



PODPIS ZAUFANY

GRAŻYNA
CZERNIAK
17.12.2025 21:07:17 GMT+1
Dokument podpisany elektronicznie
podpisem zaufanym

Politechnika Bydgoska
Wydział Technologii i Inżynierii Chemicznej
Ul. Al. Prof. S. Kaliskiego 7
85-796 Bydgoszcz
za pośrednictwem:
Rady Doskonałości Naukowej
pl. Defilad 1
00-901 Warszawa
(Pałac Kultury i Nauki, p. XXIV, pok. 2401)

Grażyna Czerniak

Politechnika Bydgoska im. Jana i Jędrzeja Śniadeckich

Wniosek

z dnia 17.12.2025 r.

o przeprowadzenie postępowania w sprawie nadania stopnia doktora habilitowanego
w dziedzinie **nauk ścisłych i przyrodniczych**, w dyscyplinie **nauki chemiczne**.

Określenie osiągnięcia naukowego będącego podstawą ubiegania się o nadanie stopnia
doktora habilitowanego

**Cykl powiązanych tematycznie artykułów naukowych, zgodnie z art.219 ust.1.pkt.2 Ustawy,
pt. *Metoda głównych składowych w rozwiązywaniu problemów analizy jakościowej i ilościowej
próbek o złożonej strukturze i składzie chemicznym – podejście niestandardowe***

Wniosuję – na podstawie art. 221 ust. 10 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie
wyższym i nauce – aby komisja habilitacyjna podejmowała uchwałę w sprawie nadania stopnia
doktora habilitowanego w głosowaniu ~~tajnym~~/jawnym^{*1}

Zostałem poinformowany, że:

*Administratorem w odniesieniu do danych osobowych pozyskanych w ramach postępowania w
sprawie nadania stopnia doktora habilitowanego jest Przewodniczący Rady Doskonałości Naukowej
z siedzibą w Warszawie (pl. Defilad 1, XXIV piętro, 00-901 Warszawa).*

*Kontakt za pośrednictwem e-mail: kancelaria@rdn.gov.pl, tel. 22 656 60 98 lub w siedzibie organu.
Dane osobowe będą przetwarzane w oparciu o przesłankę wskazaną w art. 6 ust. 1 lit. c)
Rozporządzenia UE 2016/679 z dnia 27 kwietnia 2016 r. w związku z art. 220 - 221 oraz art.
232 – 240 ustawy z dnia 20 lipca 2018 roku - Prawo o szkolnictwie wyższym i nauce, w celu
przeprowadzenie postępowania o nadanie stopnia doktora habilitowanego oraz realizacji praw i
obowiązków oraz środków odwoławczych przewidzianych w tym postępowaniu.*

*Szczegółowa informacja na temat przetwarzania danych osobowych w postępowaniu dostępna jest
na stronie www.rdn.gov.pl/klauzula-informacyjna-rodo.html*

.....
(podpis wnioskodawcy)

¹ * Niepotrzebne skreślić.

Załączniki:

Załącznik 1. Dane wnioskodawcy

Załącznik 2. Kopia dyplomu potwierdzającego posiadanie stopnia doktora

Załącznik 3. Autoreferat

Załącznik 4. Wykaz osiągnięć naukowych albo artystycznych, stanowiących znaczny wkład w rozwój określonej dyscypliny

Załącznik 5. Osiągnięcie naukowe będące podstawą o ubieganie się o stopień doktora habilitowanego (Załącznik 5a) wraz z oświadczeniami autorów (Załącznik 5b)

Załącznik 6. Kopia dokumentów potwierdzających odbycie staży naukowych

Autoreferat

do wniosku w sprawie nadania stopnia doktora habilitowanego

Dziedzina nauk ścisłych i przyrodniczych

Dyscyplina naukowa: nauki chemiczne

Grażyna Czerniak

Zakład Fizykochemii Powierzchni
Wydział Technologii i Inżynierii Chemicznej
Politechnika Bydgoska

Bydgoszcz 2025

Podziękowania

W badaniach przedstawionych w niniejszym autoreferacie byłam odpowiedzialna za projektowanie, opracowanie i interpretację części chemometrycznej. Wnioskowanie chemometryczne nie byłoby możliwe bez merytorycznego wsparcia moich współautorów, którym pragnę serdecznie podziękować. Szczególne podziękowania kieruję do współautorów prac A1, A3, A4 oraz A5, za ich wkład w część eksperymentalną - od zaprojektowania eksperymentu, przygotowania próbek, do kolekcji danych, bez których niniejsze prace by nie powstały. Pragnę również serdecznie podziękować eksperymentatorom, którzy zgromadzili i upublicznili dane, na podstawie których powstały prace A2, A6 i A7.

Spis treści

I.	Informacje o kandydacie.....	4
1.	Imię i nazwisko	4
2.	Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej	4
3.	Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych ..	4
II.	Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.)	5
1.	Tytuł osiągnięcia naukowego.....	5
2.	Wykaz publikacji wchodzących w skład osiągnięcia naukowego	5
3.	Omówienie celu naukowego w.w. prac wraz z wnioskami naukowymi	7
3.1	Wprowadzenie	7
3.2	Cel badań	9
3.3	Szczegółowe omówienie osiągnięcia naukowego	9
3.3.1	Metoda PCA w analizie ilościowej na podstawie danych spektroskopowych	10
3.3.2	Metoda PCA w analizie jakościowej na podstawie danych spektroskopowych	19
3.3.3	Metoda PCA w opracowaniu nowego statystycznego indeksu jakości	27
III.	Omówienie pozostałych osiągnięć naukowo-badawczych	34
1.	Działalność przed uzyskaniem stopnia naukowego doktora	34
2.	Działalność po uzyskaniu stopnia naukowego doktora	35
3.	Kierunek przyszłych badań	36
4.	Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.....	36
5.	Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę	37
6.	Oprócz kwestii wymienionych w pkt. 1-5, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej	39

I. Informacje o kandydacie

1. Imię i nazwisko

Grażyna Czerniak (nazwisko rodowe Balcerowska)

2. Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej

- **Dyplom magistra fizyki**, 1994, specjalność fizyka doświadczalna, uzyskany na Wydziale Fizyki i Astronomii Uniwersytetu Mikołaja Kopernika w Toruniu.
Tytuł pracy magisterskiej: *Badanie asymetrii zderzeniowej absorpcyjnej linii kadmu $\lambda=326.1\text{ nm}$* . Praca wykonana w Zakładzie Spektroskopii Fazy Gazowej pod kierunkiem prof. dr hab. Andrzeja Bielskiego.
- **Dyplom uzyskania kwalifikacji pedagogicznych**, 1994, po ukończeniu cztero-semestralnego Studium Pedagogicznego Uniwersytetu Mikołaja Kopernika w Toruniu.
- **Stopień doktora nauk fizycznych**, 2003, nadany uchwałą Rady Naukowej Instytutu Fizyki Politechniki Wrocławskiej na podstawie przedstawionej rozprawy pod tytułem: *Wybrane zagadnienia interpretacji widm elektronów za pomocą analizy głównych składowych i analizy czynnikowej*. Promotor: prof. dr hab. Ryszard Siuda. Rozprawa opierała się na części doświadczalnej oraz analizie wyników z użyciem w/w metod wielowymiarowej analizy statystycznej.

3. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych

- 1994÷2003- asystent w Zakładzie Fizyki Doświadczalnej Instytutu Matematyki i Fizyki Akademii Techniczno-Rolniczej w Bydgoszczy,
- od 2003 do 2021 – adiunkt w Zakładzie Fizyki Instytutu Matematyki i Fizyki Akademii Techniczno-Rolniczej w Bydgoszczy (do 2006), Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy (do 2021). W wymienionym okresie moja aktywność zawodowa była nieco zróżnicowana, m.in. z powodu urlopu wychowawczego (03.2009÷09.2010) oraz urlopu dla poratowania zdrowia (01.10.2016 ÷ 20.02.2017).
- od 01.07.2021 adiunkt w Zakładzie Fizykochemii Powierzchni Instytutu Matematyki i Fizyki wchodzącego w skład Wydziału Technologii i Inżynierii Chemicznej Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy (do 31.08.2021), Politechniki Bydgoskiej (od 01.09.2021),
- od 01.01.2025 adiunkt w Zakładzie Fizykochemii Powierzchni Wydziału Technologii i Inżynierii Chemicznej Politechniki Bydgoskiej w związku z likwidacją Instytutu Matematyki i Fizyki.

II. Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.)

1. Tytuł osiągnięcia naukowego

Metoda głównych składowych w rozwiązywaniu problemów analizy jakościowej i ilościowej próbek o złożonej strukturze i składzie chemicznym – podejście niestandardowe

Osiągnięcie jest udokumentowane cyklem siedmiu powiązanych tematycznie prac oznaczonych symbolami: **A1-A7**, opublikowanych w recenzowanych czasopismach znajdujących się w bazie *Journal Citation Reports (JCR)* w latach 2014÷2025. Sumaryczny współczynnik oddziaływania IF wynosi 33.64.

2. Wykaz publikacji wchodzących w skład osiągnięcia naukowego

- **A1. Balcerowska-Czerniak G. ✉, Kupcewicz B.,**
Score-based quantitative principal component analysis with application to the study of active pharmaceutical ingredients based on attenuated total reflection fourier-transform-infrared spectra.
J. Chemom. (WILEY) **2017**, 31, e2863. <https://doi.org/10.1002/cem.2863> IF(2017)=1.5
MNiSW:30 pkt (obecnie 70)

W ramach pracy badawczej opracowałam koncepcję nowej techniki chemometrycznej - SQPCA (Score-based Principal Component Analysis), służącej analizie ilościowej w sytuacjach, gdy skład próbki nie jest w pełni znany, a dane pochodzą z jednej techniki pomiarowej. Mój wkład w powstanie manuskryptu polegał na opracowaniu i opisaniu podstaw matematycznych metody SQPCA, napisaniu odpowiedniego programu do obliczeń, napisaniu manuskryptu, korespondencji z edytorem.
- **A2. Balcerowska-Czerniak G. ✉,**
Chemical information obtained from multicomponent spectra by means of Score-based Quantitative Principal Component Analysis,
Chemom. Intell. Lab. Syst. 219 (ELSEVIER) **2021**, 104438.
<https://doi.org/10.1016/j.chemolab.2021.104438> IF (2021)=4.11
MNiSW:100 pkt

Praca stanowi rozwinięcie wcześniej opracowanej metody SQPCA i dotyczy istotnego problemu oznaczania ilościowego pojedynczego analitu w złożonej próbce. Dane do pracy zostały zaczerpnięte z bazy danych <https://ucphchemometrics.com/datasets/>.
- **A3. Balcerowska-Czerniak G. ✉, Dittmar-Wituski A., Hiller T., Bukaluk A.,**
Multivariate analysis to evaluate indium behavior at the copper phthalocyanine thin film.
J. Appl. Phys. 124 (API Publishing) **2018**, 075302. <https://doi.org/10.1063/1.5027912> IF (2018)=2.43
MNiSW :35 pkt, (obecnie 100)

W pracy jestem pomysłodawcą nowego podejścia, opartego na analizie kształtów sum wektorów ładunków (loadings), które doprowadziło do wyodrębnienia informacji niedostępnych z poziomu klasycznych analiz kształtów widm. Mój wkład w powstanie tego manuskryptu polegał na opracowaniu koncepcji pracy, opracowaniu metody analizy danych, napisaniu odpowiedniego programu do obliczeń, napisaniu manuskryptu, korespondencji z edytorem.

- **A4. Balcerowska-Czerniak G.✉**, Wronkowski A., Antończak A.J., Skowroński Ł., Wronkowska A.A.,

The potential of multivariate analysis to phase identification based on X-ray diffraction patterns.

Chemom. Intell. Lab. Syst. 135 (ELSEVIER) **2014**, 126-132.

<http://dx.doi.org/10.1016/j.chemolab.2014.04.010>

IF (2014)=2.78

MNiSW:40 pkt (obecnie 100)

W pracy jestem pomysłodawcą nowego podejścia chemometrycznego polegającego na odwróceniu kolejności standardowego postępowania – w analizie PCA zbiór nowych danych jest potraktowany jako kalibracyjny, natomiast standardowy zbiór referencyjny jako zbiór testowy. Mój wkład polegał na opracowaniu koncepcji pracy, napisaniu odpowiedniego programu do analizy danych, napisaniu manuskryptu, korespondencji z edytorem.

- **A5. Balcerowska-Czerniak G.✉**, Trzcinski M., Naparty M., Bukaluk A.,
Study of surface oxidation of polycrystalline rhodium using multivariate analysis based on combined optical microscopy and X-ray photoelectron spectroscopy data.

App. Surf. Sci. 526 (ELSEVIER) **2020**, 146617. <https://doi.org/10.1016/j.apsusc.2020.146617>

IF (2020)=6.62 MNiSW:140 pkt

W pracy opracowałam koncepcję umożliwiającą integrację informacji uzyskanych z obrazu mikroskopowego próbki z wynikami analizy XPS (X-Ray Photoelectron Spectroscopy). Cel został osiągnięty poprzez użycie w metodzie PCA informacji o próbce, które dostarczyła analiza kłasterowa. Mój wkład polegał na opracowaniu metody analizy danych, napisaniu odpowiedniego programu do obliczeń, napisaniu manuskryptu, korespondencji z edytorem.

- **A6. Balcerowska-Czerniak G.✉**, Gorczyca B.,

Rapid assessment of surface water quality using statistical multivariate analysis approach:
Oder River system case study,

Sci. Tot. Environ. (ELSEVIER) **2024**, 912, 168754.

<https://doi.org/10.1016/j.scitotenv.2023.168754>

IF(2023)=8.2

MNiSW:200 pkt

W pracy jestem pomysłodawcą nowego indeksu statystycznego służącego kontroli jakości, w szczególności próbek wody. Mój wkład polegał na opracowaniu podstaw matematycznych tego indeksu, napisaniu odpowiedniego programu do obliczeń, napisaniu manuskryptu, korespondencji z edytorem.

- **A7. Balcerowska-Czerniak G.✉**, Westad F.,

A statistical index of surface water quality as a useful extension of monitoring systems for understanding ecosystems' response to adverse events.

J. Environ. Manag. (ELSEVIER) **2025**, 389: 12610423.

<https://doi.org/10.1016/j.jenvman.2025.126104>

IF(2025)=8.0 MNiSW:200 pkt

W pracy zaproponowałam modyfikację przedstawionego w artykule A6 indeksu statystycznego, umożliwiającą prognozowanie zakwitów złotych alg. Mój wkład polegał na opracowaniu koncepcji pracy, napisaniu odpowiedniego programu do obliczeń, napisaniu manuskryptu, korespondencji z edytorem.

3. Omówienie celu naukowego w.w. prac wraz z wnioskami naukowymi

3.1 Wprowadzenie

Sygnały instrumentalne uzyskane dla złożonych próbek posiadają duży zasób informacji. W efekcie, bardzo kłopotliwe staje się analizowanie złożonych danych eksperymentalnych bez użycia technik statystycznej analizy wielowymiarowej, znanych pod nazwą chemometria. Zakres chemometrii dobrze oddaje definicja zaproponowana przez prof. Jana Mazerskiego: *”Chemometria jest dziedziną nauki i techniki zajmującą się wydobywaniem użytecznej informacji z wielowymiarowych danych pomiarowych, wykorzystującą metody statystyki i matematyki.”*¹. Jak wynika z powyższego istota chemometrii leży we właściwym użyciu idei statystycznych i matematycznych, w zależności od sposobu formułowania zasadniczych problemów w danej dziedzinie. W Polsce, okazją do wyeksponowania interdyscyplinarnych zastosowań chemometrii stała się konferencja „Chemometria – Metody i zastosowania”, odbywająca się cyklicznie od 2001 roku, organizowana przez Instytut Ekspertyz Sądowych w Krakowie. Uniwersalny charakter chemometrii można również odnaleźć w przedstawionych poniżej badaniach naukowych. Mamy tu przykład przenikania się z chemometrią wielu różnorodnych idei z czterech dyscyplin nauk – chemicznych, fizycznych, farmaceutycznych oraz nauk o Ziemi i środowisku.

Spośród wielu opracowanych metod chemometrycznych najpopularniejsza jest metoda głównych składowych, nazywana w skrócie PCA (z ang. *Principal Component Analysis*). Po raz pierwszy została sformułowana przez statystyka, Karla Pearsona w 1901 roku². Od tego czasu użyteczność metody została potwierdzona w wielu różnych dziedzinach naukowych, co przyczyniło się również do stworzenia nadmiarowej terminologii; PCA występuje obecnie pod wieloma nazwami³, np. w analizie numerycznej stosuje się nazwę – SVD (*Singular Value Decomposition*), w elektronice i fizyce znana jest jako rozwinięcie Karhunen-Loève.

Obecnie w wielu procesach badawczych dane uzyskuje się jako rezultat analiz szeregu próbek. Wyniki te można zorganizować w macierz (standardowo nazywaną macierzą danych i oznaczaną X), gdzie m wierszy odpowiada liczbie mierzonych próbek, a n kolumn odpowiada mierzonym parametrom - *zmiennym*. Często zamiast pojęcia próbka używa się terminu *obiekt*, gdyż np. kolejne wiersze mogą zawierać pomiary tej samej próbki w różnych odstępach czasu.

¹ Mazerski J., Podstawy chemometrii, Wydawnictwo Politechniki Gdańskiej 2000.

² Pearson K., On lines and planes of closest fit to systems of points in space, *Philosophical Magazine*, (6) 2, 559-572, 1901.

³ Wold S., Esbensen K., Geladi P., Principal Component Analysis, *Chemom. Intell. Lab. Syst.*, 2, 37-52, 1987.

Analiza PCA macierzy X polega na jej rozkładzie na dwie macierze, co w formie równania zapisuje się jako

$$X = T \cdot P^T + E \quad (1)$$

gdzie $T(m \times k)$, nazywana macierzą udziałów (*scores*), opisuje zależności pomiędzy obiektami, natomiast $P(n \times k)$, nazywana macierzą ładunków (*loadings*), zawiera informacje o zmiennych eksperymentalnych. E jest macierzą reszt i jest miarą niedopasowania danych do modelu PCA z k -głównymi składowymi. Indeks górny T oznacza transpozycję macierzy.

Interpretacja geometryczna stanowi cenne narzędzie w zrozumieniu istoty metody PCA. Każdy obiekt stanowi wektor w n -wymiarowej przestrzeni zmiennych oryginalnych. Głównym zadaniem PCA jest redukcja wymiaru n do mniejszej liczby k , w taki sposób, aby maksymalizować ilość wariancji w pewnych kierunkach. Te kierunki tworzą nową przestrzeń – *głównych składowych*. Poprzez projekcję danych na najczęściej 2 lub 3-wymiarowe podprzestrzenie głównych składowych, uzyskuje się graficzny obraz próbek (*przestrzeń obiektów*). Ze względu na matematyczne podstawy metody PCA, zgodnie z równaniem (1), obrazowanie danych możliwe jest nie tylko w przestrzeni obiektów (*scores space*), gdzie dokonuje się wnioskowania na temat próbek, ale również w przestrzeni zmiennych (*loadings space*), w celu analizy korelacji pomiędzy badanymi zmiennymi/parametrami.

W standardowych zastosowaniach metoda PCA znajduje zastosowanie w kompresji danych, określaniu liczby chemicznie/fizycznie interpretowalnych składowych, modelowaniu i wizualizacji struktury danych, identyfikacji obiektów odstających³. Ponadto, z uwagi na hierarchiczny charakter głównych składowych (uporządkowanie informacji wg. malejącej wariancji), umożliwia ona redukcję poziomu zakłóceń/szumów w danych pomiarowych. Redukowanie zaburzeń poprzez tworzenie modelu danych z pominięciem pewnej liczby PC gwarantuje zachowanie struktury danych, pod warunkiem, że właściwie wybierze się liczbę k istotnych głównych składowych.

Niniejszy autoreferat stanowi podsumowanie badań skupionych wokół niestandardowego podejścia do modelowania matematycznego z użyciem metody głównych składowych, mającego na celu rozwiązanie konkretnych problemów z zakresu analizy jakościowej i ilościowej próbek o złożonej strukturze i składzie chemicznym.

Doświadczenia zdobyte podczas stażu w Zakładzie Chemii Organicznej, Wydziału Chemicznego Uniwersytetu w Umëa (Szwecja) w 1996 roku, pod kierunkiem prof. Paula Geladiego, oraz prof. Svante Wolda, stały się dla mnie inspiracją do dogłębnego poznania metody PCA. Prof. Svante Wold, znany jako twórca terminu „chemometria”, opracował współczesny algorytm PCA, będący rozwinięciem koncepcji pierwotnie zaproponowanej przez prof. Ronalda Fishera. W zrozumieniu matematycznych podstaw

metody PCA bardzo pomogło mi wykształcenie w dziedzinie fizyki; pojęcia wektorów bazowych, wektorów własnych, zagadnienie własne to fundamentalne zagadnienia w algebrze liniowej i mechanice teoretycznej. Prace **A1** i **A2**, opisane w niniejszym autoreferacie, stanowią mój oryginalny wkład w odkrywanie nieznanych dotąd w literaturze aspektów metody PCA, natomiast prace **A3÷A7** ukazują nowy potencjał metody, rozszerzając klasyczny schemat jej zastosowań.

3.2 Cel badań

Celem badań stanowiących podstawę cyklu powiązanych tematycznie artykułów naukowych jest wykazanie nowego, dotychczas nieopisanego potencjału metody głównych składowych. Osiągnięcie tego celu polega na opracowaniu nowych algorytmów oraz schematów postępowania w chemometrii, opartych na metodzie PCA, które znajdują zastosowanie w sytuacjach, gdy metody analizy typowe dla danej techniki pomiarowej lub dziedziny nauki wykazują istotne ograniczenia.

W cyklu przedstawiono przykłady niestandardowego zastosowania metody PCA, w tym wykorzystanie PCA do osiągnięcia trzech celów głównych:

- I. opracowania nowej metody chemometrycznej, SQPCA (*Score-based Principal Component Analysis*) do oznaczania ilościowego składników w złożonej mieszaninie, w szczególności pojedynczego analitu, na podstawie danych typu *first-order data* (**A1**, **A2**),
- II. opracowania schematów postępowania ze zbiorami danych spektroskopowych, które ze względu na swoją złożoność nie mogą być efektywnie przetwarzane przy użyciu klasycznego podejścia (**A3÷A5**),
- III. opracowania nowego statystycznego indeksu jakości, w szczególności do monitorowania wielu parametrów fizykochemicznych i biologicznych wód powierzchniowych jednocześnie, oraz do wczesnego wykrywania próbek atypowych na przykładzie danych rzeki Odry (**A6**, **A7**),

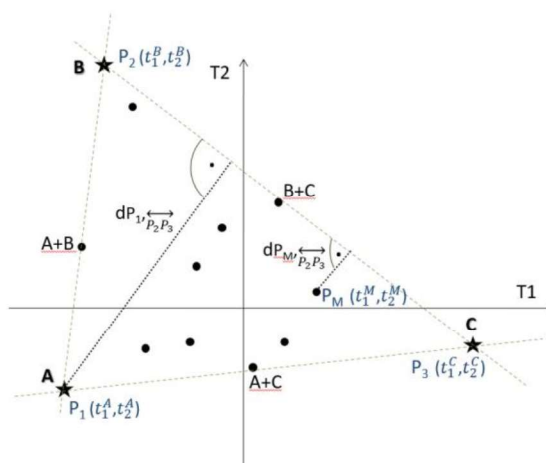
3.3 Szczegółowe omówienie osiągnięcia naukowego

Rosnąca ilość badanych materiałów o złożonej strukturze i składzie chemicznym, sprawia, że zastosowanie metod chemometrycznych nabiera szczególnego znaczenia - staje się nie tylko przydatne w zaawansowanej analizie danych, lecz często również niezbędne w rozwiązywaniu problemów badawczych. W określonych przypadkach efektywne wykorzystanie tych metod wymaga opracowania nowych, dedykowanych strategii analizy danych.

3.3.1 Metoda PCA w analizie ilościowej na podstawie danych spektroskopowych

Oznaczanie wielu składników w próbce o złożonym składzie

Analiza ilościowa z wykorzystaniem metody PCA nie była do tej pory sformułowana w sposób ogólny. W literaturze można znaleźć opracowania⁴, które pokazują, że możliwe jest wyznaczenie składu próbki metodą geometryczną poprzez obliczanie odległości euklidesowych w przestrzeni głównych składowych, na podstawie znajomości kształtów widmowych poszczególnych składników. W pewnych przypadkach można uzyskać bardzo dobre przybliżenia składu próbek, nawet gdy nie wszystkie profile spektroskopowe są znane. Gdy w tej samej przestrzeni głównych składowych, znajdzie się wiele wektorów reprezentujących widma czystych form spektralnych oraz ich mieszanin, to ich wzajemne rozmieszczenie w przestrzeni obiektów (*scores space*) przyjmuje określoną formę. Ryc.1. jest ilustracją przypadku trójskładnikowego układu po zastosowaniu odpowiedniej transformacji widm, w tym wypadku procedury centrowania danych. Jak wynika z rysunku, układ trójskładnikowy tworzy w 2-wymiarowej przestrzeni głównych składowych trójkąt, którego wierzchołki wyznaczają czyste formy – A, B oraz C. Mieszaniny dwuskładnikowe, A+B, A+C, B+C, są reprezentowane przez wektory, których końce leżą na bokach trójkąta, odpowiednio na prostych AB, AC oraz BC, natomiast końce wektorów mieszanin trójskładnikowych leżą wewnątrz trójkąta.



Ryc.1 Rozmieszczenie w przestrzeni głównych składowych (*scores space*) końców wektorów widm trzech form spektralnych A, B, C (gwiazdki) oraz ich mieszanin (koła).

⁴ Mazerski J. Chemometryczna analiza widm absorpcyjnych, II Konferencja Chemometria – Metody i Zastosowania, Wydawnictwo Instytutu Ekspertyz Sądowych, Kraków 2003.

Założmy, że przedmiotem badań jest próbka o nieznanym składzie (M), której widmo reprezentowane jest przez punkt P_M o współrzędnych (t_1^M, t_2^M) . Ilość składnika np. A można wyznaczyć dzieląc odległość punktu P_M od prostej przechodzącej przez punkty P_2 i P_3 przez odległość punktu P_1 od tej samej prostej

$$C_A = \frac{d(P_M, P_2, P_3)}{d(P_1, P_2, P_3)}$$

W pracy **A1** zaproponowałam bardziej ogólną metodę analizy ilościowej z wykorzystaniem metody PCA, którą nazwałam SQPCA (*Score-based Quantitative Principal Component Analysis*). Punktem wyjściowym metody SQPCA jest założenie, że widmo mieszaniny (M) trzech nieoddziałujących składników A, B i C można zapisać jako kombinację liniową

$$\mathbf{x} = \alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C} \quad (2)$$

gdzie α , β i γ , oznaczają ilości odpowiednio składnika A, B i C w próbce mieszaniny.

Kluczowa idea metody SQPCA, która stanowi o jej nowatorskim w chemometrii charakterze, opiera się na szczególnej właściwości modelu PCA – główne składowe zachowują proporcje (koncentracje) poszczególnych składników w mieszaninie. W literaturze znany był fakt równości skumulowanej wariancji zawartej w macierzy danych X oraz zawartej w macierzy T, natomiast nie było odniesienia do mieszanin składników. Zachowawczy charakter głównych składowych oznacza, że dla mieszaniny trójskładnikowej, reprezentowanej w modelu PCA przez wektor $\mathbf{t}_M$ spełniona jest zależność

$$\begin{cases} t_1^M = \alpha t_1^A + \beta t_1^B + \gamma t_1^C \\ t_2^M = \alpha t_2^A + \beta t_2^B + \gamma t_2^C \\ t_3^M = \alpha t_3^A + \beta t_3^B + \gamma t_3^C \end{cases} \quad (3)$$

co oznacza, że odpowiednią współrzędną mieszaniny w układzie głównych składowych wyznacza kombinacja liniowa współrzędnych czystych składników. Warto zauważyć, że wymiar przestrzeni głównych składowych w metodzie SQPCA odpowiada liczbie czystych form spektralnych. Stąd układ równań (3) rozpisany jest na trzy składowe wektora mieszaniny M - $\mathbf{t}_M(t_1^M, t_2^M, t_3^M)$. Dla porównania w omówionej wcześniej metodzie geometrycznej z wyliczaniem odległości euklidesowych mamy dwie składowe $\mathbf{P}_M(t_1^M, t_2^M)$ dla układów trójskładnikowych.

Zależność (3) zaobserwowałam w toku analizy chemometrycznej widm UV-VIS-NIR preparatów farmaceutycznych trójskładnikowych leków przeciwbólowych. Celem badań

realizowanych pod kierunkiem prof. Bogumiły Kupcewicz (Collegium Medicum UMK) było porównanie zdolności predykcyjnych trzech metod spektrofotometrycznych do jednoczesnego oznaczania paracetamolu (PAR), kwasu acetylosalicylowego (ASA) i kofeiny (CAF). Na potrzeby testowania przygotowano preparaty będące mieszaniną trzech substancji czynnych o z góry zadanych proporcjach. Technika chemometryczną, stanowiącą główne narzędzie badań była metoda kalibracji wielowymiarowej PLS (Partial Least Squares). Jednakże, praktyką jest poddawanie danych wstępnej analizie metodą PCA, którą to poprzedza standardowo transformacja danych. Zauważyłam, że gdy pominiemy etap centrowania danych, współrzędne mieszanin w modelu PCA spełniają relację (3).

Analiza ilościowa z użyciem metody SQPCA przebiega dwuetapowo. W pierwszym etapie (i) należy zbudować model PCA, wykorzystując widma mieszanin w skład których wchodzi tylko znane składniki np. A, B i C, oraz widma bazowe (widma czystych form składników A, B i C). Zaleca się aby do budowania modelu trójskładnikowego użyć widm mieszanin wygenerowanych numerycznie na podstawie widm bazowych, co znacznie ogranicza wpływ błędów pomiarowych oraz czas wymagany na przygotowanie mieszanin. Drugi etap polega na (ii) rozwiązaniu układu równań (3) dla widma mieszaniny x_M o nieznanym składzie. W tym celu można zastosować metodę Gaussa rozwiązywania układów równań liniowych. Jednoczesne ilościowe oznaczanie składników A, B i C w mieszaninie M przeprowadza się na podstawie wyrażenia

$$c = (T_{ABC}^T \cdot T_{ABC})^{-1} \cdot T_{ABC}^T \cdot t_M \quad (4)$$

gdzie wektor c zawiera szukane koncentracje składników α, β, γ , a T_{ABC} oznacza macierz współrzędnych (*scores*) widm bazowych w modelu SQPCA.

Należy podkreślić podobieństwo w sposobie wyznaczania koncentracji na podstawie wyrażenia (4) do klasycznej metody najmniejszych kwadratów (*CLS-Classical Least Squares*), w której koncentracje składników wyznacza się jako

$$c = x \cdot F \cdot (F^T \cdot F)^{-1} \quad (5)$$

gdzie macierz F jest utworzona z widm składników A, B i C. To podobieństwo jest jeszcze bardziej widoczne, gdy w wyrażeniu (4) zamiast wektorów kolumnowych użyjemy wektorów wierszowych, wówczas

$$c = t_M \cdot T_{ABC}^T \cdot (T_{ABC} \cdot T_{ABC}^T)^{-1} \quad (6)$$

Jak wynika z powyższego, różnica pomiędzy metodą SQPCA a klasycznym podejściem, wykorzystującym dekompozycję widm z użyciem metody CLS, polega na zastąpieniu przestrzeni zmiennych oryginalnych przestrzenią głównych składowych.

Otrzymane na podstawie wyrażeń (4) lub (6) wartości α, β, γ , są wiarygodne przy założeniu, że badana mieszanina nie zawiera nieznanymi składników. W takim przypadku klasyczne podejście, polegające na dekompozycji widma na znane składowe np. przy użyciu metody OLS (*Ordinary Least Squares*), daje również pożądany efekt. Metoda SQPCA została zaprojektowana głównie w celu rozwiązania innego problemu - jednoczesnych oznaczeń wybranych składników w mieszaninach wieloskładnikowych o nieznanym w pełni składzie, gdzie dodatkowo spodziewamy się nakładania form spektralnych. Klasyczna dekompozycja widm w takiej sytuacji ma ograniczenia; może być stosowana do oznaczeń ilościowych tylko tego składnika, którego forma spektralna jest ortogonalna (rozdzielona) z pozostałymi, w przeciwnym wypadku należy spodziewać się niedokładnych wyników.

Założmy, że celem jest oznaczenie ilościowe trzech znanych składników A, B i C na podstawie widma mieszaniny tych składników z nieznanym składnikiem D, co w przestrzeni zmiennych oryginalnych można zapisać jako

$$\mathbf{x}_{new} = \alpha_{new}\mathbf{A} + \beta_{new}\mathbf{B} + \gamma_{new}\mathbf{C} + \delta_{new}\mathbf{D} \quad (7)$$

natomiast w modelu SQPCA 4-wymiarowym widmo mieszaniny można przedstawić jako

$$\begin{cases} t_1^{new} = \alpha_{new}t_1^A + \beta_{new}t_1^B + \gamma_{new}t_1^C + \delta_{new}t_1^D \\ t_2^{new} = \alpha_{new}t_2^A + \beta_{new}t_2^B + \gamma_{new}t_2^C + \delta_{new}t_2^D \\ t_3^{new} = \alpha_{new}t_3^A + \beta_{new}t_3^B + \gamma_{new}t_3^C + \delta_{new}t_3^D \\ t_4^{new} = \alpha_{new}t_4^A + \beta_{new}t_4^B + \gamma_{new}t_4^C + \delta_{new}t_4^D \end{cases} \quad (8)$$

Szczegóły dalszego postępowania opisuje rozdział 2.3 pracy **A1**. Ogólnie idea polega na tym, aby przeprowadzić analizę dwuetapowo – najpierw wyliczyć przybliżone wartości $\alpha_{new}, \beta_{new}$ i γ_{new} w modelu PCA 3-wymiarowym, zbudowanym w oparciu o mieszaniny o znanym składzie, a następnie użyć tych wartości jako początkowych w modelu PCA 4-wymiarowym. W ten sposób można wyznaczyć wektor $\mathbf{t}_D$, co z kolei pozwala znaleźć widmo nieznanego składnika D. Dalej postępujemy analogicznie do sytuacji, gdy wszystkie formy spektralne są znane. Warto zauważyć, że wyznaczone widmo nieznanego składnika D nie musi oznaczać wyłącznie jednej formy spektralnej, w ogólności jest wypadkową sumą form spektralnych pozostałych nieznanymi składników próbki. Zaproponowany schemat postępowania może być uogólniony dla dowolnej liczby składników będących przedmiotem analiz ilościowych.

Efektywność metody SQPCA oceniono w pracy **A1** na podstawie widm FTIR-ATR (*Attenuated Total Reflection Fourier-Transform-Infrared*) wybranych preparatów farmaceutycznych, zawierających mieszaniny substancji czynnych takich jak: kwas

acetylosalicylowy (ASA), paracetamol (PAR), kofeina (CAF), kodeina (COD), propyfenazon (PROP). Analizę przeprowadzono na dwa sposoby. W pierwszym dokonano porównania zdolności predykcyjnych metody z metodami chemometrycznymi analizy ilościowej takimi jak OLS, PLS, MCR (*Multivariate Curve Resolution*) (tabela nr 3 w pracy A1). Do porównania użyto wskaźnika błędów predykcji (RSEP⁵ - *Relative Standard Error of Prediction*), który wyliczono niezależnie dla każdej z badanych substancji czynnych. Wyniki analizy ilościowej dla trzech substancji czynnych – ASA, PAR oraz PROP, potwierdziły zdecydowanie większą efektywność metody SQPCA w stosunku do pozostałych. Otrzymano następujące wskaźniki błędów – odpowiednio dla ASA -7,4%, PAR – 8,8%, PROP-16,6%. Dla porównania metoda PLS wskazywała na wartości dla ASA-18,4%, PAR-17,2% , natomiast PROP-309%. Tylko w przypadku oznaczeń ilościowych kofeiny metoda PLS okazała się dokładniejsza od zaproponowanej SQPCA. Natomiast żadna z zastosowanych metod nie była efektywna w ilościowym oznaczaniu kodeiny.

Drugi sposób weryfikacji metody SQPCA był ściśle ilościowy. Dla badanych dziesięciu leków (Antidol 15, Cefalgin, Dafalgan Codeine, Excedrin Migra, Nurofen Plus, Aspirin Extra, Etopiryna, Kopiryna, Modafen, Panadol Extra) otrzymane wyniki przeliczono na jednostki mg/tabletkę, a następnie porównano z informacjami o składzie podanymi przez producenta danego leku (tabela nr 4 w pracy A1). Uzyskane rezultaty wskazały na duży potencjał metody SQPCA w analizie ilościowej, przy czym zdecydowanie lepszą zgodność otrzymano dla tabletek, w których badany składnik występował w znacznych ilościach bądź był nieobecny w mieszaninie (nominalna wartość 0).

Główną zaletą zaproponowanej metody SQPCA jest jej niska kosztowość i czasochłonność w porównaniu z innymi metodami chemometrycznymi stosowanymi do oznaczania substancji w próbkach wieloskładnikowych o nieznanym składzie na podstawie danych spektralnych. Metoda ta wykorzystuje jedynie niezbędne informacje - widma referencyjne interesujących składników – dzięki czemu nie wprowadza nadmiarowości problemu. Dla porównania, metoda kalibracji wielowymiarowej PLS, chociaż nie wymaga widm referencyjnych, narzuca konieczność przygotowania wielu próbek mieszanin na potrzeby zbudowania modelu regresyjnego (model PLS do oceny jakości predykcyjnej wymaga próbek kalibracyjnych, walidacyjnych oraz testowych).

$$^5 \text{RSEP}_j (\%) = 100 \sqrt{\frac{\sum_{i=1}^m (c_{ij} - c'_{ij})^2}{\sum_{i=1}^m c_{ij}^2}}$$

Oznaczanie pojedynczego składnika w próbce o złożonym składzie

Oznaczanie ilościowe składnika w próbce wieloskładnikowej o nieznanym w pełni składzie na podstawie jej formy spektralnej jest zadaniem niełatwym, w szczególności gdy celem jest oznaczenie ilościowe pojedynczego analitu. Wraz z rozwojem technik chemometrycznych w literaturze pojawiły się algorytmy, które ten problem rozwiązują. Jeden typ algorytmów polega na budowaniu modeli kalibracyjnych na podstawie dwóch typów danych - macierzy widm X oraz wektora y , zawierającego informację o koncentracji składnika w mieszaninach użytych do zbudowania modelu np. za pomocą wspomnianej techniki PLS^{6,7}. Drugi to algorytm znany pod nazwą GRAM (*Generalized Rank Annihilation Factor Analysis*), który opiera się na badaniu rzędu macierzy danych^{8,9} po odjęciu skalowanego widma wzorcowego od widma mieszaniny. Jeśli składnik jest obecny w mieszaninie to w metodzie GRAM rząd macierzy powinien zmniejszyć się o 1. Jednak ten schemat postępowania w przypadku danych eksperymentalnych obarczonych szumami daje niejednoznaczne odpowiedzi; dodatkowe problemy stwarza też poszukiwanie właściwego sposobu skalowania widm. Trzeci typ algorytmów, wśród których metoda PARAFAC (*Parallel Factor Analysis*)^{10,11} zasługuje na wyróżnienie, wymaga użycia dwóch technik pomiarowych. W chemometrii, dane przygotowane na podstawie dwóch technik pomiarowych przyjęło się nazywać z ang. *second-order data*.

Metoda SQPCA jest jednowymiarowa; informację ilościową o próbce otrzymuje się na podstawie wektora danych (*first-order data*). Takie podejście znacznie ogranicza konieczność wykonywania dużej liczby pomiarów i stwarza możliwość szybkiej i taniej analizy. Wyniki otrzymane z wykorzystaniem metody SQPCA, w badaniach wybranych substancji czynnych preparatów farmaceutycznych przedstawione w pracy **A1** wskazywały na potencjał metody w analizie ilościowej. Stąd podjęta została próba takiej modyfikacji, aby metoda SQPCA służyła

⁶ Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* 1986; 186: 1-17.

⁷ Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* 2001; 58: 109-130.

⁸ Sanchez E, Kowalski BR. Generalized rank annihilation factor analysis. *Anal. Chem.* 1986; 58: 496-499.

⁹ Faber NM. Towards a rehabilitation of the generalized rank annihilation method (GRAM). *Anal. Bioanal. Chem.* 2002; 372: 683-687.

¹⁰ Bro R., PARAFAC. Tutorial and applications. *Chemom. Intell. Lab. Syst.* 1997, 28:2, 149-171.

¹¹ Murphy K.R., Stedmon C.A., Graeber D., and Bro R., Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Anal. Methods* 2013, 5: 6557-6566.

do oznaczeń ilościowych również pojedynczego składnika w złożonej mieszaninie na podstawie znajomości jej formy spektralnej. Badania doprowadziły do opracowania szybkiej i prostej procedury, posiadającej szeroki potencjał w analizie ilościowej, co opisano w pracy **A2**.

Gdy przedmiotem badań jest oznaczenie ilościowe tylko jednego analitu, ogólny schemat postępowania w metodzie SQPCA jest analogiczny jak w przypadku oznaczania wielu składników jednocześnie. Problemатyczny w przypadku pojedynczego składnika, może być sposób generowania widm mieszanin. Dla dwóch i więcej składników np. A, B i C generowane mieszaniny stanowiły kompozycję ich widm (*compositional data*) w taki sposób, aby suma ich koncentracji stanowiła 100%, tzn. każde widmo mieszaniny było normowane do sumy

$$c_{ij} = c_{ij} / \sum_{j=1}^k c_{ij}$$

Np. gdy $c = [56,91 \ 70,57 \ 31,92]$, po transformacji $c = [56,91 \ 70,57 \ 31,92] / 159,40 = [35,7\% \ 44,3\% \ 20\%]$ ¹²

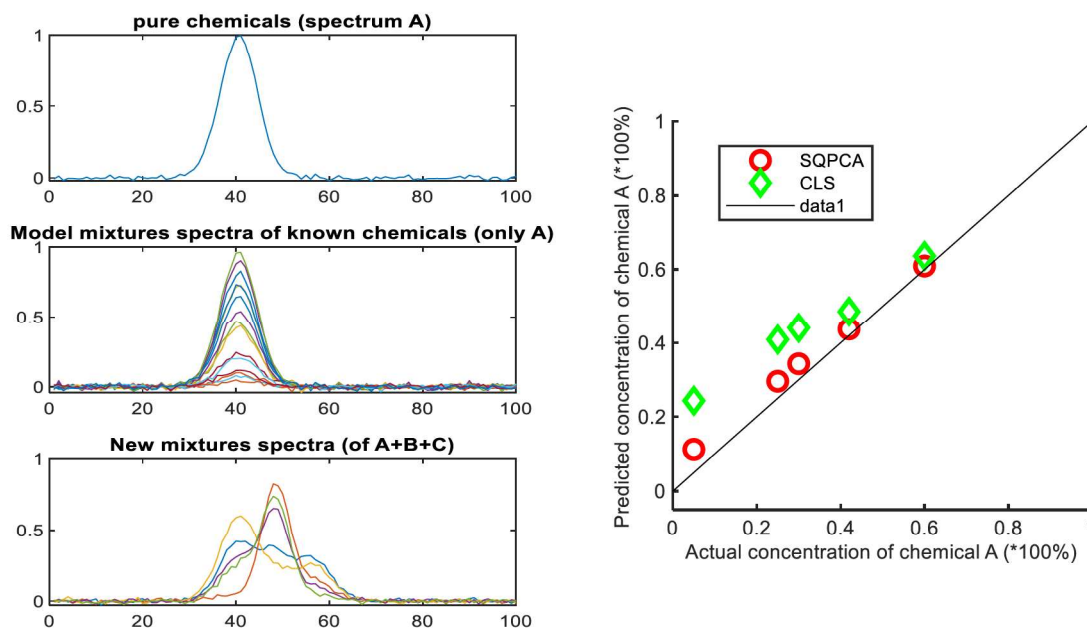
To założenie jest w ogólności poprawne; często bowiem interesują nas nie bezwzględne koncentracje, lecz jedynie proporcje pomiędzy składnikami. Jak się okazało, metoda SQPCA nie wymaga tego założenia, jednak istotne jest, aby zbiór treningowy, składający się z mieszanin znanych składników, na którym budowany jest model SQPCA, odzwierciedlał różne proporcje składnika w zakresie od 1% do 100%. To o czym należy bezwzględnie pamiętać to wybór odpowiedniego rozmiaru przestrzeni głównych składowych ($z=1$ w algorytmie). Uzyskanie poprawnych wyników wymaga przeprowadzenia obliczeń w pętli, zgodnie z diagramem zamieszczonym w pracy **A2**, który ilustruje działanie metody SQPCA i ułatwia zrozumienie poszczególnych etapów algorytmu; dla oznaczania pojedynczego składnika zaleca się trzy powtórzenia, natomiast dla większej liczby składników - dziesięć.

Przykładowe działanie metody SQPCA dla pojedynczego składnika ilustruje Ryc. 2.

Kod programu, napisany przy użyciu środowiska Matlab® (MathWorks Inc., Natick, MA, USA) jest publicznie dostępny na moim profilu na platformie Research Gate. Działanie programu jest testowane dla 1 formy spektralnej (A) w postaci krzywej Gaussa.

Zastosowanie metody SQPCA do jednoczesnego wyznaczania koncentracji wielu składników lub pojedynczego składnika może być przydatne w przypadku obecności w sygnałach spektralnych szumów na niepożądanym poziomie, lub innych zaburzeń mogących mieć wpływ na strukturę widm, wówczas chroni ono przed nadmiernym dopasowaniem modelu do danych. Co jest warte podkreślenia, z Ryc. 2 wynika, że przewaga metody SQPCA nad CLS ujawnia się szczególnie dla małych wartości koncentracji badanego składnika.

¹² Przykład zaczerpnięty z The Chemometrics Column, Brereton R.G. Principal Components Analysis: Row Scaling and Compositional Data, J. Chemom. 39:e3606, 2025, <https://doi.org/10.1002/cem.3606>



Ryc.2 Ilustracja działania metody SQPCA, gdy oznaczany jest pojedynczy składnik A w mieszaninie trójskładnikowej. Do oceny jakości otrzymanych wyników dla mieszanin (dolny rysunek) użyto typowego wykresu przedstawiającego wartości uzyskane na podstawie modelu w funkcji wartości rzeczywistych (wykres z prawej). Dla porównania przedstawiono wyniki, które otrzymuje się przy użyciu klasycznego podejścia na podstawie metody CLS.

W pracy A2 skuteczność metody SQPCA testowano dla danych pochodzących z dwóch technik analitycznych – EEM (*Fluorescence Excitation-Emission-Matrices*) oraz NMR (*Nuclear Magnetic Resonance*), udostępnionych przez grupę chemometrystów z Uniwersytetu Kopenhaskiego (<http://www.models.life.ku.dk/datasets>). Przedmiotem badań fluorescencyjnych były dwa związki organiczne z grupy fenoli – hydrochinon oraz rezorcyna, związek organiczny z grupy polifenoli – pirokatechina (katechol), dwa aminokwasy – tryptofan i tyrozyna oraz heterocykliczny związek chemiczny indol, z których przygotowano dwa zestawy próbek składające się z 16 (*dataset 1*) i 15 (*dataset 2*) mieszanin. Natomiast techniką NMR analizowano 28 próbek mieszanin przygotowanych na bazie dwóch związków peptydowych, jednego aminokwasu, oligosacharydu i alkoholu. Były to odpowiednio: walina-tyrozyna-walina (Val-Tyr-Val), tryptofan-glicyna (Trp-Gly), fenyloalanina (Phe), maltoheptaaza (Malto) i propanol (Prop).

Dla zachowania schematu 1-wymiarowego metody SQPCA ($z=1$), obliczenia dla każdego ze składników wykonywano oddzielnie. Porównanie wartości błędów predykcji otrzymanych trzema metodami analizy ilościowej – SQPCA, CLS oraz PARAFAC dla danych

fluorescencyjnych (EEM) prezentuje w pracy A2 tabela nr 1. Szczególnie interesujące jest porównanie wyników otrzymanych metodą SQPCA i metodą PARAFAC¹³, ponieważ PARAFAC należy do grupy metod analizy wykorzystujących dane zebrane w postaci macierzy trój-wymiarowej (*second-order-data*) i jest uogólnieniem metody PCA na więcej wymiarów. Dla ilustracji w Tabeli 1 przedstawiono porównanie błędów predykcji (RSEP) dla trzech fluoroforów. W przypadku tryptofanu metoda SQPCA dostarczała nieco dokładniejszych wartości stężeń w porównaniu z modelem PARAFAC, co dodatkowo potwierdza jej potencjał jako metody analizy ilościowej.

Tabela 1. Porównanie błędów predykcji (RSEP) otrzymanych dla analiz ilościowych danych EEM.

Fluorophores	SQPCA RSEP (%)	PARAFAC RSEP (%)
Resorcinol	21.94	13.84
Tryptophane	13.29	15.86
Tyrosine	31.24	17.63

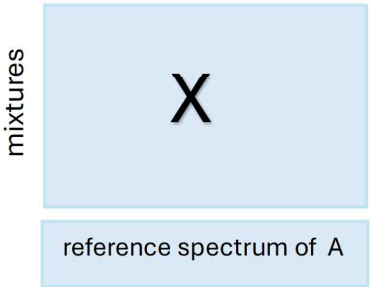
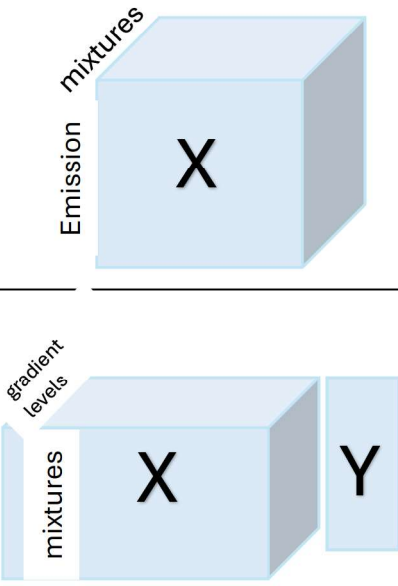
W przypadku danych NMR porównanie wyników analizy ilościowej otrzymanych metodą SQPCA z wartościami nominalnymi, użytymi podczas przygotowywania mieszanin, wypadło jeszcze bardziej korzystnie, co ilustruje rysunek nr 5 w pracy A2. Wyniki te były porównywalne z rezultatami uzyskanymi metodą PARAFAC¹⁴. W cytowanej pracy autorzy przedstawili analogiczny rysunek, stąd forma prezentacji wyników w postaci wykresów zamiast wartości wskaźników statystycznych.

Podsumowując, w obu analizowanych zestawach widm próbek wieloskładnikowych metoda SQPCA okazała się niewiele ustępować technice PARAFAC, co ze względu na fakt wymagań znacznie mniejszej ilości danych (porównanie ilustruje poniższy diagram) czyni ją alternatywnym narzędziem szybkich analiz ilościowych w celach np. przesiewowych.

W oparciu o wyniki przedstawione w pracach A1 i A2 można wnioskować, że nowa metoda chemometryczna SQPCA wypełnia istniejącą lukę badawczą, związaną z brakiem narzędzia umożliwiającego szybką i niedrogą analizę ilościową na podstawie danych spektroskopowych typu *first-order data*. W ten sposób cel badawczy nr I został osiągnięty.

¹³ Elcoroaristizabal S., Bro R., Garcia J.A., Alonso L., PARAFAC models of fluorescence data with scattering: A comparative study. Chemom. Intell. Lab. Sys. 2015, 142: 124-130.

¹⁴ Acar E., Papalexakis E.E., Gurdeniz G., Rasmussen M.A., Lawaetz A.J., Nilsson M., Bro R., Structure-Revealing Data Fusion, BMC Bioinformatics 2014, 15: 239-256.

SQPCA	PARAFAC
1 technika pomiarowa	2 techniki pomiarowe
	

3.3.2 Metoda PCA w analizie jakościowej na podstawie danych spektroskopowych

W przypadku, gdy zbiór danych stanowią dane spektroskopowe, widma, będące kombinacją liniową stałych chemicznie/fizycznie interpretowalnych czynników, metoda głównych składowych (PCA) stanowi wydajne narzędzie ich analizy, dostarcza bowiem informacji o strukturze danych. W zastosowaniach mniej standardowych PCA można stosować do wyznaczania położenia energetycznych pików, wnioskowania o korelacjach chemicznych/fizycznych składników odpowiedzialnych za obecność tych pików. Bardzo dobrze może się również sprawdzić użycie modelu głównych składowych w celu wyeliminowania niepożądanych informacji takich jak szumy, co oznacza wyodrębnianie macierzy E w równaniu (1) i użycie w dalszej analizie macierzy modelu X_{PCA} zamiast macierzy X danych oryginalnych. Te zastosowania metody PCA są bardzo dobrze udokumentowane w literaturze.

Poniżej przedstawione badania pokazują, że w sytuacjach, w których klasyczne zastosowanie metody PCA napotyka ograniczenia, zaproponowane modyfikacje dają dobre podstawy do dalszego wnioskowania. Realizacja II celu głównego, związanego z niestandardową analizą danych spektroskopowych z wykorzystaniem metody PCA, obejmowała szereg szczegółowych zadań:

- i. opracowanie nowego podejścia analizy 1-wymiarowych wektorów ładunków dla zbioru danych spektroskopowych w celu wyodrębnienia pasm o niskiej intensywności,
- ii. opracowanie schematu chemometrycznego dla dużych, zróżnicowanych zbiorów danych referencyjnych w celu ułatwienia identyfikacji związków chemicznych w próbkach o nieznanym składzie,
- iii. opracowanie schematu postępowania chemometrycznego umożliwiającego integrację informacji uzyskanych z obrazu mikroskopowego próbki z wynikami analizy spektroskopowej, w celu wyznaczenia widm referencyjnych poszczególnych obszarów/ domen próbek.

PCA jako narzędzie identyfikacji pasm o niskiej intensywności

W układach z metalami materiał organiczny taki jak ftalocyjanina miedzi (CuPc), wykazuje właściwości półprzewodnikowe, dzięki czemu stała się ona materiałem szeroko wykorzystywanym w technologii przyrządów elektronicznych oraz optoelektronicznych. Układy typu metal/CuPc na różnych podłożach są opisywane w literaturze, ale jak się okazuje wyniki nie są spójne i niejednoznaczne. Na łamach czasopisma *Physical Review B* ukazały się w odstępie pięciu lat dwie prace, które prezentowały sprzeczne ustalenia dotyczące reakcji chemicznej pomiędzy indem a warstwą ftalocyjaniny podczas osadzania się indu na warstwie organicznej. Na podstawie analizy widm UPS (*Ultraviolet Photoelectron Spectroscopy*) autorzy pierwszej pracy¹⁵ przyjęli hipotezę, że widoczny dublet z pikami przy 18,5eV i 19,5eV świadczy o obecności aktywnej chemicznie fazy, którą przypisano związkowi In₂CuPc. Natomiast autorzy drugiej pracy¹⁶ już w tytule podkreślili, że atomy indu nie reagują z ftalocyjaniną, a jako przyczynę wyżej wymienionego dubletu wskazano formowanie się klastrów, złożonych z atomów indu. Na potwierdzenie hipotezy załączono obraz próbki uzyskany metodą skaningowej mikroskopii elektronowej SEM (*SEM -Scanning Electron Microscope*).

¹⁵ Aristov V.Y., Molodtsova O.V., Zhilin V.M., Vyalikh D.V. and Knapfer M., Chemistry and electronic properties of a metal-organic semiconductor interface: In on CuPc, *Phys. Rev. B* 2005, 72, 165318-1-165318-7.

¹⁶ Ivanco J., Toader T., Frisov A., Brzhezinskaya M., Sperling M., Braun W. and Zahn D.R.T., Indium on a copper phthalocyanine thin film: Not a reactive system, *Phys. Rev B* 2010, 81, 115325-1÷115325-8.

Celem badań opisanych w pracy **A3** było powtórzenie i zweryfikowanie przyjętej metodyki analizy układu In/CuPc, a następnie - poprzez zastosowanie metody PCA - jednoznaczne wskazanie przyczyny powstania spornego dubletu w widmie UPS. Materiał doświadczalny stanowiły dwie próbki otrzymane w wyniku dwuetapowego osadzania na dwóch różnych podłożach: grafitu pirolitycznego o wysokiej orientacji (HOPG) oraz monokryształu antymonku indu o orientacji InSb (111). W pierwszym etapie osadzono warstwę ftalocyjaniny, a następnie na warstwę CuPc nanoszono warstwę indu. Obserwacje wstępne widm UPS, przeprowadzone bez użycia metod chemometrycznych, były nieco zaskakujące z dwóch powodów. Po pierwsze sporny dublet nie był widoczny. Po drugie, podczas formowania warstwy indu zaobserwowano wyraźny trend przesuwania się linii dubletu In 4d - od pozycji 17eV i 17,9eV do pokrycia się z fazą metaliczną indu (16,8eV i 17,7eV). Podobny, choć subtelny trend dał się zauważyć w wynikach prezentowanych w obu wymienionych pracach, jednakże autorzy w żadnym miejscu nie odnieśli się do tego faktu.

W analizie danych spektroskopowych z użyciem metody PCA szczególne znaczenie mają 1-wymiarowe wektory ładunków (*loadings*), odpowiadające kolejnym głównym składowym, ponieważ często na ich podstawie wnioskuje się o liczbie form spektralnych. W literaturze zwyczajowo nazywa się je widmami abstrakcyjnymi (*abstract spectra*), ze względu na brak jednoznacznej interpretacji fizycznej/chemicznej. Pierwsze widmo abstrakcyjne ma charakter widma uśrednionego (*obiekt typowy*), natomiast kolejne mają charakter widm różnicowych¹⁷. Zwykle dalszej analizie poddaje się te widma abstrakcyjne, które wizualnie wykazują nieprzypadkowy charakter, tzn. różnią się od sygnału szumu losowego. Należy zauważyć, że w przypadku wykonaniu procedury centrowania widm wymiar przestrzeni k istotnych PC zmniejsza się o jeden (stąd spotykane w literaturze oznaczenie PC0 dla widma typowego), a PC1, PC2, ... informują o różnicach w strukturze danych w stosunku do widma uśrednionego.

Zastosowanie powyższego schematu postępowania pozwoliło w przypadku obu próbek, In/CuPc/HOPG oraz In/CuPc/InSb, wskazać na obecność trzech istotnych głównych składowych¹⁸. Wniosek ten w postaci dendrogramu potwierdziła również analiza skupień (*Cluster Analysis*). W związku z tym przyjęto hipotezę, że do prawidłowej interpretacji zestawu widm UPS należy założyć obecność trzech faz indu - metalicznej, fazy aktywnej chemicznie i

¹⁷ Beattie J.R., Esmonde-White F.W.L., Exploration of Principal Component Analysis: Deriving Principal Component Analysis Visually Using Spectra, *Applied. Spectr.* 2021,75(4),361-375.

¹⁸ Balcerowska-Czerniak G., Dittmar-Wituski A., Hiller T., and Bukaluk A., Efficiency of chemometric methodology for characterization of In/CuPc films on HOPG and InSb based on ultraviolet photoelectron spectra, *Inżynieria Materiałowa* 2016-5.

fazy nieaktywnej. Jednakże uzyskane informacje były na tyle niepełne, że nie pozwalały na jednoznaczne przypisanie poszczególnym fazom położenia spektralnych. Rozwiązaniem problemu okazało się niestandardowe podejście, polegające na tworzeniu nowych profili widm abstrakcyjnych, uzyskanych poprzez skumulowane sumowanie efektów przypisywanych kolejnym głównym składowym.

Schemat postępowania, który zaproponowałam w celu uzyskania nowych informacji z kształtu widm, jest adaptacją na grunt przestrzeni głównych składowych podejścia stosowanego w analizie sygnałów cyfrowych; przeprowadza się sumowanie sygnałów w celu wygładzenia szumów i uwypuklenia nieprzypadkowych kształtów przy użyciu np. filtrów cyfrowych.

Zaproponowany w pracy **A3** schemat tworzenia skumulowanych sum, tzn. $PC1+PC2$, $PC1+PC2+PC3$, itd., na podstawie wektorów *loadings* nazwano wskaźnikiem RSSL (*Root of the Sum of the Squared Loadings*). Analiza kształtu wskaźnika RSSL, uzyskanego poprzez sumowanie efektów z ośmiu pierwszych głównych składowych, ujawniła obecność spornego dubletu (rysunek nr 2 w pracy **A3**), którego pozycja energetyczna pozostawała stała niezależnie od liczby głównych składowych włączanych do obliczania wskaźnika RSSL oraz rodzaju podłoża. Dodatkowo można było potwierdzić, że za wspomniany wcześniej trend przesuwania się dubletu In4d w miarę narastania warstwy indu na ftalocyjaninie odpowiada jednoczesna obecność blisko położonych dwóch dubletów. Pozycja energetyczna jednego z dubletów niemal idealnie odpowiadała wartościom literaturowym dla fazy metalicznej indu (16,8eV i 17,7eV). O pochodzeniu drugiego dubletu, przesuniętego o 0,2eV w stronę wyższych energii wiązania w stosunku do fazy metalicznej, wnioskowano w oparciu o metodę MCR-ALS (*Multivariate Curve Resolution-Alternating Least Squares*). Wyniki analizy ilościowej (rysunek 4 b w pracy **A3**) potwierdzały finalną dominację tej fazy w próbkach na podłożu HOPG, natomiast z określenia widm form spektralnych wynikało, że jest to faza niezależna, nieskorelowana z warstwą organiczną (kształt fazy A na rysunku 4a). W związku z powyższym uznano, że powstanie tej fazy jest wynikiem formowania się klastrów indu, zgodnie z propozycją przedstawioną w jednej z wcześniejszych prac.

Dzięki wizualizacji chemometrycznej w postaci standardowego biplotu (rysunek nr 3 w pracy **A3**), można było wykazać, że podczas osadzania indu na warstwie organicznej ftalocyjaniny, w początkowej fazie preferowane jest powstawanie związku In_2CuPc , dalszy proces osadzania sprzyja tworzeniu klastrów indu, na których to tworzy się warstwa metaliczna. Ponadto wykazano, że istotną rolę w procesie formowania się układu metal-warstwa organiczna, a co za tym idzie stopniu zdominowania przez jedną z faz, duże znaczenie odgrywa podłoże. Podłoże HOPG sprzyjało powstawaniu klastrów indu, co z kolei nie sprzyjało

formowaniu się finalnie warstwy metalicznej, natomiast faza reagująca tworzyła się na obydwu podłożach w podobny sposób. Rezultatem przeprowadzonych analiz było potwierdzenie faktu, że wyniki uzyskane przez obydwie grupy badaczy nie są sprzeczne, ponieważ odnoszą się do różnych faz, dominujących w mniejszym lub większym stopniu, którym odpowiadają różne wartości energii wiązania.

Niestandardowy charakter opracowanego podejścia wynika nie tylko z zastosowania nowego wskaźnika, lecz także ze sposobu doboru liczby głównych składowych, uwzględnianych przy jego obliczaniu. W analizie danych spektroskopowych zwykle pomija się te główne składowe, którym odpowiadają stosunkowo niewielkie wartości własne, uznawane za niosące znikomą ilość informacji związaną z zaburzeniami aparaturowymi. W omówionym przykładzie pierwsze trzy główne składowe wyjaśniały odpowiednio 77%, 17% oraz 3% wariancji. Wprowadzenie nowego wskaźnika opartego na większej niż wskazywały kryteria liczbie głównych składowych, w tym przypadku ośmiu, pozwoliło na przeprowadzenie wiarygodnego wnioskowania. W konsekwencji przeprowadzonych badań osiągnięto założony cel nr II; wykazano, że zastosowanie analizy PCA w niestandardowy sposób pozwala pozyskać nowe, niedostępne informacje, w tym wypadku zwiększające zdolność metody PCA do wykrywania składników widm o niewielkim udziale.

PCA w analizie próbek z rozbudowanymi zbiorami danych

Współczesne analizy coraz częściej obejmują duże, zróżnicowane zestawy danych – zarówno pomiarowych, jak i pochodzących z rozbudowanych baz referencyjnych. Tradycyjna analiza bywa wówczas czasochłonna i skomplikowana, co w konsekwencji może prowadzić do niejednoznacznych wniosków. W takich sytuacjach kluczowe staje się zastosowanie metod takich jak analiza głównych składowych, pozwalających uporządkować i zsyntetyzować nadmiar informacji. Jednak w przypadku niektórych problemów analitycznych zastosowanie PCA wymaga odejścia od typowego schematu postępowania. W szczególności dotyczy to próbek o złożonej strukturze i nieznanym składzie, które były przedmiotem badań opisanych w pracach **A4** i **A5**.

Celem badań opisanych w pracy **A4** była identyfikacja na podstawie dyfraktogramów XRD (*X-ray Diffraction*) związków chemicznych odpowiedzialnych za powstawanie danego koloru w próbkach tytanu poddanych procesowi laserowego kolorowego znakowania. Doniesienia literaturowe wskazywały na powstawanie znacznie zróżnicowanej grupy

związków chemicznych towarzyszących procesowi wytwarzania próbek, głównie tlenków, choć odnotowywano także obecność azotków^{19,20}. Ze względu na warunki wytwarzania próbek lista składników, które wymagały szczegółowego przeanalizowania na podstawie bazy ICDD (*International Centre for Diffraction Data*) obejmowała ponad 600 potencjalnych związków tytanu, głównie z tlenem oraz azotem. Jakikolwiek próby dopasowania pozycji dziesięciu charakterystycznych pików widocznych w dyfraktogramach (rysunek 5 w pracy A4) do związków chemicznych z bazy referencyjnej nie przyniosły jednoznacznych odpowiedzi.

W znanej chemometrycznej metodzie SIMCA (*Soft Independent Modeling of Class Analogy*)²¹, wykorzystującej metodę głównych składowych do modelowania wielu grup/klas jednocześnie, stosuje się schemat postępowania polegający na budowaniu modelu PCA indywidualnej grupy na podstawie danych referencyjnych. Konstrukcja jakiegokolwiek modelu wymaga zdefiniowania celu modelowania – zwykle celem tym jest przypisanie nowej próbki do danej klasy na podstawie analizy odległości nowej próbki od centrum danej klasy. W praktyce oznacza to użycie przestrzeni danych referencyjnych do testowania nowej próbki. W przypadku próbek tytanu takie postępowanie okazało się zupełnie nieskuteczne. Ze względu na dużą różnorodność danych referencyjnych, model dwóch głównych składowych zbudowany na ich podstawie opisywał mniej niż 30% wariancji. Rozwiązaniem problemu stało się odwrócenie kolejności etapów postępowania w „schemacie SIMCA”. Model dwóch głównych składowych (PC1-PC2) został zbudowany w oparciu o dane badawcze (dyfraktogramy) próbek tytanu, następnie dane referencyjne z bazy ICCD zostały użyte jako testujące a dalszy schemat związany z wnioskowaniem na podstawie mapy obiektów przebiegał dokładnie tak jak w metodzie SIMCA.

Aby ograniczyć liczbę potencjalnych związków chemicznych użytych do testowania dokonano ich wstępnej selekcji wykorzystując w tym celu znane w PCA wskaźniki statystyczne – statystykę Hotellinga T^2 oraz statystykę Q reszt. Szczegółowo wskaźniki te będą omówione w dalszej części niniejszego opracowania. Należy podkreślić, że selekcja danych referencyjnych również wymagała niestandardowego podejścia. Symulacje numeryczne dla

¹⁹ Pérez del Pino A., Serra P., Morenza J.L., Coloring of titanium by pulsed laser processing in air, *Thin Solid Films* 2002, 415: 201–205.

²⁰ Lavis L., Sahour M.C., Jouvard J.M., Pillon G., Marco de Lucas M.C., Bourgeois S., Grevey D., Growth of titanium oxynitride layers by short pulsed Nd:YAG laser treatment of Ti plates: influence of the cumulated laser fluence, *Appl. Surf. Sci.* 2009, 255:5515–5518.

²¹ Wold S., Pattern recognition by means of disjoint principal components models, *Pattern Recognition* 1976, 8 (3), 127-139.

widm w postaci krzywych Gaussa wskazały, że za potencjalnie istotne związki należy uznać te, dla których wartości statystyki Hotellinga T^2 przyjmują najwyższe wartości. W standardowej analizie PCA obiekty o odległościach Mahalanobisa (odpowiednik statystyki T^2) od centrum modelu przekraczających wartość progową nie są uwzględniane w modelu²². W przypadku statystyki Q reszt zastosowano klasyczny schemat eliminacji obiektów przekraczających wartość progową.

Efektywność zaproponowanego schematu postępowania została w pierwszym etapie sprawdzona dla układu próbek cienkich warstw złota oraz złota z cyną, naniesionych na podłoże ze szkła (rysunek nr 4 w pracy **A4**). Próbkę te były podstawą innych szczegółowych badań²³, stąd wnioskowanie mogło być zweryfikowane z wcześniejszymi wynikami. W związku z tym, że analiza składu próbki na podstawie jej dyfraktogramu polega na przypisaniu charakterystycznych pików do składników chemicznych, odwrócony „schemat SIMCA” zastosowano do standardowego biplotu. W rezultacie możliwe było skorelowanie składników referencyjnych z pikami charakterystycznymi widocznymi w dyfraktogramach próbek za pomocą jednego grafu w przestrzeni PC1-PC2.

W przypadku próbek tytanu opracowany schemat postępowania chemometrycznego (rysunek nr 8 w pracy **A4**) pozwolił zidentyfikować związki chemiczne odpowiedzialne za powstanie w dyfraktogramach charakterystycznych pików widocznych na rysunku nr 5, oznaczonych literami A÷U; szczegółowy opis zawiera tabela nr 3 w pracy **A4**. Podsumowując, analiza na podstawie biplotu pozwoliła wyodrębnić związki dominujące podczas procesu laserowego „kolorowania” próbek tytanu. Podłoże, oprócz dominującej fazy β -Ti, wykazywało również obecność wodoru tytanu (TiH_2). W próbce postrzeganej wizualnie jako żółta dominował związek Ti_3O , w próbce pomarańczowej - Ti_6O , natomiast w próbce czerwonej dominującą fazą był związek Ti_2O_3 . Zaskakujące okazało się wskazanie obecności azotku tytanu (TiN) w próbce, której barwę oceniono jako niebieską; w literaturze²⁴ związek ten

²² Daszykowski M., Walczak B., Analiza czynników głównych i inne metody eksploracji danych, w Zuba D., Parczewski A., Chemometria w analityce, IES, Kraków, 2008.

²³ Wronkowska A.A., Czerniak G., Wronkowski A., Skowroński Ł., Optical and microstructural characterisation of Au-Sn and Cu-Sn diffuse layers, Appl. Surf. Sci. 2013, 281:30–37.

²⁴ Barhai P.K., Kumari N., Banerjee I., Pabi S.K., Mahapatra S.K., Study of the effect of plasma current density on the formation of titanium nitride and titanium oxynitride thin films prepared by reactive DC magnetron sputtering, Vacuum 2010, 84: 896–901.

wiązano z kolorem żółtym próbek. Późniejsze badania potwierdziły słuszność wnioskowania chemometrycznego opartego na metodzie PCA²⁵.

W wyniku przeprowadzonych badań opracowano uniwersalny schemat, którego zastosowanie wykracza poza technikę XRD, co może znaleźć zastosowanie także w innych technikach spektroskopowych opartych na rozbudowanych bazach danych referencyjnych.

W recenzji pracy zwrócono uwagę na wysoki potencjał opracowanego schematu do implementacji w oprogramowaniu zintegrowanym z przyrządami pomiarowymi.

Zastosowanie schematu SIMCA jest szczególnie wartościowe w sytuacjach wymagających precyzyjnego modelowania zmienności poszczególnych klas oraz ich rozróżnienia na podstawie odrębnych modeli PCA. Podejście to okazało się kluczowe w badaniach, opisanych w pracy A5, których celem była charakterystyka oraz ocena ilościowa tlenków, jakie tworzą się na powierzchni polikrystalicznego rodu w domenach o różnej orientacji krystalograficznej. Pierwszy etap badań dotyczył szczegółowego „mapowania” próbki z wykorzystaniem techniki XPS (*X-ray Photoelectron Spectroscopy*), w wyniku czego otrzymano 2000 widm z charakterystycznymi pikami fotoelektronów. Przeprowadzenie wiarygodnej analizy ilościowej przy użyciu oprogramowania CASAXPS® zintegrowanego z układem pomiarowym okazało się problematyczne. W efekcie nie udało się potwierdzić obecności pięciu domen krystalograficznych zidentyfikowanych na obrazie próbki uzyskanym metodą mikroskopii konfokalnej. Ze względu na niewielkie zróżnicowanie widm XPS, również zastosowanie metod chemometrycznych, takich jak analiza PCA czy analiza klastrowa (*CA - Cluster Analysis*) nie przyniosło oczekiwanych rezultatów w postaci jednoznacznego pogrupowania danych. W związku z tym pojawiła się konieczność pogrupowania widm XPS zgodnie z obrazem mikroskopowym próbki, przedstawionym na rysunku 2a w pracy A5.

Początkowe podejście polegające na wybraniu kilku punktów na powierzchni próbki, a następnie ich analiza techniką XPS, ze względu na brak pewności co do miejsca/domeny której analiza dotyczy wymagało modyfikacji. Ponieważ dysponowano obrazem próbki z mikroskopu konfokalnego narodził się pomysł, aby do analizy XPS wybrać fragment obszaru, który zostanie poddany dokładnemu mapowaniu w kierunkach x i y (rysunek 2a w pracy A5), w rezultacie czego otrzymano 2000 widm. Pozostał jedynie problem integracji dwóch typów informacji o wybranym obszarze próbki. Efektem moich badań było opracowanie metody fuzji informacji z obrazu z danymi spektroskopowymi.

²⁵ Antończak A., Skowroński Ł., Trzcinski M., Kinzhybalo V., Łazarek Ł., Abramski K., Laser-induced oxidation of titanium substrate: Analysis of the physicochemical structure of the surface and sub-surface layers, Appl. Surf. Sci. 2015, 325: 217-226.

Obraz mikroskopowy próbki, po konwersji do układu liczb RGB, poddano analizie klasterowej z wykorzystaniem aglomeracyjnej metody opartej na odległości euklidesowej. Uzyskany wzór w postaci jednoznacznego przypisania odpowiedniego piksela próbki do klasy/domeny, zastosowano do pogrupowania widm XPS. Informacje o przynależności widm do określonych domen pozwoliły następnie na przeprowadzenie odrębnych analiz PCA dla każdej z pięciu domen wyodrębnionych metodą CA. Wynikiem tych analiz było uzyskanie widm typowych – jednowymiarowych wektorów ładunków dla pierwszej głównej składowej – pełniących funkcje widm referencyjnych dla poszczególnych domen. O jednorodności widm w ramach domeny (każda zawierająca od 285 do 493 widm) wnioskowano na podstawie wartości wariancji odtwarzanej przez PC1 (tabela 1 w pracy **A5**) oraz na podstawie analizy kształtów widm resztowych. Kolejnym etapem była analiza ilościowa przy użyciu zintegrowanego z przyrządem oprogramowania. Jako dane wejściowe wykorzystano informacje o położeniach pików oraz liczbie form spektralnych wynikające z ponownej analizy PCA dla wyodrębnionych pięciu widm referencyjnych. Na tej podstawie dokonano analizy ilościowej poszczególnych domen krystalograficznych (tabela 3 w pracy **A5**).

Niestandardowość zastosowanego w pracy **A5** podejścia nie wiąże się z modyfikacją samego schematu analitycznego, lecz jest związana z innowacyjną integracją danych z dwóch technik pomiarowych analizowanych przy użyciu uzupełniających się metod chemometrycznych.

Przedstawione powyżej przykłady trzech tematów badawczych związanych z analizą danych spektroskopowych dowodzą, że analiza chemometryczna stanowi cenne wsparcie, a czasami jest kluczowa, w rozwiązywaniu rozmaitych problemów w badaniach próbek o złożonej strukturze i składzie chemicznym, co potwierdza osiągnięcie celu badawczego nr II.

3.3.3 Metoda PCA w opracowaniu nowego statystycznego indeksu jakości

Interdyscyplinarność prowadzonych badań naukowych staje się obecnie kluczem do rozwiązywania problemów w obszarach związanych z eksploracją danych eksperymentalnych. W sytuacji gdy techniki pozyskiwania danych są dobrze określone, użycie nowej strategii analizy może przyspieszyć znacznie wnioskowanie.

Metoda PCA stała się podstawą opracowania nowego indeksu statystycznego nazwanego TQ_PCA, służącego kontroli jakości, co jest istotną częścią mojego osiągnięcia naukowego. Praca badawcza nad stworzeniem relatywnie prostego narzędzia do monitorowania ogólnego stanu wód była odpowiedzią na potrzeby instytucji rządowych odnośnie wczesnego ostrzegania

przed potencjalnymi zagrożeniami dla środowiska. Motywacją do badań stała się katastrofa ekologiczna na Odrze, która miała miejsce na przełomie lipca i sierpnia 2022 roku.

W Polsce stan wód powierzchniowych bada i ocenia w ramach państwowego monitoringu Inspekcja Ochrony Środowiska; pod uwagę brane jest wiele parametrów w tym fizykochemicznych, biologicznych oraz hydromorfologicznych. Po szczegółowej analizie eksperckiej dokonuje się oceny stanu próbki poprzez nadanie jej jednej z 5-stopniowej skali rang jakości²⁶.

Moim głównym celem stało się opracowanie rozwiązania, na wzór stosowanego wskaźnika WQI (*Water Quality Indeks*)²⁷, które integrowałoby informacje niesione przez każdy parametr z osobna w jednym indeksie statystycznym. Procedura powinna być na tyle prosta, aby można ją było zastosować na stacjach pomiarowych w ramach monitoringu on-line.

Główny problem ze stosowaniem indeksu WQI polega na jego niespójnej definicji – jest wyliczany jako średnia ważona, a wartość zależy od liczby uwzględnianych parametrów. Dodatkowy problem, moim zdaniem, stanowi jego przyjęta interpretacja w zakresie wartości od 0 (bardzo zła jakość) do 100 (bardzo dobra jakość); przy katastrofach, taka skala może być niewystarczająca.

Idea indeksu TQ_PCA oparta jest na teście Hotellinga T^2 , który jest uogólnieniem znanego rozkładu testowego t-Studenta i pozwala obrazować wielowymiarowe obserwacje na jednym wykresie w postaci jednego wskaźnika. Podejście wielowymiarowe oparte jest na założeniu, że monitorowane parametry są zwykle skorelowane, co skutkuje tym, że zmiana jednego parametru wpływa na pozostałe w określony sposób i nawet małe wzajemne zmiany wnoszą wkład do zsumowanej kowariancji.

Dla każdej próbki/obserwacji $\mathbf{x}$ (wiersz macierzy $\mathbf{X}$) wylicza się parametr T^2 na podstawie wzoru

$$T^2 = (\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{Z}^{-1} (\mathbf{x} - \bar{\mathbf{x}}) \quad (9)$$

gdzie $\bar{\mathbf{x}}$ oznacza obiekt średni a $\mathbf{Z}$ - macierz kowariancji, litera T w indeksie górnym oznacza transpozycję macierzy. Tak wyliczony parametr przedstawia się graficznie w funkcji numeru próbki, z zaznaczoną granicą ufności, obliczoną z użyciem rozkładu F jako granicą kontrolną²⁸.

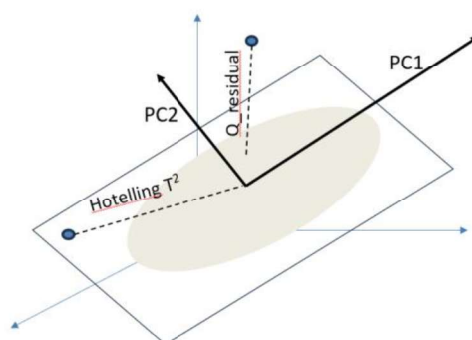
²⁶<https://wody.gios.gov.pl/pjwp/publication/568>

²⁷ W literaturze istnieje wiele sposobów wyliczania wskaźnika WQI; pomimo gotowych kalkulatorów ogólnie wymagana jest wiedza ekspercka do jego obliczenia.

²⁸ W zagadnieniach kontroli jakości produkcji znane są karty Shewharta, z linią centralną (LC), która odpowiada wartości oczekiwanej oraz granicami kontrolnymi - UCL (*upper control limit*) oraz LCL (*Lower Control Limit*).

Pomimo zalet karty kontrolnej Hotellinga T^2 , ma ona pewne wady; może prowadzić do wskazań zarówno fałszywie pozytywnych, jak i fałszywie negatywnych. Aby poprawić wydajność testu T^2 zaproponowano dwa typy podejść. Jeden typ opiera się na pewnej modyfikacji tej karty i jest zalecany, gdy liczba próbek jak i zmiennych jest mała. Drugi typ, integruje indeks T^2 w ramach analizy głównych składowych (PCA).

W celach statystycznej kontroli jakości z użyciem metody PCA w trybie on-line należy zastosować znany w literaturze schemat uczenia maszynowego, co wiąże się z budową modelu PCA w oparciu o próbki referencyjne (dobrej jakości). Zbudowanie modelu, który byłby stabilny tzn. nie zaburzony obiektami/próbkami odległymi, wymaga ich weryfikacji pod względem dopasowania do modelu. Metoda PCA jest techniką rzutowania/projekcji danych na wybrane podprzestrzenie, zwykle dwuwymiarowe, więc sama inspekcja map obiektów nie jest odporna na obiekty odległe w prostokątnym do płaszczyzny wymiarze. Standardowo w celach weryfikacji wykorzystuje się dwa indeksy statystyczne - statystykę Hotellinga T^2 oraz statystykę Q (reszt od stabilnego modelu). Takie podejście pozwala na analizę danych w więcej niż dwóch wymiarach przestrzeni PC. Oba indeksy wyrażają odległości w przestrzeni głównych składowych a ich ideę przedstawia Ryc.3.



Ryc. 3 Rysunek poglądowy przedstawiający diagnostykę odległości dwóch obiektów od modelu PCA (elipsa).

Aby ułatwić diagnostykę obiektów odległych, wartości obu statystyk przedstawia się graficznie tak jak obrazuje to rysunek nr 4 w pracy A6. Ogólnie, statystyka Q jest znacznie bardziej wrażliwa na zmiany atypowe w obserwacjach w stosunku do statystyki T^2 . W metodzie opartej na PCA, aby ułatwić optymalną wydajność i zachować ten sam wpływ na wykrywanie obiektów odległych indeksy te są normalizowane. Zwykle statystyki są dzielone przez odpowiadające im 95% granice ufności obliczone na podstawie konkretnych wartości danych modelu i nazywane odpowiednio T^2 zredukowane i Q-reszta zredukowana. Próbką, dla której T^2 zredukowane > 1 (duża odległość próbki od środka modelu PCA), i jednocześnie Q-reszta zredukowana > 1 (odległość między próbką a jej prostokątną projekcją do modelu PCA) jest

wskazywana jako odstająca/atypowa. Za próbki regularne uznaje się te poniżej przyjętego limitu odcięcia.

W budowaniu modelu głównych składowych kluczowe są dwa aspekty - wybór metody wstępnej transformacji danych/parametrów (szczególnie ważny, gdy dane mają zróżnicowany charakter pod względem zakresu i jednostek) oraz wybór liczby k istotnych głównych składowych. Po zbudowaniu stabilnego modelu PCA, kontrola jakości nowej próbki, zgodnie z przedstawioną w pracy **A6** koncepcją, opiera się na wyliczeniu indeksu TQ_PCA dla tej próbki na podstawie wyrażenia

$$TQ_PCA = (\sqrt{(T^2)^2 + Q^2})/\sqrt{2} \quad (10)$$

przy czym T^2 oraz Q odnoszą się do wartości zredukowanych.

Pilotażowe badanie przeprowadzone w 2018 roku na danych Ocean Observatory (Norwegia)²⁹, wskazywało na potencjał metody PCA w automatycznym monitoringu jakości wód, wykorzystującym 12 sensorów dla danych z okresu od 16 maja do 31 sierpnia. Autorzy zastosowali schemat uczenia maszynowego: zbudowano model kalibracyjny na podstawie danych z maja-lipca, który następnie wykorzystano do testowania danych z sierpnia. Wyniki przedstawiono w postaci mapy obiektów w przestrzeni PC1-PC2. Dodatkowo, w pracy wykorzystano statystyki – T^2 oraz Q , jako wskaźniki wspomagające walidację wyników oraz identyfikację parametrów odbiegających od danych ze zbioru kalibracyjnego. Walidację zaprezentowano oddzielnie dla każdej ze statystyk, bez użycia procedury normalizacji, co skutkowało znacznymi różnicami w ich zakresach wartości.

Z praktycznego punktu widzenia obserwacja dwóch parametrów, czy to na osobnych kartach kontrolnych czy na dwu-wymiarowym wykresie jest kłopotliwa. Stąd moja propozycja połączenia dwóch statystyk w jeden indeks TQ_PCA .

Wzór (10) stanowi analogię między statystykami T^2 i Q , traktowanymi jako miary odstępstw od modelu PCA, a miarami niepewności stosowanymi w metrologii. Jeśli dwie wielkości, x i y , zostały zmierzone niezależnie i obie podlegają rozkładowi normalnemu, wówczas niepewność wielkości będącej ich sumą ($q=x+y$) oblicza się jako pierwiastek z sumy kwadratów dwóch miar niepewności x i y na podstawie wzoru

$$u(q) = \sqrt{u(x)^2 + u(y)^2}$$

Podzielenie przez $\sqrt{2}$ we wzorze (10) sprawia, że wartość granicy kontrolnej wynosi 1, co jest najprostsze z praktycznego punktu widzenia.

²⁹ Eide I., Westad F., Automated multivariate analysis of multi-sensor data submitted online: Real-time environmental monitoring. PLoS ONE 2018, 13(1):e0189443.

Postać wyrażenia (10) budziła wątpliwości jednego z recenzentów pracy **A6**; nieuzasadnione wydaje się podnoszenie do kwadratu wielkości już będących kwadratem. Jednak główny zarzut dotyczył braku analizy parametrów biologicznych wód. W związku z tym wystąpiłam z pisemną prośbą do stacji we Frankfurcie nad Odrą o udostępnienie danych odnośnie parametrów fizykochemicznych oraz biologicznych, które obejmowały również okres sprzed katastrofy. Po uzupełnieniu analiz zwróciłam się do prof. Franka Westada, autora wspomnianego pilotażowego badania, o wyrażenie opinii na temat treści manuskryptu. On również zgłosił zastrzeżenia co do postaci wzoru (10). Jednakże, z mojego doświadczenia podczas analiz danych dla Odry wynikało, że wzór w postaci zwykłej sumy podzielonej przez 2, jak sugerowano, prowadził do zbyt wysokiego oszacowania indeksu TQ_PCA. Oznaczało to np. ciągłe alarmy zgłaszane przez kryterium, gdy wartości parametrów odbiegały bardzo nieznacznie od referencyjnych, co z kolei ze względu na niepewności pomiarowe jest dopuszczalne przy danych eksperymentalnych. Od prof. Franka Westada otrzymałam nie tylko uwagi dotyczące pracy, lecz także skrypt (Matlab) umożliwiający wyznaczanie „contribution plots”, czego pierwotna wersja pracy nie uwzględniała. W skrypcie dokonałam modyfikacji w celu dostosowania do postaci indeksu TQ_PCA. Nawiązany kontakt zaowocował współpracą przy przygotowywaniu pracy **A7**.

Użyteczność zaproponowanego indeksu TQ_PCA została przetestowana i opisana w pracach **A6** i **A7** dla zbiorów danych publicznie dostępnych, dotyczących jakości wód powierzchniowych rzeki Odry.

Analiza dystrybucji indeksu TQ_PCA, zastosowanego do zbioru danych fizykochemicznych i biologicznych ze stacji niemieckiej/Frankfurt obejmujących okres katastrofy, wykazała typowy dla modeli epidemii kształt krzywej Gaussa. Wynik ten przedstawiono na rysunku nr 9a w pracy **A6**. Jest to wynik interesujący, ponieważ uzyskanie go z punktu widzenia klasycznych analiz nie jest możliwe. Jednakże, szczególnie istotna z punktu widzenia zapobiegania podobnym wydarzeniom w przyszłości okazała się inna obserwacja. Indeks TQ_PCA przekraczał kryterium alarmowe miesiąc wcześniej niż to wynikało z oficjalnych komunikatów i utrzymywał się powyżej niego aż do momentu załamania się ekosystemu rzeki, co pokazuje rysunek 9b w pracy **A6**. Takie zachowanie potwierdza fakt, że indeks statystyczny rzeczywiście jest wrażliwy na niewidoczne dla oka inżyniera obsługującego stację synergie, zachodzące między zmiennościami różnych parametrów fizyko-chemicznych.

Te ustalenia wzbudziły zainteresowanie mediów; m.in. ukazały się artykuły w Gazecie Wyborczej pt. „Katastrofę na Odrze można było wykryć z miesięcznym wyprzedzeniem. Naukowcy z Bydgoszczy mają metodę” oraz Zielonej Interii pt. „Katastrofy ekologiczne są do uniknięcia. Polacy opracowali skuteczną metodę”.

Otrzymane na podstawie wzoru (10) wyniki wartości indeksu TQ_PCA wymagały weryfikacji zewnętrznej. W związku z dostępnymi od początku katastrofy zdjęciami

satelitarnymi zakwitów chlorofilu (<https://www.igb-berlin.de/en/news/new-analysis-satellite-data-confirm-massive-algal-bloom-oder-river>), zaistniała możliwość porównania tych obrazów z otrzymanymi wynikami obliczeń. Chlorofil jest uznanym indeksem hydrobiologicznym obecności w Odrze złotych alg³⁰. Na wzór przypisywanych kategorii klasyfikujących próbki pod względem ich jakości dla stacji polskich przyjęto następujące rangi ryzyka zdarzeń niepożądanych:

- $TQ_PCA \leq 1$ niskie ryzyko
- $1 < TQ_PCA \leq 3$ średnie ryzyko
- $TQ_PCA > 3$ wysokie ryzyko

Rezultat w postaci map przestrzennych prezentuje rysunek nr 6 w pracy **A6**. Jak można się przekonać korelacja pomiędzy tymi dwoma sposobami prezentacji sytuacji na Odrze w dniach 11÷26 sierpnia 2022 roku jest znacząca. Charakter indeksu TQ_PCA dobrze odzwierciedla spływającą w okolicach Odry środkowej falę wód złej jakości.

Obecnie na stronie <https://badania.gios.gov.pl/odra/> dostępna jest interaktywna mapa klasyfikująca miejsca pobranych próbek wody z rzeki Odry pod kątem ryzyka zakwitów złotych alg, których toksyny uznano za przyczynę masowego śnięcia ryb. W tym celu opracowano specjalną procedurę klasyfikacyjną³¹. Prowadzenie ciągłych badań laboratoryjnych fitoplanktonu (badania mikroskopowe) jest jednak kosztowne i czasochłonne, a w literaturze opisywane są przypadki pojawiania się złotych alg w miejscach, w których wcześniej nie były rejestrowane. W związku z tym, w przypadku zagrożenia potencjalnym zakwitem, zwłaszcza toksycznym, przewidywanie warunków jego wystąpienia jest kluczowe dla służb monitorujących.

Celem pracy **A7** stało się przystosowanie indeksu TQ_PCA do informowania o potencjalnym zakwicie fitoplanktonu. Wyniki z pracy **A6** w postaci map przestrzennych (rysunek nr 6) były na tyle obiecujące, że osiągnięcie założonego celu wydawało się możliwe. W pracy **A7** zaproponowano budowanie modelu PCA na potrzeby wyliczeń indeksu TQ_PCA do przewidywania zakwitów fitoplanktonu w oparciu o wyselekcjonowane parametry fizykochemiczne. Wśród tych parametrów znalazły się: temperatura wody, tlen rozpuszczony (DO), odczyn pH, TN/TP (stosunek wagowy azotu do fosforu, ważny z punktu widzenia

³⁰ Absalon, D., Matysik, M., Woźnica, A., Janczewska, N., Detection of changes in the hydrobiological parameters of the Oder River during the ecological disaster in July 2022 based on multi-parameter probe tests and remote sensing methods. Ecol. Indic. 2023, 148, 110103.

³¹ Procedura monitorowania interwencyjnego *Prymnesium Parvum* „złotej algi” (GIOŚ)

limitacji fitoplanktonu) oraz nowy indeks zasolenia (*salinity_idx*), określający udział jonów siarczanów do sumarycznej ilości jonów siarczanów oraz chlorków.

W trakcie badań nad optymalnym zestawem parametrów zauważyłam pewną zależność – bezwzględne wartości zarówno przewodności elektrolitycznej jak i stężenia siarczanów czy chlorków, czyli wskaźników charakteryzujących zasolenie, niekoniecznie korelowały z podanymi ilościami fitoplanktonu, stąd pojawiła się potrzeba stworzenia nowego indeksu, który byłby mniej wrażliwy na zmiany wartości stężeń. Jak się okazało duże zakwity występowały tylko w pewnym wąskim przedziale wartości zaproponowanego indeksu zasolenia.

Wprowadzenie nowego indeksu *salinity_idx* pozwoliło zinterpretować fakt występowania zakwitów również przy niższych wartościach ogólnego zasolenia. Przedstawiony w postaci graficznej charakter indeksu *salinity_idx* (rysunek nr 2 w pracy A7) pozwolił wyróżnić dwa typy wód - jeden związany z głównym nurtem Odry (I), drugi dla obszaru Kanału Gliwickiego (II). Podobny charakter z podziałem na dwa typy można znaleźć w pracy³².

Analiza przebiegu funkcji indeksu TQ_PCA, wyliczonego na podstawie pięciu wybranych parametrów fizykochemicznych z okresu od stycznia do września 2024 (rysunek nr 3 w pracy A7), wskazuje na jego przydatność w ocenie ryzyka wystąpienia zakwitu. Wysoka korelacja indeksu z występowaniem zakwitu złotych alg jest szczególnie widoczna dla stacji Wrocław-Łany w ujęciu jakościowym. Indeks dla stacji Pławniowice z drugiej połowy lata 2024 roku również dobrze koreluje z zakwitem. W tym czasie na zbiorniku Dzierżno Duże dochodziło do masowego śniecia ryb oraz innych organizmów wodnych.

Z założenia indeks jakości TQ_PCA podaje informację ogólną o stanie wód. Jednak w przypadku przekroczenia wartości progowej informacje o parametrze lub parametrach odpowiedzialnych za takie zachowanie stają się istotne z punktu widzenia instytucji nadzorujących stan środowiska; można podjąć działania ograniczające czasowo zasilanie wód pod kątem substancji odpowiedzialnych za wzrost poziomu odpowiednich parametrów.

W celu oszacowania wkładu poszczególnych parametrów w wartość indeksu TQ_PCA w pracy A6 zastosowano tzw. rysunki udziałów (*contribution plots*), które w pracy A7 stanowią jej istotną część. Z wykresów udziałów (rysunek nr 4 w pracy A7) można dokonać oceny wpływu poszczególnych parametrów na wartość indeksu TQ_PCA. Dla polskich stacji największy wkład ma wzrost wartości pH wody, co jest zgodne z udokumentowanym danymi

³² Ruman, M., Janczewska, N., Kosek, K., Artichowicz, W., Nasiek, M., Matysik, M., Fluvial ecology disasters: the impact of the Gliwice Canal on the ecological crisis in the Oder River basin, Poland. *Acta Geophysica*. 2024, 73:775-798.

literaturowymi związkami pomiędzy odczynem pH a rozwojem *Prymnsium Parvum*³³. W miejscach, gdzie rejestrowano najwyższe wartości fitoplanktonu (rzędu 10^8 osobników na litr) na stacjach Pławniowice oraz Kanał Kędzierzyński istotne są również zwiększone wartości natlenienia, temperatury oraz indeksu zasolenia, z którymi ujemnie koreluje TN/TP.

Podsumowując, można wskazać na następujące zalety indeksu TQ_PCA

- szybka klasyfikacja jakości próbki w trybie on-line,
- brak potrzeby informacji o limitach kontrolnych dla poszczególnych parametrów fizyko-chemicznych, biologicznych, hydromorfologicznych,
- czytelna wartość limitu kontrolnego, odpowiadająca poziomowi referencyjnemu,
- wzrost wartości przy obniżonej jakości, a co za tym idzie brak ograniczeń co do wartości maksymalnej,
- możliwość śledzenia wzajemnych zmian parametrów w postaci wykresów udziałów,
- możliwość ograniczenia szczegółowych analiz tylko do przypadków z wskazanym przez indeks poziomem ryzyka,
- możliwość użycia w dowolnych badaniach obejmujących duże zbiory parametrów monitorowanych w dowolnym procesie.

Opracowanie koncepcji indeksu TQ_PCA, umożliwiającego jednoczesne monitorowanie wielu parametrów fizykochemicznych i biologicznych wód powierzchniowych oraz wczesne wykrywanie próbek atypowych na przykładzie danych rzeki Odry, stanowi realizację celu badawczego nr III.

III. Omówienie pozostałych osiągnięć naukowo-badawczych

1. Działalność przed uzyskaniem stopnia naukowego doktora

Po ukończeniu studiów rozpoczęłam pracę w Zakładzie Fizyki Doświadczalnej Instytutu Matematyki i Fizyki Akademii Techniczno-Rolniczej (ATR) w Bydgoszczy. Prowadzona pod kierunkiem prof. Ryszarda Siudy działalność dotyczyła spektroskopii elektronów Augera, badania procesów zachodzących na powierzchniach metali oraz w głównej mierze zastosowania metod wielowymiarowej analizy statystycznej, znanych pod nazwą chemometria, takich jak analiza głównych składowych i analiza czynnikowa, do interpretacji wyników pomiarów w spektroskopii elektronowej. Szkolenie w dziedzinie chemometrii oraz

³³ Caron, D.A., Lie, A.A.Y., Buckowski, T., Turner, J., Frabotta, K., The effect of pH and salinity on the toxicity and growth of the golden alga, *Prymnesium parvum*. Protist 2023, 174 (1), 125927.

kurs programowania w środowisku MATLAB otrzymałam w ramach 3 miesięcznego stażu na Uniwersytecie w Umea w Szwecji w 1996 roku.

Zagadnienia, którymi się zajmowałam można podzielić na dwie kategorie. Jedna kategoria obejmowała problemy istotne z punktu widzenia interpretacji złożonych linii widmowych. Druga dotyczyła wpływu niedoskonałości układu pomiarowego na wyniki analizy jakościowej widm, w szczególności na strukturę modelu głównych składowych analizowanych widm i wyciągane na jej podstawie wnioski. Wyniki badań, prezentowane były przeze mnie na wielu konferencjach, m.in. w 2001 roku podczas międzynarodowej konferencji ECASIA'01 (*European Conference on Application of Surface and Interface Analysis*) w Avinionie we Francji. W sumie powstało 8 współautorskich prac o zasięgu międzynarodowym, które stały się podstawą mojej pracy doktorskiej.

2. Działalność po uzyskaniu stopnia naukowego doktora

Od 2003 roku stałam się członkiem nowo powstałej grupy, pod kierunkiem prof. Ryszarda Siudy, zajmującej się tematyką agrofizyki. Grupa nawiązała współpracę z Katedrą Fitopatologii Wydziału Rolniczego oraz Katedrą Biochemii i Biotechnologii Zwierząt Wydziału Zootechnicznego Akademii Techniczno-Rolniczej w Bydgoszczy. Efektem wspólnych starań było przygotowanie wniosku grantowego na podjęcie prac w celu opracowania postępowania, które na podstawie analizy widm odbiciowych metodami chemometrycznymi, pozwoliłoby oznaczać poziom mikotoksyn w ziarnie pszenicy. Moim zadaniem było wykonywanie pomiarów optycznych z użyciem spektrofotometru Cary 5000 z przystawką DRA-2500 do materiałów sypkich, oraz opracowanie procedury chemometrycznej analizy danych jako alternatywy dla drogich i czasochłonnych metod chromatograficznych. Projekt grantowy został zakończony w 2006 roku. W międzyczasie w latach 2004-2005 realizowany był program umowy międzynarodowej z Instytutem Analizy Roślin w Quedlinburgu (Niemcy). Tematem współpracy był rozwój metod spektroskopii optycznej służącej szybkiej detekcji wybranych szkodliwych substancji w roślinach i nasionach. Tematyce chorób ziarna pszenicy poświęcony był również drugi projekt grantowy realizowany w latach 2006-2007 jako efekt współpracy z wyżej wymienionymi jednostkami ATR oraz dodatkowo z Katedrą Chemii Rolnej Wydziału Rolniczego ATR oraz Instytutem Agrofizyki PAN w Lublinie. Efektem badań z moim udziałem była prezentacja wyników na wielu konferencjach krajowych oraz podczas Międzynarodowego Sympozjum „Recent Advaces in Food Analysis” w Pradze w 2007 roku.

W 2011 roku, po powrocie z urlopu macierzyńskiego, dołączyłam do grupy zajmującej się badaniem właściwości optycznych cienkich warstw metodą clipsometrii spektroskopowej,

kierowanej przez doc. dr Aleksandrę Wronkowską. W ramach tej współpracy powstała praca A4, w której po raz pierwszy zaproponowałam niestandardowe wykorzystanie metody PCA. W tym czasie rozpoczęłam bliską współpracę z prof. Bogumiłą Kupcewicz z Katedry Chemii Nieorganicznej i Analitycznej Wydziału Farmaceutycznego Collegium Medicum UMK. Wspólne badania dotyczyły wykorzystania technik chromatograficznych i spektroskopowych w analizie leków w połączeniu z metodami chemometrycznymi. Współpraca ta zaowocowała opracowaniem nowej metody chemometrycznej (praca A1) - SQPCA, służącej do oznaczania ilościowego wybranego analitu w złożonej mieszaninie o nieznanym składzie na podstawie tylko jednej techniki pomiarowej.

Od 2015 uczestniczę w procesie recenzowania publikacji głównie w dwóch międzynarodowych czasopismach - *Chemometrics and Intelligent Laboratory Systems* (ELSEVIER) oraz *Journal of Chemometrics* (Wiley).

Obecnie moja działalność naukowa koncentruje się na tworzeniu i rozwijaniu nowych algorytmów chemometrycznych oraz ich zastosowaniu w rozwiązywaniu problemów badawczych z obszaru współczesnej chemii, fizyki powierzchni, nauk farmaceutycznych i środowiskowych, a w ostatnim czasie również nauk medycznych i społeczno-ekonomicznych.

3. Kierunek przyszłych badań

Przyszła praca badawcza będzie się koncentrować głównie na rozwijaniu potencjału metody SQPCA dla przypadków silnego nakładania form spektralnych oraz małych ilości substancji <5%. Poza tym jako specjalista w dziedzinie wielowymiarowej analizy statystycznej będę kontynuowała dotychczasowe działania w ramach współpracy z różnymi ośrodkami.

4. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej

- (I) Odbyłam 3 miesięczny staż naukowo-dydaktyczny w Zakładzie Chemii Organicznej, Wydziału Chemicznego Uniwersytetu w Umea w Szwecji w 1996 roku pod kierunkiem prof. Paula Geladiego. Środki pochodziły z programu TEMPUS, otrzymałam grant indywidualny (Individual Mobility Grant-95-PL-2140).
- (II) W czerwcu 2006 roku wzięłam udział w warsztatach szkoleniowych z analizy chemometrycznej danych spektralnych podczas „2nd European User Meeting on Multivariate Data Analysis”, Como, Włochy.
- (III) Jestem współwykonawcą wielu projektów międzynarodowych. W latach 1995-98 byłam członkiem zespołu realizującego projekt TEMPUS-a SJEP 09418-95 “Modernisation

and Curricula Development of Physical Measurement Technology Courses”. Brałam udział w dwóch projektach w ramach współpracy naukowej z ośrodkami naukowymi w Niemczech (5769/R05/07) oraz Austrii (OeAD PL-12/2018 WTZ Polen). W ramach projektu umowy międzyrządowej o współpracy naukowej i naukowo-technicznej między Rządem Polskiej Rzeczypospolitej Ludowej a Rządem Republiki Federalnej Niemiec realizowanej w latach 2004-2005 (Nr 5769/R05/07) odbyłam krótkoterminowy staż (maj 2005) podnoszący kwalifikacje z zakresu opracowywania danych agrofizycznych w Instytucie Analizy Roślin w Quedlinburgu, pod kierownictwem prof. Hartwiga Schulza.

(IV) Współpracuję z naukowcami z ośrodków krajowych i zagranicznych. Część badań naukowych, w tym przedstawionych w pracy **A1**, była prowadzona w ramach współpracy z Collegium Medicum UMK (w sumie powstały 3 współautorskie prace). Praca **A6** została przygotowana w ramach współpracy z dr Beatą Gorczycą z Department of Civil Engineering, University of Manitoba (Kanada), natomiast praca **A7** powstała w ramach współpracy z prof. Frankiem Westadem z Department of Engineering Cybernetics, Norwegian University of Science and Technology.

5. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę

- Umiejętności, które nabyłam na Uniwersytecie w Umea w Szwecji stały się podstawą opracowania materiałów do ćwiczeń projektowych z przedmiotu „*Metody opracowywania wyników pomiarów i obserwacji*” dla nowo powstałego na Uniwersytecie Technologiczno-Przyrodniczym w Bydgoszczy kierunku - fizyka techniczna.

Zajęcia te zyskały uznanie Zespołu Wizytującego Państwowej Komisji Akredytacyjnej, który w raporcie z maja 2005 zawarł następującą opinię „Szczególnie podobały się natomiast zajęcia laboratoryjne z metod opracowywania pomiarów i obserwacji, które prowadzone były w sposób zindywidualizowany, na materiale powiązanym z wynikami badań prowadzonych w Instytucie”.

- W trakcie prowadzenia przez Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy kierunku fizyka techniczna (do 2017 roku) byłam opiekunem naukowym 3 prac magisterskich (2006, 2007, 2008) i 2 prac inżynierskich (2014, 2015) na tym kierunku. Byłam też członkiem Rady Programowej kierunku Fizyka Techniczna oraz pełniłam funkcję opiekuna roku.
- Od początku zatrudnienia prowadzę liczne zajęcia dydaktyczne w formie ćwiczeń i zajęć laboratoryjnych, a od 2003 roku również w formie wykładów m.in. dla kierunków:
 - ✓ fizyka techniczna (wykład z optyki instrumentalnej oraz wykład z metod opracowywania wyników pomiarów i obserwacji),

- ✓ budownictwo, inżynieria środowiska, informatyka, technologia chemiczna, technologia żywności i żywienie człowieka (wykład z fizyki)
- ✓ technologia żywności i żywienie człowieka (zajęcia projektowe - statystyka w doświadczalnictwie)
- ✓ bioinżynieria (zajęcia projektowe – statystyka medyczna)
- W związku z moją specjalizacją w dziedzinie analizy danych byłam członkiem trzyosobowego zespołu prowadzącego szkolenie dla pracowników naukowo-dydaktycznych Wydziału Budownictwa Architektury i Inżynierii Środowiska Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy (01.03.2019) nt. "Właściwa interpretacja rezultatów badań i ich analiza statystyczna". Moim zakresem tematycznym była „Wielowymiarowa analiza statystyczna danych pomiarowych”.
- W ramach doskonalenia zawodowego brałam udział w cyklicznych konferencjach pt. Nauczanie Fizyki w Wyższych Szkołach Technicznych. W 2014 roku konferencja ta była organizowana przez Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy, gdzie pełniłam funkcję sekretarza konferencji.
- Moja aktywność pedagogiczna dotyczy również szkolnictwa średniego, przez kilka lat współorganizowałam zawody finałowe Wojewódzkiego Konkursu Fizycznego dla uczniów szkół ponadgimnazjalnych. Byłam też powoływana do pełnienia funkcji obserwatora podczas egzaminu maturalnego z fizyki. Jestem jednym z głównych autorów kalendarza na 2024 rok, dedykowanego młodzieży szkolnej, pod hasłem „Fizyka jest spoko”. Brałam udział w projektach dydaktycznych dla młodzieży, których celem było przypomnienie zdobytej już wiedzy z fizyki i jej poszerzenie.
- Od początku zatrudnienia w Instytucie Matematyki i Fizyki byłam aktywnym członkiem Polskiego Towarzystwa Fizycznego pełniącym funkcję skarbnika (1994-2023), za co zostałam uhonorowana medalem 100-lecia Polskiego Towarzystwa Fizycznego. W tym okresie prowadziłam obsługę finansową dwóch międzynarodowych konferencji – EURODIM_2018 (XIII Europejska fizyczna konferencja „Defekty w materiałach izolatorów”) oraz LUMDERT_2021 (X edycja Międzynarodowej Konferencji „Detektory luminescencyjne i przetworniki promieniowania jonizującego”), które odbyły się w Bydgoszczy pod patronatem PTF. W 2021 byłam członkiem komitetu organizacyjnego 47 Zjazdu Fizyków Polskich i również byłam odpowiedzialna za obsługę finansową zjazdu.
- Jestem członkiem Rady Dyscypliny: nauki chemiczne na Wydziale Technologii i Inżynierii Chemicznej Politechniki Bydgoskiej.

6. Oprócz kwestii wymienionych w pkt. 1-5, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej

Inne ważne osiągnięcia i nagrody

- Nagroda w postaci certyfikatu recenzenta, przyznana w 2019 roku przez zespół redakcyjny czasopisma *Journal of Chemometrics* (WILEY), podpisana przez redaktora naczelnego Age K. Smilde
- Medal Stulecia Polskiego Towarzystwa Fizycznego przyznany w 2021 roku za pracę na rzecz środowiska fizyków
- Medal Komisji Edukacji Narodowej (2021)
- Wielokrotne nagrody Rektora za osiągnięcia naukowe, dydaktyczne i organizacyjne:
 - stypendium Rektora dla wyróżniających się asystentów (1.07.1999 ÷ 31.03.2000)
 - nagroda zespołowa II stopnia za osiągnięcia w działalności naukowej (1999, 2002, 2004, 2006, 2009, 2014, 2016)
 - nagroda zespołowa II stopnia za osiągnięcia w działalności dydaktycznej (1997, 2004, 2011, 2012, 2016)
 - nagroda zespołowa II stopnia za wyróżniające osiągnięcia w działalności organizacyjnej (2013, 2014, 2017)
 - nagroda indywidualna za wyróżniające osiągnięcia w działalności naukowej (2003, 2017, 2018, 2020, 2021)
 - nagroda indywidualna za wyróżniające osiągnięcia w działalności dydaktycznej (2020)
 - nagroda indywidualna za wyróżniające osiągnięcia w działalności organizacyjnej (2021)
- Nominacja do tytułu Osobowość Roku 2024 w kategorii nauka Kapituły Redakcji Gazety Pomorskiej, Ekspresu Bydgoskiego i Nowości Dziennika Toruńskiego – za wkład w życie naukowe regionu, w szczególności za opracowanie nowego systemu analizy parametrów czystości wody na stacjach monitorujących stan rzeki.

.....
(podpis wnioskodawcy)



Wykaz osiągnięć naukowych albo artystycznych, stanowiących znaczny wkład w rozwój określonej dyscypliny

Dziedzina nauk ścisłych i przyrodniczych

Dyscyplina naukowa: nauki chemiczne

dr Grażyna Czerniak

Politechnika Bydgoska

Wydział Technologii i Inżynierii Chemicznej

SPIS TREŚCI

I.	WYKAZ OSIĄGNIĘĆ NAUKOWYCH ALBO ARTYSTYCZNYCH, O KTÓRYCH MOWA W ART. 219 UST. 1. PKT 2 USTAWY	3
II.	WYKAZ AKTYWNOŚCI NAUKOWEJ ALBO ARTYSTYCZNEJ	4
III.	WSPÓLPRA Z OTOCZENIEM SPOŁECZNYM I GOSPODARCZYM	12
IV.	DANE NAUKOMETRYCZNE	13

I. WYKAZ OSIĄGNIĘĆ NAUKOWYCH ALBO ARTYSTYCZNYCH, O KTÓRYCH MOWA W ART. 219 UST. 1. PKT 2 USTAWY

1. Monografie naukowe, zgodnie z art. 219 ust. 1. pkt 2a ustawy; lub cykl powiązanych tematycznie artykułów naukowych, zgodnie z art. 219 ust. 1. pkt 2b ustawy;

W skład cyklu wchodzi 7 opublikowanych artykułów naukowych, których celem było wskazanie nowych możliwości zastosowań metody głównych składowych w rozwiązywaniu konkretnych problemów z zakresu chemii analitycznej, fizyko-chemii powierzchni oraz nauk środowiskowych:

A1. Balcerowska-Czerniak G., Kupcewicz B.; Score-based quantitative principal component analysis with application to the study of active pharmaceutical ingredients based on attenuated total reflection fourier-transform-infrared spectra. J. Chemom. (WILEY) **2017**, 31, e286.

A2. Balcerowska-Czerniak G.; Chemical information obtained from multicomponent spectra by means of Score-based Quantitative Principal Component Analysis, Chemom. Intell. Lab. Syst. (ELSEVIER) **2021**, 219, 104438.

A3. Balcerowska-Czerniak G., Dittmar-Wituski A., Hiller T., Bukaluk A.; Multivariate analysis to evaluate indium behavior at the copper phthalocyanine thin film. J. Appl. Phys. (API Publishing) **2018**, 124, 075302.

A4. Balcerowska-Czerniak G., Wronkowski A., Antończak A.J., Skowroński Ł., Wronkowska A.A.; The potential of multivariate analysis to phase identification based on X-ray diffraction patterns Chemom. Intell. Lab. Syst. (ELSEVIER) **2014**, 135, 126-132.

A5. Balcerowska-Czerniak G., Trzcinski M., Naparty M., Bukaluk A.; Study of surface oxidation of polycrystalline rhodium using multivariate analysis based on combined optical microscopy and X-ray photoelectron spectroscopy data. App. Surf. Sci. (ELSEVIER) **2020**, 526, 146617.

A6. Balcerowska-Czerniak G., Gorczyca B.; Rapid assessment of surface water quality using statistical multivariate analysis approach: Oder River system case study, Sci. Tot. Environ. (ELSEVIER), **2024**, 912, 168754.

A7. Balcerowska-Czerniak G., Westad F.; A statistical index of surface water quality as a useful extension of monitoring systems for understanding ecosystems' response to adverse events. J. Environ. Manag. (ELSEVIER) **2025**, 389:12610423.

Podsumowanie cyklu powiązanych tematycznie artykułów		
	Zgodnie z rokiem publikacji	Zgodnie z wykazem MNiSW z 2024 roku
IF sumaryczny	33,64	
Liczba punktów MNiSW		910

II. WYKAZ AKTYWNOŚCI NAUKOWEJ ALBO ARTYSTYCZNEJ

1. Wykaz recenzowanych prac naukowych lub artystycznych, w szczególności publikowanych w czasopismach międzynarodowych – nie włączonych do cyklu dokumentującego osiągnięcie

1.1 Przed uzyskaniem stopnia naukowego doktora:

[I.1] Siuda R., **Balcerowska G.**, Chojnacki S., Aberdam D., Rozwadowski M., Adjustment of the set of imperfect data to principal component and factor analyses. The case of Auger spectra, *Chemia Analityczna* 43(6):975-985, 1998.

[I.2] Siuda R., **Balcerowska G.**, Application of principal component and factor analyses in electron spectroscopy, *Electron Tech.*, 31, 3/4, 487-494, 1998.

[I.3] Siuda R., **Balcerowska G.**, Aberdam D., Spurious principal components in the set of spectra subjected to disturbances I. Presentation of the problem, *Chemom. Intelligent Lab. System*, 40, 193, 1998.

[I.4] **Balcerowska G.**, Siuda R., On the application of principal component analysis and factor analysis in electron spectroscopy: the limit of detectability of factors in a set of noisy spectra, *Appl. Surf. Sci.*, 144, 83, 1999.

[I.5] **Balcerowska G.**, Siuda R., Inelastic background subtraction from a set of angle-dependent XPS spectra using PCA and polynomial approximation, *Vacuum* 53, 195, 1999.

[I.6] **Balcerowska G.**, Bukaluk A., Siuda R., Seweryn J., Rozwadowski M., Grain boundary diffusion of Pd through Ag thin layers evaporated on polycrystalline Pd in high vacuum studied by means of AES, PCA and FA, *Vacuum* 53, 93, 1999.

[I.7] **Balcerowska G.**, Siuda R. and Engelhard H., Application of PCA and FA in electron spectroscopy II. The minimum energy shifts of factors that can be resolved in a set of noisy spectra, *Surface Interface Anal.* 29, 492, 2000.

[I.8] **Balcerowska G.**, Cyrankiewicz M., Siuda R., Engelhard H., Buergener M. and Goldman A., Instrumental effects in principal component structure of a set of spectra, *Surf. Interface Anal.* 34, 248, 2002.

1.2 Po uzyskaniu stopnia naukowego doktora:

[II.1] Siuda R., Engelhard H., Goldman A., Cyrankiewicz M., **Balcerowska G.**, PCA-based analysis of X-ray-excited Auger spectra from non-ordered Ag (110), *Czechoslovak Journal of Physics*, 53:63-67, 2003.

[II.2] **Balcerowska G.**, Siuda R., Czarnik-Matusiewicz H., Orthogonal signal correction to PLS modelling in application to spectral data. *Acta Agrophysica* 6(1), 7-18, 2005.

[II.3] Siuda R., **Balcerowska G.**, Sadowski Cz., Comparison of the usability of different spectral ranges within the near ultraviolet, visible and near infrared ranges (UV-VIS-NIR) region for the determination of the content of scab-damaged component in blended samples of ground wheat, *Food Add. Contamin.*, 23(11): 1201-1207, 2006.

- [II.4] Siuda R., **Balcerowska G.**, Kupcewicz B., Lenc L., A modified approach to evaluation of DON content in scab-damaged ground wheat by use of Diffuse Reflectance Spectroscopy, *Food Anal. Methods* 1(4):283-292, 2008.
- [II.5] **Balcerowska G.**, Siuda R., Sadowski C.: Modelowanie zawartości materiału porażonego fuzariozą w rozdrobnionym ziarnie pszenicy na podstawie odbiciowych widm optycznych rejestrowanych w różnych zakresach długości fali. *Acta Agrophysica* 11, s. 323-334, 2008.
- [II.6] **Balcerowska G.**, Siuda R., Skrzypczak B.J., Łukanowski A., Sadowski C., Effect of particle size and spectral sub-range within the UV-VIS-NIR range using diffuse reflectance spectra on multivariate models in evaluating the severity of fusariosis in ground wheat, *Food Add. Contamin.* 26, s. 726-732. 2009.
- [II.7] Wronkowska A.A., **Czerniak G.** Wronkowski A., Kuświk P., Stobiecki F., Ellipsometric characterisation of ARr⁺ -induced changes in optical and electronic properties of Ni₈₀Fe₂₀/Au/Co/Au multilayers. *Acta Physicae Superficierum*, vol XII, s.196-204., 2012.
- [II.8] Wronkowska A.A., **Czerniak G.** Wronkowski A., Skowroński Ł., Optical and microstructural characterization of Au-Sn and Cu-Sn diffusive nanolayers, *Appl. Surf. Sci.*, 281, 30-37, 2013.
- [II.9] Kupcewicz B., Ronowicz J., **Balcerowska-Czerniak G.**, Walasek A., Budzisz E.: Evaluation of impurities in simvastatin drug products with the use of FT-IR spectroscopy and selected chemometric techniques. *Central European Journal of Chemistry*, 11(8), 1320-1329, 2013.
- [II.10] Kupcewicz B., **Balcerowska-Czerniak G.**, Małecka M., Paneth P., Krajewska U., Różalski M., Structure-cytotoxic activity relationship of 3-arylidene flavanone and chromanone (E,Z isomers) and 3-arylflavones. *Bioorganic & Medicinal Chemistry Letters* 23, 4102-4106, 2013.
- [II.11] Kupcewicz B., Ronowicz J., **Balcerowska-Czerniak G.**, Walasek A., Budzisz E., Evaluation of impurities in simvastatin drug products with the use of FT-IR spectroscopy and selected chemometric techniques, *Central Eur. J. Chem.* 11(8): 1320-1329, 2013.
- [II.12] Wronkowska A.A., **Czerniak G.** A, Wronkowski A., Czerwosz E., Kowalska E., NIR-Vis-UV permittivity of nanoporous C-Pd thin films determined using spectroscopic ellipsometry. *Proc. of SPIE 8903, Photonic Applications in Astronomy, Communications Industry, and High-Energy Physics Experiments*, 89030F, 2013.
- [II.13] **Balcerowska-Czerniak G.**, Dittmar-Wituski A., Hiller T., Bukaluk A.; Efficiency of chemometric methodology for characterization of In/CuPc films on HOPG and InSb based on ultraviolet photoelectron spectra, *Inżynieria Materiałowa* 1(5): 27-33, 2016.
- [II.14] Trzcinski M., **Balcerowska-Czerniak G.**, Bukaluk A., XPS studies of the initial oxidation of polycrystalline Rh surface, *Catalysts* 10(6): 617, 2020.
- [II.15] Goryńska-Goldman E., Murawska A., **Balcerowska-Czerniak G.**, Consumer Profiles of Sustainable Fruit and Vegetable Consumption in the European Union, *Sustainability* 15(21):15512, 2023.

2. Wykaz wystąpień na krajowych lub międzynarodowych konferencjach naukowych lub artystycznych, z wyszczególnieniem przedstawionych wykładów na zaproszenie i wykładów plenarnych

2.1 Przed uzyskaniem stopnia naukowego doktora

2.1.1 Zaprezentowane na konferencji międzynarodowej

- **Balcerowska G.**, Siuda R., Buergener M., Engelhard H., Goldmann A.; *Instrumental effects in principal component structure of the set of spectra*, European Conference on Applications of Surface and Interface Analysis (ECASIA), Avignon, France, 2001.

2.1.2. Współautorstwo komunikatów lub posterów na konferencjach krajowych i międzynarodowych:

- Siuda R., **Balcerowska G.**; *Metody analizy głównych składowych i analizy czynnikowej w spektroskopii elektronów*. V Seminarium PTP: Powierzchnia i Struktury Cienkowarstwowe, Ustroń, 1997.
- Siuda R., **Balcerowska G.**, Aberdam D., Baudoing-Savois R., Gauthier Y.; *PCA based analysis of incidence angle dependent Auger spectra from Ni (110) surface*, 18th International Seminar on Surface Physics, Polanica Zdrój, 1997.
- **Balcerowska G.**, Siuda R.; *Inelastic background subtraction from a set of angle-dependent XPS spectra using PCA and polynomial approximation* oraz **Balcerowska G.**, Bukaluk A., Siuda R., Seweryn J., Rozwadowski M.; *Grain boundary diffusion of Pd through Ag thin layers evaporated on polycrystalline Pd in high vacuum studied by means of AES, PCA and FA*, 19th International Seminar on Surface Physics, Polanica Zdrój, 1997.
- **Balcerowska G.**, Siuda R., *On the application of principal component analysis and factor analysis in electron spectroscopy: the limit of detectability of factors in a set of noisy spectra*, 10th International Conference on Quantitative Surface Analysis (QSA-10), Birmingham, UK, 1998.
- **Balcerowska G.**, Siuda R. and Engelhard H.; *On application of PCA and FA in electron spectroscopy II. The minimum energy shifts of factors that can be resolved in a set of noisy spectra*, European Conference on Applications of Surface and Interface Analysis, (ECASIA), Spain, 1999.
- **Balcerowska G.**, Siuda R., *A relationship between basic properties of a set of composed spectra*, 7th International Conference on Chemometrics and Analytical Chemistry CAC, Antwerp, Belgium, 2000.
- **Balcerowska G.**, Rozwadowski M., Siuda R., *Search for nonmetal impurities in nominally pure polycrystalline Ag and Pd*, 20th International Seminar on Surface Physics, Kudowa Zdrój, 2000.
- **Balcerowska G.**, Cyrankiewicz M., Siuda R., Engelhard M., Buergener M., Goldman A.; *Efekty aparaturowe w strukturze głównych składowych zespołu widm*. I Konferencja Chemometria - Metody i Zastosowania, Zakopane, 2001.

2.2 Po uzyskaniu stopnia naukowego doktora

2.2.1. Wygłoszone referaty

- *Porównanie efektywności różnych metod wielowymiarowej regresji liniowej w zastosowaniu do wykrywania porażenia ziarna pszenicy wołkiem zbożowym na podstawie pomiarów w bliskiej podczerwieni*, konferencja naukowa: Procesy fizyczne w kształtowaniu środowiska i jakości surowców żywnościowych, Lublin, 2006, organizator Instytut Agrofizyki PAN w Lublinie.
- *Modelowanie zawartości materiału porażonego fuzariozą w rozdrobnionym ziarnie pszenicy na podstawie odbiciowych widm optycznych rejestrowanych w różnych zakresach długości fal*, konferencja naukowa: Agrofizyka w inżynierii produkcji i ochronie środowiska, Krasiczyn, 2007, organizator Instytut Agrofizyki PAN w Lublinie.
- *Zagadnienie uniwersalnego dla ziarna pszenicy ozimej modelu oceny porażenia fuzariozą na podstawie odbiciowych widm optycznych IV Zjazd Naukowy Polskiego Towarzystwa Agrofizycznego*, Nałęczów, 2008, organizator Instytut Agrofizyki PAN w Lublinie.
- *Dramat Odry – chemometria daje szansę na zapobieganie tragediom rzek*, Polska Konferencja Chemii Analitycznej, Gdańsk, 2025.

2.2.2 Zaprezentowane postery

- **Balcerowska G.**, Siuda R., Czarnik-Matusiewicz H., *Orthogonal Signal correction to PLS modelling in application to spectral data*, The International Workshop: Characterization of granular agro-materials and food powders, Lublin, 2004.
- **Balcerowska G.**, Siuda R., *Comparison of the efficiency of three variable selection methods for use in Multiple Linear Regression (MLR)*, 2nd European User Meeting on Multivariate Analysis, Como, Italy, 2006.
- **Balcerowska G.**, Siuda R., Wiliński W., Łukanowski A., Sadowski Cz., *Efficiency of chosen methods for wavelength selection in improvement of multivariate models for severity of fusariosis in ground wheat based on diffuse reflectance spectra*, 3rd International Symposium on Recent Advances in Food Analysis, Praga, Republika Czeska, 2007.
- **Balcerowska-Czerniak G.**, Kuśtra M., Trzcinski M., Hiller T., Bukaluk A., *Using two-dimensional correlation spectroscopy and principal component analysis to enhance spectral information available from XPS spectra for Ti-O-N system*, 7th International Workshop on Surface Physics. Molecular Nanostructures. Trzebnica, 2015.
- **Balcerowska-Czerniak G.**, Dittmar-Wituski A., Hiller T., Bukaluk A., *Efficiency of chemometrics methodology for characterization of In/CuPc films on HOPG and InSb based on ultraviolet photoelectron spectra*, XXI Physical Metallurgy and Materials Science Conference Advanced Materials and Technologies AMT 2016, Rawa Mazowiecka, 2016.
- **Balcerowska-Czerniak G.**, Kupcewicz B., Trzcinski M., Hiller T., Naparty M.K., Dittmar-Wituski A., Bukaluk A., *Wykorzystanie metody SQPCA w analizie ilościowej*

na podstawie danych spektroskopowych. 60 Zjazd Polskiego Towarzystwa Chemicznego, Wrocław, 2017.

- **Balcerowska-Czerniak G.**, Dittmar-Wituski A., Hiller T., Trzcinski M., Bukaluk A. *Investigation of interface formation of indium on an organic thin films deposited on different substrates using photoelectron spectroscopy combined with chemometrics methodology*. 9th International Workshop on Surface Physics. Nanostructured surfaces. Trzebnica, 2019.
- **Balcerowska-Czerniak G.**, *Statystyczne monitorowanie wielu parametrów dla potrzeb kontroli jakości próbek/procesu na przykładzie danych z katastrofy ekologicznej na Odrze w 2022 roku*. 48 Zjazd Fizyków Polskich, Gdańsk, 2023.

2.2.3. Współautorstwo komunikatów lub posterów w konferencjach krajowych i zagranicznych:

- Siuda R., Balcerowska G., Skrzypczak J., Łukanowski A., Sadowski Cz.; *Effect of particle size and spectral range on performance of multivariate calibration models for severity of fusariosis in ground wheat based on diffuse reflectance spectra*, 3rd International Symposium on Recent Advances in Food Analysis, Praga, Republika Czeska, 2007.
- Siuda R., Balcerowska G., Grabowski A., Hiller T., Sadowski Cz., *Some physical methods for fusariosis detection in wheat grain*, 1st Monika International Conference "Increasing Trust in Rapid Analysis for Food Quality and Safety, Rome, Włochy, 2008.
- Kupcewicz B., Balcerowska G.; *Comparison of the spectrophotometric methods (UV, IR) with multivariate calibration for simultaneous determination of binary and ternary mixtures of analgesic drugs*, IV Konferencja "Chemometria – metody i zastosowania", Zakopane 2008.
- Kupcewicz B., **Balcerowska-Czerniak G.**, Sobiesiak M., Sobiesiak K., Kacprowicz A., *Zastosowanie technik spektrofotometrycznych w zakresie UV-VIS-IR z kalibracją wielowymiarową metodą PLS (Partial Least Squares Regression) w analizie wieloskładnikowych leków przeciwbólowych*. Ogólnopolskie Sympozjum Naukowe: Ocena jakości preparatów farmaceutycznych z uwzględnieniem metod badania w fazie stałej. Gdańsk, 2008.
- Kupcewicz B., Ronowicz J., **Balcerowska-Czerniak G.**, Piekuś N., Budzisz E., *Profilowanie zanieczyszczeń produktów leczniczych zawierających ibuprofen z wykorzystaniem HPLC, FT-IR oraz wielowymiarowych technik eksploracji danych*, Ogólnopolskie Sympozjum Naukowe: Ocena jakości preparatów farmaceutycznych z uwzględnieniem metod badania w fazie stałej. Gdańsk, 2011.
- Wronkowski A., **Czerniak G.**, Skowroński Ł.: *Optical and microstructural properties of Au-Sn and Cu-Sn diffusive layers*. 5th International Workshop on Surface Physics "Surfaces and Nanostructures", Łądek-Zdrój, 2011.
- Wronkowska A.A., **Czerniak G.**, Skowroński Ł., Wronkowski A. *Optical properties of the Cu, β -Sn and In nanoscale films on glass substrates*, 24th General Conference of the Condensed Matter Division of The European Physical society (CMD-24), Edinburgh, UK, 2012.
- **Czerniak G.**, Skowroński Ł., Wronkowski A. and Wronkowska A.A., *Use of multivariate analysis in optimization process of black electroplated Cu-Ni coatings*

preparation, 24th General Conference of the Condensed Matter Division of The European Physical society (CMD-24), Edinburgh, UK, 2012.

- Wronkowska A.A., Wronkowski A., Bodzek H., **Czerniak G.**, Skowroński Ł., Naparty M.K., *IR-Vis-UV optical response of ultra-thin Pd films determined using spectroscopic ellipsometry*, 6th International Workshop on Surface Physics. Functional Materials. Niemcza, 2013.
- **Czerniak G.**, Skowroński Ł., Wronkowski A., Antończak A.J., Ziółkowska D., Wójcik M., Wronkowska A.A. *Multivariate modelling of laser coloured titanium oxides coatings based on their X-ray diffraction patterns*, 6th International Workshop on Surface Physics. Functional Materials. Niemcza, 2013.
- Hiller T., Dittmar-Wituski A., Kustra M., **Czerniak G.**, Trzcinski M., Bukaluk A., *Formation of CuPc and Indium/CuPc films on the HOPG and InSb substrates studied by electron spectroscopies*, 7th International Workshop on Surface Physics. Molecular Nanostructures, Trzebnica, 2015.
- **Balcerowska-Czerniak G.**, Dittmar-Wituski A., Hiller T., Naparty M.K., *Formation of different In layer on CuPc/HOPG and CuPc/InSb investigated by UPS and chemometric methodology*, VI Kongres Polskiego Towarzystwa Próżniowego, Trzebnica, 2016.
- **Balcerowska-Czerniak G.**, Trzcinski M., Hiller T., Naparty M.K., Dittmar-Wituski A., Bukaluk A., Suchorski Y., *Chemometrics approach to the spatially-resolved XPS studies of heterogeneous surfaces*, 8th International Workshop on Surface Physics. Surface is Alive. Trzebnica, 2017.
- Dittmar-Wituski A., **Balcerowska-Czerniak G.**, Hiller T., Trzcinski M., Bukaluk A., *In/Alq3/HOPG studied by UPS and XPS*. 9th International Workshop on Surface Physics. Nanostructured surfaces. Trzebnica, 2019.
- Gazdecki M., Goryńska-Goldmann E. Murawska A., **Balcerowska-Czerniak G.**; *Consumer Profiles for Sustainable Fruit and Vegetable Consumption in the European Union*, OECD-COLEAD “The Fruits and Vegetable Industry, Series – Session 7, 2024.
- Sierakowska K., Husejko J., **Czerniak G.**; *The use of telemedicine solutions in assesment of elderly patients in the perioperative period – results of a pilot study using the mobile application “Drogowskaz”*, EUROMED, Tangier, Maroko, 2025.

2.2.4. Opublikowane materiały konferencyjne

- Siuda R., Cyrankiewicz M., **Balcerowska G.**, Wpływ niedoskonałości aparaturowych na strukturę widm w modelu głównych składowych. w: II Konferencja “Chemometria – metody i zastosowania”, Wydawnictwo Instytutu Ekspertyz Sądowych, Kraków 2003, s. 404-410.
- Chmielewski J., Siuda R., **Balcerowska G.**, Tyrakowska B and Kozioł J., Flavour discrimination by principal component analysis of mass spectra, XLVII Zjazd PTChem i SITPChem, Wrocław, 2004 w: Annals of the Polish Chemical Society, 3, part2, 923-926, 2004.
- Jankowski J., **Balcerowska G.**, Czarnik-Matusiewicz H., Siuda R., Compositional analysis of rapeseeds meal from NIR spectra modeled with sample variants of PCR, Annals of the Polish Chemical Society, 3, part2, 927-930, 2004.

- Siuda R., **Balcerowska G.**, Kupcewicz B., Sadowski Cz., The dependence of the performance of calibration models on the method of reduction of the effect of multiple scattering of radiation in application to determination of deoxynivalenol content in wheat grain damaged with fusariosis, w: III Konferencja “ Chemometria – metody i zastosowania”, Wydawnictwo Instytutu Ekspertyz Sądowych, Kraków 2006, s. 367-373.
- **Balcerowska G.**, Siuda R., Grundas. S., The dependence of the performance of calibration models on the method of reduction of the effect of multiple scattering of radiation in application to detection of infestation of wheat grain with insects, w: III Konferencja “ Chemometria – metody i zastosowania”, Wydawnictwo Instytutu Ekspertyz Sądowych, Kraków 2006, s. 375-380.
- **Balcerowska G.**, Siuda R., Grundas S., The dependence of the performance of calibration models on the method of reduction of the effect of multiple scattering of radiation in application of infestation to detection of infestation of wheat grain with insects, III Konferencja “ Chemometria – metody i zastosowania”, Chemometrics Methods and Applications, s.375-379, 2006.
- Piekus N., Badura A., **Balcerowska-Czerniak G.**, Kupcewicz B. Wykorzystanie HPLC, FT-IR oraz wybranych technik kalibracji danych do analizy zanieczyszczeń produktów leczniczych zawierających ibuprofen. Chemometria metody i zastosowania. V Konferencja Naukowo-Szkoleniowa "Zrozumieć podstawy metod chemometrycznych" Kraków s. 84-85, 2012.
- Kupcewicz B., Różalski M., **Balcerowska-Czerniak G.**, Paneth P.: Ocena zależności cytotoksyczności izomerów E, z 3-arylidenoflwanonów i 3-arylidenochromanonów od struktury z wykorzystaniem wybranych metod chemometrycznych., Chemometria w rozwiązywaniu problemów nauki i praktyki, Wydawnictwo Instytutu Ekspertyz Sądowych, Kraków, s. 253-259, 2014.

3. Wykaz referatów wygłoszonych w instytucjach naukowych na zaproszenie

- Uniwersytet Przyrodniczy w Poznaniu, Katedra Chemii, 2006, temat wystąpienia: *Wybrane metody interpretacji danych wielowymiarowych.*
- Uniwersytet Mikołaja Kopernika, Instytut Fizyki, Zakład Fizyki Atomowej, Molekularnej i Optycznej, 2008, temat wystąpienia: *Metody fizyczne w wykrywaniu fuzariozy ziarna pszenicy.*
- Bydgoski Oddział Polskiego Towarzystwa Chemicznego, Bydgoszcz, 2021, temat wystąpienia: *Techniki chemometryczne w analizie jakościowej i ilościowej.*
- Wydział Technologii i Inżynierii Chemicznej Politechniki Bydgoskiej, 2022, wykład w ramach cyklu „nie Taka Chemia Straszna”, temat wystąpienia: *Chemometria – statystyka nie tylko dla chemika.*
- Instytut Informatyki Teoretycznej i Stosowanej Polskiej Akademii Nauk (IITiS PAN), Gliwice, 2023, temat wystąpienia: *Metoda głównych składowych w analizie danych wielowymiarowych: podejście niestandardowe.*
- Politechnika Śląska, Instytut Fizyki, Gliwice, 2023, temat wystąpienia: *Chemometryczna analiza katastrofy ekologicznej na Odrze w 2022 roku.*

- Kujawsko-Pomorskie Centrum Edukacji Nauczycieli we Włocławku, Konferencja z okazji światowego dnia wody „Zatrzymać wodę”, 2024, temat wystąpienia: *Dramat Odry: statystyka daje szansę na zapobieganie tragediom rzek.*
 - Collegium Medicum UMK, Katedra Biofizyki, 2024, temat wystąpienia: *Metody chemometryczne w analizie danych.*
 - Instytut Ochrony Środowiska -Państwowy Instytut Badawczy, marzec 2025, temat wystąpienia: *Dramat Odry – statystyka daje szansę na zapobieganie tragediom rzek.*
 - Spotkanie Grupy Roboczej przedstawicieli Ministerstwa Klimatu i Środowiska, maj 2025, temat wystąpienia: *Dramat Odry – statystyka daje szansę na zapobieganie tragediom rzek.*
 - Instytut Chemii Fizycznej PAN, Warszawa, 2025, temat wystąpienia: *Metoda głównych składowych w rozwiązywaniu problemów analizy ilościowej próbek o złożonej strukturze i składzie chemicznym: podejście niestandardowe.*
4. Wykaz udziału w komitetach organizacyjnych i naukowych konferencji krajowych lub międzynarodowych, z podaniem pełnionej funkcji:
- Nauczanie Fizyki w Wyższych Szkołach Technicznych, 2014, organizator: Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy, pełniona funkcja – sekretarz konferencji.
 - EURODIM_2018-XIII Europhysical Conference on Defects in Insulating Materials, organizator- Uniwersytet Kazimierza Wielkiego oraz Polskie Towarzystwo Fizyczne, pełniona funkcja – skarbnik odpowiedzialny za rozliczenia finansowe konferencji.
 - LUMDETR_2021-11th International Conference on Luminescent Detectors and Transformers of Ionizing Radiation, organizator- Uniwersytet Kazimierza Wielkiego oraz Polskie Towarzystwo Fizyczne, pełniona funkcja – skarbnik odpowiedzialny za rozliczenia finansowe konferencji.
 - 47 Zjazdu Fizyków Polskich, 2021, organizator – Polskie Towarzystwo Fizyczne, Politechnika Bydgoska, Uniwersytet Kazimierza Wielkiego, Collegium Medicum UMK, pełniona funkcja – skarbnik odpowiedzialny za rozliczenia finansowe konferencji.
5. Wykaz uczestnictwa w pracach zespołów badawczych realizujących projekty finansowane w drodze konkursów krajowych lub zagranicznych, z podziałem na projekty zrealizowane i będące w toku realizacji, oraz z uwzględnieniem informacji o pełnionej funkcji w ramach prac zespołów.

Projekty zrealizowane:

- TEMPUS-a SJEP 09418-95 “Modernisation and Curricula Development of Physical Measurement Technology Courses” – wykonawca.
- 5769/R05/07 „Rozwój metod spektroskopii optycznej służącej szybkiej detekcji wybranych szkodliwych substancji w roślinach i nasionach” – projekt międzynarodowy realizowany we współpracy z Institut für Pflanzenanalytik, Quedlinburg, Niemcy, wykonawca.

- OeAD PL-12/2018 WTZ Polen PL_2018-20 “Role of the metal-oxide interaction in the catalytic hydrogen oxidation”, - projekt międzynarodowy realizowany w ramach polsko-austriackiego konkursu wymiany naukowej z Technische Universität Wien, wykonawca, liczba punktów 20/M =5.
- KBN PB1337/P03/96/10 „Analiza zjawisk fizycznych towarzyszących procesowi dyfuzji w wielowarstwowych strukturach metali, wykonawca.
- 2P03B 135 10 „Analiza zjawisk towarzyszących procesowi dyfuzji w wielowarstwowych strukturach metali” – wykonawca.
- 2P06R 113 26 „Opracowanie metody oznaczania stężeń mikotoksyn fuzaryjnych w ziarnie pszenicy za pomocą spektroskopii optycznej” – wykonawca.

6. Wykaz członkostwa w międzynarodowych lub krajowych organizacjach i towarzystwach naukowych wraz z informacją o pełnionych funkcjach.

- członek Polskiego Towarzystwa Fizycznego – skarbnik oddziału bydgoskiego (1994-2023).

7. Wykaz staży w instytucjach naukowych lub artystycznych, w tym zagranicznych, z podaniem miejsca, terminu, czasu trwania stażu i jego charakteru.

- Otrzymałam 3 miesięczny staż naukowo-dydaktyczny w okresie 06.03-01.06.1996 w Zakładzie Chemii Organicznej, Wydziału Chemicznego Uniwersytetu w Umea w Szwecji pod kierunkiem prof. Paula Geladiego. Środki pochodziły z programu TEMPUS, otrzymałam grant indywidualny (Individual Mobility Grant-95-PL-2140).
- Otrzymałam krótkoterminowy staż (22-27 maj 2005) podnoszący kwalifikacje z zakresu opracowywania danych agrofizycznych w Instytucie Analizy Roślin w Quedlinburgu (Niemcy) pod kierownictwem prof. Hartwiga Schulza.

8. Wykaz udziału w zespołach badawczych, realizujących projekty inne niż określone w pkt. II.4.

Projekty zrealizowane:

- BS-70/1991 „Badanie właściwości powierzchni polikrystalicznych metali i cienkich warstw otrzymywanych metodami próżniowymi i elektrochemicznymi”, wykonawca.
- BS-21/2001: „Współczesne metody interpretacji danych spektroskopowych i ich zastosowania”, wykonawca.
- BS-16/2005 „Metody interpretacji widm i budowy modeli kalibracji wielowymiarowej”, wykonawca.
- Wronkowska A.A., Wronkowski A., **Czerniak G.**: Wykonanie analizy właściwości optycznych nanowarstw węglowo-metalicznych metodą spektroelipsometrii – raport_BZ09IMiF. 2011.

III. WSPÓŁPRA Z OTOCZENIEM SPOŁECZNYM I GOSPODARCZYM

W związku z opracowaniem nowego indeksu statystycznego, dotyczącego jakości wód nawiązałam kontakty oraz odbyłam spotkania w celu prezentacji indeksu TQ_PCA z następującymi instytucjami w kraju:

- Instytut Ochrony Środowiska – Państwowy Instytut Badawczy w Warszawie
- Eksperci dla Odry – facebookowa grupa ekspertów i naukowców
- Grupa robocza przedstawicieli Ministerstwa Klimatu i Środowiska
(<https://www.facebook.com/SZTABdlaODRYpl/>, marek.kajs@gdos.gov.pl;
Louis.COURSEAU@klimat.gov.pl)
- Firma Waterly - producent automatycznych boi do monitoringu jakości wód powierzchniowych (aktualnie 9 takich urządzeń jest zainstalowanych na terenie naszego kraju, a wyniki można śledzić w czasie rzeczywistym na stronie <https://waterly.cloud/>).

IV. DANE NAUKOMETRYCZNE

1. ORCID <https://orcid.org/0000-0002-8180-2536>
2. Web of Science ResearcherID: AAM-5429-2021
3. Indeks Hirscha : 8 (Web of Science)
4. Liczba publikacji: 22 (Web of Science)
5. Lista cytowań publikacji: 166 (Web of Science, Sum of Times Cited)
6. Lista cytowań publikacji bez autocytowań: 153 (Web of Science, Sum of Times Cited)

.....

(podpis wnioskodawcy)