

Analysis of clustering methods for crop type mapping using satellite imagery

Antonio J. Rivera ^{a,*}, María D. Pérez-Godoy ^a, David Elizondo ^b, Lipika Deka ^b, María J. del Jesus ^a

^a Department of Computer Science, Universidad de Jaén, Spain

^b Department of Computer Science and Informatics, De Montfort University, UK

ARTICLE INFO

Article history:

Received 12 March 2021

Revised 21 February 2022

Accepted 3 April 2022

Available online 6 April 2022

2000 MSC:

0000

1111

Keywords:

Unsupervised learning
Clustering
Crop mappings
Time series
Classification
Satellite imagery

ABSTRACT

With the current challenges in population growth and scarceness of food, new technologies are emerging. Remote sensing in general and satellite imagery more specifically are part of these technologies which can help provide accurate monitoring and classification of cultivars. Part of the increase in the use of these technologies has to do with the ongoing increment on the spatial-temporal resolution together with the free availability of some of these services.

Typically time series are used as a pre-processing technique and combined with supervised learning techniques in order to build models for crop type identification in remote images. However, these models suffer from the lack of labelled data sets needed to train them. Unsupervised classification can overcome this limitation but has been less frequently used in this research field.

This paper proposes to test and analyse the performance of several unsupervised clustering algorithms towards crop type identification on remote images. In this manner combinations of clustering algorithms and distance measures, a key element in the behaviour of these algorithms, are studied using an experimental design with more than twenty datasets built from the combinations of five crops and more than 45000 parcels. Results highlight better clustering methods and distance measures to create accurate and novel crop mapping models for remote sensing images.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The world population has reached 7.8 billion and it keeps growing, making food security a real challenge. To meet the needs of an ever growing population agricultural production must increase within the available resources. In other words smart farming practices must be adopted in order to meet, what seems like, a very ambitious aim of the United Nation's Sustainable Development Goal No. 2 of zero hunger by 2030 [1]. As of 2019, 47 million children aged 5 or below are malnourished and the current Covid-19 pandemic has only worsened the already bleak scenario. Smart agricultural practices have shown not only to improve production but also to be of great benefit for the environment. One of the most important tasks in smart agriculture is the creation of crop type maps.

As the spatial and temporal resolution of satellite images keep improving [2], they continue to be successfully used to solve diverse agricultural issues including: monitoring growth, predicting productivity, disease detection, etc. The combination of remote and ground sensing, powered with machine learning based tools, defines a new area known as Precision Agriculture. This area can be defined as a management strategy that collects, analyzes and processes temporal, spatial and individual data. This data is then combined with other information to support managerial decisions. As a consequence, this improves resource use, efficiency, productivity and quality leading to a more sustainable agricultural production [3].

Satellite missions can obtain multi-spectral data/images, each for a band in the different parts of the spectrum: visible, near infrared, short wave infrared, etc. The data is either directly analysed as images or surface reflectance information is combined in one or more bands to derive certain vegetation property through Vegetation Indices (VIs) [4]. Normalized Difference Vegetation Index (NDVI) [5] is one of the most used VI and it is defined as the differenced ratio of reflectance in the red and near-infrared wavelengths.

* Corresponding author.

E-mail addresses: arivera@ujaen.es (A.J. Rivera), lperez@ujaen.es (M.D. Pérez-Godoy), elizondo@dmu.ac.uk (D. Elizondo), lipika.deka@dmu.ac.uk (L. Deka), mjjesus@ujaen.es (M.J. del Jesus).

Machine Learning methods have been successfully used to solve problems related to agriculture and satellite images [6]. These methods include supervised and unsupervised algorithms.

Supervised learning algorithms require that every data sample is labelled and models learn according these labels. These methods are extensively used in precision agriculture with satellite data. However they present several limitations which need to be addressed before they can be massively adapted. One such limitation is the need for large amounts of labelled data. Without such data, supervised methods tend to produce rather low generalisation levels [2]. However, when accessing satellite data, it can be observed that most of the areas and croplands are unlabeled. It is true that there are some labeled databases, but they are expensive and difficult to maintain, so they are often out of date [7]. Thus, with the current availability of satellite data, there is a need to develop techniques that do not require data labeling, unsupervised learning approaches. They remain relatively unexplored in this field of application. Despite their potential drawbacks, the main advantage of them, if properly implemented, is clear: crop type maps can be created where little or no information exists on the ground the field level. Therefore, these techniques can be applied, *a priori*, satellite images anywhere in the world.

In *unsupervised learning* there are no labels and data samples are clustered together according to their similitude. Clustering algorithms [8] are amongst the most popular unsupervised methods currently used. They have proven to be useful in crop type mapping [7] as well as in soil and its associated properties. Nevertheless, in the bibliography a deep comparative study about characteristics and performance of clustering method applied to crop type mapping can not be found.

The main objective of this paper is to analyse the behaviour that clustering methods can offer towards the problem of crop type mapping based on satellite images. For this purpose, external validation indices, that measure the similarity between clusters (predicted labels) obtained by the clustering algorithms and the real labels (crops) of the instances, are used. The proposed analysis is to perform a comparative using some of the most popular clustering methods to perform crop type mapping. The distance measure is one of the most important parameters of any clustering algorithm. Therefore, this study will perform an analysis combining several clustering algorithms and distance measures.

This paper is an extension of a preliminary study presented in [9]. In the preliminary paper, an experimental study comparing traditional two clustering algorithms for the classification of olive trees and cereal crops from time series remote sensing in one parcel data was presented. The main contributions of this extended paper are:

- The study of 16 clustering methods with the combination of 4 clustering algorithms and 4 distance measures.
- The use of multi-spectral satellite data with 5 crops over more than 45000 parcels.
- A preprocessing data stage to clean and to impute missing values.
- The creation of 26 datasets which combine a total of 5 crops.
- The determination of:
 - the best performing algorithms for the problem of land usage with a varying number of crops
 - the best performing algorithms for the problem of land usage with a fixed number of crops (2, 3 and 4)
 - the best performing distance measure for each clustering algorithm.

The rest of the paper is organised as follows: Section 2 details the state of the art in the area of using remote sensing data in agriculture and highlights commonly used clustering techniques. Section 3 describes the adopted methodology. Section 4 explains the

experimental design and show the results and their analysis. Finally, Section 5 outlines the main conclusions.

2. Background

Remote sensing data, i.e sensing with a piece of equipment that is not touching the study object, is derived from either satellite hosted sensors or aerial sensors mounted on drones, unmanned aerial vehicles or other light-weight air crafts. Non-satellite based approaches are resource intensive and severely limited in terms of coverage. This work concentrates on data acquired from satellites. Information provided by satellite based data can be applied into a number of fields including: agriculture, meteorology and geology.

Satellites have been used for agricultural applications since the 1970s. However, data was limited to only a few key visible and near infrared bands in the early days. Today data is largely available from a much larger portion of the spectrum ranging from ultraviolet all the way to microwave. The focus is on hyperspectral and thermal sensing. Information provided by the satellites come in different resolutions and can provide a variety of information depending on the sensors that the satellite is equipped with. High resolution satellites can provide images with resolutions between 30 cm and 5 m per pixel. Low resolution satellite images cover over 60 m per pixel [10]. Detailed cultivars information is provided by satellites within the different spectral bands. This information includes plant nutrient levels and soil moisture contents [11]. There are public and private domain satellites. Amongst the most popular public satellites there is Galileo, Lansat, MODIS, Meteosat and Sentinel. Within the private ones it is worth mentioning Spot, ImageSat and China Siwei. A key contribution in this field is that of the High Resolution Land Cover (HRLC) project which was developed by the European Space Agency (ESA) [12]. Its main objective is the accurate description of land cover, as well as the analysis of its changes, using high-resolution images, mainly in areas such as: Amazonia, West Africa and West Siberia. Important applications of this project are in the field of energy and climate modelling.

This section introduces the use of remote sensing data in agriculture. More specifically, the use of vegetation indices is discussed. A summary of some of the most common machine learning methods applied to remote sensing is provided. These include both supervised and unsupervised methods and their use in agricultural based applications.

2.1. Remote sensing applied to agriculture

Remote Sensing, in the form of satellite images, has been used in crops to measure productivity, vigour, chlorophyll, foliar damage from pests or disease and presence of weeds [2]. This information plays a key role towards crop management. It provides producers with ways of predicting yields and optimising both the use of agro chemicals and of irrigation strategies.

An example of the use of satellite images to optimise the use of irrigation systems can be found in Khanal et al. [13]. Here the authors refer to the use of normalized difference vegetation index (NDVI) images and soil adjusted vegetation indexes (SAVI) to diagnose water stress and soil moisture conditions for a variety of crops.

Remote sensing is also used to estimate evapotranspiration [14,15]. Deductions made are used to estimate crop water requirements and schedule irrigation.

Several studies have used satellite images to calculate soil moisture [14] [16] [17]. These authors document the use of multiple spectral bands which include optical, thermal and microwave

information. Microwave data provides the most efficient way to predict soil moisture.

Plant chlorophyll is highly correlated with photosynthetic activity and plant productivity. This provides critical information in terms of crop nutrients. NDVI and SAVI images have been successfully used to predict nutrient levels [18–20].

There has also been a number of successful applications of satellite images to disease control [21]. The earlier potential diseases can be detected, the better the farmer can control them. A combination of visible light, thermal, fluorescence and multi spectral images, have been successfully used to identify crop diseases.

Weed management is another area where the use of satellite images brings great benefits. By identifying the areas of the plantation which have higher concentrations of weeds, producers can better select the areas which need treatment [3]. A significant amount of progress has been made in this area with the use of artificial intelligence [22,23].

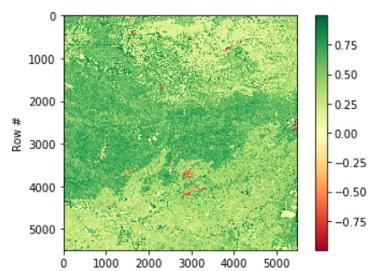
Last but not least, research has also been done towards the use of satellite images for crop monitoring and yield. This enables producers to efficiently plan their farm activities. Biomass and Leaf Area Index (LAI) are used to predict the development and the status of crops. LAI is also used as the basis for many yield predicting models [24].

2.1.1. Vegetation indices

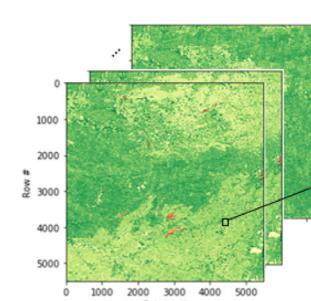
Vegetation indices (VI) provide useful information on land vegetation [11]. The most commonly used index is the NDVI. This corresponds to the differenced ratio of reflectance in the red and near-infrared wavelength [5] (see Eq. 1). This index provides information in terms of the growth, spatial density and phenology stages of the plants. It can also be used to estimate moisture, nitrogen and growth stages of plants. One of the main limiting factors of NDVI is the interference of soil reflectance. It is also limited when it comes to changes in leaf chlorophyll content in mature canopies [25]. Other indices used are Difference Vegetation Index (DVI), Enhanced Difference Vegetation Index (EDVI), Green Chlorophyll Index (GCI), Modified Soil Adjusted Vegetation Index 2 (MSAVI2) etc.

$$\text{NDVI} = \frac{(\text{NIR} - \text{RED})}{(\text{NIR} + \text{RED})} \quad (1)$$

The left hand side of Fig. 1 shows a plot of a single instance of an NDVI image. The right hand side illustrate a time series of NDVI images together with the evolution on time of NDVI values over a single image pixel.



(a) Single NDVI instance



(b) Time Series NDVI

Fig. 1. Single instance and time series plots for NDVI values.

Time series from vegetation indices. Time series is a common technique used in combination with vegetation indices. It consists of measuring the reflectance level per pixel. This reflectance level depends on the type of plant, level of chlorophyll, etc. Therefore, it reflects changes throughout the growth cycle of the plant. An example can be seen on the right side of Fig. 1, where the reflectance level has been collected in the NDVI index.

In Schultz et al. [26], the authors present a study on the performance of vegetation indices from Landsat time series in deforestation monitoring. The authors managed to improve the spatial accuracy level of their model by fusing VI on feature level. This also had an effect on the reduction of the overestimation of vegetation change. The best results were obtained by using NDVI and NDFI.

Another example for the measure of seasonal pattern vegetation indices using time series was presented by Panuju et al. [27]. In this paper, the authors present a study modelling seasonal patterns of vegetative cover based on NDVI time series. The authors used SPOT-VGT images to predict phenology stages. They compared NDVI range variations amongst several crops. Some of these variations might be related to gaps in plantations.

The authors in Kalisa et al. [28], present a study on the assessment of climate impact on vegetation dynamics over East Africa. They used a combination of radiometer and climate data sets to correlate climate change and agricultural practices towards severe vegetation degradation. By analysing vegetation levels within El Niño years, the authors showed a weak connection between this meteorological event and vegetation growth changes in croplands. They also highlighted a stronger correlation between NDVI and precipitation than with temperature.

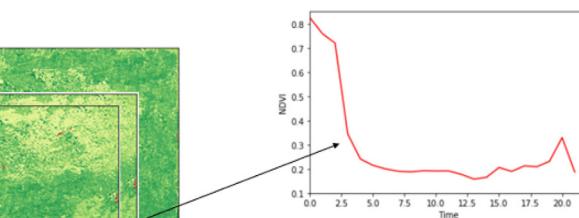
2.2. Machine learning methods for remote sensing

Most of the machine learning techniques applied to the field of remote sensing are either supervised or unsupervised. The following two subsections introduce both approaches together with some examples of their use towards remote sensing.

2.2.1. Supervised learning

Supervised learning techniques require for the data samples to be labelled. Some examples of the use of these methods towards remote sensing are provided in this sub section.

Conrad et al. [29] uses NDVI data from the MODIS satellite to classify crop rotation patterns using classification and regression trees. Both methods propose the dividing and aggregating of data over a temporal dimension. This allows for temporal differences to be accounted for.



Support Vector Machine (SVM) is a popular method used for crop classification [30–32]. Low et al. [33] use SVM while recognising that most alternative methods fall short due to the large number of features available. Hence, it uses data from the multispectral RapidEye system and employs the Random Forest method to reduce the dimension of the feature matrix. k-nearest neighbour (kNN) and its variations have also been popular algorithms for the classification of remote sensing images. For instance in [34], kNN has been combined with SVM to improve the speed and accuracy of the models through local partitioning.

Two types of wheat have been classified in India using multi-temporal images from the Indian Remote Sensing satellite (IRS)-1B with variations of Maximum Likelihood Classification and Artificial Neural Network (ANN) [35]. Neural Network models have been quite extensively used either alone or in combination with other techniques such as seen above [36–39]. Murmu et al. [40] discusses the use of fuzzy logic combined neural networks for crop classification using remotely sensed data.

Decision Tree classification techniques have also shown good results [41] [42] [43]. For example, Yang et al. [41] use classification and regression trees to classify silage and grain corn. Their system is based on hyper-spectral images of experimental farming plots. They use a hyperspectral camera mounted on the top of a lightweight aircraft to achieve good levels of generalisation.

Omkar et al. [44], present a comparison study between two biologically-inspired techniques and a statistical technique for crop coverage identification from high resolution satellite images. The methods explored include Maximum Likelihood Classifier, Particle Swarm Optimisation (for evolving the Multi-layer Perceptron Neural Networks) and Ant Colony Optimisation technique (for extracting the rules of classification). The data sources come from images produced by the QuickBird Satellite Sensor with a resolution of 2.4 m [44]. Results show that there is not a one-size-fits-all crop classifier and proposes the multi-classifier solution.

Ji et al. [45], apply deep-learning techniques in the form of 3D Convolutional Neural Network to crop-classification. This builds on the research of Kussul et al. [46] where they used a 2D Convolutional Neural Network to learn spatial features and a 1D CNN to learn temporal features. The 3rd dimension in the presented work [45] captures the temporal information of crop evolution. However, such methods are extremely data hungry. [45] used the data from the Chinese high-resolution Earth imaging satellites, Gaofen 2 and Gaofen 1, with 4 m and 15 m ground resolution respectively, for training and testing an image classification model. Their model used the image directly without extracting information in the form of vegetation indices. This may lead to the incorporation of unnecessary information, sometimes acting as noise.

2.2.2. Unsupervised learning

In *unsupervised learning* the data samples contain no labels and they are clustered together based on their similarities [8]. These algorithms classify or group objects based on common features and pattern recognition [47,48]. Clustering usually follows the process of feature extraction or selection. This involves defining a measure or function, leading on to cluster validation where the results are subject to an evaluation criteria and interpretation of the results obtained. These methods aim to group data samples according to their similarity, so that the samples within a cluster are more similar than those that fall outside the cluster.

From an intuitive point of view, this problem has a very clear objective: to properly cluster an unlabeled dataset. Despite its intuitiveness, the notion of “clustering” cannot be precisely defined [48], since there are different ways of defining the concept of similarity. Thus, many algorithms and similarity measures have been developed that differ from each other according to what is meant by cluster (which, in essence, is given by how we define that two

objects are similar or different) and by the computational efficiency in achieving the final grouping.

Clustering algorithms[8,48] can be categorised into the following types:

- *Partition clustering methods*: In these methods a 2-step loop drives the formation of the clusters. First, the membership of each instance to a cluster is determined using a distance measure and then the center or centroid of the cluster is calculated in each iteration averaging the instances of the dataset. K-means [49] and K-medoids [50] are the most well-known examples of this category.
- *Agglomerative clustering*: Here clusters are built by merging (bottom-up approach) or splitting (top-down approach) patterns in an iterative way. Examples of this include: Dlvisive ANALysis (DIANA) [51] and Agglomerative [52]
- *Fuzzy clustering algorithms*: This kind of algorithms can be considered as a generalization of the Partition Clustering algorithms where a data point can belong to two clusters as a membership function in the interval [0, 1] [53]
- *Clustering algorithms based on density*: The core idea of these methods is to place data in regions with high density which are considered part of the same cluster. DBSCAN [54] and OPTICS [55] are representatives of this category.
- *Clustering based on graph theory*: In this case, the clustering is done by regarding nodes as data points and edges as relationships amongst these data points. Spectral clustering [56] is an example of this type of method. It groups data samples using eigenvectors of matrices derived from the data.
- *Affinity Propagation algorithms*: For this kind of algorithms [57] affinity is defined as the negative distance between two data points. All the data are considered as possible centers of a cluster. The bigger the affinity of one data sample is, the bigger is its probability of being a cluster center.

Clustering algorithms have been widely used in image processing tasks for the classification of part or entire images including satellite images [58]. More specifically, when applied to the problem of crop type mapping, most of the methods use extracted vegetation indices such as the NDVI and EVI [26]. They combine readings on different spectral bands which provide a physical characteristic that is unique to each plant over a temporal space.

3. Methodology

The objective of this paper is to test and to analyse the performance level of several clustering algorithms in combination with different distance measures, applied to the task of building crop type maps from remote sensing images. This experimentation can support the usefulness of clustering methods for unsupervised crop classification. Data for the experimentation was retrieved from the BreizhCrops dataset [59], a multispectral satellite dataset.

The use of clustering algorithms to create crop type maps is very attractive because they can be directly applied without any kind of previous learning phase or without having to have any previous model available. This is important when dealing with the creation of crop map models because there is a lack of labelled datasets. Furthermore, building datasets based on the ground truth is expensive and very resource intensive.

Two approaches will be followed in order to carry out a deeper analysis. The first consists of utilising the data set containing all crops identified on the hyper-spectral images. The second will include datasets that represent combinations of the different crops.

The results will allow to analyse how clustering algorithms respond to different data features, for example, to combinations

limited to only two crops, when the period of growth of the crops in the combination is either similar or different; or how these unsupervised methods behave when new crops are considered.

The algorithms used in the experimentation are described in the following sections.

3.1. Clustering

In the experimentation, clustering algorithms belonging to well-known categories are used. Two hierarchical algorithms (Agglomerative and DIANA), a partition clustering algorithm (PAM) and one based on graph theory (Spectral) are considered. A few other algorithms such as DBSCAN, Optics, Affinity propagation and Mean Shift clustering, were tested but discarded in the final analysis due to poor initial results. Additionally, these algorithms were chosen as they allow the use of different distance measures. They are briefly explained below.

- Hierarchical clustering: This group of methods creates a hierarchy of clusters that is represented by a tree structure known as a dendrogram. The root of the tree represents a single cluster in which all the data would be placed, while the leaves of the tree each represent clusters with a single element. Depending on the strategy followed to create the clusters, there are two types of Hierarchical clustering:

-Agglomerative [52] (bottom-up): In this case clustering starts from the leaves of the tree, where each element is a cluster and in each iteration the two closest clusters are grouped, so that they are combined until they converge into a single central branch.

-Divisive ANAlysis Clustering(DIANA) [51] (top-down): This algorithm starts with a single cluster with all the data. In each iteration divisions are performed on the cluster with the largest diameter i.e., the one with the largest difference between two of its observations. The data pair with the greatest average distance, with respect to the rest, is selected. This pair is used then to create a new cluster. The observations that are closest to this new cluster are then added.

- PAM (Partitioning Around Medoids) [50]: It is based on the K-medoids method within the partition clustering category. In the K-medoids method the data are grouped into k clusters where k is a parameter that has to be set. In this sense it is similar to k-means. The main difference between the two methods, is that in the k-medoids there is an element (medoid) within the cluster that represents each cluster, while in k-means a cluster is represented by the average of the elements belonging to that cluster (centroid).

The medoid is the element within a cluster whose average distance between itself and the rest of the elements within the cluster is the smallest possible. The use of medoids instead of centroids makes the method less susceptible to noise and outliers, which makes it more robust.

- Spectral clustering [56]: takes into account the relative position of the data in the data space. It uses the similarity matrix between the data to perform a dimensionality reduction before proceeding to cluster the data.

The similarity matrix containing the distances between each data pair is constructed from the dataset. Once obtained, a threshold is set to obtain the adjacency matrix. The values are set to 1 if the distance in the similarity matrix exceeds the threshold and to 0 otherwise. The adjacency matrix is used to construct the degree matrix which is a diagonal matrix. Each element of the diagonal is calculated based on the sum of the corresponding row in the adjacency matrix. The difference between the degree matrix and the adjacency matrix forms

the laplacian matrix.

Once a laplacian matrix is obtained, their components are used by the algorithm to cluster the data.

3.1.1. Distance measures and cluster quality indices

Distance measure or the measure of how similar data points are to each other, is a key component of a clustering algorithm. Some of the most well-known distance measures are used within the presented work, such as Canberra, Euclidean, Manhattan and Minkowski [60].

Other distance measures, including specific time series distance measures such as *TSclust* [60], were considered in this research but were discarded. Time series distance measures, such as the classical DTW (Dynamic Time Warping), did not provide good results. This could be due to facts such as: time series blue in the dataset used for this research only contains data for one year or season and therefore the typical seasonal repetitions on time series cannot be found. Furthermore, as will be explained below, in the preprocessing stage all the time series have been built with the same length (same dates), avoiding sample displacements. Normally, certain time series distances exploit the two previous characteristics, namely absence of seasonal variation and displacement, but in this case they are not met.

In the case of clustering, real labels, indicating to which cluster each observation belongs to, may or may not be available. Our objective is to demonstrate the reliability of the clustering methods when the crop type mapping task is addressed in the absence of real or ground truth labels. To support the evaluation of this objective, external validation indices that measure the similarity between the labels predicted by the clustering algorithms and the real labels (crops) are used.

As external validation measures of cluster quality, we have considered two of the best-known measures: Jaccard [61], a classical measure and Adjusted Mutual Information (AMI), a newer measure designed against chance [62].

3.2. Statistical test methodology

Hypothesis testing techniques are used to statistically support confidence in the results. The first option is to apply parametric tests. However, initial conditions that guarantee the reliability of these tests may not be satisfied in this problem. Therefore, statistical analyses carried out may lose credibility [63]. For this reason non-parametric tests, such as that stated in [64] are applied.

Following [64], first a multiple comparison test, Friedman [65], is applied to detect significant differences in results between two or more algorithms. As a result of this test, a *p*-value is obtained. If this *p*-value is lower than 0.1 or 0.05, the existence of significant differences can be established with a 90% or 95% level of confidence respectively.

The next step is to detect the specific pairs of methods between which significant differences exist. With this aim, three well-known methods will be applied: Nemenyi [66], Holm [67] and Shaffer [68]. Similarly, a *p*-value with the same interpretation as above is obtained.

4. Experimentation

In this section, the conducted experimentation will be described. First, traits of the dataset, such as the location, environmental conditions, sampling process, parcels or crop categories are detailed. Then, the preprocessing phase is described, in which the NDVI is calculated, the time series and some outliers are cleaned and missing or removed values are imputed. After that, the exper-

imental design is clarified. Following which, results are shown and finally these results are analysed using statistical tools.

4.1. Datasets

A multispectral satellite dataset, BreizhCrops dataset [59], built from Sentinel-2 mission imagery is used in the experiments. The monitored area is located in the northwest of France, specifically in the Brittany region. Breizh, as this region is popularly known as, has a typically continental climate with a moderate temperature range from 5.6 degrees centigrade in winter to 17.5 degrees centigrade in summer and around 650 mm of annual precipitation.

Data are sampled, once a week, from January 1, 2017 to December 31, 2017, and are available in two processing levels: Level 1C and Level 2A. Level 1C provides 13 spectral bands with raw reflectance at the top-of-atmosphere and Level 2A disposes 10 spectral bands with atmospherically corrected surface reflectances at the bottom-of-atmosphere. Level 2A data is chosen in our experimentation, as this kind of preprocessing may improve the results of the machine learning models when large areas are monitored in time [69].

For both levels, the reflectance values over the bounds of each parcel are averaged, where a parcel is the term as defined in the land registry [59] and, therefore, for each parcel of the dataset, a time series per reflectance band is available. The number of data points per time series varies between 27 and 67 points, due to some images, for example with more than 80% cloud cover, are discarded. In any case the existence of clouds in the sample effects the values of the reflectances (bands) and an additional preprocessing step is required.

The BreizhCrops dataset is mainly organized in 4 regions: FRH01, FRH02, FRH03 and FRH04. Due to large number of parcels (between 180,000 and 120,000 approximately) in each region, even after the preprocesing conducted by the authors of the dataset, only the region FRH01 was selected for this study. The number of original parcels of the region FHR01 is 221,095 and when the data is pre-processed as per the method used for the Level 2A data, the number of total observations is 178,632. In Fig. 2 the location FHR01 and its cultivated parcels are shown.

The parcels of the original dataset are labelled with 328 different crop identifiers grouped into 23 groups or categories. In our experimentation, 8 crop labels (categories): barley, wheat, rapeseed, corn, sunflower, orchards, nuts and permanent meadows,

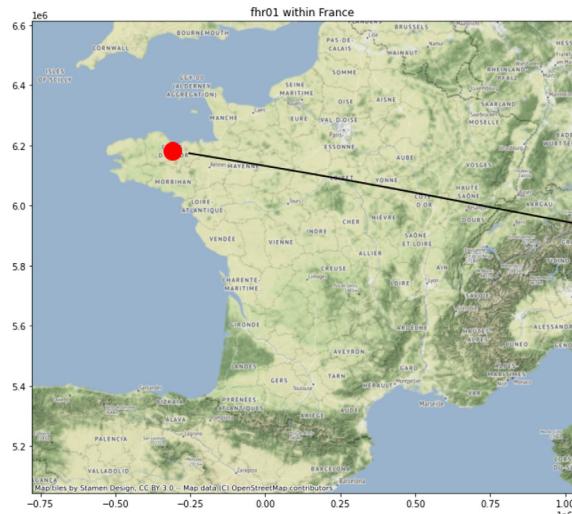


Fig. 2. Region FHR01.

Table 1
Crops and parcels selected for the experimentation.

Crop	Original parcels	Parcels after our preprocessing
Barley	13051	4918
Wheat	30380	8833
Rapeseed	5596	1560
Corn	44003	14254
Sunflower	1	0
Orchards	937	142
Nuts	10	4
Permanent meadows	32641	15443

have been selected in order to work with clearly defined classes avoiding categories such as, fodder crops, according to [59]. The number of parcels for each crop is in Table 1. With closer inspection of the data, crops such as sunflower, nuts and orchards with fruits are discarded due to low observations.

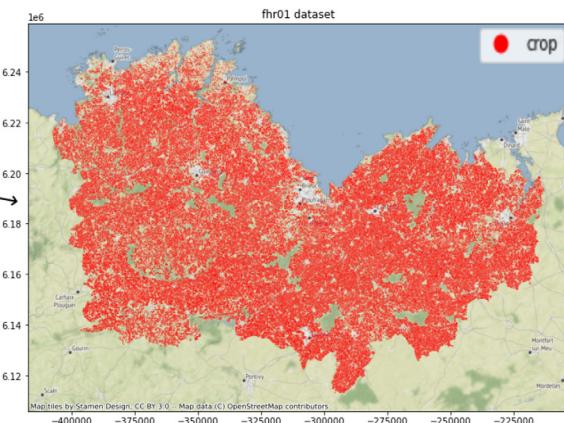
Considering all the combinations of the remaining crop categories namely, barley, wheat, rapeseed, corn and permanent meadows were used to build the datasets. Thus, 26 datasets were obtained: 10 datasets for the combinations of 2 crops, 10 datasets for the combinations of 3 crops, 5 datasets for the combinations of 4 crops and 1 dataset with the 5 crops.

4.2. Data preprocessing

The objective of this preprocessing step is to remove outliers and to achieve that all time series have the same number of points [59]. First of all the NDVI vegetation index is calculated from the Level 2A bands RED and NIR. As mentioned, the presence of clouds in samples, results in bias in the observations and cause outliers in the time series data. Due to the high number of parcels available in the region FHR01, time series with more than 6 (10%) negative points are removed. Thus, regarding our crop selection and the preprocessing phase, the remaining number of parcels was 45 153. The number of parcels for each crop is described in Table 1. Regarding these data, sunflower, orchards and nuts are removed from the experimentation because of its low number of parcels.

Next, points with a NDVI below 0.1 are considered outliers and are erased in the remaining time series. These low values of NDVI correspond to barren rock, sand, or snow.

Finally, linear interpolation to impute missed or removed data points is used, with the aim of obtaining time series with the same



length. This length was set to 67 points, as this is the maximum length of any of our time series.

The parameters of this preprocessing phase have been obtained in a heuristic manner after carry out different test with interpolation algorithms, value parameters, etc.

4.3. Experimental design and results

As mentioned in Section 3 our objective is to analyse the performance of clustering algorithms and distance measures in order to study the performance of these methods in creating crop maps in an unsupervised way.

With this aim a review of software implementations have been carried out. A set of clustering methods that allows for using pre-computed distances and software packages specialized in the calculation of distance measures were identified. The software packages used in this paper are *scikit-learn* from Python and *cluster* from R. The list of clustering methods are: Agglomerative, DIANA, PAM and Spectral. The Spectral implementation were taken from the *scikit-learn* package and the remaining from the *cluster* package. The distance measures used are: Canberra, Euclidean, Manhattan and Minkowsky, all from the *cluster* package of R.

In summary, 16 clustering methods obtained from the combinations of clustering algorithms (Agglomerative, DIANA, PAM and Spectral) and distance measures (Canberra, Euclidean, Manhattan, and Minkowsky) were tested on 26 datasets obtained from all the possible combinations of the mentioned crops: barley, wheat, rapeseed, corn and permanent meadows. All the methods have been applied with their default configuration values.

As cluster validation indices, the well-known Adjusted Mutual Info (AMI) and Jaccard measures were used.

Due to the extensive experiments conducted on the large data set, all the results and analysis could not be presented due to space constraints. However, those interested can find the detailed results in [70].

4.4. Analysis of the results

In this section the results obtained are analysed. More specifically, in the first subsection, 4.4.1, an exploratory analysis of the behaviour of clustering methods when addressing different crop combinations is illustrated and clarified.

After that, hypothesis testing techniques were used to provide statistical support to the analysis of the results. In particular, they were applied in order to elucidate the best clustering algorithm or distance measure in different scenarios such as:

1. Determining the best clustering algorithms, including distance measures, for all crop combinations (subSection 4.4.2). Therefore, this study considers all datasets together.
2. Determining the best clustering algorithms, including distance measures, taking into account the number of crops in the dataset (subSection 4.4.3). Therefore, three studies are carried out such that, one for datasets with two crops, one for datasets with three crops and a final study for datasets with three crops.
3. Determining the best distance measure for each clustering algorithm (subSection 4.4.4).

For reasons of space, pairwise clustering methods comparison tables of post hoc tests have been given in [71].

4.4.1. Exploratory analysis of the clustering methods performance

The objective in this section is to describe the behaviour of the clustering methods when crop type mapping problem is addressed.

First of all, it is necessary to point out the characteristics of data in the addressed problem. The data come from satellite sensors

that are affected by atmospheric phenomenons, where clouds are the most important ones and that causes a number of missing values, outliers, etc. This fact together with the similarity in the growth cycle of some crops significantly hampers the clustering task. Similarity in growth cycles results in similar NDVI signature and hence difficult to identify each crop uniquely.

Fig. 3 illustrates the ideal cluster prototypes or typical NDVI cultivars signatures calculated averaging the entire set of time series for each crop analysed within this research. Despite the preprocessing steps carried by the authors of the dataset and in the presented methodology, there are some discontinuities caused by periods of clouds cover.

As can be seen, prototypes for barley and wheat crops are very similar because their growth cycle is very similar in the peak around December and mainly in April, May and June. These facts determine the results obtained as is explained in more details below.

The outcomes of the methods for two crops with similar growth cycle, barley and wheat, are shown in **Table 2**. As can be observed, clustering methods show a low performance in this case, with best results are depicted in bold.

A visual analysis of these results is shown in **Fig. 4**, where the clusters obtained by the DIANA clustering algorithm (with Manhattan distance) is compared with the map of the ground truth of crop type (barley and wheat) in a central zone of the region FHR01. Some errors of the clustering algorithms are depicted by the circled areas.

On the other hand, the prototype for corn, with a main peak around August, is very different to the wheat one, and clustering methods obtain a very high performance for the dataset as can be seen in **Table 3**, with best results depicted in bold. In this case for both cluster validation external indices, AMI or Jaccard measures are used. The best method is Spectral_Canberra followed by Spectral_Manhattan, Spectral_Minkowski and Spectral_euclidean, that implies a good behaviour of the Spectral method.

In **Fig. 5**, a comparison between the map of real or ground truth information of crop type (wheat and corn) and that obtained by the Spectral clustering algorithm with the Canberra distance is represented. For better visualisation, the figure shows only the central area of the region FHR01. As can be seen there are few differences between the two map.

The behaviour of the clustering methods applied for the crop types mapping task, that can be clearly seen with two crops, is maintained for the remaining crop combination data set. For example, **Fig. 6** shows a comparison between real crop type map or the ground truth data and the obtained by Spectral clustering method (with Canberra distance) when all crops are taken into account. Again, it is observed that the clustering algorithm has problems distinguishing between barley and wheat crops, given the similarity of their NDVI time series signature. Thus, on the left side of **Fig. 6**, the barley (red) and wheat (brown) plots can be seen, while on the right side of the figure most of these plots have been clustered as barley plots. While the remaining crops, specially permanent meadows and corn, are properly clustered.

As a summary of the preliminary analysis, clustering algorithms show more difficulty grouping crops with a similar NDVI time series (growth cycle) into different clusters. However, clustering algorithms show good results when the NDVI time series of crops are different.

4.4.2. Determining the best clustering algorithms for all crop combinations

The objective of this subsection is to determine the best combination of clustering algorithm and distance measures for any crop combination. Therefore, this study includes the 26 datasets obtained from all the possible combinations of the crops.

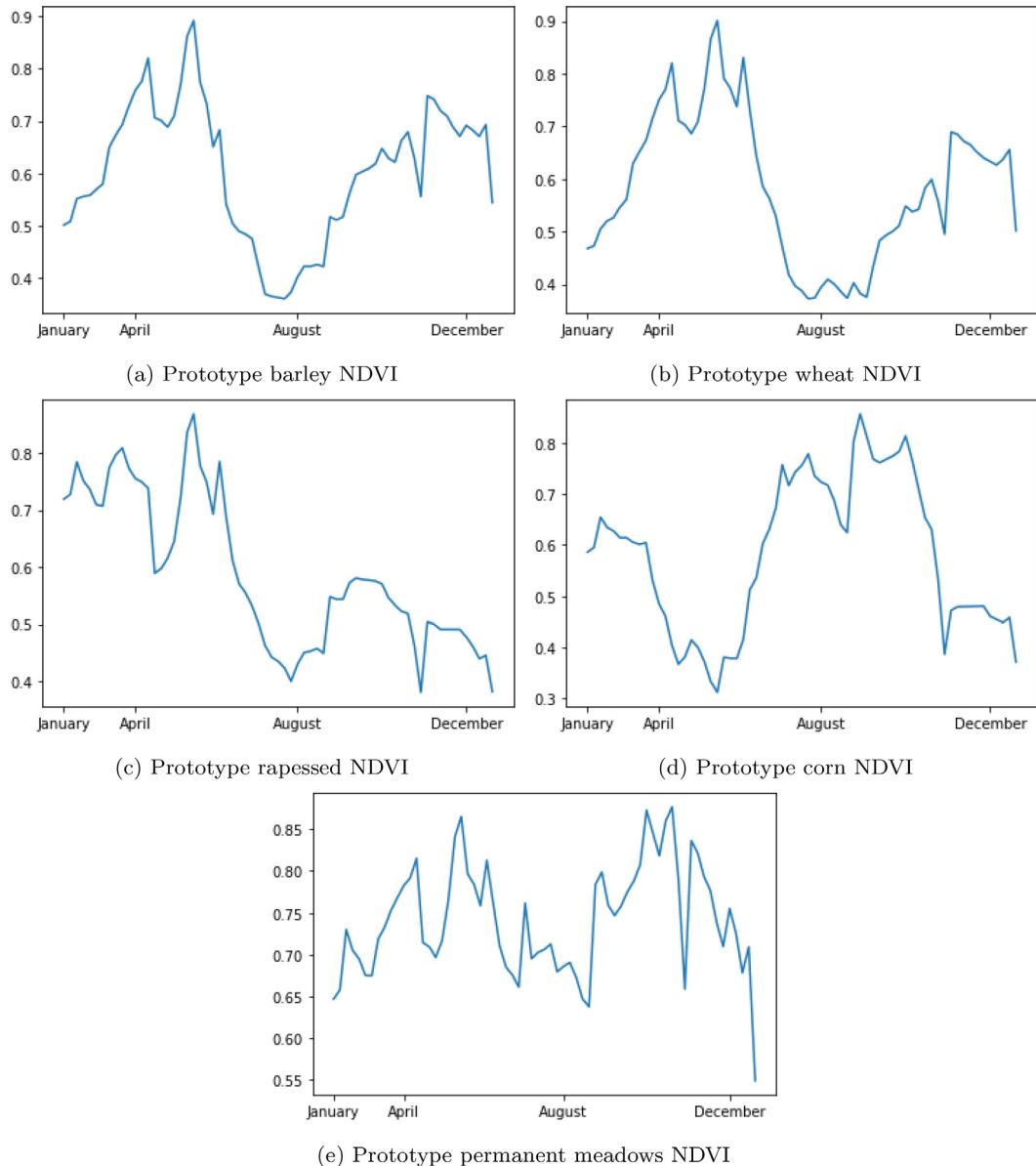


Fig. 3. Ideal prototypes of crops.

Table 2
Results with barley and wheat crops.

Algorithm_distance	AMI	Jaccard
Agglomerative_Canberra	0.02177	0.37917
Agglomerative_Euclidean	0.02148	0.49954
Agglomerative_Manhattan	0.00397	0.49869
Agglomerative_Minkowski	0.00148	0.53648
DIANA_Canberra	0.03233	0.36996
DIANA_Euclidean	0.03897	0.37867
DIANA_Manhattan	0.04356	0.38285
DIANA_Minkowski	0.04209	0.38164
PAM_Canberra	0.01452	0.36889
PAM_Euclidean	0.01766	0.37573
PAM_Manhattan	0.02324	0.36264
PAM_Minkowski	0.02437	0.36361
Spectral_Canberra	0.00054	0.54049
Spectral_Euclidean	0.00031	0.54049
Spectral_Manhattan	0.00124	0.54049
Spectral_Minkowski	0.00124	0.54049

To provide more meaningfully results, statistical techniques were used over a simple error means analysis.

First, in Table 4 and (Fig. 7) the average ranking computed for all approaches according to the AMI measure is shown. Following which, the Friedman test is applied and a *p*-value of 9.5E-11 is obtained, which implies that there are significant differences among the results of the different clustering methods. Thus, post hoc tests (Nemenyi, Holm and Shaffer) were applied to detect significant differences between the methods. As mentioned the post hoc tables are show in [71].

As can be seen, significant differences between the methods that rank higher and those that rank lowest can be found. These results advise to use the Spectral method along with the distance measures, Euclidean, Manhattan or Minkowski when the crop mapping task is addressed with clustering techniques.

The same calculations are repeated for the Jaccard measure. In this case, the Friedman test obtained a *p*-value of 5.22E-11, indicating the existence of significant differences. A ranking of these methods is shown in Table 5 (Fig. 7). The *p*-values obtained for

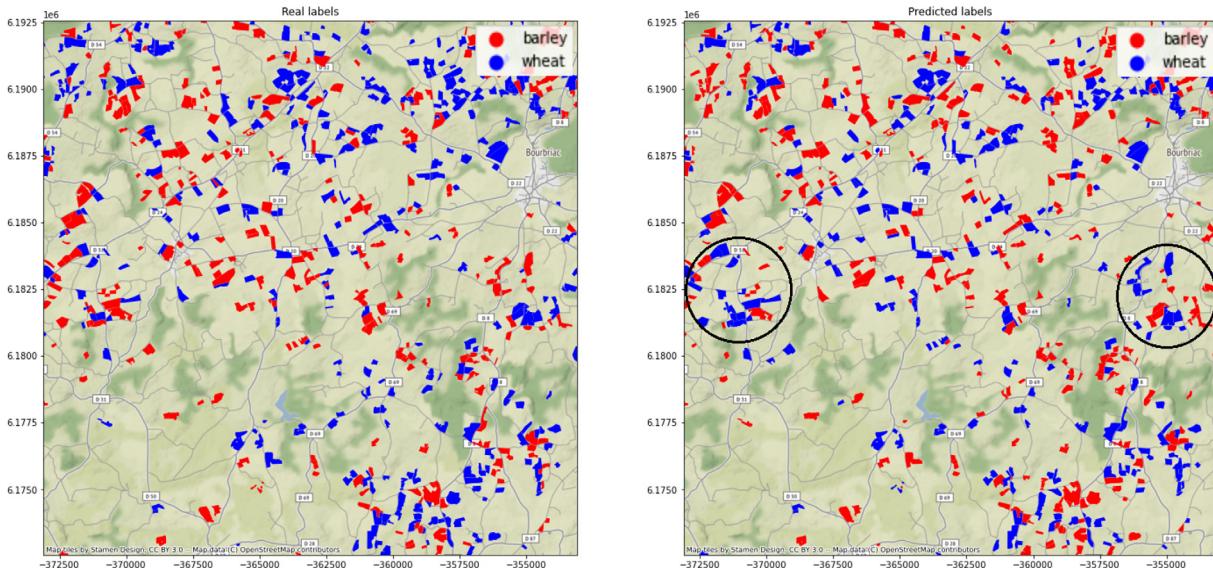


Fig. 4. Map of real crop types (left) vs. obtained by the clustering algorithm (right) for barley and wheat crops.

Table 3
Results with wheat and corn crops.

Algorithm_distance	AMI	Jaccard
Agglomerative_Canberra	0.77011	0.85990
Agglomerative_Euclidean	0.90318	0.95438
Agglomerative_Manhattan	0.89668	0.94756
Agglomerative_Minkowski	0.91599	0.96149
DIANA_Canberra	0.90793	0.95700
DIANA_Euclidean	0.90532	0.95625
DIANA_Manhattan	0.91759	0.96291
DIANA_Minkowski	0.91269	0.96027
PAM_Canberra	0.88775	0.94712
PAM_Euclidean	0.83953	0.90804
PAM_Manhattan	0.89775	0.95243
PAM_Minkowski	0.89181	0.94907
Spectral_Canberra	0.92938	0.96983
Spectral_Euclidean	0.92669	0.96841
Spectral_Manhattan	0.92816	0.96920
Spectral_Minkowski	0.92782	0.96904

the post hoc tests methods show that again, an important number of significant differences can be found between the first methods of the ranking and the last ones, therefore, the use of the Spectral method with distances measures such as Manhattan, Minkowski or Euclidean is advised to deal with crop mapping task addressed in an unsupervised way.

This study shows that when considering any combination of crops, the best clustering algorithm is the Spectral one when combined with either the Manhattan, Euclidean and Minkowski distances. The second best results are obtained by using the DIANA clustering method using the Manhattan. These results are supported by the statistical tests applied, where significant differences can be found between these two clustering methods and the remaining ones. The worth results are obtained with the PAM method using the Euclidean distance.

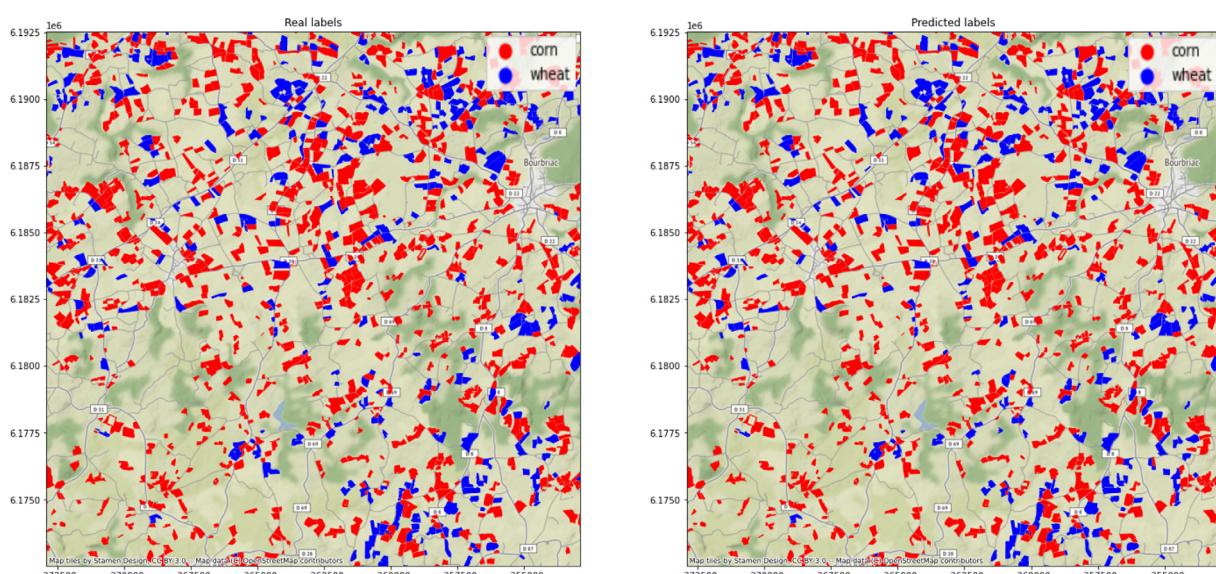


Fig. 5. Map of real crop types (left) vs. obtained by the clustering algorithm (right) for wheat and corn crops.

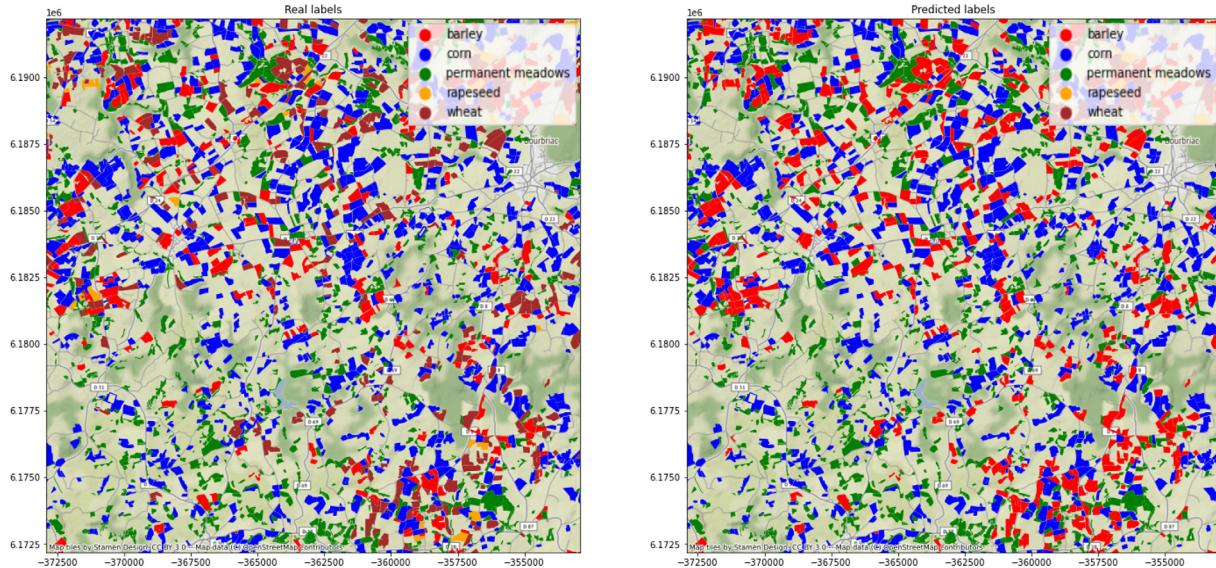


Fig. 6. Map of real crop types (left) vs. obtained by the clustering algorithm (right) for all crops.

Table 4

Friedman ranking (AMI measure) of clustering methods for all datasets.

Algorithm	Ranking
Spectral_Euclidean	3.5769
Spectral_Manhattan	4.2692
Spectral_Minkowski	4.3269
DIANA_Manhattan	5.3846
Agglomerative_Manhattan	6.7692
DIANA_Minkowski	6.9615
Agglomerative_Minkowski	7.1923
PAM_Manhattan	8.6154
DIANA_Euclidean	9
Agglomerative_Euclidean	10
PAM_Canberra	10.8846
DIANA_Canberra	10.9231
Spectral_Canberra	10.9808
PAM_Minkowski	11.5
Agglomerative_Canberra	11.8846
PAM_Euclidean	13.7308

4.4.3. Determining the best clustering algorithms taking into account the number of crops in the dataset

In this section the objective is to elucidate the best clustering methods taking into account the number of crops addressed. Thus, three groups are built: datasets with 2 crops, datasets with 3 crops and datasets with 4 crops. Therefore, better clustering methods are determined when the crop mapping task is addressed for locations with 2 crops, 3 crops or 4 crops. The same method as above was followed where, ranking of the methods by performance is obtained and, then, existence of significant differences were determined by means of the appropriate statistical tools.

Beginning with the datasets of 2 crops, Tables 6 and 7 show the average ranking computed for all approaches according to the AMI and Jaccard measures. In both cases, the best option is the Spectral algorithm, occupying the first three positions. However, the specific position depends on the distance measure chosen. Euclidean, Manhattan and Minkowski, in this order, is the best option for the AMI measure and Manhattan, Minkowski and Euclidean is the ranking for the Jaccard measure. For the AMI measure, DIANA_Manhattan and DIANA_Minkowski are the following methods and for the Jaccard measure, Agglomerative_Minkowski and Spectral_Canberra obtains the third and fourth positions.

Then, the Friedman test is applied and low *p*-values (2.12E-8 and 3.06E-7) are respectively obtained for the AMI and Jaccard measure. Therefore, post hoc tests are applied and the significant difference results are obtained. From these data, significant differences can be deduced between the Spectral method with the Euclidean, Manhattan, and Minkowski distances and the methods that occupy the last positions of the ranking tables.

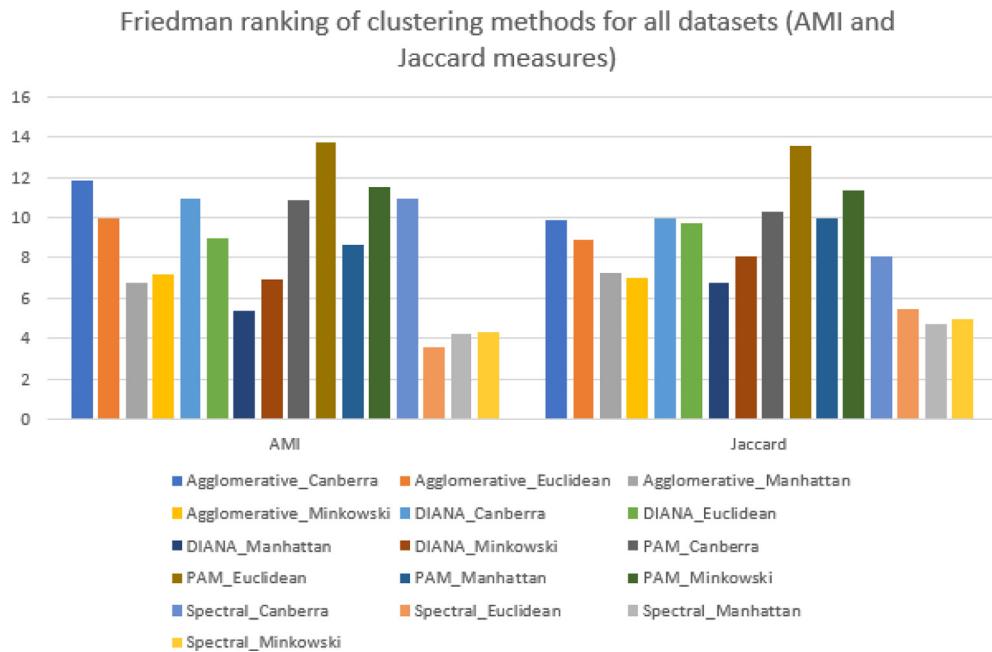
For the combinations/datasets with 3 crop types, Tables 8 and 9 show an average ranking of the clustering methods for the AMI and Jaccard measures respectively. As with the experiment with two crops, the Spectral algorithm heads the table with the Minkowski, Manhattan and Euclidean measures respectively. DIANA_Manhattan is the clustering method next in performance for the two measures.

p-values of 1.52E-7 and 1.83E-6 are obtained in the the Friedman test and applied for AMI and Jaccard measures respectively. In the same manner, post hoc tests obtain *p*-values lower than 0.1, that implies significant differences between the mentioned first ranking methods and the last ones.

Rankings based on models developed using datasets with 4 crop types are presented in Tables 10 and 11 respectively. Both AMI and Jaccard measures were used. In both cases, the Spectral methods produced the best results when combined with the Euclidean, Minkowski or Manhattan distances.

Following the Friedman test, *p*-values of 1.12E-5 and 0.24 were obtained for the AMI and the Jaccard measures respectively. This implies that there are significant differences for the AMI measure but not for the Jaccard distance. Thus, only post hoc tests were applied to the AMI measures. *p*-values lower than 0.1 for these tests imply the existence of significant differences amongst the two methods. They occupy the first and last positions respectively in the ranking table.

The main objective of this research was to identify the best performing clustering algorithms taking into account the number of crops in the dataset. In general, the best generalisation results were obtained when using the Spectral clustering algorithm with different distance measures. After performing a basic ranking of the distances involving two crops, it becomes clear that the best results were obtained when using, in order of best performance, the Manhattan, Euclidean and Minkowsky distances. When expanding the number of crops from two to three, the best results were obtained with the Minkowsky distance followed by the Manhattan and the

**Fig. 7.** Friedman ranking of clustering methods for all datasets (AMI and Jaccard measures).**Table 5**

Friedman ranking (Jaccard measure) of clustering methods for all datasets.

Algorithm	Ranking
Spectral_Manhattan	4.6923
Spectral_Minkowski	4.9808
Spectral_Euclidean	5.4231
DIANA_Manhattan	6.8077
Agglomerative_Minkowski	7.0385
Agglomerative_Manhattan	7.2692
Spectral_Canberra	8.0577
DIANA_Minkowski	8.0769
Agglomerative_Euclidean	8.8846
DIANA_Euclidean	9.7308
Agglomerative_Canberra	9.8462
DIANA_Canberra	9.9615
PAM_Manhattan	9.9615
PAM_Canberra	10.3077
PAM_Minkowski	11.3846
PAM_Euclidean	13.5769

Table 7

Friedman ranking (Jaccard measure) of clustering methods for all two-crop combination datasets

Algorithm	Ranking
Spectral_Manhattan	3.7
Spectral_Minkowski	4.2
Spectral_Euclidean	5.05
Agglomerative_Minkowski	5.9
Spectral_Canberra	6.15
DIANA_Manhattan	6.7
Agglomerative_Manhattan	8
DIANA_Canberra	8.4
DIANA_Minkowski	8.4
DIANA_Euclidean	10.1
Agglomerative_Euclidean	10.4
PAM_Manhattan	10.4
PAM_Canberra	11.2
PAM_Minkowski	11.8
Agglomerative_Canberra	12.2
PAM_Euclidean	13.4

Table 6

Friedman ranking (AMI measure) of clustering methods for all two-crop combination datasets.

Algorithm	Ranking
Spectral_Euclidean	3.25
Spectral_Manhattan	4.4
Spectral_Minkowski	4.8
DIANA_Manhattan	4.9
DIANA_Minkowski	6.5
Agglomerative_Minkowski	7.9
Agglomerative_Manhattan	8.2
DIANA_Canberra	8.5
PAM_Manhattan	8.5
DIANA_Euclidean	8.8
Spectral_Canberra	10.05
PAM_Canberra	10.3
PAM_Minkowski	11.2
Agglomerative_Euclidean	11.7
PAM_Euclidean	12.7
Agglomerative_Canberra	14.3

Table 8

Friedman ranking (AMI measure) of clustering methods for all three-crop combination datasets.

Algorithm	Ranking
Spectral_Minkowski	4.4
Spectral_Manhattan	4.8
Spectral_Euclidean	4.9
DIANA_Manhattan	5.5
Agglomerative_Manhattan	6.1
DIANA_Minkowski	6.7
Agglomerative_Minkowski	6.7
DIANA_Euclidean	8.3
PAM_Manhattan	8.8
Agglomerative_Euclidean	9
Spectral_Canberra	10.1
Agglomerative_Canberra	11.1
PAM_Canberra	11.5
PAM_Minkowski	11.9
DIANA_Canberra	12.2
PAM_Euclidean	14

Table 9

Friedman ranking (Jaccard measure) of clustering methods for all three-crop combination datasets.

Algorithm	Ranking
Spectral_Minkowski	4.3
Spectral_Manhattan	5
Spectral_Euclidean	6.2
DIANA_Manhattan	6.2
Spectral_Canberra	6.4
Agglomerative_Manhattan	6.9
Agglomerative_Euclidean	7.4
DIANA_Minkowski	7.6
Agglomerative_Minkowski	7.8
Agglomerative_Canberra	8.6
DIANA_Euclidean	9.7
PAM_Canberra	10.4
PAM_Manhattan	10.8
DIANA_Canberra	11.2
PAM_Minkowski	12.8
PAM_Euclidean	14.7

Table 10

Friedman ranking (AMI measure) of clustering methods for all four-crop combination datasets

Algorithm	Ranking
Spectral_Euclidean	2.1
Spectral_Minkowski	3.1
Spectral_Manhattan	3.4
Agglomerative_Manhattan	5.4
DIANA_Manhattan	5.6
Agglomerative_Minkowski	7.4
Agglomerative_Canberra	7.8
DIANA_Minkowski	8
PAM_Manhattan	8.8
Agglomerative_Euclidean	10
DIANA_Euclidean	10.4
PAM_Canberra	11
PAM_Minkowski	11
DIANA_Canberra	13
Spectral_Canberra	14
PAM_Euclidean	15

Table 11

Friedman ranking (Jaccard measure) of clustering methods for all four-crop combination datasets.

Algorithm	Ranking
Spectral_Manhattan	5
Spectral_Euclidean	5.5
Spectral_Minkowski	6.5
DIANA_Manhattan	6.8
Agglomerative_Manhattan	6.8
PAM_Manhattan	7.2
DIANA_Minkowski	8.2
DIANA_Euclidean	8.4
Agglomerative_Minkowski	8.4
PAM_Minkowski	8.4
Agglomerative_Canberra	9
PAM_Canberra	9
Agglomerative_Euclidean	9.6
PAM_Euclidean	11.2
DIANA_Canberra	12.2
Spectral_Canberra	13.8

Euclidean distances. Finally, when the number of crops considered were extended to four, the best results were obtained using the Euclidean distance with both the Manhattan and the Minkowsky distances providing the second and third best results respectively.

In addition to this, it can be observed that the level of performance of the DIANA clustering method, when combined with the Manhattan distance, comes second from the Spectral method for all crop combinations. These methods often produce significant statistical differences with respect to the worst methods in the ranking tables. This is specially the case for the PAM method when using the Euclidean distance.

4.4.4. Determining the best distance measure for each clustering algorithm

The best suited distance measure for each of the clustering methods will now be analysed. Thus, a study will be carried out for each clustering algorithm including the 26 datasets (all crop combinations), where the performance of each of the distance measures will be examined.

Beginning with the Agglomerative clustering method in [Tables 12 and 13](#), a ranking by the AMI and Jaccard measures is shown. In both rankings the worst performance is obtained by using the Euclidean and Canberra distances. The best performers are the Manhattan and the Minkowski. This order depends on the quality measure considered i.e AMI or Jaccard measure.

The Friedman Test returns a value of 0.000271 for the AMI measure and 0.241 for the Jaccard measure. This implies that only the the AMI measure shows significant differences. According to post hoc tests based on *p*-values, significant differences between the Manhattan and the Canberra and Euclidean distances can be found. Similar results can be seen between the Minkowski and the Canberra distances.

Taking into account the DIANA clustering methods, [Tables 14 and 15](#) show the rankings obtained based on the distance measure used for the AMI an Jaccard measure respectively. In both cases the order of this ranking is Manhattan and Minkowski followed by Euclidean or Canberra depending on the measure used.

Computing the Friedman test for these two measures, a *p*-value of 5.21E-11 is obtained for the AMI measure and 1.84E-6 for Jaccard, which involves significant differences in both cases. As can

Table 12

Friedman ranking (AMI measure) of distance measures for Agglomerative method.

Algorithm	Ranking
Agglomerative_Manhattan	1.8846
Agglomerative_Minkowski	2.0385
Agglomerative_Euclidean	2.8846
Agglomerative_Canberra	3.1923

Table 13

Friedman ranking (Jaccard measure) of distance measures for Agglomerative method.

Algorithm	Ranking
Agglomerative_Minkowski	2.1538
Agglomerative_Manhattan	2.3462
Agglomerative_Euclidean	2.7308
Agglomerative_Canberra	2.7692

Table 14

Friedman ranking (AMI measure) of distance measures for DIANA method.

Algorithm	Ranking
DIANA_Manhattan	1.1538
DIANA_Minkowski	2.1154
DIANA_Euclidean	3.1538
DIANA_Canberra	3.5769

be seen from the post hoc tests, for both measures, significant differences between the best distances (Manhattan and Minkowski) and the worst ones (Euclidean and Canberra) can be found. Also, there exist significant differences between Manhattan and Minkowski for AMI measure.

Regarding the PAM clustering method, the average ranking results of its distances for measures AMI and Jaccard are shown in [Tables 16 and 17](#). In both cases the order is Manhattan, Canberra, Minkowski, followed by Euclidean. An can be observed, PAM is the only clustering method that Canberra distance is not in the last position in a ranking achieving a second position.

When the Friedman test is applied, a *p*-value of 3.12E-10 is obtained for the AMI measure and a *p*-value 1.06E-6 is obtained for the Jaccard measure. These values report the existence of significant differences for both measures. The post hoc tests indicate that

Table 15
Friedman ranking (Jaccard measure) of distance measures for DIANA method.

Algorithm	Ranking
DIANA_Manhattan	1.5
DIANA_Minkowski	2.2308
DIANA_Canberra	3.0769
DIANA_Euclidean	3.1923

Table 16
Friedman ranking (AMI measure) of distance measures for PAM method.

Algorithm	Ranking
PAM_Manhattan	1.2308
PAM_Canberra	2.5
PAM_Minkowski	2.5769
PAM_Euclidean	3.6923

Table 17
Friedman ranking (Jaccard measure) of distance measures for PAM method.

Algorithm	Ranking
PAM_Manhattan	1.7692
PAM_Canberra	2.0769
PAM_Minkowski	2.5385
PAM_Euclidean	3.6154

Table 18
Friedman ranking (AMI measure) of distance measures for Spectral method.

Algorithm	Ranking
Spectral_Euclidean	1.8462
Spectral_Minkowski	2.1731
Spectral_Manhattan	2.4231
Spectral_Canberra	3.5577

Table 19
Friedman ranking (Jaccard measure) of distance measures for Spectral method.

Algorithm	Ranking
Spectral_Euclidean	2.3077
Spectral_Manhattan	2.4615
Spectral_Minkowski	2.5192
Spectral_Canberra	2.7115

the distance which mainly establish significant differences is Manhattan, since it establishes these differences with the remaining ones for the AMI measure and with Euclidean and Minkowski for the Jaccard measure.

Finally, the Spectral clustering method is analysed. Thus, the average ranking of their distances are shown in [Tables 18 and 19](#) for the AMI measure and for the Jaccard measure respectively. For both cases, the best results were obtained using the Euclidean distance. The worst results were obtained using the Canberra distance. However, for the AMI distance the second and third positions were obtained using the Minkowski and the Manhattan distances. The opposite is true for the Jaccard measure where the second and third positions were obtained by using the Manhattan and the Minkowski measurements respectively.

For the Friedman test, *p*-values of 1.01E-5 and 0.728 were obtained for the AMI and the Jaccard measures respectively. This implies the existence of significant differences within the AMI measure but not within the Jaccard one. Significant differences between any distance and the Canberra distance are established according to *p*-values obtained by post hoc tests.

The studies carried out in this section can be summarised as:

- For the Agglomerative clustering method, the distance measures that produce best results are the Manhattan and the Minkowski. For these distances even, statistically significant differences can be found with respect to the Euclidean and the Canberra distances.
- In the case of the DIANA clustering method, the Manhattan and the Minkowsky are also the best distance measures. Statistically significant differences can be found with respect to the Euclidean and the Canberra distances.
- The best distance measures for the PAM clustering method are the Manhattan and the Canberra distances. The Euclidean distance yields the worst results. There is statistically significant differences with the remaining distances.
- The Euclidean distance produces the best results for the Spectral clustering algorithm. However, there are no significant differences with respect to the Manhattan and the Minkowski distances. The worst results are obtained when using the Canberra distance with statistically significant differences.

4.5. Guidelines on using clustering methods for crop mapping tasks

Based on this study (with the datasets, clustering methods and distances measures considered), some guidelines on the use of clustering methods for the purpose of crop mapping are presented below.

When the number of crops in the study area is undetermined the best results are obtained when using the Spectral clustering method with either the Euclidean, Manhattan or Minkowski distances. DIANA clustering with Manhattan distance is the next best alternative to the Spectral clustering method. The worst results are obtained when using the PAM method. This is particularly the case when using the Euclidean distance.

When the number of crops is known for a land surface the best performance is obtained with the Spectral clustering method combined with the Euclidean, Manhattan or Minkowski distance methods. Statistical analysis suggests that this applies to any number of crops. The combinations of Agglomerative_Canberra or PAM_Euclidean methods is discouraged for land surfaces with 2 crops. The PAM_Euclidean, DIANA_Canberra or PAM_Minkowski methods are not recommended for land surfaces with 3 crops. PAM_Euclidean, Spectral_Canberra and DIANA_Canberra are also not recommended for land cover with 4 crops.

The choice of distance methods plays a critical role on the performance of the clustering algorithms. For the Agglomerative

method, the best performing distances are the Manhattan and the Minkowski. The worst is Canberra. For the DIANA method the best working distance is the Manhattan and the worst are the Canberra and Euclidean Distance. For the PAM method, the distance that work best is the Manhattan and the worst is the Euclidean. Finally, for the Spectral method and according to the rankings, the Euclidean distance is the best performing distance. The worth results are systematically obtained when using the Canberra distance measure.

In summary, the best clustering method for crop type mapping is the Spectral one when used in combination with the Euclidean or the Manhattan distances. The Euclidean distance, when combined with the Spectral method, provides the largest number of best results but with low differences. However, when Euclidean distance under performs, it does so with high differences. This fact causes some oscillations in the statistical tests. In conclusion, the best general option is to choose the Spectral clustering method with the Manhattan distance which shows a more robust and consistent behaviour.

5. Conclusions

With a growing population and reduced food production, food security is one of the most critical challenges for humanity. Climate change, the degradation of the soil and the quality of water are one of the primary causes of this food scarcity. Smart agriculture is a key area that has demonstrated positive results in improving food production. It has also shown a reduction on the use of resources such as agro-chemicals. This has brought about significant environmental benefits. Smart agricultural techniques are empowered through data obtained from both ground and remote sensors. Within the field of remote sensing, satellite imagery is one of the most powerful tools which can provide data for the machine learning methods. These methods allow for the analyses at a much larger scale covering larger land areas.

Unsupervised machine learning methods and more specifically clustering methods, have been successfully used for agricultural tasks involving satellite image data. These methods are on increased demand within a wide range of applications when crop labels are not available, which is often the case.

In this paper, a number of clustering methods were tested and compared in the domain of crop type mapping. Extensive experiments were done with different combinations of clustering algorithms and distance measures on a large number of datasets containing different crop types. The results obtained with the different algorithms have been analysed from a predictive point of view, with appropriate external validation clustering measures.

Specifically, a total of 4 clustering algorithm with 4 distance measures were compared using 26 datasets containing 5 crops across more than 45000 parcels. After the experimentation some hypothesis testing techniques were explored and the results were summarised. The behaviour of the clustering methods was explored in terms of the similarity of the growth cycle of the crops. Based on this, it can be deduced that the different clustering methods perform very well when the growth cycle of crops is different. However, the level of performance is lower when the cycles are very similar.

Regarding land coverage with any number of crops, the Spectral clustering algorithm, in combination with the Manhattan or the Euclidean distances, are best suited to address the crop mapping task. In contrast, the results obtained using the PAM method together with the Euclidean distance are not very encouraging.

A study of land covers with a fixed number of crops was also done. For 2, 3 and 4 crops, the best performing method was found to be the Spectral clustering method using either the Euclidean or

the Manhattan distances. The best distance measures for each clustering algorithm was also analysed. The best choice of the distance method depends on the clustering algorithm used: Manhattan for the DIANA and the PAM, Euclidean and Manhattan for the Spectral and Manhattan and Minkowski for the Agglomerative clustering methods. The behaviour of the Camberra distance measure produced the worst results for all clustering methods.

After carrying on all the experiments, it is clear that the Spectral clustering method combined with the Manhattan distance shows the most robust and consistent behaviour.

CRediT authorship contribution statement

Antonio J. Rivera: Methodology, Investigation, Formal analysis, Data Curation, Software, Writing – review & editing. **María D. Pérez-Godoy:** Methodology, Investigation, Formal analysis, Data Curation, Software, Writing – review & editing. **David Elizondo:** Conceptualization, Methodology, Writing – review & editing. **Lipika Deka:** Methodology, Investigation, Writing – review & editing. **María J. del Jesus:** Conceptualization, Methodology, Writing – review & editing.

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 partially supported by the Spanish Ministry of Science and Technology under project TIN2015-68454-R and by the Spanish Ministry of Science and Innovation under project PID2019-107793 GB-I00 funded by AEI/ 10.13039/501100011033.

References

- [1] End hunger, achieve food security and improved nutrition and promote sustainable agriculture, <https://sdgs.un.org/goals/goal2>, accessed: 2021-03-03 (2021).
- [2] R.P. Sishodia, R.L. Ray, S.K. Singh, Applications of remote sensing in precision agriculture: a review, *Remote Sens.* 12 (3136) (2020), <https://doi.org/10.3390/rs12193136>.
- [3] M. Weiss, F. Jacob, G. Duveiller, Remote sensing for agricultural applications: a meta-review, *Remote Sensing of Environment* 236 (111402) (2020). doi:10.1016/j.rse.2019.111402.
- [4] A. Bannari, D. Morin, F. Bonn, A. Huete, A review of vegetation indices, *Remote Sens. Rev.* 13 (1–2) (1995) 95–120, <https://doi.org/10.1080/02757259509532298>.
- [5] C.J. Tucker, J.E. Pinzon, M.E. Brown, D.A. Slayback, E.W. Pak, R. Mahoney, E.F. Vermote, N.E. Saleous, An extended avhr 8-km ndvi dataset compatible with modis and spot vegetation ndvi data, *Int. J. Remote Sens.* 26 (20) (2005) 4485–4498, <https://doi.org/10.1080/014311605001686868>.
- [6] J.M. Zhang, M. Harman, L. Ma, Y. Liu, Machine learning testing: survey, landscapes and horizons, *IEEE Transactions on Software Engineering* (2020) 1–1. doi:10.1109/TSE.2019.2962027.
- [7] S. Wang, G. Azzari, D.B. Lobell, Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques, *Remote Sens. Environ.* 222 (2019) 303–317, <https://doi.org/10.1016/j.rse.2018.12.026>.
- [8] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O.P. Patel, A. Tiwari, M.J. Er, W. Ding, C.-T. Lin, A review of clustering techniques and developments, *Neurocomputing* 267 (2017) 664–681, <https://doi.org/10.1016/j.neucom.2017.06.053>.
- [9] A. Rivera, M. Pérez-Godoy, D. Elizondo, L. Deka, M. del Jesus, A preliminary study on crop classification with unsupervised algorithms for time series on images with olive trees and cereal crops, in: 15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020), Springer International Publishing, Cham, 2021, pp. 276–285.
- [10] E.O. System, Satellite data: what spatial resolution is enough?, <https://eos.com/blog/satellite-data-what-spatial-resolution-is-enough-for-you/>, accessed: 18-02-2021 (2019).
- [11] S. Milrad, Synoptic Analysis and Forecasting, *Science Direct*, 2018, Ch. 11, pp. 145–161.

- [12] E.S.A. (ESA), High resolution land cover project (H.R.L.C), <https://climate.esa.int/en/projects/high-resolution-land-cover/about/>, accessed: 02-03-2022.
- [13] S. Khanal, J. Fulton, S. Shearer, An overview of current and potential applications of thermal remote sensing, *Computers and Electronics in Agriculture* 139 (2017) 22–32, <https://doi.org/10.1016/j.compag.2017.05.001>.
- [14] W. Verstraeten, F. Veroustraete, J. Feyen, Assessment of evapotranspiration and soil moisture content across different scales of observation, *Sensors* 8 (1) (2008), <https://doi.org/10.3390/s8010070>.
- [15] J.B. Barker, C.M. Neale, D.M. Heeren, A.E. Suyker, Evaluation of a hybrid reflectance-based crop coefficient and energy balance evapotranspiration model for irrigation management, *Transactions of the ASABE* 61 (2018), <https://doi.org/10.13031/trans.12311>.
- [16] L. Zho, N. Chen, Z. Chen, C. Xing, Roscc: an efficient remote sensing observation-sharing method based on cloud computing for soil moisture mapping in precision agriculture, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (12) (2016) 5588–5598, <https://doi.org/10.1109/JSTARS.2016.2574810>.
- [17] D. Zhang, G. Zhou, Estimation of soil moisture from optical and thermal remote sensing: a review, *Sensors* 16 (8) (2016), <https://doi.org/10.3390/s16081308>.
- [18] M.M. Ali, A. Al-Ani, D. Eamus, D.K.Y. Tan, Leaf nitrogen determination using non-destructive techniques—a review, *J. Plant Nutr.* 40 (7) (2017), <https://doi.org/10.1080/01904167.2016.1143954>.
- [19] D. Franzen, N. Kitchen, K. Holland, J. Schepers, W. Raun, Algorithms for in-season nutrient management in cereals, *Agronomy Journal* 108 (2016) 1775–1781, <https://doi.org/10.2134/agronj2016.01.0041>.
- [20] P.C. Scharf, D.K. Shannon, H.L. Palm, K.A. Sudduth, S.T. Drummond, N.R. Kitchen, L.J. Mueller, V.C. Hubbard, L.F. Oliveira, Sensor-based nitrogen applications out-performed producer-chosen rates for corn in on-farm demonstrations, *Agronomy Journal* (2011) 1684–1691, <https://doi.org/10.2134/agronj2011.0164>.
- [21] A. Mahlein, Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping, *Plant Dis.* 100 (2016) 241–251, <https://doi.org/10.1094/PDIS-03-15-0340-PE>.
- [22] H. Huang, Y. Lan, A. Yang, Y. Zhang, S. Wen, J. Deng, Deep learning versus object-based image analysis(obia) in weed mapping of uav imagery, *Int. J. Remote Sens.* 41 (2020) 3446–3479, <https://doi.org/10.1080/01431161.2019.1706112>.
- [23] V. Partel, S.C. Kakarla, Y. Ampatzidis, Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence, *Computers and Electronics in Agriculture* 157 (2019) 339–350, <https://doi.org/10.1016/j.compag.2018.12.048>.
- [24] J. Yeom, J. Jung, A. Chang, A. Ashapure, M. Maeda, A. Maeda, J. Landivar, Comparison of vegetation indices derived from uav data for differentiation of tillage effects in agriculture, *Remote Sens.* 11 (13) (2019), <https://doi.org/10.3390/rs11131548>.
- [25] E.D.P. Prasad, S. Thenkabail, Ronald B. Smith, Hyperspectral vegetation indices and their relationships with agricultural crop characteristics, *Remote Sens. Environ.* 71 (2) (2000) 158–182, [https://doi.org/10.1016/S0034-4257\(99\)00067-X](https://doi.org/10.1016/S0034-4257(99)00067-X).
- [26] M. Schultz, J.G. Clevers, S. Carter, J. Verbesselt, V. Avitabile, H.V. Quang, M. Herold, Performance of vegetation indices from landsat time series in deforestation monitoring, *Int. J. Appl. Earth Obs. Geoinf.* 52 (2016) 318–327, <https://doi.org/10.1016/j.jag.2016.06.020>.
- [27] D.R. Panjuji, B.H. Trisasongko, Tropical Forests, Intechopen, 2012, Ch. Seasonal Pattern of Vegetative Cover from NDVI Time-Series. doi:10.5772/30344.
- [28] W. Kalisa, T. Igbarawa, M.H. et.al., Assessment of climate impact on vegetation dynamics over east africa from 1982 to 2015, *Scientific Reports* 9 (16865) (2019), doi:10.1038/s41598-019-53150-0.
- [29] C. Conrad, R.R. Colditz, S. Dech, D. Klein, P.L.G. Vlek, Temporal segmentation of modis time series for improving crop classification in central asian irrigation systems, *Int. J. Remote Sens.* 32 (23) (2011) 8763–8778, <https://doi.org/10.1080/01431161.2010.550647>.
- [30] J. Brown, J. Kastens, A. Coutinho, D. Victoria, C. Bishop, Classifying multiyear agricultural land use data from mato grosso using time-series modis vegetation index data, *Remote Sens. Environ.* 130 (2013) 39–50, <https://doi.org/10.1016/j.rse.2012.11.009>.
- [31] T. Murakami, S. Ogawa, N. Ishitsuka, K. Kumagai, G. Saito, Crop discrimination with multitemporal spot/hrv data in the saga plains, *Japan International Journal of Remote Sensing* 22 (7) (2001) 1335–1348, <https://doi.org/10.1080/01431160151144378>.
- [32] F. Giacco, C. Thiel, L. Pugliese, S. Scarpetta, M. Marinaro, Uncertainty analysis for the classification of multispectral satellite images using svms and soms, *IEEE Trans. Geosci. Remote Sens.* 48 (10) (2010) 3769–3779, <https://doi.org/10.1109/TGRS.2010.2047863>.
- [33] F. Löwa, U. Michelb, S. Dechac, C. Conrada, Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines, *ISPRS Journal of Photogrammetry and Remote Sensing* 85 (2013) 102–119, <https://doi.org/10.1016/j.isprsjprs.2013.08.007>.
- [34] E. Blanzieri, F. Melgani, Nearest neighbor classification of remote sensing images with the maximal margin principle, *IEEE Trans. Geosci. Remote Sens.* 46 (6) (2008) 1804–1811, <https://doi.org/10.1109/TGRS.2008.916090>.
- [35] C. Murthy, P. Raju, K.V.S. Badrinath, Classification of wheat crop with multi-temporal images: performance of maximum likelihood and artificial neural networks, *Int. J. Remote Sens.* 24 (23) (2003) 4871–4890, <https://doi.org/10.1080/0143116031000070490>.
- [36] P. Kumar, D.K. Gupta, V.N. Mishra, R. Prasad, Comparison of support vector machine, artificial neural network, and spectral angle mapper algorithms for crop classification using liss iv data, *Int. J. Remote Sens.* 36 (6) (2015) 1604–1617, <https://doi.org/10.1080/2150704X.2015.1019015>.
- [37] P. Kumar, R. Prasad, V.N. Mishra, D.K. Gupta, S.K. Singh, Artificial neural network for crop classification using c-band risat-1 satellite datasets, *Russian Agricultural Sciences* 42 (2016) 281–284.
- [38] C. Yang, S. Prasher, J.L. DiTommaso, Application of artificial neural networks in image recognition and classification of crop and weeds, *Canadian Agricultural Engineering* (2000).
- [39] H. Wang, J. Zhang, K. Xiang, Y. Liu, Classification of remote sensing agricultural image by using artificial neural network, in: 2009 International Workshop on Intelligent Systems and Applications, Wuhan, China, 23–24 May 2009, 2009, pp. 1–4.
- [40] S. Murmu, S. Biswas, Application of fuzzy logic and neural network in crop classification: a review, *Aquatic Procedia* 4 (2015) 1203–1210, <https://doi.org/10.1016/j.aqpro.2015.02.153>.
- [41] C. Yang, S.O. Prasher, P. Enright, C. Madramootoo, M. Burgess, P.K. Goel, I. Callum, Application of decision tree technology for image classification using remote sensing data, *Agric. Syst.* 76 (3) (2003) 1101–1117, [https://doi.org/10.1016/S0308-521X\(02\)00051-3](https://doi.org/10.1016/S0308-521X(02)00051-3).
- [42] J.M. Peña, P.A. Gutiérrez, C. Hervás-Martínez, J. Six, R.E. Plant, F. López-Granados, Object-based image classification of summer crops with machine learning methods, *Remote Sens.* 6 (6) (2014) 5019–5041.
- [43] B. Deschamps, H. McNairn, J. Shang, X. Jiao, Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier, *Canadian Journal of Remote Sensing* 38 (1) (2014), <https://doi.org/10.5589/m12-012>.
- [44] S.N. Omkar, J. Senthilnath, D.M. et.al, Crop classification using biologically-inspired techniques with high resolution satellite image, *Journal of the Indian Society of Remote Sensing* 36 (2008) 175–182, doi:10.1007/s12524-008-0018-y.
- [45] S. Ji, C. Zhang, A. Xu, Y. Shi, Y. Duan, 3d convolutional neural networks for crop classification with multi-temporal remote sensing images, *Remote Sens.* 10 (1) (2018), <https://doi.org/10.3390/rs10010075>.
- [46] N. Kussul, M. Lavreniuk, S. Skakun, A. Shelestov, Deep learning classification of land cover and crop types using remote sensing data, *IEEE Geosci. Remote Sens. Lett.* 14 (5) (2017) 778–782, <https://doi.org/10.1109/LGRS.2017.2681128>.
- [47] R. Xu, D. Wunsch, Survey of clustering algorithms, *IEEE Trans. Neural Networks* 16 (3) (2005) 645–678, <https://doi.org/10.1109/TNN.2005.845141>.
- [48] D. Xu, Y. Tian, A comprehensive survey of clustering algorithms, *Annals of Data Science* 2 (2) (2015) 165–193.
- [49] J. Macqueen, Some methods for classification and analysis of multivariate observations, in: In 5-th Berkeley Symposium on Mathematical Statistics and Probability, 1967, pp. 281–297.
- [50] H.-S. Park, C.-H. Jun, A simple and fast algorithm for k-medoids clustering, *Expert Systems with Applications* 36 (2, Part 2) (2009) 3336–3341.
- [51] K.L. R.P.J., Finding groups in data, Wiley, New-York, 1990.
- [52] F. Murtagh, A survey of recent advances in hierarchical clustering algorithms, *The Computer Journal* 26 (4) (1983) 354–359, <https://doi.org/10.1093/comjnl/26.4.354>.
- [53] J.C. Bezdek, R. Ehrlich, W. Full, Fcm: the fuzzy c-means clustering algorithm, *Computers and Geosciences* 10 (2) (1984) 191–203.
- [54] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, AAAI Press, 1996, pp. 226–231.
- [55] M. Ankerst, M.M. Breunig, H.-P. Kriegel, J. Sander, Optics: ordering points to identify the clustering structure, *SIGMOD Rec.* 28 (2) (1999) 49–60, <https://doi.org/10.1145/304181.304187>.
- [56] A.Y. Ng, M.I. Jordan, Y. Weiss, On spectral clustering: analysis and an algorithm, in: *Advances in Neural Information Processing Systems*, MIT Press, 2001, pp. 849–856.
- [57] B.J. Frey, D. Dueck, Clustering by passing messages between data points, *Science* 315 (5814) (2007) 972–976, <https://doi.org/10.1126/science.1136800>.
- [58] L. Khialia, M. Ndiathia, S. Alleaumea, D. Iencoab, K. Osea, M. Teisseirea, Detection of spatio-temporal evolutions on multi-annual satellite image time series: a clustering based approach, *Int. J. Appl. Earth Obs. Geoinf.* 74 (2019) 103–119, <https://doi.org/10.1016/j.jag.2018.07.014>.
- [59] M. Rußwurm, C. Pelletier, M. Zollner, S. Lefèvre, M. Körner, Breizhcrops: a time series dataset for crop type mapping, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* XLIII-B2-2020-1545–2020.
- [60] P. Montero, J.A. Vilar, Tsclust: an r package for time series clustering, *Journal of Statistical Software, Articles* 62 (1) (2014) 1–43, <https://doi.org/10.18637/jss.v062.i01>.
- [61] J.P., Nouvelles recherches sur la distribution florale, *Bulletin de la Société Vaudense des Sciences Naturelles* 44 (1908) 223–270.
- [62] N. Vinh, J. Epps, J. Bailey, Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance, *Journal of Machine Learning Research* 11 (2010) 2837–2854.
- [63] J. Demšar, Statistical comparisons of classifiers over multiple data sets, *J. Mach. Learn. Res.* 7 (2006) 1–30.

- [64] J. Derrac, S. García, D. Molina, F. Herrera, A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, *Swarm and Evolutionary Computation* 1 (1) (2011) 3–18, <https://doi.org/10.1016/j.swevo.2011.02.002>.
- [65] M. Friedman, A comparison of alternative tests of significance for the problem of m rankings, *Ann. Math. Stat.* 11 (1) (1940) 86–92.
- [66] P. Nemenyi, Distribution-free multiple comparisons, Ph.D. thesis, Princeton University (1963).
- [67] S. Holm, A simple sequentially rejective multiple test procedure, *Scand. J. Stat.* 6 (2) (1979) 65–70.
- [68] J.P. Shaffer, Modified sequentially rejective multiple test procedures, *Journal of the American Statistical Association* 81 (395) (1986) 826–831, <https://doi.org/10.1080/01621459.1986.10478341>.
- [69] C. Song, C.E. Woodcock, K.C. Seto, M.P. Lenney, S.A. Macomber, Classification and change detection using landsat tm data: when and how to correct atmospheric effects?, *Remote Sens Environ.* 75 (2) (2001) 230–244, [https://doi.org/10.1016/S0034-4257\(00\)00169-3](https://doi.org/10.1016/S0034-4257(00)00169-3).
- [70] A. Rivera, M. Pérez-Godoy, D. Elizondo, L. Deka, M. del Jesus, Supplementary material (results), <https://ajriverar.github.io/ClusteringCrops/ClusteringMethodsMeasuresWeb.xlsx>, accessed: 02-02-2022.
- [71] A. Rivera, M. Pérez-Godoy, D. Elizondo, L. Deka, M. del Jesus, Supplementary material (analysis), <https://ajriverar.github.io/ClusteringCrops/TablesLink.pdf>, accessed: 02-02-2022.



Antonio J. Rivera received the M.Sc. and Ph.D. degrees in Computer Science from the University of Granada, Granada, Spain, in 1995 and 2003, respectively. He is an Associate Professor with the Department of Computer Science at University of Jaén (Spain). He has been the supervisor of 2 Ph.D. students. She has published more than 25 journal papers and he is coauthor of the book “Multilabel Classification. Problem analysis, metrics and techniques” (Springer, 2016). His current main research interests include computational intelligence (including evolutionary algorithms and deep learning), data science (including data preprocessing, prediction, multi-label classification and big data) and information fusion.



M. Dolores Pérez-Godoy received the M.Sc. in Computer Science from the University of Granada in 1993 and the Ph.D. degree of Computer Science from the University of Jaén in 2010. She is an Associate Professor with the Department of Computer Science at University of Jaén (Spain). She has published 14 journal papers and more than 35 conference papers. Her current main research interests include computational intelligence (evolutionary algorithms and deep learning), data science (including data preprocessing, prediction, classification, and big data).



David Elizondo is a Professor in Intelligent Transport Systems at the Department of Computer Technology at De Montfort University (Directions from Train Station to DMU). After completing his BA in Computer Science from Knox College, Galesburg, Illinois, USA, he worked as a software engineer/lab manager for a latinoamerican agronomical research and teaching institute based in Costa Rica (CATIE). This institute, through a Swiss project, sponsored him to do a MS in Artificial Intelligence at the Department of Artificial Intelligence and Cognitive Computing of the University of Georgia, Athens, Georgia, USA. After this he obtained a PhD in computer science from the University of Strasbourg, France in cooperation with the Swiss Dalle Molle Institute for Perceptual Artificial Intelligence (IDIAP). He then

worked for Neuvoice, formerly Neural Systems, a spin off company of the University of Plymouth, UK. As a senior researcher he worked in the development of an intelligent monitoring system for the petroleum industry. This system was based on neural network techniques. Later, he worked as a software architect for ACTERNA, an international company which supplies software/hardware solutions to telecom companies. He was part of the team developing QMS, a quality of service management system for leased lines. In parallel to this work, he was a part time lecturer at the University of Plymouth where he taught database, and data structures and algorithms.



Lipika Deka is a Senior Lecturer in Computer Science at De Montfort University, Leicester, UK and a member the University's ITS research group (DIGITS). She received the BEng degree in Computer Science and Engineering followed by the MTech degree in Computer Science respectively from Dibrugarh University and Jawaharlal Nehru University, India. In 2013 she received her PhD degree in Computer Science and Engineering from Indian Institute of Technology Guwahati, India specializing in Concurrency Control Techniques for File Systems. Her research interest lies in the area of Smart Agriculture, Intelligent Transportation System, Obstacle Detection and Path Planning of Autonomous Vehicles, Concurrency Control Techniques, Transactional File Systems and Dynamic Software Updates.



María J. del Jesus received the M.Sc. and Ph.D. degrees in Computer Science from the University of Granada, Granada, Spain, in 1994 and 1999, respectively. She is a Professor with the Department of Computer Science at University of Jaén (Spain). She has been the supervisor of 8 Ph.D. students. She has published more than 65 journal papers and she is coauthor of the book “Multilabel Classification. Problem analysis, metrics and techniques” (Springer, 2016). Her current main research interests include computational intelligence (including fuzzy modeling, evolutionary algorithms and deep learning), data science (including data preprocessing, prediction, multi-label classification, subgroup discovery, emerging pattern mining, and big data) and information fusion.