Resources Optimization by Homogenization of Agricultural Fields

Gordana Đuraš

UN/Austria Symposium
"Space Applications for Food Systems"
7-9 September 2021
Current Agricultural Situation

- Optimization of resources due to constantly growing population, sustainability and environmental protection.

- Management: soil conditions within the fields are not taken into account → over-fertilization or undersupply.

- Sowing seeds: at proper distances and depth, but selected quantity uniformly for the entire field → affect plant development and yield.

- Soil samples: taken at 25 different field points, obtained information is averaged and standardized for the entire field → high loss of information.
Aims and Objectives

- Optimization of resources
  - Exact knowledge of the soil characteristics needed
  - Differences between and within fields have to be identified

- Characterization of soil differences
  - Identification of homogeneous field areas (clusters) based on georeferenced soil sensor data

- Identification of representative points
  - Basis for additional soil samples
    → analysis in labor
How to Split a Field Into Homogeneous Areas?
Prerequisites for Zoning

- Extensive data base
  - Georeferenced data matrix, point related to geographical location (longitude, latitude, altitude)
- Collection of diverse soil parameters
  - Data should describe soil characteristics
- Technologies for data generation
  - Diverse sensor technologies
  - Satellite information
  - Results from chemical analysis of soil data
  - Drone-collected data
- Statistical methods to extract relevant information from recorded data

Soil sensor generates thousands of georeferenced data records per field
**Data Recording by Veris Soil Sensor**

- **RTK GPS**
  - Permanent measurements

- **Electrical conductivity**
  - 0 – 30 cm
  - 0 – 90 cm
  - Permanent measurement

- **Infrared and red-radiation**
  - Every second measurement

- **pH measurements**
  - Every 20m

Fotocredit: Maschinenring Steiermark
**Data Recording on a Field**

**Step 1:** drive along the edge of the field

**Step 2:** drive inside the field, distances between two tracks constant

10m – 20m

*Missing values*
Data Preparation

1. Plausibility check to eliminate non-positive values and outliers
2. Data discretization on 6mx6m grid to ensure data completeness
3. Data aggregation (by mean calculation) in order to get one value per cell
Spatial Interpolation

- Prerequisite for identification of homogeneous areas
  - Prediction of missing values (empty cells) in order to get complete data
  - Model of choice: Generalized Additive Model (GAM)
Data Smoothing

Smoothing of all values by GAM proved to be more beneficial
Identification of Homogeneous Field Areas

- Application of cluster analysis
  - Approach that groups a set of objects similar to each other in the same clusters
  - Objects within the same cluster are dissimilar to the objects in other clusters
  - Number of clusters $n$ must be specified in advance

- Hierarchical clustering performed in statistical software R
  - `hclust(d, method)`, where $d$ denotes the Euclidean distance (dissimilarity measure)

- Cluster analysis is based on $m$ explanatory variables
  - Our case ($m = 2$): soil conductivity parameters (in 30cm and 90cm depth)
Identification of Representative Points

1. Determination of overall-mean for each explanatory variable and cluster
   \(\rightarrow\) for \(m=2, n=6\): Zone1(\(\bar{x}_1, \bar{x}_2\)), …, Zone6(\(\bar{x}_1, \bar{x}_2\))

2. Computation of the Euclidian distance between overall-mean and cell values (\(x_1, x_2\)) of the respective cluster
   \[d_l = \sqrt{(\bar{x}_1 - x_1)^2 + (\bar{x}_2 - x_2)^2}\] for \(l = 1, \ldots, 6\)

3. For each cluster the selection of the cell with minimum \(d_l\)
   \(\rightarrow\) Representative point

Fotocredit: Maschinenring Steiermark
Methodology Advantages

- Obtained zone information basis for determination of sowing quantities and application maps
- Optimal points for additional soil data collection and analysis in labor
  → Data are characteristic for the entire zone, not only one point
- Zone information and results from chemical analysis basis for statistical modeling of crop, manuring and irrigation
- Applicability does not depend on the farm size and shape of a field
Resources Optimization Diagram

- Sampling of Soil Parameters
- Explanatory Variables for Statistical Modeling
- Chemical Analysis
- Additional Soil Sampling
- Identification of Homogeneous Field Regions
- Identification of Representative Points
- Data Management & Clustering
Future Work

- Data fusion: Inclusion of remote sensing data into the existing database
  - Diverse vegetation indices as further explanatory variables
  - To what extent the zoning of homogeneous areas with similar growth conditions can be improved
- Consideration of seasonal and year-specific effects
- Consideration of weather conditions
Thank You For Your Attention!

Partners

Data Recording & Sampling on Representative Points by

Maschinenring Steiermark Nährstoffmanagement

Statistical Analysis, Database Building and Development of a Shiny-App by