



UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU

Wydział Nauk Geograficznych i Geologicznych

Kierunek studiów: Geoinformacja

Nr albumu: 444861

Błażej Kościański

**Porównanie metod (miar?) określania zmian
struktury przestrzennej kategorii pokrycia
terenu**

TODO

Praca magisterska napisana
w Instytucie Geoekologii i Geoinformacji
pod kierunkiem
dr. hab. Jakuba Nowosada

Poznań, 2023

Streszczenie

Abstrakt

Streszczenie powinno przedstawiać skrótowo główny problem pracy i jego rozwiązanie. Możliwa struktura streszczenia to: (1) 1-3 zdania wstępu do problemu (czym się zajmujemy, dlaczego jest to ważne, jakie są problemy/luki do wypełnienia), (2) 1 zdanie opisujące cel pracy, (3) 1-3 zdania przedstawiające użyte materiały (dane) i metody (techniki, narzędzia), (4) 1-3 zdania obrazujące główne wyniki pracy, (5) 1-2 zdania podsumowujące; możliwe jest też określenie dalszych kroków/planów.

Słowa kluczowe: (4-6 słów /zwrotów opisujących treść pracy, które nie wystąpiły w tytule)

Abstract

The abstract must be consistent with the above text.

Keywords: (as stated before)

Spis treści

Streszczenie	3
1 Wprowadzenie	5
2 Metody	9
2.1 Struktury przestrzenne	9
2.2 Metryki krajobrazowe	10
2.3 Sygnatury przestrzenne	10
2.4 Metody analiz różnic pokrycia terenu	12
2.5 Miary odległości i niepodobieństwa	13
2.6 Przykład obliczenia niepodobieństwa rastrów	14
2.7 Symulowanie rastrów o określonej kompozycji i konfiguracji przestrzennej .	15
3 Materiały	19
3.1 Przygotowanie pierwszej ankiety	19
3.2 Przygotowanie drugiej ankiety	22
3.3 Ankieta w formie aplikacji shiny???	22
4 Wyniki	23
4.1 Wyniki pierwszej ankiety	23
4.2 Wyniki drugiej ankiety	27
5 Porównanie wyników	29
6 Podsumowanie	31

Rozdział 1

Wprowadzenie

Informacje geograficzne stanowią wyniki selekcji i przetwarzania danych **dotyczących aspektów** otaczającej nas przestrzeni geograficznej. Pozwalają na bardziej zrozumiałe i efektywne analizowanie, modelowanie oraz interpretowanie złożonych zjawisk i procesów zachodzących w naszym otoczeniu. Informacje geograficzne i ich aspekty nie stanowią niepodważalnych faktów, lecz często powstają w wyniku działań jednostek, jak i wspólnych wysiłków grup ekspertów, którzy zajmują się wyborem, analizą i klasyfikacją danych geograficznych (Comber et al., 2005). W procesie tworzenia informacji geograficznych istnieje zatem pewien stopień subiektywności, który może wpłynąć na ostateczny kształt (ostateczną postać?) tych informacji, ich interpretację, jak i na ich użyteczność w kontekście innych zastosowań. Przykładem informacji geograficznej, której ostateczna postać zależy od założeń przyjętych w trakcie tworzenia danych przestrzennych, jest pokrycie terenu.

Przyjmuje się, że termin pokrycie terenu obejmuje zbiór wszelkich elementów obecnych na powierzchni Ziemi (**dodać ref**). W elementy pokrycia terenu włączają się obiekty związane z działalnością człowieka, skutkami sił przyrody oraz wszelkie inne istniejące obiekty, które mogą znaleźć się w przestrzeni geograficznej (**dodać ref**). Tworzenie dokładnych i wiarygodnych danych dotyczących pokrycia terenu jest niezbędne w kontekście wielu zastosowań, takich jak planowanie przestrzenne (**dodać ref**), ochrona środowiska (**dodać ref**), czy analiza zmian klimatycznych (**dodać ref**). Ostateczna

forma tych danych jest jednak w dużej mierze determinowana przez wybory i założenia dokonywane w procesie ich tworzenia (**ewentualnie dodać rycinę z przykładem mapy pokrycia terenu**). W tym kontekście, analiza pokrycia terenu staje się istotnym polem badań, które skupia się na zarówno na technicznych aspektach zbierania danych, jak i na ich semantycznej interpretacji (**dodać ref**).

w tym akapicie gdzieś wspomnieć o rastrach Dane oraz wynikowe mapy pokrycia terenu są rezultatem złożonego procesu przetwarzania i analizy danych przestrzennych najczęściej w postaci obrazów satelitarnych (**dodać ref**). Na początku tego procesu, satelity wyposażone w sensory rejestrują obrazy Ziemi z różnych zakresów widmowych. Uzyskane obrazy mogą być interpretowane manualnie przez grupy specjalistów. Pozwala to na uzyskanie map pokrycia terenu o wysokiej dokładności, kosztem długiego procesu ich tworzenia (**dodać ref**). Dużo mniej czasochłonną metodą jest przetwarzanie przy użyciu algorytmów. Umożliwiają one względnie szybką, półautomatyczną identyfikację i klasyfikację różnych typów powierzchni kosztem mniejszej dokładności mapy wynikowej (**dodać ref**). Ostatecznie, dane przekształcone w mapy pokrycia terenu mogą posłużyć do analiz zmian pokrycia terenu (**dodać ref**).

Celem analiz zmian pokrycia terenu jest przede wszystkim monitorowanie i pogłębienie aktualnej wiedzy na temat ewolucji otaczającego nas krajobrazu (**dodać ref**). Jest to istotne w kontekście ochrony przyrody (**dodać ref**), planowania przestrzennego (**dodać ref**), oceny wpływu inwestycji (**dodać ref**) i infrastruktury na środowisko (**dodać ref**), a także w badaniach naukowych dotyczących zmian klimatycznych (**dodać ref**), bioróżnorodności (**dodać ref**) oraz innych procesów ekologicznych (Lu et al., 2004). Dzięki analizie zmian pokrycia terenu można identyfikować obszary zagrożone degradacją, monitorować skutki urbanizacji, deforestacji czy erozji, co umożliwia podejmowanie odpowiednich działań w celu zrównoważonego zarządzania środowiskiem i zachowaniem jego integralności.

W badaniach nad zmianami pokrycia terenu wykorzystuje się różnorodne metody analityczne. Niemniej jednak, wiele z tych technik koncentruje się na analizie zmian na poziomie indywidualnych komórek w siatce rastra (Lu et al., 2004). Choć podejście to może dostarczać użytecznych informacji dotyczących trendów zmian pokrycia terenu

na niewielkich obszarach, charakteryzuje się ono istotnymi ograniczeniami w kontekście interpretacji wyników. Szczególnie w przypadku badań obejmujących rozległe terytoria, takie jak kraje czy nawet kontynenty, bardziej efektywne staje się zastosowanie metod opartych na analizie struktur przestrzennych (Netzel et al., 2015). Głównym założeniem tych metod jest przekształcenie danych z postaci pojedynczych wartości komórek rastra w sygnatury przestrzenne, a następnie porównanie ich ze sobą za pomocą miar odległości i niepodobieństwa.

Sygnatury przestrzenne stanowią statystyczny opis struktur przestrzennych kategorii pokrycia terenu na mniejszych, wydzielonych obszarach w obrębie całego zbioru danych. W celu porównania ze sobą dwóch sygnatur przestrzennych, wykorzystywane są miary niepodobieństwa. Umożliwiają one określenie w jakim stopniu dwa analizowane obszary się od siebie różnią pod względem kompozycji oraz konfiguracji przestrzennej. Opracowane zostało wiele różnych miar niepodobieństwa, takich jak odległość euklidesowa, odległość Canberra, metryka Wave Hedges, współczynnik podobieństwa Jaccarda, odległość Jensena-Shannona czy dywergencja Pearsona (Cha, 2007). Współcześnie jednak nie określono, która z tych miar jest najbardziej zgodna zarówno z postrzeganiem przez człowieka, jak i wpływem zmian na procesy środowiskowe.

wady i zalety obu podejść, np. że percepcja jest subiektywna i nie skaluje się

Celem tej pracy było porównanie metod określania zmian struktury przestrzennej kategorii pokrycia terenu w kontekście ich korelacji z postrzeganiem zmian przestrzennych przez ludzi. W celu realizacji tego zadania przeprowadzone zostały ankiety, w których zadaniem respondentów było określenie stopnia podobieństwa między parami rastrów. Badania przeprowadzone zostały na rastrach składających się wyłącznie z dwóch lub trzech kategorii. Wyniki ankiety zestawione zostały z wartościami 46 miar niepodobieństwa. Na tej podstawie, do dalszej analizy wybrane zostało 8 miar niepodobieństwa charakteryzujących się największą zgodnością z ludzką percepcją zmian przestrzennych.

Rozdział 2

Metody

W tym rozdziale zostaną kolejno opisane wszystkie zagadnienia kluczowe do zrozumienia tematyki określania zróżnicowania struktur przestrzennych w kontekście zmian pokrycia terenu. W pierwszej kolejności wyjaśnione zostanie zagadnienie struktur przestrzennych. Następnie omówione zostaną wskaźniki umożliwiające określanie charakterystyk struktur przestrzennych, czyli metryki krajobrazowe oraz sygnatury przestrzenne. Omówione zostaną koncepcje kompozycji oraz konfiguracji przestrzennej, a także miary umożliwiające ich obliczenie. Opisana zostanie także idea reprezentacji rastrów w postaci macierzy i wektorów współwystępowania. W następnej kolejności zostaną przedstawione dwie metody analiz zmian pokrycia terenu. Pierwsza, opierająca się na analizie różnic na poziomie indywidualnych komórek w siatce rastra, oraz druga, oparta na analizie struktur przestrzennych występujących wewnątrz rastra. Wyjaśnione zostaną zagadnienia miar odległości i niepodobieństwa oraz ich wykorzystania w analizach przestrzennych. Przedstawiony zostanie także proces obliczenia niepodobieństwa między parą rastrów. W końcowej części rozdziału opisany zostanie sposób symulacji danych rastrowych o określonej kompozycji i konfiguracji przestrzennej.

2.1 Struktury przestrzenne

Znaczna część dziedziny ekologii krajobrazu opiera się na paradygmacie płatów (McGarigal et al., 2009). Według tej idei krajobrazy składają się z jednostek (płatów), które

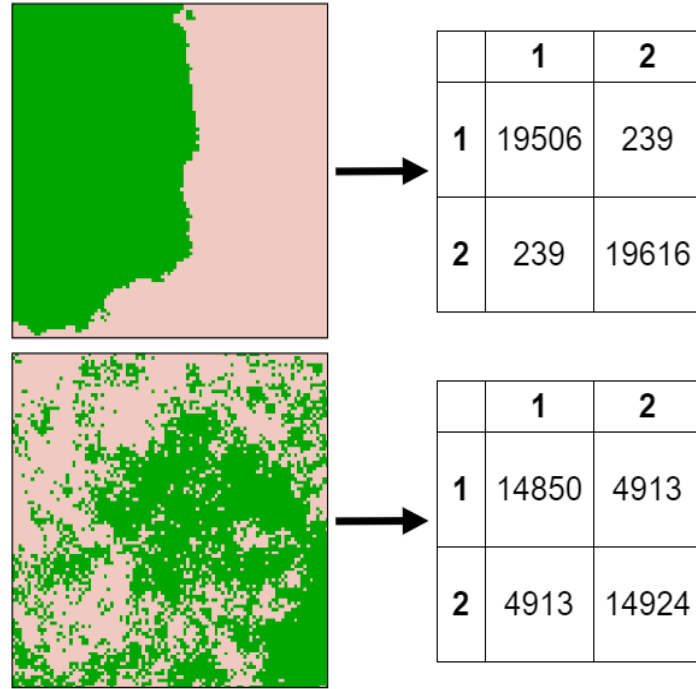
charakteryzuje się jako wyodrębnione obszary o jednakowej kategorii pokrycia terenu (Forman, 1995; Solon, 2002). Każdy krajobraz natomiast cechuje się pewną strukturą przestrzenną, której różne cechy mogą następnie być opisywane na przykład za pomocą metryk krajobrazowych lub sygnatur przestrzennych.

2.2 Metryki krajobrazowe

Metryki krajobrazowe umożliwiają analizę struktury przestrzennej krajobrazu na podstawie danych o pokryciu terenu (Pukowiec-Kurda et al., 2016). Można je podzielić na dwie grupy: wskaźniki kompozycji i konfiguracji (Solon, 2002; Pukowiec-Kurda et al., 2016). Celem pierwszej grupy metryk jest opis kompozycji (udziału) rastrów, czyli różnicowania i liczby pól poszczególnych kategorii pokrycia terenu bez uwzględniania informacji o ich lokalizacji w przestrzeni (Gustafson, 1998; Solon, 2002; Kozak et al., 2014). Konfiguracja (ułożenie) natomiast w sposób liczbowy opisuje sąsiedowanie ze sobą poszczególnych pól (Gustafson, 1998; Solon, 2002; Kozak et al., 2014). Szczególną zaletą takiego podejścia jest to, że metryki krajobrazowe mogą być obliczone dla różnorodnych jednostek przestrzennych, które mogą być zdefiniowane na podstawie wszelkich aspektów administracyjnych, geograficznych, biogeograficznych lub umownych (Pukowiec-Kurda et al., 2016). Jednostki te mogą obejmować zarówno granice gmin, zlewni, ekoregionów czy nawet abstrakcyjne obszary na mapie, co z kolei umożliwia wszechstronne stosowanie tych metryk w badaniach związanych z analizą krajobrazu.

2.3 Sygnatury przestrzenne

Często w celu lepszej reprezentacji analizowanych rastrów możemy wykorzystywać także bardziej skomplikowane sygnatury przestrzenne. Przykładem sygnatury łączącej zarówno kompozycję, jak i konfigurację przestrzenną jest macierz współwystępowania. Jest to macierz o wymiarach k na k , gdzie k reprezentuje liczbę kategorii pokrycia terenu obecnych w analizowanym rastrze (Haralick et al., 1973; Jasiewicz et al., 2015). Macierz tą możemy skonstruować poprzez zliczanie kolejno wszystkich par sąsiadujących ze sobą komórek w rastrze. Wewnątrz tej macierzy, wartości ułożone na przekątnej odnoszą się do kompozycji rastra, natomiast pozostałe do jego konfiguracji. Przykład dwóch



Rycina 2.1: Przykład dwóch macierzy współwystępowania dla rastrów o zbliżonej kompozycji i różnej konfiguracji przestrzennej.

macierzy współwystępowania dla rastrów o zbliżonej kompozycji i różnej konfiguracji przestrzennej przedstawia Rycina 2.1.

Na podstawie macierzy współwystępowania mogą zostać obliczone różne miary wywodzące się z dziedziny teorii informacji. Przykładem takiej miary jest entropia brzegowa **PRZYKŁADY, JAK JEST LICZONA, WZÓR** która opisuje zróżnicowanie kompozycji rastra, czyli udziałów każdej z kategorii w rastrze.

$$H(y) = - \sum_{j=1}^K p(y = c_j) \log_2 p(y = c_j)$$

$$H(y|x) = \sum_{i=1}^K \sum_{j=1}^K p(x = c_i, y = c_j) \log_2 p(y = c_i | x = c_j)$$

$$I(y, x) = H(y) - H(y|x)$$

$$U = I(y, x) / H(y)$$

Kolejną miarą jest względna informacja wzajemna **PRZYKŁADY, JAK JEST LICZONA, WZÓR**. Reprezentuje ona stopień sąsiedowania ze sobą kategorii w rastrze, czyli jego konfigurację przestrzenną. W celu porównywania ze sobą sygnatur dwóch rastrów w postaci dwuwymiarowej macierzy należy je sprowadzić do postaci jednowymiarowego wektora, a następnie przeprowadzić jego normalizację, tak aby wszystkie wartości sumowały się do 1. Taka postać pozwala na obliczanie miar odległości lub podobieństwa, pozwalających na porównywanie histogramów wartości (Cha, 2007). Miary te następnie pozwalają określić stopień odmienności dwóch rastrów. Podejście to może być także wykorzystane w innych analizach przestrzennych, jak wyszukiwanie obszarów o podobnej strukturze przestrzennej, wykrywanie ich zmian oraz grupowanie obszarów o podobnej strukturze przestrzennej (Jasiewicz et al., 2015; Nowosad, 2021).

2.4 Metody analiz różnic pokrycia terenu

Różnice w pokryciu terenu można analizować przy użyciu różnych metod. Wiele z nich koncentruje się na analizie różnic na poziomie indywidualnych komórek w siatce rastra. Najbardziej podstawowym przykładem takiego podejścia jest analiza ilościowa różnic w pokryciu terenu. Zaletą tego podejścia jest przede wszystkim łatwość w wykonaniu analizy. Wystarczy zliczyć wszystkie komórki należące do poszczególnych kategorii dla wybranych rastrów, a następnie porównać ze sobą te wartości, aby otrzymać wynik informujący nas o ilościowych różnicach między analizowanymi rastrami. Analiza ilościowa najczęściej wykorzystywana jest w celu wskazania ogólnych trendów zmian pokrycia terenu dla określonego obszaru badań jak na przykład zmniejszanie się obszarów leśnych lub wzrost terenów zurbanizowanych.

Wszelkie metody analiz zmian pokrycia terenu opierające się na analizie poszczególnych komórek w siatce rastra są użyteczne na obszarach, gdzie zmiany między indywidualnymi komórkami dostarczają istotnych informacji. Ich przydatność jednak maleje, gdy informacja na poziomie pojedynczej komórki przestaje być tak istotna, na przykład dla rastrów o wysokiej rozdzielczości lub znacznym zasięgu przestrzennym (Jasiewicz et al., 2015). W takiej sytuacji bardziej efektywne staje się zastosowanie metod opartych na analizie struktur przestrzennych (Netzel et al., 2015).

Metody oparte na analizie struktur przestrzennych wywodzą się z dziedziny ekologii krajobrazu. Pozwalają one przede wszystkim na opis oraz obliczenie podobieństwa struktur przestrzennych. Głównym zamysłem tych metod jest przekształcenie danych z postaci dużych rastrów zbudowanych z wielu indywidualnych komórek zawierających pojedyncze informacje w sygnatury przestrzenne, a następnie porównanie ich za pomocą miar odległości lub niepodobieństwa. Sygnatura przestrzenna stanowi statystyczny opis pewnych struktur przestrzennych występujących wewnątrz rastra (Jasiewicz et al., 2015; Nowosad, 2021). Dokładnie w tym samym celu wykorzystywane są także metryki krajobrazowe. Mają one jednak istotną wadę w kontekście kompleksowych analiz przestrzennych. Jako że pojedyncza metryka krajobrazowa reprezentuje wyłącznie jedną, konkretną charakterystykę analizowanego obszaru, to nie jest w stanie zdefiniować całej charakterystyki przestrzennej danego rastra. W tym celu bardziej efektywne może być zastosowanie sygnatur przestrzennych. Są one dwuwymiarową reprezentacją kompozycji i konfiguracji przestrzennej rastra, czyli jego najbardziej podstawowych charakterystyk. Oznacza to też, że można je ze sobą porównywać przy użyciu szerokiej gamy istniejących miar odległości i niepodobieństwa. Umożliwia to także wykonywanie bardziej skomplikowanych analiz przestrzennych, jak wyszukiwanie, wykrywanie zmian, grupowanie i segmentacja (Nowosad, 2021).

2.5 Miary odległości i niepodobieństwa

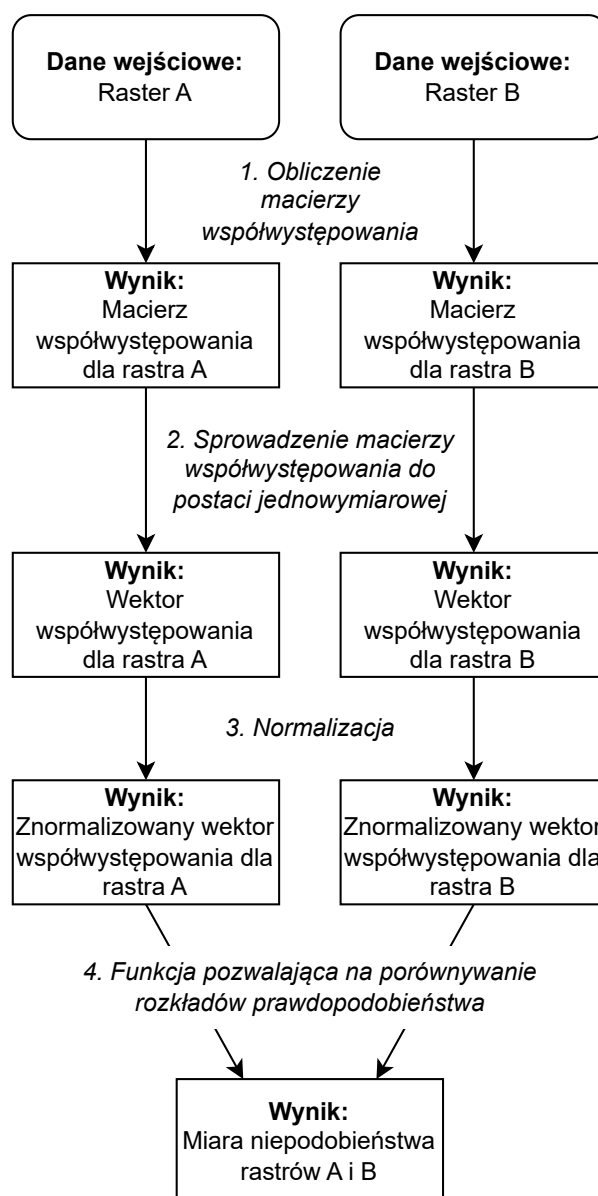
Odległość i rozbieżność (inaczej niepodobieństwo) stanowią pewien policzalny stopień zróżnicowania pary obiektów. Największą różnicą między nimi jest to, że odległości są symetryczne, podczas gdy rozbieżności są niesymetryczne. Oznacza to, że wyłącznie dla miar odległości otrzymujemy identyczny wynik przy porównywaniu par obiektów A i B, jak i par B i A.

Niepodobieństwo jest przeciwieństwem podobieństwa. Ponadto, miary podobieństwa można łatwo przekształcić w miary niepodobieństwa (Niesterowicz et al., 2016). W związku z tym, w celu uproszczenia terminologii, wszystkie miary odległości, podobieństwa oraz te wywodzące się z dziedziny teorii informacji, które zostały wykorzystane w tej pracy, będą dalej zbiorowo nazywane miarami niepodobieństwa.

Wybór odpowiedniej miary niepodobieństwa zależy między innymi od rodzaju pomiaru lub sposobu reprezentacji obiektów (Cha, 2007). Na podstawie podobieństw syntaktycznych, wyróżnia się kilka grup rodzin miar niepodobieństwa (Cha, 2007): rodzina Minkowski (odległość euklidesowa, odległość Minkowskiego, odległość Manhattan), rodzina L1 (Canberra, Sorensen, Kulczynski), rodzina Intersection (Intersection, Wave Hedges, Ruzicka), rodzina Inner Product (Jaccard, Harmonic mean), rodzina Squared-chord (Fidelity, Matusita), rodzina Squared L2 (Clark, Pearson X2, Neyman X2), rodzina Shannon's Entropy (Jensen-Shannon, Kullback-Leibler), a także miary będące połączeniem innych miar (Taneja, Kumar-Johnson) oraz miary wywodzące się z teorii informacji (informacja wzajemna, entropia Shannona).

2.6 Przykład obliczenia niepodobieństwa rastrów

Pierwszym krokiem, jaki należy podjąć w celu obliczenia podobieństwa struktur przestrzennych danych rastrowych z wykorzystaniem metod opartych o sygnatury przestrzenne jest sprowadzenie rastrów wejściowych do postaci macierzy współwystępowania. Proces utworzenia macierzy współwystępowania polega na zliczeniu wartości każdej indywidualnej komórki rastra, a także przylegających do niej komórek (najczęściej czterech lub ośmiu). Przykładowe macierze współwystępowania widoczne są na Rycinie 2.1. Następnie, dwuwymiarową macierz należy sprowadzić do postaci jednowymiarowej, czyli wektora współwystępowania. Kolejnym etapem analizy jest normalizacja wektora współwystępowania. Po dodaniu do siebie wszystkich wartości tego wektora powinniśmy otrzymać wynik równy 1. Po wykonaniu powyższych czynności otrzymujemy reprezentację rastrów wejściowych, która umożliwia porównanie ich ze sobą przy użyciu miar niepodobieństwa między rozkładami prawdopodobieństwa, takich jak rozbieżność Jensena-Shannona. Proces obliczenia niepodobieństwa dwóch rastrów w postaci schematu przedstawia Rycina 2.2.



Rycina 2.2: Schemat procesu obliczenia niepodobieństwa dwóch rastrów.

2.7 Symulowanie rastrów o określonej kompozycji i konfiguracji przestrzennej

Najważniejszym założeniem przy tworzeniu zbioru rastrów do pierwszej ankiety było przygotowanie ich w sposób umożliwiający uzyskanie pełnej reprezentacji wszystkich możliwych wartości kompozycji, jak i konfiguracji przestrzennej. Zbiór rastrów został przygotowany w języku programowania R (R Core Team, 2023), w oparciu o wykorzystanie funkcji *nlm_fbm* z pakietu NLMR (Sciaini et al., 2018). Funkcja ta pozwala na

symulację rastrów przy użyciu ułamkowych ruchów Browna, będących uproszczeniem ruchów Browna (Schlather et al., 2015). W tej funkcji poziom korelacji między kolejnymi krokami jest kontrolowany za pomocą parametru "fract_dim". W kontekście tego badania, parametr ten reguluje konfigurację przestrzenną. Oznacza to, że w przypadku, gdy "fract_dim" przyjmuje niską wartość, wartości w generowanym rastrze rozmieszczone są w sposób losowy, zbliżony do szumu. Natomiast w przypadku wysokiej wartości "fract_dim", na wynikowym rastrze tworzą się skupiska najwyższych i najniższych wartości, a przejścia pomiędzy nimi mają płynny, wygładzony charakter. Powyższa funkcja umożliwia uzyskanie danych rastrowych wypełnionych wartościami zmiennoprzecinkowymi mieszczącymi się w zakresie od 0 do 1 oraz dowolnymi parametrami konfiguracji przestrzennej.

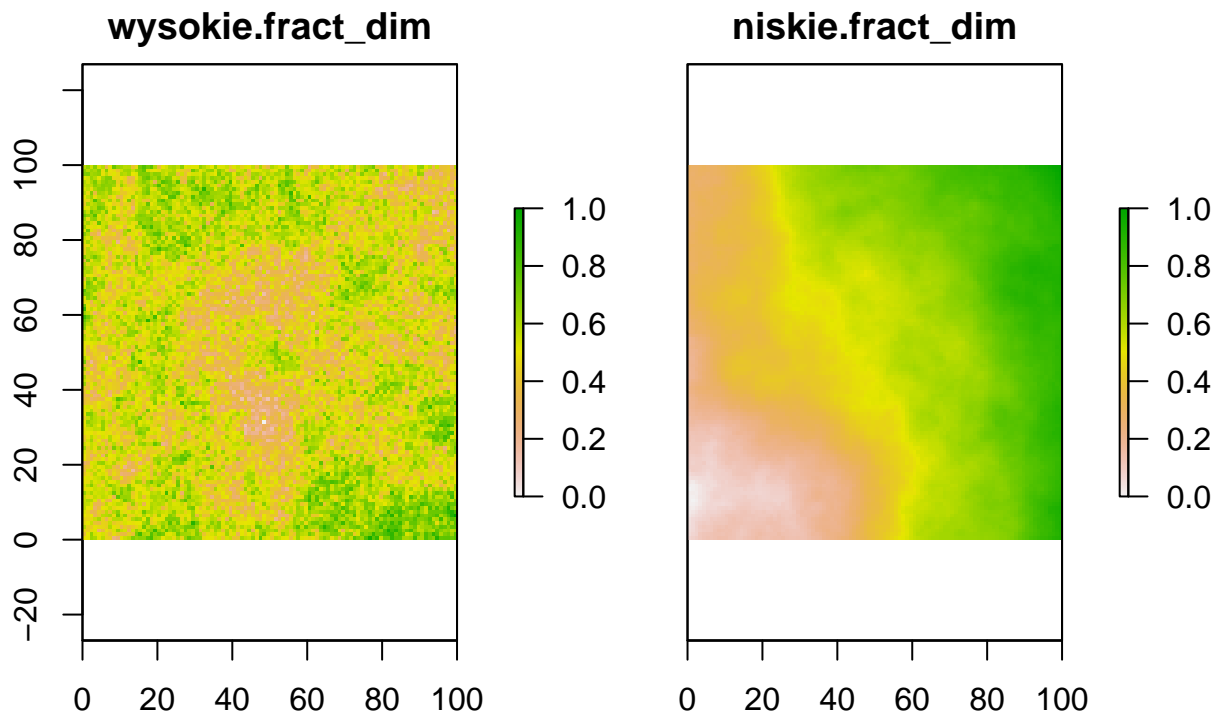
```
library(RandomFields)
library(NLMR)

sim1 = nlm_fbm(100, 100, fract_dim = 0.3)
sim2 = nlm_fbm(100, 100, fract_dim = 1.6)

sim_stack1 = raster::stack(sim1, sim2)
names(sim_stack1) = c("wysokie fract_dim", "niskie fract_dim")

plot(sim_stack1)
```

Następnie, aby otrzymać zbiór rastrów uwzględniający także pełen przekrój kompozycji należy przeprowadzić proces reklasyfikacji. Procedura ta polega na podziale każdego dotychczas utworzonego rastra na kategorie pokrycia terenu w różnych proporcjach, na przykład 90:10, 70:30 oraz 50:50 w przypadku rastrów zawierających wyłącznie dwie kategorie pokrycia terenu. Wykonanie tego procesu ułatwia między innymi funkcja *util_binarize* z pakietu *landscapetools* (Sciaini et al., 2018). W tej funkcji proporcje kategorii pokrycia terenu kontrolowane są za pomocą parametru "breaks". Przykładowo, ustawienie parametru "breaks" na poziomie 0.2 skutkuje otrzymaniem rastra o kategoriach pokrycia terenu w proporcjach 20 do 80. Oznacza to, że jedna z kategorii



będzie obejmowała swoim zasięgiem 20% komórek rastra, podczas gdy druga kategoria wypełni pozostałe 80% komórek.

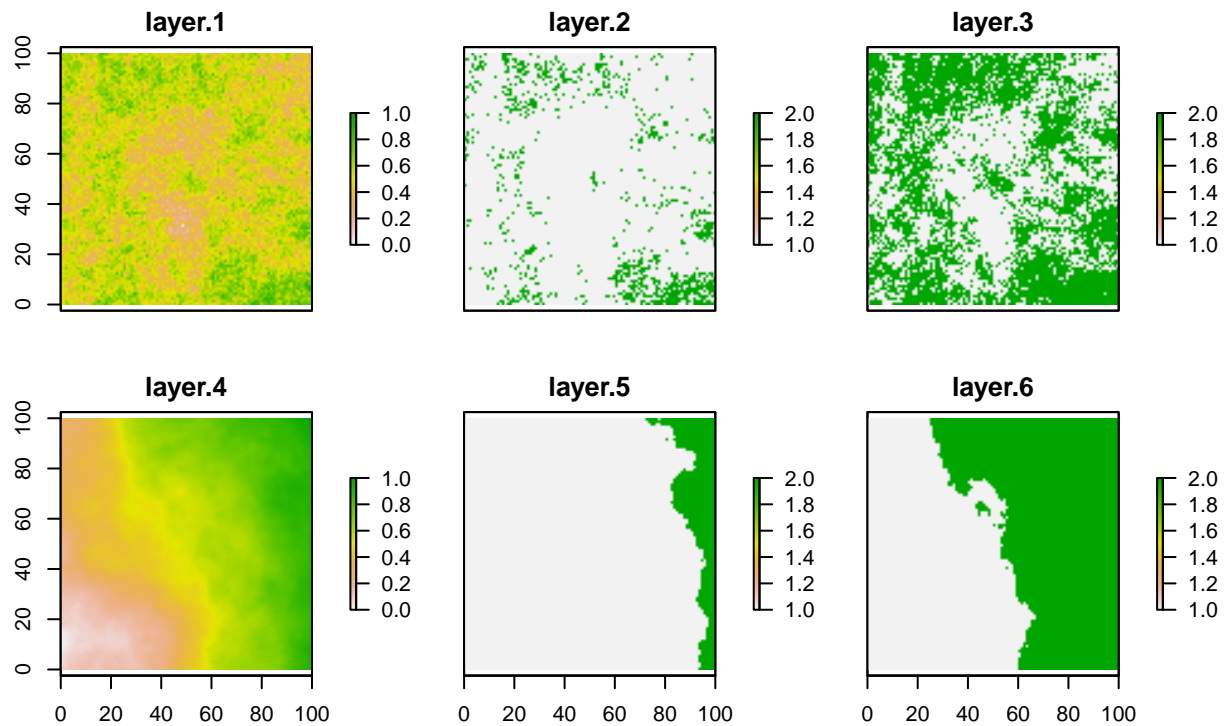
```
library(landscapetools)

sim1_1 = landscapetools::util_binarize(sim1, breaks = 0.1)
sim1_2 = landscapetools::util_binarize(sim1, breaks = 0.5)

sim2_1 = landscapetools::util_binarize(sim2, breaks = 0.1)
sim2_2 = landscapetools::util_binarize(sim2, breaks = 0.5)

sim_stack2 = raster::stack(sim1, sim1_1, sim1_2,
                           sim2, sim2_1, sim2_2)

plot(sim_stack2)
```



```
# library(rasterVis)
# gplot(sim_stack2) +
#   geom_tile(aes(fill = value)) +
#   facet_wrap(~ variable) +
#   scale_fill_gradientn(colours = rev(terrain.colors(225))) +
#   coord_equal()
#
# # wizualizacja symulowanych rastrów
# tm_shape(sim_stack2) +
#   tm_raster(style = 'cont', breaks = c(0,2), palette = terrain.colors(255)) +
#   tm_layout(asp = 1, legend.show = FALSE, panel.show = FALSE)
```

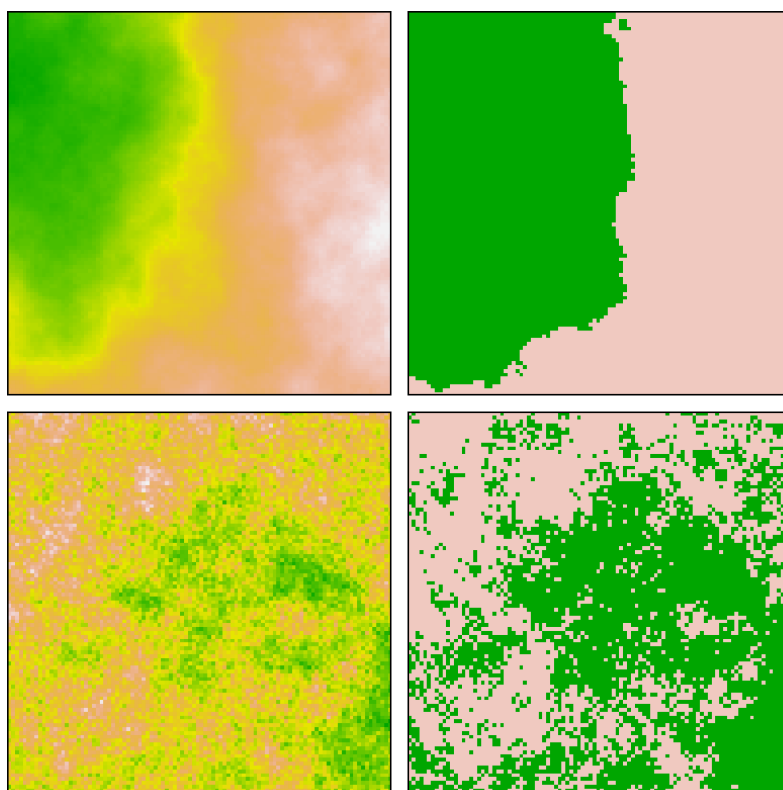
Rozdział 3

Materiały

W tym rozdziale zostanie omówiony cel przeprowadzenia pierwszej ankiety oraz istotne aspekty związane z jej formą. Następnie opisany zostanie sposób przygotowania danych wraz z przyjętą metodą doboru pytań do ankiety.

3.1 Przygotowanie pierwszej ankiety

Najważniejszym założeniem przy tworzeniu zbioru rastrów do pierwszej ankiety było przygotowanie ich w sposób umożliwiający uzyskanie pełnej reprezentacji wszystkich możliwych wartości kompozycji, jak i przestrzennej konfiguracji. Zbiór rastrów został przygotowany w oparciu o wykorzystanie funkcji *nlm_fbm* z pakietu NLMR (Sciaini et al., 2018). Funkcja ta pozwala na symulację rastrów przy użyciu ułamkowych ruchów Browna, będących uproszczeniem ruchów Browna (Schlather et al., 2015). W tej funkcji poziom korelacji między kolejnymi krokami jest kontrolowany za pomocą parametru “frac_dim”. W kontekście tego badania, parametr ten reguluje konfigurację przestrzenną. Oznacza to, że w przypadku, gdy “frac_dim” przyjmuje niską wartość, wartości w generowanym rastrze rozmieszczone są w sposób losowy, zbliżony do szumu. Natomiast w przypadku wysokiej wartości “frac_dim”, na wynikowym rastrze tworzą się skupiska najwyższych i najniższych wartości, a przejścia pomiędzy nimi mają płynny, wygładzony charakter. **tutaj dodać kawałek kodu pokazujący jak działa funkcja?** Kolejnym krokiem w procesie tworzenia danych było dodanie zróżnicowania rastrów na podstawie ich



Rycina 3.1: Wizualizacja procesu symulacji rastrów do pierwszej ankiety albo placeholder zamiast pokazania jak działa funkcja

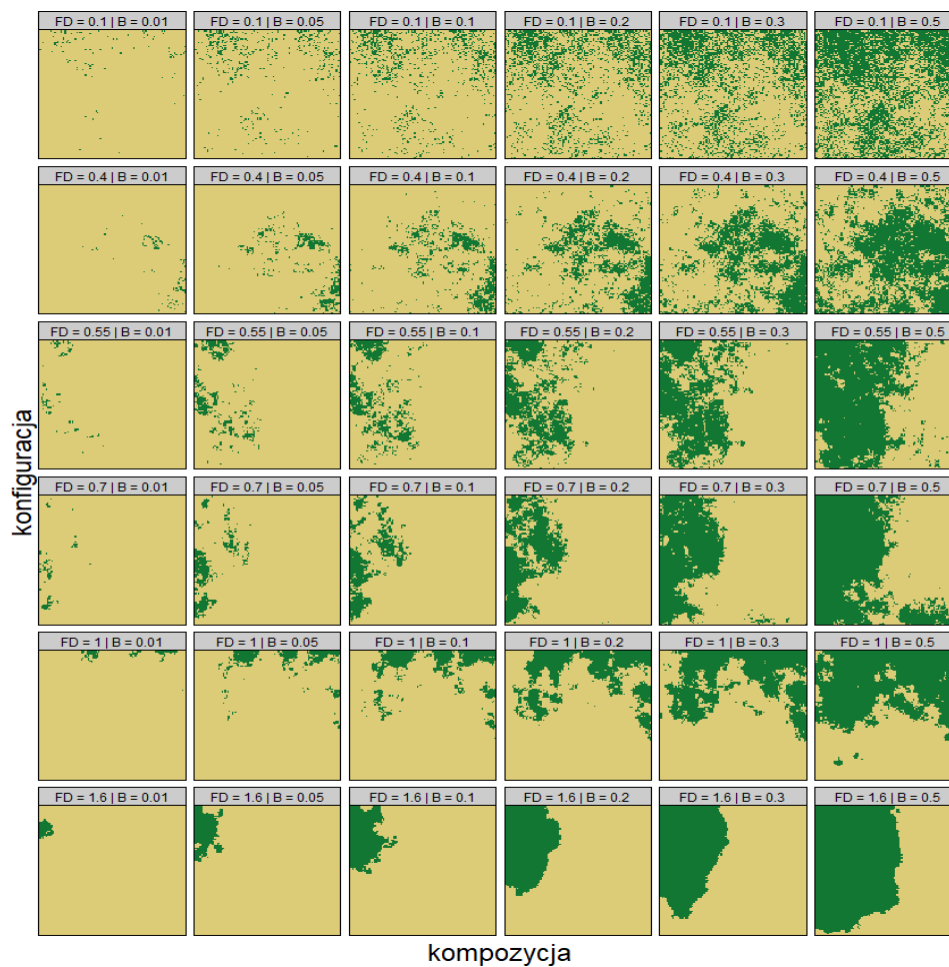
przestrzennej kompozycji. Zostało to osiągnięte poprzez podział każdego z wynikowych rastrów na kategorie w oparciu o stałe przedziały wartości dla wszystkich rastrów. Proces

W następnym etapie przygotowania danych obliczone zostały wybrane miary opisujące struktury przestrzenne: entropia (ent) oraz względna informacja wzajemna (relmutinf). Miary te umożliwiły potwierdzenie uzyskania oczekiwanego rozkładu kompozycji i konfiguracji wewnątrz zbioru rastrów.

Ostatecznie wygenerowane zostały zbiory rastrów składających się wyłącznie z dwóch lub trzech kategorii pokrycia terenu. Przykład jednego ze zbiorów rastrów przedstawia Rycina 3.2. Rastry zawierające trzy kategorie zostały uwzględnione w badaniu w celu próby wskazania czy liczba kategorii na rastrach ma wpływ na odpowiedzi udzielane przez ankietowanych.

Ostatnim krokiem przygotowania danych było wybranie par rastrów tworzących poszczególne pytania. W tym celu, pytania w ankietach podzielone zostały na dwie grupy,

wewnątrz których znalazły się po trzy podgrupy pytań. W pierwszej kolejności respondenci zetknęli się z 24 pytaniami dotyczącymi rastrów uwzględniających wyłącznie dwie kategorie pokrycia terenu, a następnie z 24 pytaniami uwzględniającymi trzy kategorie pokrycia terenu. Pierwsza podgrupa pytań (6 par rastrów) składała się z par rastrów różniących się między sobą wyłącznie entropią. Podgrupa druga (6 par rastrów) zawierała wyłącznie rastry różniące się względną informacją wzajemną. Ostatnia podgrupa (12 par rastrów) składała się z pytań zróżnicowanych zarówno pod względem entropii, jak i względnej informacji wzajemnej. Taki sposób doboru pytań pozwolił na zredukowanie liczby odpowiedzi wymaganych od respondentów, jak i ograniczenie wpływu błędu selekcji, który powstałby w wyniku niewłaściwego doboru pytań. Respondenci celowo nie zostali poinformowani o występujących różnicach pomiędzy kolejnymi pytaniami, ponieważ mogłoby mieć to wpływ na udzielane przez nich odpowiedzi, co z kolei mogłoby wpłynąć na ostateczne wyniki badania.



Rycina 3.2: Przykład zbioru wygenerowanych rastrów (2 kategorie pokrycia terenu)

3.2 Przygotowanie drugiej ankiety

3.3 Ankieta w formie aplikacji shiny???

Rozdział 4

Wyniki

W tym rozdziale przeprowadzona zostanie analiza i podsumowanie wyników uzyskanych z ankiet. Wyniki zostaną zestawione z miarami niepodobieństwa uwzględnionymi w analizie.

4.1 Wyniki pierwszej ankiety

Pierwsza ankieta została przeprowadzona w celu uzyskania wstępnych informacji, mających na celu zrozumienie wpływu różnic w entropii oraz względnej informacji wzajemnej między parami analizowanych rastrów na wizualne postrzeganie zmian pokrycia terenu przez ludzi. Głównym celem badania było wykazanie potencjalnych związków między percepcją zmian w pokryciu terenu przez ludzi a miarami niepodobieństwa, które te zmiany kwantyfikują. Przeprowadzenie ankiety pozwoliło także na wyznaczenie dalszego kierunku badań, które miałyby zostać zrealizowane w kolejnej ankiecie Section 4.2.

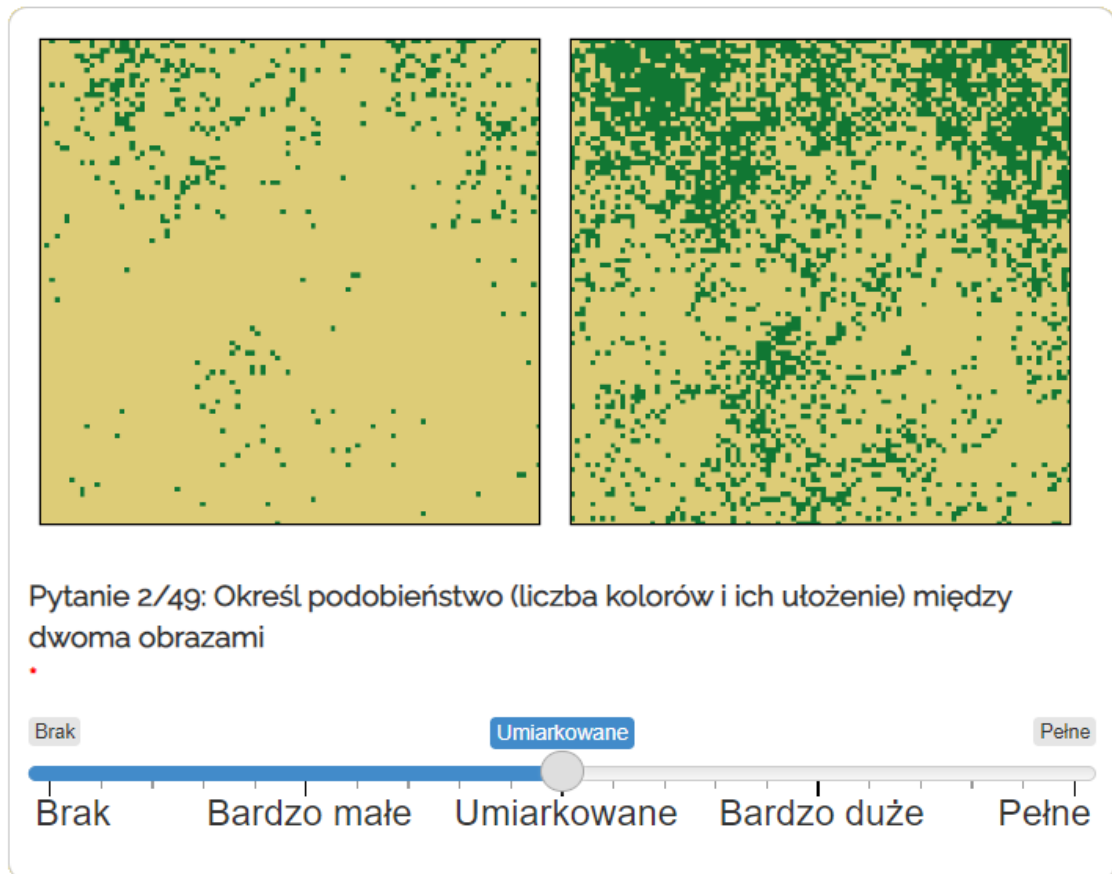
Badanie zostało realizowane w terminie od 21 do 24 listopada 2022 roku. Proces zbierania odpowiedzi respondentów przyjął formę ankiety online, co pozwoliło respondentom na wygodny udział w badaniu przy użyciu komputera lub urządzenia mobilnego. Ankieta stworzona została w formie aplikacji internetowej za pomocą języka programowania R, na podstawie pakietów **shiny** (2022) oraz **shinysurveys** (2021) **NAPRAWIĆ BO NIE POKAZUJE SIĘ CYTOWANIE W BIBLIOGRAFII**. Sama aplikacja umieszczona została na platformie shinyapps.io. Przeprowadzenie ankiety w formie online umożliwiło

systematyczne gromadzenie oraz przechowywanie odpowiedzi w formie tabelarycznej, ułatwiając tym samym dalszą analizę i interpretację danych. Respondenci stanowili grupę 50 studentów drugiego roku studiów inżynierskich na kierunku Geoinformacja na Wydziale Nauk Geograficznych i Geologicznych Uniwersytetu im. Adama Mickiewicza. Wybór tej grupy respondentów oznacza, że byli oni już zaznajomieni z tematyką tworzenia i analiz map w formie rastrowej oraz pojęciem zmian pokrycia terenu.

Każdy z ankietowanych otrzymał do wypełnienia jeden z dwóch wcześniej przygotowanych zbiorów pytań. Każdy ze zbiorów składał się z 48 pytań, przy czym część pytań między zbiorami się pokrywała. Oznacza to, że łącznie uzyskano odpowiedzi na 93 unikatowe pytania. W każdym z pytań zadaniem respondentów było określenie podobieństwa na podstawie dwóch załączonych rastrów. W ramach badania respondenci mieli możliwość wyrażania swoich odpowiedzi za pomocą pięciostopniowej skali Likerta (Likert, 1932), która obejmowała poziomy od "Brak" przez "Bardzo małe", "Umiarkowane", "Bardzo duże" aż po "Pełne". Wykorzystanie skali Likerta o nieparzystej liczbie przedziałów, pozwoliło na zastosowanie przedziału środkowego, którego celem było reprezentowanie odpowiedzi neutralnych lub trudnych do określenia. Początkowo, zamiast skali Likerta planowano wykorzystać skalę liczbową, w zakresie mieszczącym się od 1 do 100, jednakże zrezygnowano z tego pomysłu, jako że znaczenie wartości na skali liczbowej może być interpretowane inaczej przez każdego respondenta oraz skala ta nie pozwala na uwzględnienie wspomnianej wcześniej odpowiedzi neutralnej. Przykład pytania przedstawionego respondentom ilustruje Rycina 4.1.

Łącznie uzyskane zostało 2400 odpowiedzi na pytania z ankiety. Podsumowanie uzyskanych odpowiedzi przedstawione zostało w Tabeli 1.

Według ankietowanych prawie 36% par rastrów charakteryzowała się brakiem podobieństwa, 32,6% uzyskanych odpowiedzi wskazywało na bardzo małe podobieństwo, 18,1% na umiarkowane, 11,6% bardzo duże, natomiast mniej niż 2% wskazywało na pełne podobieństwo. Warto tutaj także zwrócić uwagę, że zestawienia wszystkich odpowiedzi w zależności od liczby kategorii widocznych na rastrach nie wskazują na znaczące różnice w liczbie odpowiedzi dla danej kategorii. Największą różnicę stanowi w tym przypadku



Pytanie 2/49: Określ podobieństwo (liczba kolorów i ich ułożenie) między dwoma obrazami

Brak Umiarkowane Pełne

Brak Bardzo małe Umiarkowane Bardzo duże Pełne

Rycina 4.1: Przykładowe pytanie z pierwszej ankiety

kategoria “Bardzo duże”, dla której liczba odpowiedzi dla rastrów z dwoma i trzema kategoriami pokrycia terenu różni się zaledwie o 2,7%. Najmniejszą różnicą charakteryzuje się kategoria “Pełne”, gdzie liczba odpowiedzi pomiędzy zestawami różni się o jedyne 0,9%.

Poziom zgodności każdego pytania został obliczony jako stosunek najczęściej udzielonej odpowiedzi względem całkowitej liczby odpowiedzi. Całkowity poziom zgodności ankietowanych został oszacowany na 55%. Oznacza to, że 1321 z 2400 udzielonych odpowiedzi znajdowało się w grupie najczęstszej odpowiedzi dla danego pytania.

Pytania różniące się zarówno entropią, jak i względną informacją wzajemną cechowały się najwyższym poziomem zgodności odpowiedzi wynoszącym 61%, podczas gdy pytania różniące się wyłącznie entropią uzyskały wynik 53%, a pytania różniące się wyłącznie względną informacją wzajemną uzyskały wynik 52%. Najwyższy poziom zgodności odpowiedzi osób ankietowanych wyniósł 92% i dotyczył pytania różniącego się zarówno

entropią, jak i względną informacją wzajemną (dopisać jak bardzo różniły się entropie). Najniższy poziom zgodności, czyli zaledwie 38%, osiągnęło aż 8 pytań. Wyłącznie dwa z nich dotyczyły pytań zawierających trzy kategorie pokrycia terenu, podczas gdy pozostałe - dwóch. 75% z tych pytań uwzględniało rastry różniące się pod względem względnej informacji przestrzennej (inna konfiguracja) przy zachowaniu tej samej entropii (identyczna kompozycja).

W dalszym etapie analizy wyniki uzyskane z przeprowadzonej ankiety zostały zestawione z wynikami 45 różnych miar niepodobieństwa dla par rastrów uwzględnionych w ankiecie. Aby ocenić relacje między najczęstszymi odpowiedziami uczestników ankiety a miarami niepodobieństwa, przeprowadzono analizę korelacji Spearmana. W ten sposób uzyskano ostateczny ranking miar, który został przedstawiony w Tabeli 1.

Na podstawie analizy korelacji można wywnioskować, że spośród 45 uwzględnionych w badaniu metod obliczania podobieństwa między parami rastrów, jedynie 6 z nich wykazuje dodatnią korelację z wynikami uzyskanymi z ankiety. Spośród wszystkich metod, najwyższą wartość korelacji osiągnęła metoda Fidelity ze współczynnikiem korelacji równym 0,34. Ponadto, miary "harmonic mean" i "intersection" również wykazały dość wysoki współczynnik korelacji. Jak wspomniano wyżej, nie wszystkie miary odległości charakteryzowały się dodatnim współczynnikiem korelacji. Miary "canberra", "divergence" i "clark" wykazały współczynniki korelacji poniżej -0,4, co wskazuje na ich umiarkowaną, ujemną zależność z wynikami ankiety. Oznacza to, że miary te wskazują na wysokie podobieństwo par rastrów, które ludzie określiliby jako zupełnie niepodobne i vice versa. Pytania dotyczące rastrów uwzględniających wyłącznie dwie kategorie pokrycia terenu z reguły charakteryzują się wyższym współczynnikiem korelacji z najczęstszymi odpowiedziami ankietowanych niż pytania z podziałem na trzy kategorie pokrycia terenu. Jedyny wyjątek stanowią w tej sytuacji wyłącznie miary, które osiągnęły najniższe wartości współczynnika korelacji. Dla rastrów z trzema kategoriami osiągnęły one jeszcze niższy współczynnik na poziomie nawet -0,77. Ich wyniki wskazują na silne, ujemne związki między analizowanymi zmiennymi.

Otrzymane wartości współczynników korelacji nie pozwalają na definitywne stwierdzenie na istnienie liniowych relacji pomiędzy metodami określania niepodobieństwa a opinią

człowieka, przynajmniej na podstawie analizy symulowanych danych rastrowych. Dotychczasowe wyniki dają ogólny zarys relacji miar niepodobieństwa do ludzkiej percepcji, w związku z czym w kolejnym etapie sporządzona została ankieta w pełni opierająca się o rzeczywiste dane rastrowe.

4.2 Wyniki drugiej ankiety

Rozdział 5

Porównanie wyników

Rozdział 6

Podsumowanie

Bibliografia

- Cha, SH (2007). Comprehensive Survey on Distance/Similarity Measures Between Probability Density Functions. *Int. J. Math. Model. Meth. Appl. Sci.* **1**.
- Chang, W, J Cheng, J Allaire, C Sievert, B Schloerke, Y Xie, J Allen, J McPherson, A Dipert, and B Borges (2022). *shiny: Web Application Framework for R*. R package version 1.7.4. <https://CRAN.R-project.org/package=shiny>.
- Comber, A, P Fisher, and R Wadsworth (2005). What is Land Cover? *Environment and Planning B: Planning and Design* **32**, 199–209.
- Forman, R (1995). *Land Mosaics: The Ecology of Landscapes and Regions*. Cambridge University Press. https://books.google.pl/books?id=sSRNU_5P5nwC.
- Gustafson, EJ (1998). Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems* **1**(2), 143–156.
- Haralick, RM, K Shanmugam, and I Dinstein (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**(6), 610–621.
- Jasiewicz, J, P Netzel, and TF Stepinski (2015). GeoPAT: A toolbox for pattern-based information retrieval from large geospatial databases. *Comput. Geosci.* **80**, 62–73.
- Kozak, J, M Luc, K Ostapowicz, and E Ziolkowska (2014). “Pozyskiwanie i analiza danych o pokryciu terenu a badania struktury przestrzennej krajobrazu”. In: pp. 63–84.
- Likert, R (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Lu, D, P Mausel, E Brondízio, and E Moran (2004). Change Detection Techniques. *International Journal of Remote Sensing* **25**.
- McGarigal, K, S Tagil, and S Cushman (2009). Surface metrics: an alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology* **24**, 433–450.

- Netzel, P and TF Stepinski (2015). Pattern-Based Assessment of Land Cover Change on Continental Scale With Application to NLCD 2001–2006. *IEEE Transactions on Geoscience and Remote Sensing* **53**(4), 1773–1781.
- Niesterowicz, J and T Stepinski (2016). On using landscape metrics for landscape similarity search. *Ecological Indicators* **64**, 20–30.
- Nowosad, J (2021). Motif: an open-source R tool for pattern-based spatial analysis. *Landscape Ecology* **36**, 29–43.
- Pukowiec-Kurda, K and M Sobala (2016). Nowa metoda oceny stopnia antropogenicznego przekształcenia krajobrazu na podstawie metryk krajobrazowych. *Prace Komisji Krajobrazu Kulturowego* **nr 31**, 71–84.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- Schlather, M, A Malinowski, P Menck, M Oesting, and K Strokorb (2015). Analysis, Simulation and Prediction of Multivariate Random Fields with Package Random Fields. *Journal of statistical software* **63**, 1–25.
- Sciaini, M, M Fritsch, C Scherer, and CE Simpkins (2018). NLMR and landscapetools: An integrated environment for simulating and modifying neutral landscape models in R. *Methods in Ecology and Evolution* **00**, 1–9.
- Solon, J (2002). *Ocena różnorodności krajobrazu na podstawie analizy struktury przestrzennej roślinności*. Prace Geograficzne - Polska Akademia Nauk. PAN IG i PZ. <https://books.google.pl/books?id=00dcAwAAQBAJ>.
- Trattner, J and L D'Agostino McGowan (2021). *shinysurveys: Create and Deploy Surveys in 'Shiny'*. R package version 0.2.0. <https://CRAN.R-project.org/package=shinysurveys>.