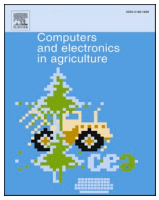
Computers and Electronics in Agriculture 197 (2022) 106894

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



Patch cropping- a new methodological approach to determine new field arrangements that increase the multifunctionality of agricultural landscapes

Marco Donat [a,b,*], Jonas Geistert [c], Kathrin Grahmann [a], Ralf Bloch [a,d], Sonoko D. Bellingrath-Kimura [a,b]

[a] Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, Germany
[b] Department of Agronomy and Crop Science, Humboldt-University of Berlin, 14195 Berlin, Germany
[c] Faculty Electrical Engineering and Computer Science, Technical University of Berlin, 10623 Berlin
[d] Eberswalde University for Sustainable Development, University of Applied Sciences, Eberswalde, Germany

ARTICLE INFO

Keywords:
Soil Management Zone Delineation
Yield Productivity Zones
Python
Clustering

ABSTRACT

Agricultural intensification decreased land cover complexity by converting small complex arable field geometries into large and simple structures which then were managed uniformly. These changes have led to a variety of negative environmental effects and influence ecosystem services. We present a novel small-scale and site-specific cropping system which splits a large field into small homogeneous sub-fields called 'patches' grouped in different yield potentials. A detailed workflow is presented to generate new spatially arranged patches with special focus on preprocessing and filtering of multi-year yield data, the variation in patch sizes and the adaptation of maximum working width to use available conventional farm equipment and permanent traffic lanes. The reduction of variance by the used cluster algorithm depends on the within-field heterogeneity. The patch size, the number of growing seasons (GS) used for clustering and the parallel shift of the patch structure along the permanent traffic lane resulted in a change in relative variance. Independent cross validation showed an increased performance of the classification algorithm with increasing number of GS used for clustering. The applied cluster analysis resulted in robust field segregation according to different yield potential zones and provides an innovative method for a novel cropping system.

1. Introduction

In recent decades, agriculture has intensified through the cultivation of fewer crop species, narrower crop rotations, and the steadily increasing use of fertilizers and plant protection products (Landis 2017). The use of larger and more powerful agricultural machinery, as well as the creation of large fields through farmland designs, led to a continuous higher productivity per unit area (Wik et al., 2008).

However, this intensification also created numerous problems: loss of insect diversity (Hallmann et al., 2017) and bird diversity (Donal et al., 2001), soil degradation due to wind and water erosion (Govers et al., 2017), soil compaction (van Ouwerkerk and Soane, 1994), habitat fragmentation and reduction of semi-natural habitats (Tscharntke et al., 2005). These impacts on agroecosystems can be explained mainly by the conversion from small-scale traditional to modern intensive agricultural production with increased field sizes in simplified landscapes and thus increased landscape homogeneity (Landis, 2017; Tscharntke et al., 2005). Agriculture is one of the most important factors shaping and changing the landscape, and in the past has often strongly modified the mosaic of crop area and natural and semi-natural habitats in key crop-producing areas (Haan et al., 2021). As a result, the compositional (amount of different habitat types) and configurational (spatial arrangement of individual habitat) heterogeneity of the landscape have been reduced.

Precision agriculture is a reaction to heterogeneous field conditions in space and time and thus implies variable application of inputs. To maximize local yield, resources are allocated according to the site variability and crop requirements of the field. This makes it necessary to divide the field into appropriate management zones. A strong focus is currently being placed on variable application of inputs such as site-

* Corresponding author.
E-mail address: donat@zalf.de (M. Donat).

https://doi.org/10.1016/j.compag.2022.106894
Received 29 November 2021; Received in revised form 22 February 2022; Accepted 19 March 2022
Available online 12 April 2022

specific nitrogen management (Mittermayer et al., 2021), variable seeding rate (da Silva et al., 2021), variable seeding depth (Coronel et al., 2020) and irrigation (Fontanet et al., 2020). Precision agriculture has increased resource efficiency and thus farmers net returns by shifting from a uniform application of inputs to a variable application. This variety of management strategies that have been developed focus on within-field inputs but have rarely been adapted to the best fitting land use for these respective zones. Previous concepts like site-specific conservation agriculture have been proposed, where certain crops or naturel habitats should only be grown in those zones within a field that are best suited for them (Basso, 2003). The question about the optimal size or shape of the zones, the variables used for management zone delineation or its implementation with conventional farm equipment, remains unsolved. Recent approaches of spot-farming (Wegener et al., 2019) , pixel cropping (Ditzler and Driessen, 2022) or patch cropping (Grahmann et al., 2021) conceptually address this demand to design sustainable cropping systems of the future. However, for the successful implementation of new cropping system designs, their possible benefits on ecosystem services and biodiversity need to be considered.

When large fields are heterogeneous in their soil characteristics or topography, it is more likely that yields in a field result extremely variable, especially under unstable and unfavorable climatic conditions. Profit mapping analyses showed that up to 14% of a field are causing costs due to its low yield potential and about 50% of the total area remained below the expected minimum revenues (Capmourteres et al., 2018). Hotspot analyses of total cropland area in Iowa, USA have shown that about 5% of the total area represents highly unprofitable land. Planting perennial plants in these unprofitable sub-areas would reduce establishment costs and increase the profitability for the farmer (Brandes et al., 2016). These analyses showed that land use changes within single management zones can be profitable for the farmer. There is an increasing demand to use the unproductive parts of arable land for different purposes. To reduce the farmers environmental impact and enhance ecosystem service and economic benefits, several authors propose to combine alternative species or perennial crops in the unproductive sub-areas within fields and annual crops in the more productive sub-areas (Basso, 2021; Khanna et al., 2021).

In this study, new field arrangements based on management zones are investigated. These zones are defined as small-structured field units with homogeneous site characteristics and hereinafter named ‘patches’. Patches differ in crop performance and are therefore managed differentially within one field. Patch cropping aims to cultivate individual small patches with crops and crop rotations that are best suited for these patches. Smaller field sizes were reported to increase pollinators (Garibaldi et al., 2016), improve natural pest regulation (Baillod et al., 2017) and enhance biodiversity (Šálek et al., 2018). The division into smaller units with differences in land management practice will lead to significantly smaller field sizes and new field shapes that replace large uniform and sole cropped fields. Incorporating this new cropping system design to fields within landscape–scales will result in changes in landscape structures with increasing compositional as well as configurational heterogeneity.

The aim of this study is to present a step by step procedure to detect less productive land with low yield potential for site adapted management and decreasing field sizes. Therefore, the objectives are to 1) apply the concept of patch cropping, a new field design approach to provide solutions for small-scale and site-specific farming, to 2) describe an automated workflow in the programming language Python for the division of fields into areas with different yield potentials based on multi-annual yield maps and to 3) analyze the effects of the individual patch size and the number of growing season (GS) used for the patch generation on the clustering success. The focus in this study is drawn on the implementation with today’s technology which takes into account maximum working widths, permanent traffic lanes and simple geometries that can be processed by conventional agricultural machinery.

2. Materials and methods

2.1. Patch clustering

2.1.1. Patch generation

Patches of the same yield potential consist of a variable number of N subpatches. Subpatches are rectangular polygons with a defined edge length. The edge length corresponds to a multiple of the maximum working width. In order to generate patches that can be managed precisely by the farmer and correspond to the permanent traffic lanes the farmer has used so far, the maximum working width and the georeferenced permanent traffic lane was used. The maximum working width is required, as the subpatch width can only be a multiple of this working width. The crop protection sprayer is usually the widest agricultural machinery with a size of 10 to 36 m. In this study, the maximum working width was 36 m. Secondly, the farmer’s permanent traffic lanes are required. This lane is a simple line string consisting of a geographic starting point (A) and an end point (B) in each field and is used for the defined permanent lanes in controlled traffic farming. To generate subpatches which are aligned parallel to any permanent controlled traffic lane, each subpatch is defined by its corners $C_{\{i,j\}}$, $C_{\{i+1,j\}}$,$C_{\{ij+1\}}$ and $C_{\{i+1j+1\}}$ for all $i,j \in \mathbb{N}$. Each corner $C_{\{i,j\}}$ is calculated using the equation:

$$C_{\{i,j\}} = A + d^{*}i^{*}V_{\{1\}} + d^{*}j^{*}V_{\{2\}} + W \tag{1}$$

where d is the edge length of subpatches and $V_{\{1\}}$ is a vector with the same direction as the traffic lane $\overline{AB}$ but a normalized length of one, hence it is calculated by the equation:

$$V_{\{1\}} = \frac{A-B}{|A-B|} \tag{2}$$

The vector $V_{\{2\}}$ is orthogonal to $V_{\{1\}}$ with the same length calculated with the equation:

$$V_{\{2\}} = V_{\{1\}}{}^{*}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3}$$

Since the traffic lane indicates the track on which the tractor drives, the subpatch structure must be shifted by half of the maximum working width to match with the equipment width (Fig. 1).

This shift can be described as an additional offset to each corner in the direction W calculated by the equation:

$$W = V_{\{2\}}{}^{*}\frac{w}{2} \tag{4}$$

where w is the working width. Without the offset W, the edges of individual subpatches would be identical to the farmer’s permanent traffic lane, preventing optimal management and causing field border overrun by farm equipment. The subpatch corners $C_{\{i,j\}}$, $C_{\{i+1,j\}}$,$C_{\{ij+1\}}$ and $C_{\{i+1j+1\}}$ are used to create individual subpatch polygons.

2.1.2. Clustering analysis

Fuzzy C-means algorithm is an unsupervised clustering method which forms cluster centers in an iterative procedure based on the distance between input data (Bezdek, 1981). The clustering method minimizes the within-group sum of square errors and assigns each subpatch to a continuous class membership value from zero to one for every cluster center (fuzzy partition). If subpatches were clustered into two groups (cluster center A and cluster center B), the fuzzy partition will

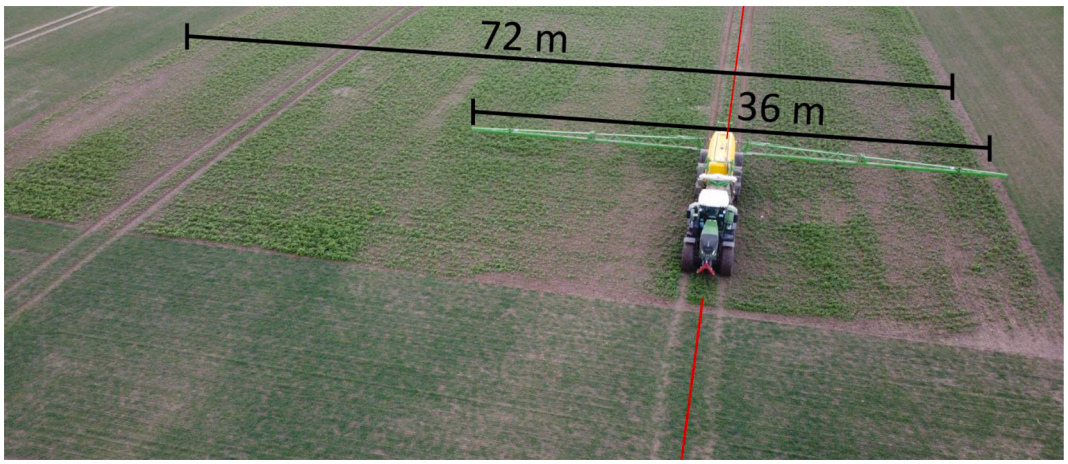


Fig. 1. Permanent traffic lane (red line) and maximum working width (36 m). When creating the patch structures, the subpatch borders must be moved parallel to the lane by halving the maximum working width © Marco Donat/ZALF.

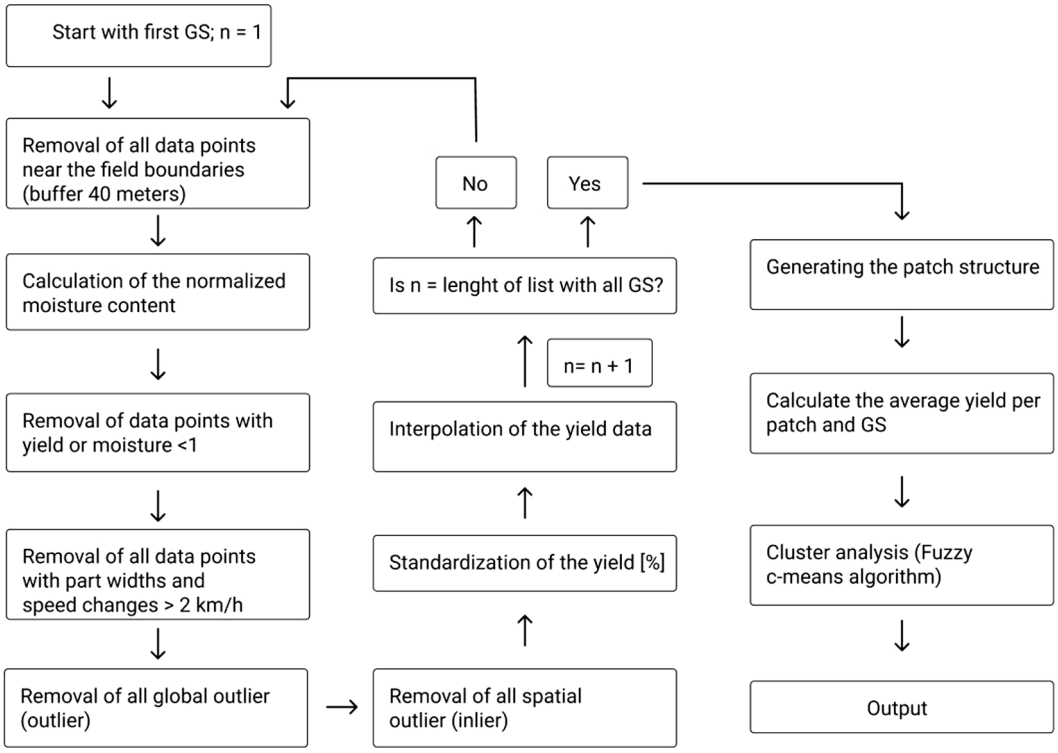


Fig. 2. Workflow of the automated data preprocessing, outlier removal, interpolation and subpatch clustering.

contain two continuous class membership values for each subpatch. Subpatches with a low class membership value for cluster center A only belong to this cluster to a low degree. Subpatches with a high class membership value for cluster center A, belong to a high degree to the cluster center A. To divide the field into two groups (cluster group class 1 and cluster group classs 2), a hard constraint must be set to assign the membership of every subpatch to one of the two groups. To defuzzy our fuzzy partition, we used the 'max membership method' with a threshold of 0.5 (Ross, 2010). If the class membership value for cluster center A for a subpatch is greater than 0.5, this subpatch is assigned to cluster group class 1, but if the value is < 0.5, it is assigned to cluster group class 2.

Cluster analyses were carried out with Fuzzy C-means algorithm implemented in the *Scikit-Fuzzy* package. Before clustering, the mean and the variance of all data points per subpatch were calculated. When using several variables at the same time to cluster field heterogeneity (soil nutrients, elevation, apparent electrical conductivity, etc.), it is necessary to consider the weighting of the individual data layers, as the influence on the yield potential varies depending on the variable (Khosla et al., 2010). For the selection and weighting of the input variables, user experience and expert guidance should be used. Since this study is focusing only on yield data, yield data from several GS were weighted equally.

2.2. Workflow

Fig. 2 provides an aggregated workflow of the development pipeline for the patch generation. This workflow is written in Python programming language. Python is an open source general purpose programming language with a diverse ecosystem and a strong increase of a wide variety of libraries which offers almost infinite possibilities in the field of data science. The open source programming language R is widely used within the scientific community since it has powerful tools for statistical computing. Analysis of spatial data and geostatistical modelling can be done by a wide range of packages which evolved in the last three decades. Both programming languages were combined within this automated workflow.

We used Anaconda Distribution (Anaconda Navigator 1.10.0) for package management and Jupyter Notebook (Version 6.1.4) as an interactive development environment.

2.3. Site description and data sampling

In this study, multi-year yield maps from an agricultural farm in Brandenburg 60 km east from Berlin were used. Yield data were collected using a yield monitoring system CLASS Quantimeter (CLAAS, Harsewinkel, Germany) mounted on a CLAAS 580 from GS 2010 to 2013 and since GS 2014 on a CLAAS Lexion 770 TT combine harvester (CLAAS, Harsewinkel, Germany). A total number of 12 different fields (field_a to field_l) with different numbers of available yield maps was tested (Fig. 3). Yield maps of 7 GS were available for field_a, 8 GS for field_b, 9 GS for field_c, 8 GS for field_d, 4 GS for field_e, 7 GS for field_f, 7 GS for fielg_g, 8 GS for field_h, 7 GS for field_i, 8 GS for field_j, 8 GS for field_k and 9 GS for field_l. The fields have a total size of 383 ha with a mean size of 31.9 ha. For the results and description of the presented method, one exemplary field was chosen (field_d). Field_d has a size of 44.5 ha, is cultivated conventionally and yield maps are available for a total of 8 GS. The crop rotation for field_d consisted of rapeseed (*Brassica napus* L.) in 2012, 2015, 2018 and winter rye (*Secale cereale* L.) in 2010, 2011, 2013, 2014, 2016, 2017, 2019. Harvesting was performed by 2 combine harvester at a time in all GS except 2013. Due to data incompleteness, the GS 2012 and 2014 were excluded from the data analysis of field_d. The heterogeneous terrain of this agricultural landscape consists mainly of strongly sandy glacial till from ground moraine material and sandy, predominantly fine-grained, partly fine silty deposits caused by meltwater from the Southernmost Weichselian ice margin at the Brandenburg stadium. The long-term average annual temperature at this site is 9.6 °C with an average annual precipitation of 581 mm.

2.4. Input data

Yield maps were provided as vector based spatial data sets (ESRI-Shape file) and were converted in a geodataframe with *Geopandas* in Python. Each data point contains information about the Global Navigation Satellite System position, yield, grain moisture, working width and operating speed. The coordinate reference system was converted into Universal Transverse Mercator coordinate system (33rd meridian strip system northern hemisphere).

When large fields are managed uniformly, harvesting usually takes place with two or more combine harvesters in order to increase efficiency. Due to possible errors caused by different calibrated sensors of the combined harvesters, data sets from one GS were analyzed separately. However, this is only possible if harvest followed a regular pattern and the distances between the rows of a single harvester do not become too large (Fig. 4). Rows are defined as yield data points generated within a combine harvester path through the stand during harvesting. Yield maps were examined for completeness, as yield maps covering only parts of the field lead to erroneous interpolation. Incomplete yield maps or yield maps in which the spacing between individual tracks of one harvester was greater than three rows were not considered for the analysis.

2.5. Data preprocessing

Working with yield data from combine harvesters requires a cleaning procedure before processing because of erroneous data points like header width error, end-of-row-error and extreme yield values (Sudduth et al., 2012). The removal of data points near field edges is an important step to remove the end-of-row erroneous data points. The buffer area around the field has to be larger than the maximum working width and was set to 40 m.

The next processing step was the calculation of the yield to standard moisture content since every data point contains a measured yield with a particular grain moisture content. Yield data with standard moisture content xs_i was calculated at 10 %, independent of the crop species with the following equation:

$$xs_i = x_i * \frac{(100 - MC_i)}{(100 - MC_s)} \tag{5}$$

with x_i is the yield of data element i, MC_i is the initial moisture content of x_i [%] and MC_s is the standard moisture content of 10 %. Data points with a moisture content or a yield value of <1 were removed as this is considered implausible from a plantś physiological perspective.

The yield monitoring system converts the measured crop volume flow into yield [Mg/ha]. The calculation is only correct if the entire

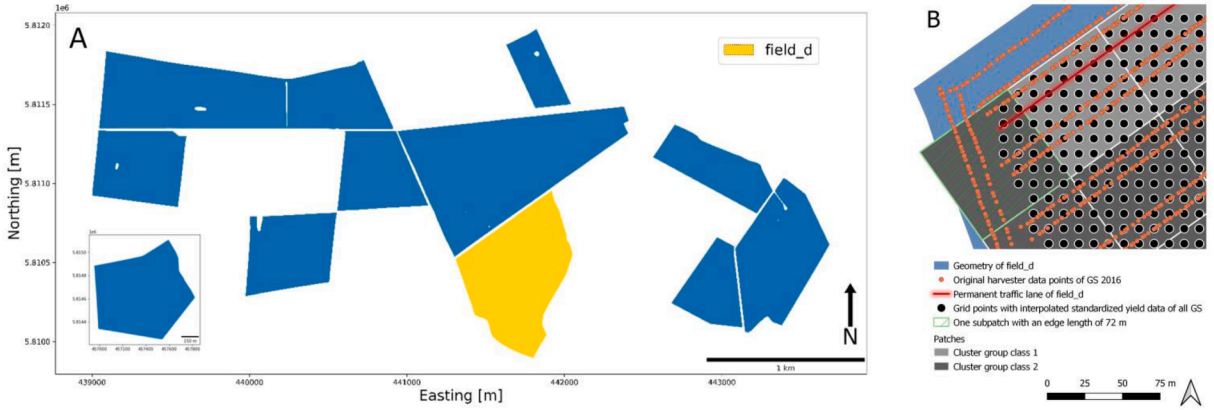


Fig. 3. Spatial arrangements and dimensions. A) Spatial distribution of field_a to field_l and B) Spatial arrangement and its relationship among different definitions.

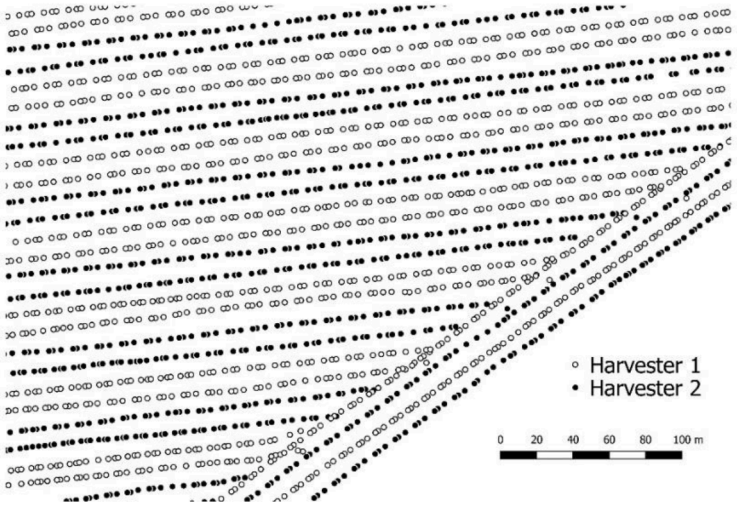


Fig. 4. Two combined harvesters in operation during the harvest of one GS with a regular pattern. Data sets for one GS should be processed individually due to possible calibration errors.

header width is used. Although the header part-width can be measured during the harvesting process, the division into 6 header part-widths is too imprecise. Hence, data points with part widths were removed. Abrupt speed changes increase yield monitoring errors compared to constant speed of the combined harvester (Arslan and Colvin, 2002). Therefore, for all data points the speed changes to the previous measured data point were calculated and the data points with a speed change greater than 2 km/h were removed.

2.6. Outlier removal

Global outlier are data points of yield values that lie outside of the general pattern of the distribution of all yield data points. Removing these extreme values is important, as they might lie outside the range of optimal yield monitor performance or could depict harvesting artifacts and thus affect the analysis. At the same time, removed data elements can also be valid data (Córdoba et al., 2016; Simbahan et al., 2004). This is not intentional, because the objective of outlier removal is to remove isolated erroneous data elements outside the true yield distribution. However, to ensure comparability to filtering techniques of other authors, we have used a commonly used criterion. To detect and remove global outliers, a Z-score Z_i for each data element was calculated by the equation:

$$Z_i = \frac{X_i - \overline{X}}{\sigma} \tag{6}$$

with X_i is the yield value of data point i, $\overline{X}$ is the mean and σ is the standard deviation. This expresses Z_i as a deviation from the mean, relative to the standard deviation. Data points with yield values plus minus 3 times the standard deviation were removed (Simbahan et al., 2004; Vega et al., 2019).

Spatial outlier are data points of yield values which significantly differ from the yield values of their neighboring data points. For spatial outlier detection, we used the local indicator of spatial association (LISA) which indicates the significance of the spatial clustering of equal values around the observed points (Anselin, 1995). A relatively large neighborhood of a radius of 40 m around each data point was selected to be checked (Vega et al., 2019). The advantage of a large neighbourhood definition is that data elements from several neighboring rows are included. This is especially important if more than one combine harvester was used in the same field.

In our workflow, all local neighbors had to be identified for each data element by creating a spatial weights matrix using *SciPy KDTree* calculation. Weigths W_{ij} were generated using data points within the radius around the data point of interest with weights of 1 and data points outside the radius with weights of 0. Then, for each data point the locale Moran's Index was calculated using the equation integrated in *PySAL* (Rey and Anselin, 2010):

$$I_i = \frac{X_i - \overline{X}}{\sigma^2} \sum_{j,j \neq i}^{n} \left[W_{ij} \left(X_j - \overline{X} \right) \right] \tag{7}$$

X_i is the value of the element of interest, $\overline{X}$ is the average of all elements X with a sample number of n, X_j is the value of all elements with restriction $j \neq i$ and σ^2 is the variance of all elements X. To identify local outliers, a conditional permutation approach was used to calculate empirical reference distribution which is required to apply statistics testing the null hypothesis. After standardization of the local Moran's Index, the significance level was tested based on the reference distributions with a pseudo p-value ≤ 0.05 (95% confidence level) (Anselin, 2019). Spatial outliers have significant negative high Local Moran's Index values. These data points were identified and removed from the dataset. Before interpolation, the yield values were standardized to relative yield values for the comparison between different crop species (Shannon et al., 2020) using the following equation:

$$X_{rel} = \frac{X_i}{\overline{X}} \times 100\% \tag{8}$$

2.7. Interpolation

For the production of area-covering yield maps, interpolation methods were applied to estimate the yield at a specific point on a regular grid between various yield points. The distance between the regular grid points distributed on the field was set to 10 m. Firstly, an experimental semivariogram was created for all yield maps and different theoretical variograms were fitted. Ordinary kriging was used to interpolate yield data on a regular grid. A spherical and exponential variogram model was fitted and the sum of squared errors (SSErr) was calculated. The model with the lowest SSErr was used for final kriging. Variogram calculations, fitting of a spherical and exponential variogram model and final predictions were done with the package *gstat* in the statistical software R. To execute the R-script within the Python environment for automated analyses, the module *subprocess* was used. The interpolated relative yields of all GS are merged with the created subpatch structure. Since there may remain subpatches that have only a small number of yield points due to complex field geometries, a threshold of 30% of available data points of a completely filled subpatch was set. Thus, only subpatches that had more than 30% of data points

within the subpatch polygon were considered in the subsequent analysis. To visualize the interannual variability of relative yields, we used the technique of Blackmore (2000) to calculate the standard deviation (σ) and generated a 'classified management map' to identify spatial and temporal yield trends. Interpolated standardized yield data points will be higher yielding if the mean yield (μ) of all GS ≥ 100, lower yielding if μ < 100 and unstable if the σ greater than 30.

2.7.1. Georeferenced output

The generated output consists of georeferenced rectangular polygons, which are orientated towards the permanent traffic lanes. The group mean has to be determined for the final assignment of cluster group classes. In the case of two groups, the group with the lower group mean is attributed to the group with lower average normalized yields. The output is a polygon ESRI shapefile and thus corresponds to a standard vector-based spatial data format.

2.8. Evaluation of clustering

2.8.1. Cross-validation

To evaluate the proposed method, the Leave-One-Out-Cross Validation (LOOCV) procedure (Thorp et al., 2007) was applied. For this purpose, independent data sets were used which were not included in the cluster group creation to determine the cluster predictive performance expressed in root mean squared error of prediction (RMSEP) calculated by the equation:

$$RMSEP = \sqrt{\frac{\sum_{j=1}^{n}(Y_{mi,j} - Y_{si-j})^2}{n}} \quad (9)$$

Here, Y_{mi} is the mean standardized yield of the respective subpatch from the GS used for LOOCV in the *i* cluster group and the *j* GS and $Y_{Si\text{-}j}$ is the mean standardized yield of the *i* cluster group with all GS except the *j* GS.

Since temporal yield variability is largely due to weather variability, we assume that the performance of the predictive power should improve with an increasing number of GS. This depends mainly on the available water for plant growth, assuming equal management and optimal supply of micro- and macronutrients. Therefore, the performance of the proposed cluster method was tested with a different number of yield maps from different GS as well as different combinations of GS. Several sub datasets were created for every field starting with two GS to *n* GS with *n* being the total number of all GS for one field. For each individual dataset, the LOOCV procedure was performed and the corresponding RMSEP was calculated. The possible combinations depend on the number of available GS for each individual field. Since there are different numbers of GS per field, the number of possibilities can be calculated using the binomial coefficient:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (10)$$

Here, *n* corresponds to the number of all GS and *k* to the number of GS used for clustering. The number of possible combinations for the LOOCV can be seen in the Pascal's triangle in Table 1 and therefore specifies the number of sub datasets required to perform the LOOCV procedure. If yield maps from four GS were available for one field (e.g. fild_e), a total of 14 sub datasets was created. From these 14 sub datasets, four datasets were used for the calculation of the RMSEP to evaluate the performance when only one GS was used for clustering. Six datasets out of these 14 sub datasets were used to calculate the RMSEP to evaluate two GS used for clustering. Another four datasets from that 14 sub datasets were used to calculate the RMSEP to evaluate three GS used for clustering (Table 1). The RMSEPs of the combinations having the same number of GS used for clustering were averaged.

Table 1
Possibilities of different combinations with different numbers of available yield maps from the respective GS.

	GS used for cluster analyses							
	1	2	3	4	5	6	7	8
Yield maps of 8 GS	8	28	56	70	56	28	8	
Yield maps of 7 GS	7	21	35	35	21	7		
Yield maps of 6 GS	6	15	20	15	6			
Yield maps of 5 GS	5	10	10	5				
Yield maps of 4 GS	4	6	4					
Yield maps of 3 GS	3	3						
Yield maps of 2 GS	2							

2.8.2. Relative variance

To evaluate the effectiveness of the proposed clustering method depending on subpatch sizes and the number of GS used for clustering, the relative variance (RV) was used (Ortega and Santibanez, 2007). The RV reflects the proportion of the variances of all GS used, explained by different subpatch sizes and clustering algorithm using the equation:

$$RV = 1 - \frac{\sigma_W^2}{\sigma_T^2} \quad (11)$$

Here σ_W^2 is the variance of the relative yield within cluster groups and σ_T^2 is the total variance without clustering. The clustering process has greatly reduced the variance within cluster groups when the RV is close to one. The RV was calculated for all possible combinations of available fields, subpatch sizes and GS.

Site heterogeneity can vary greatly within a few meters. Using today's technology requires certain maximum working widths. If the subpatch size is reduced to a edge length of two working widths, the within zone homogeneity is higher than a subpatch size of more than two working widths. To test this hypothesis, the workflow described above was applied with different GS and subpatch sizes. To do so, patches were created with varying edge lengths from 12 m to 108 m with a stepwise increase of the subpatch edge length by 6 m.

Although the patches are aligned parallel to the traffic lane and thus have a fixed orientation in space, the patches can be shifted along the traffic lane. Especially in the case of large patches, it is possible that the heterogeneity is not well reflected if the starting point for generating the patch structure is always identical by default. In order to test whether changes in RV can be observed with identical subpatch sizes but with patches shifted parallel to the traffic lane, four different starting coordinates for generating the patch structure were created for all fields and all GS. These different starting points resulted from the parallel shift along the traffic lane by ¼, ½ and ¾ of the corresponding edge lengths of the rectangular subpatches.

3. Results and discussion

3.1. Workflow processing

The number of original data points and removed data points were different between GS (Table 2). The mean distance between data points within a row was 7.5 m in the GS 2010 to 2013 (harvest with CLAAS 580) and 3.6 m in the GS 2015 to 2019 (harvest with CLAAS Lexion 770 TT). A large proportion of the removed data points were within the buffer area. The removal of the global outliers had only a small share and varied from 0 data points in GS 2010 to 51 in GS 2013. Overall, the number of removed data points varied between years. While only 33.9% of the data points were removed in GS 2010, 45.1% had to be removed in GS 2013.

During the preprocessing of field_d and GS 2015, about 340 of 548 raw yield data points with values between 0 and 0.2 Mg/ha were removed (Appendix A). The rapeseed yield data points were not normally distributed and the histogram of 2015 GS showed one peak at around 2 Mg/ha and another peak at around 7 Mg/ha. The positive and

Table 2
Number of original yield data points of field_d per GS and number of removed data points in the respective data cleaning steps.

	GS 2010	GS 2011	GS 2013	GS 2015	GS 2016	GS 2017	GS 2018	GS 2019
No. Original data points	2648	2692	7699	6671	7055	6380	5328	8239
Data pre-processing (except Buffer removal)	256	196	932	548	249	835	356	522
Buffer removal	578	619	1714	1927	1793	1090	1292	2012
Global Outlier	0	2	51	3	35	44	1	11
Local Outlier	64	122	778	213	391	401	318	449
Total data removed [%]	33.9	34.9	45.1	40.3	35.0	37.1	36.9	36.3

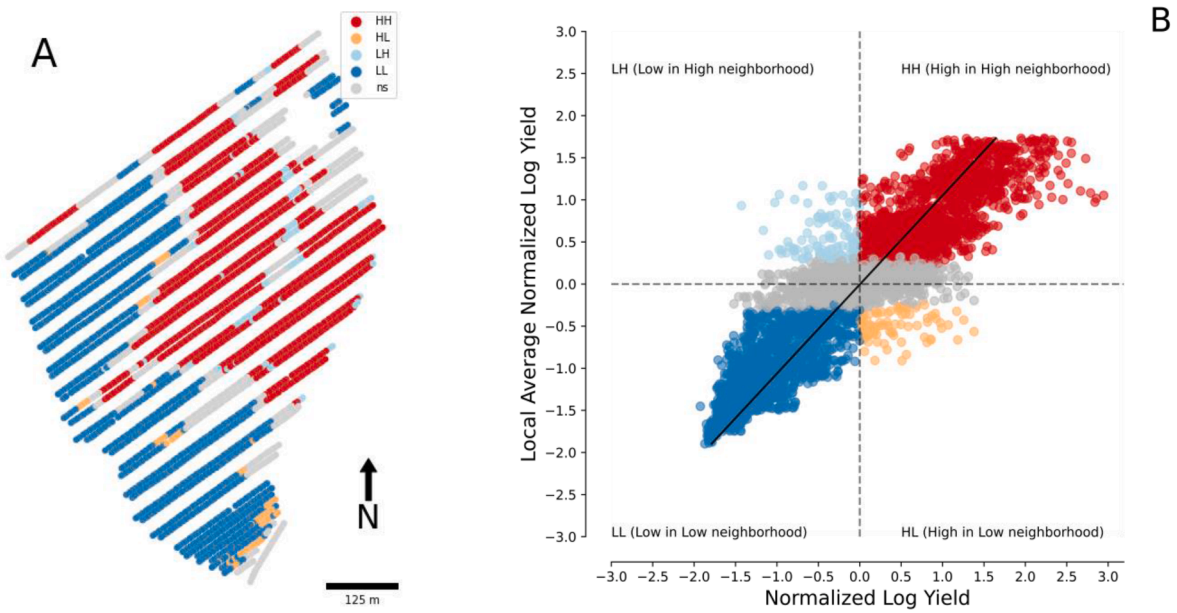


Fig. 5. Local Outlier analyses from GS 2015 rapeseed yield data of field_d A) Neighborhood plot of the original yield data points with its location within the field boundaries: Red dots symbolize significant high yield data points in close proximity to high yield data points (HH), blue data points symbolize significant low yield data points in close proximity to low yield data points (LL), orange data points symbolize data points that have significant high yield and are in the neighborhood of low yield (HL) and light blue data points symbolize significant low yield data points that are in the neighborhood of high yield data points. Grey data points are not significantly equal nor significantly unequal to the data points in its neighborhood (ns). B) Moran scatter plot.

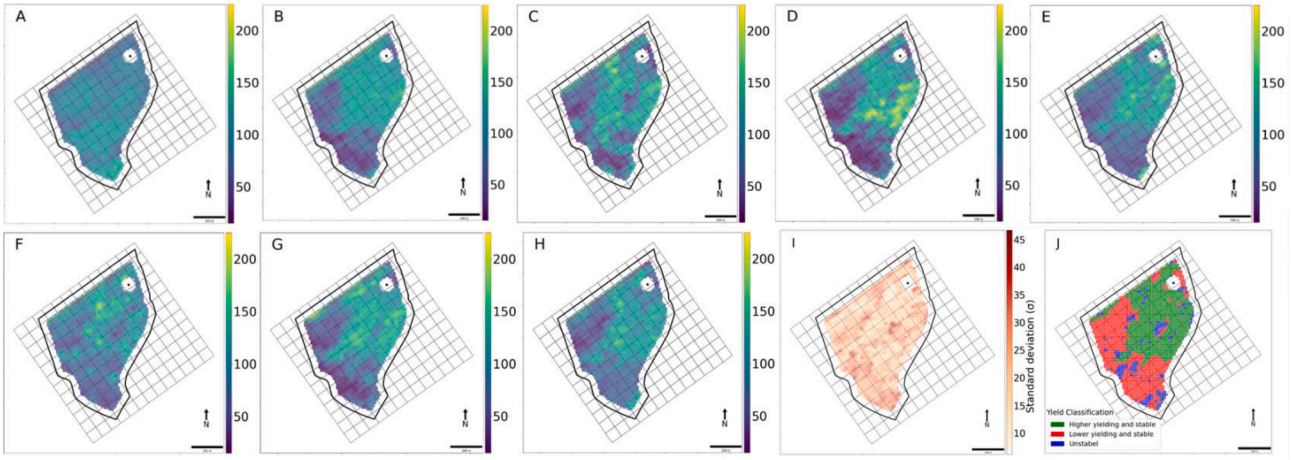


Fig. 6. Plots of interpolated relative yields of field_d overlaid with the created patch structure and a subpatch size of 72 m × 72 m oriented parallel to the traffic lane of Permanent Traffic Control (red line); A) Winter rye in 2010, B) Winter rye in 2011, C) Winter rye in 2013, D) Rapeseed in 2015, E) Winter rye in 2016, F) Winter rye in 2017, G) Rapeseed in 2018, H) Winter rye in 2019 I) Temporal stability map of all 8 GS with standard deviation of all data points and J) classified management map of all 8 GS with the classification 'higher yielding and stable', 'lower yielding and stable' and 'unstable'

significant global Moran index was 0.764 ($p < 0.001$) for the field_d in the GS 2015, implying positive autocorrelation of the yield data points (Anselin, 1995). The local Moran analysis performed on all data points identified 213 local outliers. The data points are shown in Fig. 5A and 5B as orange dots which represent high yield in low yield neighborhood (HL) and light blue dots which depict low yield data points in high yield neighborhood (LH). The red dots over a large contiguous area are hot spots for high yields, the blue dots are hot spots for low yields. The

Fig. 7. Output of the cluster analysis shows A) the patch structures with the continuous class membership value for cluster center A presented in greyscale. Cluster group class 1 corresponds to high average relative yield and therefore high yield potential. White subpatches have a high membership value of cluster center A. Black subpatches have a low membership value of cluster center A. B) The final output with 72x72 m subpatches and an assignment of the individual subpatches to a cluster group class.

Moran scatterplot (Fig. 5B) visualizes gray data points in the LH and HL quadrants which are not treated as local outliers because they are not significantly (ns) different from their neighboring data points.

Fig. 6 shows all interpolated standardized yield maps of 8 GS of field_d (Figs. 6A-H) and the interannual variability of relative yields. The spatial distribution of yield differences can be seen by eye (Figs. 6A-H). Zones with higher yields are located in the northern, central and eastern part of the field, whereas zones with lower yields can be found in the southern and western part. This spatial yield pattern can also be recognized in the classified management map (Fig. 6J). There are several areas within the field that showed a high degree of yield variability. These unstable zones (Figs. 6I-J) are distributed across the field but occur more frequently in the central and southern area.

By hardening the fuzzy partition (threshold of 0.5), all subpatches with a high continuous class membership value of cluster center A belong to cluster group class 1 (Fig. 7A). Cluster group class 1 obtained high average relative yields and is therefore assigned to the high yield potential. By setting the threshold to 0.5, a total of five patches were created for field_d (Fig. 7B), from which two of the patches consist of several connected individual subpatches, another patch consists of two subpatches and two patches consist of only a single subpatch each. Two subpatches are close to the threshold of 0.5 with a continuous class membership value of cluster center A of 0.49 and 0.51 (Fig. 8). Different cluster group classes were assigned to both subpatches, knowing that both differ only slightly from each other and both single subpatches belong neither to cluster group class 1 nor to cluster group class 2 to a

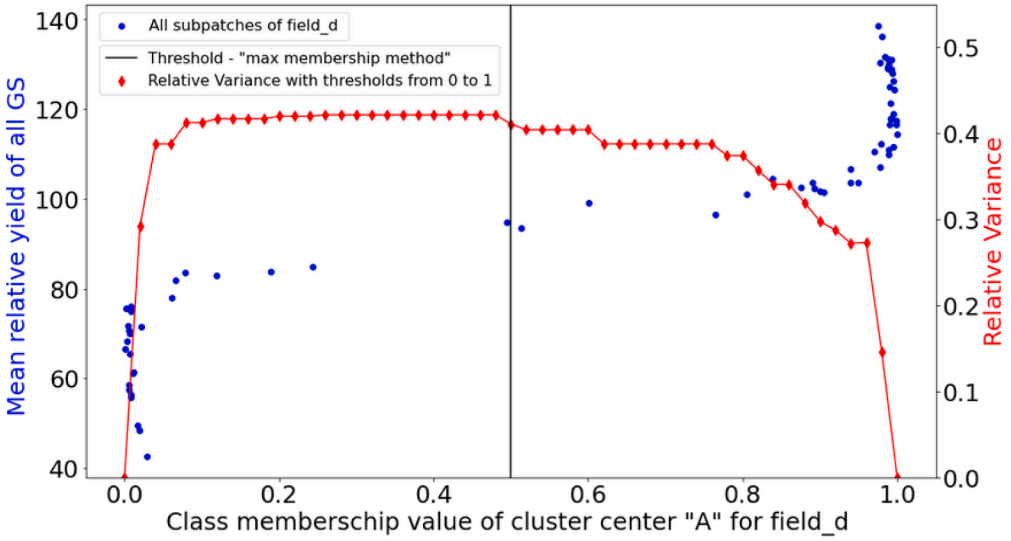


Fig. 8. Scatterplot of all Subpatches of field_d, their mean relative yield of all 8 GS and class membership value of cluster center A (blue points). The vertical black line visualizes the threshold of 0.5 to defuzzy the fuzzy partition. The red dots show the calculated RV for 50 different thresholds between 0 and 1 to defuzzy the fuzzy partition.

high degree.

In Fig. 8, all subpatches of field_d are plotted with their mean relative yield over 8 GS and their class membership value of cluster center A. All subpatches to the right of the vertical bar belong to cluster group class 1, all subpatches to the left of the vertical bar belong to cluster group class 2. The red dots within this plot represent the RV of 50 different thresholds between 0 and 1 for hardening the fuzzy partition. A threshold of zero or one would result in an RV of zero, since all subpatches are in one group and thus no variance would be reduced. The highest RV of 0.42 is obtained with a threshold between 0.26 and 0.48. With a threshold of 0.5, a slightly lower RV of 0.41 was obtained, assigning a total of 26 subpatches to cluster group class 2 and 42 subpatches to cluster group class 1. The choice of the threshold for hardening the fuzzy partition thus has a direct influence on the number of individual subpatches in the cluster group classes as well on the RV. If the total number of subpatches within one cluster group class should be modified this can be done within a certain range without major changes to the RV by threshold selection. Choosing thresholds between 0.08 and 0.80 would change the cluster group class assignment for a total of 8 subpatches, but would have relatively small effects on the RV (Fig. 8).

If field_d is clustered with the same variables but Fuzzy C-means algorithm settings of three or more cluster center, these two subpatches with membership vales of 0.49 and 0.51 would belong to the middle cluster center. In precision agriculture application, the search for the optimal number of cluster groups and thus management zones is often in the focus (Córdoba et al., 2016; Reyes et al., 2019; Xiang et al., 2007). However, since the objective of this study was to design a new cropping system with two cluster groups, several subpatches belong neither to one nor to the other cluster group to a high degree. Furthermore, the presented framework does not directly take into account temporal variability of yields due to the soil-climate interactions, which leads not only to zones with stable and high yield zones or stable and low yields zones, but also to unstable yield zones (Fig. 6J).

The largest average RV of 0.72 was obtained with a subpatch edge lengths of 12 m and one GS (Fig. 9). If an additional GS was used for clustering, the RV was reduced to 0.57. If all GS were used, the RV decreased to 0.51. The high RV values with only one GS are explained by the one-dimensionality of the data set. If data from an additional GS are added, the dimensionality of the dataset increases and with it the difficulty to reduce the variance for both GS. As a result of the increase in dimensionality through additional GS, the calculated total RV decreases.

A reduction of the RV also takes place with an increase in subpatch size. Here, the RV was reduced to 0.47 when one GS was used and the subpatch size increased from 12 m to 108 m. If all 8 GS were used, the RV diminished by 0.12 to an RV of 0.39 when the subpatch size increased from 12 m to 108 m. Xiang et al. (2007) detected a decrease of RV with an increase of the grid size used for the post-classification majority filter. The larger the subpatches become, the more difficult it gets to reflect homogeneity within the subpatches due to the sometimes small-scale site differences resulting in low RV.

Shifting the patch structure parallel along the traffic line while keeping the number of GS and the edge length of the subpatches constant, caused an alteration in the RV (Fig. 10). The RV of the initial patch structure with seven GS and a subpatch size of 108 m was 0.39, the RV of the clustering with identical parameters but shifted parallel by ½ was 0.34. The difference in RV caused by this parallel shift of the patches is 0.05 and thus 12.0 % in relation to the original RV value (Fig. 10).

However, most of the changes in the RV caused by parallel shift were smaller but as the edge length of the subpatches increased, the changes due to the parallel shift also became greater (Fig. 11).

There was a remarkable increase in RV with subpatch sizes of 78 m and 84 m (Fig. 9). In this case, the optimal subpatch size was found which, due to the spatial arrangement of the patch structure as well as their size, better reflects the heterogeneity through clustering and increases the RV. If the patch structures are moved by the parallel shift, the RV with subpatch sizes 78 m and 84 m decreased (Fig. 10). Therefore, the successful clustering with the highest possible RV for a specific subpatch size can be obtained when the patch structure is shifted parallel to the traffic lane.

3.2. Method verification by RV and RSMEP

3.2.1. Relative variance (RV)

When clustering is conducted for homogeneous within-field patches, the sum of variances of the data points within the groups is reduced compared to the total variance. The extend of RV reduction depends on the size of the subpatches. The RV of all fields gradually decreased as the edge length of the subpatches, and consequently the subpatch size became larger (Fig. 12). Clustering was less effective in reducing variance as the subpatches increased in size. In other words, subpatch sizes of single or double working width (36 m, 72 m) have a higher degree of homogeneous site conditions and yield patterns than subpatch sizes of threefold working width (108 m) and are therefore more suitable to be selected for a new site-specific cropping method.

The RV of some fields was always greater than the RV of other fields with the same number of GS and similar subpatch size (e.g RV of field_c > RV field_k > RV of field_g). Data clustering of the presented workflow will always generate two groups, although the data have low variance. Thus, a field with homogeneous site characteristics and yield pattern will be divided into two groups with different yield potential. Clustering

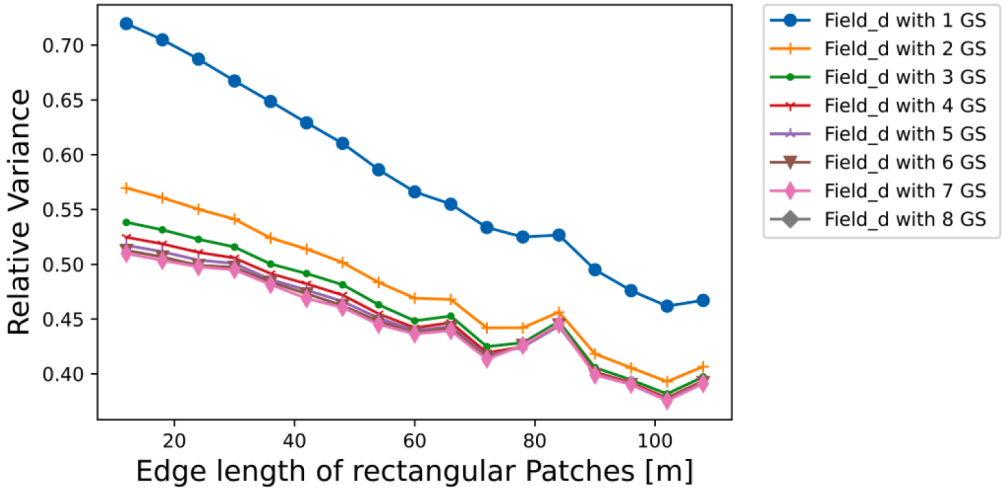


Fig. 9. Relative variance of Field_d with varying edge lengths and increasing number of GS.

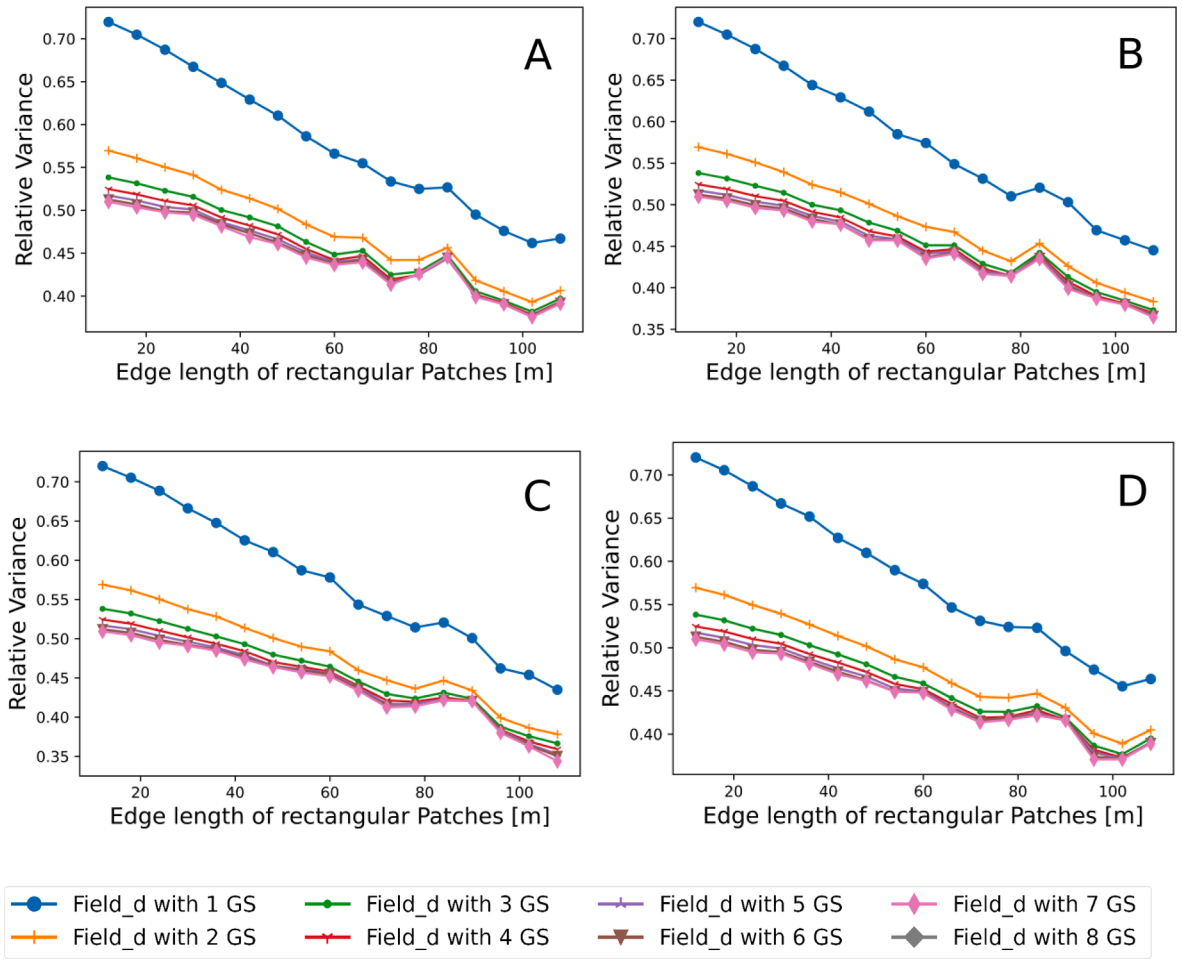


Fig. 10. Relative variance of Field_d with varying edge lengths and GS used depicting A) the relative variance with no parallel shift, B) Field_d with a parallel shift of ¼, C) Field_d with a parallel shift of ½ and D) Field_d with a parallel shift of ¾.

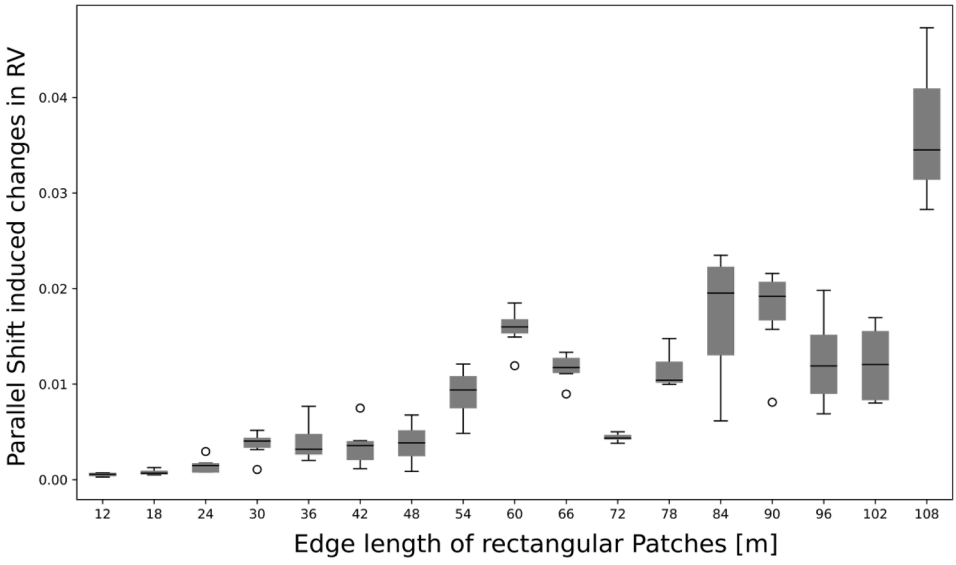


Fig. 11. Boxplot diagram of the relative variance (RV) differences caused by the parallel shifts of the patch structures in field_d. Single data points of a single boxplot consist of the difference of the minimum and maximum RV of all parallel shifts per GS and subpatch size.

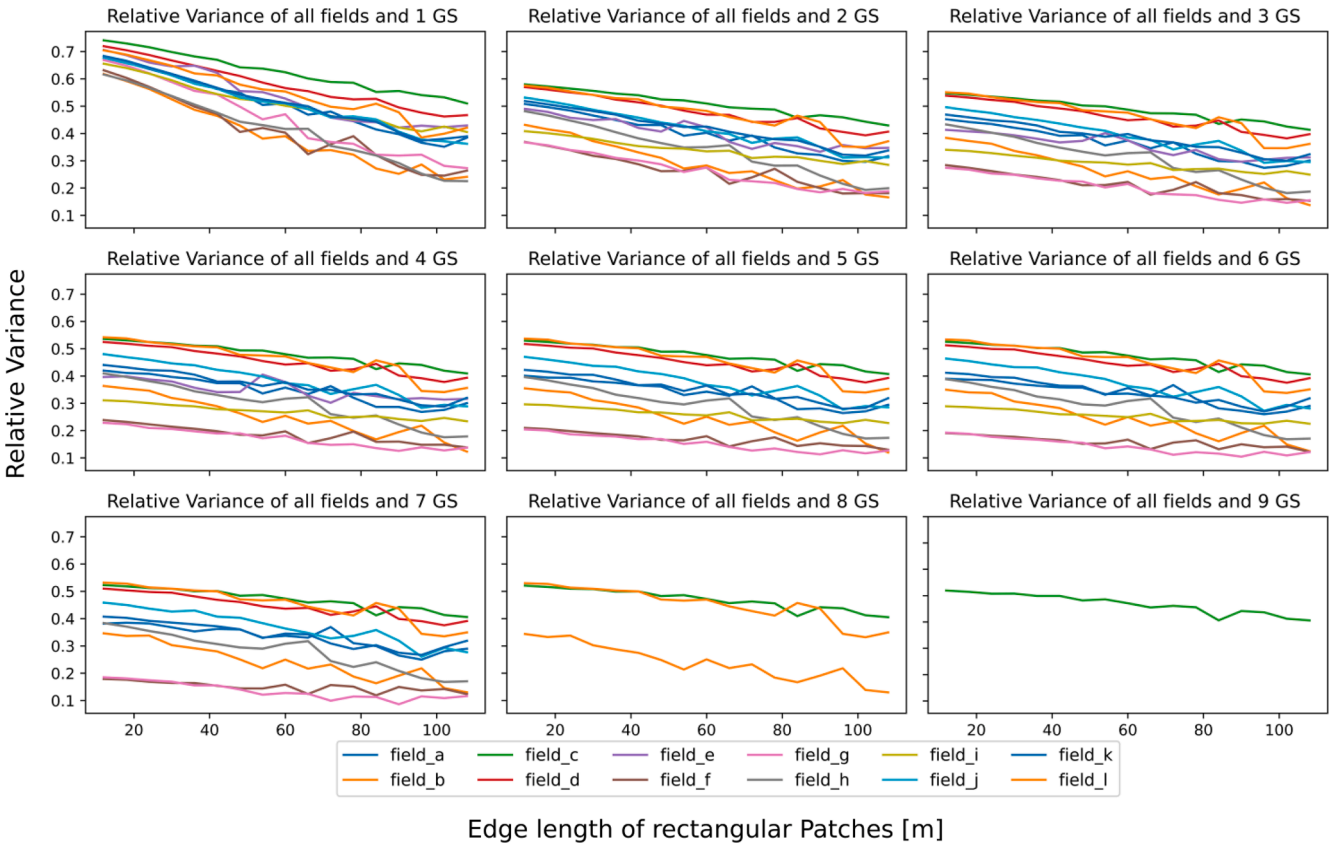


Fig. 12. Relative variance of different fields as a function of increasing subpatch sizes and increasing numbers of yield maps used from the corresponding GS.

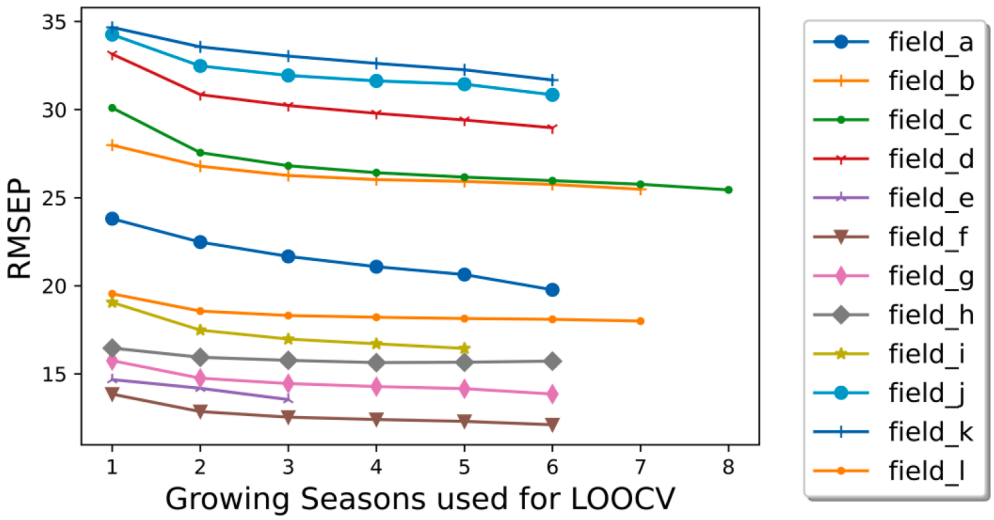


Fig. 13. RMSEP from LOOCV of all fields and all GS with subpatches with edge length of 72 *m*.

of data originating from homogeneous fields reduces the sum of variances, but not to the extend as in heterogeneous fields. The comparison of RV between fields can be used to draw conclusions about the heterogeneity of a single field. Since the RV of field_c is much higher than the RV of field_g, field_c is more heterogeneous in the distribution of yields than field_g.

3.2.2. LOOCV

In LOOCV analysis of multi-year yield maps, a decreasing RSMEP value was found with increasing number of GS used for simulating corn yield (Thorp et al., 2007). In this study, the results of LOOCV showed that the RMSEP for all fields and constant subpatch size decreased with an increasing number of GS (Appendix C). If more GS were used to cluster the data, the division into two yield potential zones was more robust.

RMSEP calculation showed differences between the individual fields (Fig. 13). While the RMSEP for field_k with one GS and subpatches with an edge length of 72 m was 34.7, it was only 14.1 for field_f with identical subpatch size and number of GS. This low RMSEP value for field_f can only be reduced by 1.7 (12.1%) to 12.7 by adding another 5 GS. With field_c, on the other hand, the RMSEP can be reduced by 15.6% from 30.1 to 25.4 by adding 7 more GS.

3.3. Patch cropping structure

The patch structures created according to the presented workflow divide previously uniformly cultivated fields into two different machine-manageable zones (Fig. 7B). Based on the concept of patch cropping, these zones would differ in future management. This will reduce the mean field size and increases the length of field boundaries per area. A number of positive changes in ecosystem services can be induced by reducing field sizes. Landscapes with smaller average field size compared to landscapes with large field size were found to have higher within-field biodiversity with higher abundance, and alpha, beta and gamma diversity of different taxa (Fahrig et al., 2015). The longer the field boundaries within a given area, the higher the species richness of all analyzed bird, plant, spider and bee taxa (Concepción et al., 2012). The positive effect of smaller fields on biodiversity was also observed with conventional agricultural management whereas in large fields with organic management, lower biodiversity was measured. This was mainly due to the fact that most farmland species were found in the area of the field boundaries (Batáry et al., 2017).

Our approach differs fundamentally from previous approaches in precision agriculture in general and from Management Zone Delineation (MZD) in particular, since they aim to identify zones with similar site conditions, similar yields or a combination of both and respond to this heterogeneity with one crop variety in all zones but differences in operational activities or inputs (Mulla and Khosla, 2015) whereas patch cropping relies on the change of the entire field geometry, size and crop selection to diversify agricultural production. Frequently, yield limiting factors or variables which lead to different crop responses in space and time (chemical or physical soil properties, landscape attributes or crop properties, etc.) are taken into account for MZD. The management zones thus correspond to a homogeneous combination of yield limiting factors that can lead to site-specific crop management recommendations (Khosla et al., 2010). Only a few of these factors are potentially manageable such as nutrients (N, P and K), lime, irrigation water, seeding rate and seeding depth and herbicides.

We have used multi-annual yield maps that quantify the spatial and temporal yield variability of a field without including the individual factors that are responsible for the differences in crop performance. However, multi-annual yield maps have significant advantages for the identification of homogeneous zones. Using a high number of yield maps over multiple GS, not only permanent environmental factors are considered (soil properties or landscape attributes) but also dynamic factors (weather, crop type and crop rotation, disease, weeds or unusual management decisions) and the interactions between both factors are taken into account (Blasch et al., 2020). The higher the number of different GS used for clustering, the more likely it is that a wide range of different crops, crop rotations and climatic interactions, and thus spatial and temporal variation will be represented. This assumption is also supported by our decreasing RMSEP with increasing numbers of GS for LOOCV (Fig. 8). The identified patches can be planted with crops most suitable for that cluster group class. Which crop species or land use is most suitable for each cluster group class must be decided individually, taking into account local agronomic expertise.

Furthermore, our approach limits the number of cluster centers to two compared to traditional MZD and does not consider the optimal number of cluster centers per field. Normally, the optimal number of cluster centers is either set by default to three yield classes (Khosla et al., 2010) or six yield classes (Ping and Dobermann, 2003) identified by the fuzziness performance index or normalized classification entropy (Fridgen et al., 2004), the relative variance reduction (Xiang et al., 2007), the silhouette width (Reyes et al., 2019) or a combination of Xie-Beni index, Fukuyama Sugeno index and proportion exponent (Córdoba et al., 2016).

Another substantial difference between the patch cropping framework and MZD approaches is the fact that patch sizes and the alignment of the patches along the permanent traffic line were considered which allows the use of the farmerś existing machinery and working widths. In traditional MZD approaches, fragmented smaller zones are often simplified into spatially contiguous zones after clustering by applying spatial filtering algorithm like Median filter or Majority filter (Ping and Dobermann, 2003) or the application of spatial correlation analysis before clustering to generate larger management zones with smoother boundaries (Gavioli et al., 2016).

Field heterogeneity is not expressed in regular local patterns, leading to single cases where individual patches of a cluster group class are isolated next to patches of another cluster group class. As depicted in Fig. 7, a single subpatch in the north-eastern area belongs to cluster group class two. This subpatch is surrounded by patches of cluster group class one. The cultivation of this single subpatch would lead to enormous additional operating costs as this subpatch requires to be accessed separately each time. To optimize the economic aspects, this subpatch should be assigned to cluster group class one or surrounding patches should be assigned to cluster group class two to increase the resulting patch area that can be cultivated uniformly. However, this assignment would reduce the RV and thus weaken the power of the cluster analysis. But even if individual patches are too large, this can limit certain applications. An approach to consider the total size of individual patches could be an iterative process after cluster analysis that determines the size of connected subpatches of the same cluster group classes. If these connected subpatches are too small or too large (min. and max. size thresholds), the patches are either merged if they are too small or split if they are too large. The possibilities of merging and splitting should be tested in future studies.

Patch cropping is an interesting use case for crop production using agricultural robots. Changing field size and shape affects the time ratio of typical field operations such as working in the main field and turning in the headland. Due to more complex geometries and the increased headland per total area, field productivity decreases and fuel consumption increases compared to large fields (Janulevičius et al., 2019). Cost analyses of autonomous vehicles in agriculture showed that fieldwork is the most time-consuming operation. With a high autonomy rate of fieldwork, the operating costs can be reduced compared to manned tractors (Lagnelöv et al., 2021). Swarm robotics can have a significant influence on the economic profitability of small-scale farming. Lowenberg-DeBoer et al. (2021) showed that the use of swarm robotics would increase the gross margin even when managing small and irregular shaped fields.

4. Conclusion

This paper presents a framework to design novel cropping systems that increase diversity in agricultural landscapes. Large fields with heterogeneous site conditions were divided into smaller homogeneous units called ‘patches’ according to their yield potential. A step by step description to divide fields into high yield and low yield potential zones of machine manageable units using multi-year yield maps is provided with an automated workflow in the Python programming language, taking into account maximum working widths and permanent traffic lanes. The reduction of variance by the used cluster algorithm depends on the within-field heterogeneity. The subpatch size, the number of GS used for clustering, and the parallel shift of the patch structures along the permanent traffic lane resulted in a change of the relative variance. Independent cross validation showed that the predictive performance increased with increasing number of GS used for clustering. Patch cropping has the potential to increase biodiversity and resource use efficiency through site-specific farming, reduces mean field sizes and is a potential use case for future autonomous robotic fieldwork.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

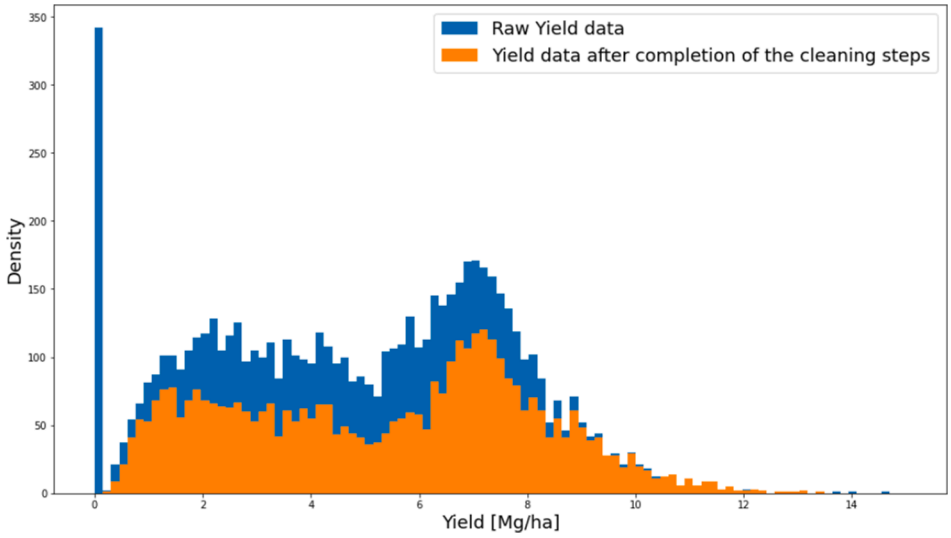
Acknowledgements

This work was made possible through funding from the Digital Agriculture Knowledge and Information System (DAKIS) Project (031B0729A), financed by the German Federal Ministry of Education and Research (BMBF). The authors acknowledge the support from the German Research Foundation under Germany's Excellence Strategy, EXC-2070 – 390732324 – PhenoRob.

Appendix

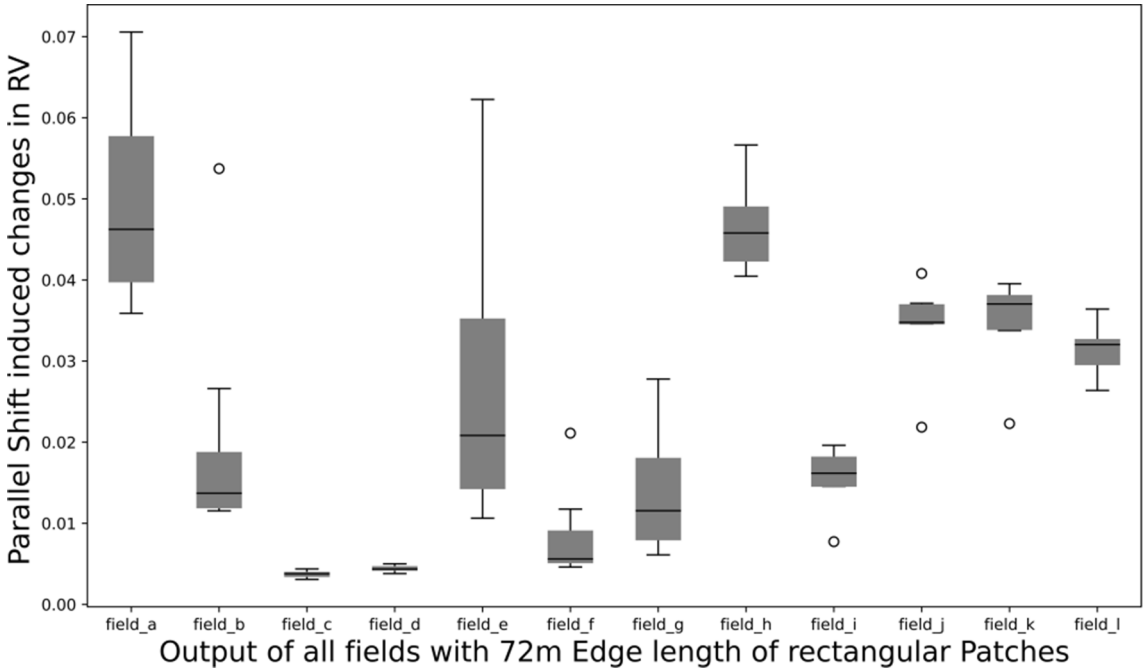
Appendix A

Histogram of 2015 Rapeseed yield data from field_d; Blue bins are raw yield data with standard moisture content; orange bins are yield data with standard moisture content after completion of the cleaning steps



Appendix B

Boxplot of all maximum RV differences by parallel shift for all fields and all GS for a patch size of 72 *m*



Appendix C

Results of the Root mean squared error of prediction (RMSEP) of standardized yield [%] of 12 fields and maximum of 8 GS from LOOCV for three different subpatch sizes

GS_used	subpatch_size	Field_a	Field_b	Field_c	Field_d	Field_e	Field_f	Field_g	Field_h	Field_i	Field_j	Field_k	Field_l
1	36	24.1	28.7	30.8	33.2	15.3	14.7	16.7	16.2	19.6	34.7	35.7	19.4
	72	23.8	28.0	30.1	33.2	14.7	13.9	15.8	16.5	19.1	34.3	34.7	19.5
	108	23.6	29.0	30.8	32.3	14.4	13.5	15.3	16.8	19.5	35.3	35.8	20.6
2	36	22.6	26.8	27.5	30.4	14.3	13.2	15.2	15.5	17.6	32.2	33.5	18.0
	72	22.5	26.8	27.6	30.9	14.2	12.9	14.8	15.9	17.5	32.5	33.6	18.6
	108	22.7	28.1	28.6	30.7	13.6	12.8	14.5	16.4	17.8	34.4	34.8	19.7
3	36	21.9	26.0	26.5	29.6	13.4	12.9	14.8	15.2	16.9	31.2	32.8	17.5
	72	21.7	26.3	26.8	30.2	13.5	12.5	14.5	15.8	17.0	31.9	33.0	18.3
	108	22.2	27.8	27.9	30.2	12.9	12.6	14.3	16.2	17.2	33.9	34.5	19.5
4	36	21.4	25.6	26.0	29.0		12.7	14.6	15.0	16.6	30.8	32.4	17.3
	72	21.1	26.0	26.4	29.8		12.4	14.3	15.6	16.7	31.6	32.6	18.2
	108	21.5	27.6	27.5	29.9		12.4	14.1	16.1	16.9	33.6	34.2	19.3
5	36	20.8	25.2	25.7	28.6		12.5	14.4	14.9	16.2	30.6	32.1	17.1
	72	20.6	25.9	26.2	29.4		12.3	14.2	15.7	16.4	31.4	32.3	18.1
	108	20.9	27.4	27.2	29.6		12.3	14.0	15.9	16.6	33.2	33.9	19.2
6	36	19.9	24.9	25.5	27.9		12.2	14.0	14.8		30.1	31.6	16.9
	72	19.8	25.8	26.0	29.0		12.1	13.9	15.7		30.8	31.7	18.1
	108	20.3	27.2	27.0	29.2		12.1	13.6	16.0		32.4	33.3	19.0
7	36		24.4	25.4								35.7	16.7
	72		25.5	25.8								34.7	18.0
	108		26.7	26.9								35.8	18.8
8	36			25.1									
	72			25.4									
	108			26.7									

References

Anselin, L., 1995. Local indicators of spatial association—LISA. Geographical analysis 27, 93–115.

Anselin, L., 2019. A local indicator of multivariate spatial association: extending Geary's C. Geographical Analysis 51 (2), 133–150.

Arslan, S., Colvin, T.S., 2002. An evaluation of the response of yield monitors and combines to varying yields. Precis. Agric. 3, 107–122.

Baillod, Bosem, Tscharntke, Teja, Clough, Yann, Batáry, Péter, et al., 2017. Landscape-scale interactions of spatial and temporal cropland heterogeneity drive biological control of cereal aphids. J. Appl. Ecol. 54 (6), 1804–1813. https://doi.org/10.1111/1365-2664.12910.

Basso, B., 2003. Perspectives of Precision Agriculture in Conservation Agriculture. In: García-Torres, L., Benites, J., Martínez-Vilela, A., Holgado-Cabrera, A. (Eds.), Conservation Agriculture: Environment, Farmers Experiences, Innovations, Socio-economy, Policy. Springer, Netherlands, Dordrecht, pp. 281–288.

Basso, B., 2021. Precision conservation for a changing climate. Nature Food 2 (5), 322–323.

Batáry, P., Gallé, R., Riesch, F., Fischer, C., Dormann, C.F., Mußhoff, O., Császár, P., Fusaro, S., Gayer, C., Happe, A.-K., Kurucz, K., Molnár, D., Rösch, V., Wietzke, A., Tscharntke, T., 2017. The former Iron Curtain still drives biodiversity-profit trade-offs in German agriculture. Nat Ecol Evol 1 (9), 1279–1284.

Bezdek, J.C., 1981. "Pattern Recognition with Fuzzy Objective Function Algorithms," 1/Ed. Springer, Boston, MA.

Blackmore, S., 2000. The interpretation of trends from multiple yield maps. Comput. Electron. Agric. 26 (1), 37–51.

Blasch, G., Li, Z., Taylor, J.A., 2020. Multi-temporal yield pattern analysis method for deriving yield zones in crop production systems. Precis. Agric. 21 (6), 1263–1290.

Brandes, E., McNunn, G. S., Schulte, L. A., Bonner, I. J., Muth, D., Babcock, B. A., Sharma, B., and Heaton, E. A. (2016). Subfield profitability analysis reveals an economic case for cropland diversification. *Environmental Research Letters* **11**, 014009.

Capmourteres, V., Adams, J., Berg, A., Fraser, E., Swanton, C., Anand, M., 2018. Precision conservation meets precision agriculture: A case study from southern Ontario. Agric. Syst. 167, 176–185.

Concepción, E.D., Díaz, M., Kleijn, D., Báldi, A., Batáry, P., Clough, Y., Gabriel, D., Herzog, F., Holzschuh, A., Knop, E., Marshall, E.J.P., Tscharntke, T., Verhulst, J., 2012. Interactive effects of landscape context constrain the effectiveness of local agri-environmental management. J. Appl. Ecol. no-no.

Córdoba, M.A., Bruno, C.I., Costa, J.L., Peralta, N.R., Balzarini, M.G., 2016. Protocol for multivariate homogeneous zone delineation in precision agriculture. Biosyst. Eng. 143, 95–107.

Coronel, E.G., Alesso, C.A., Bollero, G.A., Armstrong, K.L., Martin, N.F., 2020. Field-specific yield response to variable seeding depth of corn in the Midwest. Agrosystems, Geosci. Environ. 3, e20034.

da Silva, E.E., Baio, F.H.R., Teodoro, L.P.R., Campos, C.N.S., Plaster, O.B., Teodoro, P.E., 2021. Variable-rate seeding in soybean according to soil attributes related to grain yield. Precis. Agric. 1–17.

Ditzler, L., Driessen, C., 2022. Automating Agroecology: How to Design a Farming Robot Without a Monocultural Mindset? J. Agric. Environ. Ethics 35, 1–31.

Donal, P.F., Gree, R.E., Heath, M.F., 2001. Agricultural intensification and the collapse of Europe's farmland bird populations. Proc Biol Sci 268 (1462), 25–29.

Fahrig, L., Girard, J., Duro, D., Pasher, J., Smith, A., Javorek, S., King, D., Lindsay, K.F., Mitchell, S., Tischendorf, L., 2015. Farmlands with smaller crop fields have higher within-field biodiversity. Agric. Ecosyst. Environ. 200, 219–234.

Fontanet, M., Scudiero, E., Skaggs, T. H., Fernàndez-Garcia, D., Ferrer, F., Rodrigo, G., and Bellvert, J. (2020). Dynamic management zones for irrigation scheduling. *Agricultural Water Management* **238**, 106207.

Fridgen, J.J., Kitchen, N.R., Sudduth, K.A., Drummond, S.T., Wiebold, W.J., Fraisse, C. W., 2004. Management Zone Analyst (MZA) Software for Subfield Management Zone Delineation. Agron. J. 96, 100–108.

Garibaldi, L.A., Carvalheiro, L.G., Vaissière, B.E., Gemmill-Herren, B., Hipólito, J., Freitas, B.M., Ngo, H.T., Azzu, N., Sáez, A., Åström, J., An, J., Blochtein, B., Buchori, D., García, F.J.C., Oliveira da Silva, F., Devkota, K., Ribeiro, M.d.F., Freitas, L., Gaglianone, M.C., Goss, M., Irshad, M., Kasina, M., Filho, A.J.S.P., Kiill, L. H.P., Kwapong, P., Parra, G.N., Pires, C., Pires, V., Rawal, R.S., Rizali, A., Saraiva, A. M., Veldtman, R., Viana, B.F., Witter, S., Zhang, H., 2016. Mutually beneficial pollinator diversity and crop yield outcomes in small and large farms. Science 351 (6271), 388–391.

Gavioli, A., de Souza, E.G., Bazzi, C.L., Guedes, L.P.C., Schenatto, K., 2016. Optimization of management zone delineation by using spatial principal components. Comput. Electron. Agric. 127, 302–310.

Govers, G., Merckx, R., van Wesemael, B., Van Oost, K., 2017. Soil conservation in the 21st century: why we need smart agricultural intensification. Soil 3, 45–59.

Grahmann, K. R., Moritz; Hernandez-Ochoa,Ixchel; Ewert,Frank (2021). An agricultural diversification trial by patchy field arrangements at the landscape level: The landscape living lab "patchCROP". *Aspects of Applied Biology* **146**.

Haan, N. L., Iuliano, B. G., Gratton, C., and Landis, D. A. (2021). Designing agricultural landscapes for arthropod-based ecosystem services in North America. *In* "Advances in Ecological Research", Vol. 64, pp. 191-250. Elsevier.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., de Kroon, H., Lamb, E. G., 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLoS ONE 12 (10), e0185809.

Janulevičius, A., Šarauskis, E., Čiplienė, A., Juostas, A., 2019. Estimation of farm tractor performance as a function of time efficiency during ploughing in fields of different sizes. Biosyst. Eng. 179, 80–93.

Khanna, M., Chen, L., Basso, B., Cai, X., Field, J.L., Guan, K., Jiang, C., Lark, T.J., Richard, T.L., Spawn-Lee, S.A., Yang, P., Zipp, K.Y., 2021. Redefining marginal land for bioenergy crop production. GCB Bioenergy 13 (10), 1590–1609.

Khosla, R., Westfall, D.G., Reich, R.M., Mahal, J.S., Gangloff, W.J., 2010. Spatial Variation and Site-Specific Management Zones. In: Oliver, M.A. (Ed.), Geostatistical Applications for Precision Agriculture. Springer, Netherlands, Dordrecht, pp. 195–219.

Lagnelöv, O., Dhillon, S., Larsson, G., Nilsson, D., Larsolle, A., Hansson, P.-A., 2021. Cost analysis of autonomous battery electric field tractors in agriculture. Cost analysis of autonomous battery electric field tractors in agric. *biosystems eng.* 204, 358–376.

Landis, D.A., 2017. Designing agricultural landscapes for biodiversity-based ecosystem services. Basic Appl. Ecol. 18, 1–12.

Lowenberg-DeBoer, J., Franklin, K., Behrendt, K., Godwin, R., 2021. Economics of autonomous equipment for arable farms. Precis. Agric. 22 (6), 1992–2006.

Mittermayer, M., Gilg, A., Maidl, F.-X., Nätscher, L., Hülsbergen, K.-J., 2021. Site-specific nitrogen balances based on spatially variable soil and plant properties. Precis. Agric. 22 (5), 1416–1436.

Mulla, D., Khosla, R., 2015. Historical Evolution and Recent Advances in Precision Farming. In: Soil-Specific Farming Precision Agriculture. Rattan Lal , B.A. Stewart, Boca Raton, p. 36.

Ortega, R.A., Santibáñez, O.A., 2007. Determination of management zones in corn (Zea mays L.) based on soil fertility. Comput. Electron. Agric. 58 (1), 49–59.

Ping, J.L., Dobermann, A., 2003. Creating spatially contiguous yield classes for site-specific management. Agron. J. 95 (5), 1121–1131.

Rey, S.J., Anselin, L., 2010. PySAL: A Python library of spatial analytical methods. *In* "Handbook of applied spatial analysis". In: Fischer, M.M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 175–193.

Reyes, J., Wendroth, O., Matocha, C., Zhu, J., 2019. Delineating Site-Specific Management Zones and Evaluating Soil Water Temporal Dynamics in a Farmer's Field in Kentucky. Vadose Zone J. 18 (1), 1–19.

Ross, T.J. (Ed.), 2010. Fuzzy Logic with Engineering Applications. Wiley.

Šálek, M., Hula, V., Kipson, M., Daňková, R., Niedobová, J., Gamero, A., 2018. Bringing diversity back to agriculture: Smaller fields and non-crop elements enhance biodiversity in intensively managed arable farmlands. Ecol. Ind. 90, 65–73.

Shannon, D.K., Clay, D.E., Kitchen, N.R., 2020. Precision agriculture basics. John Wiley & Sons.

Simbahan, G.C., Dobermann, A., Ping, J.L., 2004. Screening yield monitor data improves grain yield maps. Agron. J. 96 (4), 1091–1102.

Sudduth, K., T Drummond, S., and Brenton Myers, D. (2012). Yield Editor 2.0: Software for Automated Removal of Yield Map Errors. *In* "2012 Dallas, Texas, July 29 - August 1, 2012". ASABE, St. Joseph, MI.

Thorp, K.R., Batchelor, W.D., Paz, J.O., Kaleita, A.L., DeJonge, K.C., 2007. Using cross-validation to evaluate CERES-Maize yield simulations within a decision support system for precision agriculture. Trans. ASABE 50, 1467–1479.

Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I., Thies, C., 2005. Landscape perspectives on agricultural intensification and biodiversity - ecosystem service management. Ecol. Lett. 8, 857–874.

van Ouwerkerk, C., and Soane, B. D. (1994). Chapter 26 - Conclusions and Recommendations for Further Research on Soil Compaction in Crop Production. *In* "Developments in Agricultural Engineering" (B. D. Soane and C. van Ouwerkerk, eds.), Vol. 11, pp. 627-642. Elsevier.

Vega, A., Córdoba, M., Castro-Franco, M., Balzarini, M., 2019. Protocol for automating error removal from yield maps. Precis. Agric. 20 (5), 1030–1044.

Wegener, J.-K., Urso, L.-M., Hörsten, D. v., Hegewald, H., Minßen, T.-F., Schhattenberg, J., Gaus, C.-C., Witte, T. d., Nieberg, H., Isermeyer, F., Frerichs, L., and Backhaus, G. F. (2019). Spot farming - an alternative for future plant production Spot Farming *Journal für Kulturpflanzen* **71**, 70-89.

Wik, M., Pingali, P., Broca, S., 2008. Background Paper for the World Development Report 2008: Global Agricultural Performance: Past Trends and Future Prospects. World Bank, Washington, DC.

Xiang, L.I., Yu-chun, PAN, Zhong-qiang, G.E., Chun-jiang, ZHAO, 2007. Delineation and scale effect of precision agriculture management zones using yield monitor data over four years. Agric. Sci. China 6 (2), 180–188.