# Integrating spatial heterogeneity to enhance spatial temporal crop yield predictions

Name: Francis Michael Msangi
Course: MSc Geospatial technologies
Student number: 542565
Email: fmsangi@uni-muenster.de

Supervisors:
Main: Prof. Dr. Edzer Pebesma- IFGI
Co-supervisor: Prof.Dr. Ana C. Costa- NOVA(IMS)
Co-supervisor: Dr. Francis Muthoni (IITA)

NOVA IMS Information Management School
Universität Münster
UNIVERSITAT JAUME I
Erasmus Mundus
Funded by the European Union

# Presentation outline

01. Motivation and Introduction

02. Objective and research question

03. Study area and Data

04. Methodology

05. Results

06. Conclusion

01

# Motivation and Introduction

# Motivation

- Spatial mapping and monitoring of crop yields is crucial for supporting decision making and ensuring food security
- Machine learning(ML), GIS and remote sensing have been integrated to make spatial mapping possible.
- However, in the prediction of crop yields the common ML algorithms often overlook the spatial heterogeneity inherent in landscapes leading to suboptimal estimations

02

# Objective & Research questions

# Main Objective

- The main objective of this research is to improve the accuracy and reliability of spatially explicit crop yield predictions in Zambia and Malawi by addressing spatial heterogeneity and effectively determining the areas in which predictions are reliable

# Research Questions

- Can addressing spatial heterogeneity by applying GWRF trained under target- oriented cross-validation strategy enhance spatial-temporal crop yield predictions?

- How can estimating the area of applicability of crop yield prediction models contribute to the effective extrapolation of agricultural technologies in Zambia and Malawi?

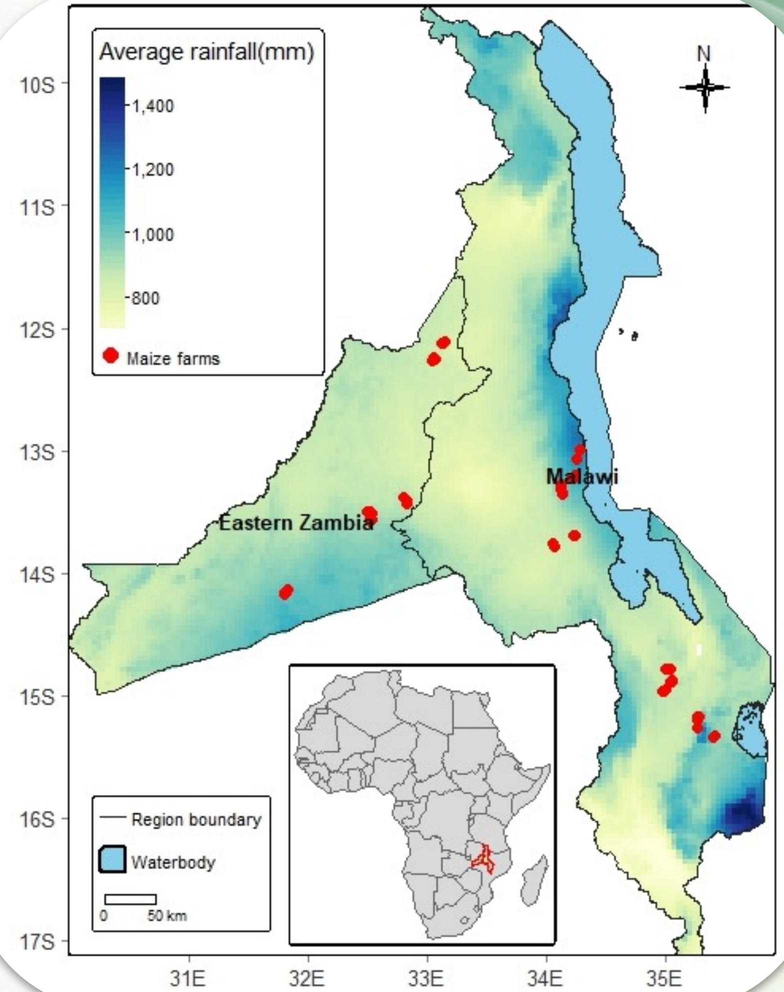

03

# Study area & Data

# Study area

Details

- located in Southern Africa
- Area coverage of approximately 170,000 square kilometres
- Characterized by uni-modal rainfall that spans from October to April



# Data

## Maize Crop yield data

Data collected in Malawi and Zambia with temporal resolution 2008/2009 to 2019/2020 season.
Divided into two groups Conservation agriculture(CA) farms and Conventional practices(CP)

## Remote sensing variables

Environmental conditions for the growing season
Vegetation productivity
Soil information
Terrain information
Socio-economic variables

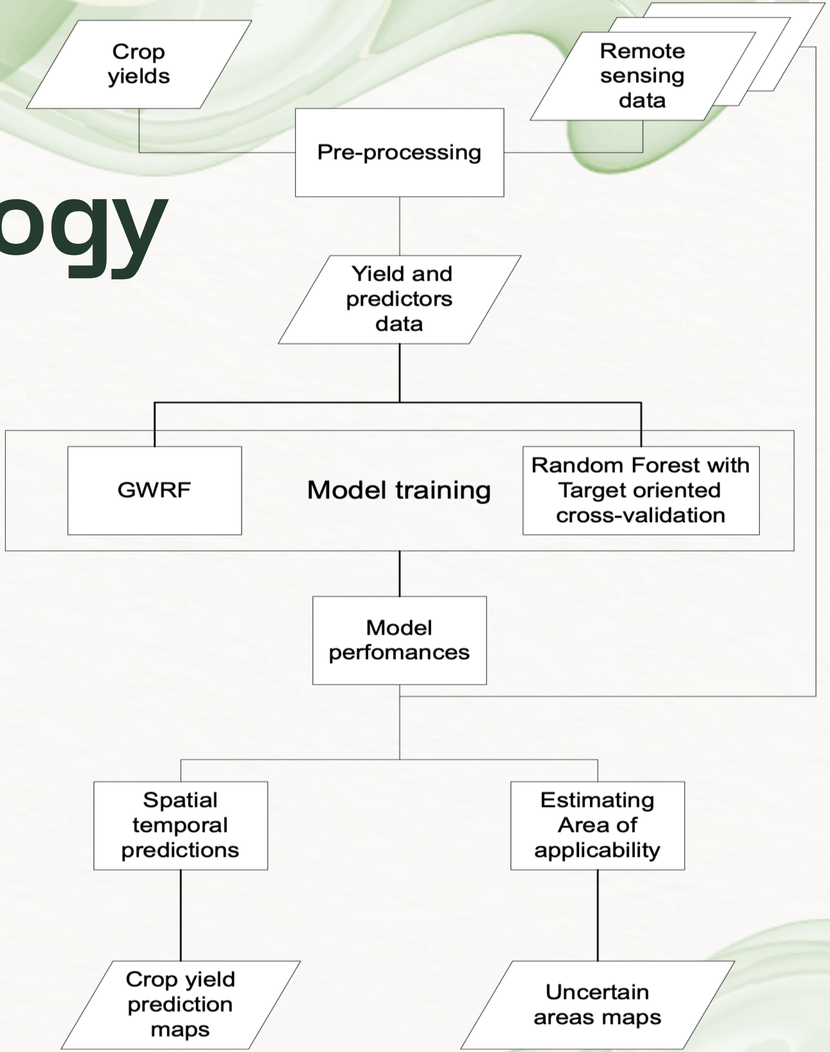
# Summary of remote sensing variables

| Environmental conditions (Growing season) | Soil variables | Socio-economic variables | Vegetation productivity | Terrain Variables |
|---|---|---|---|---|
| Rainfall | Total Nitrogen, Soil Organic carbon, Bulk density, Cation exchange capacity, pH | Cattle density | Enhanced vegetation index | Digital Elevation model |
| Temperature(Maximum and Minimum temperatures) | Extractable Boron, Aluminum, Zinc, calcium, sodium, potassium, | Market access | Absorbed photosynthetically Active Radiation(FPAR) | |
| Actual evapotranspiration | Soil Texture, Clay content, Silt content, Sand content | | | |

03

# Methodology

# Methodology workflow




KEY

**GWRF**- Geographically weighted random forest

Target oriented cross validation strategy- Environmental blocking- 4 clusters using K-means

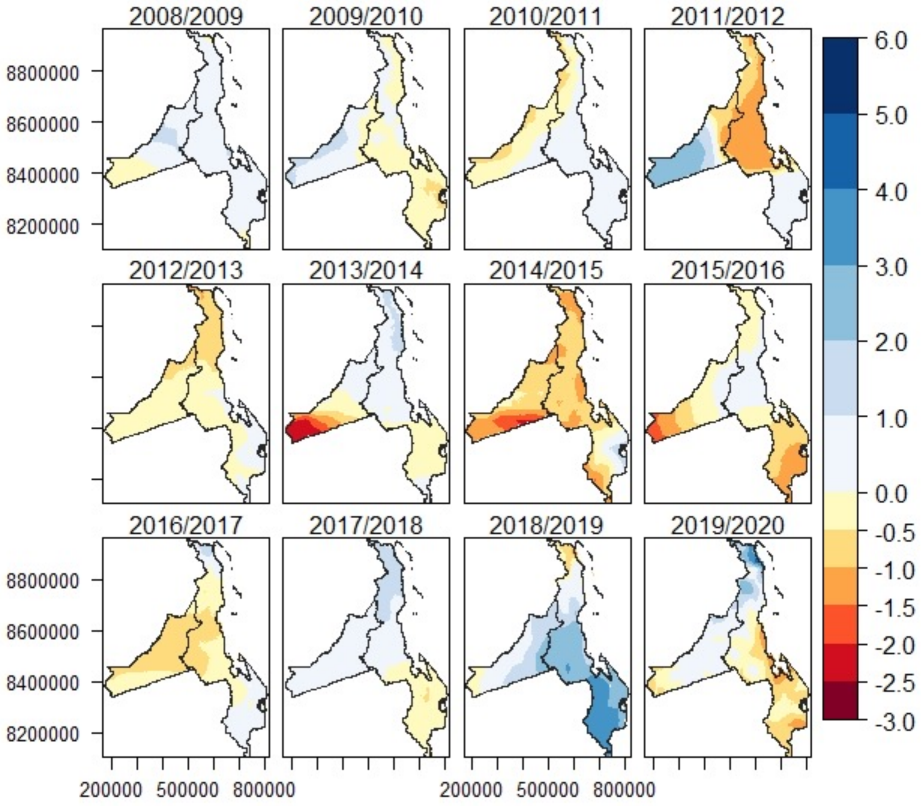
**CA**-Conservation agriculture
**CP**-Conventional practice

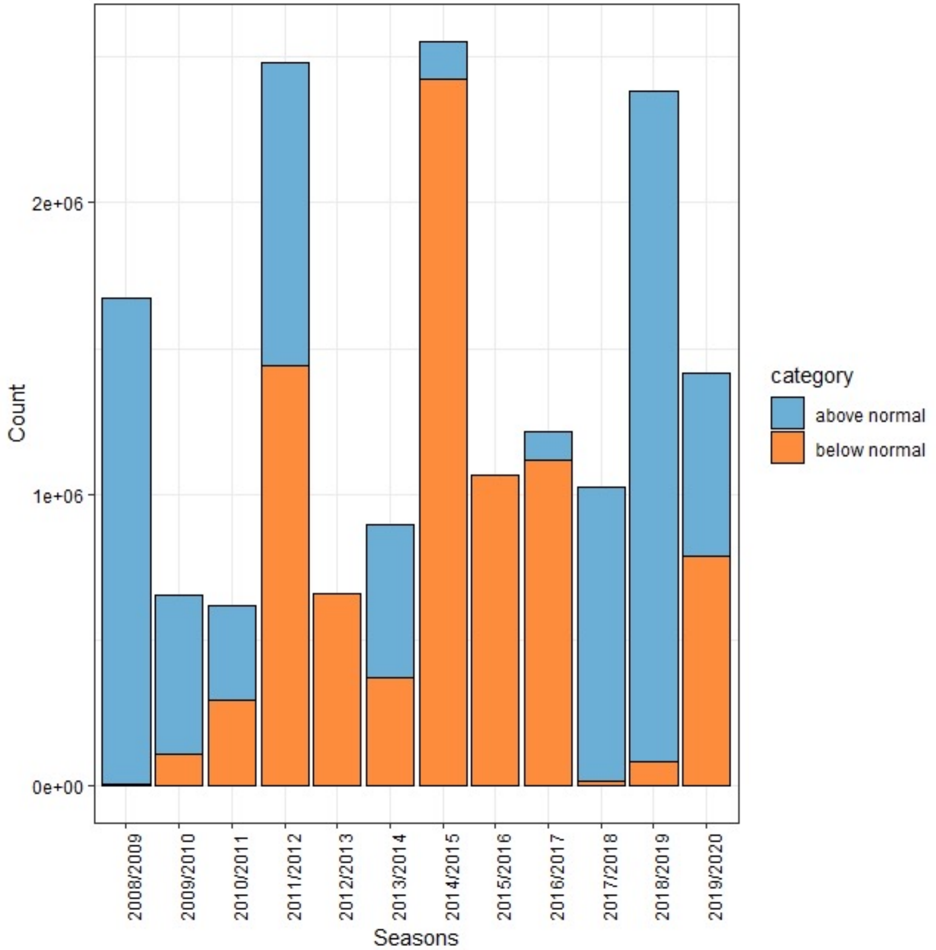
**GWRF shows better performances**

| | Model | RMSE | $R^2$ |
|---|---|---|---|
| CP | RF | 1409.902 | 0.013 |
| | **GWRF** | **1389.206** | **0.234** |
| CA | RF | 1547.705 | 0.037 |
| | **GWRF** | **1587.731** | **0.171** |

# Selecting seasons for spatial predictions

Standerdized rainfall anomalies



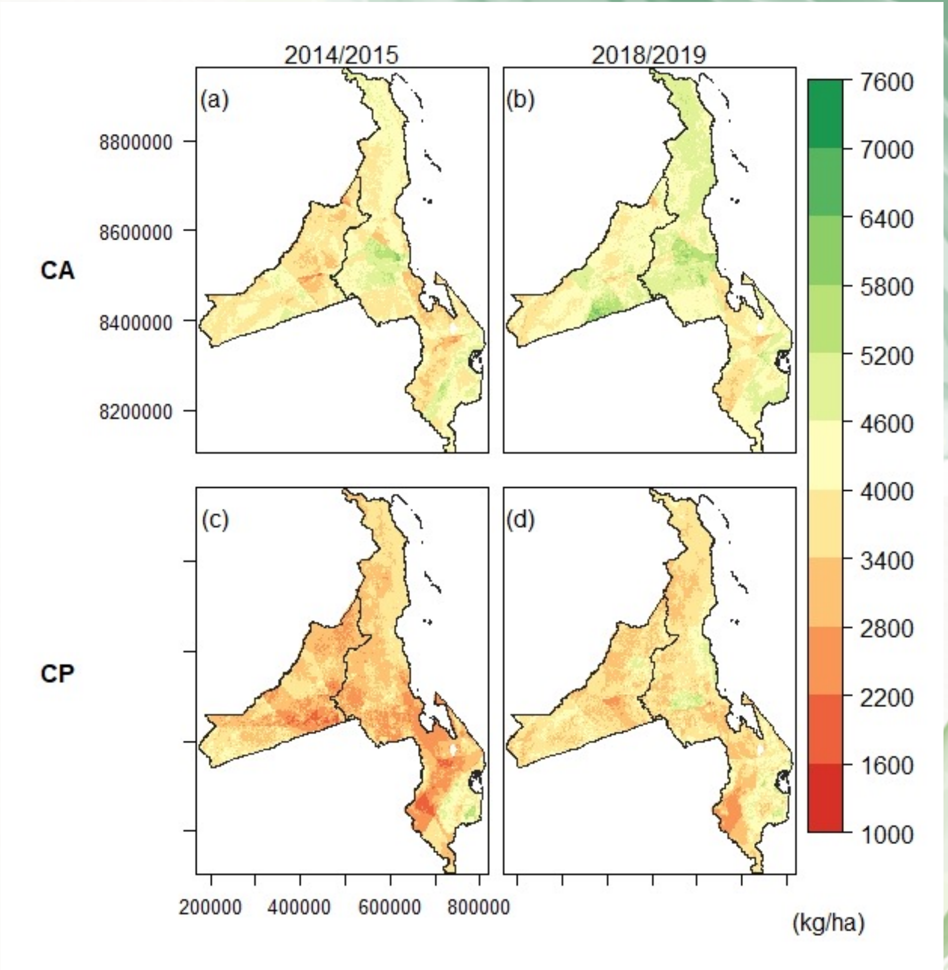

05

# Results

# Spatial temporal predictions

Details

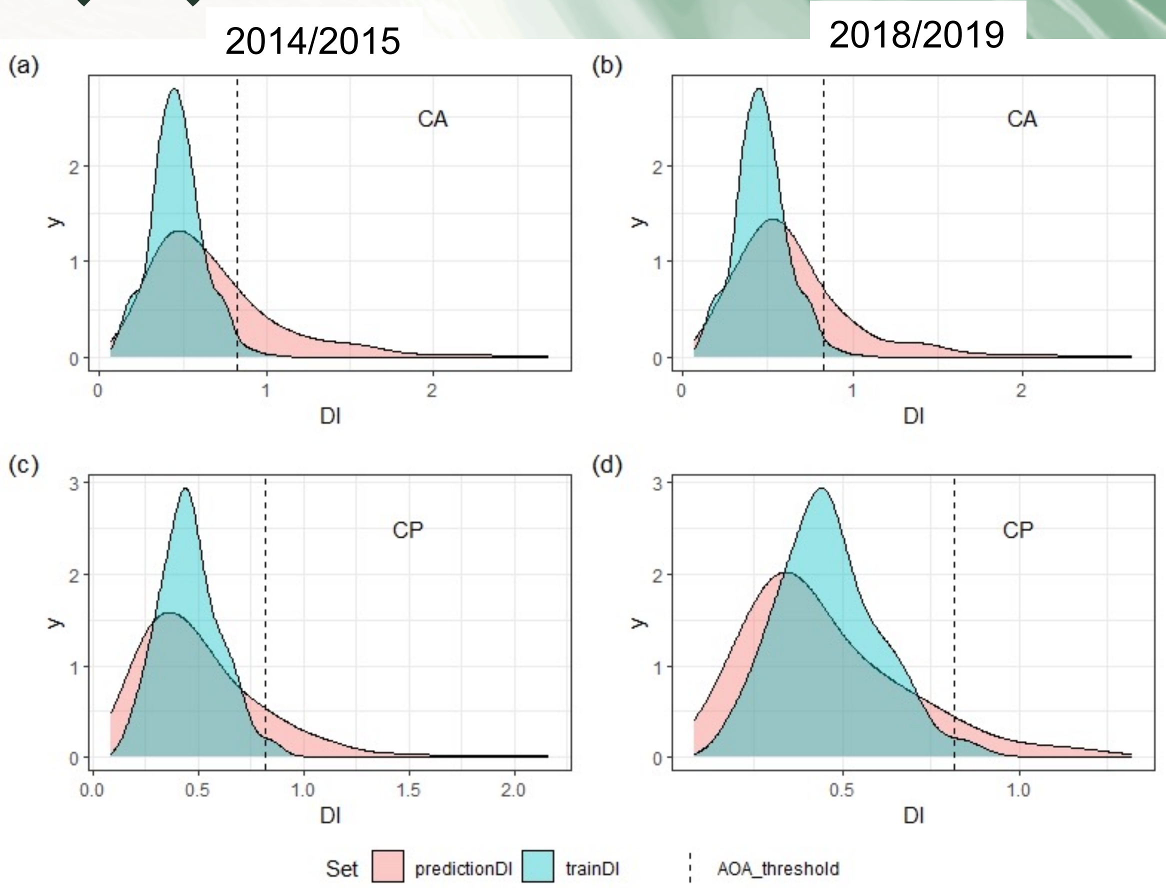
- Higher yields in 2018/2019 compared to 2014/2015 season
- Higher yields for CA compared to CP



# Dissimilarity index(DI) for training data & new locations

Details

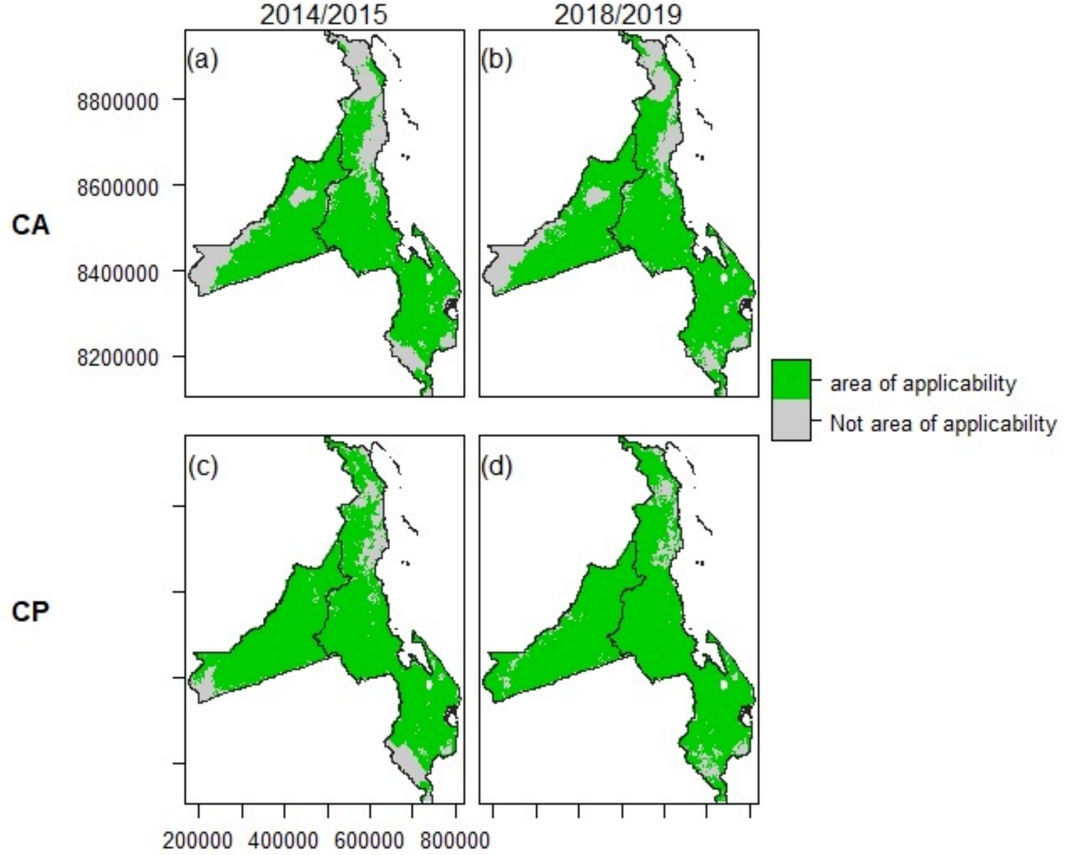
- Threshold values
- CA= · 0.825
- CP= ·0.819



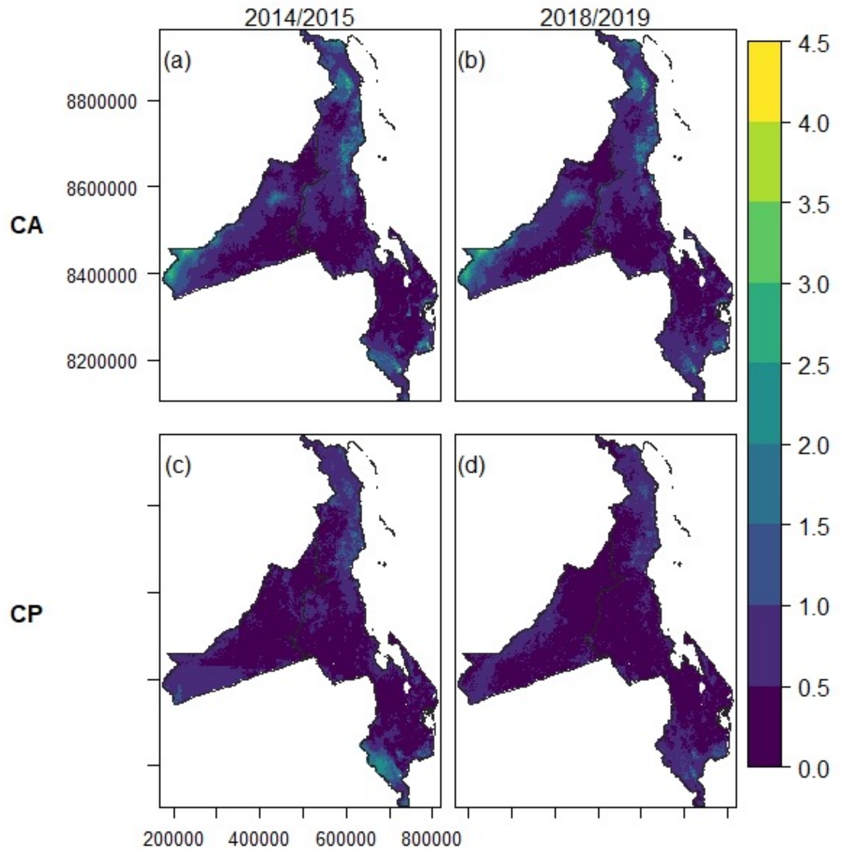
# Area of Applicability

Details

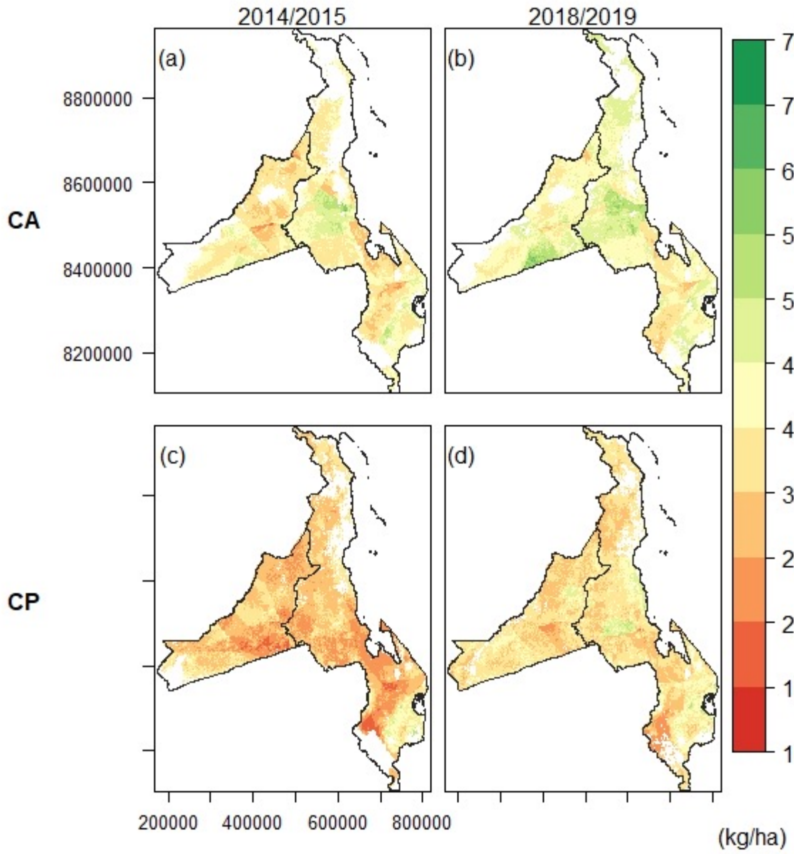
- Model did not learn relationships in the north eastern part of Malawi and western part of Eastern Zambia



# DI spatial distribution




# Final predictions



06

# Conclusion

# Takeaway

### Spatial heterogeneity & crop yields predictions

Indeed accounting for spatial heterogeneity can enhance spatial temporal crop yields.

### CA and Maize yields

Conservation agriculture practices increase maize yields

### The Area of applicability

Can effective highlight areas where a ML model can make predictions reliably. This facilitates effective extrapolation of agricultural technology

# Thanks!

## Do you have any questions?

francismich196@gmail.com
+255 688 960 833
+49 1522 806 5584
+34 624 696 788

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**