Towards improved land use mapping of irrigated croplands: performance assessment of different image classification algorithms and approaches

Amit Kumar Basukala	extsuperscript{a}, Carsten Oldenburg	extsuperscript{a}, Jürgen Schellberg	extsuperscript{b,ab}, Murodjon Sultanov	extsuperscript{c} and Olena Dubovyk	extsuperscript{a}

\textsuperscript{a}Centre for Remote Sensing of Land Surfaces, University of Bonn, Bonn, Germany; \textsuperscript{b}Institute of Crop Science and Resource Conservation, University of Bonn, Bonn, Germany; \textsuperscript{c}Department of Geodesy, Cartography, Geography, Faculty of Natural Sciences, Urgench State University/KRASS NGO, Khorezm, Uzbekistan

ABSTRACT

Accurate agricultural land use (LU) map is essential for many agro-environmental applications. With advances in technology, object-based image classification and non-parametric machine learning algorithms evolved. Still, no particular method has universal applicability. This paper compares robust non-parametric machine learning algorithms, random forest (RF) and support vector machine (SVM), and a common parametric algorithm maximum likelihood (MLC) based on multiple Landsat 8 images. We have also assessed the classifier performance relative to the choice either pixel-based (PB) or field-based (FB) approach. The study area, a semi-desert irrigated region, lies in Khorezm province and Republic of Karakalpakstan in Uzbekistan. Accuracy assessment showed higher overall accuracy (OA) and kappa index (KI) of the nonparametric machine learning FB-RF and FB-SVM algorithms over the PB-RF, PB-SVM and PB-MLC algorithms. The lowest OA and KI occurred with the parametric FB-MLC. Based on the results, the FB machine learning non-parametric algorithms are recommended for mapping irrigated croplands.

ARTICLE HISTORY

Received 14 March 2017
Accepted 15 March 2017

KEYWORDS

Land use (LU) mapping; random forest; support vector machine; maximum likelihood; field-based; Uzbekistan

Introduction

Management of agricultural resources and its planning requires knowledge of the current state of the cropping system such as accurate information on the current and past agricultural land use (LU). Assessment of current cropping patterns and their intra- and inter-seasonal dynamics is usually achieved through agriculture LU mapping, a process that quantifies current land resources into a series of thematic categories representing the specific crop type. Agriculture LU maps, resulting from the digital image classification, provide important baseline information for agricultural decisions and monitoring applications, such as land degradation assessments and identifying LU change options (Dubovyk, Christopher, Khamzina, & Menz, 2013a; Dubovyk et al., 2013b), field-based (FB) estimations of crop yield (Fritsch, 2013) and sustainable planning of irrigation water distribution (Conrad et al., 2014).

Remotely sensed (RS) imagery and image classification methods provide cost-effective and accurate means to derive agriculture LU maps. There are many examples of using these LU maps by agriculturist, developers, GIS professionals, consultants and local planners without labour-intensive costly fieldwork in remote areas where access to training data is limited (Löw, Conrad, & Michel, 2015; Olofsson et al., 2014; Peña-Barragán, Ngugi, Plant, & Six, 2011; Thenkabail & Wu, 2012). Hence, RS has become a vital source of information in natural resource management including agriculture LU mapping. Since the beginning of the development of supervised RS classification methods in the 1970s, the vigorous effort for effective and meaningful use of satellite-based information has started and this process is still ongoing.

Image classification is a most commonly used technique to extract LU information from satellite imagery. In general, image classification can be performed using pixel-based (PB) and object-based (OB) approaches. The PB classification is currently the most commonly used method for classifying RS data (Whiteside, Boggs, & Maier, 2011). However, LU maps derived from PB classifications are reportedly having misclassifications due to the impact of mixed pixels from a combination of the reflectance from two or more LU types (Blaschke, 2010). Sub-PB classification approaches have potential to deal with the mixed pixels problem but are unable to deal with spectral variation in LU classes (Jawak, Devliyal, & Luis, 2015).

The OB classification works on the objects. The objects are commonly derived based on image segmentation (Jawak et al., 2015). The derived objects allow to use more features as an input for image classification. In agricultural LU mapping, the objects’ unit is the agricultural field. The borders of the fields could be derived via both image segmentation and digitisation from the available cadastral maps (Dubovyk et al., 2015). We refer, in this paper, to the FB classification approach when the GIS layer of the existing field borders was
used as an input for image classification. Similar to OB classification approach, the PB classification approach is also claimed to be most suitable for mapping mosaicked landscapes of irrigated croplands (Conrad et al., 2013; Löw et al., 2015). During the FB classification, each pixel is assigned to a final class of the entire field according to their statistical and other properties, instead of determining the class label for each pixel separately. Therefore, FB methods, being an OB method, also eliminate effect of spectral variability within fields and mixed pixels falling on field boundaries (De Wit & Clevers, 2004; Dubovyk et al., 2013b).

Whether pixels or objects/fields are used as a basis for the classifications of satellite imagery, there are still many classification algorithms that can be used (Duro, Franklin, & Dubé, 2012). Based on assumptions on data distribution, parametric and non-parametric classifiers are distinguished. Parametric methods assume that RS data have a normal distribution and homogenous variance. Non-parametric methods assume that RS data have any distribution and any variance (Jawak et al., 2015). Even though much has been done to improve classification accuracy (Chen & Stow, 2002), no single procedure has universal applicability as well as acceptability (Lu & Weng, 2007). No pattern classification method is inherently superior to any other (Jensen, 2015). The nature of the classification problem, the biophysical characteristics of the study area, the distribution of RS data and a priori knowledge determine which classification algorithm will yield useful results (Duro et al., 2012). These scenarios urge the need for more studies on the comparison of different algorithms using both PB and OB approaches to provide recommendations of algorithm applicability for specific applications and geographic area.

There is a very less research conducted on the comparison of PB and FB (based on the digitised fields) image classification methods with different classification algorithms. There are studies which are carried out based on fields created with image segmentation instead of digitised fields (Alganci, Sertel, Ozdogan, & Ormeci, 2013; Conrad, Fritsch, Zeidler, Rücker, & Dech, 2010; Duro et al., 2012; Duveiller & Defourny, 2010; Jawak et al., 2015; Löw et al., 2015; Löw, Michel, Dech, & Conrad, 2013; Myint, Gober, Brazel, Grossman-Clarke, & Weng, 2011; Peña-Barragán et al., 2011). Despite these, previous studies often focused on use on combination of two or more spatial resolution imageries and commercial imageries (Alganci et al., 2013). Further, there are studies which evaluated different classification algorithms using both PB and OB approaches without using any vegetation index as an input feature for classifier (Myint et al., 2011) or used a single classification method to measure the effects of spatial resolution on cultivated area (Conrad et al., 2010, 2013; Löw et al., 2015, 2013; Peña-Barragán et al., 2011). Comparison of the performance of different classification algorithms, when implemented with PB and FB approaches, had sought less attention. Most of the previous studies used either simple classification algorithms (e.g., K-Nearest Neighbour (K-NN), maximum Likelihood (MLC)) or relatively modern and robust supervised machine learning algorithms (e.g. decision tree, random forest (RF), support vector machine (SVM)) (Duro et al., 2012; Jawak et al., 2015).

In addition, the comparison of the most up-to-date parametric and non-parametric classifiers is rare for the PB and FB approaches.

The overall goal of this study was to provide explicit recommendations on the applicability of different classification methods and approaches for accurate mapping of irrigated croplands where digital cadastral information is available. In this context, the specific objectives of this study were:

(i) to map agriculture LU using both the PB and FB common parametric MLC algorithms and modern and robust supervised non-parametric machine learning algorithms: RF and SVM;
(ii) to compare the performance of the FB and PB classification results for mapping agricultural LU; and
(iii) to compare the performance of the RF, MLC and SVM algorithms for mapping agricultural LU.

The novelty of this study is in both methodological (development of the classification routines) and content-wise (recommendation on the use of classifiers for agricultural LU mapping) aspects.

**Study area**

The Khorezm Region of Uzbekistan and southern part of Autonomous Republic of Karakalpakstan (SKKP) are located in the northern part of Uzbekistan in the lower reaches of the Amu Darya River (Figure 1). Geographically, the area lies in a semi-desert zone in Central Asia characterised by an extreme continental climate, spreading between latitude 40°62’ to 42°71’ N and longitude 60°02’ to 62°44’ E. The study area is about 225 km south of the former shore of the Aral Sea. The natural sandy desert Karakum and Kyzylkum borders the study region in the south and east. Annual average precipitation of 100 mm falls outside crop-growing season (April–October), whereas mean annual evaporation is around 1500 mm (Conrad et al., 2014). Therefore, agriculture depends solely and heavily on irrigation water. The water is supplied from the Amu Darya River through a dense network of irrigation canals. The study area is among the final receivers of the irrigated water due to its downstream location. The water supply for irrigation from the Amu Darya River in the region has been recently not stable (Conrad et al., 2014). It is because the water flow had significantly decreased in the river due to the increasing upstream water use,
unpredictable climate and frequent drought events. These events had already triggered the major crop failures in 2000, 2001 and 2008 (Simonett & Novikov, 2010). Khorezm area has a population of 1.5 million of which 70% are engaged in crop production and in animal husbandry and horticulture.

Cotton and winter wheat are the two major crops cultivated in this area (UZSTAT, 2010). Cotton is the major export crop of the Uzbekistan, while wheat is a main grain crop of the country. Other cultivated crops are rice, watermelons, alfalfa, maize, sorghum and grapes. Fruit gardens are scattered throughout the region. The growing season of most of the crops is from April to September except for winter wheat which is grown from October to July (Figure 2). The cropping calendar was collected from regional agricultural land management authority of Khorezm region, Uzbekistan.

We selected the study area in Uzbekistan where intensive irrigated agriculture was conducted since the late 1950s resulted in dramatic environmental and socio-economic changes (Dubovyk et al., 2015). As such, improving water use efficiency in the agricultural sector is critical. Monitoring agricultural water use by crop type at different spatial scales could assist in formulating water resources management plans and policies and formulating proper LU plans (Zheng, Myint, Thenkabail, & Aggarwal, 2015). Such assessments require accurate and frequent site-specific information about crop distribution in the study area (Conrad et al., 2014).

**Materials**

**Earth observation data**

Multispectral cloud-free 30 m Landsat 8 Operational Land Imager (OLI) level 1 T data (Path 160/Row 31) of Khorezm Region was acquired for the year of 2013 (http://glovis.usgs.gov/). We selected Landsat images due to its quite high spatial and temporal resolution as well as their cost-effectiveness due to open policy of the Landsat archive (Wulder et al., 2016). The image selection was done in such a way that the acquired multi-temporal images covered the key crop-growth stages of the major crops (Figure 2). The imageries from 16 June, 18 July and 4 September were used in the analysis.

**Ancillary data**

Accurate ground truth data are crucial for training and validation of classification algorithms (Olofsson et al., 2014). A field sampling was conducted twice to integrate rotations of winter wheat with summer crops: May–June 2013 and September–October 2013. Summer crops included cotton, rice, maize, sorghum, vegetables, alfalfa and sunflower, whereas winter wheat is the only winter crop grown in the region. All classes observed during the field trips were sampled using a global positioning system (GPS) device following a previously developed field protocol procedure. Altogether 133 fields were sampled that represented 11 classes (Figure 3) (Table 1). For the FB classification methods, field boundaries were derived from the shapefile produced by digitising the cadastral maps field borders of the Khorezm area (Dubovyk et al., 2015).

**Table 1.**

| Month | Cotton | Wheat | Rice | Maize | Alfalfa | Sunflower | Watermelon |
|-------|--------|-------|------|-------|---------|-----------|------------|
| Jan   |       |       |      |       |         |           |            |
| Feb   |       |       |      |       |         |           |            |
| Mar   |       |       |      |       |         |           |            |
| Apr   |       |       |      |       |         |           |            |
| May   |       |       |      |       |         |           |            |
| Jun   |       |       |      |       |         |           |            |
| Jul   |       |       |      |       |         |           |            |
| Aug   |       |       |      |       |         |           |            |
| Sep   |       |       |      |       |         |           |            |
| Oct   |       |       |      |       |         |           |            |
| Nov   |       |       |      |       |         |           |            |
| Dec   |       |       |      |       |         |           |            |

**Figure 2.** Cropping calendar for study region of Uzbekistan. The highlighted column shows the month in which cloud-free Landsat imagery were available and downloaded for classification.
Methods

The methodology is illustrated in the flowchart in Figure 4. The methods involved were divided into sub-topics as described in the figure.

Datasets and processing

RS imagery

Landsat images were geometrically adjusted to an already geo-corrected 2.5 m SPOT-5 image and projected to the Universal Transverse Mercator (UTM) coordinate system (zone 41) based on differential GPS points (Conrad et al., 2010). On average, a total of 530 number of tie points per image were selected via an automated point matching in Earth Resource Data Analysis System (ERDAS) Imagine 2014 with the function Autosync (ImagineAutosync, 2008). Cubic convolution routine with the polynomial model (second degree) was used to co-register the images, resulting in an overall positional error of less than 0.5 pixels. Atmospheric correction was done on the June 2013 image using the ATCOR-2 (version 7.1) module for flat terrain (Richter & Schläpfer, 2007). The iteratively re-weighted multivariate alteration detection transformation (IR-MAD) was used for radiometric normalisation of Landsat image (Canty & Nielsen, 2008).

Ancillary data

Due to the unrepresentative number of samples and to keep a balanced number of samples per class, some classes were reorganised. Similar methods were already used in other crop mapping studies (De Wit & Clevers, 2004; Dubovyk et al., 2013b; Martinez-Casasnovas, Martin-Montero, & Casterad, 2005). In total, six classes were used for classification: “cotton”, “wheat”, “rice”, “fallow land”, “fruit trees” and class “other”. “Fruit trees” included mulberry and apple gardens. The “other” class included the sampled crop fields with less than seven samples per class and not rotated with “wheat” (“tugay forest”, “water” and remaining class defined as “other”). “Wheat” included winter wheat, winter wheat rotations with rice, maize and sorghum (Table 1). The training data were then sub-divided into two halves as training and validation data sets.

Image classification

For both FB and PB classifications, the input features were as follows: mean of the seven spectral bands (1–7) of Landsat 8 OLI images, mean of the Normalized Developed Vegetation Index (NDVI) and standard
deviation of the NDVI. The analysis was conducted in the software packages EnMAP Box 2.1, ENVI/IDL 5.1 and Arc GIS 10.2. The PB classifications were carried out with three algorithms:

1. non-parametric RF, SVM via EnMAP Box and
2. parametric MLC via ENVI (Section 4.3).

The FB classification method was also carried out for these three algorithms (Section 4.3). While the EnMAP Box does not support vector files they must be changed to validation and classification images. For this and to extract object features, Interactive Data Language (IDL)-based programs were developed and ran in ENVI.

**Classification algorithms used**

**RF algorithm**

RF algorithm is a non-parametric machine learning algorithm (Breiman, 2001). Many classification decision trees are grown in RF algorithm classification. The unclassified pixel or object is assigned with its associated attributes in a class by running through each of \( x \) decision trees. Each decision tree votes and classifies the pixel into one of \( y \) classes. The forest assigns the pixel to the class having the most votes from all the trees in the forest. The individual decision trees are grown from bagging training sets. By randomly resampling \( N \) times with replacement, two-third of the original data set is used for training a single tree, \( N \) being the number of samples in the original training set. The remaining one-third is put down the tree and left out of the bag for an internal accuracy assessment (Breiman & Cutler, 2014). RF runs efficiently on very large data sets, which is a useful characteristic when analysing satellite data. It also provides estimates of what variables are most important in the classification. For this study, the RF implemented in EnMAP box was used.

**SVM algorithm**

SVM algorithm is another non-parametric machine learning algorithm that can be used for image classification (Duro et al., 2012; Jensen, Im, Hardin, & Jensen, 2009; Löw et al., 2015). SVM algorithms discriminate the classes by fitting an optimal separating hyperplane (OSH) between classes using the training samples within feature space and to maximise the margins between OSH and the closest training samples (Foody & Mathur, 2006; Van Der Linden & Hostert, 2009). The points lying on the boundaries are called support vectors, and the middle of the margin is the OSH (Meyer, 2014). Training data points on the wrong side of the discriminant margin are given negative weights to reduce their influence. For linearly not separable classes, the training data are implicitly mapped by a kernel function into a higher dimensional space, wherein the new data distribution enables a better fitting of a linear hyperplane (Jensen, 2015).

In this study, training and parameterisation of the SVM model were performed with radial basis function (RBF) kernel. Generation of the rule images were
performed using SVM implementation in EnMAP Box which is based on the Library for Support Vector Machine (LIBSVM) (Chang & Lin, 2001). The algorithm parameterisation requires the definition of the kernel parameter γ and regularisation parameter C. Ideal values for these parameters depend on the distribution of the classes in the feature space. It was hence useful to test ranges of parameters using a grid search with internal performance estimation. By doing so, pairs of γ and C are tested and those parameters with the best performance were automatically used for the training of the final SVM model (Linden, Rabe, Held, & Wirth, 2014). After the parameterisation, this model was applied for classifying the images and used for accuracy assessment with validation data.

MLC algorithm
The MLC algorithm continues to be one of the most widely used supervised image classification algorithms (Campbell & Wynne, 2011; Myint et al., 2011). The algorithm is based on probability function and assumes that training data for each class in each band are normally distributed. The probability of a pixel belonging to each of a predefined set of m classes is calculated, and the pixel is then assigned to the class for which the probability is the highest (Dalponte, Bruzzone, & Gianelle, 2008; Lo & Yeung, 2007). We used ENVI for MLC image classification.

Results
Crop development curves
The resulting NDVI profiles representing the development of the mapped crops for 3 months are shown in Figure 5. The phenological patterns of “cotton”, “winter wheat” and “rice” were similar and corresponded with Landsat TM and MODIS.

![Temporal NDVI profiles](image-url)

Figure 5. Temporal NDVI profiles for the mapped classes sampled in 2013 at field level calculated from multi-temporal Landsat 8 OLI data. The error bars show mean NDVI and its standard deviation.
NDVI temporal profile reported in the previous studies from the same region (Conrad et al., 2014; Dubovyk et al., 2013b). The visual examination of the profiles showed that they allowed the trustworthy discrimination of all classes, taking into account simple phenological criteria like the minimum and maximum values of the NDVI profile and the range of NDVI over 3 months. The fluctuating maximum and minimum values of NDVI were shown by irrigated crops and “fruit trees”, and lower constant value of NDVI was shown by classes “fallow” and “other”.

Agriculture LU maps based on FB and PB image classification methods

The calculated LU maps based on both FB classification methods and PB classification methods using the three algorithms are presented in Figures 6 and 7, respectively. In all maps, the class “cotton” was evenly distributed across the entire region. Its estimated percentage coverage area in all six classified images was higher (38–53%) than rest classes (Table 2), which reflect the overall importance of cotton as the dominant export commodity. Rice fields (4–5%) were mainly located along the Amu Darya River and main irrigated canals. Wheat (9–25.6%) and fruit trees (12–29%) were mapped nearly evenly throughout the region. The class “other” (0.1–9%) appeared northeast of the study area. Fallow fields (12–29%) were mostly identified in the boundary of the irrigated region near the border to the desert. The cropping density was highest in the centre of the study region. Major differences among the results of the algorithms were found in the northern part of the study area for the class “fruit trees”, “wheat”, “fallow land” and “other”. The northern part of the study area is dominated by abandoned fallow cropland (Dubovyk, 2013d; Dubovyk et al., 2013c). Fallow lands were well captured by most of classification methods. Negligible fallow land is only depicted by FB-MLC. Considering the actual situation of the study area, the FB-RF map estimated a better agriculture LU map than the other methods. There were only minor visual differences between the three applied PB classification algorithms. In the PB-MLC map, the area under fruit trees (29%) was larger compared to the other PB analysis (Table 2).

Figure 6. Comparison of FB classification on the basis of agricultural LU map of the study area in 2013, a) RF classification; b) MLC; and c) SVM classification.
In general, all applied classification methods showed a reasonably accurate visual depiction of the agricultural LU of the area except for the FB-MLC classification.

### Accuracy assessment of OB crop classifications

With reference to overall accuracies (OAs), FB classification methods for SVM and RF algorithms performed well and similar, whereas MLC showed lower precision results. In the FB-RF confusion matrix, the OA was 87.69%, while the user’s accuracy (UA) for all classes was higher than 80%, except for the class “other” (75%). Producer’s accuracy (PA) for all classes were higher than 80% except for the class “other” (50%) (Table 3). In the FB-SVM confusion matrix, the OA was 89.23%, UA for all classes was higher than 80%, except for the class “other” (75%) and the class “wheat” (61.54%). The PA for all classes was higher than 80% except for the class “other” (50%). For the FB-MLC, the results of accuracy assessment were substantially lower compared to the FB-RF and FB-SVM: OA was equal to 66.87%, PA was very low for the class “fallow land” (18.52%) and the class “other” (23.08%). The UA for classes “wheat” (46.43%) and “fruit trees” (56.76%) were also very low. In response to kappa accuracies (KA), FB classification methods for SVM (0.86) and RF (0.84) algorithms
Table 3. Confusion matrices and associated classifier accuracies based on the validation data.

|        | PB-RF       |          |          |          |          |          | Total UA |          |          |          |          |          |          |
|--------|-------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|        | A           | B        | C        | D        | E        | F        |          |          |          |          |          |          |          |
| A      | 534         | 0        | 1        | 38       | 1        | 14       | 588      | 90.82%   |          |          |          |          |          |
| B      | 0           | 566       | 0        | 30       | 1833     | 2        | 2434     | 23.38%   |          |          |          |          |          |
| C      | 5           | 0        | 1023     | 2        | 10       | 8        | 1048     | 97.61%   |          |          |          |          |          |
| D      | 6           | 1        | 11       | 2243     | 25       | 34       | 2320     | 96.68%   |          |          |          |          |          |
| E      | 70          | 10       | 3        | 26       | 3307     | 20       | 3436     | 96.25%   |          |          |          |          |          |
| F      | 135         | 0        | 1        | 53       | 8        | 782      | 979      | 79.88%   |          |          |          |          |          |
| Total  | 750         | 590      | 1039     | 2392     | 5184     | 860      | 10,805   |          |          |          |          |          |          |
| PA     | 71.20%      | 98.10%   | 98.46%   | 93.77%   | 63.79%   | 90.93%   |          |          |          |          |          |          |          |
| OA     | 78.28%      |          |          |          |          |          |          |          |          |          |          |          |          |
| KA     | 71.7%       |          |          |          |          |          |          |          |          |          |          |          |          |

|        | FB-RF       |          |          |          |          |          | Total UA |          |          |          |          |          |          |
| A      | 8           | 0        | 0        | 0        | 0        | 0        | 8        | 100.00%  |          |          |          |          |          |
| B      | 0           | 6         | 0        | 0        | 0        | 0        | 6        | 100.00%  |          |          |          |          |          |
| C      | 0           | 0         | 4        | 0        | 0        | 0        | 4        | 100.00%  |          |          |          |          |          |
| D      | 0           | 1         | 1        | 23       | 3        | 0        | 28       | 82.14%   |          |          |          |          |          |
| E      | 0           | 0         | 0        | 0        | 3        | 1        | 4        | 75.00%   |          |          |          |          |          |
| F      | 2           | 0         | 0        | 0        | 0        | 13       | 15       | 86.67%   |          |          |          |          |          |
| Total  | 10          | 7         | 5        | 23       | 6        | 14       | 65       |          |          |          |          |          |          |
| PA     | 80.00%      | 85.71%    | 100.00%  | 50.00%   | 92.86%   |          |          |          |          |          |          |          |          |
| OA     | 87.69%      |          |          |          |          |          |          |          |          |          |          |          |          |
| KA     | 87.78%      |          |          |          |          |          |          |          |          |          |          |          |          |

|        | PB-SVM      |          |          |          |          |          | UA       |          |          |          |          |          |          |
| A      | 604         | 0        | 27       | 1825     | 16       | 2473     | 24.42%   |          |          |          |          |          |          |
| B      | 0           | 564       | 0        | 29       | 3        | 596      | 94.63%   |          |          |          |          |          |          |
| C      | 0           | 0        | 999      | 1        | 9        | 5101     | 98.52%   |          |          |          |          |          |          |
| D      | 3           | 12        | 26       | 2285     | 35       | 23       | 2389     | 95.85%   |          |          |          |          |          |
| E      | 7           | 3         | 7        | 3309     | 7        | 3366     | 99.19%   |          |          |          |          |          |          |
| F      | 136         | 1         | 64       | 47       | 3        | 809      | 78.44%   |          |          |          |          |          |          |
| Total  | 750         | 590      | 1039     | 2392     | 5184     | 860      | 10,805   |          |          |          |          |          |          |
| PA     | 80.53%      | 97.24%    | 96.15%   | 95.33%   | 63.83%   | 94.07%   |          |          |          |          |          |          |          |
| OA     | 78.28%      |          |          |          |          |          |          |          |          |          |          |          |          |
| KA     | 73.06%      |          |          |          |          |          |          |          |          |          |          |          |          |

|        | PB-MLC      |          |          |          |          |          | UA       |          |          |          |          |          |          |
| A      | 609         | 0        | 0        | 28       | 16       | 11       | 664      | 91.72%   |          |          |          |          |          |
| B      | 0           | 576       | 0        | 31       | 57       | 0        | 664      | 86.75%   |          |          |          |          |          |
| C      | 0           | 0        | 1025     | 6        | 1475     | 1        | 2507     | 40.89%   |          |          |          |          |          |
| D      | 0           | 0        | 2231     | 40       | 15       | 2292     | 97.34%   |          |          |          |          |          |          |
| E      | 0           | 0        | 4        | 3209     | 0        | 3209     | 100.00%  |          |          |          |          |          |          |
| F      | 139         | 4         | 10       | 96       | 367      | 833      | 1449     | 56.71%   |          |          |          |          |          |
| Total  | 750         | 590      | 1039     | 2392     | 5184     | 860      | 10,805   |          |          |          |          |          |          |
| PA     | 81.20%      | 99.31%    | 98.65%   | 93.27%   | 61.90%   | 96.86%   |          |          |          |          |          |          |          |
| OA     | 78.51%      |          |          |          |          |          |          |          |          |          |          |          |          |
| KA     | 72.01%      |          |          |          |          |          |          |          |          |          |          |          |          |

A = wheat, B = rice, C = fallow, D = cotton, E = other, F = fruit trees; OA = overall classification accuracy, PA = producer's accuracy, UA = user's accuracy and KA = kappa accuracy.

performed well and similar, whereas MLC (0.60) showed lower results.

**Accuracy assessment of the PB crop classifications**
All PB classification methods for SVM, RF and MLC algorithms performed good and similar. In the PB-RF confusion matrix the OA was 78.28 %, UA for all classes was higher than 80%, except for the class “other” (79.88 %) and for the class “rice” (23.38%). The PA for all classes were higher than 71.2%, except for the class “other” (63.79%) (Table 3). In the PB-SVM confusion matrix, the OA was 80 %, UA for all classes was higher than 80%, except for the class “wheat” (24.42%), and PA for all classes were higher than 80% except for the class “other” (63.83%) (Table 3). In the PB-MLC confusion matrix, the OA was 78.51%, UA for all classes was higher than 86%, except for the class “fallow land” (40.89%) and the class “fruit trees” (61.9%). PA for all classes were higher than 80% except for the class “other” (61.90%) (Table 3). In response to KA, all classification methods, SVM (0.73), MLC (0.72) and RF (0.71) algorithms, performed well and similar.

**Discussion**
The results showed that non-parametric FB classification (RF and SVM) yielded in the highest OA and KA compared to PB methods and produced visually appealing agricultural LU maps. This result is consistent with findings from other studies performed in other environments (Castillejo-Gonzalez et al., 2009; Gao & Mas, 2008; Gao, Mas, Maathius, Xiangmin, & Van Dijk, 2006; Whiteside et al., 2011). Review paper of Jawak et al. (2015) also concluded that satellite images of medium resolution and high resolution can be best classified with the OB approaches. The improved classification results using FB analysis in the study area can be attributed primarily to the use of field objects that reduced the spectral variability of the complex LU types of this irrigated area. The FB also eliminated the “salt and pepper” effect by considering mean pixel values within field objects as opposed to individual pixel values (Figure 8). Our study also confirmed the reported disadvantage of the PB approaches which do misclassify pixels, particularly in spectrally heterogeneous areas (Whiteside et al., 2011). It is acknowledged that the use of the field object features enhances the classification accuracy and image (Conrad et al., 2013).

In contrary, the parametric FB-MLC showed very low OA and KA and also visually inaccurate agriculture LU maps. This is due to fact that the effectiveness of the MLC algorithm depends upon a reasonably accurate estimation of the mean vector and the covariance matrix for each spectral class. Thus, it relies on a sufficient number of training samples per class and it performs unsatisfactorily with a low number of training samples as was the case in the studies of Richards and Jia (2006) and Cord, Conrad, Schmidt, and Dech (2010).

When comparing algorithms, RF and SVM performed better than MLC in FB analysis while all three algorithms performed equally well (but still worse than the FB classifications) in the case of the PB classifications. The reason for this could be that RF and SVM algorithms work very well on a small number of
training samples as well as with data without Gaussian distribution (Jensen, 2015), which is not the case for MLC (Waske, 2007). MLC shows fair results for the PB classification as there is a higher number of training pixels per class (580–5184) compared to the low number of training objects per class (26–30) for FB analysis.

When comparing OA and KA, FB-SVM and FB-RF performed best producing the highest accuracies among tested methods (Figure 9(a)). Considering UA, FB-RF showed good results of above 75% for all classes (Figure 9(c)). FB-SVM also showed good results for all classes above 75%, except for class “wheat” (61.54%) (Table 3 and Figure 9(c)). This signifies an overestimation of wheat by FB-SVM. Similarly, the low UA accuracies in FB-MLC for class “wheat” and “fruit trees”, in PB-RF for class “rice”, in PB-SVM for class “wheat” and in PB-MLC for classes “fallow land” and “fruit trees” indicated an overestimation of these classes in the respective algorithm. In relation to PA, FB-RF, FB-SVM, PB-SVM and PB-MLC showed better results above 80% for all classes except for the class “other” (Figure 9(b)). In PB-RF, “wheat” and “other” showed fair PA. In FB-MLC class, “fallow land” (18.52%) and “other” (23.08%) had an unacceptable low PA. The low and fair result of PA

Figure 9. Comparison of the (a) OA and KA, (b) Producer’s accuracy and (c) User’s accuracy of FB- and PB-RF, SVM and MLC methods.
signified the wrong identification of these classes (Figure 9b). Due to this, in LU map derived from FB-MLC, the mapped area of the class “fallow land” (0.3%) and the class “other” (0.1%) was less in comparison to other classification methods’ results (Table 2). With high OA, KA, UA and PA for all classes, RF and SVM can be recommended as an effective and accurate classifier for agriculture LU mapping where the constraints of training samples occur, confirming previous findings (Waske & Braun, 2009), also for the irrigated croplands of Central Asia (Löw et al., 2015).

The “rice”, “cotton” and “fallow land” were properly identified by all methods and the UA and PA were also acceptable (Figure 9). The lowest accuracies were found for the class “other”, “wheat” and “fruit trees” (Table 3). This is partially due to the fact that these classes are very heterogeneous encompassing multiple crops with different spatial and spectral characteristics and also due to a lower number of samples for these classes (Table 1). An adequate number of training samples and one or two additional acquisition dates of satellite images would have been essential for a better discrimination of these classes and a better integration into the classification rules.

The same problem with a collective class for crops covering minor area portions (“other” crops) was observed by Conrad et al. (2010). For a successful mapping of the entire diversity of crop classes, it is recommendable that the sampling strategy in the future should be more focused on including also minor crops to avoid the combination of the minor classes in classes like “other” or mixed crops like wheat. In our study, this led to an increased confusion among these collective classes, which was also reported in the study by De Wit and Clevers (2004) and Conrad et al. (2014). The distinct NDVI profiles of each class (Figure 5) indicated the potential of the spatial resolution of Landsat imagery for including minor classes into the classification.

Implications for land and water management
The availability of agricultural LU maps is of high interest for mid- to long-term land and water management applications in irrigated croplands of northern Uzbekistan and similar environments. Accurate maps allowed for identifying actually cultivated crops in the study area such as cotton, fruit trees and rice fields. Under environmental conditions where water is permanently scarce, rice cultivation can be reflected critically as the water requirement for rice is very high. Instead, other crops with low water requirements or drought-resistant rice variety could be introduced.

Another important aspect of using such maps is that the combination of the produced RS-based agricultural LU maps with the crop models offers fast and low-cost information about the amount of water needed for agricultural production for large areas at different scales (Conrad et al., 2013). This can be used to increase the efficiency of irrigated water use as well as to promote sustainable land management.

The area of fallow land in the obtained maps showed that the land abandonment is increasing at an alarming rate in this area. The main reason for this is water scarcity and widespread land degradation (Dubovyk et al., 2015). In Khorezm, agriculture LU maps were already derived for supporting degradation assessments and identifying LU change options (Conrad et al., 2010, 2013; Dubovyk et al., 2013a, 2013b) or for FB estimations of crop yield (Fritsch, 2013). For future prospects, calculation of crop water requirement and crop irrigation requirement could be carried out based on the resulted accurate LU maps for each year using various models (e.g. SWAT (soil and water assessment tool), SEBAL (surface energy balance model)). Thus, water demand and supply difference could be assessed by comparing the water needed by crops and the total amount of water officially allocated for irrigation. Therefore, mapping spatially distributed of crop water/irrigation requirement and mapping of deficit water requirement could be now possible in the study area.

Conclusions
Classifications of satellite imagery using PB and OB approaches were performed using three algorithms (RFs, SVM and MLC) providing the methodological and practical basis for conducting a statistically rigorous comparison between agricultural LU maps. The study was based on freely available satellite imagery (Landsat 8 OLI images) using open-source software (IDL-based program, EnMAP Box). Hence, the methodology followed in this study suits for developing countries with the limited funds available for agro-environmental assessments.

The non-parametric FB-RF and SVM classifiers yielded the highest overall and KAs among the tested methods (RF, SVM and MLC) and produced visually appealing agriculture LU maps for the irrigated croplands in Uzbekistan. We, therefore, conclude that the FB classification methods were better suited for mapping heterogeneous irrigated cropland. These methods also have potential as alternative to commonly used PB approaches for crop mapping from medium to high-resolution satellite imagery.

Among three classifiers used, the SVM and RF performed equally well, suggesting their usefulness, especially for the cases when the number of training samples is limited. The MLC is recommended for PB classifications when the size of training data set is sufficiently large. The low user’s accuracies and producer’s accuracies were found in class “wheat”, class “fruit trees” and class “other” which had the lowest number of training
samples and were most spectrally confusing classes. For more successful mapping of the entire agriculture cropland classes, sampling strategies need to be more focused on minor crops to avoid the combination of minor classes and decrement of accuracy by mixed crops.

The derived LU maps could be used as guidance for agricultural LU planning characterised by less pressure on water sources within irrigated agricultural areas that could be economically diverted from the Amu Darya River in Uzbekistan. The methodology, as well as recommendations developed in this study, could be also applicable for similar irrigated agricultural environments in other Central Asian countries and beyond.

Acknowledgements

The authors highly appreciate the support of the German Academic Exchange Service (DAAD) (Germany) for providing scholarship during this study period. We would like to acknowledge the contribution of Mr. Joseph Scepan of Medford, Oregon, USA, in producing the document.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Deutscher Akademischer Austauschdienst.

References

Alganci, U., Sertel, E., Ozdogan, M., & Ormeci, C. (2013). Parcel-level identification of crop types using different classification algorithms and multi-resolution imagery in southeastern Turkey. Photogrammetric Engineering & Remote Sensing, 79, 1053–1065. doi:10.14358/PERS.79.11.1053

Blaschke, T. (2010). Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 65(Issue 1), 2–16. doi:10.1016/j.isprsjprs.2009.06.004

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. doi:10.1023/A:1010933404324

Breiman, L., & Cutler, A. (2014). Random Forests. http://www.stat.berkeley.edu/~bREiMAN/RandomForests/cc_home.htm.

Campbell, J.B., & Wynne, R.H. (2011). Introduction to remote sensing (5th ed., pp. 684). New York: Guilford.

Canty, M., & Nielsen, A.A. (2008). Automatic radiometric normalization of multispectral satellite imagery with the iteratively re-weighted MAD transformation. Remote Sensing of Environment, 112, 1025–1036. doi:10.1016/j.rse.2007.07.013

Castillejo-Gonzalez, I.L., Lopez-Granados, F., Garcia-Ferrer, A., Pe’ Na-Barragan, J.M., Jurado-Expósito, M., Sanchez-Dela Orden, M., & González-Audicana, M. (2009). Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. Computers and Electronics in Agriculture, 68, 207–215. doi:10.1016/j.compag.2009.06.004

Chang, C.C., & Lin, C.J. (2001). LIBSVM: A library for support vector machines. Software available at (last access: Sep. 2007) http://www.csie.ntu.edu.tw/~cjlin/libsvm

Chen, D., & Stow, D. (2002). The effect of training strategies on supervised classification at different spatial resolutions. Photogrammetric Engineering & Remote Sensing, 68, 1155–1161.

Conrad, C., Dech, S., Dubovyk, O., Fritsch, S., Klein, D., Löw, F., … Zeidler, J. (2014). Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multitemporal RapidEye images. Computer and Electronics in Agriculture, 103, 63–74. doi:10.1016/j.compag.2014.02.003

Conrad, C., Fritsch, S., Zeidler, J., Rücker, G., & Dech, S. (2010). Per-field irrigated crop classification in arid Central Asia using SPOT and ASTER data. Remote Sensing, 2, 1035–1056. doi:10.3390/rs2041035

Conrad, C., Rahmann, M., Machwitz, M., Stulina, G., Paeth, H., & Dech, S. (2013). Satellite-based calculation of spatially distributed crop water requirements for cotton and wheat cultivation in Fergana Valley, Uzbekistan. Global and Planetary Change, 110, 88–98. doi:10.1016/j.gloplacha.2013.08.002

Cord, A., Conrad, C., Schmidt, M., & Dech, S. (2010). Standardized FAO-LCCS land cover mapping in heterogeneous tree savannas of West Africa. Journal of Arid Environments, 74, 1083–1091. doi:10.1016/j.jare.2010.03.012

Dalponte, M., Bruzzone, L., & Gianelle, D. (2008). Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. IEEE Transactions on Geoscience and Remote Sensing, 46, 1416–1427. doi:10.1109/TGRS.2008.916480

De Wit, A.J.W., & Clevers, J.G.P.W. (2004). Efficiency and accuracy of per-field classification for operational crop mapping. International Journal of Remote Sensing, 25, 4091–4112. doi:10.1080/01431160310001619580

Dubovyk, O. (2013d). Multi-scale targeting of land degradation in northern Uzbekistan using satellite remote sensing. (Doctoral dissertation). Retrieved from http://hss.ulb.uni-bonn.de/2013/3398/3398.htm

Dubovyk, O., Christopher, C., Khamzina, A., & Menz, G. (2013a). Object-Based Retro-Classification of Agricultural Land Use: A Case Study of Irrigated Croplands. ESA Living Planet Symposium 2013. Edinburgh, UK.

Dubovyk, O., Menz, G., Conrad, C., Kan, E., Machwitz, M., & Khamzina, A. (2013c). Spatio-temporal analyses of cropland degradation in the irrigated lowlands of Uzbekistan using remote-sensing and logistic regression modeling. Environmental Monitoring and Assessment, 185(6), 4775–4790. doi:10.1007/s10661-012-2904-6

Dubovyk, O., Menz, G., Conrad, C., Lamers, J.P.A., Lee, A., & Khamzina, A. (2013b). Spatial targeting of land rehabilitation. A Relational Analysis of Cropland Productivity Decline in Arid Uzbekistan. Erikkund, 67, 167–181.

Dubovyk, O., Menz, G., Lee, A., Schellberg, J., Thonfeld, F., & Khamzina, A. (2015). SPOT-based sub-field level monitoring of vegetation cover dynamics: A case of irrigated croplands. Remote Sensing, 7, 6763–6783. doi:10.3390/rs70606763.

Duro, D.C., Franklin, S.E., & Dubé, M.G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sensing of Environment, 118, 259–272. doi:10.1016/j.rse.2011.11.020
Duveiller, G., & Defourny, P. (2010). A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing. *Remote Sensing of Environment, 114*, 2637–2650. doi:10.1016/j.rse.2010.06.001

Foody, G.M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by SVM. *Remote Sensing of Environment, 103*, 179–189. doi:10.1016/j.rse.2006.04.001

Fritschi, S. (2013). Spatial and temporal patterns of crop yield and marginal land in the Aral Sea Basin: derivation by combining multi-scale and multi-temporal remote sensing data with a light use efficiency model (Doctoral dissertation). Retrieved from https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/7186.

Gao, Y., & Mas, J.F. (2008). A comparison of the performance of pixel based and object based classifications over images with various spatial resolutions. *Online Journal of Earth Sciences, 2*(8701), 27–35. http://doc.sdrive.com/pdfs/medwelljournals/ojesi/2008/27-35.pdf

Gao, Y., Mas, J.F., Maathuis, B.H.P., Xiangmin, Z., & Van Dijk, P.M. (2006). Comparison of pixel-based and object oriented image classification approaches – a case study of a coal fire area, Wuda, inner Mongolia, China. *International Journal of Remote Sensing, 27*, 4039–4055. doi:10.1080/0143116060070072632

ImaginAutosync. (2008). ImaginAutosync user’s guide. The earth to business company. Retrieved from https://faculty.unr.edu/cas/szemans/r/docs/AutoSync.pdf

Jawak, S.D., Devliyal, P., & Luis, A.J. (2010). Classification by SVM. *International Journal of Applied Earth Observation and Geoinformation, 11*, 1301–1316. doi:10.1016/j.rse.2010.01.009

Richards, J.A., & Jia, X. (2006). Remote sensing digital image analysis: An introduction (4th ed.). Berlin: Springer-Verlag Berlin Heidelberg.

Richter, R., & Schläpfer, D. (2007). Atmospheric/Topographic Correction for Satellite Imagery (ATCOR-2/3 user guide). Retrieved from http://www.researchgate.net/publication/220154316Topographic_Correction_for_Satellite_Imagery.

Simonett, O., & Novikov, V. (2010). Land degradation and desertification in Central Asia: Central Asian countries initiative for land management analysis of the current state and recommendation for the future. Zoi Environment Network website. Retrieved from http://www.zoinet.org/earth-to-business Baldan (2010). Impact of a coal fire area, Wuda, inner Mongolia, China. *International Journal of Remote Sensing, 11*, 2298–2305. doi:10.1016/j.rse.2009.06.004

Waske, B. (2007). Classifying multisensor remote sensing data: concepts, algorithms and applications (Doctoral dissertation). Retrieved from http://hss.ulb.uni-bonn.de/diss-online.

Waske, B., & Braun, M. (2009). Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*(5), 450–457. doi:10.1016/j.isprsjprs.2009.01.003

Whiteside, T.G., Boggs, G.S., & Maier, S.W. (2011). Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation, 13*, 884–893. doi:10.1016/j.jag.2011.06.008

Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B.,… Roy, D.P. (2016). The global Landsat archive: Status, consolidation, & direction. *Remote Sensing of Environment, 185*, 271–283. doi:10.1016/j.rse.2015.11.032

Zheng, B., Myint, S.W., Thenkabail, P.S., & Aggarwal, R.M. (2015). A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. *International Journal of Applied Earth Observation and Geoinformation, 34*, 103–112. doi:10.1016/j.jag.2014.07.002
Appendix 1

Acronyms

ATCOR          Atmospheric Correction
FB             Field based
FB-MLC         Field-based Maximum Likelihood Classification
FB-RF          Field-based random forests classification
FB-SVM         Field-based support vector machine classification
GIS            Geographic information system
GPS            Global positioning system
IDL            Interactive data language
IR-MAD         Iteratively re-weighted multivariate alteration detection transformation
KA             Kappa accuracy
K-NN           K-Nearest neighbour
LIBSVM         Library for support vector machine
LU             Land use
LULC           Land use land cover
MLC            Maximum likelihood classifier
NDVI           Normalized Developed Vegetation Index
OA             Overall accuracy of image classification
OB             Object-based
OLI            Operational land imagery
OSH            Optimal separating hyper plane
PA             Producer’s accuracy of image classification
PB             Pixel-based
PBIA           Pixel-based image analysis
PB-MLC         Pixel-based maximum likelihood classification
PB-RF          Pixel-based random forests classification
PB-SVM         Pixel-based support vector machine classification
RBF            Radial basis function
RF             Random forests
RS             Remotely sensed
SEBAL          Surface energy balance model
SKKP           Southern part of autonomous Republic of Karakalpakstan
SWAT           Soil and water assessment Tool
SVM            Support vector machines
UTM            Universal Transverse Mercator
UA             User’s accuracy of image classification