A rule-based approach for crop identification using multi-temporal and multi-sensor phenological metrics

Gohar Ghazaryan a,b, Olena Dubovyk a,b, Fabian Löw c, Mykola Lavreniuk d, Andrii Kolotii e, Jürgen Schellberg e and Nataliia Kussul a

a Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn, Bonn, Germany; b Remote Sensing Research Group (RSRG), University of Bonn, Bonn, Germany; c MapTailor Geospatial Consulting GbR, Bonn, Germany; d Department of Space Information Technologies and Systems, Space Research Institute of National Academy of Sciences of Ukraine and State Space Agency of Ukraine, Kyiv, Ukraine; e Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Bonn, Germany

ABSTRACT
Accurate classification and mapping of crops is essential for supporting sustainable land management. Such maps can be created based on satellite remote sensing; however, the selection of input data and optimal classifier algorithm still needs to be addressed especially for areas where field data is scarce. We exploited the intra-annual variation of temporal signatures of remotely sensed observations and used prior knowledge of crop calendars for the development of a two-step processing chain for crop classification. First, Landsat-based time-series metrics capturing within-season phenological variation were preprocessed and analyzed using Google Earth Engine cloud computing platform. The developmental stage of each crop was modeled by fitting harmonic function. The model’s output was further used for the automatic generation of training samples. Second, several classification methods (support vector machines, random forest, decision fusion) were tested. As input data for crop classification, composites based on Sentinel-1 and Landsat images were used. Overall classification accuracies exceeded 80% when the seasonal composites were used. Winter cereals were the most accurately classified, while we observed misclassifications among summer crops. The proposed approach offers a potential to accurately map crops in the areas where in situ field data are scarce or unavailable.

Introduction
Due to a growing world population and decreasing land and water resources, there is a need for enhancing agricultural productivity to ensure food security. Accurate crop maps from Earth Observation can build the basis for agricultural monitoring and decision-making at wider spatial scales (Kussul, Lemoine, Gallego, Skakun, & Lavreniuk, 2015a; Löw & Duveiller, 2014) to support sustainable agricultural land management. Improving the resiliency of food production systems and implementing sustainable agricultural practices have been recognized as one of the most important development goals (United Nations, 2015). Spatially explicit information on the distribution of croplands and crop types can also assist accurate statistical estimation such as yield prediction and crop area estimation (Kogan et al., 2013; Kussul et al., 2015b; Pan et al., 2012) and so support policy making (Davidson, Fisette, Mcnairn, & Daneshfar, 2017).

The recent advancement of satellite remote sensing, such as the free distribution of satellite archives and the availability of new sensors, gives more opportunities to derive land cover and land use information. Also, the frequency of data acquisition is nowadays rather high, which allows not only to discriminate different crops but also to assess their growth stage.

Several studies discussed the use of multi-temporal imagery for classification (Hentze, Thonfeld, & Menz, 2016; Liu, Heiskanen, Aynekulu, Maeda, & Pellikka, 2016; Siachalou, Mallinis, & Tsakiri-Strati, 2016; Simonetti, Simonetti, Szantoii, Lupi, & Eva, 2015; Yan, Wang, Lin, Xia, & Sun, 2015) often using vegetation indices, such as Normalized Difference Vegetation Index (NDVI), for the identification of croplands and crop types from other types of land cover (Hao, Wang, Zhan, & Niu, 2016; Marais Sicre et al., 2016). Previous studies also discussed the use of knowledge-based temporal features for crop identification based on several optical sensors. These studies applied a phenology-based approach for the identification of areas cultivated with different crop types (Bargiel, 2017; Waldner, Canto, & Defourny, 2015; Zhong, Hu, Yu, Gong, & Biging, 2016). However, these multi-temporal classification approaches either rely on a high number of
training data and/or cannot be applied in years without training data. Moreover, the integration of several data sources and the discrimination of crop types are still challenging as most of the studies are based on the use of remote sensing data coming from one sensor.

Hence, the present study proposes an approach to automatically generate training samples based on phenological metrics derived from both optical and radar image time series. We applied this method to the study site in Central Ukraine as this region has undergone profound changes during the last decades in the extent and intensity of agricultural land use (Kuemmerle et al., 2011). For the study area, Kussul et al. (2015b) demonstrated successful classification of the agricultural land use. However, the crop identification methods applied were relying on extensive field data collection which was further used for calibration of the classification model.

Although a number of studies have addressed the use of satellite data products for crop mapping (Loosvelt, Peters, & Skriver, 2012; Löw, Michel, Dech, & Conrad, 2013; Wardlow & Egbert, 2008), there is still an ongoing discussion regarding the choice of input data. Further, when using only optical remote sensing data, unfavorable atmospheric conditions such as cloud cover can reduce the quality of classification results (Forkuor, Conrad, Thiel, Ullmann, & Zoungrana, 2014; Joshi et al., 2016; Whitcraft, Becker-Reshef, & Justice, 2015). Therefore, analyses are needed which strive for a better understanding of the effect of type and number of input variables on classification accuracy. Such analysis should also consider an increase in amount and quality of freely distributed data especially when data from multiple sensors are combined. In our study, we thus tested the performance of several nonparametric classification algorithms (support vector machines (SVMs), random forest (RF) and decision fusion (DF)) and applied them to multi-temporal datasets. As these algorithms are widely used for classification, the comparison of these approaches is important to test their applicability for crop type mapping.

In this light, the objectives of the research were: (1) to develop a two-step procedure for crop identification based on analysis of crops’ within-season temporal dynamics and application of nonparametric classification algorithms, (2) to identify a favorable selection of input data derived from optical and Synthetic Aperture Radar (SAR) sensors and (3) to compare the applicability of classification methods and a training sample generation approach across years.

**Study site**

We selected the Vasylkovsky district in Kiev region as the test site (Figure 1). This area is located in Central Ukraine in the semi-steppe zone, comprising 1184 km². In 2015, the total sown area was 58,618.36 ha, while in 2016 it comprised only 46,092.52 ha (State Statistical Department of Kiev region, 2015, 2016). The test site is characterized by a temperate climate and a heterogeneous agricultural land use. Agricultural production in the region is highly dependent on rainfall, as the area is not irrigated.

In this region, two major cropping seasons are distinguished. Winter crops are grown from October to July of the following year, whereas summer crops are grown from April to September. In this area, during one growing season, one crop is planted and after the harvest of winter crops, no other crop is planted. Although the area is located in the temperate zone,

![Figure 1. Location of the study area (Sentinel-1 multi-temporal composite, RGB bands: VH images from May, August and October in 2016).](image)
agroclimatic conditions were different during the study years. Particularly, in 2015 the weather was drier lacking rain from mid-July. Although growing conditions in 2016 were more favorable, the average temperature and precipitation deviated from the long-term mean.

The agricultural field size in this area varies mostly from 30 to 100 ha. Small fields have a size of less than 5 ha (5% of the fields in the area) and large fields can even reach up to 250 ha. Such differences occur due to several factors such as different types of cropland tenure, ownership (agricultural holdings, household farms).

Wheat is the main winter cereal, which is typically planted in late September or October. The seeding date depends on weather conditions as plants start growing when temperature exceeds 1–2°C, while the optimal temperature for growth ranges from 14°C to 20°C. The active vegetative period ends in November but starts again in the following spring (Day of the year (DOY) 60–85). Winter wheat usually reaches the flowering stage in the second part of May; and the crop is consequently harvested in July and early August (DOY 11–123) (Figure 2). Besides winter wheat, other winter cereals, such as barley, are also sown, but usually the cultivated area is smaller compared to winter wheat. Another common winter crop is rapeseed, which is one of the main oilseed crops grown in this area. It is also planted in fall. The distinctive flowering period is in late April, which usually lasts for 35 days. The duration of the growing season of rapeseed is around 300–320 days.

Maize, soybean and sunflower are the main summer crops cultivated in the study site. Maize is sown between late April and May: the crop reaches the peak of the vegetative phase in July and is harvested till mid-September. Maize typically reaches the flowering stage in mid-July (DOY 190–202). Soybean is usually sown in May, reaching a flowering stage in July and maturing in August (the overall growing period is 80–240 days). Sunflower is another important oilseed crop, which exhibits similar crop development stages during the growing season: seeding is in late April, flowering in July and harvest is in early September (growing period: 95–120 days) (State Hydrometeorological Services of Ukraine, 2007).

The other crops which are cultivated in this area are sugar beet, buckwheat, oat, spring wheat and barley. However, due to relatively small area (0.88–10% of cultivated area in the study region), they were not included in the study. In the following section, these classes were designated as class “other”.

Datasets

Satellite data

Landsat 8 Operational Land Imager time series were used as the primary source of data. Top of Atmosphere collection was accessed using Google Earth Engine (GEE) (Google Inc, 2016; Lessel & Ceccato, 2016). The Landsat archive was chosen due to its 30m spatial resolution, which is suitable to characterize land use at field level given the average field size in the study area. As the study area is in the overlapping section of two Landsat paths (181/25 and 182/25), four optical acquisitions are available per month. In Figure 3, the selected Landsat scenes are illustrated (several Landsat scenes were excluded due to extensive cloud cover). Some of the Landsat bands,
such as costal blue (Band 1), thermals (Band 10 and 11), panchromatic (Band 8) and cirrus (Band 9) were not used in the study. Following initial tests, Blue (Band 2) and Green (Band 3) bands were excluded from the classification framework as well.

In addition to the optical images, Sentinel-1A C-band SAR images with VH (Vertical transmit, Horizontal receive) polarization were used (Interferometric Wide swath mode (IW), Processing level 1 high-resolution products, Ground Range Detected). As SAR backscatter is sensitive to both geometric (e.g., crop structure, roughness) and dielectric properties (permittivity, which highly depends on water content) of the targets, it provides additional information to optical data (Inglada, Vincent, Arias, & Marais-Sicre, 2016; Weng, 2011). When using multi-temporal SAR data, it is important to use the data with the same viewing configuration. For this, we used the images in descending mode, and we filtered images by the orbit, as the difference in orbit might cause changes in backscattering that are not related to (bio)-physical properties of the target.

**In situ data**

*In situ* phenological observations acquired from the State Hydrometeorological Services of Ukraine were used for the assessment of seasonality parameters derived from remote sensing. This data represents long-term observations of dates of the main phenological events for each crop. These key dates of crop development can be grouped into the following temporal classes: initial (low vegetation, beginning of emergence), development of leaves (greening up), mid-season peak and flowering, late season (grain filling and senescence) and harvest. According to the collected *in situ* crop phenological data, among recent years these plant growth stages generally corresponded to crop calendar (State Hydrometeorological Services of Ukraine, 2016).

Land cover information was collected using a mobile device with the GPS during the growing season of each year. The field data collection was carried out during the entire growing season of all crops, which enabled the recognition of winter crops (cereals and rapeseed) and summer crops (maize, sunflower and soybean). The reference dataset for each crop is listed in Table 2.

**Methods**

**Preprocessing**

Prior to the analysis, remotely sensed time series went through several preprocessing steps. We filtered the Landsat data collection by the time interval corresponding to the monitoring period and to the extent of the study area using GEE. The Fmask (Function of mask) method was applied for cloud and cloud-shadow detection in Landsat imagery (Zhu, Wang, & Woodcock, 2015). Temporal SAR composites were created and, subsequently, included into the classification scheme. Prior to compositing, it is important to filter the speckle noise in SAR images. We used GammaMap filter for this purpose (Lopes, Nezry, Touzi, & Laur, 1990) with the $7 \times 7$ kernel size. The Sentinel-1 and Landsat imagery were orthorectified to the same map projection and resampled to a common 30 m pixel resolution (Figure 4).

The near-infrared and red bands of Landsat were used to compute NDVI images for each individual time step.

**Two-step approach**

A two-step approach for crop identification was developed which included generation of training samples and a pixel-wise classification per se.

**Training sample generation**

To automatically generate training samples, a processing chain was developed to separate the main crops in the study area. First, a harmonic function was fitted to the time series of NDVI data. By fitting a harmonic function (Equation 1), it has been possible to model temporal signal of NDVI as a sum of additive term and harmonic term, which were then expressed by phase and amplitude. Phase term represents the angle at which the peak occurs (time represented in radians). The amplitude corresponds to the amount of the NDVI change (Dubovyk, Landmann, Dietz, & Menz, 2016; Jakubauskas, Legates, & Kastens, 2002).

![Figure 4. Methodological workflow.](image-url)
where NDVI<sub>t</sub> represents the reconstructed series, \( A \) is the amplitude and \( \varphi \) is the phase of the \( n \)th harmonic term, \( f \) is the frequency and \( t \) represents the time. \( \beta_0 \) can be viewed as the coefficient at zero frequency (intercept), which is the average of the series. Phase and amplitude terms were further used to extract the decision rules. The phenometrics were calculated for each pixel-wise time series (Table 1). As the main variance in NDVI series is captured in first harmonic terms and their additive component, we have tested our approach based on these terms. The first harmonic term represents the annual cycle, while the second harmonic term captures semiannual variation (Jakubauskas et al., 2002). We filtered the Landsat time series based on the crop calendar, which means that for winter crops (wheat, rapeseed) it was essential to use the observations from previous fall to have correct fitting. In situ observations of crop phenology were used to check their agreement with remotely sensed seasonality metrics by the comparison of key phenological dates.

Based on the harmonic regression fit, we identified the essential temporal intervals for the derivation of the Landsat composites. Using statistical aggregates such as mean, median and maximum values during each phenological stage, we generated the image composites (Table 1). Although all pixels in such composites were not acquired during the same day, they still represent the same phenological state of the crop. Thus, such composites from different periods of the growing season (start, mid-season, end of the season) can be further used for differentiation of crops. “Start”, “mid-season” and “end of the season” terms correspond to period 1, period 2 and period 3, respectively and through the manuscript are used interchangeably.

Afterwards, we used these composites, derivatives from harmonics and the reference phenological information to determine the range of the values of the peak of vegetation index and the timing of the peak, the range values of base and mid-season NDVI values. In this way, the decision rule set was developed to separate crop types. Based on these decision rules, we generated the masks which we labeled as “probable” class for each crop (Figure 5). The winter cereals (wheat, barely) were aggregated in one class due to their spectral similarities. To reduce the effect of mixed pixels in the generated crop masks, they were filtered one more time. The final training set was randomly selected each time from the crop masks. The selection of the sample points was based on the lowest standard deviation of NDVI values, where the “lowest” means those standard deviations that are within 95% confidence interval, derived from observed standard deviations. The number of generated training samples was proportional to the validation set. The values that were out of the range for the main crop classes were used to generate a class labeled “other”.

We run image segmentation to generate field boundaries and identify only cropland areas in the study area. Elimination of non-cropped areas and generation of field boundaries was performed by the intersecting several segmentation results based on both Landsat and Sentinel-1 images. The elimination of non-cropped areas was possible when using the distinctive characteristics of harmonic derivatives of different land cover classes such as masking the areas with low yearly average (zero harmonic component) for bare land and high values for the forest. Another assumption used for the delineation of cropland areas was the higher amplitude compared to non-cropped areas (Figure 6). In the map, we can clearly differentiate non-cropped areas, illustrated in green, and cropped areas, in the shades of purple. The visible differences in the cropland area are due to characteristic distinctions in temporal metrics of different crop types.

**Classification**

The second step in the classification procedure was the selection of the features that can improve the class separability and the training of the classifier. The main input features for the classifier were four spectral bands of Landsat, NDVI, derived from three temporal composites, and Sentinel-1 monthly composites (Table 1). We tested three machine learning algorithms for pixel-based classifications RF, SVM and one DF approach.

The first classification algorithm employed was the RF which is an ensemble machine learning technique that combines multiple trees (Breiman et al., 2001). One of the advantages of the RF is that the increase in the number of trees does not result in over-fitting (Hao, Zhan, Wang, Niu, & Shakir, 2015a; Ozdarici-Ok, Akar,
We have selected the optimal number of trees in the ensemble as 500. For each scenario of classification, the square root of the number of features was used for generation of the trees at each node.

SVM was the second method applied for classification, which uses higher dimensional features space for a class separation (Foody, 2004b). SVM is a non-parametric method, and the parametric distribution of the input data is not required. SVM is reported to work well with small training datasets (Mathur & Foody, 2008). For the SVM test, we used radial basis function kernel.

The third classification approach was based on DF by combining the results of several classifier algorithms (Kittler, Hatef, Duin, & Matas, 1998), similar to the method proposed by Waske and Benediktsson (2007). Two classifier algorithms (here: RF and SVM) were trained based on randomly selected feature subsets. While Waske, Van Der Linden, Benediktsson, Rabe and Hostert (2010) recommended 20–30% of all input features to get accurate results, such a low number of features decreased classification accuracy in our study. Hence, the subsets were created by
randomly drawing 75% of all input variables. This process was repeated 50 times, based on recommendation of Waske et al. (2010) and initial tests in our study. Through this procedure, we created 100 outputs (50 by the RF and 50 by the SVM) which were 100 land cover maps. These 100 maps were then combined using a majority vote (Löw, Conrad, & Michel, 2015), and as a result, one final map was the output from the DF (Figure 7).

### Accuracy assessment

The accuracy of the initial crop masks and final crop maps was assessed based on confusion matrices (Congalton, 1991), calculated with the independent validation set collected during the field visits (Table 2). From confusion matrices, we derived overall accuracy (OA), as well as producer’s (PA) and user’s (UA) accuracy. We used McNemar’s test (Foody, 2004a) for the comparison of the classification results. Further, we calculated the area of each crop class and compared it with the official agricultural census statistics (State Statistical Department of Kiev region, 2015, 2016). For the spatially explicit representation of uncertainty of crop maps, we divided the training data into multiple overlapping collections by randomly selecting each time 75% of data. Afterwards we trained the classifier with each subset, classified the input with each classifier. After generation of the classification results, we have normalized the values of pixels in each array to a common scale and estimated the variance of classification for each pixel. All the analyses were carried out with the use of R (R Core Team, 2016) and GEE (Google Inc., 2016).

### Results and discussion

According to the approach described in the previous section, crop maps were produced using one of the supervised algorithms (RF, SVM and DF) and different input time series derived by single Landsat or Sentinel-1 sensors and their combination. The study site was characterized with different agricultural landscapes and its crop variability therein (Figure 8). As SVM resulted in accuracies lower than 75%, we did not consider it for further analysis.

Crop maps indicated that the general patterns of crop distribution were similar when using different classification methods. An exception was fields in the southwestern region of the study area, which were classified in 2015 as “winter cereal” when using DF, whereas with RF those fields were assigned to the class “other”. Based on a visual inspection of the generated

---

**Table 2. The samples used for training and validation.**

| Class            | 2015 Training | 2015 Validation | 2016 Training | 2016 Validation |
|------------------|---------------|-----------------|---------------|-----------------|
| Winter cereals   | 19            | 19              | 20            | 23              |
| Winter rapeseed  | 7             | 7               | 5             | 4               |
| Maize            | 18            | 20              | 20            | 19              |
| Sunflower        | 7             | 8               | 13            | 13              |
| Soy              | 20            | 21              | 19            | 19              |
| Other            | 10            | 6               | 7             | 6               |

---

**Figure 7.** The workflow for decision fusion approach.
crop maps, we found out that several fields of the same crop were in the direct neighborhood or in the close distance. In general, we can observe that the output using different input data (Landsat or Sentinel-1 sensors plus their combination) was similar for several fields. Particularly, this is visible for “winter cereal” class, where the area in the disagreement class comprised 4–14% of the classified “winter cereal” area. The explanation for this could be the fact that winter cereals have distinctive temporal characteristics which can be detected in both SAR and optical temporal composites.

The disagreement was higher for summer crops, which can be explained by their spectral similarities and the allocation of these fields to different classes during several classification tests (Figure 8, class “disagreement among different scenarios” represented in red). For instance, soy had higher portion of pixels with higher disagreement among the classifiers, reaching up to 33% with DF and 41% with RF in 2016. Furthermore, looking at the accuracy metrics (Table 3), we can observe, that with the elimination of the input variables, PA and UA for crops such as soy decrease respectively. The reason for this is that with less data it is harder to distinguish the summer crops which have similar development throughout a growing season.

We compared the results when using three seasonal composites from Landsat 8 and Sentinel-1 using McNemar’s test. The difference between the classification results of RF and DF was statistically significant in both years (for 2015 $\chi^2$: 4.5; $p = 0.034$ and for 2016 $\chi^2$: 5; $p = 0.025$). Similar results were observed in the study of Löw et al. (2015) and Hao, Wang and Niu (2015b) who report that the fusion approach had better results.

Among the classification algorithms tested, DF performed the best in terms of OA using both Landsat and Sentinel as input data (88% in 2015 and 85% in 2016) and resulted in more accurate crop maps with an increase in OA of 5.7–6.6% over the RF classification’s accuracy. DF yielded acceptable accuracies (>80%) when using optical and SAR data separately as well. The only exclusion was the SAR-based classification of 2016, where the OA was 78% which was close to the desired accuracy metric. This can be due to the fact that the map produced based on DF was a product of several classification results based on combination of the classifiers. In addition to the traditional accuracy

Figure 8. Crop maps for 2015 obtained using (a) random forest and (b) decision fusion, and 2016 obtained using (c) random forest (d) decision fusion with the input data from Landsat 8 and Sentinel-1 composites from start, mid and end of the growing season.
Table 3. Accuracy measures of area dominant crops using different classification approaches and data (The results are based on classification tests using Landsat 8 composites from start, mid and end of the growing season and Sentinel-1 monthly composites from April to October).

| Time  | OA  | PA  | UA  | OA  | PA  | UA  | OA  | PA  | UA  | OA  | PA  | UA  | OA  | PA  | UA  |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2015  | 88  | 86  | 100 | 82  | 95  | 94  | 80  | 90  | 94  | 70  | 95  | 93  | 66  |
| 2016  | 85  | 90  | 86  | 90  | 100 | 93  | 71  | 82  | 86  | 82  | 90  | 95  | 100 | 71  |

Figure 9. The uncertainty of classification results in using (a) random forest (b) decision fusion and 2016 obtained using (c) random forest (d) decision fusion with the input data from Landsat 8 and Sentinel-1 composites from start, mid and end of the growing season.
assessment methods, we have also visualized pixel-based uncertainty of classifications (Figure 9). In general, the variance was higher when using RF, whereas with DF we can observe good agreement of class allocation among multiple classification results.

We have evaluated the efficiency of training sample generation by assessment of the accuracies of the "possible crop masks" with the independent validation data. For both years, the accuracies were high for winter crops (80%), whereas lower accuracies were found for soybean and maize (~60%). This leads us to infer that the masks should be further filtered and only the samples which exhibited low standard deviation were selected for final classification.

Figure 10 illustrates the temporal profiles of main crops derived from the generated training set. Some classes indicate a high variability of NDVI signal, although training set was reduced to the samples with the lowest standard deviation.

Overall, results indicate that the temporal patterns of crops can be modeled with harmonic function especially when using the first and second components. For summer crops, the main variation was derived using only first harmonic component as they exhibited a strong unimodal behavior. In contrast, to derive the best fit for winter crops, we used the function generated from the first and second harmonic components. These crops are usually planted in fall of the previous year and due to their distinct phenological development, the second component is characterized with substantial variability.

Per-class separability

In general, the winter cereal class was most accurately classified, whereas we observed misclassifications with maize and soybean classes. This could be explained by the fact that winter cereals exhibit distinctive development stages and early growth, which make them easily separable from spring crops. In the case of spring/summer crops, we observed misclassifications between maize and soybean in 2015. In 2016, sunflower was misclassified with maize. Soybean also exhibited slightly lower accuracy and was misclassified together with maize and sunflower. Relatively poor classification results for soybean (both accuracies (Table 3) and the areal estimates

Figure 10. Temporal profiles of the main crops. (a) Sunflower, (b) maize, (c) winter cereals, (d) winter rapeseed, (e) soybean (error bars are illustrating the standard deviations and the temporal windows show the phenological phases).
(Figure 11) can be explained by confusion with other summer crops.

The comparison of the sown areas derived from classification results and the official agricultural census statistics published by the State Statistical department of Kiev region (Figure 11) showed that the area estimated for each class was in acceptable agreement with reference agricultural census statistics and the deviation for four crops was ranging from 7% to 35%. We observed large error estimates among spring crops which can be accounted to misclassifications between them. We observed the lowest accuracies in the class other and winter rapeseed. This could be the result of high heterogeneity of the class as it includes several summer crops, which have different spectral response. Another reason for this could be the relatively small number of training samples compared to the rest of classified crop classes. The estimated area for winter rapeseed was significantly larger than recorded in the official statistics.

**Evaluation of input features**

To check the effect of selected input features for the classification, we applied a backward elimination approach, and estimated the decrease of prediction accuracy due to the exclusion of some input variables. The most successful classifications were conducted using Landsat bands along with NDVI and monthly backscatter composites. The integration of both Landsat features (spectral bands, NDVI) and Sentinel-1-based composites at multiple time steps is advantageous. Overall, these results are in line with several studies (Fontanelli et al., 2014; Kussul et al., 2015b) where the integration of optical and SAR data improved crop identification. These results demonstrate the applicability of Landsat time series and fitted harmonic function for the derivation of phenological features. When adding Sentinel-1 composites, the cropped areas were quantified with higher accuracies. The integration of multispectral and SAR data improved the classification accuracy by 2–5%. The increase in class-specific accuracies was observed in both studied years for winter cereals. In addition, there was an improvement for the separation of spring/summer crops (Table 3).

We have also noted the increase in OAs when using several temporal composites. Despite this, with mono-temporal analysis, acceptable results can be achieved, if using mid-season composites (Period 2). Mono-temporal classifications from early season and end of the growing season resulted in significantly lower accuracies. To achieve high separability among crops, the use of imagery from at least two periods within the crop growing season is required (Figure 12).

The use of temporal composites was justified, as when averaging different images, we decreased the contribution of speckle which resulted in within-class variance decrease and easy separation of the
crops. In case of optical imagery, composting enabled the creation of cloud-free images covering the study area.

It is worth mentioning, that the use of GEE profoundly decreased the time for data access and analysis, as it was possible to use the satellite time series from repository and run computationally intensive pixel-wise analysis using Google’s distributed power.

Considerations and outlook

The mapping of different crop types based on temporal variation is a demanding and complex process. First, several factors influence the process of crop identification. Individual fields can be managed with different seeding and harvesting times for the same crop species, which in turn has an impact on spectral signatures. Second, different crops can have analogical stages of development throughout a growing season. This can introduce a bias to the selection of training samples and classification. The timing and frequency of image acquisition influence the accuracy of the derived sample set, as it has a significant effect on the fitting accuracy and subsequent estimation of phenological parameters (Eklundh & Jönsson, 2016; Kuenzer, Dech, & Wagner, 2015; Wardlow & Egbert, 2008). An important aspect of this approach is that crop development throughout the growing season highly depends on several factors such as temperature, the amount and temporal distribution of precipitation that can vary from year to year. This variation plays a key role in the development of the decision rules of several years. One solution for this was making the range of values in decision rule set quite wide, but with the consideration that too large range can lead to lower classification accuracies. Another solution was creation of the composites and fitting the harmonic function, which removed sharp differences and noise in time series due to atmospheric conditions and cloud cover. Although the development of the decision rule set was straightforward for winter crops due to distinctive spatio-temporal characteristics, the identification of summer crops was challenging due to spectral similarities of these crops. For spatial upscaling and transferring of this approach to other areas, the variation in phenological development from one region to another should be considered. For this, in situ data of phenology, the knowledge of crop development and the integration of the weather data may lead to accurate results for large area classification.

Conclusions

In the presented paper, we discussed the delineation of main crop types in the central region of Ukraine using multisource remote sensing data. We have proposed and evaluated a two-step approach that constitutes of training sample generation based on harmonic regression, RF and DF classification. The approach developed and tested herein has the potential for cases especially when field data are insufficient. We have also demonstrated the potential use of optical and SAR data for the rule-based training sample generation and subsequent crop classification in complex cropping systems. The derivation of key phenological and temporal metrics enabled the creation of image composites that allow mapping the crop types accurately. In addition, we demonstrated that the harmonic fit and derived metrics can be very useful for representation of the seasonal variations of crops.

The obtained results confirmed the feasibility of the herein developed two-step approach for crop mapping in the study area in Ukraine. OA exceeded 80% when seasonal composites were used. Among the crop species, the class of winter cereals was the most accurately identified, while we observed misclassifications between soybean and maize. We showed that class separability is depending on the selected features and their temporal distribution of the input data. The integration of multispectral and SAR data improved the classification accuracy.

The method was tested for 2 years which enabled us to study the effect of interannual meteorological differences on crop phenological development that in turn impacts crop classification results. Based on our results, we recommend the use of seasonal composites in a two-step approach of classification to create accurate crop maps over several years. Overall, our remotely sensed crop patterns were confirmed by quantitative (via accuracy assessment and comparison with state crop statistical data) and qualitative tests including agroecological knowledge of local experts. Therefore, we concluded that the proposed approach could be implemented in the areas where field observations are not available or scarce.

Acknowledgments

This study was conducted in the scope of EvI DENz project which is supported by the German Federal Ministry for Economic Affairs and Energy (BMWi; grant No. 50EE1541). The field research in 2016 was supported by DAAD. The authors would like to thank GEE community for the support and shared knowledge. We are thankful to Guido Lüchters and Valerie Graw for their help with the analyses. The authors greatly acknowledge the Space Research Institute of National Academy of Sciences of Ukraine and Tatyana Adamenko for the support during the field work.

Disclosure statement

No potential conflict of interest was reported by the authors.
Funding
This study was conducted in the scope of EvIDENz project which is supported by the German Federal Ministry for Economic Affairs and Energy [BMWi; grant No. 50EE1541]. The field research in 2016 was funded by DAAD.

ORCID
Andrii Kolotii @ http://orcid.org/0000-0002-6972-4483

References
Bargiel, D. (2017). A new method for crop classification combining time series of radar images and crop phenology information. Remote Sensing of Environment, 198, 369–383.
Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32.
Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37(1), 35–46.
Davidson, A.M., Fisette, T., Menairn, H., & Daneshfar, B. (2017). Detailed crop mapping using remote sensing data (crop data layers). In J. Delince (Ed.), Handbook on remote sensing for agricultural statistics (chapter 4). Rome: Handbook of the Global Strategy to improve Agricultural and Rural Statistics (GSARS).
Dubovyk, O., Landmann, T., Dietz, A., & Menz, G. (2016). Quantifying the impacts of environmental factors on vegetation dynamics over climatic and management gradients of Central Asia. Remote Sensing, 8(7), 600.
Eklundh, L., & Jönsson, P. (2016). TIMESAT for processing time-series data from satellite sensors for land surface monitoring. In Multitemporal remote sensing: Methods and Applications, edited by Y. Ban. Remote Sensing and Digital Image Processing 20 (pp. 177–194). Cham: Springer. doi:10.1007/978-3-319-47037-5_9
Fontanelli, G., Crema, A., Azar, R., Stroppiana, D., Villa, P., & Boschetti, M. (2014). Agricultural crop mapping using optical and SAR multi-temporal seasonal data: A case study in Lombardy region, Italy. In 2014 IEEE Geoscience and Remote Sensing Symposium (pp. 1489–1492).
Foody, G.M. (2004a). Thematic map comparison. Photogrammetric Engineering & Remote Sensing, 70 (5), 627–633.
Foody, G.M. (2004b). A relative evaluation of multiclass image classification by support vector machines. Geoscience and Remote Sensing, 42(6), 1335–1343.
Forfuor, G., Conrad, C., Thiel, M., Ullmann, T., & Zoungarra, E. (2014). Integration of optical and synthetic aperture radar imagery for improving crop mapping in Northwestern Benin, West Africa. Remote Sensing, 6(7), 6472–6499.
Google Inc. (2016). Earth engine documentation. Retrieved 12 14, 2016 from https://sites.google.com/site/earthengineapidocs/
Hao, P., Wang, L., & Niu, Z. (2015b). Comparison of hybrid classifiers for crop classification using normalized difference vegetation index time series: A case study for major crops in North Xinjiang, China. PLoS ONE, 10, 9.
Hao, P., Wang, L., Zhan, Y., & Niu, Z. (2016). Using moderate-resolution temporal NDVI profiles for high-resolution crop mapping in years of absent ground reference data: A case study of Bole and Manas counties in Xinjiang, China. ISPRS International Journal of Geo-Information, 5(5), 67.
Hao, P., Zhan, Y., Wang, L., Niu, Z., & Shakir, M. (2015a). Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. Remote Sensing, 7(5), 5347–5369.
Hentze, K., Thonfeld, F., & Menz, G. (2016). Evaluating crop area mapping from MODIS time-series as an assessment tool for Zimbabwe’s “fast track land reform programme”. Plos One, 11(6), e0156630.
Ingliada, J., Vincent, A., Arias, M., & Marais-Sicre, C. (2016). Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series. Remote Sensing, 8(5), 362.
Jakubauskas, M.E., Legates, D.R., & Kastens, J.H. (2002). Crop identification using harmonic analysis of time-series AVHRR NDVI data. Computers and Electronics in Agriculture, 37(1–3), 127–139.
Joshi, N., Baumann, M., Ehhammer, A., Fensholt, R., Grogan, K., Hostert, P., … Waske, B. (2016). A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sensing, 8(1), 70.
Kittler, J., Hatif, M., Duin, R.P.W., & Matas, J. (1998). On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell., 20, 226–239.
Kogan, F., Kussul, N., Adamenko, T., Skakun, S., Kravchenko, O., Kryvobok, O., & Lavrenyuk, A. (2013). Winter wheat yield forecast in Ukraine based on Earth observation, meteorological data and biophysical models. International Journal of Applied Earth Observation and Geoinformation, 23, 192–203.
Kuemmerle, T., Olofsson, F., Chaskovskiy, O., Baumann, M., Ostatowicz, K., Woodcock, C.E., & Radehoff, V.C. (2011). Post-Soviet farmland abandonment, forest recovery, and carbon sequestration in western Ukraine. Global Change Biology, 17(3), 1335–1349.
Kuenzer, C., Dech, S., & Wagner, W. (2015). Remote sensing time series: Revealing land surface dynamics. Springer, Cham.
Kussul, N., Lemoine, G., Gallego, J., Skakun, S., & Lavrenyuk, M. (2015a). Parcel based classification for agricultural mapping and monitoring using multi-temporal satellite image sequences. IEEE, 165–168. doi:10.1109/IGARSS.2015.7325725
Kussul, N., Skakun, S., Shelestov, A., Lavrenyuk, M., Yailymov, B., & Kussul, O. (2015b, May). Regional scale crop mapping using multi-temporal satellite imagery. In G. Schreier, P.E. Skrovseth, & H. Staudenrausch (Eds.) ISPRS - International archives of the photogrammetry, remote sensing and spatial information sciences, XL-7/W3 (pp. 45–52). Copernicus Publications, Göttingen.
Lessel, J., & Ceccato, P. (2016). Creating a basic customizable framework for crop detection using Landsat imagery. International Journal of Remote Sensing, 37(24), 6097–6107.
Liu, J., Heiskanen, J., Aynekulu, E., Maeda, E.E., & Pellikka, P.K.E. (2016). Land cover characterization in West Sudanian savannas using seasonal features from annual Landsat time series. Remote Sensing, 8(5), 365.
Loosvelt, L., Peters, J., & Skriver, H. (2012). Impact of reducing polarimetric SAR input on the uncertainty of crop classifications based on the random forests algorithm. IEEE Transactions on Geoscience and Remote Sensing, 50, 4185–4200.
Lopes, A., Nezry, E., Touzi, R., & Laur, H. (1990). Maximum a posteriori speckle filtering and first order texture models in SAR images. In 10th Annual International Symposium on
