

AUTOREFERAT

Dr DARIUSZ GOZDOWSKI

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Wydział Rolnictwa i Biologii

Katedra Doświadczalnictwa i Bioinformatyki

Warszawa 2016

1. Imię i Nazwisko

Dariusz Gozdowski

2. Posiadane dyplomy

2005 – stopień doktora nauk rolniczych w zakresie agronomii, Wydział Rolnictwa i Biologii, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie. Tytuł rozprawy doktorskiej: *Struktura i architektura łanu a plonowanie dwóch odmian jęczmienia jarego w zależności od wybranych czynników agrotechnicznych*. Promotor: prof. dr hab. Zdzisław Wyszyński.

1999 – magister inżynier na kierunku rolnictwo, Wydział Rolniczy, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie. Tytuł pracy magisterskiej: *Wpływ zmiennego pola magnetycznego o niskiej częstotliwości na wartość siewną nasion buraka cukrowego*. Promotor: dr Małgorzata Rochalska.

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych

01.12.2005 – obecnie: Katedra Doświadczalnictwa i Bioinformatyki (do roku 2008 Katedra Biometrii), Wydział Rolnictwa i Biologii, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, adiunkt.

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

a) Tytuł osiągnięcia naukowego:

Zastosowanie metod statystyki przestrzennej w badaniach rolniczej przestrzeni produkcyjnej

b) (autor/autorzy, tytuły publikacji, rok wydania, nazwa wydawnictwa, recenzenci wydawniczy)

1. **Gozdowski D.**, Samborski S., Sioma S. 2007. *Rolnictwo precyzyjne* (monografia). Wydawnictwo SGGW, s. 136 [MNiSW₂₀₀₇=12 pkt, mój udział procentowy szacuję na 45%].

2. **Gozdowski D.**, Samborski S., Dobers E.S. 2010. Evaluation of methods for the detection of spatial outliers in the yield data of winter wheat. *Colloquium Biometricum*, 40, 41-51. [MNiSW₂₀₁₀= 4 pkt, mój udział procentowy szacuję na 50%]

3. **Gozdowski D.**, Roszkowska-Mądra B. 2010. Detection of spatial outliers in rural communes of Podlasie province (Poland) on the basis of socio-economic attributes. *Journal of Central European Agriculture*, 11, 409-414. [MNiSW₂₀₁₀=8 pkt, mój udział procentowy szacuję na 70%]

4. Stępień M., **Gozdowski D.**, Samborski S. 2013 A case study on the estimation accuracy of soil properties and fertilizer rates for different soil-sampling grids. *Journal of Plant Nutrition and Soil Science*, 176, 57-68. [IF₂₀₁₃=1,663, MNiSW₂₀₁₃=30 pkt, mój udział procentowy szacuję na 30%].

5. **Gozdowski D.**, Stępień M, Samborski S. Dobers S., Chormański J., Szatyłowicz J. 2014. Determination of the most relevant soil properties for the delineation of management zones in production fields. *Communications in Soil Science and Plant Analysis*, 45, 2289-2304. [IF₂₀₁₄=0,390, MNiSW₂₀₁₄=15 pkt, mój udział procentowy szacuję na 40%]

6. **Gozdowski D.**, Stępień M., Samborski S., Dobers E.S, Szatyłowicz J., Chormański J. 2015. Prediction accuracy of selected spatial interpolation methods for soil texture at farm field scale: *Journal of Soil Science and Plant Nutrition*, 15, 639-650. [IF₂₀₁₅=1,600, MNiSW₂₀₁₅=20 pkt, mój udział procentowy szacuję na 50%].

7. **Gozdowski D.** 2016. Relationships between selected soil properties examined by the LUCAS project and satellite-derived vegetation indices for Poland. *Fresenius Environmental Bulletin*, 25, 641-646. [IF₂₀₁₅=0,372, MNiSW₂₀₁₅=15 pkt]

Łączny IF ww. publikacji zgodny z rokiem wydania (lub w przypadku najnowszych prac zgodny z najnowszym, tj. IF₂₀₁₅) wynosi 4,025. Łączna liczba punktów wg punktacji MNiSW zgodna z rokiem opublikowania wynosi 104 pkt.

c) Omówienie celu naukowego ww. prac i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania

Wprowadzenie

Badania nad rolniczą przestrzenią produkcyjną mogą być prowadzone w różnych skalach przestrzennych, począwszy od skali pojedynczego pola uprawnego (rolnictwo precyzyjne) po zasięg obejmujący regiony lub większe obszary świata. Pełniejszy opis zjawisk charakteryzujących warunki rolniczej przestrzeni produkcyjnej jest możliwy przy zastosowaniu analiz przestrzennych, które uwzględniają położenie geograficzne obiektów i pozwalają na ocenę ich zmienności oraz wzajemnego powiązania między różnymi zjawiskami. Zmienne losowe, które są powiązane z ich geograficznym zlokalizowaniem nazywamy zmiennymi zregionalizowanymi. Takie zmienne podlegają zwykle tzw. pierwszemu prawu geografii Toblera (1970), które mówi, że *„wszystko jest związane z wszystkim innym, ale rzeczy znajdujące się w pobliżu są bardziej związane niż rzeczy odległe”* (ang. *„everything is related to everything else, but near things are more related than distant things”*). Oznacza to, że różnego rodzaju zjawiska dotyczące przestrzeni geograficznej zmieniają się stopniowo i w miejscach bliskich często obserwowane warunki (np. glebowe lub klimatyczne) są podobne, natomiast wraz z zwiększaniem odległości występują większe różnice. Powiązania te możemy wyrazić w sposób ilościowy wykorzystując metody statystyki przestrzennej, które pozwalają na pełniejszą ocenę zmienności zjawisk opisywanych przez zmienne zregionalizowane (Haining, 2003). Najczęściej metody statystyki przestrzennej są wykorzystywane do:

- scharakteryzowania zmienności zjawisk, m.in. poprzez wykorzystanie wariogramów określających ich zmienność jako funkcję odległości;
- scharakteryzowania powiązań przestrzennych (geograficznych), m.in. poprzez wykorzystanie współczynników autokorelacji;
- modelowania zjawisk przestrzennych, m.in. poprzez wykorzystanie różnych metod interpolacji, w tym stochastycznych metod geostatystycznych takich jak kriging lub co-kriging.

Zastosowanie metod statystyki przestrzennej jest coraz częstsze w wielu dziedzinach życia. W przypadku badań rolniczych jednym z możliwych zastosowań jest wykorzystanie metod analizy przestrzennej do oceny zmienności pól uprawnych, co jest ważne przy wprowadzaniu rolnictwa precyzyjnego. Zastosowanie rolnictwa precyzyjnego w uprawie roślin polega na dostosowaniu technologii uprawy do warunków (głównie glebowych) w poszczególnych częściach pola (ang. *site-specific crop management*) (Robert i in., 1994; Srinivasan, 2006). Najczęściej stosowanymi technikami rolnictwa precyzyjnego są zmienne w obrębie pola uprawnego nawożenie fosforem, potasem, azotem oraz wapnowanie. Ponadto rozwijającymi i przyszłościowymi technikami są zastosowanie zmiennej gęstości siewu oraz zmienne dawki środków ochrony roślin. Wdrażanie rolnictwa precyzyjnego w tym zakresie wymaga dokładnej oceny zmienności właściwości fizyko-chemicznych gleby oraz łanu, co z kolei jest powiązane z gromadzeniem danych przestrzennych w różnej formie, danych punktowych (np. lokalizacje próbek gleby), poligonów (np. obszary obejmujące różne typy gleb) oraz rastrowych (zobrazowania lotnicze lub satelitarne przedstawiające np. zmienność wskaźników wegetacyjnych charakteryzujących stan łanu roślin). Gromadzenie i analiza tego typu danych odbywa się zazwyczaj z wykorzystaniem oprogramowania GIS, czyli systemów informacji geograficznej (ang. *geographic information system*). Metody statystyki przestrzennej pozwalają na ilościową ocenę zmienności przestrzennej właściwości fizyko-chemicznych w obrębie pól uprawnych oraz ich szacowanie z użyciem metod interpolacji przestrzennej w oparciu o dane punktowe. Dzięki temu możliwe jest otrzymanie zgodnych w dużym stopniu z rzeczywistością, map właściwości gleby ważnych z agronomicznego punktu widzenia. Ma to duże znaczenie praktyczne, gdyż pozwala ograniczyć liczbę pobieranych próbek gleby oraz otrzymać wiarygodny obraz zmienności pola, co może być wykorzystane w wdrażaniu rolnictwa precyzyjnego. Metodami statystyki przestrzennej, które mają duże znaczenie w rolnictwie precyzyjnym są również metody detekcji obserwacji odstających w przestrzeni (ang. *spatial outliers*), które często występują w danych z urządzeń pomiarowych wykorzystywanych w rolnictwie precyzyjnym (np. mierników plonu montowanych na kombajnach). Detekcja obserwacji odstających w przestrzeni geograficznej może być przeprowadzona z użyciem lokalnych współczynników autokorelacji, dzięki którym możemy wskazać, które z obserwacji znacznie odbiegają od innych znajdujących się w bliskim ich otoczeniu.

W przypadku danych przestrzennych w skali regionalnej stosuje się podobne metody statystyczne, jak stosowane na potrzeby analiz na niewielkich obszarach w skali pojedynczych gospodarstw. Podstawową różnicą jest to, że zwykle w skali regionalnej interesujemy się innymi typami zjawisk. Może być to np. regionalne zróżnicowanie rolnictwa pod względem wybranych wskaźników ekonomicznych i produkcyjnych. Ze względu na skalę badań obiektami podlegającymi ocenie mogą być nie dane punktowe, lecz np. dane reprezentujące obszary o określonym zasięgu (np. jednostki administracyjne). Wykorzystanie metod statystyki przestrzennej w przypadku tego typu danych pozwala na określenie powiązań między sąsiadującymi obiektami, ocenę występowania zależności przestrzennych pod względem wybranych zmiennych oraz ocenę zależności między różnymi zmiennymi.

Syntetyczne omówienie wyników badań

Przeprowadzone badania składające się na osiągnięcie naukowe dotyczą w dużym stopniu rolnictwa precyzyjnego. **Publikacja nr 1** jest monografią, która opisuje podstawy teoretyczne oraz możliwości wdrażania technik rolnictwa precyzyjnego z uwzględnieniem specyfiki gospodarstw rolnych w Polsce. Monografia ma charakter przeglądu literatury wraz z krytycznymi uwagami dotyczącymi możliwości i ograniczeń wdrażania wybranych rozwiązań rolnictwa precyzyjnego z uwzględnieniem struktury agrarnej gospodarstw w różnych regionach Polski. Jednym z rozdziałów ściśle związanych z metodami statystyki przestrzennej i ich praktycznym wykorzystaniem na potrzeby rolnictwa precyzyjnego, jest rozdział dotyczący oceny zmienności pól uprawnych z użyciem metod tradycyjnych oceny właściwości fizyko-chemicznych gleby. W rozdziale tym zostały opisane metody oceny zmienności wybranych właściwości gleby na podstawie wariogramów (semiwariogramów) (McBratney i Webster, 1986). Powołano się m.in. na badania McBratney'a i Pringle'a (1999), w których została oceniona zmienność przestrzenna gleby pod względem zawartości fosforu dostępnego oraz pH gleby. Zmienność ta była wyrażona w formie graficznej tj. wariogramów teoretycznych, które przedstawiają wariancję ocenianych zmiennych (zawartości fosforu oraz pH) między punktami odległymi o określoną odległość. Wariogramy teoretyczne są funkcjami różnej postaci, np. funkcjami sferycznymi lub wykładniczymi. Funkcje te dopasowane są do danych empirycznych, którymi są wartości semiwariancji dla par punktów odległych od siebie o określoną długość.

Semiwariancja jest natomiast definiowana jako:

$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (z_i - z_{i+h})^2$$

gdzie $\gamma(h)$ jest semiwariancją dla odległości h między parami punktów, N jest liczbą par punktów odległych od siebie o odległość h , z_i oraz z_{i+h} są wartościami zmiennej w pewnym i -tym punkcie oraz w punkcie odległym od niego o odległość h .

Na podstawie parametrów wariogramu opisujących przebieg funkcji, takich jak przecięcie z osią Y czyli tzw. efekt samorodka (ang. *nugget effect*), próg (ang. *sill*) oraz zasięg (ang. *range*) możemy określać zmienność przestrzenną badanego zjawiska oraz występowanie lub brak zależności przestrzennych. Przecięcie wariogramu z osią Y , czyli wartość funkcji dla $h=0$ informuje o zmienności losowej spowodowanej np. błędem pomiaru lub występującej w rzeczywistości zmienności zjawiska. Jest to tzw. efekt samorodka, czyli informacja o różnicy między wartościami badanej zmiennej w punktach bardzo bliskich sobie, czyli np. różnicy między pH gleby dla dwóch próbek gleby pobranych w tym samym punkcie. Wysoka wartość efektu samorodka świadczy o bardzo dużej zmienności przestrzennej atrybutów zjawiska lub też wysokiej zmienności spowodowanej błędem analizy chemicznej (często niemożliwe jest określenie przyczyn tej zmienności). Próg wariogramu (ang. *sill*) jest wartością, która informuje, jaka jest zmienność między wartościami badanej zmiennej w punktach bardzo odległych. Natomiast, zasięg wariogramu (ang. *range*) informuje o odległości występowania zależności przestrzennej (czyli np. propagacji jakiegoś zjawiska). Na podstawie przebiegu funkcji wariogramu teoretycznego możemy optymalizować rozmieszczenie próbek gleby, odległości między nimi, dla zapewnienia możliwie małej ich liczby przy jednoczesnym poprawnym zmapowaniu powierzchni pola uprawnego. W przypadku właściwości fizyko-chemicznych gleby, które charakteryzują się występowaniem silnych zależności przestrzennych na znaczne odległości, możliwe jest ograniczenie gęstości pobrania próbek gleby, natomiast w przypadku występowania słabych zależności przestrzennych (niewielki zasięg wariogramu) konieczne jest zwiększenie liczby pobranych próbek gleby.

Poza aspektami dotyczącymi optymalizacji rozmieszczenie próbek gleby w ww. monografii została podjęta tematyka interpolacji przestrzennej z wykorzystaniem dwóch różnych metod tj. metody odwrotnych odległości (ang. IDW – *inverse distance weighting*)

oraz krigingu, jako rodziny metod przestrzennej interpolacji stochastycznej (Longley i in. 2006). Podana jest główna zaleta interpolacji z użyciem krigingu, tj. uzyskanie, poza wartością szacowaną badanej zmiennej, również oszacowania błędu, co jest ważne dla oceny wiarygodności uzyskanej mapy np. pH gleby. Jest to bardzo istotna informacja, mająca duże praktyczne znaczenie, gdyż określa wiarygodność wyników interpolacji (szacowanych wartości właściwości fizyko-chemicznych gleby) w poszczególnych częściach pola. Ponadto w przypadku interpolacji z użyciem krigingu możliwe jest uwzględnienie anizotropii, tj. różnej zmienności kierunkowej, co często występuje w warunkach pól uprawnych (np. zmienność wzdłuż kierunku: północ-południe może być większa niż ze wschodu na zachód lub odwrotnie). Należy dodać, że kriging jest nazwą dla pewnej grupy metod interpolacji geostatystycznej a nie jedną metodą. Najpowszechniej używanym typem krigingu, jest kriging zwyczajny (ang. *ordinary kriging*), jednak występują m.in. takie typy krigingu jak: kriging uniwersalny (ang. *universal kriging*), kriging prosty (ang. *simple kriging*) oraz kriging wskaźnikowy (ang. *indicator kriging*). Najczęściej stosowany jest jednak kriging zwyczajny lub też jego modyfikacja uwzględniająca przy interpolacji zmienne towarzyszące (ang. *covariates*), które możemy wykorzystać do poprawienia jakości szacowania przestrzennego. Właściwy dobór metod interpolacji jest szczególnie ważny w rolnictwie precyzyjnym, gdyż stosując różne metody interpolacji możemy na podstawie tych samych danych uzyskać znacząco różniące się wyniki w postaci szacowanych powierzchni badanej zmiennej. Rozwinięcie problemu interpolacji danych przestrzennych w skali pól uprawnych znajduje się w publikacji nr 6.

W **publikacji nr 2** składającej się na osiągnięcie podejmowana jest tematyka dotycząca detekcji obserwacji (przypadków, jednostek) odstających w przestrzeni (ang. *spatial outliers*) (Dobermann i in., 2003). Badania przedstawione w artykule zostały przeprowadzone w skali pojedynczego pola uprawnego. Metodami użytymi do detekcji obserwacji odstających były: metoda oparta na rozkładzie zmiennej losowej, który został przedstawiony w formie histogramu oraz metoda uwzględniająca położenie geograficzne badanych obiektów, lokalny współczynnik autokorelacji przestrzennej I Morana wg następującego wzoru (Mitchell, 2005).

$$I = \frac{N \sum_i \sum_j w_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{(\sum_i \sum_j w_{i,j}) \sum_i (x_i - \bar{x})^2} \quad (2.1)$$

gdzie: N – liczebność próby, x_i – wartość zmiennej dla i -tej lokalizacji, x_j – wartość zmiennej dla j -tej lokalizacji, \bar{x} – średnia wartość zmiennej, w_{ij} – waga wynikająca z położenia lokalizacji i względem lokalizacji j . Wagi te dla par obiektów są odwrotnościami odległości euklidesowych między tymi obiektami.

Interpretacja wartości I Morana dla dużych liczebności analizowanych obserwacji jest podobna do współczynnika korelacji prostej, dodatnie wartości oznaczają występowanie obok siebie obiektów podobnych pod względem wartości badanej zmiennej, natomiast wartości ujemne oznaczają występowanie obiektów znacznie różniących się pod względem wartości badanej zmiennej. Wartość I bliska 0 oznacza losowe rozmieszczenie obiektów (Anselin, 1995). Testowanie istotności współczynnika autokorelacji wykonano na podstawie wartości empirycznej Z , która ma rozkład normalny standardowy $Z \sim N(0; 1)$.

$$Z(I) = \frac{I - E(I)}{S_{E(I)}} \quad (2.2)$$

gdzie $E(I)$ jest wartością oczekiwaną współczynnika autokorelacji I Morana

$$E(I) = \frac{-1}{N-1} \quad (2.3)$$

a $S_{E(I)}$ jest jego odchyleniem standardowym, liczonym według wzoru:

$$S_{E(I)} = \sqrt{\frac{N^2 \sum_{ij} w_{ij}^2 + 3(\sum_{ij} w_{ij})^2 - N \sum_i (\sum_j w_{ij})^2}{(N^2 - 1)(\sum_{ij} w_{ij})^2}}$$

Badaną zmienną była wielkość plonu pszenicy ozimej, która została zarejestrowana z użyciem miernika plonu zamontowanego na kombajnie zbożowym współpracującego z odbiornikiem GPS. Urządzenia takie szacują wielkość plonu w czasie zbioru na podstawie ilości ziarna przemieszczającego się w danym momencie w przenośnikach ziarnowych. Duża niedoskonałość oceny powoduje, że występuje wiele obserwacji odstających, które znacząco odbiegają od wielkości plonu rzeczywistego. Wynika to z wielu czynników, a w szczególności zbioru ziarna przy wykorzystaniu niepełnej szerokości roboczej, wykonywania nawrotów na polu, błędów pomiaru miernika plonu (Arslan and Colvin, 2002). W związku z tym, że nie zawsze istnieje możliwość jednoznacznego wskazania, które z pomiarów są

zgodne z rzeczywistymi plonami, a które z nich silnie odbiegają od rzeczywistych plonów, to takie obserwacje odstające mogą powodować otrzymanie map plonów, które nie odpowiadają faktycznej zmienności plonowania w obrębie pola. Liczba punktów z zarejestrowanymi plonami jest zwykle bardzo duża, zwykle od kilkudziesięciu do kilkuset na hektar powierzchni pola. W przypadku dużych pól, kilkunasto-kilkudziesięcio-hektarowych liczba punktów z zarejestrowanymi plonami wynosi od kilku do kilkunastu tysięcy. Przy tak dużej liczbie punktów nie jest możliwe analizowanie poszczególnych punktów. Oczyszczanie danych z obserwacji odstających odbywa się zwykle metodami pozwalającymi na wskazanie takich obserwacji w całym zbiorze jednocześnie. W przeprowadzonych badaniach stwierdzono, że wykorzystanie metody opartej jedynie na rozkładzie (przedstawionym w formie histogramu) dla całego pola jednocześnie, nie wykrywa obserwacji, które są odstające od sąsiadujących, natomiast nie są odstające dla całego rozkładu wartości. W związku z tym zaproponowano wykorzystanie lokalnych współczynników autokorelacji przestrzennej I Morana. Jako obserwacje odstające przyjęto takie, których współczynnik autokorelacji przestrzennej był ujemny, czyli charakteryzujące się znacznie innymi (mniejszymi lub większymi) wielkościami plonu w porównaniu z sąsiednimi punktami. Podejście takie jednak nie powodowało usunięcia obserwacji odstających, które zostały zarejestrowane w bardzo bliskim sąsiedztwie do innych obserwacji odstających, co często zdarzało się podczas nawrotów. Zatem, zaproponowano inne podejście, w którym jako obserwacje odstające potraktowano zarówno te punkty pomiarowe, dla których stwierdzono ujemne lokalne współczynniki autokorelacji przestrzennej, jak i te, dla których te współczynniki były bardzo wysokie (jako poziom graniczny przyjęto wartość 3 – należy tu wspomnieć, że wartość współczynnika autokorelacji I Morana może przyjmować wartości mniejsze niż -1 i większe niż 1). Takie podejście wydaje się obiecujące, gdyż usuwa większość obserwacji odstających, zarówno tych pojedynczych, otoczonych obserwacjami, które można uznać za typowe oraz odstających zgrupowanych w grupy punktów (np. bardzo niskich plonów obok siebie). We wnioskach ww. pracy uznano, że komplementarne podejście z wykorzystaniem klasycznych metod statystycznych oraz metod statystyki przestrzennej z wykorzystaniem współczynnika autokorelacji I Morana powinno pozwolić na otrzymanie bardziej wiarygodnych map plonu, charakteryzujących się wysoką zgodnością z rzeczywistymi plonami.

Publikacja nr 3 dotyczy podobnej tematyki do publikacji nr 2, gdyż głównym celem jest wskazanie obserwacji nietypowych w przestrzeni. Głównymi różnicami jest jednak znacznie

inna skala przestrzenna, obejmująca w przypadku publikacji nr 3 region obejmujący województwo Podlaskie. Podstawową jednostką obserwacyjną w tym przypadku jest obszar obejmujący jedną gminę. Ponadto ważną różnicą jest to, że analizowano jednocześnie 8 zmiennych, a nie tak jak w przypadku publikacji nr 2, jedną zmienną. Zmiennymi były wybrane wskaźniki charakteryzujące produkcję rolniczą i jej uwarunkowania w poszczególnych gminach, takie jak m.in. średnia wielkość gospodarstwa rolnego, obsada bydła, wartość produkcji rolniczej w przeliczeniu na 1 ha itp.. Celem badań było wskazanie obserwacji nietypowych pod względem wielu zmiennych jednocześnie. W odróżnieniu od publikacji nr 2, zmienne uwzględniane w analizach były oceniane na podstawie Powszechnego Spisu Rolnego przeprowadzonego w roku 2002, który obejmował wszystkie gospodarstwa rolne i cały obszar gminy, można zatem uznać, że nie są to wartości szacowane a określone bez błędu lub też z niewielkim błędem, co odróżnia również te dane od danych z publikacji nr 2.

Dla umożliwienia zastosowania metod statystyki przestrzennej, 8 zmiennych uwzględnianych w analizie poddano analizie składowych głównych (ang. *PCA – principal component analysis*) (Filipiak i Wilkos, 1998). Pierwsza składowa główna wyjaśniała blisko 50% zmienności całego zbioru danych, dalsze analizy przeprowadzono z użyciem lokalnego współczynnika autokorelacji *I* Morana. Jako wskaźnik obserwacji nietypowej, tj. gminy odbiegającej od sąsiednich pod względem uwzględnianych w analizie zmiennych, przyjęto ujemny współczynnik autokorelacji, a dokładniej wartość statystyki *Z* (*Z-score*). Im niższa wartość tego parametru, tym bardziej dana gmina różni się od gmin z nią sąsiadujących. W przypadku badanego regionu stwierdzono ujemne wartości statystyki *Z* dla 9 gmin, z czego dla dwóch gmin (Sejny i Rudka) wartości te były najniższe (poniżej -0,2), co oznacza, że gminy te najbardziej odbiegają od sąsiednich pod względem struktury agrarnej gospodarstw i innych zmiennych charakteryzujących warunki naturalne i socjoekonomiczne rolnictwa. Metoda detekcji obserwacji nietypowych wśród jednostek administracyjnych samorządu terytorialnego może mieć zastosowanie w ocenie równomierności rozwoju gospodarczego pod względem wybranych wskaźników. Umożliwia wskazanie jednostek nietypowych, czyli zarówno znacznie silniej, jak i znacznie słabiej rozwiniętych pod pewnymi względami, co może posłużyć do kształtowania polityki mającej na celu zrównoważony rozwój regionów. Zastosowane metody mogą być wykorzystywane w bardzo różnej skali przestrzennej, np. w skali województw, tak jak to miało zastosowanie w omawianej pracy, ale również w skali

kontynentu, gdzie celem może być wskazanie odbiegających krajów pod względem wybranych zmiennych. W przypadku, gdy pierwsza składowa główna nie wyjaśnia wystarczająco dużo zmienności badanego zbioru danych, możliwe jest rozwinięcie tej metody, poprzez dwuetapową analizę polegającą np. na oddzielnej analizie pierwszej i drugiej składowej głównej (PC1 i PC2), a następnie detekcję obserwacji nietypowych z wykorzystaniem sumy dwóch współczynników autokorelacji przestrzennej. Należy dodać, że lokalny współczynnik autokorelacji przestrzennej *I* Morana jest jednym z tego typu współczynników. Cała ich grupa jest określana skrótem LISA (ang. *local indicators of spatial autocorrelation*) (Anselin, 1995). Innym lokalnym współczynnikiem autokorelacji przestrzennej jest m.in. współczynnik *C* Geary'ego oraz statystyka Getis-Ord (Ord i Getis, 2001). Warunkiem niezbędnym do zastosowania tych metod jest występowanie dodatniej autokorelacji globalnej, jest on jednak bardzo często spełniony, ze względu na charakter badanych zjawisk.

Publikacja nr 4 przedstawia porównanie różnych wariantów pobrania próbek gleby na obszarze pojedynczego pola uprawnego o powierzchni około 45 ha. Porównywano pobranie próbek gleby w różnych rozdzielczościach przestrzennych, tj. w siatce dzielącej całe pole na powierzchnie 1-hektarowe, 2-hektarowe oraz 4-hektarowe. Celem pracy było określenie na ile pobieranie próbek gleby z większych obszarów (siatka 2 lub 4 ha), powoduje niezgodność w stosunku do próbobrania w siatce 1 ha, traktowanej jako referencyjna. Jest to ważne ze względu na optymalizację rozmieszczenia próbek gleby na potrzeby mapowania przy wdrażaniu rolnictwa precyzyjnego (Mallarino i Wittry, 2004). Badanymi zmiennymi były: zasobność gleby w dostępny fosfor, potas i magnez oraz pH gleby, jak również związane z tą zasobnością zalecenia nawożenia tymi pierwiastkami i zalecenia dotyczące wapnowania. Ponadto porównano również wyniki uzyskane dla siatki 1 ha z potraktowaniem całego pola, jako jednorodnego obszaru gleby. Oceniono, w jakim stopniu uzyskane wyniki z siatki 1 ha różnią się w stosunku do wyników otrzymanych z siatek o większych powierzchniach komórek. W analizach wykorzystano m.in. ważone estymatory parametrów statystycznych, takich jak średnie i odchylenia standardowe (Gatz i Smith, 1995). Zastosowanie wag dla poszczególnych obserwacji wynikało z tego, iż wyznaczone siatki dla różnych gęstości próbobrania gleby nie miały idealnie takich samych rozmiarów komórek. Na przykład w siatce o powierzchni komórki 1 ha występowały komórki w pewnym stopniu odbiegające od tej powierzchni (mniejsze lub większe niż 1 ha). Podobnie było dla siatek 2 ha i 4 ha, gdzie

powierzchnia poszczególnych komórek była często nieco mniejsza lub nieco większa niż zakładana. W obliczaniu estymatorów przyjęto zatem jako wagi, powierzchnie poszczególnych komórek. Do obliczeń wykorzystano funkcje *wtd.** z pakiet *Hmisc* działającego w środowisku R (Harrell, 2011).

Na podstawie porównania wyników zasobności i pH gleby dla poszczególnych komórek siatki 2 ha i 4 ha z siatką 1 ha określono średnie błędy, ich zakres oraz odchylenie standardowe. Błędy były definiowane jako różnice między wartością zasobności dla siatki 2 ha lub 4 ha względem obserwowanych wartości dla siatki 1 ha. Na podstawie uzyskanych wyników stwierdzono podobne wielkości błędów dla obu siatek, tj. siatki 2 ha i 4 ha dla zasobności w P, K i Mg. W przypadku pH gleby uzyskano znacznie lepszą zgodność z siatką 1 ha dla siatki 2 ha niż dla siatki 4 ha. Było to odpowiednio 74% i 55% komórek odpowiednio siatki 2 ha i 4 ha. Wskazuje to na konieczność zastosowania możliwie gęstego próbobrania w przypadku oceny pH gleby dla uzyskania mapy w dużym stopniu zgodnej z rzeczywistym pH. Jest to ważna informacja przy wdrażaniu technologii rolnictwa precyzyjnego, a w tym przypadku dotyczy stosowania zmiennej dawki wapnowania, dostosowanej do pH gleby w poszczególnych punktach pola. Stosowanie siatki 2 ha a w szczególności 4 ha, powodowałoby przy stosowaniu zmiennego wapnowania zbyt niskie lub zbyt wysokie wapnowanie na dość dużej powierzchni pola (dla siatki 2 ha – ok. 40% a dla siatki 4 ha – ok. 50%), przy założeniu, że siatka 1 ha niesie informacje o zmienności pH zgodne z rzeczywistością. W związku z tym, że badania były przeprowadzone tylko na jednym polu uprawnym, mimo jego dużej powierzchni, nie można uznać, że są one reprezentatywne dla wszystkich typów gleb, a nawet dla gleb o tym samym typie, gdyż uzyskane wyniki mogą być zależne od wielu czynników, w tym od stosowanej technologii uprawy w ostatnich kilkunastu latach, która wpływa na zasobność i pH gleby. Zatem, konieczne jest w przyszłości rozszerzenie tego typu badań na różne typy gleb, w różnych regionach Polski.

Publikacja nr 5 dotyczy problem wydzielenia stref produkcyjnych (ang. *management zones*) wydzielanych w obrębie pól uprawnych na potrzeby rolnictwa precyzyjnego, czyli na potrzeby dostosowania technologii uprawy (np. dawki nawożenia, gęstości siewu itp.) w poszczególnych strefach pola (Khosla i in., 2010). Głównym celem artykułu jest wskazanie spośród wybranych zmiennych charakteryzujących właściwości fizyko-chemiczne gleby i warunki topograficzne, tych spośród nich, które w największym stopniu wpływają na potencjał plonowania. Jest to ważne, gdyż nie wszystkie właściwości fizyko-chemiczne gleby

można uznać za tak samo ważne w takim procesie (Fleming i in., 2000; Mzuku i in., 2005) Badania przeprowadzono na dwóch polach (oznaczonych jako D2 i D5) o powierzchni 22 i 45 ha. Potencjał plonowania określono na podstawie wielkości plonu zarejestrowanego z użyciem mierników plonu współpracujących z GPS w sześciu latach poprzedzających badanie. W związku z tym, że plony były rejestrowane w różnych latach, a tym samym w różnych warunkach pogodowych, dla dwóch różnych gatunków roślin uprawnych (rzepaku ozimego i pszenicy ozimej), to w analizach wykorzystano plony standaryzowane. Na podstawie średnich standaryzowanych plonów wydzielono dwie strefy w obrębie każdego z dwóch badanych pól uprawnych. Jedną strefę stanowiły obszary pola, dla których stwierdzono średni standaryzowany plon większy, niż 0 (czyli plony powyżej średniej) a drugą strefę obszary pola, dla których średni standaryzowany plon był mniejszy od 0 (czyli mniejszy od średniej). Główną zastosowaną metodą statystyczną była analiza regresji logistycznej wielokrotnej (Hosmer i Lemeshow, 2004), gdzie zmiennymi przyczynowymi były m.in. zawartość frakcji piasku i łu, zawartość węgla organicznego, nachylenie terenu, pH i przewodność elektryczna gleby. Na podstawie przeprowadzonych analiz stwierdzono, że w przypadku obu badanych pól uprawnych zmiennymi istotnie determinującymi wydzielone strefy o różnym potencjale plonowania były: zawartość frakcji piasku oraz zawartość węgla organicznego. Pozostałe zmienne wpływały istotnie jedynie w przypadku jednego pola spośród dwóch badanych lub też nie stwierdzono ich statystycznie istotnego wpływu. W opisywanym artykule oceniono również zmienność badanych cech oraz występowanie powiązań przestrzennych z użyciem globalnego współczynnika *I* Morana.

Globalny współczynnik *I* Morana (Moran 1950) jest parametrem statystycznym pozwalającym na określenie występowania (lub braku) autokorelacji przestrzennej, czyli występowania (lub nie występowania) powiązań między zmiennymi ilościowymi w bliskich geograficznie punktach. Jest on obliczany według następującego równania:

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$

gdzie \bar{x} jest średnią badanej zmiennej x , w_{ij} są wagami dla par obserwacji i oraz j ; S_0 jest sumą wag: $S_0 = \sum_i \sum_j w_{ij}$. Wagi dla par obserwacji i oraz j są określane na podstawie odległości między tymi parami obserwacji, czym mniejsza odległość, tym większa waga.

Interpretacja globalnego współczynnika I Morana jest podobna, jednak nie taka sama, jak współczynnika korelacji (np. Pearsona). Ujemny współczynnik autokorelacji oznacza, że sąsiednie obserwacje bardzo się różnią (taka sytuacja bardzo rzadko występuje w rzeczywistości). Dodatni współczynnik oznacza występowanie dodatniej autokorelacji przestrzennej, czyli występowanie powiązania między sąsiadującymi obserwacjami. W związku z tym, że oczekiwana wartość współczynnika I Morana jest mniejsza od zera: $E(I) = -1/n - 1$, to przy niewielkiej liczbie obserwacji, nawet współczynnik Morana bliski 0 może wskazywać na występowanie dodatniej autokorelacji przestrzennej.

Największą zmienność, wyrażoną jako współczynnik zmienności, stwierdzono dla nachylenia terenu w przypadku obu badanych pól. Spośród właściwości gleby, dla obu pól dużą zmienność zaobserwowano dla zawartości ilu oraz zawartości węgla organicznego. Statystycznie istotne zależności przestrzenne wyrażone w postaci współczynnika autokorelacji przestrzennej I Morana stwierdzono dla obu pól dla węgla organicznego i pH gleby, co świadczy o tym, że sąsiadujące punkty miały zbliżone wartości tych zmiennych, a więc ich zmiana w obrębie pola następuje stopniowo. Dość silne zależności stwierdzono również dla wielkości plonu, czyli miejsca o niskich plonach sąsiadowały ze sobą (podobnie było dla miejsc o wyższych plonach), czyli plony w sąsiadujących punktach były zbliżone. Siła zależności przestrzennych większości badanych zmiennych była znacznie większa w przypadku pola D2, niż w przypadku pola D5. Świadczy to o dużym zróżnicowaniu warunków glebowych w przypadku pola D5.

Uzyskane wyniki badań umożliwiły wskazanie właściwości fizyko-chemicznych gleby, które można uznać za najważniejsze przy wydzielaniu stref produkcyjnych na potrzeby rolnictwa precyzyjnego.

Publikacja nr 6 dotyczy porównania różnych metod interpolacji przestrzennej na potrzeby szacowania frakcji gleby (piasku, pyłu i ilu) w obrębie pojedynczych pól uprawnych. Problem interpolacji przestrzennej frakcji gleby jest szczególnie ważny, ze względu na czasochłonność a tym samym wysoki koszty wykonania klasycznej oceny laboratoryjnej, a tym samym należy ograniczyć liczbę analizowanych próbek gleby do

minimum. Ważne jest zatem, aby dobór metody interpolacji do tego typu danych był poprawny, pozwalający na uzyskanie szacowanej zawartości poszczególnych frakcji, jak najbardziej zgodnych z rzeczywistością. Ponadto uziarnienie gleby jest szczególnie ważną właściwością, którą powinno się uwzględniać w rolnictwie precyzyjnym (He et al., 2013; Dec i Dörner, 2014). Badania przeprowadzono z wykorzystaniem danych dla czterech pól o powierzchni od 10 do 45 ha, położonych na Mazowszu oraz na Pomorzu. Dane wejściowe stanowiły zawartości frakcji gleby dla punktów zlokalizowanych w obrębie badanych pól. Liczba punktów wynosiła od 21 do 60, w zależności od pola. Na potrzeby oceny metod interpolacji zbiory danych zostały podzielone na zbiór treningowy – około 70% obserwacji (służący do interpolacji) oraz zbiór testowy – około 30% obserwacji (służący do walidacji metody). Ocenie poddano następujące metody interpolacji: metoda odwrotnych odległości (IDW – *inverse distance weighting*), radialnej funkcji bazowej (RBF - *radial basis function*), krigingu zwyczajnego (*ordinary kriging*) oraz cokrigingu zwyczajnego (*ordinary cokriging*) (Robinson i Metternicht, 2006; Li i Heap, 2011). W przypadku trzech pierwszych wymienionych metod interpolacji w analizie uwzględniane są tylko zmienne interpolowane, natomiast metoda czwarta – cokriging jest metodą uwzględniającą zmienne towarzyszące (ang. *covariates*) (Ersahin, 2003). W przypadku tych badań, jako zmienną towarzyszącą uwzględniono przewodność elektryczną gleby (EC – *electrical conductivity*), która w dużym stopniu jest związana z uziarnieniem gleby, czyli zawartością poszczególnych frakcji glebowych (piasku, pyłu i iłu) (Corwin i Lesch, 2005). Walidację metod interpolacji przeprowadzono z użyciem dwóch różnych podejść. Pierwszym podejściem było zastosowanie metody LOOCV (ang. *leave-one-out crossvalidation*), gdzie przeprowadzana jest interpolacja z użyciem całego zbioru danych z wyłączeniem jednej obserwacji a następnie określone są wartości błędów między wartością szacowaną na podstawie interpolacji a obserwacją, która była wyłączona z analizy (Refaeilzadeh i in., 2009). Procedura jest powtarzana tyle razy, ile wynosi liczba obserwacji. Drugie podejście było przeprowadzone w oparciu o podział zbioru danych na zbiór treningowy i testowy. W tym celu losowo wybrano około 70% obserwacji, na których przeprowadzono interpolację z użyciem z każdej z metod podlegających ocenie. Walidacja, czyli ocena błędów między wartościami szacowanymi z użyciem interpolacji a wartościami ze zbioru testowego była prowadzona na około 30% obserwacji. Takie procedury walidacyjne były przeprowadzone dla każdej z metod interpolacji, dla każdej z badanych frakcji gleby, czyli piasku, pyłu i iłu. Wartości błędów

średniokwadratowych (RMSE – *root mean square error*) posłużyły do określenia przydatności poszczególnych metod interpolacji. Mniejsza wartość RMSE oznacza lepszą zgodność szacowanych wartości z obserwowanymi (Mouder i Algina, 2002). Na podstawie uzyskanych wyników stwierdzono, że najlepszą metodą interpolacji w przypadku wszystkich badanych pól uprawnych i ocenianych frakcji gleby był *cokriging*, czyli metoda interpolacji uwzględniająca zmienną towarzyszącą, którą była przewodność elektryczna gleby. Wartość RMSE była przeciętnie niższa o około 10% w porównaniu z pozostałymi metodami interpolacji. Pozostałe trzy metody interpolacji (IDW, RBF i kriging zwyczajny) dawały zbliżone wartości błędów, co wskazuje na podobną ich przydatność do interpolacji uziarnienia gleby w skali pojedynczych pól uprawnych. Podobne wyniki uzyskał również Karydas et. al. (2009) w badaniach na terenie Grecji.

Poza oceną metod interpolacji, w publikacji przedstawiono również zmienność przestrzenną badanych zmiennych (frakcji gleby). W tym celu obliczono parametry wariogramów teoretycznych dla każdej z frakcji gleby, oddzielnie dla każdego z pól oraz przedstawiono te wariogramy dla wybranej frakcji (zawartości pyłu). We wszystkich analizach wykorzystano funkcję sferyczną, jako dobrze opisującą charakter ocenianych zależności przestrzennych. Ponadto obliczono globalne współczynniki autokorelacji przestrzennej *I* Morana dla każdej z frakcji dla poszczególnych badanych pól uprawnych. Na podstawie wyników tych analiz stwierdzono występowanie statystycznie istotnych dodatnich autokorelacji dla frakcji piasku (dla trzech spośród czterech badanych pól) oraz dla frakcji pyłu (dla dwóch pól). W przypadku ilu zależności przestrzenne były słabsze, gdyż jedynie dla jednego pola stwierdzono statystycznie istotną autokorelację przestrzenną. Uzyskane wyniki świadczą o tym, że dla uzyskania map glebowych przedstawiających uziarnienie wskazane jest wykorzystanie zmiennych towarzyszących, które są skorelowane z zawartością różnych frakcji gleby. Wykorzystanie do tego celu przewodności elektrycznej jest rozwiązaniem, które może być w przyszłości wykorzystywane w praktyce, gdyż coraz powszechniej są oferowane usługi mapowania przewodności elektrycznej pól uprawnych na potrzeby rolnictwa precyzyjnego. Łączne wykorzystanie danych z klasycznych analiz laboratoryjnych (np. zawartości frakcji glebowych) wraz z danymi uzyskanymi z mapowania z urządzeń służących do szybkiej oceny (np. pomiar przewodności elektrycznej gleby) przy zastosowaniu interpolacji z użyciem *cokrigingu* powinno pozwolić na uzyskanie dokładniejszych map

zmienności gleby. Jest to szczególnie ważne dla uziarnienia gleby, gdyż jest to właściwość gleby ważna z agronomicznego punktu widzenia, a jednocześnie mało zmienna w czasie.

Publikacja nr 7 przedstawia ocenę zależności między różnymi właściwościami gleby a zmiennymi uzyskanymi z satelitarnego sensora MODIS dotyczącymi wybranych kanałów światła i wskaźników wegetacyjnych. Właściwości gleby były oceniane w ramach ogólnoeuropejskiego projektu LUCAS (skrót od ang. *Land Use/Cover Area frame Statistical Survey*), koordynowanego przez instytucje badawcze Komisji Europejskiej. Dane te są publicznie dostępne i dotyczą m.in. następujących właściwości gleby: zawartość frakcji gleby (piasek, pył, il oraz części szkieletowe), zawartość węgla organicznego, azotu całkowitego, zawartości fosforu i potasu wymiennego (Carre i in., 2013). Do analiz wykorzystano 967 obserwacji dla próbek gleby położonych na obszarze Polski na terenie użytków rolnych (obszary zurbanizowane, lasy inne typy użytków wykluczono z analiz). Źródłem danych satelitarnych były średnie miesięczne wartości wybranych kanałów i wskaźników wegetacyjnych z satelity MODIS (Gu i in., 2008). Wybrano dane dla czerwca roku 2013, czyli w okresie intensywnej wegetacji roślin. Uwzględniono w analizach kanały w zakresie czerwieni i koloru niebieskiego oraz podczerwień (średnia i bliska) oraz wskaźniki wegetacyjne NDVI (skrót od ang. *Normalized Difference Vegetation Index* - znormalizowany różnicowy wskaźnik wegetacji) i EVI (skrót od ang. *Enhanced Vegetation Index*) (Huete i in., 2002). Indeksy wegetacyjne są powiązane ze stanem roślinności, a w szczególności z pokryciem gleby roślinnością i zielonością liści (Haboudane i in. 2004). Wartości zmiennych uzyskanych z danych satelitarnych zostały odczytane z punktów pobrania próbek gleby a następnie na tak sparowanych danych zostały przeprowadzone analizy korelacji i regresji mające na celu określenie występowania związków między właściwościami gleby a danymi satelitarnymi. Na podstawie współczynników korelacji stwierdzono umiarkowanie silną, dodatnią korelację między zawartością frakcji pyłu i łu w glebie a wskaźnikami wegetacyjnymi EVI i NDVI oraz bliską podczerwienią (współczynniki korelacji w zakresie od 0,25 do 0,44). Jednocześnie stwierdzono ujemną korelację między EVI, NDVI i bliską podczerwienią a zawartością piasku (współczynniki korelacji w zakresie od -0,46 do -0,27). Pozostałe zależności były znacznie słabsze, jednak należy zwrócić uwagę również na dodanie, statystycznie istotne korelacje między zawartością potasu i azotu w glebie a wskaźnikami wegetacyjnymi EVI i NDVI. Poza współczynnikami korelacji przedstawiono graficznie i w

formie równań regresji liniowej wybrane zależności, tj. między zawartością piasku a EVI oraz między zawartością pyłu a EVI.

Poza określeniem zależności między wybranymi właściwościami gleby a danymi satelitarnymi przeprowadzono również analizy oceniające zmienność badanych cech i występowanie zależności przestrzennych. Stwierdzono, że większość właściwości fizykochemicznych gleby charakteryzowała się dużą lub bardzo dużą zmiennością. W przypadku danych satelitarnych współczynniki zmienności były znacznie mniejsze, zwłaszcza dla indeksów wegetacyjnych NDVI i EVI. Analizy autokorelacji przestrzennej z użyciem współczynnika *I* Morana (Moran, 1950) wykazały występowanie istotnych statystycznie autokorelacji dla wszystkich frakcji gleby (piasku, pyłu i ilu) oraz pH gleby. Również statystycznie istotne dodatnie autokorelacje stwierdzono dla wszystkich zmiennych pozyskanych z danych satelitarnych. Oznacza to, że zależności przestrzenne dla tych zmiennych, dla których takie zależności przestrzenne występują, są obserwowane na bardzo dużych odległościach, gdyż zasięg badanego obszaru obejmował całą Polskę, czyli odległości były rzędu setek kilometrów. Na podstawie uzyskanych wyników możemy stwierdzić, że zmienne uzyskane na podstawie danych satelitarnych mogą w pewnym stopniu być przydatne do poprawy predykcji niektórych właściwości gleby z użyciem np. cokrigingu.

Podsumowanie

Zastosowanie, proponowanych w moim osiągnięciu naukowym, metod statystyki przestrzennej w badaniach rolniczej przestrzeni produkcyjnej pozwala na pełniejszy opis zmienności geograficznej warunków gospodarowania w rolnictwie. W szczególności dotyczy to warunków glebowych, których zmienność oceniana może być począwszy od skali pola po skalę regionalną. Wiele z metod statystyki przestrzennej, takich jak różne metody interpolacji, zwłaszcza te uwzględniające zmienne towarzyszące, metody detekcji obserwacji odstających oraz metody oceniające zmienność zjawisk (np. wariogramy), ma szczególnie duże znaczenie w rolnictwie precyzyjnym. Rolnictwo precyzyjne wymaga wszechstronnej analizy danych, w tym modelowania zjawisk zachodzących w skali pojedynczych pól uprawnych. Metody statystyki przestrzennej są tu niezbędne zarówno na etapie badań, jak i przy tworzeniu map do celów praktycznych, na potrzeby wdrażania technologii rolnictwa precyzyjnego. Cykl publikacji składający się na osiągnięcie naukowe przedstawia możliwości zastosowania wybranych metod statystyki przestrzennej oraz wnioskowanie na ich podstawie. Szczególnie

istotne są zastosowania metod, pozwalających na ocenę autokorelacji (lokalnej i globalnej) oraz te uwzględniające powiązania między różnymi zmiennymi. Wyniki uzyskanych badań w pewnym stopniu mogą mieć praktyczne zastosowanie we wdrażaniu rozwiązań rolnictwa precyzyjnego (badania w skali pojedynczych pól uprawnych) oraz w kształtowaniu zróżnicowanego rozwoju regionalnego (badania w skali województwa lub kraju).

Literatura

1. Anselin L. (1995). Local indicators of spatial association LISA. *Geographical Analysis* 27, 93–115.
2. Arslan S., Colvin T. (2002). Grain yield mapping: yield sensing, yield reconstruction, and errors. *Precision Agriculture* 3, 135–154.
3. Carre F., Rusco E., Tóth G., Jones A., Gardi C., Stolbovoy V. (2013). 2. Soil sampling methodology. In: Tóth G., Jones A. and Montanarella L.(eds.) *LUCAS Topsoil Survey methodology, data and results*. Joint Research Centre, Report EUR 26102.
4. Corwin D.L., Lesch S.M. (2005). Apparent soil electrical conductivity measurements in agriculture. *Computers and Electronics in Agriculture* 46, 11-43.
5. Dec D., Dörner J. 2014. Spatial variability of the hydraulic properties of a drip irrigated andisol under blueberries. *Journal of Soil Science and Plant Nutrition* 14, 589-601.
6. Dobermann A., Ping J.L., Adamchuk V.I., Simbahan G.C., Ferguson R.B. (2003). Classification of crop yield variability in irrigated production fields. *Agronomy Journal* 95, 1105–1120.
7. Ersahin S. (2003). Comparing ordinary kriging and cokriging to estimate infiltration rate. *Soil Science Society of America Journal* 67, 1848-1855.
8. Filipiak K., Wilkos S., (1998). *Wybrane metody analizy wielozmiennej i ich zastosowanie w badaniach przestrzennych*. IUNG Puławy, R (349)
9. Fleming K.L., Westfall D.G., Wiens D.W., Brodahl M.C. (2000). Evaluating farmer defined management zone maps for variable-rate fertilizer application. *Precision Agriculture* 2, 201–215.
10. Gatz D.F., Smith L. (1995). The standard error of a weighted mean concentration—I. Bootstrapping vs other methods. *Atmospheric Environment* 11, 1185-1193.
11. Gu Y., Hunt E., Wardlow B., Basara J.B., Brown J.F., Verdin J.P. (2008): Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. *Geophysical Research Letters* 35, L22401.
12. Haboudane D., Miller J.R., Pattey E., Zarco-Tejada P.J., Strachan I.B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment* 90, 337-352.
13. Haining R. (1993). *Spatial data analysis in the social and environmental sciences*. Cambridge University Press.

14. Harrell F. E. (2011). *Hmisc: Harrell Miscellaneous*, <http://CRAN.R-project.org/package=Hmisc>. R package.
15. He Y., Wei Y., DePauw R., Qian B., Lemke R., Singh A., Cuthbert R., McConkey B., Wang H. (2013). Spring wheat yield in the semiarid Canadian prairies: Effects of precipitation timing and soil texture over recent 30 years. *Field Crops Research* 149, 329–337.
16. Hosmer Jr D.W., Lemeshow S. (2004). *Applied logistic regression*. John Wiley & Sons.
17. Huete A., Didan K., Miura T., Rodriguez E.P., Gao X., Ferreira L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83, 195-213.
18. Karydas C.G., Gitas I.Z., Koutsogiannaki E., Lydakis-Simantiris N., Silleos G.N. (2009). Evaluation of spatial interpolation techniques for mapping agricultural topsoil properties in Crete. *EARSeL eProceedings*. 8, 26–39.
19. Khosla R., Westfall D.G., Reich R.M., Mahal J.S., Gangloff W.J. (2010). Spatial variation and site-specific management zones. In: *Geostatistical applications for precision agriculture*, ed. M. A. Oliver, 195–219. Dordrecht, the Netherlands: Springer.
20. Li J., Heap A.D. (2011). A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. *Ecological Informatics* 6, 228-241.
21. Longley P., Goodchild M., Maguire D., Rhind D. (2006). *GIS – Teoria i praktyka*. PWN Warszawa, ss. 519
22. Mallarino A.P., Wittry D.J. (2004): Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH and organic matter. *Precision Agriculture* 5, 131–134
23. McBratney A.B., Webster R. (1986). Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *Journal of Soil Science* 37, 617-639.
24. McBratney A., Pringle M. (1999). Estimating average and proportional variograms of soil properties and their potential use in precision agriculture, *Precision Agriculture* 1, 125-152
25. McBratney A., Whelan B., Ancev T., Bouma J. (2005). Future directions of precision agriculture. *Precision Agriculture*, 6, 7–23.
26. Mitchell A. (2005). *The ESRI guide to GIS analysis. Vol.2 Spatial measurements and statistics*. ESRI Press, Redlands, California.
27. Moran P.A.P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika* 37, 17–23.
28. Moulder B.C., Algina J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling*, 9(1), 1-19.
29. Mzuku M., Khosla R., Reich R., Inman D., Smith F., MacDonald L. (2005). Spatial variability of measured soil properties across site-specific management zones. *Soil Science Society of America Journal* 69, 1572–1579.
30. Ord K., Getis A. (2001). Testing for local spatial autocorrelation in the presence of global autocorrelation. *Journal of Regional Science* 41, 411-432.
31. Refaeilzadeh P., Tang L., Liu H. (2009). Cross-validation. In *Encyclopedia of database systems* (pp. 532-538). Springer US.

32. Robert P. C., Rust R.H., Larson. W. E. 1994. Preface p. XIII. w: Robert i in. (red.) *Site-specific management for agricultural systems*. American Society of Agronomy, Madison, WI.
33. Robinson T.P.I., Metternicht G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture* 50, 97-108.
34. Srinivasan A. (ed.) (2006). *Handbook of precision agriculture: Principles and applications*. The Haworth Press Inc., New York.
35. Tobler W. (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46, 234-240.

5) Omówienie pozostałych osiągnięć naukowo-badawczych

Działalność naukowo-badawczą rozpocząłem w roku 1998 wraz z początkiem realizacji pracy magisterskiej pod kierunkiem dr Małgorzaty Rochalskiej, w Katedrze Hodowli Roślin i Nasiennictwa (obecnie włączonej do Katedry Fizjologii Roślin) na Wydziale Rolniczym (obecnie Wydział Rolnictwa i Biologii) Szkoły Głównej Gospodarstwa Wiejskiego W Warszawie. Praca badawcza dotyczyła tematu: *Wpływ zmiennego pola magnetycznego o niskiej częstotliwości na wartość siewną nasion buraka cukrowego*. W ramach tej pracy uczestniczyłem w prowadzeniu badań polowych i laboratoryjnych dotyczących oceny wartości siewnej (w tym zdolności kiełkowania i polowej zdolności wschodów oraz ich szybkości) nasion burak cukrowego. Pracę magisterską zakończyłem obroną w roku 1999 i w tym samym roku rozpocząłem studia doktoranckie na Wydziale Rolniczym Szkoły Głównej Gospodarstwa Wiejskiego. Początkowo pod kierunkiem prof. dr hab. Marianny Kalinowskiej-Zdun a następnie pod kierunkiem prof. dr hab. Zdzisława Wyszyńskiego realizowałem badania dotyczące tematu: *Struktura i architektura łanu a plonowanie dwóch odmian jęczmienia jarego w zależności od wybranych czynników agrotechnicznych* w Katedrze Szczegółowej Uprawy Roślin (w okresie późniejszym przemianowanej na Zakład Szczegółowej Uprawy Roślin w ramach Katedry Agronomii). W czasie tych badań realizowałem prace badawcze, na które składały się między innymi: prowadzenie badań polowych w Rolniczym Zakładzie Doświadczalnym w Chylicach, badania laboratoryjne związane z oceną cech kłosów i ziarniaków jęczmienia oraz statystyczna analiza danych z ww. doświadczeń. Praca doktorska została zakończona obroną z wyróżnieniem w roku 2005 i w tym samym roku podjąłem pracę na stanowisku adiunkta w Katedrze Biometrii (obecnie Katedra Doświadczalnictwa i Bioinformatyki) na Wydziale Rolnictwa i Biologii, Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie, gdzie ciągle jestem zatrudniony. Prace badawcze od roku 2005 koncentrują się przede wszystkim na wykorzystaniu różnorodnych metod statystycznych w analizie danych z obszaru nauk rolniczych, przyrodniczych i medycznych. Jednym z szczególnych pól zastosowań metod statystycznych są metody statystyki przestrzennej (geostatystyka i związane z nią analizy przestrzenne), które początkowo stosowałem głównie na potrzeby analiz związanych z badaniami dotyczącymi rolnictwa precyzyjnego a obecnie rozszerzyłem również na inne obszary. Trudny jest podział realizowanych badań na obszary tematyczne, gdyż w zależności od przyjętych kryteriów podział ten może być inny. Poniżej przyjąłem podział na obszary tematyczne dotyczące

różnych dyscyplin naukowych oraz w ramach niektórych dyscyplin podział bardziej szczegółowy. Są to następujące kategorie badań:

- a) Badania agronomiczne związane rolnictwem precyzyjnym
- b) Badania agronomiczne związane z wielo cechową oceną odmian
- c) Badania agronomiczne związane z związkami przyczynowo-skutkowymi między plonem a cechami go kształtującymi
- d) Badania medyczne związane z epidemiologią i badaniami klinicznymi dotyczące głównie stomatologii i diabetologii
- e) Badania dotyczące ochrony środowiska i ekologii
- f) Badania dotyczące hodowli drobiu
- g) Badania dotyczące oceny jakości żywności

5a) Badania agronomiczne związane z rolnictwem precyzyjnym były i są nadal realizowane przede wszystkim wspólnie z zespołem z Katedry Agronomii w ramach dwóch zakończonych projektów naukowych, tj. grantu: *Optymalizacja nawożenia azotem pszenicy ozimej z wykorzystaniem urządzenia teledetekcyjnego oraz ocena przydatności pomiaru przewodności elektrycznej gleby do wydzielania stref operacyjnych na polach produkcyjnych (2009-2012)* oraz projektu *BIOPRODUKTY, innowacyjne technologie wytwarzania prozdrowotnych produktów piekarskich i makaronu o obniżonej kaloryczności (2012-2015)*. W ramach pierwszego grantu powstały niektóre publikacje wchodzące w skład ww. opisanego osiągnięcia. Poza tym powstały inne publikacje, których tematyka dotyczy wykorzystania sensorów optycznych do oceny odżywienia roślin azotem oraz optymalizacji metod służących do mapowania zmienności glebowej pojedynczych pól uprawnych. Mój udział w ww. badaniach dotyczył przede wszystkim statystycznej analizy danych z wykorzystaniem metod geostatystycznych (interpolacja przestrzenna), walidacji metod z użyciem metod statystycznych oraz w przypadku doświadczeń czynnikowych – ocenie statystycznej efektów badanych czynników. Efektem było powstanie kilku publikacji, w tym trzech w czasopismach z listy JCR dotyczących wykorzystania wskaźnika wegetacyjnego NDVI w optymalizacji nawożenia azotem oraz pośrednio wykorzystanie tego wskaźnika, jako źródła informacji o zmienności glebowej w ramach pola uprawnego.

Ponadto we współpracy z zespołem z Katedry Gleboznawstwa i Ochrony Gleb w Uniwersytecie Technologiczno-Przyrodniczym w Bydgoszczy oceniałem zmienność wybranych właściwości gleby (m.in. zasobności w fosfor, potas i sól) z użyciem

współczynników autokorelacji przestrzennej, czego efektem było powstanie dwóch publikacji w czasopismach z listy JCR (posiadających IF).

Poza działalnością naukową prowadziłem i prowadzę działalność popularyzatorską z tego zakresu (rolnictwa precyzyjnego), czego efektem były m.in. jeden artykuł w prasie rolniczej (czasopismo Farmer, rok 2008), dwa artykuły w portalu Agropolska.pl (rok 2015). Ponadto przeprowadziłem szkolenia dla doradców i uczniów rolniczych szkół średnich z zakresu rolnictwa precyzyjnego, a w szczególności wykorzystania systemów GIS w rolnictwie, organizowane przez Centrum Doradztwa Rolniczego w Brwinowie.

5b) Badania nad wielocechową oceną odmian dotyczyły przede wszystkim odmian pszenicy ozimej (we współpracy z Katedrą Agronomii SGGW), jak również genotypów traw i żyta (we współpracy z IHAR-PIB w Radzikowie oraz Ogrodem Botanicznym – CZRB PAN w Powsinie). Badania te były prowadzone m.in. w ramach ww. projektu BIOPRODUKTY oraz wcześniej realizowanego projektu badawczego MNiSW pt. „*Badanie uwarunkowania plonu ziarna pszenicy ozimej przez cechy plonotwórcze roślin*” (2009 – 2012), w którym pełniłem rolę wykonawcy. W ramach tego projektu powstało kilka publikacji z *impact factor*. W publikacjach tych przedstawiono m.in. klasyfikację odmian pszenicy ozimej pod względem różnego uwarunkowania plonu ziarna przez jego składowe (liczbę kłosów na jednostce powierzchni, liczbę ziaren w kłosie i masę tysiąca ziaren) oraz klasyfikację odmian pod względem cech wpływających na jakość mąki i pieczywa. W analizach wykorzystano wielowymiarowe metody statystyczne takie jak analiza składowych głównych (PCA) i analiza skupień, co pozwoliło na wyodrębnienie grup odmian podobnych pod względem ważnych cech z agronomicznego punktu widzenia lub ważnych w technologii żywności. Poza klasyfikacją odmian, jednym z celów badań nad ich zróżnicowaniem była ocena wpływu środowiska, czyli ocena interakcji genotypowo-środowiskowej, jak również wpływ technologii uprawy na badane cechy.

5c) Inną tematyką badań, lecz powiązaną z ww. tematami są badania dotyczące zależności przyczynowo-skutkowych między plonem a cechami plonotwórczymi. Część badań z tego zakresu była kontynuacją badań prowadzonych w ramach pracy doktorskiej. Badania te dotyczyły jęczmienia jarego, wpływu składowych plonu oraz wybranych cech struktury łanu na zmienność plonu ziarna w różnych warunkach agrotechnicznych u różnych odmian. Do oceny zależności wykorzystywano analizę ścieżek wykorzystującą modele regresji wielokrotnej. W zależności od zestawu cech traktowanych jako przyczynowe

tworzone były jedno- lub wielosekwencyjne modele przyczynowo-skutkowe. Metody te zastosowano również do badań na uwarunkowaniu plonu ziarna różnych odmian pszenicy oraz żywicy trwałej.

5d) Badania medyczne dotyczyły przede wszystkim stomatologii (współpraca z Warszawskim Uniwersytetem Medycznym) oraz endokrynologii (współpraca z Centralnym Szpitalem Klinicznym MSWiA). Badania dotyczyły przede wszystkim epidemiologii, w tym czynników ryzyka powodujących występowanie określonych chorób. Moją rolą był dobór metod statystycznych, przeprowadzenie analiz statystycznych oraz wstępna interpretacja uzyskanych wyników. W różnych badaniach były wykorzystywane m.in. takie metody statystyczne jak: regresja logistyczna, analiza przeżycia (metoda Kaplana-Meiera i regresja Coxa), ogólne modele liniowe oraz różne metody nieparametryczne. W ramach tej tematyki powstało kilkanaście publikacji naukowych z *impact factor*, które dotyczyły m.in. stomatologii dziecięcej, ortodontji, periodontologii, diabetologii i kardiologii. Zespół badawczy, w którym uczestniczyłem został nagrodzony zespołową nagrodą III stopnia Warszawskiego Uniwersytetu Medycznego w roku 2011 za cykl prac dotyczących powiązań chorób sercowo-naczyniowych z chorobami przyzębia.

5e) Badania dotyczące ochrony środowiska i ekologii prowadzone były przede wszystkim we współpracy z Ogrodem Botanicznym – CZRB w Powsinie oraz Instytutem Ochrony Środowiska. Tematyka badań dotyczyła wykorzystania metod bioindykacyjnych w ocenie zanieczyszczenia środowiska metalami ciężkimi, wielopierścieniowymi węglowodarami aromatycznymi oraz wpływu zasolenia gleby w środowisku miejskim na stan zdrowotny drzew. W ramach tych badań powstało kilkanaście publikacji, w tym około połowa z *impact factor*. Moim udziałem w tych badaniach był dobór metod statystycznych, przeprowadzenie analiz oraz w wybranych pracach przygotowanie map z wykorzystaniem metod interpolacji przestrzennej i systemów informacji geograficznej (GIS). Najczęściej wykorzystywanymi metodami statystycznymi w tego badaniach były analiza regresji, analiza korelacji oraz porównania średnich lub rozkładów z użyciem metod parametrycznych i nieparametrycznych, jak również klasyfikacja wielocechowa z wykorzystaniem analizy skupień.

5f) Badania dotyczące hodowli drobiu prowadzone są wspólnie z Wydziałem Nauk o Zwierzętach SGGW. Tematyka badań dotyczyła m.in. oceny wyników produkcyjnych brojlerów indyjskich, ze szczególnym uwzględnieniem przyrostów masy ciała w zależności

od rodzaju krzyżówek różnych ras. Badane były przyrosty masy ciała zwierząt w całym okresie wzrostu oraz inne cechy charakteryzujące jakość tuszy tych zwierząt. W badaniach tych moja rolą było przeprowadzenie analiz statystycznych, a w szczególności modelowanie przyrostu zwierząt z wykorzystaniem modeli logistycznych (np. Gomperta, Richardsa, Mignon-Grasteau i Beaumonta) oraz ocena efektów genotypowych wynikających z różnego krzyżowania. Efektem prac było powstanie czterech publikacji z *impact factor*, dalsze prace są w przygotowaniu.

5g) Innymi badaniami, w których uczestniczyłem są badania dotyczące oceny jakości żywności prowadzone wspólnie z Wydziałem Nauk o Żywności oraz Wydziałem Nauk o Żywieniu Człowieka i Konsumpcji. Dotyczyły one między innymi oceny jakości olejów roślinnych w czasie przechowywania, spożycia akrylamidu oraz porównania wybranych cech jakości warzyw między gospodarstwami konwencjonalnymi a ekologicznymi. W badaniach tych wykorzystywałem różne metody statystyczne, m.in. analizę regresji i analizę składowych głównych. Efektem badań było powstanie kilku publikacji z tego zakresu, w tym jedna z *impact factor*.

Podsumowanie ilościowe dotychczasowego dorobku publikacyjnego

Mój całkowity dorobek publikacyjny stanowi 136 publikacji w czasopismach recenzowanych (w tym 5 publikacji przeglądowych), jedna monografia, dwa rozdziały w monografiach, 28 streszczeń konferencyjnych oraz 3 publikacje popularnonaukowe.

Łącznie jestem autorem/współautorem 61 prac w czasopismach z bazy Journal Citation Reports (JCR).

Sumaryczna wartość współczynnika IF (zgodnie z rokiem publikacji) wynosi: 79,506
Sumaryczna liczba punktów według listy MNiSW, zgodnie z czasem opublikowania prac wynosi 1821. Liczba cytowań według bazy Web of Science: 139, bez autocytowań: 116, Indeks Hirscha według bazy Web of Science: 7

Mój dorobek naukowy niewliczający się do osiągnięcia naukowego, stanowiącego powiązany tematycznie cykl publikacji, wynosi 1717 pkt MNiSW, sumaryczna wartość współczynnika IF wynosi: 75,481.

Łączne zestawienie ilościowe opublikowanych prac naukowych

Rodzaj publikacji	Przed uzyskaniem stopnia doktora	Po uzyskaniu stopnia doktora	Ogółem
Publikacje recenzowane w czasopismach posiadających impact factor (z bazy JCR)	-	61	61
- w tym w języku polskim	-	4*	4*
- w tym w języku angielskim	-	57	57
Publikacje recenzowane w czasopismach nie posiadających impact factor (spoza bazy JCR)	12	63	75
- w tym w języku polskim	11	35	46
- w tym w języku angielskim	1	28	29
Monografie	-	1	1
- w tym w języku polskim	-	1	1
- w tym w języku angielskim	-	-	-
Rozdziały w monografiach	-	2	2
- w tym w języku polskim	-	2	2
- w tym w języku angielskim	-	-	-
Streszczenia w materiałach konferencyjnych	2	21	23
- w tym w języku polskim	1	17	18
- w tym w języku angielskim	1	4	5
Publikacje popularnonaukowe	-	3	3
- w tym w języku polskim	-	3	3
- w tym w języku angielskim	-	-	-

* w tym jedna publikacja w języku czeskim

Gozdowski D.