



**AGH**

AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

**DZIEDZINA NAUK INŻYNIERYJNO-TECHNICZNYCH**

DYSCYPLINA AUTOMATYKA, ELEKTRONIKA, ELEKTROTECHNIKA I  
TECHNOLOGIE KOSMICZNE

# **AUTOREFERAT ROZPRAWY DOKTORSKIEJ**

Generatywne modele procesów Gaussowskich w analizie  
zjawisk funkcjonalnych i wykrywaniu usterek

Autor: Adrian Dudek

Promotor rozprawy: Prof. dr hab. inż. Jerzy Baranowski

Kraków, 2025

## Wstęp

Moja praca dotyczy zastosowania *Procesów Gaussowskich* (PG) do analizy złożonych zjawisk funkcjonalnych i diagnostyki systemów technicznych. Główne tezy pracy wskazują, że PG pozwalają skutecznie modelować nieliniowe zależności, dostarczając jednocześnie kwantyfikację niepewności, a także że modele generatywne PG mogą kompensować brak danych, tworząc syntetyczne zbiory do diagnostyki.

Skupiłem się na zastosowaniu PG w detekcji anomalii w systemach przemysłowych, gdzie niepewność predykcji odgrywa kluczową rolę w ocenie stanów przejściowych. Ponadto, badałem możliwości wykorzystania PG do generowania wiarygodnych danych w sytuacjach ograniczonej dostępności rzeczywistych pomiarów. W celu poprawy efektywności obliczeniowej zastosowałem aproksymacje Czebyszewa oraz szybką transformację Fouriera (FFT), co pozwoliło na redukcję kosztów obliczeniowych przy zachowaniu jakości wyników.

Zastosowania PG obejmowały m.in. prognozowanie zużycia baterii litowo-jonowych, modelowanie przestrzennych zmian poziomu zanieczyszczeń oraz diagnostykę anomalii w sygnałach przemysłowych. Wyniki pokazują, że PG mogą być skutecznym narzędziem w szeroko pojętej diagnostyce technicznej, oferując interpretowalne modele o wysokiej precyzji predykcji.

## Motywacja

Nowoczesny przemysł opiera się na technologiach zwiększających efektywność i automatyzację procesów produkcyjnych. Maszyny i urządzenia, takie jak silniki elektryczne, pompy czy systemy chłodzenia, odgrywają kluczową rolę w zapewnianiu ciągłości operacyjnej. Niestety, wraz z eksploatacją następuje ich stopniowe zużycie, co prowadzi do awarii, przestojów oraz wzrostu kosztów operacyjnych. Choć producenci rzadko dostarczają narzędzia do autodiagnostyki, rozwój metod analitycznych pozwala na lepsze monitorowanie stanu technicznego urządzeń, przewidywanie potencjalnych usterek i minimalizację strat produkcyjnych.

Tradycyjne podejścia w diagnostyce technicznej, takie jak regresja liniowa czy modele ARMA, zakładają stabilność systemu w określonych warunkach (Box and Jenkins 1970). Metody te sprawdzają się w prognozowaniu liniowych szeregów czasowych, lecz napotykają trudności w analizie niestacjonarnych i nieliniowych układów, które są typowe dla wielu systemów przemysłowych. Gdy pojawiają się anomalie lub nagłe zmiany w danych, skuteczność klasycznych modeli znacząco spada, co wskazuje na potrzebę bardziej elastycznych metod analizy.

Wprowadzenie technik uczenia maszynowego umożliwiło analizę dużych zbiorów danych i identyfikację ukrytych wzorców. Algorytmy takie jak lasy losowe czy sieci neuronowe (Bishop 1995) oferują wysoką skuteczność w przewidywaniu nieliniowych zależności, co znajduje potwierdzenie w badaniach (Bikku 2020). Ich największym ograniczeniem pozostaje jednak trudność w interpretacji wyników – wiele modeli działa jako „czarne skrzynki”, utrudniając zrozumienie wpływu poszczególnych zmiennych na prognozy. Dodatkowo, większość metod nie zapewnia mechanizmów oceny niepewności predykcji, co może prowadzić do ryzykownych decyzji w systemach wymagających wysokiej niezawodności.

PG stanowią nowoczesne podejście do modelowania systemów technicznych, łącząc zdolność analizy złożonych nieliniowych zależności z probabilistycznym szacowaniem niepewności. Jak wskazano w przeglądzie literatury (Dudek and Baranowski 2022), PG znajdują coraz szersze zastosowanie w diagnostyce technicznej, a ich wykorzystanie w analizie funkcjonalnej i modelowaniu systemów nieliniowych jest obszarem intensywnych badań. Ich probabilistyczny charakter pozwala na precyzyjne wykrywanie anomalii oraz generowanie syntetycznych danych, które mogą wspomagać inne algorytmy diagnostyczne, zwłaszcza w sytuacjach, gdzie rzeczywiste pomiary są ograniczone lub trudno dostępne (Gonzalvez, Lezmi, Roncalli, and Xu 2019).

W ramach projektu badawczego NCN „Prognozowanie i wykrywanie błędów procesowych” skupiłem się na detekcji błędów w procesach przemysłowych, gdzie PG umożliwiają nie tylko predykcję, ale również identyfikację rzadkich i nieoczekiwanych zdarzeń. Dzięki probabilistycznej naturze PG możliwe jest nie tylko przewidywanie przyszłego zachowania systemu, ale także oszacowanie poziomu niepewności, co jest kluczowe dla poprawy interpretowalności wyników oraz podejmowania świadomych decyzji diagnostycznych.

Zdolność PG do modelowania skomplikowanych zależności oraz generowania danych sprawia, że są one konkurencyjnym narzędziem w diagnostyce systemów technicznych. Ich zastosowanie w przemyśle może przyczynić się do lepszego monitorowania procesów, minimalizacji ryzyka awarii oraz poprawy efektywności operacyjnej, co czyni je obiecującym rozwiązaniem dla nowoczesnej inżynierii predykcyjnej.

## Tezy

Celem rozprawy jest opracowanie kompleksowego rozwiązania w zakresie diagnostyki danych i wykrywania usterek, z uwzględnieniem analizy funkcjonalnej oraz problemu niedoboru danych w modelowaniu poprzez wykorzystanie generatywnych PG. W rezultacie sformułowano następujące tezy:

**Teza 1: Procesy Gaussowskie mogą wypełnić lukę w analizie i diagnostyce zjawisk funkcjonalnych.** PG umożliwiają modelowanie złożonych nieliniowych zależności, których tradycyjne metody często nie uchwytują. Dzięki funkcjom jądrowym mogą elastycznie odwzorowywać relacje między danymi, a ich probabilistyczna natura pozwala na kwantyfikację niepewności. Ta cecha jest kluczowa w zastosowaniach wymagających wysokiej precyzji diagnostycznej. Ponadto, PG oferują interpretowalne wyniki, co odróżnia je od metod działających na zasadzie „czarnej skrzynki”. Łącząc zdolności predykcyjne z transparentnością, PG stanowią wartościowe narzędzie w diagnostyce funkcjonalnej.

**Teza 2: Generatywne modele procesów Gaussowskich mogą być użyteczne w diagnostyce w warunkach niedoboru danych.** W diagnostyce technicznej często występuje ograniczona dostępność danych, co obniża skuteczność klasycznych metod. Generatywne modele PG pozwalają na tworzenie syntetycznych danych, zachowując kluczowe zależności funkcjonalne, co poprawia działanie algorytmów diagnostycznych. Dzięki temu można zwiększyć odporność modeli na niedobór danych bez ryzyka utraty jakości prognoz. To szczególnie istotne w zastosowaniach, gdzie pozyskiwanie rzeczywistych danych jest kosztowne lub trudne. Dalsze badania nad efektywnością obliczeniową PG są jednak konieczne dla ich skutecznego wdrożenia w systemach przemysłowych.

Sformułowane tezy stanowią podstawę do zrozumienia potencjału PG w diagnostyce i wzbogacaniu zbiorów danych, szczególnie w kontekstach, gdzie dane są rzadkie lub niepełne. Mimo że modele te oferują obiecujące rozwiązania, ich praktyczne wdrożenie wiąże się z wieloma wyzwaniami, które należy rozwiązać.

## Problemy badawcze

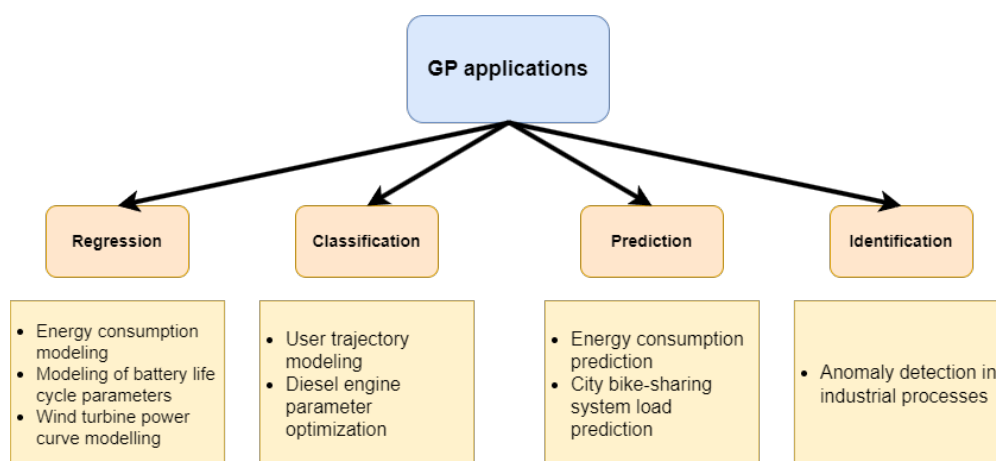
W kolejnych sekcjach zdefiniowano sześć kluczowych problemów badawczych, których rozwiązanie wspiera sformułowane tezy. Główna część rozprawy doktorskiej składa się z następujących artykułów, z których każdy dostarcza rozwiązania dla jednego z określonych problemów badawczych:

- [1] A. Dudek and J. Baranowski. "Gaussian Processes for Signal Processing and Representation in Control Engineering". In: *Applied Sciences* 12.10 (2022).
- [2] A. Dudek and J. Baranowski. "Efficient Gaussian Process Calculations Using Chebyshev Nodes and Fast Fourier Transform". In: *Electronics* 13.11 (2024).
- [3] J. Baranowski, A. Dudek, and R. Mularczyk. "Transient Anomaly Detection Using Gaussian Process Depth Analysis". In: *2021 25th International Conference on Methods and Models in Automation and Robotics (MMAR)*. 2021, pp. 221–226.
- [4] A. Dudek, K. Jarzyna, and J. Baranowski. "Mixture Based Classifier Using Gaussian Processes for Induction Motor Diagnosis". In: *Proceedings of the IECON 2024 - 50th Annual Conference of the IEEE Industrial Electronics Society*. 3-6 November, Chicago, USA, 2024.
- [5] A. Dudek and J. Baranowski. "Modelling of Li-Ion battery state-of-health with Gaussian processes". In: *Archives of Electrical Engineering* vol. 72, No 3 (2023), pp. 643–659.
- [6] A. Dudek and J. Baranowski. "Spatial Modeling of Air Pollution Using Data Fusion". In: *Electronics* 12.15 (Aug. 2023), p. 3353.

### Problem 1. Zastosowanie Procesów Gaussowskich w naukach technicznych

PG zdobyły uznanie w naukach technicznych dzięki zdolności modelowania nieliniowych zależności i kwantyfikacji niepewności (Garay, Huaman, and Vargas-Machuca 2021b; Zeng, Ho, and Yu 2020; Gönül, Kutlar, Calik, and Orcun Parlak 2021). Znajdują szerokie zastosowanie w inżynierii sterowania, przetwarzaniu sygnałów, diagnostyce systemów oraz analizie predykcyjnej. W szczególności PG wykorzystywane są do regresji, klasyfikacji, identyfikacji systemów oraz predykcji trendów (Garay, Huaman, and Vargas-Machuca 2021b; Zeng, Ho, and Yu 2020). Ich probabilistyczna natura umożliwia dokładne przewidywanie zachowań systemów, a jednocześnie dostarcza miary niepewności, co jest kluczowe w aplikacjach wymagających wysokiej interpretowalności wyników.

Mimo licznych zalet, PG mają ograniczenia związane z kosztami obliczeniowymi i skalowalnością. Ich złożoność obliczeniowa rośnie sześciennie wraz z liczbą próbek, co sprawia, że są trudne do zastosowania w dużych zbiorach danych i aplikacjach czasu rzeczywistego (Wang and Mao 2019). Rozwiązaniem



Rysunek 1: Obszary zastosowań PG: regresja, klasyfikacja, predykcja oraz identyfikacja anomalii.

tego problemu są techniki aproksymacyjne, takie jak rzadkie PG, metody punktów indukujących oraz wnioskowanie wariacyjne, które zmniejszają obciążenie obliczeniowe przy zachowaniu wysokiej jakości predykcji (Tagade, Hariharan, Ramachandran, Khandelwal, Naha, Kolake, and Han 2020). Dodatkowo, zastosowanie obliczeń na GPU pozwala na znaczące przyspieszenie operacji (Tagade, Hariharan, Ramachandran, Khandelwal, Naha, Kolake, and Han 2020).

Kolejnym wyzwaniem jest dostosowanie PG do specyficznych dziedzin. W diagnostyce i wykrywaniu anomalii kluczowy jest odpowiedni dobór jądra, aby skutecznie odróżnić normalne fluktuacje od rzeczywistych usterek (Wang and Mao 2019). W predykcji systemów energetycznych, gdzie istotne są sezonowość i wpływ warunków zewnętrznych, PG wymagają odpowiedniej struktury jądra do uchwycenia tych zależności (Zeng, Ho, and Yu 2020). Optymalizacja hiperparametrów pozostaje istotnym zagadnieniem, wpływającym na dokładność prognoz.

PG są także integrowane z innymi metodami, co poprawia ich skuteczność i rozszerza możliwości zastosowań. Połączenie PG z sieciami neuronowymi pozwala na poprawę dokładności predykcji przy jednoczesnym zachowaniu probabilistycznej interpretowalności wyników (Gönül, Kutlar, Calik, and Orcun Parlak 2021). Wykorzystanie PG w systemach hybrydowych, np. w kontroli emisji i optymalizacji parametrów silników, potwierdza ich wartość w praktycznych aplikacjach.

Postępy w metodologii PG, takie jak próbkowanie metodą łańcucha Markowa Monte Carlo (MCMC), maksymalizacja prawdopodobieństwa (metody MLE) oraz optymalizacja jąder, zwiększyły ich skuteczność w inżynierii technicznej. Badania nad metodami aproksymacyjnymi są kluczowe dla zwiększenia ich praktycznego zastosowania w dużych i dynamicznych systemach. Rozwiązanie tych problemów pozwoli na szersze wdrożenie PG w inżynierii, automatyce i analizie predykcyjnej, torując drogę do bardziej efektywnych i interpretowalnych rozwiązań.

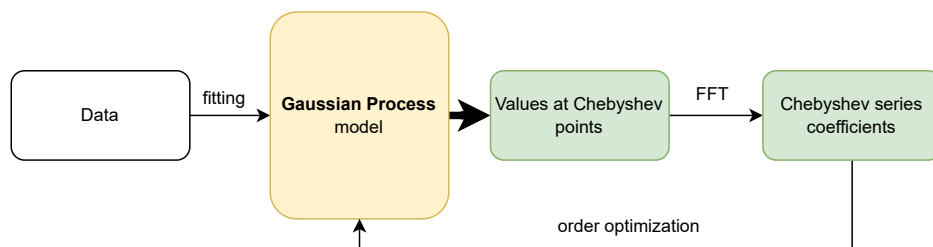
## Problem 2. Efektywność obliczeniowa generatywnych modeli procesów Gaussowskich

PG są szeroko stosowane do modelowania nieliniowych zależności i kwantyfikacji niepewności, jednak ich wysokie wymagania obliczeniowe, zwłaszcza w zastosowaniach generatywnych, stanowią istotne wyzwanie (Dudek and Baranowski 2022). Kluczowym problemem jest manipulowanie dużymi macierzami kowariancji, których wymagania pamięciowe rosną kwadratowo, a złożoność obliczeniowa sześciennie względem

liczby próbek. Jest to szczególnie dotkliwe w modelowaniu przestrzennym, prognozowaniu szeregów czasowych i aproksymacji funkcji (Garay, Huaman, and Vargas-Machuca 2021b; Dudek and Baranowski 2023).

Najbardziej wymagającą operacją w PG jest inwersja macierzy kowariancji, tradycyjnie wykonywana poprzez rozkład Choleskiego o złożoności  $O(n^3)$ , co ogranicza ich zastosowanie w dużych zbiorach danych. W modelach generatywnych dodatkowym obciążeniem jest wielokrotne próbkowanie z rozkładu posteriori, co zwiększa koszty obliczeniowe. W celu ich redukcji stosuje się techniki aproksymacyjne, takie jak rzadkie PG czy wariacyjne wnioskowanie, choć często wiąże się to z kompromisem między dokładnością a wydajnością (Trefethen 2012; Ahmed and Fisher 1968).

Jednym z obiecujących rozwiązań tego problemu jest integracja węzłów Czebyszewa i transformacji Fouriera (FFT). Węzły Czebyszewa, jako optymalne punkty interpolacyjne, pozwalają na efektywne odwzorowanie funkcji przy minimalnej utracie dokładności, natomiast FFT umożliwia szybkie przekształcenie funkcji w postać spektralną, znacząco redukując obciążenie obliczeniowe (Trefethen 2012; Ahmed and Fisher 1968).



Rysunek 2: Schemat blokowy przedstawiający metodę optymalizacji generatywnego modelowania PG za pomocą węzłów Czebyszewa i FFT.

Rys. 2, przedstawiono cały algorytm i połączenie tych metod. Pozwala on na redukcję wymiarowości oraz znaczące przyspieszenie obliczeń w generatywnych modelach. Proponowana metoda wykorzystuje strukturę wielomianów Czebyszewa do optymalizacji rozmieszczenia punktów obliczeniowych, tym samym redukując wymiarowość problemu. Przekształcając wartości funkcji w tych punktach na współczynniki szeregu Czebyszewa za pomocą FFT, uzyskujemy wysoce efektywną reprezentację funkcji, minimalizując zarówno koszt obliczeniowy, jak i wymagania pamięciowe. Technika ta nie tylko przyspiesza obliczenia PG, ale także dostarcza bogatszej reprezentacji modelowanej funkcji, uchwytyjąc bardziej subtelne szczegóły niż tradycyjne podejścia.

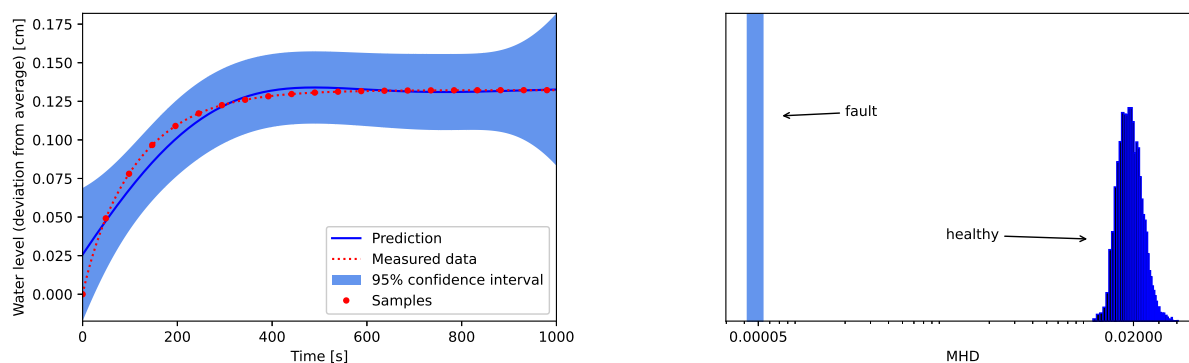
### Problem 3. Wykrywanie anomalii w sygnałach

W nowoczesnych systemach sterowania anomalie, takie jak nagłe skoki czy krótkotrwałe fluktuacje, są trudne do wykrycia, ponieważ tradycyjne metody analizy często je maskują. Standardowe podejścia oparte na uśrednianiu i transformacjach częstotliwościowych skutecznie identyfikują trwałe zmiany, ale mają trudności z wykrywaniem krótkich, nieregularnych odchyień, które mogą wskazywać na wczesne usterki systemu.

Proponujemy podejście oparte na PG, w którym cały sygnał jest modelowany jako funkcja czasu, eliminując konieczność ręcznej ekstrakcji cech. Dzięki probabilistycznemu wnioskowaniu PG pozwalają na sys-

tematyczne wykrywanie odchyleń od oczekiwanego zachowania. W połączeniu z analizą głębokości danych podejście to umożliwi rozróżnienie stanów zdrowych i uszkodzonych bez potrzeby jawnej inżynierii cech.

Walidacja została przeprowadzona na układzie hydraulicznym z trzema zbiornikami, w którym celowo wprowadzono usterki poprzez manipulację zaworami. Model PG generował probabilistyczne oszacowania oczekiwanych wzorców sygnału, a obserwacje rzeczywiste były porównywane z predykcjami w celu wykrycia anomalii. Wykorzystano funkcje głębokości Euklidesowej i Mahalanobisa do określenia prawdopodobieństwa, że dana próbka należy do normalnego reżimu operacyjnego (Tukey 1975).



Rysunek 3: Po lewej: Predykcja PG dla stanu zdrowego z rzeczywistymi pomiarami i 95% przedziałem ufności. Po prawej: Histogram głębokości Mahalanobisa pokazujący wyraźne rozróżnienie stanów zdrowych i uszkodzonych.

Jak przedstawiono na Rys. 3, model PG dokładnie odwzorowuje trajektorię sygnału zdrowego systemu, a przedziały ufności oddają naturalną zmienność. Histogram głębokości Mahalanobisa ukazuje wyraźny podział – dane zdrowe grupują się przy wyższych wartościach, podczas gdy anomalie mają znacznie niższe wartości, co pozwala na skuteczną klasyfikację usterek.

Uzyskane wyniki pokazują, że połączenie PG z analizą głębokości danych stanowi skuteczne narzędzie do wykrywania anomalii w systemach sterowania. W przeciwieństwie do metod opartych na cechach, nasze podejście nie wymaga jawnej inżynierii cech, co czyni je bardziej uniwersalnym w zastosowaniach przemysłowych.

#### Problem 4. Wyzwania w diagnostyce silników indukcyjnych

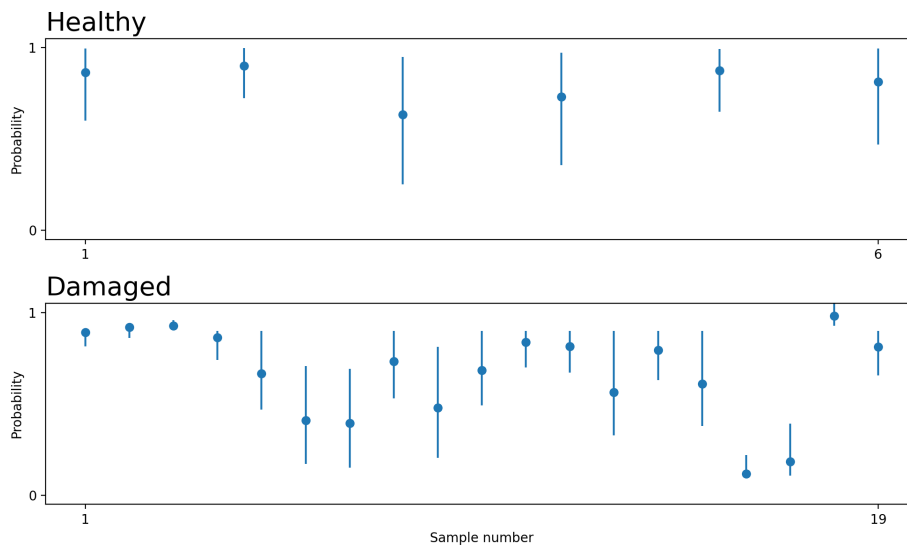
Dokładna diagnostyka silników indukcyjnych jest kluczowa dla efektywności przemysłowej i zapobiegania awariom. Tradycyjne metody mają trudności z klasyfikacją usterek w obecności szumu, zmiennych warunków pracy oraz przejściowych stanów rozruchu. Główne wyzwanie stanowi obsługa rzadkich i zaszumionych danych, ponieważ diagnostyka opiera się na profilach funkcjonalnych, a nie pojedynczych pomiarach.

Aby rozwiązać ten problem, proponujemy probabilistyczne podejście do klasyfikacji, łączące PG z modelami mieszanin Gaussowskich (GMM). PG umożliwiają solidne wnioskowanie w warunkach niepewności, a GMM poprawiają klasyfikację poprzez modelowanie struktury danych. Metoda ta skutecznie wykrywa usterki nawet przy ograniczonej ilości danych.

Badanie obejmowało cztery silniki indukcyjne o różnym stopniu uszkodzenia, testowane w różnych warunkach napięcia i obciążenia. Analiza spektrogramów ujawniła kluczowe różnice między zdrowymi i uszkodzonymi silnikami. Klasyfikacja została przeprowadzona przy użyciu GMM, gdzie każda składowa była modelowana za pomocą PG, co zwiększyło odporność na szum i brakujące dane.

Implementacja w środowisku Stan (Stan Development Team 2024) umożliwiła: - integrację wiedzy apriorycznej dla lepszej klasyfikacji, - efektywne próbkowanie Hamiltonian Monte Carlo (HMC), - elastyczne modelowanie parametrów wysokowymiarowych.

W celu poprawy uogólnienia klasyfikatora wygenerowano syntetyczne dane z modeli B-splain, co zwiększyło adaptację do rzeczywistych obserwacji.



Rysunek 4: Wyniki klasyfikacji GMM-PG: prawdopodobieństwa podziału na zdrowe i uszkodzone silniki.

Rys. 4, ukazuje klasyfikację probabilistyczną, która pozwala na ocenę stopnia pewności predykcji, co jest kluczowe w aplikacjach przemysłowych. Integracja PG z GMM znacząco poprawia diagnostykę usterek, oferując solidne i efektywne podejście do klasyfikacji. Metoda ta wspiera strategię proaktywnej konserwacji, minimalizując czas przestoju i zwiększając niezawodność systemów przemysłowych.

### Problem 5. Diagnostyka baterii litowo-jonowych

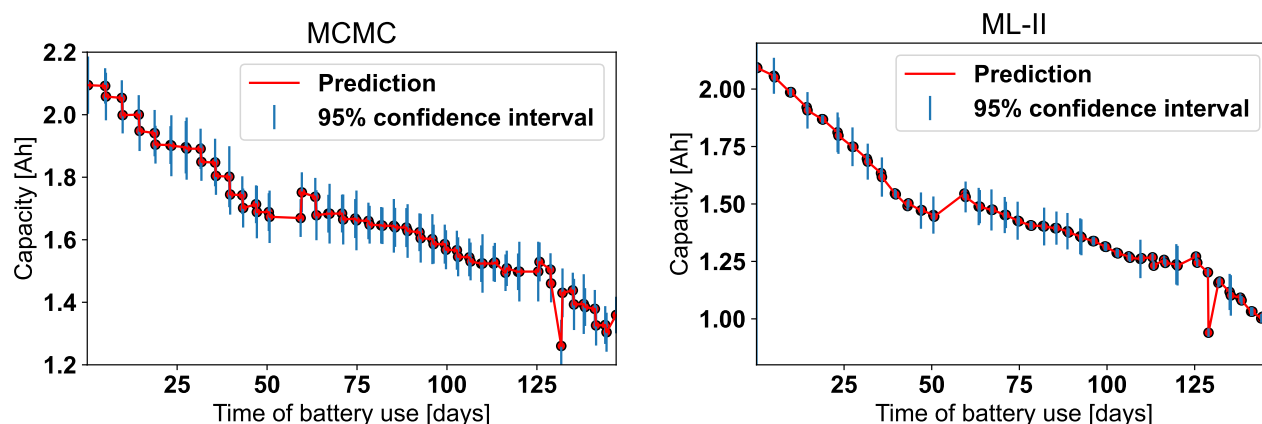
Degradacja baterii litowo-jonowych prowadzi do spadku pojemności i wzrostu oporu wewnętrznego, co wpływa na ich wydajność. Prognozowanie stanu zdrowia (SoH) (Garay, Huaman, and Vargas-Machuca 2021a) jest kluczowe, lecz trudne ze względu na złożoność procesu starzenia się ogniw.

W celu poprawy predykcji SoH zastosowaliśmy PG z metodami ML-II oraz MCMC. PG umożliwiają modelowanie nieliniowych zależności bez określonej postaci funkcji, a podejście wielojądrowe zwiększa dokładność prognoz. Analizowaliśmy dane z NASA AMES Prognostics Repository (Bole, Kulkarni, and Daigle 2014), obejmujące 28 ogniw 18650 o różnych profilach eksploatacyjnych.

Przetestowano dwa podejścia: **ML-II** – efektywne obliczeniowo, optymalizuje hiperparametry maksymalizując prawdopodobieństwo brzegowe oraz **MCMC** – zapewnia bogatszą estymację niepewności, lecz wymaga większej mocy obliczeniowej.



Metoda ML-II osiągnęła najwyższą dokładność prognoz SoH. MCMC dostarczyło szerszych przedziałów ufności, co zwiększa wiarygodność kwantyfikacji niepewności. Rysunek 5 przedstawia porównanie wyników obu metod.



Rysunek 5: Porównanie predykcji SoH przy użyciu ML-II (prawy wykres) oraz MCMC (lewy wykres). Czerwone kropki – przewidywane wartości, niebieskie linie – 95% przedział ufności.

Wyniki potwierdzają skuteczność PG w prognozowaniu SoH. ML-II zapewnia kompromis między dokładnością a efektywnością, co czyni ją przydatną w zastosowaniach czasu rzeczywistego. MCMC oferuje lepszą estymację niepewności, co może być kluczowe w systemach zarządzania bateriami. Dalsze badania skoncentrują się na doskonaleniu modeli MCMC i ich integracji z systemami zarządzania energią.

## Problem 6. Modelowanie przestrzenne zanieczyszczenia powietrza

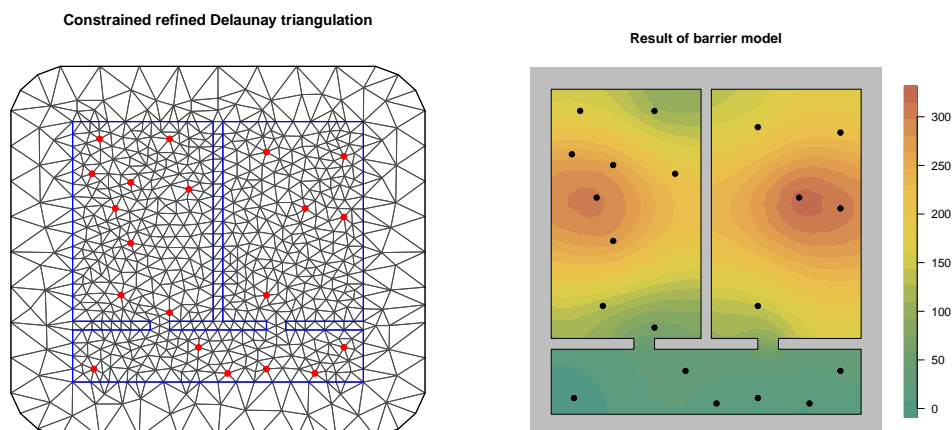
Zanieczyszczenie powietrza stanowi zagrożenie dla zdrowia publicznego i środowiska, a jego przestrzenne i czasowe zmienności utrudniają dokładne modelowanie. Tradycyjne systemy monitorowania, oparte na ograniczonej liczbie stacji pomiarowych, często dostarczają fragmentaryczny obraz rzeczywistej jakości powietrza. Prognozowanie zanieczyszczeń jest dodatkowo utrudnione przez nieliniowe zależności przestrzenne i wpływ czynników środowiskowych.

W celu poprawy dokładności modelowania zastosowaliśmy fuzję danych, integrując pomiary z niskokosztowych sensorów i profesjonalnych stacji monitorujących. Wykorzystując metodę INLA w ramach R-INLA, skutecznie odwzorowaliśmy przestrzenny rozkład zanieczyszczeń, uwzględniając zmienność w środowiskach miejskich i przemysłowych. Podejście to poprawia prognozy w obszarach o ograniczonej liczbie czujników, wspierając monitoring środowiskowy i analizy wpływu na zdrowie publiczne.

W warunkach przemysłowych rozprzestrzenianie zanieczyszczeń modelowano za pomocą dwuwymiarowych rozkładów normalnych, uwzględniając model barierowy, który odwzorowuje przeszkody fizyczne wpływające na przepływ zanieczyszczeń. W modelu miejskim analizowano rzeczywiste dane jakości powietrza z Krakowa, uwzględniając stacje rządowe oraz czujniki amatorskie. Wyniki potwierdzają skuteczność podejścia INLA w trzech scenariuszach:

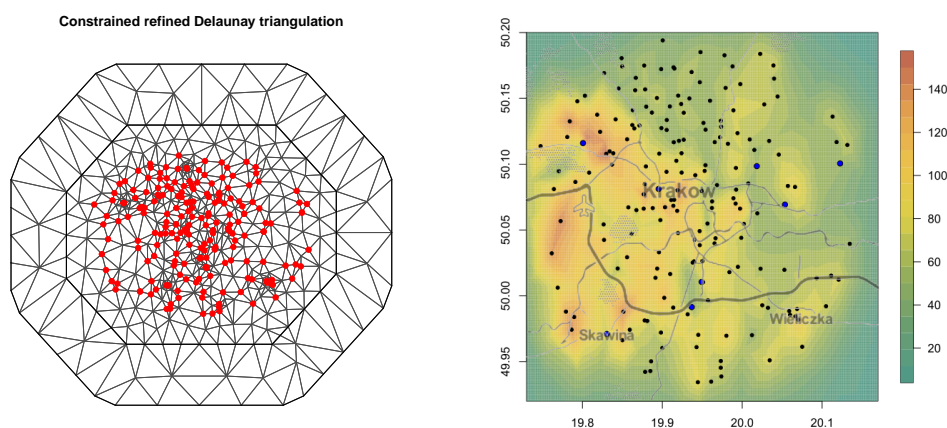
**1. Model przemysłowy:** Stężenie zanieczyszczeń osiąga najwyższe wartości w pobliżu źródeł emisji i maleje wraz z odległością. Triangulacja Delaunaya skutecznie odwzorowuje zmienność przestrzenną.

**2. Model przemysłowy z barierami:** Uwzględnienie przeszkód fizycznych poprawia dokładność predykcji. Rys. 6 przedstawia triangulację oraz wyników rozkład zanieczyszczeń.



Rysunek 6: Po lewej: Triangulacja pomieszczenia przemysłowego z barierami. Po prawej: Rozkład zanieczyszczeń, gdzie najwyższe stężenia występują w pobliżu źródeł emisji.

**3. Model zanieczyszczeń w Krakowie:** Fuzja danych ujawnia wyższe stężenia w zachodnich i centralnych obszarach miasta. Rys. 7 przedstawia przestrzenny rozkład zanieczyszczeń.



Rysunek 7: Po lewej: Triangulacja Krakowa z lokalizacjami czujników. Po prawej: Modelowany rozkład zanieczyszczeń, uwidaczniający obszary o wysokim stężeniu.

Nasze badanie potwierdza skuteczność fuzji danych w modelowaniu przestrzennym zanieczyszczeń powietrza. Integracja metod INLA zwiększa dokładność prognoz, szczególnie w obszarach o ograniczonej liczbie pomiarów. Przyszłe prace będą koncentrować się na modelowaniu trójwymiarowym oraz integracji dodatkowych czynników środowiskowych, takich jak kierunek wiatru, w celu dalszego udoskonalenia prognoz.

## Wnioski końcowe i przyszłe kierunki badań

Praca doktorska analizowała zastosowanie modeli PG w diagnostyce technicznej i systemach sterowania. Celem było rozwiązanie problemów związanych z zarządzaniem niepewnością i wykrywaniem anomalii w czasie rzeczywistym. Dzięki probabilistycznemu charakterowi PG poprawiły zarówno przewidywanie zachowań systemów, jak i kwantyfikację niepewności, co czyni je cennym narzędziem w diagnostyce usterek.

W ramach badań osiągnięto między innymi następujące rezultaty:

- Przeprowadzono przegląd zastosowań PG w inżynierii sterowania, obejmujący optymalizację, przetwarzanie obrazu i modelowanie nieliniowych systemów, co podkreśliło ich wszechstronność w analizie złożonych układów.
- Opracowano efektywne algorytmy numeryczne, w tym konwersję wartości funkcji w punktach Czebyszewa na współczynniki szeregowo przy użyciu FFT, co zoptymalizowało czas obliczeń bez utraty dokładności modelu.
- Wprowadzono nowatorskie podejście do wykrywania anomalii w stanach przejściowych poprzez połączenie PG z funkcjami głębokości danych oraz generatywne modele PG do probabilistycznej estymacji i syntezy danych w warunkach niedoboru informacji.
- Rozwinięto metody diagnostyki silników indukcyjnych poprzez integrację PG z modelem mieszanin Gaussowskich (GMM) oraz analizę częstotliwościową prądów rozruchowych, co pozwoliło na precyzyjne wykrywanie usterek i optymalizację strategii konserwacyjnych.
- Zaprojektowano modele PG do estymacji stanu zdrowia (SoH) baterii litowo-jonowych, wdrażając optymalizację ML-II oraz próbkowanie MCMC, a także zastosowano metody PG z wieloma jądrami do precyzyjnego szacowania żywotności baterii, co wspiera rozwój systemów zarządzania energią.
- Opracowano przestrzenne modele prognozowania zanieczyszczeń powietrza, łącząc metodę INLA z technikami fuzji danych, a także wdrożono zaawansowane modelowanie barierowe, które poprawiło odwzorowanie dyspersji zanieczyszczeń w środowiskach zamkniętych.

Wyniki badań potwierdziły skuteczność PG w diagnostyce technicznej, oferując precyzyjne modelowanie, wykrywanie anomalii oraz kwantyfikację niepewności w szerokim zakresie zastosowań.

Osiągnięcia tej pracy potwierdziły sformułowane tezy, skutecznie rozwiązując zidentyfikowane problemy badawcze. Badania wykazały, że PG efektywnie modelują złożone nieliniowe zależności w diagnostyce, łącząc wysoką moc predykcyjną z kwantyfikacją niepewności. Generatywne modele PG okazały się szczególnie użyteczne w warunkach ograniczonej dostępności danych, umożliwiając tworzenie syntetycznych zbiorów poprawiających dokładność diagnostyki.

Mimo tych zalet, skalowalność PG dla dużych zbiorów oraz ich wydajność w czasie rzeczywistym pozostają wyzwaniem. Dalsze badania powinny koncentrować się na metodach aproksymacyjnych, zastosowaniach online oraz rozszerzeniu generatywnych modeli na obszary takie jak opieka zdrowotna i monitorowanie środowiska. Testowanie PG w rzeczywistych warunkach, np. w smart grids i pojazdach autonomicznych, otwiera nowe perspektywy dla diagnostyki i konserwacji predykcyjnej.

Podsumowując, praca wniosła istotny wkład w rozwój diagnostyki technicznej, pokazując, że PG nie tylko są skutecznym narzędziem predykcyjnym, ale także znacząco zwiększają odporność i praktyczność algorytmów diagnostycznych poprzez obsługę niepewności i generowanie syntetycznych danych.

## Wybrana literatura

- Ahmed, N. and P. Fisher (1968). “Study of algorithmic properties of chebyshev coefficients”. In: *International Journal of Computer Mathematics* 2.1-4, pp. 307–317.
- Bikku, T. (July 2020). “Multi-layered deep learning perceptron approach for health risk prediction”. In: *Journal of Big Data* 7.1.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press.
- Bole, B., C. Kulkarni, and M. Daigle (2014). *Randomized battery usage data set*. NASA AMES Prognostics Data Repository, NASA Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA.
- Box, G. E. P. and G. M. Jenkins (1970). *Time Series Analysis: Forecasting and Control*. 1st. San Francisco: Holden-Day.
- Dudek, A. and J. Baranowski (2022). “Gaussian Processes for Signal Processing and Representation in Control Engineering”. In: *Applied Sciences* 12.10.
- (Aug. 2023). “Spatial Modeling of Air Pollution Using Data Fusion”. In: *Electronics* 12.15, p. 3353.
- Garay, F., W. Huaman, and J. Vargas-Machuca (2021a). “State of health diagnostic and remain useful life prognostic for lithium-ion battery by combining multi-kernel in Gaussian process regression”. In: *2021 IEEE XXVIII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, pp. 1–4.
- Garay, F., W. Huaman, and J. Vargas-Machuca (2021b). “State of health diagnostic and remain useful life prognostic for lithium-ion battery by combining multi-kernel in Gaussian process regression”. In: *2021 IEEE XXVIII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, pp. 1–4.
- Gönül, M., O. A. Kutlar, A. T. Calik, and F. Orcun Parlak (2021). “Prediction of oil dilution formation rate due to post injections in diesel engines by using Gaussian process”. In: *Fuel* 305, p. 121608.
- Gonzalvez, J., E. Lezmi, T. Roncalli, and J. Xu (2019). *Financial Applications of Gaussian Processes and Bayesian Optimization*. arXiv: [1903.04841 \[q-fin.PM\]](https://arxiv.org/abs/1903.04841).
- Stan Development Team (2024). *Stan User’s Guide*. <https://mc-stan.org/users/documentation/>. Accessed: 05.01.2024.
- Tagade, P., K. S. Hariharan, S. Ramachandran, A. Khandelwal, A. Naha, S. M. Kolake, and S. H. Han (2020). “Deep Gaussian process regression for lithium-ion battery health prognosis and degradation mode diagnosis”. In: *Journal of Power Sources* 445, p. 227281.
- Trefethen, L. N. (2012). *Approximation Theory and Approximation Practice*. SIAM, pp. I–VII, 1–305.
- Tukey, J. W. (1975). “Mathematics and the picturing of data”. In: *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*. Vol. 2, pp. 523–531.
- Wang, B. and Z. Mao (2019). “Outlier detection based on Gaussian process with application to industrial processes”. In: *Applied Soft Computing* 76, pp. 505–516.
- Zeng, A., H. Ho, and Y. Yu (2020). “Prediction of building electricity usage using Gaussian Process Regression”. In: *Journal of Building Engineering* 28, p. 101054.