

AUTOREFERAT

Małgorzata Charytanowicz

Instytut Badań Systemowych PAN

Warszawa 2015

SPIS TREŚCI

1. Dane osobowe.....	1
2. Posiadane dyplomy i stopnie naukowe.....	1
3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych.....	2
4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki.....	2
4.1. Tytuł osiągnięcia naukowego	2
4.2. Publikacje wchodzące w skład osiągnięcia naukowego.....	2
4.3. Omówienie celu naukowego przedłożonego osiągnięcia naukowego i otrzymanych wyników oraz ich możliwości aplikacyjnych.....	4
5. Omówienie pozostałych osiągnięć naukowo-badawczych.....	14

Warszawa, 10 sierpnia 2015 r.

dr Małgorzata Charytanowicz
Instytut Badań Systemowych
Polska Akademia Nauk

AUTOREFERAT

**przedstawiający opis dorobku i osiągnięć naukowych
w związku z ubieganiem się o nadanie stopnia doktora habilitowanego**

1. Dane osobowe

Imię i nazwisko: **Małgorzata Charytanowicz**

Adres służbowy: Instytut Badań Systemowych
Polska Akademia Nauk
ul. Newelska 6
01-447 Warszawa
e-mail: malgorzata.charytanowicz@ibspan.waw.pl

Instytut Matematyki i Informatyki
Wydział Matematyki, Informatyki i Architektury Krajobrazu
Katolicki Uniwersytet Lubelski Jana Pawła II
ul. Konstantynów 1H
20-708 Lublin

2. Posiadane dyplomy i stopnie naukowe

2005 Doktor nauk technicznych w zakresie informatyki, z wyróżnieniem
Instytut Badań Systemowych Polskiej Akademii Nauk w Warszawie
Tytuł rozprawy doktorskiej:
*Bayesowska determinizacja informacji nieprecyzyjnej w zagadnieniach
medycznych*
Promotor: prof. dr hab. inż. Piotr Kulczycki.

1989 Magister matematyki
specjalność: metody numeryczne i programowanie
Wydział Matematyki i Fizyki
Uniwersytet Marii Curie-Skłodowskiej w Lublinie
Tytuł pracy magisterskiej:
Numeryczne obliczanie dwuwymiarowej całki osobliwej Cauchy'ego.

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych

Okres	Miejsce zatrudnienia
2005-	Instytut Badań Systemowych Polska Akademia Nauk, Warszawa.
2005-	Katedra Analizy Numerycznej i Programowania Instytut Matematyki i Informatyki Wydział Matematyki, Informatyki i Architektury Krajobrazu Katolicki Uniwersytet Lubelski Jana Pawła II.
1990-2005	Międzywydziałowy Zakład Matematyki i Biostatystyki Medycznej Akademia Medyczna w Lublinie.

4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki (Dz. U. 2003 nr 65 poz. 595 ze zm.)

4.1. Tytuł osiągnięcia naukowego

Komputerowa analiza danych w zastosowaniach technicznych i medycznych.

4.2. Publikacje wchodzące w skład osiągnięcia naukowego

Sumaryczny IF ośmiu publikacji przedstawionych w ramach cyklu według bazy JCR wynosi **13,965**, a sumaryczny 5-Year IF – **14,164**. Sumaryczna liczba punktów według MNiSW, zgodnie z obowiązującym w roku 2015 wykazem czasopism naukowych, jest równa **215**.

Przedstawiony do oceny cykl publikacji powiązanych tematycznie obejmuje osiem artykułów:

H1. P. Kulczycki (50%), **M. Charytanowicz (50%)**, An Algorithm for Conditional Multidimensional Parameter Identification with Asymmetric and Correlated Losses of Under- and Overestimations. *Journal of Statistical Computation and Simulation*, przyjęte do druku, DOI: 10.1080/00949655.2015.1049606; artykuł dostępny na stronie czasopisma (10 sierpnia 2015): <http://www.tandfonline.com/doi/full/10.1080/00949655.2015.1049606>.

JCR, IF/5-Year IF: 0,635/0,760; MNiSW: Lista A, 20 pkt.

Mój wkład w powstanie tej publikacji szacuję na **50%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 40%, wybór metodyki badań 50%, projekt i implementację algorytmu 70%, analizę i interpretację wyników 60%, przygotowanie publikacji do druku 40%.

H2. H. Czachor (27%), **M. Charytanowicz (27%)**, S. Gonet (16%), J. Niewczas (10%), G. Józefaciuk (10%), L. Lichner (10%), Impact of long term mineral and organic fertilization on water stability, wettability and porosity of aggregates of two silt loamy soils. *European Journal of Soil Science*, vol. 66, nr 3, ss. 577-588, 2015.

JCR, IF/5-Year IF: 2,649/3,487; MNiSW: Lista A, 35 pkt.

Mój wkład w powstanie tej publikacji szacuję na **27%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 20%, wybór metodyki badań 25%, projekt i implementację algorytmu 40%, analizę i interpretację wyników 25%, przygotowanie publikacji do druku 25%.

- H3. P. Kulczycki (33%), **M. Charytanowicz (33%)**, A.L. Dawidowicz (33%), A Convenient Ready-to-Use Algorithm for a Conditional Quantile Estimator. *Applied Mathematics & Information Sciences*, vol. 9, nr 2, ss. 841-850, 2015.

JCR, IF/5-Year IF: 1,232/1,204; MNiSW: Lista A, 30 pkt.

Mój wkład w powstanie tej publikacji szacuję na **33%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 30%, wybór metodyki badań 30%, projekt i implementację algorytmu 50%, analizę i interpretację wyników 40%, przygotowanie publikacji do druku 30%.

- H4. M. Klatka (22%), E. Grywalska (18%), M. Partyka (18%), **M. Charytanowicz (22%)**, E. Kiszczak-Bochyńska (10%), J. Roliński (10%), Th17 and Treg cells in adolescents with Graves' disease. Impact of treatment with methimazole on these cell subsets. *Autoimmunity*, vol. 47, nr 3, ss. 201-211, 2014.

JCR, IF/5-Year IF: 2,714/2,163; MNiSW: Lista A, 20 pkt.

Mój wkład w powstanie tej publikacji szacuję na **22%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 15%, wybór metodyki badań 15%, projekt i implementację algorytmu 50%, analizę i interpretację wyników 20%, przygotowanie publikacji do druku 10%.

- H5. P. Kulczycki (50%), **M. Charytanowicz (50%)**, Conditional Parameter Identification with Different Losses of Under- and Overestimation. *Applied Mathematical Modelling*, vol. 37, nr 4, ss. 2166-2177, 2013.

JCR, IF/5-Year IF: 2,251/2,326; MNiSW: Lista A, 35 pkt.

Mój wkład w powstanie tej publikacji szacuję na **50%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 40%, wybór metodyki badań 50%, projekt i implementację algorytmu 70%, analizę i interpretację wyników 60%, przygotowanie publikacji do druku 40%.

- H6. J. Klatka (21%), M. Remer (14%), R. Dobrowolski (12%), W. Pietruszewska (12%), A. Trojanowska (10%), H. Siwiec (10%), **M. Charytanowicz (21%)**, The content of cadmium, cobalt and nickel in laryngeal carcinoma. *Archives of Medical Science*, vol. 7, nr 3, ss. 517-522, 2011.

JCR, IF/5-Year IF: 2,030/1,656; MNiSW: Lista A, 25 pkt.

Mój wkład w powstanie tej publikacji szacuję na **21%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 10%, wybór metodyki badań 15%, projekt i implementację algorytmu 50%, analizę i interpretację wyników 20%, przygotowanie publikacji do druku 10%.

- H7. P. Kulczycki (50%), **M. Charytanowicz (50%)**, A Complete Gradient Clustering Algorithm Formed with Kernel Estimators. *International Journal of Applied Mathematics and Computer Science*, vol. 20, nr 1, ss. 123-134, 2010.

JCR, IF/5-Year IF: 1,227/1,284; MNiSW: Lista A, 25 pkt.

Mój wkład w powstanie tej publikacji szacuję na **50%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 40%, wybór metodyki badań 50%, projekt i implementację algorytmu 70%, analizę i interpretację wyników 60%, przygotowanie publikacji

do druku 40%.

H8. P. Kulczycki (50%), **M. Charytanowicz (50%)**, Bayes sharpening of imprecise information. *International Journal of Applied Mathematics and Computer Science*, vol. 15, nr 3, ss. 393-404, 2005.

JCR, IF/5-Year IF: 1,227/1,284; MNiSW: Lista A, 25 pkt.

Mój wkład w powstanie tej publikacji szacuję na **50%**. Obejmował on wszystkie etapy przygotowania publikacji, w szczególności: opracowanie koncepcji i celu pracy 40%, wybór metodyki badań 50%, projekt i implementację algorytmu 70%, analizę i interpretację wyników 60%, przygotowanie publikacji do druku 40%.

W nawiasach podano procentowy wkład poszczególnych autorów w powstanie publikacji. W *Załącznikach 6, 7* przedstawiono oświadczenia współautorów ośmiu prac wchodzących w skład osiągnięcia naukowego o udziale w poszczególnych etapach przygotowania publikacji.

IF oraz pięcioletni IF (5-Year IF) został podany według bazy JCR zgodnie z dostępem z dnia 10 sierpnia 2015. Liczbę punktów przyznawaną za publikację wyszczególniono zgodnie z obowiązującym w roku 2015 wykazem czasopism naukowych ogłoszonym przez Ministra Nauki i Szkolnictwa Wyższego z dnia 31 grudnia 2014 r.

4.3. Omówienie celu naukowego przedłożonego osiągnięcia naukowego i otrzymanych wyników oraz ich możliwości aplikacyjnych

Niniejszy rozdział stanowi omówienie celu naukowego i wyników publikacji wchodzących w skład osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki (*Dz. U. 2003 nr 65 poz. 595 ze zm.*).

Ogólne przedstawienie tematyki i głównego celu prac

Przedmiotem prowadzonych badań, których wyniki zaprezentowane zostały w postaci przedstawionego do oceny cyklu publikacji powiązanych tematycznie, jest zaprojektowanie wydajnych algorytmów rozwiązujących trudne obliczeniowo problemy występujące w naukach technicznych i medycznych.

Zagadnienia badawcze rozważane przeze mnie po uzyskaniu stopnia naukowego doktora dotyczyły nowoczesnych metod komputerowej analizy i eksploracji danych. Wzrost mocy obliczeniowej i pojemności pamięci komputerów oraz algorytmizacja i globalizacja procedur przetwarzania informacji, otworzyły przed nauką szerokie możliwości, stawiając jednocześnie nowe wyzwania. Z jednej strony nastąpił wzrost ilości posiadanej informacji, możliwość jej szybkiego rejestrowania, przekazywania i przetwarzania, z drugiej – korzystanie z owych narzędzi wymaga szczegółowego zrozumienia przedmiotu badań i opracowania procedur pozwalających na analizę zebranego materiału. Ich właściwy wybór, dokonany na początku badań, skutkuje gromadzeniem danych w sposób celowy i efektywny i jest obecnie kluczowym elementem procesu badawczego.

Problem matematycznego modelowania posiadanej informacji jest jednym z najistotniejszych zagadnień występujących we współczesnej nauce. Metody klasyczne wykorzystują informację określoną jednoznacznie, zwykle dostarczaną przez urządzenia pomiarowe. Tymczasem człowiek efektywnie funkcjonuje w sytuacjach wykraczających poza możliwości takich modeli i stosuje nieściśle oceny wielkości fizycznych i stanów obiektów, posługując się określeniami języka naturalnego. We współczesnych zagadnieniach

aplikacyjnych koncepcje uwzględniające niepewność i nieprecyzyjność informacji stają się coraz bardziej konkurencyjnym – wobec metod klasycznych – sposobem opisu nieokreśloności natury, wymagając jednocześnie nowatorskich narzędzi i procedur. Koncepcje te, początkowo wykorzystywane głównie do wyrafinowanych zagadnień inżynierii, coraz częściej znajdują zastosowanie w naukach medycznych [Gerstenkorn et al., 1990].

W badaniach naukowych kładłam nacisk na opracowanie kompletnych procedur komputerowej analizy danych i ich praktyczne zastosowanie w naukach technicznych i medycznych. Prace prowadzone w Instytucie Badań Systemowych PAN w Warszawie oraz Instytucie Matematyki i Informatyki KUL w Lublinie wykazały duży potencjał możliwości aplikacyjnych nieparametrycznej metodyki estymatorów jądrowych w zagadnieniach wyostrażania informacji nieprecyzyjnej oraz identyfikacji parametrów, z uwzględnieniem różnych skutków przeszacowania i niedoszacowania ich wartości oraz dowolności rozkładu zmiennych opisujących i warunkujących. Umożliwiły również opracowanie nowoczesnych metod wielowymiarowej analizy danych oraz przetwarzania obrazów cyfrowych. W efekcie zaprojektowano i zaimplementowano wydajne obliczeniowo algorytmy służące do rozwiązywania rzeczywistych problemów występujących w naukach technicznych i medycznych. Ich wykorzystanie poprzedzone zostało szczegółową analizą formalną i numeryczną, wraz z podaniem wszystkich potrzebnych wzorów i formuł obliczeniowych.

Za oryginalny dorobek naukowy prowadzonych przeze mnie badań uważam:

- praktyczne zastosowanie procedury bayesowskiego wyostrażania informacji nieprecyzyjnej dla ustalonych czynników warunkujących różnego typu;
- zaprojektowanie i implementację nieparametrycznych algorytmów warunkowej identyfikacji parametrów z niesymetrycznymi stratami i przedstawienie możliwości ich praktycznego zastosowania w naukach technicznych;
- opracowanie i implementację kompletnej nieparametrycznej procedury wyznaczania warunkowego estymatora kwantyla, opartej na nieparametrycznej metodzie estymatorów jądrowych;
- zaprojektowanie i implementację kompletnego gradientowego algorytmu klasteryzacji, umożliwiającego, za pomocą wyodrębnionych parametrów, wpływ na liczbę klastrów oraz proporcję między ich ilościami w obszarach zagęszczenia elementów próby wobec rejonów ich rzadkiego występowania;
- zaprojektowanie i implementację algorytmu wyznaczania rozkładu wielkości porów, bazującego na analizie obrazów cyfrowych oraz opracowanie nowych metod oceny trwałości agregatów glebowych;
- opracowanie i praktyczne wykorzystanie algorytmów wielowymiarowej analizy i eksploracji danych do modelowania złożonych systemów biologicznych, trudnych do opisanego za pomocą standardowych procedur.

Szczegółowy opis rozważanych problemów badawczych i uzyskanych wyników

Moje osiągnięcie naukowe stanowi cykl publikacji powiązanych tematycznie pt. „*Komputerowa analiza danych w zastosowaniach technicznych i medycznych*”.

W ostatnich latach nastąpił znaczny wzrost badań w zakresie metod przetwarzania informacji nieprecyzyjnej oraz ich zastosowań do różnorodnych dziedzin nauki i działań

praktycznych. Informacja nieprecyzyjna, zawarta w określeniach języka naturalnego, była wykorzystywana od dawna, stanowiąc często jedyną dostępną informację. Nie była jednak stosowana w metodach opartych na klasycznej matematyce, co prowadziło do utraty lub wypaczenia zawartej w niej wiedzy. Użycie modeli dopuszczających znacznie ogólniejszą klasę informacji wymagało opracowania nowatorskich narzędzi reprezentacji danych i ich oceny jakościowej. Celem pracy [poz. H8 osiągnięcia naukowego] pt. „*Bayes sharpening of imprecise information*” było opracowanie kompletnej procedury wyznaczania na podstawie nieprecyzyjnych ocen konkretnej wartości, możliwie dobrze charakteryzującej rozważany aspekt. Dostępne w literaturze metody nie uwzględniają różnych – co do znaku i wielkości – skutków przeszacowania i niedoszacowania wyniku. Przedstawiona w niniejszej pracy metoda, oparta na bayesowskiej regule decyzyjnej, pozwala wziąć pod uwagę ten aspekt, a uzyskany wynik minimalizuje potencjalne straty. Ponadto zaproponowany algorytm umożliwia uwzględnienie wpływu różnorodnych czynników warunkujących różnego typu. Do estymacji rozkładu miary charakteryzującej informację nieprecyzyjną zastosowano teorię statystycznych estymatorów jądrowych [Kulczycki, 2005; Wand & Jones, 1995], co uniezależnia metodę od typu rozkładów informacji nieprecyzyjnej i czynników warunkujących. W konsekwencji opracowany został kompletny algorytm obliczeniowy, pozwalający na bayesowską defuzyfikację informacji nieprecyzyjnej przy ustalonych wartościach czynników warunkujących. Ze względu na uniwersalne możliwości aplikacyjne estymatorów jądrowych wyznaczono i podano postaci wszystkich użytych wzorów i formuł w ogólnym, n -wymiarowym przypadku oraz przedstawiono kompletne procedury, umożliwiające uzyskanie ostatecznych wyników w formie niewymagającej dogłębnej znajomości aspektów teoretycznych przez użytkownika. Poprawność algorytmu została zweryfikowana za pomocą numerycznego programu symulacyjnego. Przeprowadzone badania dotyczyły w szczególności:

- intensywności procedury modyfikacji parametru wygładzania estymatora jądrowego; w rezultacie została zaproponowana formuła uwzględniająca zmienną intensywność parametru wygładzania i polepszająca jakość wyników;
- występowania różnego typu rozkładów zmiennej losowej oraz rodzaju i wartości czynników warunkujących.

Wyniki symulacji numerycznej potwierdziły poprawność koncepcji opracowanej w niniejszej pracy procedury. Zastosowanie nieparametrycznej metodyki statystycznych estymatorów jądrowych pozwoliło uzyskać szereg pozytywnych cech przedstawionego algorytmu, między innymi:

- uniezależnienie metody od występującego rozkładu, także w zakresie wielomodalności i niesymetrii;
- możliwość uwzględnienia czynników warunkujących różnego typu;
- uwzględnienie wielowymiarowości zmiennych warunkujących;
- możliwość wnioskowania przy braku danych z otoczenia ustalonej wartości zmiennej warunkującej;
- możliwość wprowadzenia dodatkowych uwarunkowań, np. ograniczenia nośnika.

W każdym z rozpatrywanych przypadków, w miarę zwiększania liczności próby, uzyskiwane wartości determinizatora zmierzały do wartości teoretycznej, a odchylenie standardowe do zera. Powyższe własności mają fundamentalne znaczenie z aplikacyjnego punktu widzenia. Stanowią bowiem, iż możliwe jest uzyskanie dowolnej dokładności determinizacji, aczkolwiek wymaga to zapewnienia odpowiedniej liczności próby losowej. W praktyce oznacza to konieczność ustalenia właściwego kompromisu między tymi wielkościami. Nawet w przypadkach skrajnych, gdy wartość zmiennej warunkującej

znajdowała się w otoczeniu drugiego jej odchylenia standardowego, wyniki były poprawne, chociaż stanowią one zadania bardzo trudno uwarunkowane i w przypadku niektórych metod klasycznych są wręcz niemożliwe do zrealizowania.

Prowadzone badania wykazały duży potencjał aplikacyjny teorii statystycznych estymatorów jądrowych w zagadnieniach komputerowej analizy danych, w szczególności w zakresie analizy skupień, identyfikacji elementów nietypowych i klasyfikacji. Praca [poz. H7 osiągnięcia naukowego] pt. „*A Complete Gradient Clustering Algorithm Formed with Kernel Estimators*” prezentuje gradientowy algorytm klasteryzacji w kompletnej postaci, dogodnej do bezpośredniego zastosowania bez konieczności prowadzenia żmudnych badań przedmiotowych ze strony użytkownika. Podstawą przedstawionej w artykule koncepcji jest uznanie zbioru danych za próbę losową pozyskaną z pewnej, n -wymiarowej zmiennej losowej, następnie wyznaczenie estymatora jądrowego gęstości jej rozkładu i przyjęcie naturalnego założenia, iż poszczególne klastry odpowiadają lokalnym maksimum uzyskanego estymatora. Istotną cechą opracowanego algorytmu jest brak wymagań dotyczących ściślego ustalenia liczby klastrów, a jedynie wskazanie jej rzędu wielkości, co pozwala dopasować ich liczbę do rzeczywistej struktury danych. Powyższe jest realizowane poprzez wyodrębnienie parametru odpowiadającego za ilość klastrów, po czym wskazanie jego standardowej wartości za pomocą kryteriów optymalizacyjnych, a następnie skutków jej ewentualnej modyfikacji implikującej zmniejszenie lub zwiększenie rzędu ilości klastrów, aczkolwiek nadal bez wskazania na ich konkretną liczbę. Możliwe jest także wyodrębnienie kolejnego parametru, którego wartość wpływa na proporcję między liczbą klastrów w obszarach zagęszczenia elementów zbioru danych oraz w obszarach, gdzie są one rzadkie. Także tu określona została jego wartość oparta na kryterium optymalizacyjnym, ewentualnie poddawana modyfikacjom w celu zwiększenia ich ilości w regionach zagęszczenia danych kosztem redukcji w obszarach rzadszego ich występowania lub odwrotnie. W szczególności odpowiednia relacja między oboma powyższymi parametrami pozwala na zmniejszenie liczby lub nawet wyeliminowanie klastrów w obszarach rzadszego występowania elementów zbioru danych, co w konsekwencji umożliwia usunięcie wpływu elementów odosobnionych [Aggarwal, 2013], bez ingerencji w liczbę klastrów w obszarach ich zagęszczenia. W pracy sformułowano i szczegółowo przeanalizowano następujące aspekty algorytmu:

- formułę konstrukcji estymatora jądrowego;
- ustalenie warunku stopu i w konsekwencji liczby kroków algorytmu;
- określenie procedury tworzenia klastrów i zaliczania do nich poszczególnych elementów zbioru po wykonaniu ostatniego kroku;
- szczegółową analizę wpływu wartości poszczególnych parametrów na liczbę klastrów i na proporcję pomiędzy ich ilością w regionach zagęszczenia elementów zbioru danych oraz w obszarach, gdzie są one rzadkie.

Wykorzystując bliską intuicji interpretację samej koncepcji algorytmu podano ilustracyjną analizę znaczenia poszczególnych parametrów i ich ewentualnej modyfikacji oraz dokonano weryfikacji na rzeczywistych danych przyrodniczych [Lubischew, 1962]. Zaproponowano również metodę ulepszenia samego estymatora jądrowego funkcji gęstości poprzez klasteryzację próby losowej i zastosowanie odrębnej macierzy transformacji w postaci pełnej wobec każdej z wyróżnionych grup.

Dalsze badania pozwoliły na opracowanie nieparametrycznej procedury warunkowej identyfikacji wartości parametrów, dla których założony model najlepiej przybliża rzeczywisty system [Walter & Pronzato, 1997]. Zagadnienie tego typu, pomimo swego tradycyjnego charakteru, ma nadal bardzo duże znaczenie w nauce i technice. W miarę

wzrostu złożoności systemu zmniejsza się możliwość zastosowania metod klasycznej identyfikacji parametrycznej na rzecz metod nieparametrycznych. Szczęśliwie, gwałtowny rozwój współczesnych systemów komputerowych, wspomagany od strony teoretycznej metodami technik informacyjnych umożliwił rozwój nowoczesnych procedur nieparametrycznych. Wyniki badań znalazły swoje odzwierciedlenie w pracy [poz. H5 osiągnięcia naukowego] pt. „*Conditional Parameter Identification with Different Losses of Under- and Overestimation*”. Istotę opracowanego algorytmu stanowią następujące aspekty:

- minimalizacja wartości oczekiwanej strat wynikłych z nieuniknionych w praktyce błędów estymacji;
- niesymetria tych strat, czyli uwzględnienie sytuacji, gdy straty wynikłe z niedoszacowania w sposób istotny różnią się od strat implikowanych przez przeszacowanie;
- dowolność występujących w zagadnieniu rozkładów probabilistycznych;
- warunkowość identyfikowanego parametru, czyli jego istotna zależność od czynników, których wartości mogą być w praktycznych zastosowaniach mierzone w trakcie trwania procesu.

W celu rozwiązania tak sformułowanego zadania, użyto bayesowskiej reguły decyzyjnej. Drugi aspekt został uwzględniony poprzez przyjęcie funkcji strat wynikłych z błędów estymacji w niesymetrycznej postaci wielomianowej. Ograniczenie funkcji strat do postaci wielomianowej nie zmniejsza w praktycznych zagadnieniach ogólności rozważań, stanowiąc efektywny kompromis pomiędzy precyzją uzyskiwanych wyników i złożonością procedur obliczeniowych. Z kolei możliwość zmiany – względem wartości otrzymanej z badań fundamentalnych – stopnia wielomianu umożliwia zróżnicowanie stopnia „zabezpieczenia” przed dużymi błędami estymacji. Trzeci aspekt został zrealizowany poprzez zastosowanie do scharakteryzowania rozkładów probabilistycznych nieparametrycznej metody statystycznych estymatorów jądrowych [Kulczycki, 2005; Wand & Jones, 1995]. Odpowiedni dobór jądra estymatora jądrowego umożliwił podanie analitycznej postaci wszystkich występujących tam wielkości, a samo rozwiązanie zostało efektywnie wyznaczone z użyciem klasycznego algorytmu Newtona [Kincaid & Cheney, 2002]. Ostatecznie podana została kompletna procedura wyznaczania wartości estymatora szukanego parametru dla ustalonej wartości czynnika warunkującego, na podstawie pomiarów tego parametru, otrzymanych uprzednio dla różnych wartości warunkujących.

W efekcie kontynuacji badań problem identyfikacji wartości parametrów został uogólniony do przypadku wielowymiarowego, gdy identyfikowanych jest kilka parametrów, traktowanych jako wektor, a straty wynikłe z ich przeszacowania i niedoszacowania mogą być niesymetryczne i wzajemnie skorelowane [poz. H1 osiągnięcia naukowego, „*An Algorithm for Conditional Multidimensional Parameter Identification with Asymmetric and Correlated Losses of Under- and Overestimations*”]. Końcowym wynikiem badań był algorytm wyznaczania wartości estymatora optymalnego w sensie wartości oczekiwanej strat, z użyciem wielowymiarowej niesymetrycznej funkcji kwadratowej, przy dowolnym rozkładzie zmiennych opisujących i warunkujących. Poszczególne współrzędne zmiennych warunkujących mogą być typu ciągłego, binarnego, dyskretnego lub kategoriowego. Poprawne działanie algorytmu zostało potwierdzone za pomocą licznych badań numerycznych przy danych sztucznie generowanych dla celów ilustracyjnych, a także przebiegów symulacyjnych oraz aplikacji do zadań praktycznych z zakresu inżynierii sterowania i medycyny. Warto zwrócić uwagę na następujące, pozytywne cechy opracowanej procedury:

- wartości wyznaczanego estymatora przesunęły się w kierunkach związanych z mniejszymi stratami wynikłymi z błędów estymacji, określanymi przyjętą funkcją strat;
- ujęcie warunkowe skutkowało odpowiednią korektą wartości estymatora, zgodnie z charakterem korelacji między zmiennymi opisującymi i warunkującymi. Jeżeli parametr był dodatnio skorelowany z czynnikiem warunkującym, to zwiększenie (zmniejszenie) wartości warunkującej powodowało powiększenie (pomniejszenie) wartości estymatora tego parametru, odwrotne skutki występowały przy ujemnej korelacji;
- dzięki uśredniającym własnościom estymatorów jądrowych algorytm okazał się odporny na małą liczbę lub wręcz brak danych z otoczenia wartości warunkującej;
- w miarę wzrostu liczności próby losowej wyznaczone wartości estymatora zmierzały do wartości teoretycznej, a odchylenie standardowe błędu sukcesywnie malało do zera.

Przedstawiona koncepcja ma charakter uniwersalny i może być stosowana do szerokiego zakresu problemów badawczych. Możliwe jest również wprowadzenie indywidualnych modyfikacji oraz generalizacji, dostosowanych do nietypowych zagadnień aplikacyjnych.

Obecnie, w związku z coraz większą złożonością i różnorodnością gromadzonych danych zmniejsza się możliwość stosowania klasycznych parametrów i ich estymatorów [Efromovich, 1999]. Występowanie obserwacji oddalonych od parametrów centralnych ma niekorzystny wpływ na ich własności i może skutkować błędnymi decyzjami, podejmowanymi przy ich użyciu. W związku z tym metody nieparametryczne znajdują coraz szersze zastosowanie w problemach współczesnej analizy i eksploracji danych, a istotnym zagadnieniem staje się uzyskanie estymatora takich charakterystyk jak dystrybuanta i kwantyl. Jakość estymacji charakterystyk rozkładu probabilistycznego można efektywnie zwiększyć poprzez wyodrębnienie wielkości mających istotny wpływ na badane zjawisko i odpowiednie wprowadzenie ich do algorytmu obliczeniowego. Estymacja wartości kwantyla nie jest obecnie zagadnieniem opracowanym wyczerpująco. Dla typowych niezłożonych zastosowań na uwagę zasługują estymatory pozycyjne [Parrish, 1990] oraz jądrowe [Sheather & Marron, 1990]. W przypadku kwantyla warunkowego najczęściej stosowaną jest metoda regresji kwantylowej [Koenker, 2005]. Metoda ta daje w większości przypadków satysfakcjonujące wyniki, jednak w wielu praktycznych zastosowaniach prowadzi do trudności interpretacyjnych. Powyższe implikuje badania nad odmiennymi koncepcjami, także wykraczające poza metody statystyczne.

W ramach dalszych badań naukowych w pracy [poz. H3 osiągnięcia naukowego] pt. „*A Convenient Ready-to-Use Algorithm for a Conditional Quantile Estimator*” opracowano aparat matematyczny i zaprojektowano kompletny algorytm wyznaczenia wartości jądrowego estymatora kwantyla w ujęciu warunkowym. W odróżnieniu od stosowanego podejścia klasycznego nieparametryczny charakter metody stanowi, iż opracowana procedura jest niezależna od typu rozkładów zmiennych losowych. Dzięki prostocie samej interpretacji w artykule wskazano na możliwości tworzenia indywidualnych modyfikacji, przydatnych w praktyce w konkretnych specyficznych zastosowaniach, w szczególności ewentualnego zwiększenia dokładności estymacji oraz uogólnienia zmiennych warunkujących także o binarne, dyskretne oraz kategoriowe (uporządkowane i nieuporządkowane), jak również ich kombinacje.

Wraz z rozwojem technik obrazowania tomograficznego rozwijane są obecnie metody ukierunkowane na automatyczną analizę obrazów. Badania, których wyniki opublikowano w pracy [poz. H2 osiągnięcia naukowego] pt. „*Impact of long term mineral and organic fertilization on water stability, wettability and porosity of aggregates of two silt loamy*

soils”, umożliwiły opracowanie nowoczesnej procedury określenia rozkładu wielkości porów agregatów glebowych oraz wyodrębnienia parametrów określających ich trwałość i porowatość. Obiektem badań były dwie gleby: czarna ziemia o składzie gleby pylastej (Mollic Gleysol) i gleba płowa o składzie piasku pylastego mocnego (Haplic Luvisol), na których od wielu lat stosowany był określony płodozmian oraz stały system nawożenia na wydzielonych poletkach: kontrola (bez nawożenia), nawożenie obornikiem i nawożenie mineralne. Wykonano następujące badania agregatów pobranych z warstwy ornej sześciu poletek doświadczalnych:

- separację agregatów glebowych i pomiar ich wodoodporności przy pomocy nowej, opracowanej na potrzeby badań metody, polegającej na określeniu frakcji agregatów, które nie uległy rozpadowi podczas nawilżania w roztworze wodnym alkoholu metylowego o stężeniach 0, 20, 40 i 60%;
- analizę glebową materii organicznej dla wszystkich typów agregatów, zróżnicowanych pod względem składu, nawożenia i zastosowanego roztworu;
- pomiar sorpcyjności wodnej i kąta zwilżania wody dla wszystkich rodzajów agregatów glebowych;
- pomiar porowatości całkowitej agregatów o najwyższej i najniższej trwałości metodą parafinową;
- scharakteryzowanie międzygranulalnego menisku wodnego jako czynnika stabilności agregatów;
- badania mikrotomograficzne struktury porów agregatów o najwyższej i najniższej trwałości.

Opracowany w pracy algorytm wyznaczania rozkładu wielkości porów, oparty na rentgenowskiej mikrotomografii komputerowej, obejmuje:

- eliminację artefaktów poprzez wydzielenie do analizy fragmentu przekroju metodą ROI;
- binaryzację obrazu z automatycznie wyznaczanym progiem binaryzacji dla granicy powietrze – faza stała gleby;
- procedurę domknięcia, która pozwoliła określić powierzchnię przekroju porów o danym promieniu, odpowiadającą przyjętej wielkości elementu strukturalnego; procedura wykonywana jest iteracyjnie, aż do skwantyfikowania wszystkich porów, z dostępną wizualizacją każdego kroku pętli;
- procedurę wyznaczenia, w oparciu o kryteria optymalizacyjne, analitycznej postaci funkcji opisującej rozkład wielkości porów na podstawie pomiarów powierzchni zajmowanej przez pory o znanych rozmiarach.

Zastosowanie opracowanego algorytmu umożliwiło zidentyfikowanie typu rozkładu wielkości porów i podanie parametrów charakteryzujących ten rozkład oraz określenie czynników wpływających na stopień agregacji gleby. Powszechnie uważa się, że wśród czynników atmosferycznych najważniejszymi są opady deszczu i wilgotność podczas okresu wegetacji [Amezketta, 1999; Goldberg et al., 1988; Martinez-Mena et al., 1998]. W wyniku przeprowadzonych badań stwierdzono, że czysta woda jest najbardziej destrukcyjną cieczą dla agregatów glebowych oraz zaproponowano nowy parametr odporności całkowitej, charakteryzujący stabilność agregatów. Rodzaj nawożenia, zawartość materii organicznej, jej skład oraz związane z nią różnice w rozkładzie porów i zwilżalności istotnie różnicują odporność agregatów na destrukcyjne działanie wody. Agregaty o większej porowatości są bardziej stabilne, przy czym ich odporność jest dodatnio skorelowana z udziałem mezo- i makroporów. Otrzymane wyniki i analiza parametrów fizykochemicznych pozwoliły na wnikliwą analizę wodoodporności i

trwałości agregatów, dostarczając zarówno wskazówek praktycznych w zakresie uprawy gleby, jak i naukowych, dotyczących nowych metod badawczych, pozwalających na badanie struktury i własności agregatów. Warty odnotowania jest fakt, iż w przeciwieństwie do klasycznych metod badania porowatości opracowana procedura nie ma działania destrukcyjnego na agregaty glebowe.

Obecnie nowoczesna medycyna nie może prawidłowo funkcjonować bez komputerowej analizy danych. Rozwój informatyki sprzyja popularyzacji metodologicznego nurtu statystyki w coraz to szerszych zastosowaniach empirycznych [Jędrzychowski & Penar, 2000; Tadeusiewicz, 2004]. W wyniku prowadzonych badań opracowane zostały procedury badawcze i zaimplementowane potrzebne algorytmy, umożliwiające wyodrębnienie czynników inicjujących wybrane procesy chorobowe. Celem prac przedstawionych w artykule [poz. H6 osiągnięcia naukowego] pt. „*The content of cadmium, cobalt and nickel in laryngeal carcinoma*” było uzyskanie wiedzy na temat czynników inicjujących rozwój nowotworu krtani. Dane dotyczące struktury zachorowań na nowotwory złośliwe głowy i szyi w Polsce wskazują, iż ponad 10% nowotworów tej grupy stanowi rak krtani, przy czym badania epidemiologiczne zwracają uwagę na wzrost zachorowalności na raka krtani, zarówno wśród mężczyzn jak i kobiet [Klatka, 2002]. Ryzyko zachorowania na raka krtani zmienia się w istotny sposób w zależności od płci i wieku. Szczyt zachorowań obserwuje się w przedziale od 50 do 70 lat. Mężczyźni chorują zdecydowanie częściej niż kobiety, przy czym stwierdzono zbliżoną ekspozycję obu płci na niekorzystne czynniki środowiskowe, z których na pierwszym miejscu wymienia się palenie tytoniu i obecność metali ciężkich. Nie bez znaczenia jest też rodzaj wykonywanego zawodu oraz udział komponenty genetycznej. W pracy zaproponowano procedurę porównania zawartości kadmu, kobaltu i niklu w tkance nowotworowej z poziomem tych samych pierwiastków w tkance zdrowej wśród mężczyzn w wieku od 44 do 76 lat. Stężenie badanych pierwiastków określono metodą indukcyjnie sprzężonej plazmowej spektroskopii emisyjnej. Poziom kadmu, kobaltu i niklu był wyższy w tkance nowotworowej. Stwierdzono istotnie wyższe stężenie kadmu u pacjentów z przerzutami do węzłów chłonnych w porównaniu do pacjentów, u których nie stwierdzono przerzutów. Poziom kobaltu był istotnie wyższy w przypadku największego stopnia zaawansowania nowotworu krtani. Nierównowaga w poziomie niklu, kobaltu i kadmu może być spowodowana zmianą metabolizmu komórkowego w procesie choroby nowotworowej. Ponadto wyniki przeprowadzonych badań wykazały istotnie wyższe stężenie tych metali w grupie pacjentów pochodzących z obszarów wiejskich, w porównaniu z grupą pacjentów pochodzących z obszarów miejskich, co sugeruje związek z czynnikami środowiskowymi lub zawodowymi.

Schorzenia immunologiczne tarczycy stanowią około 30% wszystkich chorób autoimmunologicznych, wśród których jedną z najczęściej występujących jest choroba Gravesa-Basedowa [Klatka et al., 2012]. Podstawowym sposobem leczenia tej choroby są leki tyreostatyczne, z których najczęściej stosowanym w Europie jest metimazol. Aktywność supresyjna komórek T regulatorowych (Treg) wobec limfocytów autoreaktywnych zapewnia ogólnoustrojową ochronę przed reakcjami o podłożu autoimmunologicznym. Z tego względu prowadzone są obecnie badania nad mechanizmami regulacji aktywności limfocytów Treg. W pracy [poz. H4 osiągnięcia naukowego] pt. „*Th17 and Treg cells in adolescents with Graves' disease. Impact of treatment with methimazole on these cell subsets*” zaproponowano procedurę badawczą obejmującą określenie roli komórek T regulatorowych i limfocytów pomocniczych Th17 w patogenezie choroby Gravesa. Celem badania była ocena zmian w częstotliwości występowania komórek Th17 i Treg w trakcie leczenia tej choroby metimizolem w grupie

młodzieży. Stwierdzono istotnie niższe występowanie komórek Treg oraz istotnie wyższe występowanie limfocytów Th17 przed podaniem metamizolu w porównaniu z wynikami po zastosowaniu leczenia. Poziom komórek Treg po zastosowaniu leczenia wzrósł i był porównywalny z wynikami z grupy kontrolnej. Poziom limfocytów Th17 po zastosowaniu metamizolu obniżył się, ale był istotnie wyższy w porównaniu z grupą kontrolną. Wyniki potwierdziły istotną rolę komórek Treg i Th17. Limfocyty T regulatorowe mogą za pośrednictwem komórek dendrytycznych hamować aktywację limfocytów T, jednak ich udział nie jest do końca wyjaśniony, co skłania do prowadzenia dalszych badań. Ich identyfikacja jest jednym z najważniejszych osiągnięć współczesnej immunologii, otwierającym nowe możliwości interwencji terapeutycznej w chorobach autoagresywnych.

Przedstawioną tematykę mojej pracy naukowo-badawczej można podsumować następująco.

Badania z zakresu wyostrażania informacji nieprecyzyjnej umożliwiły opracowanie efektywnego algorytmu wyznaczania bayesowskiego determinizatora takiej informacji oraz zastosowanie opracowanej procedury do rzeczywistych zagadnień w naukach technicznych i medycznych. W wyniku badań dotyczących identyfikacji parametrów zaprojektowane i zaimplementowane zostały nowoczesne procedury wyznaczania wartości estymatora szukanego parametru, uwzględniające czynniki warunkujące różnego typu, niesymetrię strat błędów estymacji i dowolność rozkładu występujących zmiennych. Użyta tu metodyka statystycznych estymatorów jądrowych stała się przedmiotem dalszych badań i posłużyła do opracowania kompletnego gradientowego algorytmu klasteryzacji oraz algorytmu wyznaczania warunkowego jądrowego estymatora kwantyla. Istotą algorytmu klasteryzacji jest brak wymagań dotyczących ścisłego ustalenia liczby klastrów oraz wpływ na ich rząd wielkości jak również proporcję między liczbą klastrów w obszarach zagęszczenia elementów oraz w obszarach, gdzie są one rzadkie. W szczególności możliwe jest zmniejszenie liczby lub wyeliminowanie klastrów w obszarach rzadszego występowania elementów zbioru danych, co w konsekwencji umożliwia usunięcie wpływu elementów odosobnionych, bez ingerencji w liczbę klastrów w obszarach ich zagęszczenia. Z kolei nieparametryczny charakter metody wyznaczania estymatora kwantyla uniezależnił zaproponowaną procedurę od typu rozkładów zmiennych losowych. Ponadto możliwość uwzględnienia czynników warunkujących różnego typu, w tym binarnego, dyskretnego i kategoriowego oraz prostota samej interpretacji algorytmu istotnie zwiększyły jego możliwości aplikacyjne.

W ramach współpracy naukowej ze specjalistami z innych dziedzin przedmiotem moich badań były również specjalistyczne procedury umożliwiające analizę danych doświadczalnych i obrazowych. W wyniku realizacji w Instytucie Agrofizyki PAN w Lublinie projektu badawczego pt. „*Wodoodporność agregatów warstwy ornej gleb mineralnych*” opracowana została nowatorska procedura oceny trwałości i porowatości agregatów glebowych, oparta na mikrotomografii komputerowej. W ramach badań zostały zaprojektowane i zaimplementowane nowoczesne algorytmy przetwarzania i analizy obrazów cyfrowych, które mogą być z powodzeniem zastosowane do scharakteryzowania struktury różnorodnych materiałów porowatych. Z kolei badania prowadzone we współpracy z Uniwersytetem Medycznym w Lublinie pozwoliły na opracowanie procedur badawczych poszerzających stan wiedzy na temat czynników inicjujących rozwój nowotworu krtani oraz rolę komórek Treg i Th17 w patogenezie choroby Gravesa-Basedova. W wyniku przeprowadzonych badań zaprojektowane zostały zaawansowane algorytmy wielowymiarowej analizy danych, służące do modelowania złożonych systemów biologicznych, trudnych do opisanego za pomocą klasycznych metod. Należy zauważyć, że właściwa analiza materiału badawczego przekłada się w tym przypadku na

implementację wyników badań w działania praktyczne, wykorzystując je nie tylko w aspekcie naukowym, ale również w procesie leczenia.

Literatura

- Aggarwal, C.C. (2013). *Outlier Analysis*. Springer, New York.
- Amezqueta, E. (1999) Soil aggregate stability: a review. *Journal of Sustainable Agriculture*, vol. 14, ss. 83-151.
- Efromovich, S. (1999) *Nonparametric Curve Estimation: Methods, Theory and Applications*. Springer, New York.
- Gerstenkorn, T., Kurnatowska, A., Rakus, E. (1990) Zastosowanie teorii zbiorów rozmytych w diagnostyce i leczeniu stanów zapalnych narządów płciowych i układu moczowego kobiet. *Wiadomości Parazytologiczne*, vol. 36, ss. 251-267.
- Goldberg, S., Suarez, D.L., Glaubig, R.A. (1988) Factors affecting clay dispersion and aggregates stability of arid-zone soils. *Soil Science*, vol. 146, ss. 317-325.
- Jędrychowski, W., Penar, A. (2000) *Statystyczna analiza wyników badań naukowych w medycynie i biologii*. Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
- Klatka, J. (2002) *Ocena wartości prognostycznej antygenów P53, BCL-2 i syndekanu-1 w raku krtani oraz ich przydatności w charakterystyce odpowiedzi pierwotnych hodowli komórkowych raka krtani na cis-platynę*. Klinika Otolaryngologii Akademii Medycznej, Lublin.
- Klatka, M., Szewczyk, L., Grywalska, E. (2012) The Role of T Regulatory Cells and Dendritic Cells in Hyperthyroidism. *Pediatric Endocrinology*, vol. 11, ss. 65-72.
- Kincaid, D., Cheney, W. (2002) *Numerical Analysis*. Brooks/Cole, Pacific Grove.
- Koenker, R. (2005) *Quantile Regression*. Cambridge University Press, Cambridge.
- Kulczycki, P. (2005) *Estymatory jądrowe w analizie systemowej*, Wydawnictwo Naukowo-Techniczne, Warszawa.
- Lubischew, A.A., (1962) On the Use of Discriminant Functions In Taxonomy. *Biometrics*, vol. 18, ss. 455-478.
- Martinez-Mena, M., Williams, A.G., Ternan, J.L., Fitzjohn, C. (1998) Role of antecedent soil water content on aggregates stability in a semi-arid environment. *Soil & Tillage Research*, vol. 48, ss. 71-80.
- Parrish, R.S. (1990) Comparison of Quantile Estimators in Normal Sampling. *Biometrics*, vol. 46, ss. 247-257.
- Sheather, S.J., Marron, J.S. (1990) Kernel quantile estimators. *Journal of the American Statistical Association*, vol. 85, ss. 410-416.
- Tadeusiewicz, R. (2004) *Medical image understanding technology: artificial intelligence and soft-computing for image understanding*. Springer, Berlin.
- Walter, E., Pronzato, I. (1997) *Identification of Parametric Models*. Springer-Verlag, Berlin.
- Wand, M.P., Jones, M.C. (1995) *Kernel Smoothing*. Chapman and Hall, London.

5. Omówienie pozostałych osiągnięć naukowo-badawczych

Po ukończeniu studiów doktoranckich i obronie rozprawy doktorskiej w 2005 roku podjęłam pracę na stanowisku adiunkta w Instytucie Badań Systemowych PAN oraz w Instytucie Matematyki i Informatyki Katolickiego Uniwersytetu Lubelskiego, gdzie rozwijałam działalność badawczą w dziedzinie informatyki. Moje badania naukowe były związane z opracowaniem i praktycznym zastosowaniem algorytmów wielowymiarowej analizy i eksploracji danych do rozwiązywania złożonych zagadnień w naukach technicznych i medycznych.

Początkowe badania z zakresu bayesowskiej defuzyfikacji informacji nieprecyzyjnej wykazały duże możliwości aplikacyjne metodyki statystycznych estymatorów jądrowych. W wyniku dalszych prac badawczych w artykule [13] zaproponowałam wykorzystanie jądrowego estymatora funkcji regresji w diagnostyce medycznej. Celem niniejszej pracy było opracowanie nowoczesnej procedury ustalenia na podstawie badań biochemicznych i hematologicznych pacjentów stopnia niewydolności nerek wskazującego na początek niedokrwistości. Nieleczona niedokrwistość prowadzi do szeregu zaburzeń fizjologicznych i skraca czas przeżycia. Zasadniczą rolę w patogenezie niedokrwistości odgrywa niedobór erytropoetyny nerkowej, spowodowany niszczeniem miększu nerek, w których wytwarzane jest około 90% tego hormonu. O unikalności zebranego materiału badawczego i otrzymanych wyników stanowił fakt, iż badanej grupie chorych nie podawano w trakcie leczenia ludzkiej rekombinowanej erytropoetyny. Przedstawiona procedura uniezależniła metodę od arbitralnych założeń dotyczących stosowanej funkcji regresji i charakteryzowała się mniejszym błędem estymacji w porównaniu z metodami klasycznymi.

Z kolei kompletny gradientowy algorytm klasteryzacji [11] został pozytywnie zweryfikowany w zastosowaniu do zadania syntezy regulatorów rozmytych przy wyodrębnianiu zbioru ich reguł, klasteryzacji ziaren różnych odmian pszenicy bez apriorycznej informacji o liczbie tych odmian, a także w ramach złożonej i zróżnicowanej metodologicznie procedury wspomagania strategii marketingowej operatora telefonii komórkowej [2, 4, 9]. Celem badań było zaprezentowanie aplikacyjnych aspektów i własności algorytmu, a także ich ilustracja dla konkretnych, praktycznych zastosowań. Uzyskane wyniki potwierdziły praktyczną użyteczność metody, w szczególności brak zasadności istotnych zmian parametru wygładzania wpływających bezpośrednio na liczbę klastrów, co wskazuje na prawidłowe dopasowanie procedury do rzeczywistej struktury danych. Bardzo cenna okazała się w praktyce możliwość zmiany wartości parametru stanowiącego o relacji liczby klastrów w obszarach zagęszczenia elementów próby wobec rejonów ich rzadkiego występowania. Własność ta jest szczególnie warta podkreślenia, gdyż możliwość taka nie występuje w innych, znanych algorytmach klasteryzacji. Użyteczność opracowanego algorytmu potwierdzona również została w przypadku segmentacji obrazów cyfrowych, umożliwiając scharakteryzowanie struktury gleby na podstawie przekrojów tomograficznych agregatów glebowych [7].

Wiele współczesnych zagadnień wymaga uwzględnienia wpływu nieuniknionych w praktyce błędów identyfikacji parametrów występujących w modelu. Dalsze badania prowadzone we współpracy z prof. Piotrem Kulczyckim z Instytutu Badań Systemowych PAN dotyczyły procedury estymacji wartości parametru w ujęciu warunkowym, pozwalającym na uwzględnienie zmiennych warunkujących różnego typu [3, 10]. Omawiane zagadnienie zostało uogólnione do przypadku wielowymiarowego, gdy identyfikowanych jest kilka parametrów. W efekcie przedstawiono kompletny algorytm wyznaczania wartości estymowanych parametrów, optymalnych w sensie minimum wartości oczekiwanej strat, gdy straty wynikłe z dodatnich i ujemnych błędów estymacji są

różne i dodatkowo skorelowane pomiędzy poszczególnymi parametrami. Procedura opracowana została w ujęciu warunkowym, co pozwoliło uściślić wynik o ustaloną (np. aktualną) wartość czynników warunkujących [6].

Współpraca z dr hab. Jerzym Niewczasem z Katedry Analizy Obrazów KUL oraz dr hab. Henrykiem Czachorem z Instytutu Agrofizyki PAN i prof. Lubomirem Lichnerem z Instytut Hydrologii Słowackiej Akademii Nauk pozwoliła na opracowanie nowoczesnych metod przetwarzania i analizy obrazów cyfrowych oraz ich zastosowanie do zagadnień z dziedziny nauk technicznych. Wyniki badań dotyczące klasyfikacji odmian pszenicy na podstawie cech geometrycznych ich ziaren zaprezentowane zostały w pracy [12]. Z kolei metoda określenia struktury materiałów porowatych, oparta na rentgenowskiej mikrotomografii komputerowej, przedstawiona w pracy [8], stanowi nieniszczącą alternatywę dla posiadających istotne ograniczenia metod klasycznych, takich jak porozymetria rtęciowa i metoda adsorpcji gazowej.

W konsekwencji współpracy z prof. Januszem Klatką z Uniwersytetu Medycznego w Lublinie byłam zaangażowana w planowanie badań i opracowywanie procedur badawczych oraz analizę uzyskanych wyników. W pracy [1] opracowana została procedura zidentyfikowania czynników odgrywających ważną rolę w patogenezie choroby Gravesa oraz w trakcie jej leczenia. Celem artykułu [5] była natomiast analiza przyczyn problemów zdrowotnych osób uczących się i pracujących w wieku od 18 do 30 lat, ze szczególnym uwzględnieniem rodzaju hospitalizacji. Należy podkreślić, iż dane medyczne charakteryzuje duży stopień zmienności, zarówno międzyosobniczej, jak i wewnątrzosobniczej, a sama specyfika analizy z wykorzystaniem narzędzi informatycznych nie ma wypracowanych stałych standardów i wzorców. Dodatkowy problem stanowi uzyskanie próby o wystarczającej liczebności w przypadku rzadkich jednostek chorobowych i kosztownych procedur badawczych. W związku z tym każdorazowo wymagają one wnikliwej analizy oraz uwzględnienia dodatkowych czynników, typu ciągłego i kategorycznego, warunkujących procesy zachodzące w organizmie. Współczesna medycyna, oprócz klasycznych metod statystycznych, w coraz szerszym zakresie wykorzystuje nowoczesne metody i narzędzia do gromadzenia, przetwarzania i analizy danych.

Wyniki badań zostały opublikowane w cenionych w skali światowej czasopismach i pracach zbiorowych w dziedzinie informatyki. Mój dorobek publikacyjny obejmuje 40 prac, w tym 33 opublikowane po uzyskaniu stopnia doktora, a także 28 referatów konferencyjnych (25 po uzyskaniu stopnia doktora):

- 14 artykułów w czasopismach wymienionych w bazie JCR, w tym 13 opublikowanych po uzyskaniu stopnia doktora;
- 7 artykułów w recenzowanych czasopismach spoza bazy JCR, w tym 4 opublikowane po uzyskaniu stopnia doktora;
- 8 rozdziałów w pracach zbiorowych, wydanych przez wydawnictwo Springer, opublikowanych po uzyskaniu stopnia doktora;
- 11 rozdziałów w pracach zbiorowych wydanych przez wydawnictwa inne niż Springer, w tym 8 opublikowanych po uzyskaniu stopnia doktora;
- 28 referatów konferencyjnych, w tym 25 wygłoszonych po uzyskaniu stopnia doktora.

Suma punktów za publikacje według MNiSW wynosi 463, w tym za prace opublikowane po uzyskaniu stopnia doktora – 417.

Poniżej zostaną podane ważniejsze publikacje, opublikowane po uzyskaniu stopnia doktora i nie wymienione w sekcji 4.2, związane z tematem osiągnięcia naukowego.

Publikacje naukowe w czasopismach znajdujących się w bazie Journal Citation Reports JCR (wybrane pozycje nie wchodzące w skład osiągnięcia naukowego):

1. M. Klatka, E. Grywalska, M. Partyka, **M. Charytanowicz (25%)**, J. Roliński, Impact of Methimazole Treatment on Concentration and Lymphocytes Activation in Adolescents with Graves' Disease. *Biological Trace Element Research*, vol. 153, nr 1-3, ss. 155-170, 2013.

JCR, IF/5-Year IF: 1,748/1,699; MNiSW: Lista A, 15 pkt.

2. P. Kulczycki, **M. Charytanowicz (25%)**, P.A. Kowalski, S. Łukasik, The Complete Gradient Clustering Algorithm: Properties in Practical Applications. *Journal of Applied Statistics*, vol. 39, nr 6, ss. 1211-1224, 2012.

JCR, IF/5-Year IF: 0,417/0,591; MNiSW: Lista A, 15 pkt.

3. P. Kulczycki, **M. Charytanowicz (50%)**, Asymmetrical Conditional Bayes Parameter Identification for Control Engineering. *Cybernetics and Systems*, vol. 39, nr 3, ss. 229-243, 2008.

JCR, IF/5-Year IF: 0,840/0,968; MNiSW: Lista A, 20 pkt.

4. P.A. Kowalski, S. Łukasik, **M. Charytanowicz (25%)**, P. Kulczycki, Data-Driven Fuzzy Modeling and Control with Kernel Density Based Clustering Technique. *Polish Journal of Environmental Studies*, vol. 17, nr 4C, ss. 83-87, 2008.

JCR, IF/5-Year IF: 0,871/0,888; MNiSW: Lista A, 15 pkt.

5. R. Czabak-Garbacz, B. Cygan, A. Mizerska, J. Mosiewicz, J. Stążka, J. Sekita-Krzak, **M. Charytanowicz (20%)**, A. Burghart, Influence of University Study Course on Health of Young People. *Polish Journal of Environmental Studies*, vol. 15, nr 2B, ss. 1051-1053, 2006.

JCR, IF/5-Year IF: 0,871/0,888; MNiSW: Lista A, 15 pkt.

IF oraz pięcioletni IF (5-Year IF) podany został według bazy JCR zgodnie z dostępem z dnia 10 sierpnia 2015. Liczbę punktów przyznawaną za publikację wyszczególniono zgodnie z obowiązującym w roku 2015 wykazem czasopism naukowych, ogłoszonym przez Ministra Nauki i Szkolnictwa Wyższego z dnia 31 grudnia 2014 r.

Rozdziały w pracach zbiorowych (wybrane pozycje nie wchodzące w skład osiągnięcia naukowego):

6. P. Kulczycki, **M. Charytanowicz (50%)**, Conditional Multidimensional Parameter Identification with Asymmetric Correlated Losses of Estimation Errors. *Neural Information Processing*, C.K. Loo, K.S. Yap, K.W. Wong, A. Teoh, K. Huang (red.), Lecture Notes in Computer Science, Springer, 2014, vol. II, ss. 287-294.

MNiSW: 5 pkt.

7. **M. Charytanowicz (50%)**, P. Kulczycki, An Image Analysis Algorithm for Soil Structure Identification. *Intelligent Systems'2014*, D. Filev, J. Jabłkowski, J. Kacprzyk, I. Popchev, L. Rutkowski, V. Sgurev, E. Sotirova, P. Szyrkarczyk, S. Zadrozny (red.), *Advances in Intelligent Systems and Computing*, Springer, 2014, vol. II, ss. 681-692.

MNiSW: 5 pkt.

8. **M. Charytanowicz (100%)**, An Algorithm for the Pore Size Determination using Digital Image Analysis. *Information Technologies in Biomedicine*, E. Pietka, J. Kawa, W. Więclawek (red.), *Advances in Intelligent Systems and Soft Computing*, Springer, 2014, vol. 3, ss. 223-234.

MNiSW: 5 pkt.

9. P. Kulczycki, **M. Charytanowicz (25%)**, P.A. Kowalski, S. Łukasik, Exemplary Applications of the Complete Gradient Clustering Algorithm in Bioinformatics, Management and Engineering. *Issues and Challenges of Intelligent Systems and Computational Intelligence*, L.T. Kóczy, C. Pozna, J. Kacprzyk (red.), *Studies in Computational Intelligence*, Springer, 2014, ss. 119-132.

MNiSW: 5 pkt.

10. P. Kulczycki, **M. Charytanowicz (50%)**, Conditional Parameter Identification with Asymmetrical Losses of Estimation Errors. *Computational Collective Intelligence. Technologies and Applications*, N.T. Nguyen, K. Hoang, P. Jedrzejowicz (red.), *Lecture Notes in Artificial Intelligence*, Springer, 2012, vol. I, ss. 553-562.

MNiSW: 10 pkt.¹

11. P. Kulczycki, **M. Charytanowicz (50%)**, A Complete Gradient Clustering Algorithm. *Artificial Intelligence and Computational Intelligence*, H. Deng, D. Miao, J. Lei, F.L. Wang (red.), *Lecture Notes in Computer Science*, Springer, 2011, vol. III, ss. 497-504.

MNiSW: 10 pkt.¹

12. **M. Charytanowicz (20%)**, J. Niewczas, P. Kulczycki, P.A. Kowalski, S. Łukasik, S. Żak, Complete Gradient Clustering Algorithm for Features Analysis of X-Ray Images. *Information Technologies in Biomedicine*, E. Pietka, J. Kawa (red.), *Advances in Intelligent Systems and Soft Computing*, Springer, 2010, vol. 2, ss. 15-24.

MNiSW: 10 pkt.¹

13. **M. Charytanowicz (50%)**, P. Kulczycki, Nonparametric Regression for Analyzing Correlation between Medical Parameters. *Information Technologies in Biomedicine*, E. Pietka, J. Kawa (red.), *Advances in Soft Computing*, Springer, 2008, ss. 437-444.

MNiSW: 10 pkt.¹

¹ Publikacja indeksowana w bazie Web of Science.

Liczba punktów przyznawana za publikacje podana została zgodnie z obwieszczeniem MNiSW z dnia 5 marca 2014 r. w sprawie ogłoszenia jednolitego tekstu rozporządzenia Ministra Nauki i Szkolnictwa Wyższego w sprawie kryteriów i trybu przyznawania kategorii naukowej jednostkom naukowym.

Całość dorobku naukowego wraz z pozostałymi osiągnięciami i działalność akademicka ujęte zostały w *Wykazie opublikowanych prac naukowych lub twórczych prac zawodowych oraz informacjach o osiągnięciach dydaktycznych, współpracy naukowej i popularyzacji nauki (Załącznik 4).*

Małgorzata Charytanowicz