p=0.080$), lecz jest bliskie istotności. SampEn (-0.2381) i LF/HF (-0.1537) mają ujemne współczynniki, przy czym LF/HF jest istotne ($p=0.001$).

Zmienne asymetryczne

W przypadku zmiennych asymetrycznych, $SD1_a$ (0.0506) wykazuje dodatni wpływ bliski istotności ($p=0.071$), a $pnn30_{dec}$ (-0.0351) ujemny ($p=0.093$). $DR4$ (0.1820) i AR_{MAX} (0.2974) są dodatnie i istotne ($p = 0.001$ i $p < 0.001$). Z kolei H_{DR} (-1.9586) jest ujemne i istotne ($p = 0.003$), a H_{AR} (-0.4702) nieistotne ($p=0.077$).

5.3.3 Wyniki modelowania predykcyjnego

W tej części oceniono skuteczność modeli predykcyjnych dla różnych faz snu i przejść między nimi. Zaprezentowano macierze pomyłek oraz miary jakości (AUC, dokładność, precyzja, czułość, F1 score) dla każdej z faz i przejść, wskazując zarówno skuteczność, jak i obszary wymagające ulepszeń. Modele, jak zaznaczono w sekcji Materiały i Metody, to te same, które

zostały opisane w poprzedniej części, przy czym prognozy obliczono dla zbioru testowego.

5.3.3.1 Faza N1

Tabela 5.16: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy N1.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	407	218
	Pozytywne	105	270

Miary Skuteczności	AUC:	0.74
	Dokładność:	0.68
	Precyzja:	0.55
	Czułość:	0.72
	F1 Score:	0.63

Model uzyskuje $AUC = 0.74$, co oznacza umiarkowaną zdolność do rozróżniania fazy N1. Czułość wynosi 72%, świadcząc o efektywnym wykrywaniu dodatnich przypadków N1. Precyzja na poziomie 55% wskazuje na pewną liczbę fałszywych alarmów (inne fazy uznane za N1). Dokładność całkowita to 68%, a wynik F1 równy 0.63 obrazuje kompromis między precyzją a czułością.

5.3.3.2 Faza N2

Tabela 5.17: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy N2.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	365	113
	Pozytywne	113	409

Miary Skuteczności	AUC:	0.85
	Dokładność:	0.77
	Precyzja:	0.78
	Czułość:	0.78
	F1 Score:	0.78

Model osiąga $AUC = 0.85$, co świadczy o dobrej zdolności do rozróżniania fazy N2. Czułość i precyzja wynoszą 78%, co oznacza zrównoważoną identyfikację przypadków pozytywnych i stosunkowo niewielką liczbę fałszywych alarmów. Dokładność 77% również jest satysfakcjonująca, a F1 score na poziomie 78% odzwierciedla zbalansowaną wydajność modelu.

5.3.3.3 Faza N3

Tabela 5.18: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy N3.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	433	107
	Pozytywne	86	374

Miary Skuteczności	AUC:	0.89
	Dokładność:	0.81
	Precyzja:	0.78
	Czułość:	0.81
	F1 Score:	0.79

Model osiąga $AUC = 0.89$, co wskazuje na wysoką zdolność odróżniania fazy N3 od innych. Czułość 81% i precyzja 78% stanowią dobrą równowagę, przy dokładności ogólnej 81%. F1 score = 79% potwierdza efektywną równowagę między precyzją a czułością.

5.3.3.4 Faza R

Tabela 5.19: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy R.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	277	233
	Pozytywne	178	312

Miary Skuteczności	AUC:	0.61
	Dokładność:	0.59
	Precyzja:	0.57
	Czułość:	0.64
	F1 Score:	0.60

Model uzyskuje $AUC = 0.61$, wskazując umiarkowaną zdolność do odróżniania fazy R. Czułość na poziomie 64% oznacza średnią skuteczność w identyfikacji rzeczywistych przypadków R. Precyzja 57% sugeruje istotną liczbę fałszywych alarmów, a dokładność 59% jest relatywnie niska. F1 score to 60%.

5.3.3.5 Faza W

Tabela 5.20: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy W.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	453	107
	Pozytywne	95	345

Miary Skuteczności	AUC:	0.87
	Dokładność:	0.80
	Precyzja:	0.76
	Czułość:	0.78
	F1 Score:	0.77

AUC na poziomie 0.87 świadczy o wysokiej zdolności do odróżniania fazy W. Czułość 78% i precyzja 76% pozwalają na dobrą identyfikację prawidłowych przypadków W oraz ograniczoną liczbę fałszywych pozytywów. Dokładność 80% jest wysoka, a F1 score wynosi 77%.

5.3.3.6 Przejście N2-N3

Tabela 5.21: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N2-N3.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	498	222
	Pozytywne	66	214

Miary Skuteczności	AUC:	0.80
	Dokładność:	0.71
	Precyzja:	0.49
	Czułość:	0.76
	F1 Score:	0.60

Model osiąga $AUC = 0.80$, co wskazuje na dużą zdolność odróżniania przejścia N2-N3. Czułość 76% jest wysoka, ale precyzja 49% ujawnia znaczną liczbę fałszywych alarmów, skutkując dokładnością 71%. F1 score = 60%.

5.3.3.7 Przejście N2-R

Tabela 5.22: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N2-R.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	297	120
	Pozytywne	116	467

Miary Skuteczności	AUC:	0.85
	Dokładność:	0.76
	Precyzja:	0.80
	Czułość:	0.80
	F1 Score:	0.80

Model osiąga $AUC = 0.85$, co wskazuje na wysoką zdolność do wykrywania przejścia N2-R. Zarówno precyzja, jak i czułość wynoszą 80%, przekładając się na dokładność 76% i F1 score = 80%.

5.3.3.8 Przejście N3-N2

W tym przypadku zbiór testowy zawierał zbyt mało obserwacji, dlatego wyniki poniżej dotyczą grupy treningowej.

Tabela 5.23: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N3-N2.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	375	254
	Pozytywne	145	226

Miary Skuteczności	AUC:	0.63
	Dokładność:	0.60
	Precyzja:	0.47
	Czułość:	0.61
	F1 Score:	0.53

Model wykazuje czułość 61% dla przejścia N3-N2, co stanowi średni poziom wykrywania przypadków pozytywnych. Precyzja 47% sugeruje znaczącą liczbę fałszywych alarmów, a dokładność 60% pozostawia przestrzeń do poprawy. F1 score wynosi 53%.

5.3.3.9 Przejście N3-R

Tabela 5.24: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N3-R.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	272	163
	Pozytywne	167	398

Miary Skuteczności	AUC:	0.72
	Dokładność:	0.67
	Precyzja:	0.71
	Czułość:	0.70
	F1 Score:	0.71

Z $AUC = 0.72$, model dysponuje umiarkowaną zdolnością do rozróżniania przejścia N3-R. Czułość 70% i precyzja 71% dają zbalansowaną klasyfikację, przy dokładności 67%. F1 score wynosi 71%.

5.3.3.10 Przejście R-N2

Tabela 5.25: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia R-N2.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	515	229
	Pozytywne	55	201

Miary Skuteczności	AUC:	0.80
	Dokładność:	0.72
	Precyzja:	0.47
	Czułość:	0.79
	F1 Score:	0.59

Model osiąga $AUC = 0.80$, czyli dobrą zdolność do rozróżniania przejścia R-N2. Czułość 79% wskazuje skuteczne rozpoznawanie przypadków pozytywnych, lecz precyzja 47% oznacza sporą liczbę fałszywych alarmów. Dokładność to 72%, a F1 score = 59%.

5.3.3.11 Przejście W-N1

Tabela 5.26: Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia W-N1.

Macierz Pomyłek			
		Przewidywanie	
		Negatywne	Pozytywne
Rzeczywistość	Negatywne	535	210
	Pozytywne	52	203

Miary Skuteczności	AUC:	0.83
	Dokładność:	0.74
	Precyzja:	0.49
	Czułość:	0.80
	F1 Score:	0.61

Model uzyskuje $AUC = 0.83$, co oznacza wyraźną zdolność do rozróżniania przejścia W-N1. Czułość 80% dobrze wychwytuje pozytywne przypadki, ale precyzja 49% wskazuje liczne fałszywe alarmy. Dokładność 74% jest dość wysoka, a F1 score wynosi 61%.

5.3.4 Dyskusja

Niniejsze badanie dotyczyło możliwości wzmocnienia tradycyjnych, symetrycznych miar zmienności i złożoności, wyprowadzanych z analizy zmienności rytmu serca i złożoności, poprzez uwzględnienie asymetrycznych miar zmienności i złożoności. Celem było wykrywanie i rozróżnianie różnych faz snu. Należy podkreślić, że w pracy tej nie analizowano jednego zintegrowanego modelu, lecz osobne modele dla każdej fazy i każdego przejścia, tak więc część statystyczna ma charakter wyłącznie eksploracyjny. Uzyskane wyniki wskazują na skuteczność łączenia metryk HRA z tradycyjnymi miarami HRV i entropii w precyzyjnej detekcji faz snu oraz przejść między nimi. We wszystkich przypadkach zmienne asymetryczne zachowywały niezależną wartość predykcyjną w wykrywaniu faz snu i przejść między nimi. Zmienne, które najczęściej były zachowywane w procesie selekcji wstecznej, to asymetryczne entropie H_{DR} i H_{AR} . Wydajność modeli jako narzędzi predykcyjnych

również okazała się dobra, biorąc pod uwagę bardzo zróżnicowany charakter grupy badawczej, obejmującej osoby w różnym wieku i o różnym stanie zdrowia.

W niniejszej pracy wykorzystano trzy klasyczne zmienne HRV: *SDNN*, *SD1* oraz *LF/HF*. Dodatkowo uwzględniono jeszcze jedną zmienną symetryczną z tej grupy, *pnn30*, która jest blisko spokrewniona z popularnym deskryptorem *pnn50* [110].

Spośród tych zmiennych najlepsze wyniki w kilku modelach uzyskały *SDNN* oraz *LF/HF*. Dla *SDNN* zaobserwowano istotny dodatni wpływ w przypadku przejść N1–N2 oraz W–N1, a także dla fazy R. W przejściu N1–N2 wzrost *SDNN* zwiększa prawdopodobieństwo przejścia z N1 do N2. Podobnie, w przejściu W–N1 wyższe wartości *SDNN* zwiększają prawdopodobieństwo przejścia ze stanu czuwania do fazy N1. Z kolei w fazie R wyższe wartości *SDNN* zwiększają prawdopodobieństwo detekcji snu REM.

W przypadku *LF/HF* najlepszy wynik uzyskano w przejściu N2–N3, gdzie wzrost *LF/HF* zwiększa prawdopodobieństwo przejścia z N2 do N3. Dodatkowo, w fazie N2 wyższe wartości *LF/HF* zwiększają prawdopodobieństwo wykrycia tej fazy.

Nasze rezultaty potwierdzają znaczenie HRV w detekcji faz snu, co jest zgodne z wcześniejszymi badaniami. Na przykład, zwiększona składowa wysokoczęstotliwościowa (HF) i obniżona składowa niskoczęstotliwościowa (LF) w fazach snu non-REM odzwierciedlają wzmożoną aktywność przywspółczulną, opisywaną w [117] i [116]. Z kolei podwyższony stosunek LF/HF w fazie REM wskazuje na wzrost aktywności współczulnej, co potwierdza się w badaniach [118].

Z kolei SampEn przyniosła bardzo dobre rezultaty w wykrywaniu kilku faz i przejść. Co ciekawe, wywiera ona zarówno istotny ujemny, jak i dodatni wpływ w zależności od rozpatrywanego modelu.

W przypadku fazy N1 spadek SampEn zwiększa prawdopodobieństwo wykrycia tej fazy, co sugeruje, że niższe wartości SampEn są powiązane z wyższym prawdopodobieństwem detekcji snu N1.

W przejściu N2–N3 wzrost SampEn istotnie zwiększa prawdopodobieństwo przejścia z N2 do N3.

Podobnie, w fazie N3 wyższe wartości SampEn zwiększają prawdopodobieństwo wykrycia tej fazy.

Choć SampEn pojawia się również w innych modelach, takich jak przejście N2–R czy faza R, wyniki te nie mają istotnego charakteru statystycznego i nie wnoszą znaczącego wkładu w detekcję tych faz czy przejść.

Silny związek SampEn z głębszymi fazami snu, widoczny w naszych wynikach, jest spójny z wcześniejszymi pracami [122, 121]. Spadek SampEn w przejściach do głębszych faz snu odzwierciedla mniejszą złożoność i wyższą

regularność sygnałów fizjologicznych, wzmacniając ideę stabilnego stanu snu – co dodatkowo potwierdza wyniki opisane w [122, 121].

Nasze wyniki wskazują, że HRA, zwłaszcza w ujęciu złożoności i deskryptorów bazujących na seriach, znacząco poprawia własności statystyczne modeli w analizie snu, a zarazem zwiększa ich właściwości predykcyjne. Szczególnie deskryptory HRA, takie jak H_{DR} i H_{AR} , pełniły kluczową rolę w usprawnianiu działania modeli, zwłaszcza w detekcji niektórych faz snu i przejść pomiędzy nimi.

Spośród analizowanych zmiennych asymetrycznych, H_{DR} znacząco zwiększał prawdopodobieństwo zidentyfikowania fazy N2 i miał istotny wkład w przejścia N2–R oraz R–N2. Również H_{AR} okazał się szczególnie skuteczny w podnoszeniu prawdopodobieństwa fazy N2 oraz w przejściu z N1 do N2. Oprócz H_{DR} i H_{AR} , zmienna AR_{MAX} silnie wpływała na identyfikację faz N2 i W, obniżając prawdopodobieństwo fazy N2 przy jednoczesnym zwiększeniu prawdopodobieństwa fazy W. Ponadto, AR_{MAX} miał istotny wpływ na przejście R–N2, gdzie jego obniżone wartości redukowały prawdopodobieństwo przejścia z REM do N2. Z kolei spadek $DR4$ był najbardziej efektywny w zmniejszaniu prawdopodobieństwa zarówno fazy N2, jak i N3.

Warto odnotować, że nawet w sytuacjach, gdzie zmienne asymetryczne nie prowadziły bezpośrednio do lepszych predykcji (np. ze względu na brak istotności statystycznej), stabilizowały modele i zmniejszały wartość kryterium AIC, co dowodzi ich przydatności w modelowaniu predykcyjnym. Szczególnie parametry bazujące na seriach okazały się najskuteczniejsze w prognozowaniu i rozróżnianiu faz snu oraz przejść, oferując szczegółowość niedostępną w przypadku miar symetrycznych.

Zróznicowanie faz snu w oparciu o asymetryczne wpływy przyspieszeń (accelerations) i zwolnień (decelerations) czynności serca jest zgodne z wcześniejszymi badaniami Guzika i wsp. [127] oraz Apnea i wsp. [128]. Zmniejszona liczba krótkich ciągów deceleracji i akceleracji u pacjentów z ciężkim obturacyjnym bezdechem sennym (OSA), a także zwiększona liczba dłuższych ciągów, ilustrują wpływ zaburzeń snu na dynamikę rytmu serca. Wyniki te korespondują z ustaleniami Herzig i wsp. [117] oraz Tsunoda i wsp. [118], którzy również zaobserwowali zmienione wzorce HRV u osób z zaburzeniami snu.

Spośród wszystkich rozpatrywanych modeli sześć wyróżnia się szczególnie wysoką skutecznością i przydatnością. Obejmują one trzy fazy snu oraz trzy przejścia. Zarówno w symetrycznych, jak i asymetrycznych parametrach tych modeli można dostrzec pewną charakterystyczną strukturę.

Najlepszą wydajność wśród modeli klasyfikujących fazy snu osiągają modele dotyczące głębszych faz – N2 i N3 – oraz model fazy czuwania (W). Dla tych faz wartości AUC wynoszą odpowiednio 0,85, 0,89 i 0,87, a

inne miary skuteczności, takie jak trafność (77%, 81% i 80%) oraz czułość (78%, 81% i 78%), również prezentują się bardzo dobrze. Analiza struktury statystycznej tych modeli prowadzi do wniosku, że w każdym z nich zarówno parametry symetryczne, jak i asymetryczne odgrywają kluczową rolę. We wszystkich istotne znaczenie mają miary złożoności – zarówno symetryczna (SampEn), jak i asymetryczne (H_{DR} , H_{AR}).

Najlepsze rezultaty wśród modeli dotyczących przejść uzyskano dla przejść z N2 do REM i z powrotem (N2–R, R–N2) oraz dla momentu zasypiania (W–N1). Wartości AUC dla tych przejść wynoszą odpowiednio 0,85, 0,80 i 0,83, a pozostałe metryki, takie jak trafność (76%, 72%, 74%) i czułość (80%, 79%, 80%) również są wysokie. Struktura tych modeli statystycznych wygląda jednak inaczej. We wszystkich trzech modelach dominują parametry asymetryczne, a jako jedyne zmienne symetryczne pojawiają się: SampEn w modelu N2–R, $SD1$ w modelu R–N2 oraz $SDNN$ w modelu W–N1. Wszystkie te modele zawierają AR_{MAX} i co najmniej jedną z miar H_{AR} lub H_{DR} .

W niniejszej pracy zaproponowano użycie metody uogólnionych równań estymujących (GEE) w badaniach nad snem, uznając ją za szczególnie dobrze dopasowaną do analizy danych o charakterze klastrowym, typowych dla badań snu, gdzie każdy badany stanowi odrębny klastrowy. Zbudowane w oparciu o GEE najlepsze modele statystyczne zastosowano następnie w analizach typu machine learning. Dzięki uwzględnieniu korelacji wewnątrz każdego klastra (pacjenta), GEE czyni zarówno wnioskowanie statystyczne, jak i metody machine learning bardziej przejrzystymi i łatwiejszymi do interpretacji. Wyniki uzyskane przy pomocy GEE w naszym badaniu są zgodne z rezultatami innych autorów, korzystających z odmiennych podejść statystycznych bądź metod uczenia maszynowego, co dodatkowo potwierdza trafność naszej propozycji. Uważamy, że GEE stanowi optymalne połączenie wnioskowania statystycznego i machine learningu w kontekście badań nad fazami snu i przejściami między nimi.

Zintegrowanie metody Heart Rate Asymmetry (HRA) z tradycyjnymi miarami, takimi jak HRV i entropia, w kontekście klasyfikacji faz snu niesie ze sobą istotne konsekwencje kliniczne i naukowe. Na podstawie przedstawionych wyników można stwierdzić, że HRA dostarcza informacji o fazach snu, które są niezależne od i jednocześnie uzupełniające wobec uznanych metryk, takich jak $SDNN$, stosunek LF/HF czy SampEn. Widać to wyraźnie na przykładzie wielowymiarowych modeli statystycznych zaprezentowanych w niniejszej pracy.

Kwestia praktycznego wykorzystania HRA w urządzeniach przenośnych i mobilnych do analizy snu jest bardzo ważna. Parametry HRA można pozyskać z najłatwiej dostępnych i najtańszych w rejestracji sygnałów, np.

z EKG lub z szeregów interwałów *RR*. Pozwala to na bieżące gromadzenie danych bez konieczności zapisywania pełnego sygnału EKG.

Fazy snu i przejścia, takie jak głębsze fazy snu (N2 i N3), moment zasypiania (przejście W–N1) czy początek i koniec snu REM (N2–R i R–N2), mają istotne znaczenie zarówno w zastosowaniach klinicznych, jak i w badaniach naukowych, inżynierskich oraz w zakresie mobilnego monitorowania zdrowia. Modele opracowane dla wymienionych faz i przejść okazały się najlepsze ze wszystkich opisanych w niniejszej pracy. Można je potencjalnie zastosować w praktyce klinicznej w celu ulepszenia diagnostyki zaburzeń snu, w badaniach naukowych dla uzyskania dokładniejszych danych na temat faz snu oraz w inżynierii, przy opracowywaniu przystępnych cenowo, mobilnych i przenośnych systemów monitorowania snu.

5.3.5 Ograniczenia badania

Mimo że niniejsza praca wnosi nowe informacje w zakresie analizy faz snu i przejść między nimi, należy uwzględnić kilka istotnych ograniczeń.

Najważniejszym z nich jest heterogeniczność badanej populacji. Uczestnikami byli pacjenci w różnym wieku i z różnymi jednostkami chorobowymi. Obecność wielu współistniejących schorzeń może wprowadzać dodatkową zmienność w dynamice rytmu serca i potencjalnie zaburzać interpretację wyników.

Dotyczy to w szczególności takich chorób jak nadciśnienie tętnicze czy obturacyjny bezdech senny (OSA), które w istotny sposób wpływają na równowagę autonomiczną i zmienność rytmu serca [139, 127, 140]. Nadciśnienie wiąże się z zaburzoną równowagą autonomiczną i obniżoną zmiennością rytmu serca, natomiast OSA – z okresowymi epizodami hipoksji i wzmożoną aktywnością współczulną, wpływając na dynamikę rytmu serca i asymetrię rytmu serca [141, 127]. Włączenie w jedną grupę pacjentów z nadciśnieniem i z zaburzeniami oddychania podczas snu może więc maskować lub wzmacniać pewne charakterystyczne cechy związane z HRV i HRA, typowe dla określonych faz i przejść snu. Dodatkowe schorzenia, takie jak cukrzyca, przebyte udary czy otyłość, także mogą wpływać na autonomiczną regulację układu krążenia.

W przyszłych badaniach warto rozważyć skupienie się wyłącznie na osobach zdrowych lub bardziej jednorodnych grupach pacjentów w celu lepszego wyizolowania efektów związanych z asymetrią rytmu serca w fazach snu, bez zakłóceń wprowadzanych przez liczne choroby współistniejące. Pozwoli to stwierdzić, na ile zaobserwowane w niniejszej pracy związki są zależne od samych faz snu, a w jakim stopniu wynikają z kondycji zdrowotnej badanych.

Kolejnym ograniczeniem jest stosunkowo niewielka liczba uczestników (31 nagrań polisomnograficznych – klastrów), chociaż łącznie dostarczyły one wielu próbek faz snu. Niewielka liczebność próby może jednak ograniczać możliwość uogólnienia wyników, gdyż w większych populacjach mogą pojawić się bardziej zróżnicowane wzorce zmienności.

Ponadto, choć modele oparte na GEE zapewniły solidne oszacowania, natura korelacji w obrębie poszczególnych nagrań pozostaje nadal złożona i nie do końca zrozumiała. W dalszych badaniach warto przeanalizować różne struktury korelacji oraz ich wpływ na jakość modeli. Głębsze zrozumienie tych wewnętrznych zależności może jeszcze bardziej poprawić dokładność i użyteczność modeli GEE w przyszłych pracach.

Pomimo wskazanych ograniczeń uważamy, że nasze wyniki stanowią wartościowy wkład w badania nad snem, choć należy je interpretować z pewną ostrożnością. Konieczne są dalsze prace z większymi i bardziej jednorodnymi grupami uczestników.

Niniejsza praca ukazuje przydatność analizy Heart Rate Asymmetry (HRA) w klasyfikacji faz snu oraz wykrywaniu przejść między tymi fazami, potwierdzoną wysokimi wartościami metryk efektywności dla kluczowych faz i przejść. Włączanie parametrów HRA umożliwiło uzyskanie wysokich poziomów dokładności, czułości i pola pod krzywą ROC (AUC) zarówno w modelach statystycznych, jak i predykcyjnych.

W kontekście klasyfikacji faz snu modele z wykorzystaniem H_{DR} i H_{AR} osiągnęły wartość AUC na poziomie 0,85, czułość 78% i trafność 77% dla fazy N2, a także AUC równe 0,89, trafność 81% i czułość 81% dla fazy N3. Model wykrywający stan czuwania (W) charakteryzował się AUC=0,87, trafnością 80% i czułością 78%. W strukturze modeli statystycznych dla tych faz zarówno parametry symetryczne, jak i asymetryczne odgrywały równorzędną rolę.

W przypadku przejść modele uwzględniające H_{DR} cechowały się wysoką skutecznością w przejściach N2–R i R–N2, uzyskując AUC=0,85, trafność 76% i czułość 80% (N2–R) oraz AUC=0,80, trafność 72% i czułość 79% (R–N2). Model wykrywający przejście W–N1, czyli moment zasypiania, osiągnął natomiast AUC=0,83, trafność 77% i czułość 80%. W strukturze modeli statystycznych dla przejść zaznacza się wyraźna dominacja zmiennych asymetrycznych.

Uzyskane wyniki sugerują, że miary HRA mogą stanowić cenne narzędzie w klasyfikacji faz snu i detekcji przejść, uzupełniając metody tradycyjnej analizy zmienności rytmu serca. Podejście to może przyczynić się do doskonalenia diagnostyki zaburzeń snu, wspierać badania naukowe dzięki precyzyjniejszej analizie przebiegu snu oraz znaleźć zastosowanie w opracowywaniu przenośnych, ekonomicznych urządzeń do monitorowania snu.

5.4 Detekcja migotania przedsionków w oparciu o entropie serii

Niniejsze badanie stanowi kontynuację dotychczasowego cyklu prac poświęconych detekcji migotania przedsionków z wykorzystaniem parametrów zmienności rytmu zatokowego. W wcześniejszych publikacjach zaprezentowano między innymi metody i wyniki badań związane ze statystycznymi oraz diagnostycznymi właściwościami parametrów $pRRx$ w detekcji AF [142], wpływem długości sygnału elektrokardiograficznego (EKG) na własności diagnostyczne miar HRV w kontekście rozpoznawania AF [143] oraz nowatorskie podejście wykorzystujące statystyki zliczeń względnych zmian między kolejnymi odstępami RR [144].

W szczególności niniejsza praca stanowi bezpośrednią kontynuację wyników przedstawionych w [142]. Podobnie jak w tamtym badaniu, podstawą analiz jest metodologia bazująca na budowie modelu regresji logistycznej opartej na parametrach policzonych z szeregu odstępów RR. Nowym aspektem wprowadzonym w opisywanym tu podejściu jest włączenie do analiz entropii serii monotonicznych odstępów RR. Ma to na celu zbadanie, czy wzbogacenie istniejących modeli o te dodatkowe miary entropii może przyczynić się do zwiększenia ich zdolności predykcyjnej w kontekście detekcji AF.

Koncepcja ta łączy w sobie zalety stosunkowo prostych do wyznaczenia parametrów zliczeniowych z bardziej złożonymi miarami entropijnymi, co w założeniu może prowadzić do lepszego rozdziału przypadków zdrowych od tych z migotaniem przedsionków oraz potencjalnie poprawić oceny prawdopodobieństwa występowania AF.

Zakładaliśmy, że entropia rytmu serca uzyskana z fragmentów spowolnień, przyspieszeń i neutralnych może różnić się między zapisami EKG pochodzącymi z rytmu zatokowego a migotaniem przedsionków. Jeśli tak jest, wówczas pomiar asymetrycznej entropii może okazać się przydatny w rozróżnianiu segmentów EKG z rytmem zatokowym i migotaniem przedsionków. W niniejszym badaniu porównano opisy oparte na asymetrycznej entropii z miarami $pRR30$ oraz $pRR3.25\%$ i ich możliwymi połączeniami w kontekście detekcji AF w jednominutowych segmentach odstępów RR.

Po uzyskaniu i przetworzeniu danych, nasze badanie składało się z kolejnych kroków:

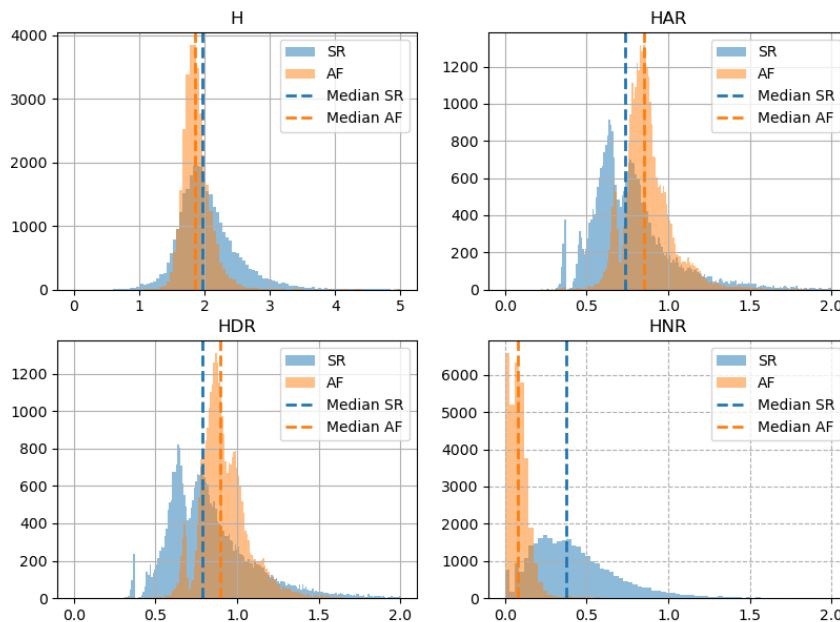
- Zbadanie rozkładów entropii asymetrycznej i parametrów zliczających ($pRR30$ i $pRR3.25\%$) dla segmentów oznaczonych jako pochodzące z nagrania z AF lub nagrań pochodzących z rytmu zatokowego (sinus rhythm, SR);

- Zbadanie diagnostycznych pojedynczych parametrów wykorzystując analize ROC;
- Budowa i ewaluacja modeli regresji logistycznej opartej na kombinacji badanych parametrów

Mając już opisany plan badania możemy przejść do opisu i dyskusji wyników.

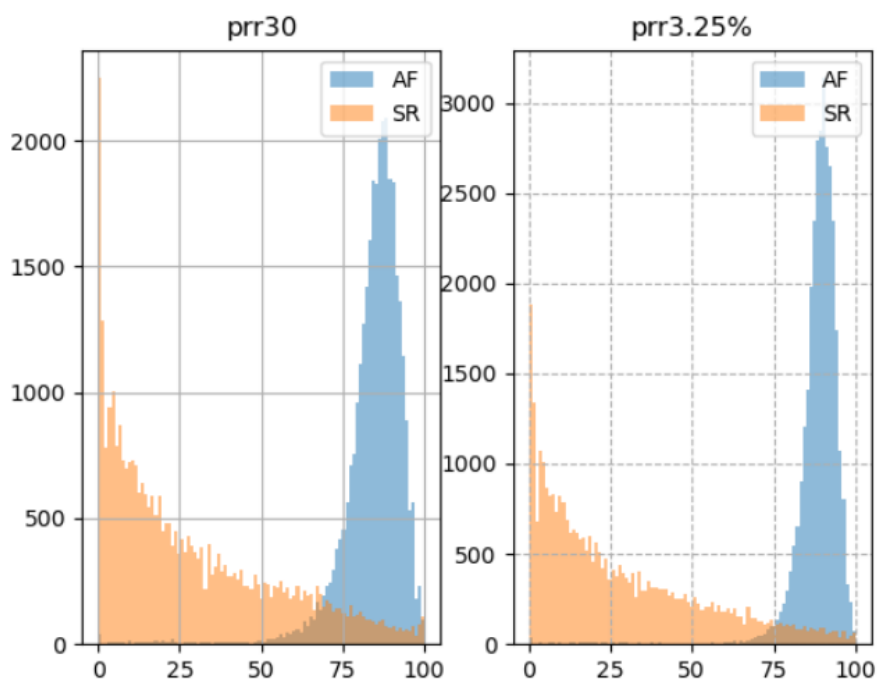
5.4.1 Opis danych

Poniższy rysunek przedstawia histogramy wszystkich analizowanych rodzajów entropii asymetrycznej: całkowitej entropii (H), entropii przebiegów przyspieszeń (HAR), entropii przebiegów spowolnień (HDR) oraz entropii przebiegów neutralnych (HNR). Dla migotania przedsionków (AF) wszystkie te entropie charakteryzują się węższymi i wyższymi rozkładami niż dla rytmu zatokowego (SR). Szczyty histogramów dla H częściowo się pokrywają w przypadku AF i SR, natomiast przy HAR, HDR i HNR rozdziela się. Średnie wartości median HAR i HDR są wyższe przy AF w porównaniu z SR, natomiast dla H i HNR wartości median są niższe przy AF.



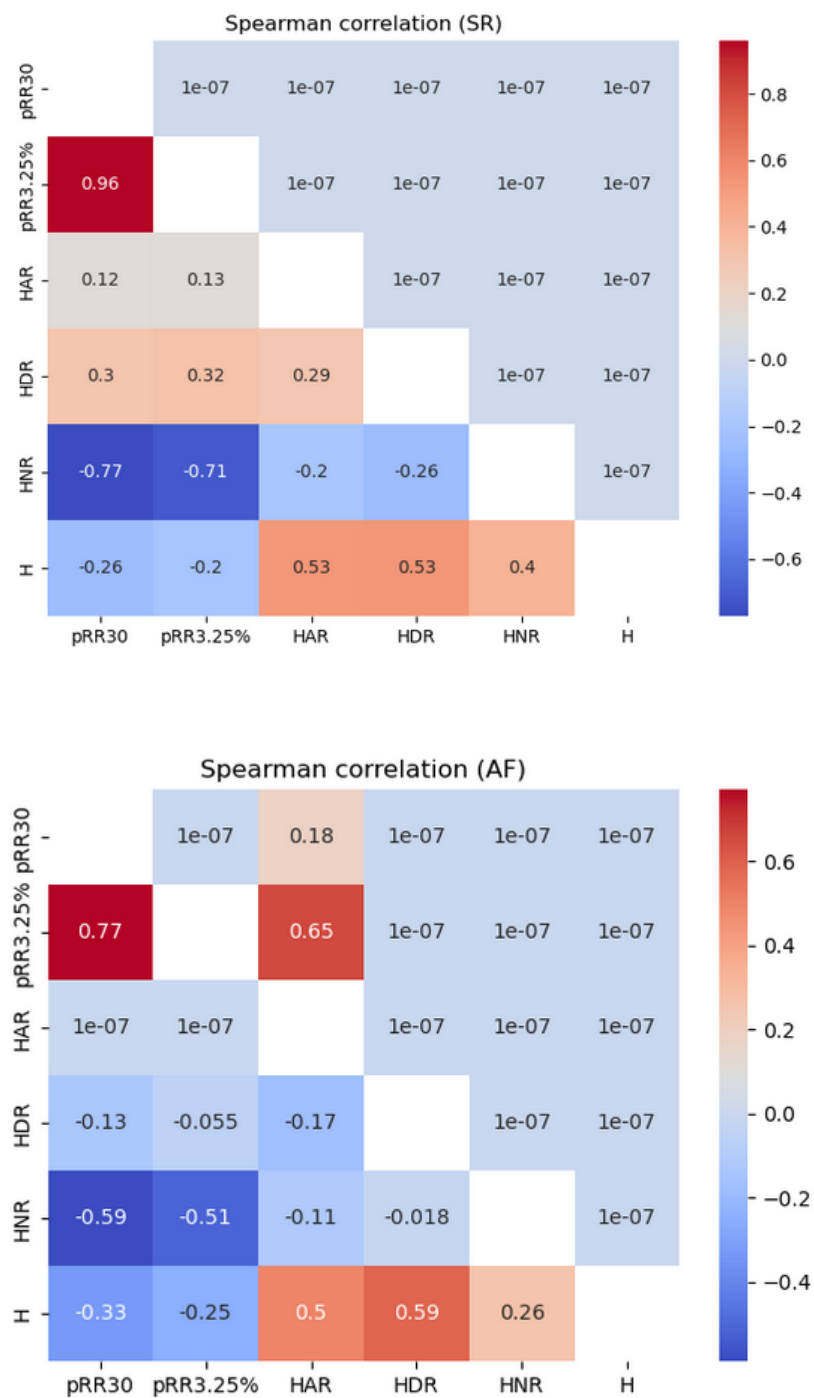
Rysunek 5.4: Histogramy różnych typów entropii dla SR i AF

Rysunek ilustruje, jak rozkłady pRR30 i pRR3.25% różnią się w migotaniu przedsionków (AF) i rytmie zatokowym (SR). Dla SR najbardziej typowe wartości pRR30 i pRR3.25% są zbliżone do 0%. Wraz ze wzrostem pRR30 i pRR3.25% częstość ich występowania stopniowo maleje. Natomiast w przypadku AF rozkłady przesuwają się w prawo, a najczęstsze wartości pRR30 i pRR3.25% wynoszą odpowiednio około 85% i 90%.



Rysunek 5.5: Histogramy pRR30 i pRR3.25% dla SR i AF

Zaprezentowano wartości korelacji Spearmana dla wszystkich analizowanych cech, osobno dla danych SR i AF. W przypadku SR, HNR silnie i ujemnie (ciemnoniebieski kolor) korelował z pRR30 (-0.77) oraz pRR3.25% (-0.71). Dla AF korelacje między HNR a pRR30 i pRR3.25% były słabsze, przy czym najsilniejsza dotyczyła HNR i pRR30 (-0.59).



Rysunek 5.6: Macierze współczynników korelacji Spearmana dla parametrów HRV dla SR i AF. Dolny trójkąt współczynniki korelacji, górny trójkąt wartości p

Tabela 1 opisuje własności diagnostyczne wszystkich badanych parametrów w rozróżnianiu AF od SR w jednodominutowych fragmentach EKG. Jako że pRR3.25% cechuje się najwyższą wartością AUC, jest on przyjęty jako punkt odniesienia do porównań z innymi parametrami i włączony do wszystkich wieloczynnikowych modeli diagnostycznych. Wartości AUC pozwalające odróżnić AF od SR w jednodominutowych odcinkach EKG odstępów RR były istotnie większe od 0.5 (p co najmniej $< 1 \times 10^{-114}$). Trzy najskuteczniejsze parametry w detekcji AF to pRR3.25% (AUC = 0.9727), pRR30 (0.9596) oraz HNR (0.9315).

Tabela 5.27: Pole pod krzywą ROC (AUC) i optymalne wartości progowe dla zbioru treningowego – część 1.

Parametr	AUC	Próg dla AF	Dokładność (%)	Czułość (%)
pRR3.25%	0.9727	< 72.3684%	95.75	99.60
pRR30	0.9596	< 66.8874%	91.00	95.62
HNR	0.9315	< 0.1884	84.86	95.45
HAR	0.6951	< 0.7546	67.85	86.83
HDR	0.6686	< 0.7726	67.68	91.49
H	0.6186	< 2.0647	60.03	85.68

Tabela 5.28: Pole pod krzywą ROC (AUC) i optymalne wartości progowe dla zbioru treningowego – część 2.

Parametr	Swoistość (%)	PPV (%)	NPV (%)	DOR
pRR3.25%	92.25	92.12	99.60	2931.82
pRR30	86.80	86.82	95.62	143.70
HNR	75.26	77.83	94.79	63.82
HAR	50.59	61.52	80.86	6.76
HDR	46.00	60.66	85.61	9.16
H	36.69	55.18	73.80	3.47

5.4.2 Wartości diagnostyczne modeli tworzonych z pRR30, pRR3.25% oraz miarami entropii asymetrycznej

Tabela 5.29 przedstawia wyniki detekcji AF z wykorzystaniem uogólnionego modelu regresji logistycznej, używającego pojedynczych lub wielu parametrów, na zestawie danych testowych. Najwyższą wartość AUC osiągnięto w modelu, w którym wykorzystano trzy zmienne: pRR30, pRR3.25%

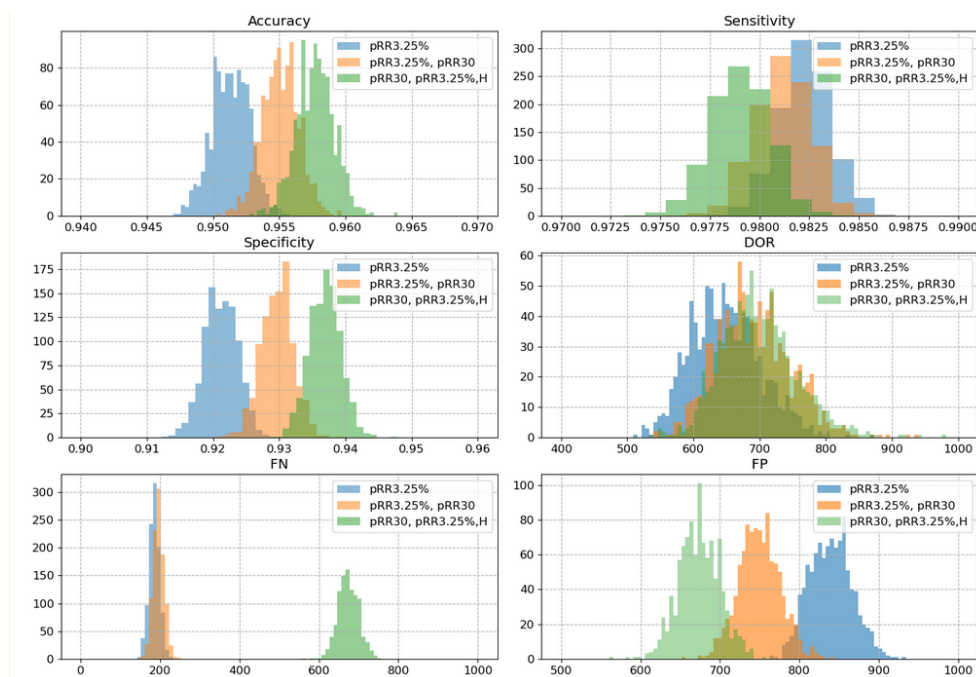
oraz H, natomiast połączenie pRR30, pRR3.25% i HAR osiągnęło najwyższy DOR. Wśród modeli jednoczynnikowych najlepsze wyniki uzyskał pRR3.25%, który cechował się najwyższą wartością AUC i pozostałych wskaźników.

Na podstawie Tabeli 5.29 wybrano trzy modele do dalszej weryfikacji: najlepszy (pod względem AUC) jednoczynnikowy model zawierający wyłącznie pRR3.25%, najlepszy dwuczynnikowy (pRR3.25% & pRR30) oraz najlepszy trójczynnikowy (pRR3.25%, pRR30 & H). Test tzw. bootstrapu z 1000 replik wykazał istotną statystycznie różnicę w AUC ($p = 0.02$) między modelem z jedną zmienną a modelem z dwiema zmiennymi. Ponadto porównania modelu jedno- vs trójczynnikowego oraz dwu- vs trójczynnikowego dały wartości p mniejsze od 10^{-6} , co wskazuje na bardzo znaczące różnice. Tę samą metodologię bootstrapu zastosowano w ocenie dokładności, czułości, swoistości, DOR, FN i FP. We wszystkich przypadkach różnice między modelami były istotne ($p < 10^{-5}$). Rysunek (5.7) pokazuje histogramy będące wizualnym odzwierciedleniem tych wyników.

Tabela 5.29: Porównanie różnych metryk dla detekcji migotania przedsionków (AF) przy użyciu modeli regresji logistycznej jedno- i wielowymiarowej. Trzy najlepsze metryki z najwyższymi wartościami AUC zostały wyróżnione pogrubieniem. Parametr pRR3.25% wyróżnia się jako najsukceszniejsza pojedyncza metryka, osiągając najwyższe wartości AUC. Co istotne, pRR3.25% regularnie pojawia się w najlepiej działających parach i trójkach metryk, wykazując doskonałą skuteczność w rozróżnianiu AF od rytmu zatokowego (SR).

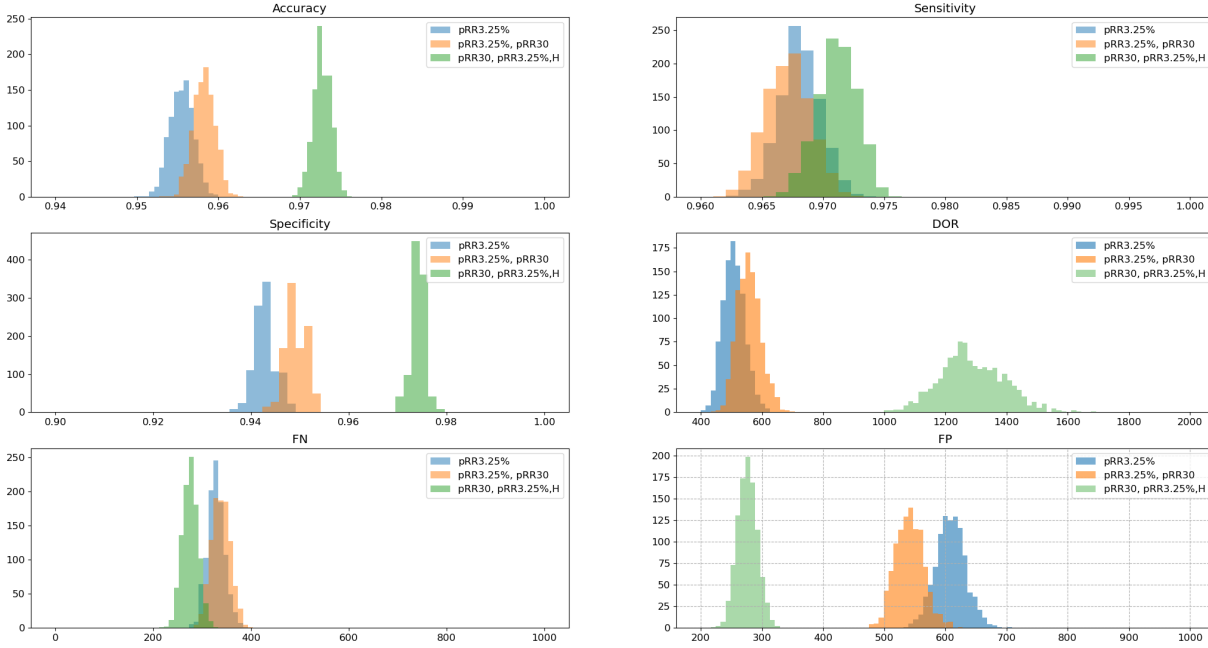
Cechy	AUC	Dokładność (%)	DOR	FP+FN (%)	FP (%)	FN (%)
pRR3.25%	0.972	95.13	643.0	4.87	3.98	0.88
pRR30	0.959	92.77	233.5	7.23	5.57	1.65
H	0.613	61.18	2.6	38.82	24.30	14.52
HAR	0.700	65.36	3.6	34.64	16.11	18.53
HDR	0.670	61.55	2.6	38.45	16.91	21.54
HNR	0.930	88.42	84.5	11.58	9.03	2.55
pRR3.25% & pRR30%	0.978	95.50	687.1	4.50	3.57	0.93
pRR3.25% & H	0.986	95.66	705.6	4.34	3.38	0.96
pRR3.25% & HAR	0.982	95.49	693.0	4.51	3.59	0.92
pRR3.25% & HDR	0.977	95.17	623.3	4.83	3.90	0.93
pRR3.25% & HNR	0.974	95.16	632.3	4.84	3.93	0.91
pRR30 & H	0.973	93.27	252.3	6.73	5.01	1.72
pRR30 & HAR	0.969	93.12	240.5	6.88	5.13	1.76
pRR30 & HDR	0.963	92.79	235.8	7.21	5.57	1.64
pRR30 & HNR	0.960	92.84	232.9	7.16	5.47	1.69
pRR30 & pRR3.25% & H	0.988	95.77	695.5	4.23	3.20	1.03
pRR30 & pRR3.25% & HAR	0.985	95.81	732.4	4.19	3.22	0.97
pRR30 & pRR3.25% & HDR	0.982	95.41	633.7	4.59	3.58	1.00
pRR30 & pRR3.25% & HNR	0.979	95.37	631.3	4.63	3.65	0.99
HNR & HAR & HDR	0.935	88.38	82.8	11.62	9.01	2.61

Warto dodać, że skale osi w Rysunku 5.19 dla dokładności, czułości i swoistości zostały odpowiednio dopasowane, aby lepiej uwidocznić niewielkie różnice. W praktyce jednak rozpiętość tych odchyłeń jest niewielka. Optymalizacja modeli pod kątem AUC zwiększa ryzyko niewykrywania AF. Mediana FN (liczba epizodów AF nierozpoznanych) rośnie z wartości poniżej 200 do niemal 700 po uwzględnieniu trzech parametrów (pRR3.25%, pRR30 i H).



Rysunek 5.7: Histogramy przedstawiają wyniki 1000-krotnego bootstrappingu dla modelu zbudowanego na podstawie pRR3.25% w porównaniu do pRR30 oraz pRR3.25% w porównaniu do modelu zbudowanego na podstawie pRR30, pRR3.25% i H . Największa różnica dotyczy liczby FN (fałszywych negatywów), gdzie dodanie parametru H do modelu zwiększa liczbę FN ponad trzykrotnie w porównaniu z modelami bez parametru H .

Po opublikowaniu opisanych wcześniej badań, rozważono możliwość zmiany algorytmu do trenowania modeli. Wybrano algorytm XGBoost dla którego powtórzono całą procedurę. Wyniki z 1000-krotnego bootstrappingu najlepszego modelu zawierającego parametry pRR30, pRR3.25% oraz H znajdują się na poniższym wykresie.



Rysunek 5.8: Histogramy przedstawiają wyniki 1000-krotnego bootstrapingu dla modelu zbudowanego na podstawie pRR3.25% w porównaniu do pRR30 oraz pRR3.25% w porównaniu do modelu zbudowanego na podstawie pRR30, pRR3.25% i HAR w oparciu o algorytm XGBoost. Widać tutaj że w przeciwieństwie do modeli opartych o regresję logistyczną, w modelu z H nie ma tutaj problemu ze znacznym wzrostem wyników fałszywie negatywnych.

5.4.3 Dyskusja

Zaobserwowaliśmy, że zarówno pRR3.25%, jak i pRR30 przewyższają pod względem diagnostyki wszystkie parametry oparte na entropii asymetrycznej, jeśli chodzi o detekcję AF w jednonominutowych zapisach EKG. Najlepszym pojedynczym wskaźnikiem w detekcji AF jest pRR3.25%, z AUC wynoszącym 0.972 i DOR = 643; druga w kolejności jest pRR30 (AUC = 0.959, DOR = 233.5), a trzecia – HNR (AUC = 0.93, DOR = 84.5). Wszelkie kombinacje, które dały wyższy AUC lub DOR niż sam pRR3.25%, zawierały właśnie pRR3.25%. Najlepsze połączenie (według wartości AUC) to pRR3.25%, pRR30 i H, a według DOR – pRR3.25% z H. Niemniej jednak modele zbiorcze o wyższej wartości AUC lub DOR wiązały się z

większą liczbą wyników fałszywie ujemnych. Przyjmując pRR3.25% jako punkt odniesienia, wzrost AUC był klinicznie niewielki (do 0.006, czyli mniej niż 1%). Co więcej, wybór najlepszego zestawu parametrów (pRR3.25%, pRR30 i H) pociągał za sobą ryzyko prawie czterokrotnie większej liczby nierozpoznanych epizodów AF.

Nie jest to pierwsze opracowanie, w którym opisano wykorzystanie miar opartych na entropii do detekcji AF. W pracach [85] Richman JS i Moorman JR zaproponowali tzw. entropię próbkową (sample entropy), jako nową metodę porównywania dwóch szeregów czasowych w fizjologii. Z kolei Liu C et al. i Zhao L et al. [145, 146] porównywali różne miary entropii i zaproponowali znormalizowaną entropię rozmytą (normalized fuzzy entropy) w detekcji AF, osiągając dokładność 0.86 dla 60-uderzeniowych fragmentów odstępów RR. Ich podejście, podobnie jak inne analizowane wskaźniki, opierało się na entropii z próby [85].

Zaobserwowaliśmy ujemną korelację między entropią przebiegów neutralnych (HNR) a pRR30. Rozkład HNR różni się pomiędzy AF a SR, a klasyfikator AF oparty wyłącznie na HNR osiąga dokładność ponad 0.88, co stanowi najlepszy wynik wśród innych miar entropii asymetrycznej opartych na seriach. Serie neutralne bywają często pomijane, usuwane z analizy lub zastępowane sztucznym szumem [8, 147, 80], jednak w kontekście detekcji AF mogą zawierać przydatne informacje. Zwykle ich udział w HRV bywa znikomy albo przypadkowy (zależy też od częstotliwości próbkowania EKG). Niemniej mniejszy udział przebiegów neutralnych w AF wydaje się istotny.

Charakter AF sprzyja losowym, mało stabilnym odstępom RR, co ogranicza szansę wystąpienia kolejnych odstępów o takiej samej długości. Tłumaczy to mniejszą liczbę neutralnych serii w AF. Z kolei zmiany odstępów RR w SR są stopniowe, a regulują je liczne mechanizmy fizjologiczne wpływające na depolaryzację węzła zatokowego. W AF szybkie fale (ponad 350 depolaryzacji na minutę) resetują i osłabiają aktywność węzła zatokowego, a zmiany kolejnych odstępów RR są mniej kontrolowane i bardziej gwałtowne. W poprzednich pracach wykazaliśmy wyższe wartości pRRx i pRRx% w AF w porównaniu z SR [143]. Z tego powodu w AF maleje prawdopodobieństwo niewielkich zmian kolejnych odstępów RR, a w SR jest ono większe. Z powyższych obserwacji wynika, że informacje wynikające z przebiegów neutralnych mogą wnieść pewną wartość diagnostyczną, choć nie jest pewne, czy należy je bezwzględnie uwzględniać w algorytmach detekcji AF. Wszystkie badane miary entropii asymetrycznej oparte na seriach – zarówno samodzielne, jak i w połączeniu z pRR3.25% czy pRR30 – nie okazały się szczególnie skuteczne w rozróżnianiu AF od SR. Są też bardziej złożone obliczeniowo niż pRR3.25% lub pRR30.

Zgodnie z wcześniejszymi ustaleniami, rodzina parametrów pRRx i

pRRx% (zwłaszcza pRR30 i pRR3.25%) wykazuje atrakcyjne właściwości diagnostyczne w odróżnianiu odcinków EKG z AF od SR. Najpierw dowiedziono, że pRR30 przewyższa inne często stosowane miary HRV (SDNN, SD1, SD2, SD2/SD1, współczynnik zmienności) w podobnych zastosowaniach [143]. Następnie wykazano, że pRR30 ustępuje miejsca pod względem skuteczności rodzinie pRRx%, w szczególności pRR3.25%. W niniejszym badaniu stwierdzamy, iż pRR30 oraz pRR3.25% lepiej od wszelkich wskaźników entropii asymetrycznej typu serii rozpoznają AF, a dodanie miar opartych na entropii asymetrycznej do modeli zawierających pRR30 lub pRR3.25% przynosi jedynie niewielki wzrost AUC lub DOR, jednocześnie podnosząc liczbę wyników fałszywie ujemnych.

Znaczenie wyników fałszywie dodatnich i ujemnych jest zróżnicowane i zależy od charakteru badanej choroby oraz kontekstu badania. W przypadku AF za wynik groźniejszy można uznać wynik fałszywie ujemny, ponieważ pacjent z nierozpoznanym AF nie otrzyma leczenia przeciwzakrzepowego i pozostanie w grupie ryzyka udaru niedokrwinnego. Fałszywie dodatni wynik pociąga za sobą dodatkową diagnostykę, lecz bez bezpośrednich negatywnych konsekwencji [148, 149].

Głównym ograniczeniem naszych analiz jest to, że opierały się one wyłącznie na odstępach RR z jednominutowych zapisów EKG, co może nie w pełni oddawać dłuższych zmian rytmu serca. Nie można więc uogólnić tych wniosków na dłuższe bądź krótsze fragmenty EKG. Kolejne ograniczenie to częstotliwość próbkowania na poziomie 128 Hz, co technicznie zwiększa liczbę neutralnych odstępów RR. W badaniu zastosowano klasyczny, prosty model regresji logistycznej. Choć osiągnięto wysoką dokładność (0.96), można rozważyć inne, bardziej złożone modele uczenia maszynowego. Jest możliwe, iż algorytmy ML (np. *random forest*) osiągną lepsze wyniki niż regresja logistyczna [150, 109], chociaż przy bardzo wysokiej wartości AUC pRR3.25% (bliskiej 1) pole do poprawy pozostaje niewielkie. Warto byłoby jednak zbadać, czy np. XGBoost lub *Random Forest* pozwoliłyby na lepszą detekcję AF z wykorzystaniem entropii asymetrycznej przebiegów neutralnych, wychwytyjąc bardziej złożone zależności.

Parametr pRR3.25% wyróżnia się jako wysoce skuteczny i niezawodny przy identyfikacji epizodów AF, przewyższając pod tym względem entropię asymetryczną bazującą na serii monotonicznych. Zauważono, że bardziej złożone modele, takie jak pRR3.25% + pRR30 + H, co prawda osiągają nieco wyższe wyniki pod względem oceny ogólnej (AUC), jednak korzyść ta jest równoważona niemal czterokrotnie większą liczbą wyników fałszywie ujemnych w analizie opartej na regresji logistycznej. Wszystkie modele o wyższej efektywności włączają w swoje obliczenia parametr pRR3.25%, co potwierdza jego podstawowe znaczenie w dokładności diagnostycznej.

W praktyce klinicznej jednoczynnikowy model oparty na pRR3.25% może zatem zapewnić optymalną równowagę między jakością detekcji a liczbą przypadków nierozpoznanych.

Dalsza analiza z wykorzystaniem algorytmu XGBoost. Po zakończeniu i publikacji niniejszych badań przeprowadzono jednak dodatkowe testy, w których do klasyfikacji epizodów AF wykorzystano model oparty na algorytmie XGBoost. W przeciwieństwie do wyników uzyskanych przy użyciu prostej regresji logistycznej, nowa analiza wykazała, że uwzględnienie entropii H w modelu nie tylko nie pogarsza detekcji AF (pod kątem liczby fałszywie ujemnych), lecz wprost przeciwnie – pozwala na jej dalszą poprawę. Opracowany model z parametrami pRR30, pRR3.25% oraz H w XGBoost okazał się najlepszy spośród wszystkich badanych wariantów, osiągając najwyższe wskaźniki jakości diagnostycznej (m.in. AUC i obniżenie FN). Szczegółowe rezultaty tej analizy zostały zaprezentowane na konferencji [151]. Fakt ten dowodzi, że w zależności od zastosowanej metody modelowania (regresja logistyczna vs. XGBoost) rola entropii H przy identyfikacji AF może być różna – w przypadku algorytmu XGBoost entropia asymetryczna przebiegów neutralnych może stanowić cenne uzupełnienie miar pRR3.25% i pRR30, istotnie poprawiając wykrywalność tej arytmii.

Nowością prezentowanego badania jest szczegółowa analiza detekcji AF z wykorzystaniem świeżego połączenia entropii asymetrycznej opartej na seriach monotonicznych i parametrów pRR30 z pRR3.25%. Entropii asymetrycznej opartej na seriach monotonicznych nie rozważano wcześniej jako kandydata w detekcji AF. Drugą nowością było świadome włączenie w analizę przebiegów neutralnych zamiast ich usuwania, co może nieść pewną wartość poznawczą.

Wyniki naszego badania wskazują, że parametr HNR może okazać się pomocny w detekcji AF, a jego potencjał zasługuje na dalsze badania.

5.5 Refaktoryzacja algorytmu NCM

5.5.1 Algorytm NCM

Podstawą algorytmu NCM (Norm component matrix) jest trójkątna macierz $\hat{N}$, której elementy $n_{i,j}$ są zdefiniowane jako:

$$n_{i,j} = |u(i) - u(i + (j + 1) \cdot \tau)|, \quad (5.39)$$

gdzie $u(i)$ to wartość szeregu czasowego, a τ to opóźnienie czasowe. Dla szeregu o długości N , macierz ta ma $N - \tau - 1$ wierszy oraz $(N - 1)/\tau - 1$

kolumn. Jednak w praktycznej implementacji wystarczające jest operowanie jedynie na kilku wierszach i kolumnach w danym momencie.

Dzięki symetrii, wykorzystuje się jedynie górną część macierzy, w której indeksy i i j spełniają warunek $i + (j + 1) \cdot \tau \leq N - \tau - 1$. Elementy tej macierzy reprezentują odległości pomiędzy zerowymiarowymi wektorami osadzania:

$$n_{i,j} = |v_0^\tau(i) - v_0^\tau(i + (j + 1) \cdot \tau)|, \quad (5.40)$$

gdzie $v_0^\tau(i)$ jest wektorem zdefiniowanym w równaniu (2). Kluczową właściwością macierzy $\hat{N}$ jest możliwość obliczenia odległości w przestrzeni osadzania o wyższym wymiarze m poprzez:

$$n_{i,j}^m = \max_{k=0,1,\dots,m-1} n_{i+k\cdot\tau,j}. \quad (5.41)$$

Ta właściwość znacząco redukuje liczbę obliczeń potrzebnych do wyznaczenia sum korelacyjnych $C_m(r)$. Dzięki temu, obliczenia te są znacznie szybsze w porównaniu z klasycznymi metodami.

Dalsze kroki obejmują porównanie wartości $n_{i,j}^m$ z zestawem progów r , co pozwala na generowanie poszczególnych macierzy dla kolejnych wartości r . Każda kolumna zawierająca $n_{i,j}^m$ jest wypełniana do indeksu odpowiadającego progowi r , który spełnia warunek $r \geq n_{i,j}^m$. Aby uprościć to zadanie, zestaw progów r można zdefiniować jako funkcję liniową:

$$r = a \cdot i + b, \quad (5.42)$$

gdzie $a = -\frac{r_{\max} - r_{\min}}{n-1}$ oraz $b = r_{\max}$. Zamiast iteracyjnie porównywać wartości r , indeks pierwszego wystąpienia spełniającego warunek można wyznaczyć za pomocą prostego działania:

$$\text{indeks} = \frac{n_{i,j}^m - b}{a}. \quad (5.43)$$

Dzięki zastosowaniu opisanej metody, algorytm staje się znacznie bardziej efektywny i niezależny od gęstości próbkowania progów r . Ponadto, metoda ta może być efektywnie implementowana w sposób równoległy z wykorzystaniem technik takich jak SIMD (*Single Instruction Multiple Data*) czy MPI (*Message Passing Interface*) a także przy wykorzystaniu nowoczesnych kart graficznych. To umożliwia przetwarzanie dużych zbiorów danych w krótkim czasie, co jest szczególnie istotne w zastosowaniach klinicznych.

Metoda macierzy składowych norm stanowi szybkie i skuteczne rozwiązanie do obliczania wymiaru korelacyjnego, a także pokrewnych parametrów, takich jak entropia próbki (SampEn) oraz entropia przybliżona (ApEn).

Dzięki swojej efektywności może znaleźć zastosowanie w analizie dużych szeregów czasowych, nawet w czasie rzeczywistym.

Sekcja ta zawiera szczegółowe opisy implementacji dwóch wersji algorytmu: wersję *wprost* realizującą koncepcje algorytmu oraz wersję *wzmocnion* o dodatkowe niskopoziomowe operacje numeryczne. Dodatkowo na potrzeby testów i analiz wydajnościowych wykorzystuje się metodę *brute force*.

W kolejnych podsekcjach wyjaśnione zostaną szczegóły implementacji każdego z wymienionych podejść, w tym przyjęte założenia i stosowane optymalizacje. Dodatkowo, w następnej sekcji zostaną zaprezentowane wyniki analizy porównawczej wydajności wszystkich badanych wersji algorytmu, co pozwoli na ocenę ich efektywności w różnych warunkach.

Wszystkie opisane implementacje zostały zintegrowane w postaci pakietu Pythonowego dystrybuowanego za pomocą systemu zarządzania pakietami PiP. PiP (*Pip Installs Packages*) to narzędzie służące do instalacji i zarządzania bibliotekami w ekosystemie Python. Dzięki temu pakiet jest łatwo dostępny dla użytkowników, a jego otwartoźródłowy charakter umożliwia dowolne wykorzystanie, analizę i rozwijanie udostępnionego kodu. Przygotowanie powszechnie dostępnego pakietu dystrybucyjnego wykonano w ramach prac nad niniejszą rozprawą doktorską.

Jako język programowania wykorzystany do implementacji algorytmu *explicite* wybrano język Python w wersji 3.8. Do wzmocnienia możliwości obsługi niskopoziomowych operacji numerycznych wykorzystano pakietu NumPy w najnowszej dostępnej wersji (1.2.4).

Omówione metody obejmują trzy podejścia:

1. **Metoda "brute force"** – bezpośrednia implementacja obliczania sum korelacyjnych, która bazuje na ich matematycznej definicji:

$$C(r) = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j \neq i} H(r - \|x_i - x_j\|), \quad (5.44)$$

gdzie:

- $C(r)$ to suma korelacyjna dla zadanego promienia r ,
- N to liczba punktów w zbiorze,
- H to funkcja Heaviside'a, która przyjmuje wartość 1, gdy argument jest nieujemny, i 0 w przeciwnym przypadku,
- $\|x_i - x_j\|$ to odległość pomiędzy punktami x_i i x_j .

2. **Metoda NCM w wersji *wprost*** – implementacja algorytmu zaproponowanego w pracy [86]. Jest to podstawowa wersja algorytmu, która

realizuje obliczenia zgodnie z opisem w tej publikacji, bez dodatkowych optymalizacji niskopoziomowych.

3. **Metoda NCM w wersji *wzmocnionej*** – ulepszona implementacja algorytmu NCM, która wykorzystuje pełen potencjał wektoryzacji obliczeń przy użyciu pakietu NumPy. Ta wersja optymalizuje wykonywanie operacji poprzez ograniczenie liczby iteracji oraz zastosowanie zaawansowanych funkcji dostępnych w bibliotece NumPy.

Kod źródłowy zawierający implementacje wszystkich opisanych metod jest dostępny w repozytorium *NCM* [adres repo](#). Repozytorium to zawiera również zestaw testów, które potwierdzają zbieżność wyników uzyskanych przy użyciu każdej z przedstawionych metod. Testy te zapewniają wiarygodność wyników i pozwalają na ich dalsze zastosowanie w analizach.

W kolejnych podsekcjach zostaną przedstawione detale implementacyjne wersji *wprost* (oznaczenie w kodzie źródłowym jako wersja *plain*) i *wzmocnionej* (oznaczenie w kodzie źródłowym jako wersja *enhanced*). Przedstawiony jest tutaj też kod wersji *brute force* jednak ze względu na trywialność tej implementacji opis detali zostanie pominięty.

```
1 import numpy
2
3 NumberTypes = (int, float)
4
5
6 class Matrix:
7     def __init__(self, signal):
8         self.__class__.__name__ = "Bruteforce_correlation_sums"
9         self.signal = signal
10
11     @staticmethod
12     def heaviside(x):
13         if x >= 0:
14             return 1
15         else:
16             return 0
17
18     @staticmethod
19     def max_coordinate_diff(v1, v2):
20         return max(abs(v1-v2))
21
22     def corsum_matrix(self, m, r_range, tau=1, precision=6):
23         results = []
24         m = m[0]
25         Lm = len(self.signal) - (m - 1) * tau
26         Corr_sum = 0
27         r_range = [r_range] if type(r_range) in (int, float) else r_range
28         for r in r_range:
29             for i in range(Lm):
30                 v_i = numpy.array(self.signal[i:i + m:tau])
31                 for j in range(Lm):
32                     if i == j:
33                         continue
34                     v_j = numpy.array(self.signal[j:j+m:tau])
35                     Corr_sum += self.heaviside(r-self.max_coordinate_diff(v_i, v_j))
36                 result = Corr_sum * (1 / Lm) * (1 / (Lm - 1))
37                 results.append(result)
38         return numpy.round(results, precision)
```

Rysunek 5.9: Kod programu wersji "brute force"

5.5.2 Implementacja NCM: wersja *wprost*

```

1 import numpy as np
2 import itertools
3
4 NumberTypes = (int, float)
5
6
7 class Matrix:
8
9     def __init__(self, signal):
10         self.__class__.__name__ = "NCM_plain"
11         self.signal = signal
12
13     def corsum_matrix(self, m_range, r_range, tau:int=1, normalize=None, selfmatches=None,
14                       precision=6):
15         assert tau == 1
16         signal = self.signal
17         m_counts = max(m_range)
18         m_range = range(m_counts)
19
20         # find linear coefficient for solving algebraic equation
21         r_range = np.array([r_range]) if type(r_range) in (int, float) else r_range
22
23         corsum_matrix = np.zeros((m_counts, len(r_range)))
24         if r_range.size > 1:
25             a = -(r_range[0] - r_range[-1]) / (r_range.size - 1)
26             b = r_range[0]
27
28         # define size of Norm component matrix
29         size_X = (len(signal) - (1 - 1) * tau) - 1
30         size_Y = (len(signal) - 1 - (1 - 1) * tau)
31         # create triangular NCM matrix
32         NCM = np.zeros((size_X, size_Y))
33         for i_row in range(size_X):
34             for j_column in range(size_Y):
35                 if i_row + (j_column + 1) * tau <= len(signal) - 1:
36                     NCM[i_row, j_column] = np.abs(signal[i_row] - signal[i_row + (j_column + 1) *
37                                                         tau])
38
39         # calculate correlation sum for embedded dimension m
40         for m in m_range:
41             for current_row_idx in range(len(NCM) - m): # Careful - we need to iterate over
42                 # len(NCM) - m rows of NCM matrix, because we cant take more than one row on last row
43                 current_row = NCM[current_row_idx:current_row_idx + (m+1)]
44                 current_row = current_row[:, :size_Y - current_row_idx - m].max(axis=0)
45                 # Solve for v using the given coefficients a and b
46                 z = np.zeros(r_range.size)
47                 if r_range.size == 1:
48                     z[0] += (current_row <= r_range[0]).sum()
49                 else:
50                     v = ((current_row - b) / a).astype('int')
51                     v = np.sort(v)
52                     for k, v in itertools.groupby(v):
53                         z[k + 1] += len(list(v))
54                 corsum_matrix[m] += z
55
56         # normalize correlation sum, multiply by 2 due to property of triangular matrix and
57         # exclude duplicates
58         for m in m_range:
59             factorA = (len(signal) - (m) * tau)
60             factorB = (len(signal) - 1 - (m) * tau)
61             factor = factorA * factorB
62             corsum_matrix[m] = corsum_matrix[m] * 2 * 1/factor
63         return np.round(corsum_matrix.T, precision)

```

Rysunek 5.10: Kod programu wersji "NCM plain"

Poniżej przedstawiono detale implementacji wersji *wprost (plain)*. Kluczową strukturą tego podejścia jest trójkątna macierz $n_{i,j}$ co pozwala efektywnie generować odległości w przestrzeni o wyższym wymiarze na bazie *pojedynczej* macierzy $\tilde{N}$. Dzięki temu obliczenia sum korelacyjnych $C_m(r)$ mogą być przeprowadzane znacznie szybciej niż w klasycznych metodach.

Linie 1–5. Pierwsze instrukcje (`import`) odpowiedzialne są za wczytanie bibliotek wymaganych do poprawnego działania skryptu, w szczególności `numpy` (operacje macierzowe) oraz `itertools` (zestaw narzędzi do iteracji i grupowania danych).

Linie 6–9. Zdefiniowano klasę `Matrix`, w której konstruktor (`__init__`) przechowuje przetwarzany sygnał (szereg czasowy) w atrybucie `self.signal`. W tym fragmencie ustalana jest też nazwa klasy na `'NCM_plain'`, sugerująca, że chodzi o podstawowy wariant algorytmu bazującego na macierzy NCM.

Linie 10–15. Metoda `corsum_matrix` rozpoczyna definicję głównej procedury służącej do obliczania sum korelacyjnych (lub wartości powiązanych) dla różnych wymiarów osadzenia m oraz progów r . W tym miejscu ustalone są podstawowe parametry:

- `m_range` – zakres wymiarów m , z którego pobierana jest wartość maksymalna do faktycznej pętli `range(m_counts)`,
- `r_range` – zbiór progów r ,
- τ – opóźnienie czasowe (w kodzie z *góry* założono, że $\tau = 1$ w linii 11).

Linie 16–20. Tworzona jest tablica dwuwymiarowa `corsum_matrix`, w której wiersze będą odpowiadały kolejnym wartościom m , a kolumny – progom r . Gdy rozmiar `r_range` jest większy niż 1, w linii 18–19 obliczane są współczynniki (a, b) umożliwiające bardzo szybką konwersję wartości $n_{i,j}^m$ na indeks odpowiadający progowi w zbiorze `r_range`. Zamiast powtarzających się porównań sprawdza się więc

$$\text{indeks} = \frac{n_{i,j}^m - b}{a}. \quad (5.45)$$

Linie 21–30. W linii 21–22 określone są rozmiary `size_X` i `size_Y`, które decydują o wymiarach macierzy NCM. Zgodnie z definicją, wartości tej trójkątnej macierzy stanowią bezwzględną różnicę między próbkami sygnału, oddalonymi o τ (a dokładniej $(j+1)\tau$). Następnie w linii 23 – 24 macierz NCM jest inicjalizowana zerami, po czym w liniach 25–30 dwie zagnieżdżone pętle uzupełniają ją obliczając

$$n_{i,j} = |u(i) - u(i + (j+1)\tau)|. \quad (5.46)$$

Część warunku `if (i_row + ... <= ...)` odpowiada temu, że w praktyce przetwarzamy jedynie trójkątny fragment macierzy (bez powielania obliczeń).

Linie 31–55. Następuje właściwe wyliczanie wartości $n_{i,j}^m$ oraz sum korelacyjnych. Dla każdego wymiaru m (linia 31) w pętli `for` rozpatrywane są kolejne indeksy wierszy `current_row_idx` (linia 32). Od linii 33 do 35 pobierany jest fragment $(m+1)$ wierszy z macierzy NCM i obliczane jest `max(axis=0)`, co odpowiada operacji

$$n_{i,j}^m = \max_{k=0,\dots,m-1} n_{(i+k),j}. \quad (5.47)$$

Następnie w liniach 37–54, dla każdej uzyskanej tablicy wartości, zlicza się, ile spośród nich jest mniejszych bądź równych wartościom w `r_range`. W przypadku pojedynczego progu (linia 38) wystarczy jedno proste porównanie. Natomiast przy wielu progach (linie 40–45) stosuje się metodę konwersji do indeksu $\text{indeks} = (n_{i,j}^m - b)/a$ i grupowania (`itertools.groupby`), co oszczędza czas obliczeń. Wyniki trafień są sumowane w `corsum_matrix[m]` w linii 54.

Linie 56–62. Zwieńczeniem jest normalizacja wyników. W liniach 57–60 obliczany jest czynnik `factor` odpowiadający liczbie możliwych par punktów w sygnale (dla danej liczby wymiaru m). Aby uwzględnić fakt, że przetwarzano tylko *trójkąt górny* macierzy, każda z wartości `corsum_matrix[m]` jest mnożona przez 2 (linia 61), co zapewnia otrzymanie poprawnych sum korelacyjnych $C_m(r)$. Następnie w linii 61–62 otrzymujemy gotowe, unormowane wartości, które mogą być interpretowane w kategoriach wymiaru korelacyjnego, entropii przybliżonej czy też entropii z próby, zależnie od kontekstu. Ostateczny wynik jest zwracany po transpozycji oraz zaokrągleniu do wskazanej liczby miejsc dziesiętnych.

5.5.3 Implementacja NCM: wersja *wzmocniona*

Poniżej przedstawiono analizę implementacji algorytmu *NCM.enhanced*. Kod wykorzystuje interfejs MPI oraz bibliotekę NumPy do obliczeń równoległych, dzięki czemu może być skalowany na wielu procesach, a tym samym przyspiesza analizę dużych serii czasowych.

```

1 import numpy
2 import itertools
3
4 from mpi4py import MPI
5 comm = MPI.COMM_WORLD
6 sub_comm = MPI.COMM_WORLD.Clone()
7 rank = comm.Get_rank()
8 size = comm.Get_size()
9 NumberTypes = (int, float)
10
11 class Matrix:
12     """ Norm Component Matrix """
13
14     def __init__(self, series):
15         self.series = series
16         self.max_diff = self.series.max() - self.series.min()
17         self.N = len(self.series)
18         self._rows = {}
19
20     def windups(self, m_counts, row, tau):
21         """ A generator that iterates over m_counts successive rows,
22             starting from the now nnumber `row` in the base norm-component
23             matrix.
24
25         """
26         current_row = self.get_row(row, tau)
27         for m in range(m_counts):
28             if row + m * tau == self.N - 1:
29                 break
30             v = self.get_row(row + m * tau, tau)
31             current_row = numpy.vstack((current_row[:v.size], v)).max(axis=0)
32             yield current_row
33
34     def get_row(self, row, tau):
35         # The matrix is generated with the successive rows defined below, where
36         # each row is the difference of the [i+1:] series subset with the i-th series element
37         # *BUT* the i goes from [0, N -1], the last point is excluded
38         try:
39             return self._rows[row]
40         except KeyError:
41             self._rows[row] = self._calc_row(row, tau)
42             return self._rows[row]
43
44     def _calc_row(self, row, tau):
45         """ Returns NCM row.
46
47         Arguments:
48         - row: the row index to calculate
49         - tau: time lag value to use
50
51         The row is calculated with the matrix definition given by:
52
53         
$$n_{i,j} = |\vec{v}_{\{0,\tau\}}(i) - \vec{v}_{\{0,\tau\}}(i + 1 + j \cdot \tau)|$$

54
55         so we do  $n_{i,:}$  here, using numpy vectorized way to process the whole
56         array in one readable code line.
57         """
58         return numpy.abs(self.series[row] - self.series[row + tau::tau])
59
60     def generate_rows(self, tau):
61         """ Yields all rows in the NCM matrix. """
62         # The matrix is generated with the successive rows defined below, where
63         # each row is the difference of the [i+1:] series subset with the i-th series element
64         # *BUT* the i goes from [0, N -1], the last point is excluded
65         # for i in xrange(0, self.N - tau, tau):
66         for i in range(0, self.N - tau, 1):
67             yield self.get_row(i, tau)
68
69     def _r_range_filter(self, row, r_range, a, b):
70         # Iterates over the al row elements and check the
71         # number of the existance od the elements lower than
72         # the filter value applied.
73         counts = []
74         for r in r_range: # r_range have to be sorted!
75             row = row.compress(row <= r)
76             counts.append(len(row))
77         return numpy.array(counts)

```

Rysunek 5.11: Kod programu wersji "NCM.enhanced", część pierwsza.

```

87 def r_range_filter(self, row, r_range_size, a, b):
88     v = ((row - b)/a).astype('int')
89     if numpy.any(v<0):
90         raise ValueError
91     v.sort()
92     z = numpy.zeros(r_range_size)
93     for k, v in itertools.groupby(v):
94         z[k+1] += len(list(v))
95     return z
96
97 def corsum_matrix(self, m_range, r_range, tau, normalize=True, selfmatches=False,
precision=6):
98     m_counts = max(m_range)
99
100     r_range = numpy.array([r_range]) if type(r_range) in (int,float) else r_range
101     corsum_matrix = numpy.zeros((m_counts, len(r_range)))
102
103     if r_range.size > 1:
104         a = -(r_range[0] - r_range[-1]) / (r_range.size - 1)
105         b = r_range[0]
106
107     if r_range.max() < self.max_diff and r_range.size > 1:
108         raise ValueError("R range maximum has to be greater then the signal max diff")
109
110     for row_n in range(rank, self.N - 1, size):
111         for m_index, row in enumerate(self.windups(m_counts, row_n, tau)):
112             if r_range.size == 1:
113                 corsum_matrix[m_index] += numpy.array([(row ≤ r_range[0]).sum()])
114             else:
115                 corsum_matrix[m_index] += self.r_range_filter(row, r_range.size, a, b)
116         for key in [k for k in self._rows.keys() if k < row_n]:
117             self._rows.pop(key)
118     results = comm.gather(corsum_matrix, root=0)
119     if rank == 0:
120         summed = corsum_matrix = numpy.zeros((m_counts, len(r_range)))
121         for r in results:
122             # we scale up the results by a factor 2 due to
123             # triangularity of the NCM matrix
124             summed += r * 2
125
126         if selfmatches:
127             for row in range(m_counts):
128                 m = row + 1
129                 offset = (self.N - (m - 1) * tau)
130                 summed[row] += offset
131
132         if normalize:
133             for row in range(m_counts):
134                 m = row + 1
135                 factorA = (self.N - (m - 1) * tau)
136                 if selfmatches:
137                     factorB = (self.N - (m - 1) * tau)
138                 else:
139                     factorB = (self.N - 1 - (m - 1) * tau)
140                 factor = (factorA * factorB)
141                 summed[row] /= factor
142
143         return numpy.round(summed.transpose(), precision)
144
145     else:
146         return None
147

```

Rysunek 5.12: Kod programu wersji "NCM.enhanced", część druga

Linie 11–18. Definiowana jest klasa *Matrix*. Ta klasa grupuje wszystkie niezbędne metody do konstruowania i wykorzystywania macierzy NCM: przechowuje oryginalną serię czasową (`self.series`) i oblicza jej podstawo-

we własności, takie jak `self.max_diff` (maksymalna różnica w sygnale) czy `self.N` (liczba próbek w sygnale). Dodatkowo tworzony jest słownik `_rows` na potrzeby pamięci podręcznej obliczanych wierszy.

Linie 20–32. W tej części znajdujemy metodę `windups`, której zadaniem jest iteracyjne wyznaczanie wartości odległości w wyższych wymiarach osadzenia. Zgodnie z koncepcją NCM, aby otrzymać odległość w wymiarze m , łączymy (za pomocą `numpy.vstack` i następnie `max(axis=0)`) kolejne wiersze bazowej macierzy norm, zdefiniowanej w funkcji `get_row`. Odpowiada to operacji

$$n_{i,j}^m = \max_{k=0,\dots,m-1} n_{i+k\tau,j}, \quad (5.48)$$

która w kroku programistycznym realizowana jest iteracyjnie, co umożliwia obsługę różnych wartości m (l. 27–29). Metoda `windups` jest zatem generatorem (*generator*), który zwraca kolejne wartości max-odległości dla poszczególnych wymiarów aż do osiągnięcia końca sygnału.

Linie 34–42. Definicja metody `get_row` odpowiada bezpośrednio pozyskiwaniu wiersza z macierzy NCM o zadanym indeksie `row` i opóźnieniu `tau`. Mechanizm pamięci podręcznej (`_rows`) zapobiega wielokrotnemu przeliczaniu tych samych wierszy. Jeśli danego wiersza nie ma w `_rows`, wywoływana jest metoda `_calc_row` (l. 44) i zwracana wartość zostaje zapisana w `_rows[row]`. Dzięki temu ponowne odwołania do tego samego wiersza `row` będą natychmiast dostępne.

Linie 44–58. Metoda `_calc_row` (l. 48–54) tworzy wiersz macierzy zgodnie z definicją:

$$n_{i,j} = |u(i) - u(i + (j + 1)\tau)|, \quad (5.49)$$

przy czym w kodzie (l. 52) wykorzystano wektorową operację `numpy.abs(self.series[row] - self.series[row + tau::tau])`. Zwracana jest zatem tablica zawierająca różnice wartości `self.series` z kroku τ . Ta implementacja unika ręcznej iteracji po wszystkich indeksach i i w pełni korzysta z operacji macierzowych `numpy`.

Linie 60–67. Metoda `generate_rows` stanowi prosty generator wyliczający wszystkie kolejne wiersze NCM (poczynając od indeksu 0 do `self.N - tau`). Jest to bezpośrednie odzwierciedlenie faktu, że indeks `row` przyjmuje wartości z zakresu $[0, \dots, N - \tau - 1]$. Zamiast przechowywać całą macierz w pamięci, rozwiązanie to pozwala przetwarzać wiersze w sposób strumieniowy (*on demand*), co może ograniczyć zużycie zasobów.

Linie 69–95. W tej części znajdujemy metody `_r_range_filter` i `r_range_filter`, które służą do zliczania, ile spośród wartości w danym wierszu (z uwzględnieniem wymiaru m) spełnia warunek $\leq r$ dla kolejnych progów r . Wersja `_r_range_filter` (l. 62–70) stosuje klasyczne podejście, iterując po kolei po progach i kompresując tablicę. Natomiast `r_range_filter` (l. 73–81) wykorzystuje technikę przeliczania indeksów, czyli

$$\text{indeks} = \frac{(\text{wartość} - b)}{a}, \quad (5.50)$$

a następnie grupowania (`itertools.groupby`), co drastycznie przyspiesza proces porównania elementów z $|r_range|$ progami.

Linie 97–146. Kluczowa metoda `corsum_matrix` zawiera całościową procedurę obliczania sum korelacyjnych z równoległym przetwarzaniem. W l. 90–94 ustalane są `m_counts`, tablica wyników `corsum_matrix` oraz współczynniki (a, b) dla zadanego wektora `r_range`. Linia 96 weryfikuje, czy maksymalny próg `r_range` jest większy od `self.max_diff`, co zapobiega niewłaściwej konfiguracji.

W kolejnej pętli (l. 99–108) skanujemy kolejne wiersze `row.n` metodą `range(rank, self.N - 1, size)`. Jest to podejście *chunkowe*, w którym każdy z `size` procesów przetwarza inny podzbiór indeksów wierszy. Następnie, dla każdej wartości `m_index` (odpowiadającej wymiarowi m) wywoływana jest metoda `windups` i (l. 102–107) zliczane są wartości $\leq r$ przy użyciu albo pojedynczego progu, albo metody `r_range_filter`.

W l. 110–111 zebrane cząstkowe wyniki `corsum_matrix` są przekazywane do procesu `rank=0` (`comm.gather`). W tym miejscu (l. 112–129) proces główny sumuje wyniki i mnoży je przez 2 (linia 117) w celu uwzględnienia symetrii (*trójkątna* część macierzy). Jeśli opcja `selfmatches` jest aktywna (l. 119–125), dodawany jest offset uwzględniający pary identycznych indeksów (na potrzeby np. entropii przybliżonej). Następnie, w l. 127–129, w zależności od ustawienia `normalize`, dokonuje się normalizacji przez liczbę par $\binom{N - (m - 1)\tau}{2}$ lub odpowiednio modyfikowaną przy `selfmatches`. Po transpozycji (l. 130) i zaokrągleniu `precision` (l. 131) wyniki są zwracane z funkcji.

5.5.4 Analiza wydajności różnych implementacji algorytmów do obliczeń sum korelacyjnych

W tej sekcji przedstawione zostały wyniki analizy porównawczej wydajności trzech wersji implementacji algorytmów sum korelacyjnych: metody *brute-force*, metody *NCM.plain* oraz metody *NCM.enhanced*. Do przeprowadzenia testów wygenerowano sygnał bazujący na odwzorowaniu logistycznym.

Odwzorowanie logistyczne jest jednym z najprostszych nieliniowych układów dynamicznych, który jest definiowany równaniem:

$$x_{n+1} = r \cdot x_n \cdot (1 - x_n), \quad (5.51)$$

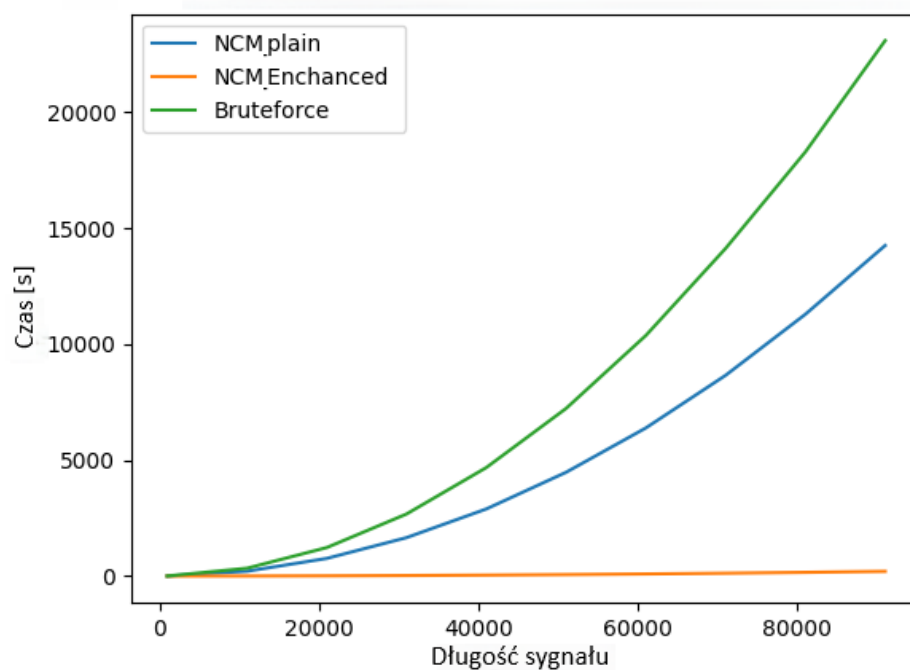
gdzie:

- x_n jest wartością zmiennej w n -tym kroku iteracji,
- r to parametr sterujący ($0 < r \leq 4$),
- $x_n \in [0, 1]$.

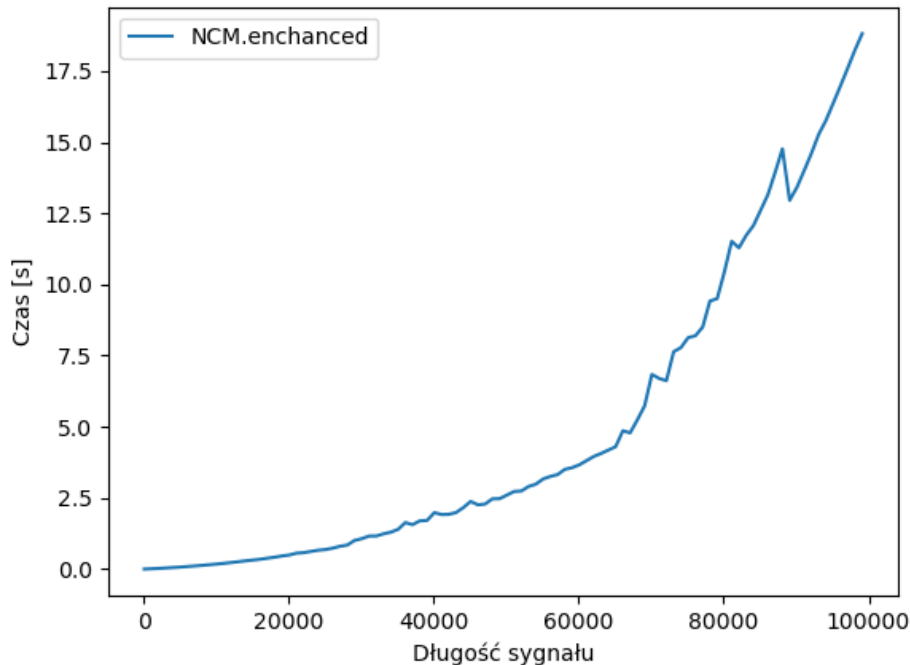
Dla odpowiednio dobranych wartości parametru r , odwzorowanie logistyczne wykazuje złożone zachowania, w tym chaos deterministyczny, co czyni je idealnym sygnałem testowym do analizy algorytmów operujących na danych czasowych.

Analiza wydajnościowa została przeprowadzona w trzech etapach, w których oceniono czas wykonania programu dla każdej z implementacji w zależności od trzech kluczowych parametrów:

1. **Długość sygnału** – w pierwszym etapie badano, jak czas wykonania każdej z metod zmienia się w funkcji długości sygnału N . Testy przeprowadzono dla sygnałów o długościach rosnących w sposób logarytmiczny, aby objąć szeroki zakres wartości.
2. **Liczba wymiarów zanurzeniowych** – w drugim etapie analizowano wpływ liczby wymiarów zanurzeniowych d na czas wykonania. Parametr ten odpowiada za złożoność przestrzeni, w której obliczane są sumy korelacyjne.
3. **Liczba parametrów r** – w trzecim etapie badano, jak liczba różnych wartości parametru promienia r wpływa na wydajność. Każdy z testów zakładał iteracyjną analizę wyników dla różnych promieni, co symuluje praktyczne zastosowania algorytmu w analizie sygnałów.

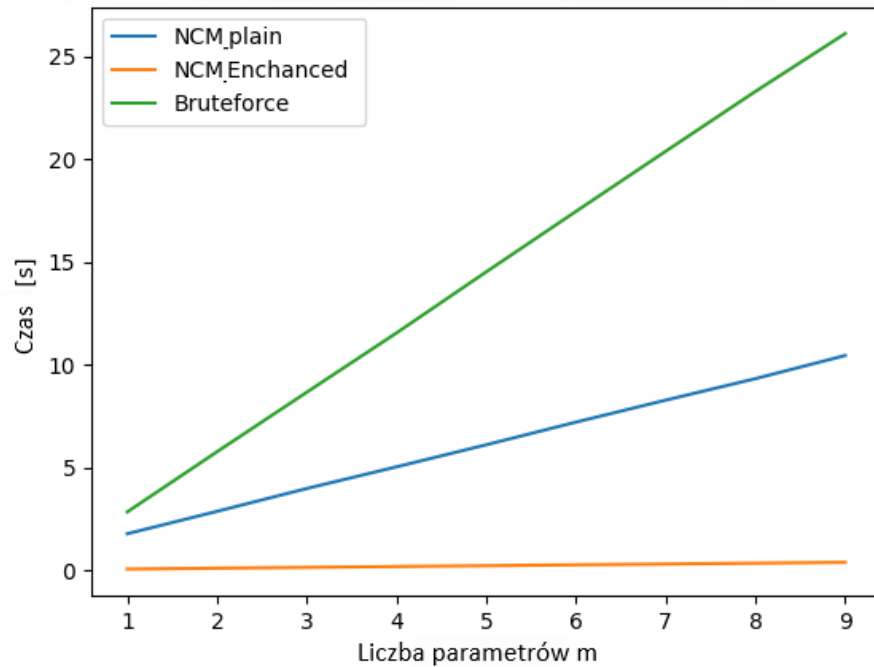


Rysunek 5.13: Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla sygnałów o różnych długościach.



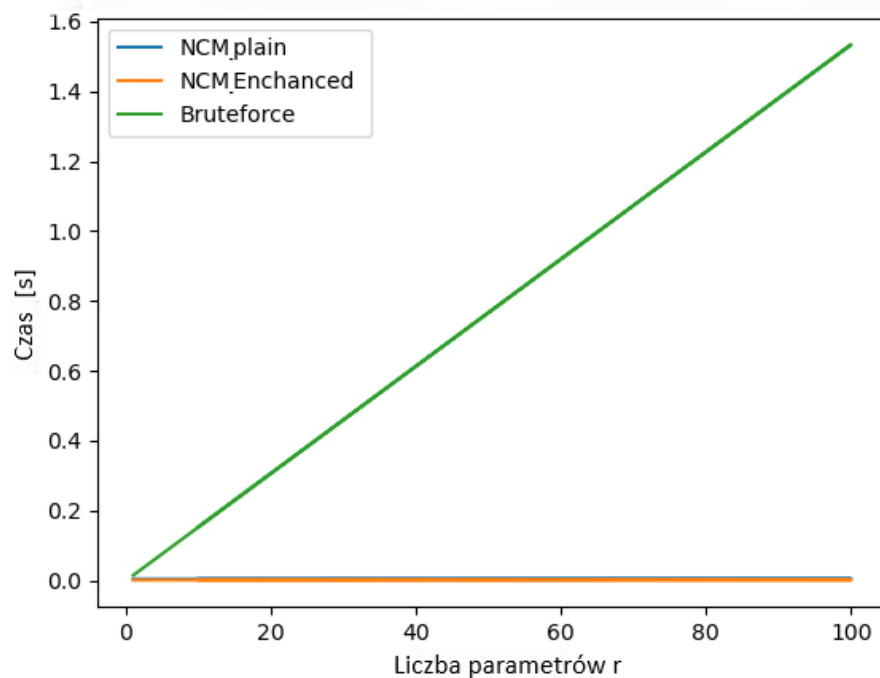
Rysunek 5.14: Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla sygnałów o różnych długościach metodą NCM.enhanced. Widać tutaj kwadratową zależność czasu wykonania programu od długości sygnału czego na wcześniejszym wykresie nie było widać ze względu na olbrzymią dysproporcję między metodą NCM.enhanced a brute-force.

Wykres 5.23 pokazuje zależność czasu wykonania programu od długości sygnału N . Dla wersji *bruteforce* i *NCM plain* zaobserwowano kwadratową zależność czasu od długości sygnału. Wersja *bruteforce* okazała się najwolniejsza, osiągając czas wykonania rzędu 20 000 sekund dla dużych długości sygnału. Wersja *NCM plain* była nieco szybsza, jednak nadal wykazywała tę samą charakterystykę wzrostu. Dla wersji *NCM.enhanced* zaobserwowano znaczną jakościową poprawę. Zależność czasu wykonania od długości sygnału jest nadal kwadratowa ale czas wykonania pozostawał bardzo niski w porównaniu do pozostałych metod, wynosząc zaledwie 20 sekund dla tych samych wartości N . Różnica w czasie wykonania między wersją *NCM.enhanced* a wersją *bruteforce* wynosiła trzy rzędy wielkości, co wskazuje na ogromną efektywność optymalizacji.

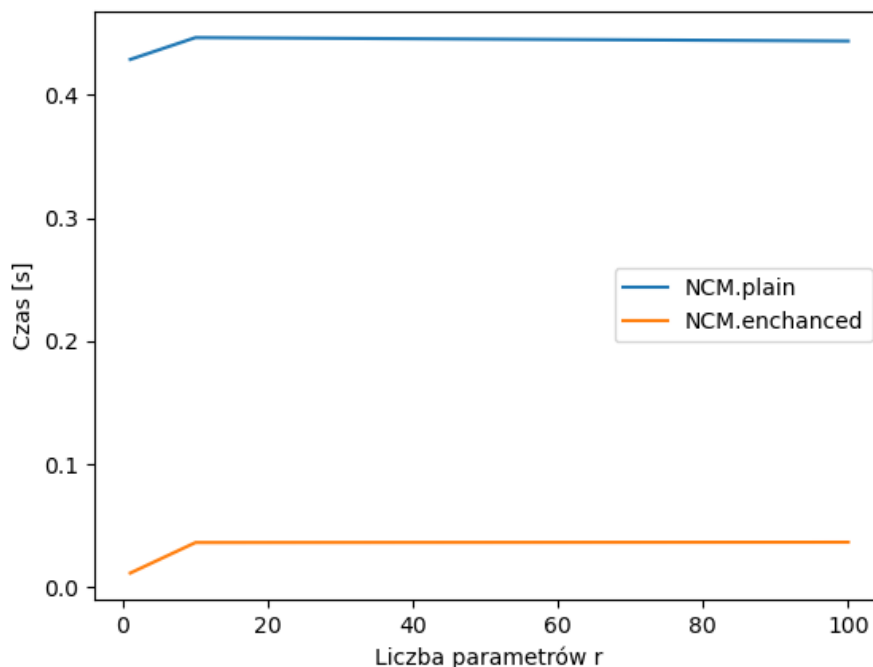


Rysunek 5.15: Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla różnych wartości wymiaru zanurzeniowego m .

Wykres 5.24 przedstawia zależność czasu wykonania programu od liczby wymiarów zanurzeniowych m . Dla wszystkich implementacji zaobserwowano liniową zależność czasu od m . Jednak różnice w wydajności pomiędzy poszczególnymi implementacjami były wyraźne. Wersja *bruteforce* była najwolniejsza, podczas gdy wersja *NCM_enhanced* okazała się najszybsza, co potwierdza przewagę wynikającą z optymalizacji i wykorzystania wektoryzacji.



Rysunek 5.16: Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla różnych wartości parametru r .



Rysunek 5.17: Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla różnych wartości parametru r dla metod NCM.plain i NCM.enhanced. Ten wykres wyraźnie obrazuje że dla obu zbadanych wersji algorytmu czas wykonania programu w zależności od liczby parametrów r jest stały.

Wykres 5.25 ilustruje zależność czasu wykonania programu od liczby parametrów promienia r . W tym przypadku zaobserwowano największą zaletę algorytmu NCM. Dla wersji *bruteforce* czas wykonania wzrastał liniowo wraz ze wzrostem liczby parametrów r . Natomiast dla wersji *NCM plain* i *NCM.enhanced* czas wykonania był praktycznie stały, niezależnie od liczby parametrów r .

Ten wynik ukazuje siłę algorytmu NCM, w którym wydajność nie zmienia się nawet dla dużej liczby parametrów r . Dzięki temu program oparty na algorytmie NCM może być tysiące, a nawet dziesiątki tysięcy razy szybszy w wielu zastosowaniach w porównaniu do wersji *bruteforce*.

Wyniki analizy wydajnościowej jednoznacznie wskazują na przewagę algorytmu NCM, szczególnie w wersji *NCM.enhanced*. Zastosowanie optymalizacji i wektoryzacji pozwala na znaczną redukcję czasu wykonania, czyniąc tę metodę znacznie bardziej efektywną i skalowalną w praktycznych

zastosowaniach.

Wszystkie opisane tutaj wyniki zostały opisane w przygotowanej do recenzji publikacji [93].

Rozdział 6

Wnioski

W niniejszym rozdziale zestawiono wnioski związane z każdą z siedmiu grup celów zdefiniowanych w poprzednim rozdziale. Przedstawione rezultaty prezentują oryginalne wkłady badawcze autora tej rozprawy doktorskiej. Opisane rezultaty zostały opublikowane w czasopismach naukowych, są przygotowane do publikacji, zostały wygłoszone jako referaty na konferencjach naukowych lub zostały opublikowane jako otwartoźródłowy kod dostępny na publicznych repozytoriach oprogramowania.

1. Studium literaturowe wykorzystania HRV, metod złożonościowych oraz HRA w badaniach nad snem

Przeprowadzone studium literaturowe potwierdziło, że:

- Zmienność rytmu serca stanowi istotne źródło informacji na temat architektury snu, zwłaszcza w kontekście detekcji faz REM i NREM oraz oceny aktywności autonomicznego układu nerwowego [110, 111].
- Parametry takie jak *SDNN*, czy wskaźniki widmowe (LF/HF) różnicują kluczowe stadia snu, co zostało szeroko opisane w wielu pracach [112, 113]. Znajdują one zastosowanie w modelach klasyfikacji faz snu i pozwalają wyróżniać wzorce związane z aktywnością przywspółczulną w głębokim śnie oraz aktywnością współczulną w fazie REM.
- Analiza złożoności (np. z wykorzystaniem *SampEn*) zapewnia dodatkową wrażliwość na przejścia między stadami snu i potrafi ujawniać zaburzone wzorce w sygnale EKG, np. u pacjentów z bezdechem sennym (OSA) [123, 124].

- Wprowadzenie koncepcji asymetrii rytmu serca, opartej na zwolnieniach i przyspieszeniach, pozwala szczegółowo opisywać mikrodynamikę rytmu serca i może stanowić użyteczne uzupełnienie tradycyjnych wskaźników HRV [127, 8]. Dotychczasowe wyniki wskazują na wyraźne różnice w długości sekwencji przyspieszeń i zwolnień u osób z obturacyjnym bezdechem sennym [128].
- Rosnąca popularność urządzeń ubieralnych i przenośnych rejestratorów EKG stwarza możliwość nieinwazyjnego oraz niedrogiego monitorowania snu, w tym analizy jedynie z sygnału EKG [42]. Jednocześnie takie uproszczone podejście może w niektórych przypadkach nie dorównywać pełnej polisomnografii, zwłaszcza u osób starszych lub z konkretnymi zaburzeniami kardiologicznymi [129, 130].

Realizacja powyższego celu badawczego, polegającego na przeglądzie piśmiennictwa dotyczącego wykorzystania HRV, metod złożonościowych oraz HRA w analizie snu, pozwoliła wyodrębnić najistotniejsze obszary badawcze oraz istniejące luki w wiedzy. Potwierdzono dużą przydatność parametrów HRV w klasyfikacji faz snu, zwrócono uwagę na wartość miar entropijnych w uzupełnieniu klasycznych wskaźników, a także wskazano na potencjał HRA w pogłębionej charakterystyce rytmu serca. Tym samym cel dotyczący przeprowadzenia studium literaturowego został w pełni zrealizowany, stanowiąc solidną podstawę do dalszych analiz i eksperymentów opisanych w kolejnych rozdziałach. Rezultaty tego studium literaturowego zostały opublikowane w [14]

2. Opisanie oraz przetestowanie, przy użyciu danych syntetycznych, teoretycznego rozkładu serii monotonicznych wraz z uwzględnieniem szumu asymetrycznego

W ramach realizacji tego celu dokonano teoretycznej analizy ogólnego procesu z asymetrycznym szumem oraz wyprowadzono nową, zwartą postać wariancji funkcji Weierstrassa, najczęściej używanej w obliczeniach numerycznych. W trakcie przeglądu literatury nie odnaleziono spójnego wzoru opisującego tę wariancję, dlatego przeprowadzono oryginalne wyprowadzenie, aby sprawdzić, czy rozważania pozostają poprawne po dodaniu asymetrycznego szumu.

Na wstępie zaprezentowano znane rozważania dotyczące wariancji sumy procesów, co pozwoliło ustalić podstawy notacyjne i metodologiczne. Następnie wyprowadzono wzór opisujący wkład asymetrycznego szumu do wariancji procesu opartego na funkcji Weierstrassa. Jest to autorski wkład doktoranta w postaci zamkniętej formy wyrażenia, wcześniej niepublikowanego w dostępnej literaturze. Ostatecznie wzory teoretyczne zostały potwierdzone symulacjami numerycznymi, które wykazały, że przy odpowiednim dobraniu parametrów proces z asymetrycznym szumem może mieć istotnie różną wariancję w zależności od kierunku i skali zaburzenia.

Zestawienie wyników w kontekście wykrywania asymetrii:

- w przypadku czysto deterministycznej funkcji Weierstrassa, bez składowej asymetrycznej, zarówno wariancyjne parametry HRA ($SD1_a, SD1_d$), jak i deskryptory entropiczne serii monotonicznych (HAR, HDR) nie wykazały asymetrii,
- po dodaniu asymetrycznego szumu do sygnału Weierstrassa oba rodzaje deskryptorów wykrywały asymetrię, potwierdzając trafność założeń teoretycznych.

Wyprowadzony wzór na wariancję dostarcza formalnych podstaw do kwantyfikacji wpływu asymetrycznego szumu na serię monotoniczną i stanowi oryginalny wkład doktoranta. Przedstawiona analiza teoretyczna i wyniki eksperymentów wskazują, że zarówno deskryptory wariancyjne, jak i entropiczne serii monotonicznych pozwalają skutecznie wykrywać asymetrię w dynamicznych systemach z niesymetrycznymi zaburzeniami. Cel związany z opisaniem teoretycznego rozkładu serii monotonicznych, zwłaszcza w kontekście asymetrycznych składników sygnału, został tym samym w pełni zrealizowany, a dokonana weryfikacja numeryczna potwierdziła adekwatność proponowanego modelu i deskryptorów. Uzyskane wyniki zostały przygotowane do wysłania do recenzji w czasopiśmie naukowym [152]

3. Zbadanie relacji między ADFA i deskryptorami serii monotonicznych

W porównaniu metod ADFA (z uwzględnieniem asymetrycznych wykładników skalowania) z deskryptorami opartymi na analizie serii monotonicznych stwierdzono znaczącą zbieżność obu podejść w detekcji asymetrii w rytmie serca. Analiza korelacji między wykładnikami α^+ oraz α^- a miarami entropijnymi wyznaczonymi z serii spowolnień i przyspieszeń (DR/AR) wykazała, że:

- Obie metody wrażliwe są na podobne elementy struktury sygnału EKG, zwłaszcza w kontekście długich serii monotonicznych.
- Rozbieżności w wynikach pojawiają się w przypadku krótkich serii monotonicznych (≤ 3), co sugeruje, że długofalowe trendy są lepiej uchwycone przez obie metody niż lokalne, szybsze zmiany rytmu.

Otrzymane rezultaty potwierdziły hipotezę, że pomimo odmiennych formalizmów matematycznych, ADFA i analiza serii monotonicznych opisują zbliżone zjawiska fizjologiczne. Uzyskane wyniki zostały opublikowane w czasopiśmie naukowym [135]

4. Zbadanie predykcyjnych zdolności miar entropijnych serii monotonicznych szeregu RR, we współpracy z pozostałymi miarami zmienności rytmu serca, do detekcji poszczególnych faz snu

W przeprowadzonych badaniach nad detekcją faz snu, parametry entropijne wyznaczone z serii monotonicznych (wspomagane innymi miarami HRV) umożliwiły skuteczną klasyfikację stadium snu. W szczególności:

- Wyodrębnione miary entropijne uwzględniające średnią długość serii spowolnień i przyspieszeń okazały się szczególnie wrażliwe na przejścia między fazami snu (np. N2–N3 czy REM–NREM).
- Modele zawierające parametry asymetryczne szczególnie dobrze sprawdziły się w detekcji głębokich faz snu (N2, N3), stanu czuwania (W) oraz przejść N2–R i R–N2.
- Współpraca tych miar z tradycyjnymi wskaźnikami HRV (takimi jak parametry częstotliwościowe czy czasowe) pozytywnie wpływa na dokładność klasyfikacji.

Przedstawione wyniki potwierdziły przydatność proponowanych deskryptorów w analizie wielokanałowych sygnałów rejestrowanych w polisomnografii, zwiększając możliwości zrozumienia neurofizjologii snu.

5. Zbadanie predykcyjnych zdolności miar entropijnych serii monotonicznych szeregu RR, w połączeniu z innymi miarami zmienności rytmu serca, do detekcji migotania przedsionków

Przedstawione analizy jednoznacznie wskazały, że w detekcji AF (migotania przedsionków) na podstawie krótkich, jednodominutowych zapisów EKG parametry pRR3.25% i pRR30 przewyższają miary oparte na entropii asymetrycznej serii monotonicznych. Najskuteczniejszym wskaźnikiem okazał się pRR3.25%, z bardzo wysoką wartością AUC (0.972) i DOR (643). Kolejne miejsca zajęły pRR30 (AUC = 0.959) oraz HNR (AUC = 0.93).

Wszelkie kombinacje, w których zwiększała się ogólna skuteczność (AUC lub DOR), zawsze obejmowały pRR3.25%, ale łączyły się jednocześnie z zwiększoną liczbą wyników fałszywie ujemnych. Z punktu widzenia ryzyka udaru, wynik fałszywie ujemny niesie poważniejsze konsekwencje niż fałszywie dodatni, co ma znaczenie w praktycznej ocenie wartości parametru diagnostycznego. Analiza dodatkowo wykazała, że parametry entropii asymetrycznej bazujące na seriach monotonicznych (m.in. HNR) są bardziej złożone obliczeniowo i w zastosowaniu do prostej regresji logistycznej nie zapewniają wyższej skuteczności niż pRR3.25%.

Po opublikowaniu głównej części badań [138] zweryfikowano także inny model uczenia maszynowego (XGBoost), który pokazał, że uzupełnienie pRR30 i pRR3.25% parametrem H może istotnie poprawić jakość klasyfikacji AF, obniżając liczbę wyników fałszywie ujemnych. Wskazuje to na istotną rolę metody modelowania w wykorzystaniu entropii asymetrycznej przebiegów neutralnych i potwierdza, że przy doborze odpowiedniego algorytmu miara ta może stanowić wartościowe uzupełnienie pRR3.25% i pRR30 [151].

6. Zbadanie algorytmu do obliczeń sum korelacyjnych oraz jego reimplementacja

W toku badań nad wydajnością obliczeniową parametrów entropijnych zidentyfikowano kluczowy element — sumy korelacyjne, niezbędne do obliczania entropii z próby. Wynikiem tej analizy było:

- wskazanie wąskich gardeł algorytmicznych wpływających na czas ob-

liczeń (szczególnie istotne przy dużej liczbie punktów w szeregu lub przy analizie wielu zapisów jednocześnie),

- reimplementacja algorytmu w języku Python, z refaktoryzacją kodu i wykorzystaniem struktur danych oraz bibliotek numerycznych przyspieszających obliczenia.

Przeprowadzone testy wydajnościowe wykazały zauważalne przyspieszenie obliczeń, co czyni zaproponowane rozwiązanie bardziej praktycznym w zastosowaniach, zwłaszcza w analizie dużych baz danych. Napisane oprogramowanie zostało upublicznione w repozytorium [153]

7. Opracowanie metodologii badawczej służącej przystosowaniu sygnałów dostępnych w technologii komercyjnej do badań naukowych

W kontekście badań nad snem i wykorzystywaniem sygnałów fizjologicznych zarejestrowanych przez komercyjne urządzenia, opracowano i upubliczniono zestaw narzędzi otwartoźródłowych:

- Aplikację do obliczeń serii monotonicznych w języku Python, wraz z funkcjami analizy parametrów entropijnych [154].
- Oprogramowanie do przetwarzania zapisów PSG z urządzeń komercyjnych, umożliwiające zarówno interaktywną analizę pojedynczych nagrań, jak i automatyczne przetwarzanie wielu zapisów w trybie wsadowym [132].

W kontekście algorytmicznym, prace związane z obliczaniem entropii z próby i sum korelacyjnych przełożyły się na refaktoryzację oraz poprawę wydajności narzędzi programistycznych. Przygotowane rozwiązania są dostępne w repozytorium otwartoźródłowym [132]. Działania te umożliwiły skrócenie czasu wstępnej analizy zapisów polisomnograficznych z około 4 godzin do dwóch minut.

Podsumowanie realizacji celów badawczych

Zrealizowane w pracy cele pozwoliły na dogłębną analizę możliwości wykorzystania serii monotonicznych i związanych z nimi miar entropijnych w kontekście rytmu serca, zarówno w aspekcie podstawowym (teoretycznym), jak i aplikacyjnym (analiza snu, detekcja arytmii). Opracowane metody oraz oprogramowanie przyczyniają się do rozwoju narzędzi badawczych łączących podejścia typowe dla HRV, HRA oraz entropijnych analiz z długo- i krótkoterminowymi obserwacjami EKG. Ich otwarty charakter stwarza możliwości dalszego rozwoju i modyfikacji metod analizy na potrzeby konkretnej problematyki badawczej.

Rozdział 7

Podsumowanie

Przedstawiona rozprawa doktorska dotyczyła szeroko pojętej problematyki analizy rytmu serca, ze szczególnym uwzględnieniem roli serii monotonicznych oraz opisującej je entropii asymetrycznej. Ukazano zarówno teoretyczne podstawy zjawisk leżących u podstaw asymetrii rytmu serca, jak i praktyczne możliwości wykorzystania opracowanych metod do detekcji arytmii i klasyfikacji faz snu.

Pierwsze rozdziały stanowiły wprowadzenie do zagadnień związanych z analizą zmienności rytmu serca, metod złożonościowych oraz asymetrii (HRA). Dokonano przeglądu literatury naukowej [14], wskazując kluczowe obszary i luki badawcze w kontekście zastosowań w badaniach snu. Istotnym elementem było przybliżenie podstaw teoretycznych serii monotonicznych wraz z wyprowadzeniem teoretycznego rozkładu tych serii oraz matematycznych własności związanych z ich entropią.

W ramach pracy opracowano model oparty na funkcji Weierstrassa, który pozwolił na teoretyczne opisanie wpływu asymetrycznego szumu na wariancję i właściwości statystyczne sygnału. Wyprowadzono nowy, zwarty wzór wartości wariancji funkcji Weierstrassa w obecności zaburzeń asymetrycznych, wcześniej nieobecny w literaturze. Autorskie wyprowadzenie zostało następnie zweryfikowane numerycznie, a uzyskane wyniki potwierdziły, że zarówno miary wariancyjne (takie jak $SD1_a$ czy $SD1_d$) oraz miary entropijne oparte na seriach monotonicznych (np. HAR , HDR) w sposób spójny i trafny wykrywają asymetrię w rozmaitych dynamicznych procesach [152].

W ramach analiz praktycznych porównano także deskryptory serii monotonicznych z metodą ADFA (uwzględniającą asymetryczne wykładniki skalowania α^+ i α^-). Stwierdzono istotną zbieżność obu podejść w uchwyceniu asymetrycznej struktury rytmu serca, zwłaszcza w obszarze dłuższych serii monotonicznych. Uzyskane rezultaty wskazują, że, pomimo różnych formalizmów matematycznych, metody ADFA oraz analiza serii monoto-

nicznych opisują zbliżone zjawiska fizjologiczne i mogą być wykorzystywane komplementarnie [135].

Przeprowadzone badania aplikacyjne pozwoliły ocenić przydatność opisywanych miar w klasyfikacji faz snu [14] i detekcji migotania przedsionków (AF) [138, 151]. W badaniach nad snem wykazano, iż parametry oparte na entropii serii monotonicznych (wspomagane klasycznymi miarami HRV) stanowią efektywny sposób rozróżniania stadiów (m.in. N2–N3, REM–NREM), co może w przyszłości sprzyjać rozwojowi bardziej dostępnych i nieinwazyjnych metod rejestracji snu. W przypadku AF udowodniono, że deskryptory entropijne do modelu zbudowanego z parametrów pRR3.25% czy pRR30 są w stanie poprawić skuteczność detekcji AF. Ponadto interesującym odkryciem było pokazanie że entropia serii neutralnych sama w sobie jest skutecznym parametrem do detekcji AF.

Dodatkowym atutem pracy jest opracowanie i udostępnienie otwartego oprogramowania umożliwiającego zarówno bardzo szybkie obliczanie sum korelacyjnych oraz związanych z nimi parametrów entropijnych jak *ApEn* i *SampEn* [153], jak i przetwarzanie zapisów polisomnograficznych [132]. Przeprowadzone analizy dla oprogramowania NCM przedstawiają olbrzymi, jakościowy zysk, jeśli chodzi o czas obliczeniowy poświęcony na wykonanie obliczeń sum korelacyjnych nawet dla bardzo długich sygnałów [93]. Podobnie jak implementacja biblioteki pozwala znacząco skrócić czas wstępnej analizy sygnału — w niektórych przypadkach z około czterech godzin do zaledwie dwóch minut. Ma to szczególne znaczenie w praktyce klinicznej i w badaniach na dużych zbiorach danych, gdyż przyspiesza diagnozę i analizę stanów fizjologicznych lub patologicznych.

Dalsze kierunki badań

Uzyskane wyniki otwierają szereg potencjalnych kierunków rozwoju i udoskonaleń:

- **zastosowanie w telemedycynie i urządzeniach mobilnych:** ze względu na coraz szersze wykorzystanie przenośnych rejestratorów EKG i urządzeń ubieralnych, metody analizy serii monotonicznych mogą służyć do ciągłego monitorowania rytmu serca w warunkach domowych lub ambulatoryjnych,
- **poszerzenie analizy o inne sygnały biomedyczne:** analogiczne podejścia można wykorzystać do przetwarzania danych zarejestrowanych w EEG czy EMG. Badanie asymetrii lub innych charakterystyk

monotonii w tych sygnałach może ujawnić nowe aspekty regulacji układów fizjologicznych,

- **wielowymiarowe rozszerzenie serii monotonicznych:** dotychczasowe prace koncentrowały się na jednowymiarowym zapisie RR. Możliwe jest przystosowanie metod do analizy wielokanałowych zapisów EKG albo sygnałów łączonych (np. EKG + sygnał oddechowy), co pozwoliłoby na głębsze poznanie korelacji między różnymi układami fizjologicznymi,
- **integracja z algorytmami sztucznej inteligencji:** wskazano, że w przypadku wykrywania AF różne modele (np. regresja logistyczna, XGBoost) mogą korzystnie lub niekorzystnie wykorzystywać pewne deskryptory entropijne. Rozwój bardziej złożonych algorytmów (np. sieci głębokich) mógłby prowadzić do lepszego wykorzystania informacji zawartych w szeregach monotonicznych,
- **walidacja kliniczna na większych populacjach:** parametry z zakresu entropii serii monotonicznych powinny zostać przetestowane na szerokiej grupie pacjentów, aby ostatecznie ocenić ich wartość diagnostyczną i prognostyczną w detekcji arytmii, zmian faz snu czy innych stanów patologicznych.

Podsumowując, zrealizowane badania łączą teorię modelowania asymetrycznych procesów z praktycznymi zastosowaniami w analizie rytmu serca. Rezultaty przedstawione w rozprawie wskazują na wysoki potencjał deskryptorów serii monotonicznych do wspomagania klasycznych metod HRV, HRA i *SampEn*, a opublikowane oprogramowanie ułatwia ich integrację w badaniach naukowych oraz praktyce klinicznej. Dzięki temu opracowane narzędzia są nie tylko nowym wkładem w dziedzinę nauk o zdrowiu, ale mogą znacząco usprawnić diagnostykę, klasyfikację i monitorowanie stanów fizjologicznych oraz patologicznych.

Spis rysunków

1.1	Przykład schemat pojedynczej ewolucji serca na EKG; źródło [21]	16
1.2	Przykład EKG z migotaniem przedsionków; źródło [32]	18
1.3	Przykład hipnogramu-wykresu przedstawiającego zmiany faz snu oraz diagram kołowy procentowego udziału czasowego danej fazy snu do całości snu; źródło [43]	21
2.1	Przykład wykresu Poincaré odstępów RR uzyskany z całonocnego zapisu EKG. Grafika wyeksportowana z oprogramowania <i>HRAexplorer</i> [53].	25
2.2	Grafika A) przedstawia sposób konstrukcji deskryptorów asymetrii krótkoterminowej, natomiast obrazek B) pokazuje proces budowy deskryptorów asymetrii długoterminowej. W obu przypadkach kluczową rolę odgrywa linia identyczności, która określa, czy dany punkt przyczynia się do części wariancji ($SD1^2$, $SD2^2$ lub $SDNN^2$) związanej ze zwolnieniami lub przyspieszeniami. $r^{\perp;d}128$ oznacza prostopadłą odległość punktu odpowiadającego zwolnieniu o numerze 128 na wykresie Poincaré od linii identyczności, natomiast $r^{\parallel;a}23$ reprezentuje prostopadłą odległość punktu przyspieszenia o numerze 23 od linii l_2 [75]. Grafika na licencji Creative Commons.	34
2.3	Podział fragmentu tachogramu według serii monotonicznych, które oznaczone są jako DRi (seria zwolnień długości i) oraz ARi (seria przyspieszeń długości i). Symbole Ni oznaczają serie neutralne, w których kolejne odstępy RR są równe i które mogą przerywać serie zwolnień i przyspieszeń. Pełne szare punkty oznaczają początki serii zwolnień, a pełne czarne punkty rozpoczynają serie przyspieszeń – są to punkty odniesienia dla tych serii [75]. Grafika na licencji Creative Commons.	39
3.1	Diagram przedstawiający algorytm wstępnego przetwarzania danych holterowskich.	63

3.2	Zrzut ekranu przedstawiający aplikację signalweaver do automatycznej detekcji odstępów RR wraz z możliwością ich ręcznej korekty. Signalweaver stanowi jeden z oryginalnych wkładów mojej pracy, pozwolił na skrócenie wstępnej analizy zapisów polisomnograficznych z około 4 godzin do 2 minut.	65
3.3	Zrzut ekranu przedstawiający aplikację Grass technologies Twin z otwartym oknem do anotowania faz snu. Te dane były potem ręcznie przepisywane ze względu na brak możliwości ich eksportu do formatu tekstowego.	65
3.4	Poglądowe zdjęcie przedstawiające badanie polisomnograficzne. W trakcie badania wykwalifikowany technik ręcznie anotuje fazy snu pacjenta na podstawie mierzonych sygnałów oraz reakcji pacjenta. (źródło: [131])	67
3.5	Zdjęcie przedstawiające polisomnograf Comet PSG Plus XL (Grass Technologies) należący do oddziału Neurologicznego Szpitala Zielonogórskiego w Zielonej Górze.	68
5.1	Wykresy pudełkowe przedstawiające różnice w rozkładach wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa.	80
5.2	Wykresy pudełkowe przedstawiające różnice w rozkładach wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa z asymetrycznym szumem.	82
5.3	Rozkład faz snu i przejść w zapisach poszczególnych osób.	87
5.4	Histogramy różnych typów entropii dla SR i AF	113
5.5	Histogramy pRR30 i pRR3.25% dla SR i AF	114
5.6	Macierze współczynników korelacji Spearmana dla parametrów HRV dla SR i AF. Dolny trójkąt współczynników korelacji, góry trójkąt wartości p	115
5.7	Histogramy przedstawiają wyniki 1000-krotnego bootstrappingu dla modelu zbudowanego na podstawie pRR3.25% w porównaniu do pRR30 oraz pRR3.25% w porównaniu do modelu zbudowanego na podstawie pRR30, pRR3.25% i H . Największa różnica dotyczy liczby FN (fałszywych negatywów), gdzie dodanie parametru H do modelu zwiększa liczbę FN ponad trzykrotnie w porównaniu z modelami bez parametru H	118

5.8	Histogramy przedstawiają wyniki 1000-krotnego bootstrappingu dla modelu zbudowanego na podstawie pRR3.25% w porównaniu do pRR30 oraz pRR3.25% w porównaniu do modelu zbudowanego na podstawie pRR30, pRR3.25% i HAR w oparciu o algorytm XGBoost. Widać tutaj że w przeciwieństwie do modeli opartych o regresję logistyczną, w modelu z H nie ma tutaj problemu ze znacznym wzrostem wyników fałszywie negatywnych.	119
5.9	Kod programu wersji "brute force"	126
5.10	Kod programu wersji "NCM plain"	127
5.11	Kod programu wersji "NCM.enhanced", część pierwsza.	131
5.12	Kod programu wersji "NCM.enhanced", część druga	132
5.13	Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla sygnałów o różnych długościach. . . .	136
5.14	Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla sygnałów o różnych długościach metodą NCM.enhanced. Widać tutaj kwadratową zależność czasu wykonania programu od długości sygnału czego na wcześniejszym wykresie nie było widać ze względu na olbrzymią dysproporcję między metodą NCM.enhanced a brute-force.	137
5.15	Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla różnych wartości wymiaru zanurzeniowego m	138
5.16	Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla różnych wartości parametru r	139
5.17	Wykres przedstawiający czas wykonywania programów do liczenia sum korelacyjnych dla różnych wartości parametru r dla metod NCM.plain i NCM.enhanced. Ten wykres wyraźnie obrazuje że dla obu zbadanych wersji algorytmu czas wykonania programu w zależności od liczby parametrów r jest stały.	140

Spis tabel

2.1	Przykłady monotonicznych serii.	38
-----	---	----

5.1	Tabela wyników testu binomialnego wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa.	79
5.2	Tabela wyników testu binomialnego wartości $SD1a$ i $SD1d$ oraz HAR i HDR dla realizacji funkcji Weierstrassa z asymetrycznym szumem.	81
5.3	Całkowity rozkład faz snu i przejść między nimi w analizowanych zapisach.	87
5.4	Wyniki regresji GEE, faza N1.	88
5.5	Wyniki regresji GEE, faza N2.	89
5.6	Wyniki regresji GEE, przejście N1-N2.	90
5.7	Wyniki regresji GEE, przejście W-N1.	91
5.8	Wyniki regresji GEE, przejście N2-N3.	91
5.9	Wyniki regresji GEE, faza N3.	92
5.10	Wyniki regresji GEE, przejście N3-N2.	92
5.11	Wyniki regresji GEE, przejście N2-R.	93
5.12	Wyniki regresji GEE, przejście R-N2.	93
5.13	Wyniki regresji GEE, faza R.	94
5.14	Wyniki regresji GEE, przejście N3-R.	94
5.15	Wyniki regresji GEE, faza W.	95
5.16	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy N1.	96
5.17	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy N2.	97
5.18	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy N3.	98
5.19	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy R.	99
5.20	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania fazy W.	100
5.21	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N2-N3.	101
5.22	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N2-R.	102
5.23	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N3-N2.	103
5.24	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia N3-R.	104
5.25	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia R-N2.	105
5.26	Wyniki modelowania predykcyjnego: macierz pomyłek i miary skuteczności dla wykrywania przejścia W-N1.	106

5.27	Pole pod krzywą ROC (AUC) i optymalne wartości progowe dla zbioru treningowego – część 1.	116
5.28	Pole pod krzywą ROC (AUC) i optymalne wartości progowe dla zbioru treningowego – część 2.	116
5.29	Porównanie różnych metryk dla detekcji migotania przedsionków (AF) przy użyciu modeli regresji logistycznej jedno- i wielowymiarowej. Trzy najlepsze metryki z najwyższymi wartościami AUC zostały wyróżnione pogrubieniem. Parametr pRR3.25% wyróżnia się jako najskuteczniejsza pojedyncza metryka, osiągając najwyższe wartości AUC. Co istotne, pRR3.25% regularnie pojawia się w najlepiej działających parach i trójkach metryk, wykazując doskonałą skuteczność w rozróżnianiu AF od rytmu zatokowego (SR).	117

Bibliografia

- [1] Edward H. Hon and Stephen T. Lee. Electronic evaluations of the fetal heart rate patterns preceding fetal death, further observations. *American Journal of Obstetrics and Gynecology*, 87(6):814–826, 1965.
- [2] Willem Einthoven. Le télécardiogramme. *Archives Internationales de Physiologie*, 4:132–164, 1906.
- [3] Simon Akselrod, Daniel Gordon, Fred A. Ubel, David C. Shannon, Anthony C. Barger, and Robert J. Cohen. Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Science*, 213(4504):220–222, 1981.
- [4] Fred Shaffer, Rollin McCraty, and Christopher L Zerr. A healthy heart is not a metronome: an integrative review of the heart’s anatomy and heart rate variability. *Frontiers in psychology*, 5:1040, 2014.
- [5] Sylvain Laborde, Emma Mosley, and Julian F Thayer. Heart rate variability and cardiac vagal tone in psychophysiological research—recommendations for experiment planning, data analysis, and data reporting. *Frontiers in psychology*, 8:213, 2017.
- [6] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Heart and Circulatory Physiology*, 283(3):H789–H797, 2002.
- [7] J. Piskorski and P. Guzik. Asymmetric properties of long-term and total heart rate variability. *Medical & biological engineering & computing*, 49:1289–1297, 2011.
- [8] Jaroslaw Piskorski and Przemyslaw Guzik. The structure of heart rate asymmetry: deceleration and acceleration runs. *Phys Meas*, 32:1011–23, 2011.

- [9] Gerhard Hindricks, Tatjana Potpara, Nikolaos Dagres, Elena Arbelo, Jeroen J Bax, Carina Blomström-Lundqvist, Giuseppe Boriani, Manuel Castella, Gheorghe-Andrei Dan, Polychronis E Dilaveris, et al. 2020 esc guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the european association for cardio-thoracic surgery (eacts) the task force for the diagnosis and management of atrial fibrillation of the european society of cardiology (esc) developed with the special contribution of the european heart rhythm association (ehra) of the esc. *European heart journal*, 42(5):373–498, 2021.
- [10] Gregory YH Lip, Allison Keshishian, Xiaoyan Li, Melissa Hamilton, Cristina Masseria, Kiran Gupta, Xuemei Luo, Jack Mardekian, Keith Friend, Anagha Nadkarni, et al. Effectiveness and safety of oral anticoagulants among nonvalvular atrial fibrillation patients: the aristophanes study. *Stroke*, 49(12):2933–2944, 2018.
- [11] Geoffrey H Tison, José M Sanchez, Brandon Ballinger, Avesh Singh, Jeffrey E Olgin, Mark J Pletcher, Eric Vittinghoff, Emily S Lee, Shannon M Fan, Rachel A Gladstone, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA cardiology*, 3(5):409–416, 2018.
- [12] Richard B Berry, Gilbert Hill, Linda Thompson, and Valorea McLaurin. Portable monitoring and autotitration versus polysomnography for the diagnosis and treatment of sleep apnea. *Sleep*, 31(10):1423–1431, 2008.
- [13] Matthew P Walker. The role of sleep in cognition and emotion. *Annals of the New York Academy of Sciences*, 1156(1):168–197, 2009.
- [14] B. Biczuk, S. Żurek, S. Jurga, E. Turska, P. Guzik, and J. Piskorski. Sleep stage classification through hrv, complexity measures, and heart rate asymmetry using generalized estimating equations models. *Entropy*, 26(12), 2024.
- [15] Faith S Luyster, Patrick J Strollo Jr, Phyllis C Zee, and James K Walsh. Sleep: a health imperative. *Sleep*, 35(6):727–734, 2012.
- [16] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, G. B. Moody, C. K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation [Online]*, 101(23):e215–e220, 2000.

- [17] W. Einthoven. Die galvanometrische registrering des menschlichen elektrokardiogramms, zugleich eine beurtheilung der anwendung des capillar-elektrometers in der physiologie. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 99:472–480, 1903.
- [18] Ary L. Goldberger, Zachary D. Goldberger, and A. Shvilkin. *Clinical Electrocardiography: A Simplified Approach*. Elsevier Health Sciences, 9th edition, 2017.
- [19] Galen S. Wagner and Henry J.L. Strauss. *Marriott’s Practical Electrocardiography*. Wolters Kluwer Health/Lippincott Williams & Wilkins, 12th edition, 2014.
- [20] M. Żyliński and G. Cybulski. Selected features for classification of 12-lead ecgs. pages 1–4, 2020.
- [21] <https://en.wikipedia.org/wiki/Electrocardiography> (Dostęp: 27 stycznia 2025).
- [22] Marek Malik and A. John Camm. *Heart Rate Variability*. Futura Publishing Company, 1996.
- [23] Task Force of the European Society of Cardiology, the North American Society of Pacing, and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *European Heart Journal*, 17(3):354–381, 1996.
- [24] M. B. Priestley. *Non-linear and Non-stationary Time Series Analysis*. Academic Press, 1988.
- [25] A. Malliani. Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84(2):482–492, 1991.
- [26] Gari D. Clifford, Francisco Azuaje, and Patrick E. McSharry. *Advanced Methods and Tools for ECG Data Analysis*. Artech House, 2006.
- [27] J. Piskorski and P. Guzik. Filtering poincaré plots. *Computational Methods in Science and Technology*, 11(1):39–48, 2005.
- [28] A.J. Camm, P. Kirchhof, G.Y.H. Lip, U. Schotten, I. Savelieva, S. Ernst, and et al. Guidelines for the management of atrial fibrillation. *European Heart Journal*, 31(19):2369–2429, 2010.

- [29] B.P. Krijthe, A. Kunst, E.J. Benjamin, G.Y.H. Lip, O.H. Franco, A. Hofman, and et al. Projections on the number of individuals with atrial fibrillation in the european union, from 2000 to 2060. *European Heart Journal*, 34(35):2746–2751, 2013.
- [30] C.T. January, L.S. Wann, J.S. Alpert, H. Calkins, J.E. Cigarroa, J.C. Cleveland, and et al. 2014 aha/acc/hrs guideline for the management of patients with atrial fibrillation. *Journal of the American College of Cardiology*, 64(21):e1–e76, 2014.
- [31] D.L. Kasper, A.S. Fauci, S.L. Hauser, D.L. Longo, J.L. Jameson, and J. Loscalzo. *Harrison’s Principles of Internal Medicine*. McGraw-Hill Education, 19th edition, 2015.
- [32] <https://www.ecgedu.com/ecg-interpretation-examples/> (Do-step: 27 stycznia 2025).
- [33] P. Kirchhof, S. Benussi, D. Kotecha, A. Ahlsson, D. Atar, B. Casadei, and et al. 2016 esc guidelines for the management of atrial fibrillation developed in collaboration with eacts. *European Heart Journal*, 37(38):2893–2962, 2016.
- [34] A. Rechtschaffen and A. Kales. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. *Public Health Service, US Government Printing Office*, 1968.
- [35] A.A. Borbély. A two process model of sleep regulation. *Human Neurobiology*, 1(3):195–204, 1982.
- [36] M.A. Carskadon and W.C. Dement. Normal human sleep: an overview. In M.H. Kryger, T. Roth, and W.C. Dement, editors, *Principles and Practice of Sleep Medicine*, pages 13–23. Elsevier Saunders, 4th edition, 2005.
- [37] L. De Gennaro, M. Ferrara, and M. Bertini. The boundary between wakefulness and sleep: quantitative electroencephalographic changes during the sleep onset period. *Neuroscience*, 107(1):1–11, 2001.
- [38] J.M. Siegel. Clues to the functions of mammalian sleep. *Nature*, 437(7063):1264–1271, 2005.
- [39] R. Stickgold, J.A. Hobson, R. Fosse, and M. Fosse. Sleep, learning, and dreams: off-line memory reprocessing. *Science*, 294(5544):1052–1057, 2001.

- [40] C. Iber, S. Ancoli-Israel, A.L. Chesson, and S.F. Quan. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine, 2007.
- [41] J.A. Hobson and E.F. Pace-Schott. The cognitive neuroscience of sleep: Neuronal systems, consciousness and learning. *Nat. Rev. Neurosci.*, 3:679–693, 2002.
- [42] H. Kinnunen, A. Rantanen, T. Kenttä, and H. Koskimäki. Feasible assessment of recovery and cardiovascular health: Accuracy of nocturnal hr and hrv assessed via ring ppg in comparison to medical grade ecg. *Physiological Measurement*, 41:04NT01, 2020.
- [43] <https://weltory.com/stages-of-sleep/> (Dostęp: 27 stycznia 2025).
- [44] P.E. Peppard, T. Young, M. Palta, and J. Skatrud. Prospective study of the association between sleep-disordered breathing and hypertension. *New England Journal of Medicine*, 342(19):1378–1384, 2000.
- [45] L.J. Epstein, D. Kristo, P.J. Jr. Strollo, N. Friedman, A. Malhotra, S.P. Patil, and et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *Journal of Clinical Sleep Medicine*, 5(3):263–276, 2009.
- [46] R.B. Berry, R. Brooks, C.E. Gamaldo, S.M. Harding, R.M. Lloyd, C.L. Marcus, and B.V. Vaughn. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine, version 2.0 edition, 2012.
- [47] Y. Li, Y. Chen, J. Ma, D. Chen, X. Li, D. Wu, Y. Zhao, W. Hu, and H. Hu. Deep learning for sleep stage classification: a survey. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(5):1013–1024, 2018.
- [48] R.H. Shumway and D.S. Stoffer. Time series analysis and its applications. *Springer*, 2006.
- [49] P. Guzik and inni. Poincaré plot analysis of heart rate variability. *Physiological Measurement*, 2007.

- [50] M. Brennan, M. Palaniswami, and P. Kamen. Do existing measures of poincaré plot geometry reflect nonlinear features of heart rate variability? *IEEE Transactions on Biomedical Engineering*, 2001.
- [51] M.P. Tulppo and inni. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. *American Journal of Physiology-Heart and Circulatory Physiology*, 1996.
- [52] M. Brennan, M. Palaniswami, and P. Kamen. Poincaré plot interpretation using a physiological model of hrv based on a network of oscillators. *American Journal of Physiology-Heart and Circulatory Physiology*, 2002.
- [53] J. Piskorski. Hraexplorer (version 1.54), 2024. Available online: <https://github.com/jaropis/HRAexplorer> (accessed on 19 November 2024).
- [54] Kuo K.J. Liu C. Improved pnnx statistics as measures of heart rate variability in the presence of ectopic beats and noise. *IEEE Transactions on Biomedical Engineering*, 60(1):335–345, 2013.
- [55] Bruce Pomeranz, Robert J B Macaulay, Margaret A Caudill, et al. Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology*, 248(1):H151–H153, 1985.
- [56] Alberto Malliani, Massimo Pagani, Federico Lombardi, and Sergio Cerutti. Power spectral analysis of heart rate variability: a tool to explore neural regulatory mechanisms. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 348(1688):93–102, 1994.
- [57] Steven M Kay. *Modern Spectral Estimation: Theory and Application*. Prentice Hall, 1988.
- [58] E Oran Brigham. *The Fast Fourier Transform and its Applications*. Prentice Hall, 1988.
- [59] S Lawrence Marple Jr. *Digital Spectral Analysis: with Applications*. Prentice-Hall, 1987.
- [60] James A Taylor, Daniel L Carr, Charles W Myers, and Dwain L Eckberg. Prognostic value of heart rate variability after myocardial infarction in patients with left ventricular dysfunction. *Circulation*, 98(10):953–960, 1998.

- [61] Gari D Clifford, Francisco Azuaje, and Patrick E McSharry. *Advanced Methods and Tools for ECG Data Analysis*. Artech House, 2006.
- [62] N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39(2):447–462, 1976.
- [63] J. D. Scargle. Studies in astronomical time series analysis. ii – statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853, 1982.
- [64] P. Laguna, G. B. Moody, and R. Mark. Power spectral density of unevenly sampled data by least-square analysis. *IEEE Transactions on Biomedical Engineering*, 45(6):698–715, 1998.
- [65] M. Aboy, T. Thong, J. McNames, and B. Oken. Averaged lomb periodograms for nonuniform sampling. *Proceedings of [include the journal/conference if available]*, 2004.
- [66] Jarosław Piskorski, Przemysław Guzik, Tomasz Krauze, and Sebastian Żurek. Cardiopulmonary resonance at 0.1 hz demonstrated by averaged lomb-scargle periodogram. *Central European Journal of Physics*, 8:386–392, 04 2010.
- [67] C.K. Peng, S. Havlin, H.E. Stanley, and A.L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5:82–87, 1995.
- [68] K. Hu, P.Ch. Ivanov, Z. Chen, P. Carpena, and H.E. Stanley. Effect of trends on detrended fluctuation analysis. *Phys. Rev. E*, 64:011114, 2001.
- [69] Z. Chen, P.C. Ivanov, K. Hu, and H.E. Stanley. Effect of nonstationarities on detrended fluctuation analysis. *Phys. Rev. E*, 65:041107, 2002.
- [70] J. Piskorski and P. Guzik. Filtering Poincaré plots. *Comp Meth Sci Tech*, 11:39–46, 2005.
- [71] J.A. Ramirez, E. Rodrigues, and J.C. Echeverria. A dfa approach for assessing asymmetric correlations. *Physica A*, 388:2263–2270, 2009.
- [72] J. Piskorski, M. Kosmider, D. Mieszkowski, T. Krauze, A. Wykretowicz, and P. Guzik. Properties of asymmetric detrended fluctuation analysis in the time series of RR intervals. *in preparation*, 2017.

- [73] M.A. Rivera-Castroa, J.G.V. Mirandaa, D.O. Cajueirob, and R.F.S. Andradea. Detecting switching points using asymmetric detrended fluctuation analysis. *Physica A*, 391:170–179, 2012.
- [74] D. Kokosińska, J. J. Żebrowski, T. Buchner, R. Baranowski, and E. Orłowska-Baranowska. Asymmetric multiscale multifractal analysis (amma) of heart rate variability. *Physiological Measurement*, 42(8):085003, 2021.
- [75] Jarosław Piskorski. *Struktura asymetrii rytmu serca*. Wydawnictwo Naukowe Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu, Poznań, 2011.
- [76] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.
- [77] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, NJ, 2 edition, 2006.
- [78] Steven M. Pincus. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6):2297–2301, 1991.
- [79] Jarosław Piskorski and Przemysław Guzik. Geometry of the Poincaré plot of RR intervals and its asymmetry in healthy adults. *Phys Meas*, 28:287–300, 2007.
- [80] H. Levene and J. Wolfowitz. The covariance matrix of runs up and down. *The Annals of Mathematical Statistics*, 15(1):58–69, 1944.
- [81] J. Wolfowitz. Asymptotic Distribution of Runs Up and Down. *Annals of Mathematical Statistics*, 15(2):163–172, 1944.
- [82] Pincus S.M. Approximate entropy as a mesusre of system compexity. *Proc. Natl. Acad. Sci. USA*, 88:2297–2301, 1991.
- [83] Pincus S.M., Gladstone I.M., and Ehrenkranz R.A. A regularity statistic for medical data analysis. *J. Clin. Monit.*, 7:335–345, 1991.
- [84] Pincus S.M. Approximate entropy (apen) as a complexity measure. *Chaos*, 5:110–117, 1994.
- [85] J.S. Richman and J.R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.*, 278:H2039–H2049, 2000.

- [86] S. Żurek, P. Guzik, S. Pawlak, M. Kośmider, and J. Piskorski. On the relation between correlation dimension, approximate entropy and sample entropy parameters, and a fast algorithm for their calculation. *Physica A: Statistical Mechanics and its Applications*, 391(24):6601–6610, 2012.
- [87] Sassi R., Cerutti S., Lombardi F., Malik M., Huikuri H.V., Peng C.K., Schmidt G., and Yamamoto Y. Advances in heart rate variability signal analysis: joint position statement by the e-cardiology esc working group and the european heart rhythm association co-endorsed by the asia pacific heart rhythm society. *Europace*, 17:1341–1353, 2015.
- [88] Kumar M., Pachori R.B., , and Acharya U.R. Automated diagnosis of myocardial infarction ecg signals using sample entropy in flexible analytic wavelet transform framework. *Entropy*, 19:488, 2017.
- [89] Buszko K., Piątkowska A., Koźluk E., and Opolski G. Entropy in investigation of vasovagal syndrome in passive head up tilt test. *Entropy*, 19:236, 2017.
- [90] Buszko K., Piątkowska A., Koźluk E., Fabiszak T., and Opolski G. The complexity of hemodynamic response to the tilt test with and without nitroglycerine provocation in patients with vasovagal syncope. *Scientific Reports*, 8:14554, 2018.
- [91] Buszko K., Piątkowska A., Koźluk E., Fabiszak T., and Opolski G. Entropy measures in analysis of head up tilt test outcome for diagnosing vasovagal syncope. *Entropy*, 20:976, 2018.
- [92] B. Graff, G. Graff, and A. Kaczkowska. Entropy measures of heart rate variability for short ecg datasets in patients with congestive heart failure. *Acta Physica Polonica B, Proceedings Supplement*, 5, 01 2012.
- [93] S. Żurek, B. Biczuk, P. Guzik, and J. Piskorski. Ncm: A python package for fast calculation of correlation sums, approximate entropy, and sample entropy. *SoftwareX*, 2025. In progress.
- [94] David W Hosmer, Stanley Lemeshow, and Rodney X Sturdivant. *Applied Logistic Regression*. John Wiley & Sons, 2013.
- [95] Pierre-François Verhulst. Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique*, 10:113–121, 1838.

- [96] Joseph Berkson. Application of the logistic function to bio-assay. *Journal of the American Statistical Association*, 39(227):357–365, 1944.
- [97] David M W Powers. *Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation*, volume 2. 2011.
- [98] Andrew Y Ng and Michael I Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in Neural Information Processing Systems*, pages 841–848, 2002.
- [99] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.
- [100] Sandro Sperandei. Understanding logistic regression analysis. *Biochemia Medica*, 24(1):12–18, 2014.
- [101] Ewout W Steyerberg. *Clinical Prediction Models*. Springer, 2009.
- [102] Kung-Yee Liang and Scott L Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.
- [103] James W Hardin and Joseph M Hilbe. *Generalized Estimating Equations*. Chapman and Hall/CRC, 2003.
- [104] Scott L Zeger and Kung-Yee Liang. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, 44(4):1049–1060, 1988.
- [105] James A Hanley, Abdissa Negassa, Michael D deB Edwardes, and Joseph E Forrester. Statistical analysis of correlated data using generalized estimating equations: an orientation. *American Journal of Epidemiology*, 157(4):364–375, 2003.
- [106] Wei Pan. Akaike’s information criterion in generalized estimating equations. *Biometrics*, 57(1):120–125, 2001.
- [107] Anil K Jain, M Narasimha Murty, and Patrick J Flynn. Data clustering: a review. *ACM Computing Surveys (CSUR)*, 31(3):264–323, 1999.
- [108] Charu C Aggarwal and Chandan K Reddy. *Data Clustering: Algorithms and Applications*. CRC Press, 2014.

- [109] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794, 2016.
- [110] Task Force of the European Society of Cardiology, the North American Society of Pacing, and Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93:1043–1065, 1996.
- [111] P.K. Stein and Y. Pu. Heart rate variability, sleep and sleep disorders. *Sleep Medicine Reviews*, 16:47–66, 2012.
- [112] M.H. Bonnet and D.L. Arand. Heart rate variability: Sleep stage, time of night, and arousal influences. *Electroencephalography and Clinical Neurophysiology*, 102:390–396, 1997.
- [113] O. Faust, H. Razaghi, R. Barika, E.J. Ciaccio, and U.R. Acharya. A review of automated sleep stage scoring based on physiological signals for the new millennia. *Comput. Methods Programs Biomed.*, 176:81–91, 2019.
- [114] N. Surantha, T.F. Lesmana, and S.M. Isa. Sleep stage classification using extreme learning machine and particle swarm optimization for healthcare big data. *Journal of Big Data*, 8:14, 2021.
- [115] M. Xiao, H. Yan, J. Song, Y. Yang, and X. Yang. Sleep stages classification based on heart rate variability and random forest. *Biomed. Signal Process. Control*, 8:624–633, 2013.
- [116] Y. Ichimaru, K.P. Clark, J. Ringler, and W.J. Weiss. Effect of sleep stage on the relationship between respiration and heart rate variability. In *Proceedings of the 1990 Computers in Cardiology*, pages 657–660, Chicago, IL, USA, 1990. IEEE.
- [117] D. Herzig, P. Eser, X. Omlin, R. Riener, M. Wilhelm, and P. Achermann. Reproducibility of heart rate variability is parameter and sleep stage dependent. *Front. Physiol.*, 8:1100, 2018.
- [118] M. Tsunoda, T. Endo, S. Hashimoto, S. Honma, and K.I. Honma. Effects of light and sleep stages on heart rate variability in humans. *Psychiatry and Clinical Neurosciences*, 55:285–286, 2001.
- [119] C. Yan, P. Li, Y. Li, J. Li, and C. Liu. Analysis of heart rate asymmetry during sleep stages. *Proceedings of CECNet 2021*, pages 731–736, 2021.

- [120] P. Fonseca, X. Long, M. Radha, R. Haakma, R.M. Aarts, and J. Rolink. Sleep stage classification with ecg and respiratory effort. *Physiol. Meas.*, 36:2027–2040, 2015.
- [121] X. Liang, J. Xiong, Z. Cao, X. Wang, J. Li, and C. Liu. Decreased sample entropy during sleep-to-wake transition in sleep apnea patients. *Physiol. Meas.*, 42:044001, 2021.
- [122] C. Yan, P. Li, M. Yang, Y. Li, J. Li, H. Zhang, and C. Liu. Entropy analysis of heart rate variability in different sleep stages. *Entropy*, 24:379, 2022.
- [123] W.Y. Pan, M.C. Su, H.T. Wu, M.C. Lin, I.T. Tsai, and C.K. Sun. Multiscale entropy analysis of heart rate variability for assessing the severity of sleep disordered breathing. *Entropy*, 17:231–243, 2015.
- [124] H.M. Al-Angari and A.V. Sahakian. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *IEEE Trans. Biomed. Eng.*, 54:1900–1904, 2007.
- [125] C.C. Chung, J.H. Kang, and C.J. Hu. Measuring entropy in functional neuroscience: Pathophysiological and clinical applications. *Neurosci. Neuroecon.*, 5:45–53, 2016.
- [126] M. Aktaruzzaman, M. Migliorini, M. Tenhunen, S.L. Himanen, A.M. Bianchi, and R. Sassi. The addition of entropy-based regularity parameters improves sleep stage classification based on heart rate variability. *Med. Biol. Eng. Comput.*, 53:415–425, 2015.
- [127] P. Guzik, J. Piskorski, K. Awan, T. Krauze, M. Fitzpatrick, and A. Baranchuk. Obstructive sleep apnea and heart rate asymmetry microstructure during sleep. *Clin. Auton. Res.*, 23:91–100, 2013.
- [128] J. Jiang, X. Chen, Ch. Zhang, G. Wang, J. Fang, J. Ma, and J. Zhang. Heart rate acceleration runs and deceleration runs in patients with obstructive sleep apnea syndrome. *Sleep Breath*, 2016. doi:10.1899/15-9834.6.
- [129] M. de Zambotti, J. Trinder, A. Silvani, I.M. Colrain, and F.C. Baker. Dynamic coupling between the central and autonomic nervous systems during sleep: A review. *Neurosci. Biobehav. Rev.*, 90:84–103, 2018.
- [130] T. Penzel, N. Wessel, M. Riedl, J.W. Kantelhardt, S. Rostig, M. Glos, A. Suhrbier, H. Malberg, and I. Fietze. Cardiovascular and respiratory dynamics during normal and pathological sleep. *Chaos*, 17, 2007.

- [131] <https://www.kokilabenhospital.com/departments/clinicaldepartments/pulmonarymedicine/polysomnographysleepstudy.html> (Dostęp: 27 stycznia 2025).
- [132] J. Piskorski. signalweaver, 2022. Available online: <https://github.com/jaropis/signalweaver> (accessed on 19 November 2024).
- [133] K. Weierstrass. Über continuirliche functionen eines reellen arguments, die für keinen werth des letzteren einen bestimmten differentialquotienten besitzen (on continuous functions of a real argument that do not have a well-defined differential quotient). In G.A. Edgar, editor, *Classics on Fractals*, pages 3–9. Addison-Wesley Publishing Company, Boston, 1993.
- [134] J. Piskorski, M. Kosmider, D. Mieszkowski, T. Krauze, A. Wykretowicz, and P. Guzik. Properties of asymmetric detrended fluctuation analysis in the time series of rr intervals. *Physica A: Statistical Mechanics and its Applications*, 491:347–360, 2018.
- [135] J. Piskorski, M. Kośmider, D. Mieszkowski, S. Żurek, B. Biczuk, S. Jurga, T. Krauze, A. Wykretowicz, and P. Guzik. Associations between heart rate asymmetry expression and asymmetric detrended fluctuation analysis results. *Medical & Biological Engineering & Computing*, 60(10):2969–2979, October 2022. Epub 2022 Aug 24.
- [136] Hana Krakovská and Anna Krakovská. Fractal dimension of self-affine signals: Four methods of estimation, 2016.
- [137] Matthew Nelson, Alese C. Wooditch, and Lisa M. Dario. Sample size, effect size, and statistical power: a replication study of weisburd’s paradox. *Journal of Experimental Criminology*, 11:141–163, 2015.
- [138] B. Biczuk, S. Bus, S. Żurek, J. Piskorski, and P. Guzik. prr30, prr3.25% and asymmetrical entropy descriptors in atrial fibrillation detection. *Entropy*, 26:296, 2024.
- [139] J.W. Shepard. Hypertension, cardiac arrhythmias, myocardial infarction, and stroke in relation to obstructive sleep apnea. *Clin. Chest Med.*, 13:437–458, 1992.
- [140] K. Kario and et al. Disrupted sleep and nocturnal cardiac events in hypertensive patients. *Hypertension*, 41(6):1313–1320, 2003.

- [141] M. Nijima, H. Kimura, H. Edo, T. Shinozaki, J. Kang, S. Masuyama, K. Tatsumi, and T. Kuriyama. Manifestation of pulmonary hypertension during rem sleep in obstructive sleep apnea syndrome. *Am. J. Respir. Crit. Care Med.*, 159:1766–1772, 1999.
- [142] S. Buś, K. Jędrzejewski, and P. Guzik. Statistical and diagnostic properties of prrx parameters in atrial fibrillation detection. *Journal of Clinical Medicine*, 11(19), 2022.
- [143] S. Buś, K. Jędrzejewski, and P. Guzik. Impact of electrocardiogram length on diagnostic properties of heart rate variability indices in atrial fibrillation detection. In *Proceedings of the 2022 12th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 1–2, Vysoké Tatry, Slovakia, 2022.
- [144] S. Buś, K. Jędrzejewski, and P. Guzik. A new approach to detecting atrial fibrillation using count statistics of relative changes between consecutive rr intervals. *Journal of Clinical Medicine*, 12(2), 2023.
- [145] C. Liu, J. Oster, E. Reinertsen, Q. Li, L. Zhao, S. Nemati, and G.D. Clifford. A comparison of entropy approaches for af discrimination. *Physiol. Meas.*, 39:074002, 2018.
- [146] L. Zhao, C. Liu, S. Wei, Q. Shen, F. Zhou, and J. Li. A new entropy-based atrial fibrillation detection method for scanning wearable ecg recordings. *Entropy*, 20:904, 2018.
- [147] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, 1948.
- [148] A.A. Renshaw and E.W. Gould. Reducing false-negative and false-positive diagnoses in anatomic pathology consultation material. *Arch. Pathol. Lab. Med.*, 137:1770–1773, 2013.
- [149] M.S. Klinkman, J.C. Coyne, S. Gallo, and T.L. Schwenk. False positives, false negatives, and the validity of the diagnosis of major depression in primary care. *Arch. Fam. Med.*, 7:451–461, 1998.
- [150] L. Breiman. Random forests. *Mach. Learn.*, 45:5–32, 2001.
- [151] B. Biczuk. Metody uczenia maszynowego w oparciu o entropie z serii monotonicznych prr30 do detekcji migotania przedsionków. In J. Wiśniewski i in., editor, *Nowoczesne metody statystyczne w badaniach medycznych*; ISBN: 9788396802002, page 21, Bydgoszcz, Polska, 2023. ISCB, UMK Wydział Farmaceutyczny CM, Bydgoszcz.

- [152] B. Biczuk, S. Żurek, and J. Piskorski. Relation between variance-based and run-based heart rate asymmetry descriptors on the weierstrass function with asymmetric noise. *Entropy*, 2025. In progress.
- [153] S. Żurek, B. Biczuk, K. Wojtiuk, and M. Kośmider. Ncm: A python package for fast calculation of correlation sums, approximate entropy, and sample entropy. <https://github.com/sebzur/NCM-algorithm>, 2025. Dostęp: 27 stycznia 2025.
- [154] B. Biczuk. monotonic_runs: A python library for detecting monotonic runs in time series data. https://github.com/dzwiedzn7/monotonic_runs, 2025. Dostęp: 27 stycznia 2025.