



Kierunek studiów: Geoinformacja

Nr albumu: 455828

**Tomasz Matuszek**

**Measuring impact of addition of Landsat 8  
thermal band on supervised land cover  
classification results**

*Ocena wpływu zastosowania kanału termalnego  
Landsat na wyniki nadzorowanej klasyfikacji pokrycia  
terenu*

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

Poznań, 2022



# 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

<b>Streszczenie</b>	<b>3</b>
<b>1 Introduction</b>	<b>5</b>
<b>2 Source data</b>	<b>7</b>
2.1 Satellite imagery . . . . .	7
2.2 Land cover data . . . . .	7
<b>3 Methods</b>	<b>11</b>
3.1 Machine learning . . . . .	11
3.2 R language environment . . . . .	12
<b>4 Result of the model - land cover map</b>	<b>13</b>
<b>5 Assessing model quality</b>	<b>15</b>
<b>6 Evaluation of thermal band's impact on prediction results</b>	<b>17</b>
<b>7 Conclusion</b>	<b>19</b>

# Rozdział 1

## Introduction

- applications and relevance of land cover maps
- machine learning and supervised classification of satellite images as a tool for creating land cover maps
- pointing out that thermal band is often omitted in land cover classification models, exact impact of thermal factor isn't fully clear
- goal of the thesis is to create land cover map of Poznań metropolitan area and measure the impact of thermal band on the model results

---

Wprowadzenie powinno mieć charakter opisu od ogółu do szczegółu (np. trzy-pięć paragrafów). Pierwszy paragraf powinien być najbardziej ogólny, a kolejne powinny przybliżać czytelnika do problemu. Przedostatni paragraf powinien określić jaki jest problem (są problemy), który praca ma rozwiązać i dlaczego jest to (są one) ważne.

Wprowadzenie powinno być zakończone stwierdzeniem celu pracy. Dodatkowo tutaj może znaleźć się również krótki opis co zostało zrealizowane w pracy.

Pisząc ten rozdział proszę pomyśleć o osobach, które zupełnie nie znają opisywanej tematyki. Należy tutaj krok po kroku wyjaśnić podstawowe koncepcje, istotność problemu,

wyniki poprzednich podobnych badań, itd. Ten rozdział obejmuje tylko kwestie, które już zostały wykonane przez inne osoby - nowe wyniki mają swoje miejsce w rozdziale **?@sec-wyniki**.

Każda kwestia opisana w tym rozdziale powinna być cytowana. Dodatkowo cytowania odbywa się poprzez uzupełnienie pliku `thesis.bib` zapisem w formacie BibTeX, a następnie dodanie nazwy referencji poprzedzonej znakiem `@`. Przykładowo, zacytowanie książki *Geocomputation with R* odbywa się poprzez (Lovelace et al., 2019).

W przypadku, gdy cytowanie zostało poprawnie wpisane oraz istnieje w pliku `thesis.bib` to bibliografia powinna się automatycznie wygenerować na końcu pracy.

W przypadku, gdy praca dyplomowa opisuje konkretny obszar to można po tym rozdziale stworzyć kolejny rozdział opisujący "obszar badań".

Ten i kolejne rozdziały mogą mieć także podrozdziały. Tworzenie podrozdziałów polega na stworzeniu nowej linii rozpoczynającej się od znaków `##` a następnie tytułu podrozdziału. Dodatkowo w postaci `{#sec - }` można dodać skrót nazwy rozdziału/podrozdziału umożliwiający odnoszenie się do niego używając operatora `[-@sec]`.

## **Rozdział 2**

# **Source data**

Two types of spatial data were used:

- satellite images from Landsat programme downloaded using GLAD ARD software (raster data)
- training points with information about land cover acquired from LUCAS programme (vector data)

### **2.1 Satellite imagery**

- preprocessing of satellite imagery performed by GLAD team from University of Maryland
- process of downloading data using GLAD ARD software
- GLAD ARD dataset structure

### **2.2 Land cover data**

- what is LUCAS programme and how its datasets are created
- LUCAS Grid and LUCAS Micro Data
- structure of LUCAS points dataset, description of relevant columns in addition to the land cover classification



**Rycina 2.1:** *Moja druga rycina*

- processing of data in order to choose reliable training points; reclassification to chosen set of land cover categories

---

Celem tego rozdziału jest przedstawienie użytych w pracy danych. Należy wyjaśnić jakie dane zostały użyte, jakiego są one rodzaju, dla jakiego okresu zostały pobrane/stworzone, co one zawierają, etc.

W tym rozdziale warto dodać ryciny i tabele przedstawiające użyte dane.

Zwróć uwagę, że poniższe bloki kodu mają parametr `#| echo: false`; oznacza to, że będą one niewidoczne w wynikowym pliku PDF. Każdy z bloków kodu musi mieć unikalną nazwę; w przypadku rycin powinna się ona rozpoczynać od prefiksu `fig-`. Dodanie podpisu pod rycinę odbywa się używając parametru `#| fig-cap:`. Następnie do tej ryciny można się odnieść używając operatora `[-@]`.

Podobnie wygląda odnoszenie się do plików graficznych. Tutaj wewnątrz bloku kodu należy użyć funkcji `knitr::include_graphics()` (Rycina 2.1). Dodatkowo możliwa jest zmiana rozmiaru obrazka używając parametrów takich jak `#| out-width: i #| out-height:`.



**Tabela 2.1:** *Moja pierwsza tabela*

	a	b
1	a	
2	b	
3	c	
4	d	
5	e	

Odnoszenie się do tabel odbywa się poprzez operator [-@] wraz z prefiksem `tbl-`. Natomiast tworzenie podpisu nad tabelą ma miejsce używając parametru `#| tbl-cap:`. Dodatkowo możliwe jest użycie pakietu **kableExtra** (Zhu, [2021](#)) do określenia szerokości kolumn (Tabela 2.1).



## **Rozdział 3**

# **Methods**

### **3.1 Machine learning**

- what is machine learning and what are its applications
- classification vs regression algorithms
- supervised and unsupervised classification

#### **3.1.1 Random forest algorithm**

- what is a decision tree
- how random forest algorithm works

#### **3.1.2 Model quality assessment**

- idea of resampling
- measures and indices of classification model quality

#### **3.1.3 Parameter tuning**

- what is tuning of model's parameters
- idea of nested resampling

## 3.2 R language environment

Short description of R and RStudio environment. List of used libraries and packages.

---

Rozdział **Metody** zawiera opis użytych metod (np. statystycznych czy geostatystycznych) oraz technologii (np. pakiety R). Opis każdej z metod czy technologii powinien być zwarty i zawierać tylko najważniejsze informacje z punktu widzenia pracy dyplomowej.

Każda użyta metoda i technologia powinna być zacytowana. W przypadku pakietów R, wystarczy wypełnić poniższy blok kodu (zwróć uwagę, że ten blok kodu ma parametr `echo: false`; oznacza to, że będzie on niewidoczny w wynikowym pliku PDF).

... a następnie zacytować pakiet używając znaku @, po którym podać nazwę pakietu rozpoczynającą się od prefiksu R-. Przykładowe cytowanie języka R bez nawiasu to R Core Team (2021), a pakietu **kableExtra** w nawiasie to (Zhu, 2021). Więcej przykładów cytowania można znaleźć na stronie [https://rmarkdown.rstudio.com/authoring\\_bibliographies\\_and\\_citations.html#citations](https://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html#citations).

W przypadkach, gdy cytowanie istnieje, ale nie jest pakietem R to należy dodać je do pliku `thesis.bib` i użyć powyższej składni ze znakiem @. W ostateczności, gdy dana technologia nie posiada cytowania, należy podać jej adres internetowy.

## Rozdział 4

# Result of the model - land cover map

- land cover map of Poznań metropolitan area
- probability map of model results

---

Część **Wyniki** może składać się z jednego lub więcej rozdziałów. Każdy z tych rozdziałów powinien mieć tytuł adekwatny do swojej treści.

Rozdziały wynikowe powinny korzystać z wiedzy opisanej w poprzednich rozdziałach (Rozdziały **?@sec-lit**, **?@sec-dane**, **?@sec-metody**). W przypadku prac analitycznych, ich treść powinna przedstawiać kolejne etapy eksploracji i analizy danych. W przypadku prac technicznych, treść tych rozdziałów powinna opisywać stworzone narzędzia, a następnie pokazywać ich zastosowanie/a.

W przypadku prac technicznych warto pokazywać fragmenty napisanego rozwiązania lub jego wywołania używając bloków kodu.

```
moja_funkcja = function(x){  
  cat(x, "rządzi!")  
}  
moja_funkcja("Autor tej pracy")
```

Autor tej pracy rządzi!

## Rozdział 5

# Assessing model quality

Table with quality indices:

- overall accuracy (OA)
  - classification error (CE)
  - producer's and user's accuracy (PA, UA)
  - Kappa coefficient
- 

Część **Wyniki** może składać się z jednego lub więcej rozdziałów. Każdy z tych rozdziałów powinien mieć tytuł adekwatny do swojej treści.

Rozdziały wynikowe powinny korzystać z wiedzy opisanej w poprzednich rozdziałach (Rozdziały **?@sec-lit**, **?@sec-dane**, **?@sec-metody**). W przypadku prac analitycznych, ich treść powinna przedstawiać kolejne etapy eksploracji i analizy danych. W przypadku prac technicznych, treść tych rozdziałów powinna opisywać stworzone narzędzia, a następnie pokazywać ich zastosowanie/a.

W przypadku prac technicznych warto pokazywać fragmenty napisanego rozwiązania lub jego wywołania używając bloków kodu.

```
moja_funkcja = function(x){  
  cat(x, "rządzi!")  
}  
moja_funkcja("Autor tej pracy")
```

Autor tej pracy rządzi!



## Rozdział 6

# Evaluation of thermal band's impact on prediction results

- mean temperature for every predicted land cover class
- variable importance plots, variable profiles
- thermal band importance map (two methods: raster aggregation and interpolation of importance in LUCAS points)
- mean importance of thermal band on each land cover class
- difference raster map between prediction with and without thermal band included, transition matrix

---

Część **Wyniki** może składać się z jednego lub więcej rozdziałów. Każdy z tych rozdziałów powinien mieć tytuł adekwatny do swojej treści.

Rozdziały wynikowe powinny korzystać z wiedzy opisanej w poprzednich rozdziałach (Rozdziały **?@sec-lit**, **?@sec-dane**, **?@sec-metody**). W przypadku prac analitycznych, ich treść powinna przedstawiać kolejne etapy eksploracji i analizy danych. W przypadku

prac technicznych, treść tych rozdziałów powinna opisywać stworzone narzędzia, a następnie pokazywać ich zastosowanie/a.

W przypadku prac technicznych warto pokazywać fragmenty napisanego rozwiązania lub jego wywołania używając bloków kodu.

```
moja_funkcja = function(x){  
  cat(x, "rządzi!")  
}  
moja_funkcja("Autor tej pracy")
```

Autor tej pracy rządzi!

## Rozdział 7

### Conclusion

- land cover map of Poznań metropolitan area was created, impact of thermal band on classification results was measured
- despite thermal band having low overall impact on model results, there is a strong spatial auto-correlation for its importance
- land surface temperature was especially significant for land cover classification of urban areas, it helped in identify built-up areas
- it may mean that thermal band will become increasingly important in studies on urban sprawl and suburbanisation
- better land cover maps will help in better management of metropolitan areas growth and quantifying impact of urbanisation on natural environment more precisely

---

Podsumowanie pracy jest w pewnym sensie znacznie rozbudowanym abstraktem. Należy wyliczyć i opisać osiągnięcia uzyskane w pracy dyplomowej. Tutaj jednak (w przeciwieństwie do np. rozdziału **?@sec-wprowadzenie**) należy przechodzić od szczegółu do ogółu - co zostało stworzone/określone, jak zostało to zrobione, jakie ma to konsekwencje, itd.

Ten rozdział powinien też zawierać opis kwestii, których nie udało się rozwiązać w pracy dyplomowej (i dlaczego się nie udało) oraz pomysły na przyszłe ulepszenie uzyskanych wyników lub dalsze badania.

# Bibliografia

Lovelace, R, J Nowosad, and J Muenchow (2019). *Geocomputation with R*. CRC Press.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.

Zhu, H (2021). *kableExtra: Construct Complex Table with kable and Pipe Syntax*. R package version 1.3.4. <https://CRAN.R-project.org/package=kableExtra>.