Universiteit Leiden
The Netherlands

A Meta-Analysis of Machine Learning in Remote Sensing

Nina Leach

Defended on 2024-06-12

MASTER THESIS

STATISTICS AND DATA SCIENCE

UNIVERSITEIT LEIDEN

Table of contents

Abstract

IntroThe United Nations SDG report (2023) emphasises embracing new data sources and innovative methodologies, including remote sensing. To leverage these data sources and complement traditional survey data, machine learning can be applied. This growing field of research has an extensive body of literature, so selecting a suitable machine learning method poses a significant challenge. The following is a research proposal for a systematic review and meta-analysis to assess the variability of quality metrics and identify study features that may impact the performance of machine learning methods in remote sensing applications.

Methods

Results

Discustion

Chapter 1

Introduction

The Sustainable Development Goals (SDGs) are framework of 17 objectives aimed at addressing global challenges, including poverty, pollution, climate change, and the loss of biodiversity. These goals were established by the United Nations in 2015 with the aim of accomplishing them by 2030. However, as we surpass the midpoint of the SDGs' timeline, it is evident that significant setbacks have limited progress toward these goals (UN DESA 2023). Therefore, effective monitoring is crucial to the provide insights needed to understand these setbacks, measure progress accurately, and identify where interventions are the most needed (Burke et al. 2021). Traditional approaches, such as land surveying and household surveys are rich in detail but are costly, sparsity collected, and difficult to verify (Burke et al. 2021). Recognizing these problems, the United Nations emphasizes the importance of adopting new data sources and methodologies, such as remote sensing, machine learning, crowdsourcing, and citizen-generated data, to overcome these challenges and advance toward the 2030 targets(UN DESA 2023).

In contrast to the scarcity of ground-based data, the quantity and quality of data collected through remote sensing via satellites, aircraft, or drones is increasing (Burke et al. 2021). Remote sensing provides a cost-effective method for monitoring extensive geographic areas Burke et al. (2021). The potential applications are numerous. Examples include monitoring of land use and degradation, forestry, biodiversity, agriculture, disaster prediction, water resources, public health, urban planning, poverty, and the management and preservation of world heritage sites [Chen et al. (2023); Khatami, Mountrakis, and Stehman (2016a); Zhang, Liu, and Shen (2022); Thapa et al. (2023); Yin et al. (2023); Holloway and Mengersen (2018); Machicao et al. (2022);Burke et al. (2021)].

The rapid availability of remote sensing data coincides with the rapid development and availability of machine learning (ML) algorithms (Shi et al. 2022). The magnitude of possible applications and increase of availability of diverse data sources and methodology has lead to a rapid increase in amount of published research papers in this field (Khatami, Mountrakis, and Stehman 2016b; Burke et al.

Table 1.1: List and a brief summary of previous reviews related to the topic.

Author.Year	Title
Thapa et al.,	Deep Learning for Remote Sensing Image Scene Classification: A Review and Meta-A
Lavallin et al.,	Machine learning in geography–Past, present, and future
Khatami et al.,	A meta-analysis of remote sensing research on supervised pixel-based land-cover imag
Jafarzadeh et al.,	Remote Sensing and Machine Learning Tools to Support Wetland Monitoring: A Me
Anshuka et al.,	Drought forecasting through statistical models using standardised precipitation index
Hanadé Houmma et al.,	Modelling agricultural drought: a review of latest advances in big data technologies
Zhao et al.,	An Overview of the Applications of Earth Observation Satellite Data: Impacts and F
Chen et al.,	Bibliometric Analysis of Spatial Technology for World Heritage: Application, Trend a
Hall et al.,	A review of machine learning and satellite imagery for poverty prediction: Implicatio
Burke et al.,	Using satellite imagery to understand and promote sustainable development
Mukonza et al.,	Meta-Analysis of Satellite Observations for United Nations Sustainable Development
Holloway et al.,	Statistical Machine Learning Methods and Remote Sensing for Sustainable Developm
Eskandari et al.,	Meta-analysis of Unmanned Aerial Vehicle (UAV) Imagery for Agro-environmental M
Shi et al.,	Variability and uncertainty in flux-site-scale net ecosystem exchange simulations base

2021). However, most individual studies are limited to a single geographical area and have very specific focus. As there is generally more incentive for researchers to make innovative methods rather reproduce existing results or apply validated methods to various locations or contexts (Burke et al. 2021). This localized approach makes it difficult to draw general guidelines for selecting the most suitable processes to achieve high accuracy for various contexts (Burke et al. 2021; Shi et al. 2022).

1.1 Previous research

Studies have previously examined the application of remote sensing for SDG monitoring (see Yin et al. (2023); Holloway and Mengersen (2018);Burke et al. (2021)). These reviews typically focus on one context (e.g. satellite imagery for poverty prediction (Hall et al. 2023) or focus on descriptive results (e.g. (Yin et al. 2023)). Some meta-analysis make comparisons to attribute accuracy improvements to individual features (see: Khatami, Mountrakis, and Stehman (2016a)).

there is a significant gap in the literature.....

1.2 The present study

The present work is a review systematically assessing the variability of machine learning methods in SDGs monitoring applications, with a specific focus on identifying study features that may impact their performance. Review studies are important for creating a comprehensive overview, assessing the accuracy of different methods, as many factors or study features can explain heterogeneity of quality metrics among individual studies. For example, the data collecting device (e.g. satellite), image type, algorithm applied, input features, pre- and post-processing techniques, the use of ancillary data.

With is in mind the aim of this review is to recreate a guideline with universal applicability to improve model accuracy though the election of appropriate algorithms

- systematic review application of machine learning methods in remote sensing, particularly focusing on research related to the SDGs
- Explaining variability of observed accuracy (quality metrics) across different remote sensing applications. with a meta-analysis
- develop and apply a meta-regression model to identify study features (meta-features) that explain the variation in classification quality.

1.3

Chapter 2

Background?

our world in data SDGs tracker: https://ourworldindata.org/sdgs

Maybe discuss here the shortage of information around the world: Burke et al.,

2.1 Remote sensing

TO DO:

- Add some background information
- suggestions for interesting features
- explain features (bands etc)

2.2 Modeling approaches

Simple models possible however because of the nature of RS data…

indexes: computed from imagery: proportions of different wave lengths examples of non machine learning approaches: crop_5

Machine Learning

Machine Learning is

why its needed in remote sensing?

types of machine learning:

possible grouping

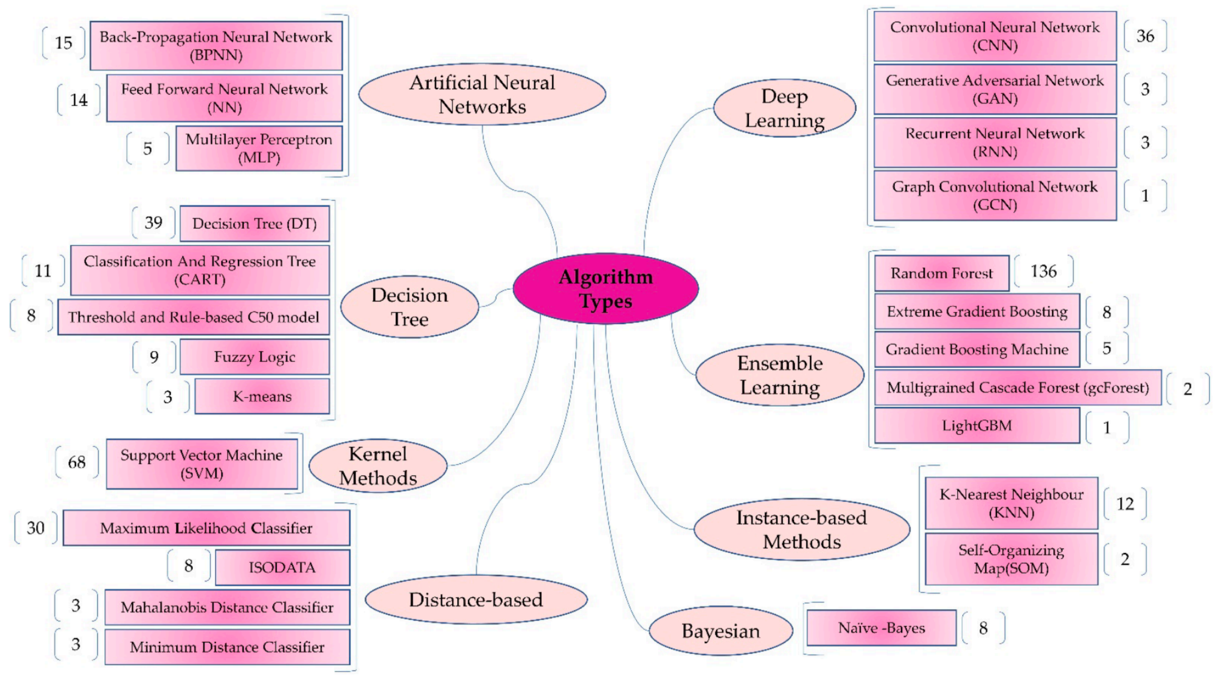


Figure 16. Number of studies employed different ML models in wetland studies.

Assessing model performance

- describe some performance metrics

2.3

Possible Additions:

- table of previous reviews in related topics

Chapter 3

Methods

This meta-analysis was conducted and reported based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al. 2021). Citation manager: Zotero and analysis using `r version$version.string`.

Eligibility criteria & Search strategy

The data collected for this report was extracted from peer-reviewed articles, published between January 2018 and December 2023. I gathered these articles from several academic research databases including the Web of Science (WOS), and ScienceDirect (see Figure 3.1 for the full list) on January 15 and 16, 2024. The search terms where “Remote sensing” AND “machine learning” AND “sustainable development goals”. The search results from the databases were downloaded to a RIS file and imported to Zotero. Duplicate articles were dealt with using Zotero’s merge duplicates function. A total of 811 articles remained after review articles and non-research articles were removed

(I will link to appendix for more info).

Selection process

A random sample of 200 articles was drawn for title and abstract screening. These potentially relevant articles were screened independently by three reviewers (). Using the R package ‘metagear‘().

The studies were selected according to the following criteria:

- publications utilizing remote sensing and machine learning techniques
- quality assessment in the from of (RMSE, Overall accuracy, ….)

From the 200 papers 57 were deemed potentially relevant by all three reviewers. [TO DO: discuss what we will do about the disputed articles].

The first 10 papers (of 57) were inspected, it was decided to start the focus on papers related to classification techniques. I screened the titles and abstracts of the 57 articles using ‘metagear‘ categorizing them to classification (40) or regression papers. Furthermore, Overall Accuracy (OA) was the most commonly reported outcome metric and therefore it was decided to include all papers that report (OA).

Data collection process & Data items

Using these papers and previous systematic reviews a list of potential study features was made and structured in a table for data collection. Various aspects of the articles were collected, including: general study details such as authors, publication year, the study objectives, the data source. Table 4.1 outlines all the extracted features.

In cases where multiple ML algorithms, data sources, etc where used each of the results were reported.

Data for all articles was extracted by the author, in cases where the features were ambiguous these were discussed....

Study risk of bias assessment

- - TO DO: look at: “utilize methodological quality and bias assessment tool (CLEAR-NPT, Coleman, Modified Coleman, CONSORT, Pedro, Cochrane, Delphi, Detsky, Downs and Black, Jadad, Level of evidence, MINORS, Newcastle-Ottawa, QUADAS, Quality Appraisal Tool, STARD, Strobe, AMSTAR, R-AMSTAR, etc.)”?
  - https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0234722#pone.0234722.s004
  - https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-023-01849-0

Effect measures & Synthesis Methods

Overall accuracy (OA) is defined as the probability of a correct classification (Magnussen 2021)

$$\text{Overall Accuracy (OA)} = \frac{\text{True Positive} + \text{True Negative}}{N}$$

proportion there for the standard error is

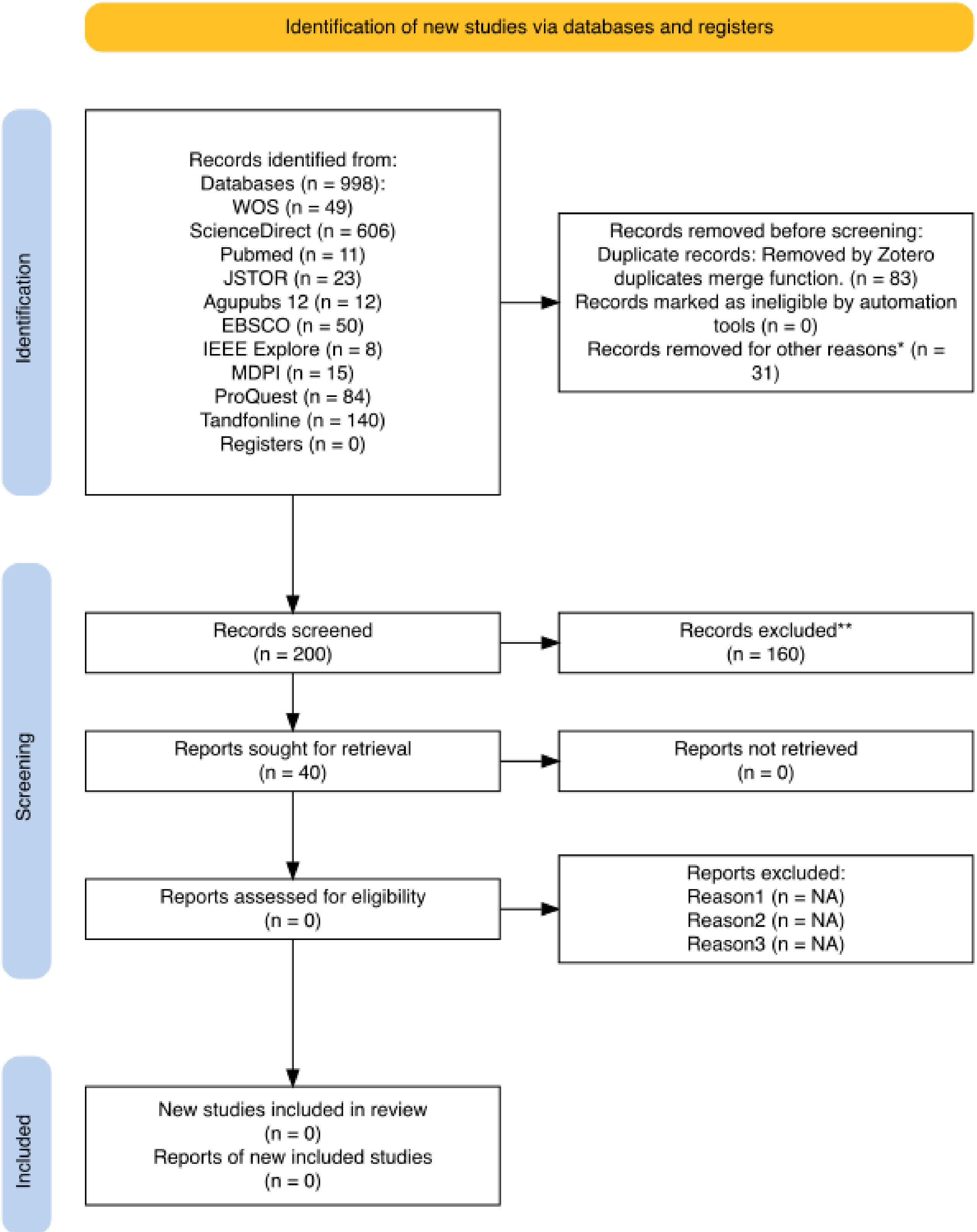


Figure 3.1: PRIMSA flow diagram for manuscript selection. Note: *number of records removed four where not journal articles and 27 were omitted for being reviews. A random sample of 200 of the total 884 was drawn and reviewed by three independent reviewers. A total of 57 records were left, 40 of which were deemed to be classification papers. WOS: Web of Science

standard error:

$$SE_p = \sqrt{\frac{p(1-p)}{N}} \text{Where p is probability of a correct classification (OA)}$$

Meta analysis

RQ: Is there heterogeneity between studies:

"It is highly advised to specify the type of model you used in the methods section of your meta-analysis report."

- expect considerable between study heterogeneity
  - use random-effects model to pool the effect size
  - TO DO: choose between methods to estimate variance τ^2
  - using {meta} package: metaprop for proportions:
    * logit-transform proportions before the meta-analysis is performed
    * apply a (generalized) logistic mixed-effect model to estimate the pooled effect
- measures of heterogeneity, forest plot
- .....

Meta-Regression model

RQ: What factors affect the probability of correct classification (Overall accuracy) of machine learning algorithms?

Mixed effects model:

$$\hat{\theta}_k = \theta + \beta_1 x_{1k} ... + \beta_n x_{nk} + \epsilon_k + \zeta_k$$

Where ϵ_k sampling error deviation from the true effect size by the study ζ_k random effect... x_i are the explanatory variables (or features):....

TO DO:

- rewrite as a probit model
- write about model fit assessments, choosing features (step-wise, or other) etc
- And which assumptions to check

Reporting bias assessment

Certainty assessment

Chapter 4

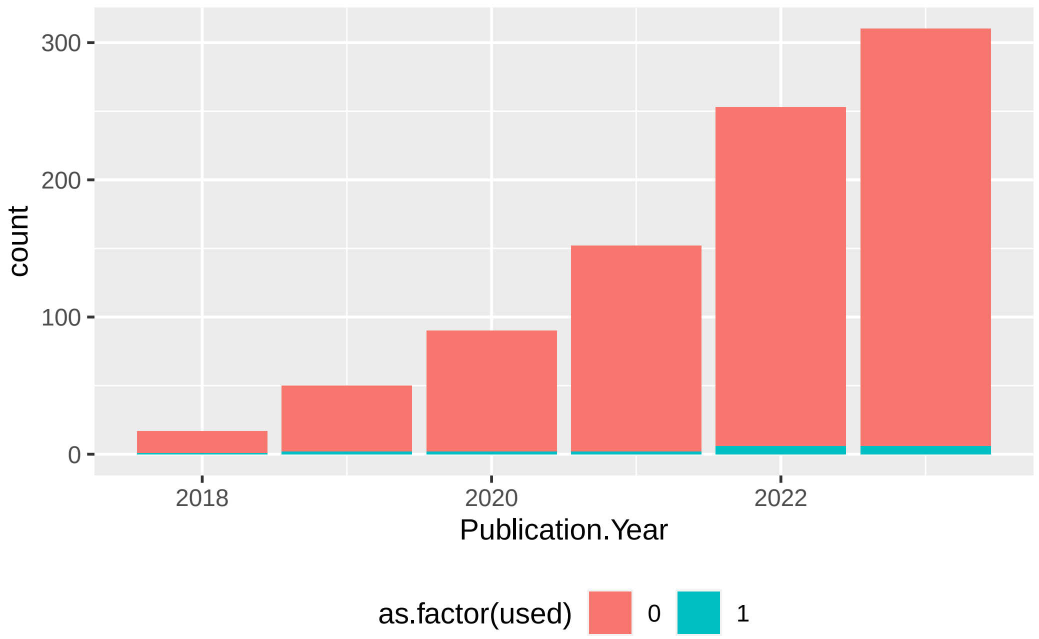
Results

```
# Packages
## TO DO: see if these can be moved to the book yaml
library(tidyverse)
library(readxl)
```

```
# Data
# All potential papers from the database search with duplicates removed by Zotero
citations<- read.csv("appendix/app1-paper_selection/citations.csv")

# Merging first auther and publication year
citations$Author.Year <- paste(sub(",.*", "", citations$Author), "et al.,", citations$Publicat

# Extracted data from selected papers
extracted_features <- read_excel("extracted_features.xlsx", sheet = "Include",  na = "NA")

# for the plot to distinugish the papers used and not used
citations$used <- ifelse(citations$DOI %in% extracted_features$DOI, 1, 0)

# Needed paper data for the extracted
citations_needed <- subset(citations, DOI %in% extracted_features$DOI,
                           select = c(DOI, Author.Year, Publication.Title))

# final dataset with extracted and article features
dat <- inner_join(citations_needed, extracted_features, by = "DOI")
```

Distribution of papers by publication year for the potential papers

```
# publication over time
ggplot(citations)+
  geom_bar(mapping = aes(x = Publication.Year, fill = as.factor(used))) +

  theme(legend.position = "bottom")
```




```
# TO DO: add caption and tilte
```

Table 4.1: Description of extracted features from the inclused papers.

feature	type	value/categories
Study ID	DOI used	(I will split these)
Auther		
Year		
Publication title		
research_theme	text	
RS_device	text/classes	satellite , drone, plane, Unmanned Aerial Vehicles (UAVs)
RS_device_specifics	text/classes	Sentinel-2, Landsat-8
ancillary	numeric	{0 = only remote sensed data, 1= additional data used} Additional (non- remote sensed data used)

feature	type	value/categories
type_	class	active vs optical
bands	class	list of the bands used
RS_spectral_bands_no	numeric	
RS_spatital_resolution_m	numeric	
model_group	class	
number_classes	numeric	
fraction_majority_class	numeric	
"sample" total	numeric	
OA	numeric	

Study characteristics

- table of article included (I think this will go in the appendix)
- total number papers
- distributions of features
- maybe pair plot to show associations between features

```
unique(dat$location)
```

```
 [1] "Ghana"        "Iran"         "Italy"        "South Africa" "Morocco"
 [6] "USA"          "Malaysia"     "Afghanistan"  "china"        "India"
[11] "East Asia"    "Kuwait"       "Mexico"       "Indonesia"    "China"
[16] "Australia"    "Pakistan"     "Saudi Arabia"
```

```
unique(dat$Author.Year)
```

```
 [1] "Aja et al., 2022"                "Pareeth et al., 2019"
 [3] "Agrillo et al., 2021"            "Mashaba-Munghemezulu et al., 2021"
 [5] "Ismaili et al., 2023"            "Tao et al., 2019"
 [7] "Shaharum et al., 2020"           "Jochem et al., 2018"
 [9] "Xia et al., 2022"                "Priyanka et al., 2023"
[11] "Wang et al., 2023"               "Al-Dousari et al., 2023"
[13] "Solórzano et al., 2023"          "Suryono et al., 2022"
```

```
[15] "Shen et al., 2023"                    "Owers et al., 2022"
[17] "Tang et al., 2022"                    "Nazir et al., 2020"
[19] "Fallatah et al., 2022"
```

Meta-Analysis Results:

- pooled effect sizes and confidence intervals.
- maybe with forest plots to visualize the individual study estimates (I have multiple results from one paper so maybe not the best method)
- overall effect size
- degree of heterogeneity

Meta-Regression

model results

- results of the model, estimated coefficients, etc (table)
- what features to add in the model
- interpretation of results in the table
- testing of assumptions (link to appendix?)

model fit

- predictions vs observations
- make table of the features and the predicted outcome (Klingwort&Toepoel)

Reporting biases

- funnel plots? TO DO: look at what is appropriate

Certainty of evidence

Chapter 5

Discussion

- Recap of main findings.
- Interpretation
- Comparison with existing literature.
- Limitations
- Implications: practical applications.
- Future Directions

5.1 Conclusion

5.2

References

Burke, Marshall, Anne Driscoll, David B. Lobell, and Stefano Ermon. 2021. "Using Satellite Imagery to Understand and Promote Sustainable Development." *Science* 371 (6535): eabe8628. https://doi.org/10.1126/science.abe8628.

Chen, Guolong, Ruixia Yang, Xiangli Zhao, Lanyi Li, Lei Luo, and Honghao Liu. 2023. "Bibliometric Analysis of Spatial Technology for World Heritage: Application, Trend and Potential Paths." *Remote Sensing* 15 (19): 4695. https://doi.org/10.3390/rs15194695.

Hall, Ola, Francis Dompae, Ibrahim Wahab, and Fred Mawunyo Dzanku. 2023. "A Review of Machine Learning and Satellite Imagery for Poverty Prediction: Implications for Development Research and Applications." *Journal of International Development* 35 (7): 1753–68. https://doi.org/10.1002/jid.3751.

Holloway, Jacinta, and Kerrie Mengersen. 2018. "Statistical Machine Learning Methods and Remote Sensing for Sustainable Development Goals: A Review." *Remote Sensing* 10 (9): 1365. https://doi.org/10.3390/rs10091365.

Khatami, Reza, Giorgos Mountrakis, and Stephen V. Stehman. 2016a. "A Meta-Analysis of Remote Sensing Research on Supervised Pixel-Based Land-Cover Image Classification Processes: General Guidelines for Practitioners and Future Research." *Remote Sensing of Environment* 177 (May): 89–100. https://doi.org/10.1016/j.rse.2016.02.028.

———. 2016b. "A Meta-Analysis of Remote Sensing Research on Supervised Pixel-Based Land-Cover Image Classification Processes: General Guidelines for Practitioners and Future Research." *Remote Sensing of Environment* 177 (May): 89–100. https://doi.org/10.1016/j.rse.2016.02.028.

Machicao, J., A. Ben Abbes, L. Meneguzzi, P. L. P. Corrêa, A. Specht, R. David, G. Subsol, et al. 2022. "Mitigation Strategies to Improve Reproducibility of Poverty Estimations From Remote Sensing Images Using Deep Learning." *Earth and Space Science* 9 (8): e2022EA002379. https://doi.org/10.1029/2022EA002379.

Magnussen, Steen. 2021. "Calibration of a Confidence Interval for a Classification Accuracy." *Open Journal of Forestry* 11 (1): 14–36. https://doi.org/10.4236/ojf.2021.111002.

Page, Matthew J, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, et al. 2021. "The PRISMA 2020 Statement: An Updated

Guideline for Reporting Systematic Reviews." *BMJ*, March, n71. https://doi.org/10.1136/bmj.n71.

Shi, Haiyang, Geping Luo, Olaf Hellwich, Mingjuan Xie, Chen Zhang, Yu Zhang, Yuangang Wang, et al. 2022. "Variability and Uncertainty in Flux-Site-Scale Net Ecosystem Exchange Simulations Based on Machine Learning and Remote Sensing: A Systematic Evaluation." *Biogeosciences* 19 (16): 3739–56. https://doi.org/10.5194/bg-19-3739-2022.

Thapa, Aakash, Teerayut Horanont, Bipul Neupane, and Jagannath Aryal. 2023. "Deep Learning for Remote Sensing Image Scene Classification: A Review and Meta-Analysis." *Remote Sensing* 15 (19): 4804. https://doi.org/10.3390/rs15194804.

UN DESA. 2023. *The Sustainable Development Goals Report 2023: Special Edition*. The Sustainable Development Goals Report. United Nations. https://doi.org/10.18356/9789210024914.

Yin, Chun, Ningyezi Peng, Yuchen Li, Yuanyuan Shi, Shujuan Yang, and Peng Jia. 2023. "A Review on Street View Observations in Support of the Sustainable Development Goals." *International Journal of Applied Earth Observation and Geoinformation* 117 (March): 103205. https://doi.org/10.1016/j.jag.2023.103205.

Zhang, Yuzhen, Jingjing Liu, and Wenjuan Shen. 2022. "A Review of Ensemble Learning Algorithms Used in Remote Sensing Applications." *Applied Sciences* 12 (17): 8654. https://doi.org/10.3390/app12178654.