

SPATIAL-TEMPORAL CONDITIONAL RANDOM FIELDS CROP CLASSIFICATION FROM TERRASAR-X IMAGES

B. K. Kenduiwo^{a,*}, D. Bargiel^a, U. Soergel^a

^a Institute of Geodesy, Technische Universität Darmstadt, Germany - (kenduiwo, bargiel, soergel)@geod.tu-darmstadt.de

Commission III, WG III/7

KEY WORDS: Conditional Random Fields (CRF), phenology, conditional probability, spatial-temporal

ABSTRACT:

The rapid increase in population in the world has propelled pressure on arable land. Consequently, the food basket has continuously declined while global demand for food has grown twofold. There is need to monitor and update agriculture land-cover to support food security measures. This study develops a spatial-temporal approach using conditional random fields (CRF) to classify co-registered images acquired in two epochs. We adopt random forest (RF) as CRF association potential and introduce a temporal potential for mutual crop phenology information exchange between spatially corresponding sites in two epochs. An important component of temporal potential is a transitional matrix that bears intra- and inter-class changes between considered epochs. Conventionally, one matrix has been used in the entire image thereby enforcing stationary transition probabilities in all sites. We introduce a site dependent transition matrix to incorporate phenology information from images. In our study, images are acquired within a vegetation season, thus perceived spectral changes are due to crop phenology. To exploit this phenomena, we develop a novel approach to determine site-wise transition matrix using conditional probabilities computed from two corresponding temporal sites. Conditional probability determines transitions between classes in different epochs and thus we used it to propagate crop phenology information. Classification results show that our approach improved crop discrimination in all epochs compared to state-of-the-art mono-temporal approaches (RF and CRF mono-temporal) and existing multi-temporal markov random fields approach by Liu et al. (2008).

1. Introduction

To monitor and estimate food production, up-to-date precise crop spatial information is required. Earth observing satellites have undergone improved spatial, spectral, and temporal resolutions. Changes in a scene can be monitored regularly and on demand. Such trend favours development of novel image classification methods that can handle temporal data (Jianya et al., 2008). This is especially true for radar sensors which overcome limitations of optical sensors: their signals can penetrate clouds and are independent of daylight (Gomez-Chova et al., 2006; Tupin, 2010). Incorporating crop growing degree day information into multi-temporal radar images from TerraSAR-X is likely to improve crop classification. Thus, Synthetic Aperture Radar (SAR) temporal data can be used to benefit crop classification given novel spatial-temporal context methods.

Use of context (spatial and temporal) in image segmentation and classification has recently gained popularity. Spatial context accounts for similarities among pixels in regard to distance from each other. It determines probability of a pixel or a group of pixels occurring at a given location based on nature of its neighbourhood (Tso and Mather, 2009). Goodchild (1992) defines spatial context as "the propensity for nearby locations to influence each other and to possess similar attributes." In contrast, temporal context defines spectral similarities of pixels with respect to different acquisition times. Use of images acquired at different times has shown significant improvement in classification e.g. in classification of crops and vegetation (Lu and Weng, 2007). However, complexity accompanying multi-temporal data requires approaches that can effectively integrate spatial and temporal data. More also, continuous increase in high temporal resolution satellites has led to a "Tsunami" of data in archives. Therefore, spatial-temporal automated classification methods are nec-

essary to bridge the gap between expensive data acquisition efforts and actual beneficial data consumption.

Markov Random Fields (MRF) (Geman and Geman, 1984) have widely been used to integrate spatial-temporal context in image classification. Introduction of Bayesian concept by Swain (1978) to classification of multi-temporal images motivated several MRF temporal studies. Examples include: a MRF approach unidirectionally passing temporal information from a classified image at a given date to a subsequent image of the same area at a later date by Jeon and Landgrebe (1992); Solberg et al. (1996) later extended in (Melgani and Serpico, 2003) to allow bidirectional exchange of temporal information. Liu et al. (2006) uses temporal correlation and temporal exclusion to control certain changes in forest disease spread monitoring. In (Moser and Serpico, 2011) MRF is applied for multi-scale multi-temporal high resolution image classification with transitional matrix determined using Expectation Maximization (EM) algorithm. These studies, except (Moser and Serpico, 2011), used one generalized class transition matrix in MRF determined heuristically. Leite et al. (2011) used a combination of expert knowledge and training data. In this approach, the matrix globally assumes stationary class transitions over all pixels neglecting changes that may exist in the image (Liu et al., 2008). In addition, MRF's assumption of conditional independence in observed data adopted for computational tractability neglect spatial context inherence in images (Lafferty et al., 2001; Kumar, 2006; Zhong and Wang, 2007a; Parikh and Batra, 2008). Remotely sensed images exhibit a coherent scene because neighbouring sites are spatially correlated. This concept is modelled by Conditional Random Fields (CRF) (Lafferty et al., 2001) introduced for one-dimensional text classification and extended to two-dimensional image classification (Kumar, 2006). The Framework provided by CRF integrates spatial context both in class labels and data.

Conditional random fields ability to model context in class la-

*Corresponding author.

bels and data has triggered studies in several applications. So far CRF has been used for: classification of settlements and urban areas (Zhong and Wang, 2007a,b; Hoberg and Rottensteiner, 2010; Niemeyer et al., 2011, 2013; Kenduywo et al., 2014), estimation of ground heights from LiDAR data (Lu et al., 2009), building extraction (He et al., 2008; Wegner et al., 2011b,a) and interpretation of terrestrial images (He et al., 2004; Korc and Förstner, 2008; Gould et al., 2008). However, these are mono-temporal CRF studies. Despite the benefits of multi-temporal images, a few CRF studies exist: crop type classification using RapidEye images by (Hoberg and Müller, 2011), land-cover classification from IKONOS and RapidEye images (Hoberg et al., 2010) and multi-scale multi-temporal study using IKONOS and Landsat images (Hoberg et al., 2011). The studies incorporate temporal context by passing temporal information through empirically determined global transition probability matrix. However, the transition matrix does not optimally represent all site changes between epochs. Hoberg et al. (2015) notes that incorrect determination of transition matrix leads to erroneous transfer of information into other epochs subsequently reducing classification accuracy. Developing an approach of determining transition probabilities for each site can minimize such errors. More also, according to our knowledge, no studies have used multi-temporal radar data for CRF classification of crops. We develop a spatial-temporal CRF classification approach exploiting site-wise crop phenological transitions from multi-temporal TerraSAR-X images.

Multi-temporal SAR images of a given season will improve crop classification. Crops show varied backscatter radar signal in time (at different epochs¹). We intend to maximize feature separation by exploiting this temporal crop phenology. For instance, crops that may not be resolved in one epoch spectrally can be resolved in another. Notably crops undergo phenological changes between different epochs resulting in varied spectral properties (Bargiel et al., 2010). Integration of spatial-temporal crop phenology information from different epochs of a vegetation season using site-wise transition matrix between a pair of epochs in temporal potential of CRF is the contribution of this study. In our study, Terra-SAR images are acquired at different phenological stages of a crop season. Therefore, we determine the site-wise transition matrix using Bayes' theorem of conditional probability. The conditional probability – representing class transitions in the matrix – are computed from a pair of class probability vectors estimated by random forest classifier at corresponding sites in different epochs. We derive this approach from a MRF forest change detection study using optical images in (Liu et al., 2008) and extend it to this study. So far no studies have considered incorporating phenology information into CRF classification using SAR images in this manner.

The rest of the paper is organized as follows. Section 2 introduces the approach we adopted and discusses how CRF is designed using site-wise conditional probabilities for bi-directional temporal information flow. Section 3.4 describes classes (crop types), data and features used, details of experiments conducted by our approach and other state-of-the-art methods and results obtained. In Section 4 a discussion of results from our approach compared to state-of-the-art is made leading to conclusions in Section 5.

2. Methods

2.1 Conditional random fields

A supervised image classification assigns class labels to image sites given user defined examples known as *training sites*. Train-

¹An epoch in this case corresponds to a particular image acquisition date within a vegetation season.

ing sites, represented by a vector of features (numeric attributes computed from user defined image sites), and corresponding class labels serve as an input to an algorithm that infers labels of all other image sites. Mono-temporal CRF classification aims to estimate an optimal label configuration $\hat{\mathbf{c}}$ of a vector of class labels $\mathbf{c} = (c_1, c_2, \dots, c_m)^T$ where m is number of image sites, from image data \mathbf{x} , i.e. $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$, by maximizing posterior probability $P(\mathbf{c}|\mathbf{x})$ thus $\hat{\mathbf{c}} = \arg \max_{\mathbf{c}} P(\mathbf{c}|\mathbf{x})$. A mono-temporal CRF models $P(\mathbf{c}|\mathbf{x})$ with a graph structure in which nodes are linked to image sites and edges bear relationships between a pair of adjacent sites. Thus, $P(\mathbf{c}|\mathbf{x})$ is modelled as CRF as:

$$P(\mathbf{c}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp \left\{ \sum_{i \in S} A(c_i, \mathbf{x}) + \sum_{i \in S} \sum_{j \in N} I(c_i, c_j, \mathbf{x}) \right\} \quad (1)$$

where A denotes association potential, I is spatial interaction potential, S is a set of all image sites, i is a site in the image, j is a neighbour of site i , N is a set of neighbours of i and $Z(\mathbf{x})$ is a normalizing constant called partition function.

2.2 Spatial-temporal CRF: Problem formulation

In spatial-temporal contextual image classification, image sites are dependent random variables that form a random field (region) with correlated spatial and temporal neighbours. For computational tractability the spatial and temporal random fields are factorized into neighbourhood systems (Li, 2009).

For spatial-temporal classification, we consider p co-registered images with the same spatial resolution and extent acquired respectively at epoch $t \in T$ such that $T = \{t_0, t_1 \dots, t_p\}$. Assume $\mathbf{c}^{t_0} = \{c_1^{t_0}, \dots, c_m^{t_0}\}$ and $\mathbf{c}^{t_1} = \{c_1^{t_1}, \dots, c_n^{t_1}\}$ where m and n represent number of image sites, are a set of random labels over a set of multispectral image features \mathbf{x}^{t_0} and \mathbf{x}^{t_1} corresponding to epochs t_0 and t_1 respectively. The set of classes in any epoch can differ in both composition and number. For instance, in multi-scale classification different number of land-cover classes can be defined subject to image resolution. Normally the image of high resolution may contain sub-classes of a class in the lower resolution image. In such a case the transition matrix would be rectangular. In our case, the image resolution and number of classes is the same within a season and thus the transition matrix is square. To estimate $P(\mathbf{c}|\mathbf{x})$ in spatial-temporal classification we extend Equation (1) as:

$$P(\mathbf{c}|\mathbf{x}) = \frac{1}{Z(\mathbf{x}^{t_0})} \exp \left\{ \sum_{i \in S} A(c_i^{t_0}, \mathbf{x}^{t_0}) + \lambda_1 \sum_{i \in S} \sum_{j \in N} I(c_i^{t_0}, c_j^{t_0}, \mathbf{x}^{t_0}) + \lambda_2 \sum_{i \in S} \sum_{t \in T} \sum_{k \in K} TP(c_i^{t_0}, c_k^{t_1}, \mathbf{x}^{t_0}, \mathbf{x}^{t_1}) \right\} \quad (2)$$

where TP is temporal interaction potential, K is a set of image sites in epoch t_1 that are temporal neighbours of site i , i.e. i is a spatial neighbour to j and a temporal neighbour to k in epoch t_1 (i and k correspond to the same ground location). Parameters λ_1 and λ_2 represents weights used to regulate spatial and temporal information respectively. Equation (2) provides a general spatial-temporal CRF classification framework (Hoberg et al., 2015). However, our TP is developed on site-wise transition matrix computed using initial class probability estimates at each site from training sites and data as opposed to an empirical

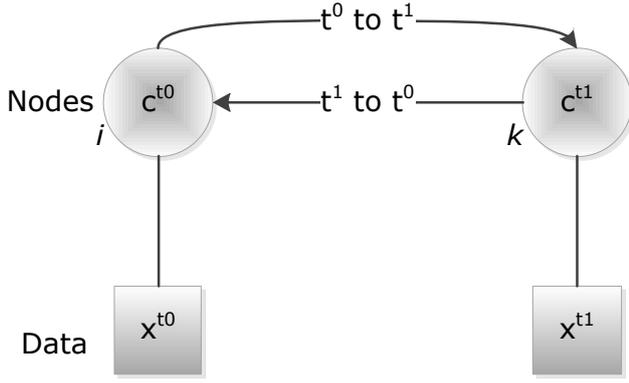


Figure 1. Designed CRF graph structure; node links are edges.

stationary transition matrix by Hoberg et al. (2015) as explained in subsequent sections.

Spatial-temporal CRF models posterior probability in Equation (2) with a graph structure where nodes are connected to image sites by A , I represent edges bearing relationships of adjacent sites i and j in one epoch, and an additional potential, TP , representing edges containing relationships between temporal sites i and k in different epochs as shown in Figure 1. In CRF framework, A , I and TP can be regraded as arbitrary local classifiers. This property enables use of domain-specific discriminative classifiers in structured data rather than restricting the potentials to a certain form (Zhong and Wang, 2007a,b). We omit weight notations λ_1 and λ_2 in subsequent sections for clarity in CRF potential equations.

2.3 Association Potential

It determines how likely an image site i takes a label c_i^{t0} in epoch $t0$ given the data \mathbf{x}^{t0} : $A(c_i^{t0}, \mathbf{x}^{t0}) = P(c_i^{t0} | \mathbf{f}_i(\mathbf{x}^{t0}))$, $\mathbf{f}_i(\mathbf{x}^{t0})$ is a site-wise feature vector (Kumar, 2006). We adopt random forest (RF) (Breiman, 2001) to determine A by independent classification of different epochs assuming class conditional independence in them. A RF conducts classification by casting votes from a number of decision trees D_T generated during training. If the number of votes cast for a given class c by RF is V_c , then our A is $P(c_i^{t0} = c | \mathbf{f}_i(\mathbf{x}^{t0})) = \frac{V_c}{D_T}$. We set $D_T = 200$ because at that value RF stabilizes (Hastie et al., 2011) and set tree depth as 25.

2.4 Interaction Potential

It measures the influence of data and neighbouring labels on site i in epoch $t0$. It ensures that site i , as initially determined by association potential, is labelled to its corresponding "true class" given data evidence \mathbf{x}^{t0} and neighbourhood dependency N where $j \in N$. This study models I using contrast sensitive Potts model designed based on Euclidean distance d_{ij} of adjacent node features \mathbf{f}_i and \mathbf{f}_j :

$$d_{ij} = \frac{\|\mathbf{f}_i(\mathbf{x}^{t0}) - \mathbf{f}_j(\mathbf{x}^{t0})\|}{\mathbf{R}} \quad (3)$$

where \mathbf{R} is the number of features. Then, model of I is:

$$I(c_i^{t0}, c_j^{t0}, \mathbf{x}^{t0}) = \begin{cases} \beta \cdot \exp(-\eta \cdot d_{ij}^2) \cdot h(c_i^{t0}, c_j^{t0}) & \text{if } c_i^{t0} = c_j^{t0} \\ h(c_i^{t0}, c_j^{t0}) & \text{if } c_i^{t0} \neq c_j^{t0} \end{cases} \quad (4)$$

where β is a spatial interaction parameter that regulates smoothness, parameter η regulates the contrast-sensitive term and $h(c_i^{t0}, c_j^{t0})$ is a histogram matrix count bearing co-occurrence of labels of neighbouring sites i and j . We normalize the histogram by row in order to minimize bias of dominant classes in training data. Therefore, the model is different from contrast sensitive Potts model because transitions of classes is now governed by their frequency in training data (Kosov et al., 2013).

2.5 Temporal interaction potential

It models interactions between data \mathbf{x}^{t0} and \mathbf{x}^{t1} and labels c_i^{t0}, c_k^{t1} of site i and k :

$$TP(c_i^{t0}, c_k^{t1}, \mathbf{x}^{t0}, \mathbf{x}^{t1}) = P(c_i^{t0}, c_k^{t1}, \mathbf{x}^{t0}, \mathbf{x}^{t1}) \quad (5)$$

We consider crop classification within one season. Consequently, spectral changes in classes at a given epoch are a result of phenology rather than transitions to other classes in another epoch. This is because a particular crop type is observed in the entire season from planting to harvest. We exploit the fact that crop phenology varies temporally and also spectrally to enhance crop discrimination. Spectral changes due to crop phenology is expressed by joint probability in Equation (5).

Transitional probabilities between similar and different classes can be represented in a transitional matrix T_{ik} . The matrix can be determined by expert opinion, empirically from existing data sources or computed (Liu et al., 2008). We compute T_{ik} for each site in the image using conditional probability computed from class membership probability vectors for node i and k estimated by RF for each site based on spectral observation in training sites. The matrix is then used to introduce a temporal directed edge between node i and k as illustrated in Figure 1.

2.5.1 Site-wise transition probability matrix Conditional probability expresses the likelihood of a class label to take up a site i in epoch $t0$ given class label information from a site k in epoch $t1$. It represents intra- and inter-class transitions. Consider a set of classes a and b such that $\alpha \in \alpha_1, \alpha_2, \dots, \alpha_a$ and $\omega \in \omega_1, \omega_2, \dots, \omega_b$ in epoch $t0$ and $t1$ respectively, using Bayes' theory class transitions from $t0$ to $t1$, i.e. $t0 \Rightarrow t1$, can be expressed as:

$$P(c_k^{t1} = \omega | c_i^{t0} = \alpha) = \frac{P(c_i^{t0} = \alpha | c_k^{t1} = \omega) P(c_k^{t1} = \omega)}{P(c_i^{t0} = \alpha)} \quad (6)$$

where prior probabilities $P(c_i^{t0} = \alpha)$ and $P(c_k^{t1} = \omega)$ are marginal distributions in Table 1. From rules of probability, sum rule

$$P(c_i^{t0} = \alpha) = \sum_{\omega} P(c_i^{t0} = \alpha, c_k^{t1} = \omega) \quad (7)$$

and product rule,

$$P(c_i^{t0} = \alpha, c_k^{t1} = \omega) = P(c_i^{t0} = \alpha | c_k^{t1} = \omega) P(c_k^{t1} = \omega) \quad (8)$$

Equation (6) can be transformed into:

$$P(c_k^{t1} = \omega | c_i^{t0} = \alpha) = \frac{P(c_i^{t0} = \alpha, c_k^{t1} = \omega)}{\sum_{\omega} P(c_i^{t0} = \alpha, c_k^{t1} = \omega)} \quad (9)$$

Similarly following Bayes' theory transitions in the reverse direction, i.e. $t1 \Rightarrow t0$, are computed as:

$$P(c_i^{t0} = \alpha | c_k^{t1} = \omega) = \frac{P(c_i^{t0} = \alpha, c_k^{t1} = \omega)}{\sum_{\alpha} P(c_i^{t0} = \alpha, c_k^{t1} = \omega)} \quad (10)$$

Therefore, from Equations (9) and (10) we compute site-wise transition probability matrix by dividing the $a \times b$ joint probability matrix of joint distributions in Table 1 by row sum and column sum (marginal distributions) for crop phenology transitions in $t_0 \Rightarrow t_1$ and $t_1 \Rightarrow t_0$ respectively.

	3	ω_1	ω_2	...	ω_b	Sum
2	ω_1	$P(\alpha_1, \omega_1)$	$P(\alpha_1, \omega_2)$		$P(\alpha_1, \omega_b)$	$P(\alpha_1)$
	α_2	$P(\alpha_2, \omega_1)$	$P(\alpha_2, \omega_2)$		$P(\alpha_2, \omega_b)$	$P(\alpha_2)$
	\vdots					
	α_a	$P(\alpha_a, \omega_1)$	$P(\alpha_a, \omega_2)$		$P(\alpha_a, \omega_b)$	$P(\alpha_a)$
	Sum	$P(\omega_1)$	$P(\omega_2)$		$P(\omega_b)$	1

Table 1. Computation of site-wise conditional probability matrix.

2.6 Training, Inference and parameter estimation

Solution to Equation (2) is obtained by maximizing probabilities, spectral (A), spatial (I) and temporal (TP) using Bayes' Maximum A Posterior (MAP) estimate. This requires an inference algorithm and we employ sum-product Loopy Belief Propagation (LBP). The association potential probabilities used in both I and TP are trained using RF implemented in OpenCV (OpenCV, 2014). We determine I parameters β and η with help of Powell's search method (Kramer, 2010) and set $\beta = 5$ and $\eta = 1$ in all experiments. All potentials (A , I , and TP) were given an equal weight, i.e. $\lambda_1 = \lambda_2 = 1$.

3. Implementation

3.1 Study site and data

The study was conducted in northern Germany (52.26°N, 9.84°E) (Fig. 6). The region is characterized by intensive agriculture, with large field sizes. The average annual precipitation in the area is 656 mm and the average annual temperature is 8.9°C (January 0.6°C, July 17.5°C) (Deutscher Wetterdienst, 2012).

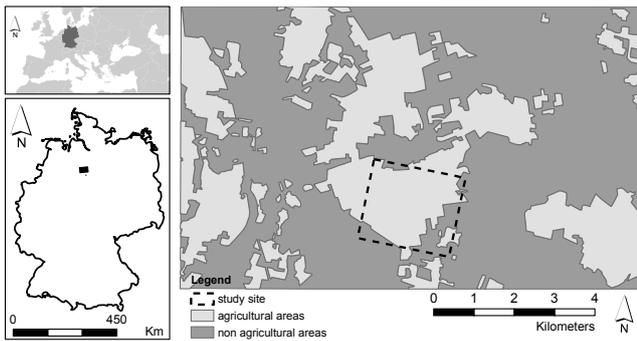


Figure 2. Location of the study site.

We use two dual polarized (HH and VV) TerraSAR-X High Resolution Spotlight images acquired on 11th March 2009 and 18th June 2009 at an incidence angle of 34.75° with a ground resolution of 2.1 m in ground range direction and 2.4 m in azimuth direction. The images were delivered as ground range products (MGD) with equidistant pixel spacing. All images are co-registered to an extent of $7.1 \times 11.8 \text{ km}^2$ and projected to WGS 1984 UTM Zone 32N. Experiment site used covers an extent of $2.3 \times 2 \text{ km}^2$.

$${}^2 P(c_k^{t1} = \omega | c_i^{t0} = \alpha)$$

$${}^3 P(c_i^{t0} = \alpha | c_k^{t1} = \omega)$$

Ground reference data campaign was conducted concurrently with image acquisition. We divided ground reference polygons for separate use in training and validation as shown in Table 2. Separation of the ground reference polygons was done using stratified random sampling in sampling design tool in ArcGIS 10.0 (Buja and Menza., 2013).

Class	Validation (ha)	Training (ha)
Grain	66.28	40.20
Broad Leaves	56.80	52.76
Unploughed	10.51	10.21

Table 2. Data used for validation and training in hectares (ha).

3.2 Classes

The study area has the following crops:

1. Grain: oat, barley, wheat and rye,
2. Broad leaved (BL) crops: maize, potatoes, canola and sugar beets,
3. Unploughed (Unp): grassland, ruderal and hedges.

We adopt the three main classes for classification using our approach in Equation (2).

The listed crops go through different phases at different times a fact that can enhance discrimination. A crop phenology calendar is shown in Figure 3. Four phases, preparation, seeding, growing, harvesting and post harvest, are considered. Preparation phase involves ploughing and soil grooming processes before seeding. In seeding phase, crop seeds are placed in the soil. Growing phase includes the period between crop germination to ripening. After ripening, harvesting starts where mature crops are gathered from the fields using relevant methods. The last stage is post harvest phase where the field could be fallow or with some remaining ripe crops.

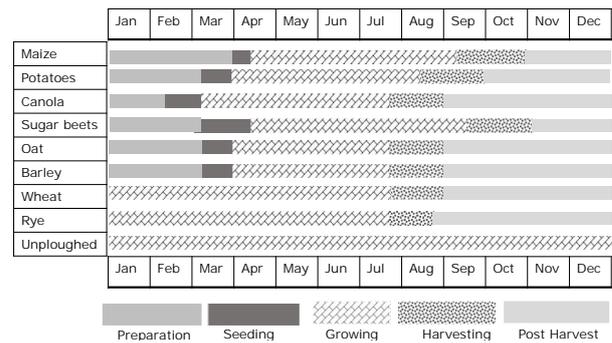


Figure 3. Phenology stages of crops considered for classification.

3.3 Features

Classification using Equation (2) requires definition of site-wise feature vectors $\mathbf{f}_i(\mathbf{x}^{t0})$ used in both A and I . We compute eight texture features (mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation) from Gray Level Co-occurrence Matrix (GLCM) (Haralick et al., 1973) using a 3×3 window with 0° direction. The features were computed from the dual polarized TerraSAR-X amplitude values in each epoch. Thus $\mathbf{f}_i(\mathbf{x}^{t0})$ used in RF (association potential) is a 16 dimensional feature vector. We use the same features in I . Thus,

the distance function, d_{ij} , in I is a 16 dimensional feature vector, i.e. $R = 16$ in Equation (3). All the features were normalized between 0 and 1 to minimize undue influence by features with high values.

3.4 Experiments

To evaluate our approach we conducted experiments using the test site and data described in Section 3.1. The experiments were done using four **approaches**:

3.4.1 Approach 1 In this approach we use RF in Section 2.3 for mono-temporal classification of crops. Approach 1 also forms A in CRF.

3.4.2 Approach 2 experiments are based on mono-temporal CRF in Section 2.1. Basically, it considers spatial context only with no temporal interactions hence "mono-temporal".

3.4.3 Approach 3 is an existing state-of-the-art spatial-temporal MRF approach in (Liu et al., 2008). We implement this approach for comparison with spatial-temporal CRF developed in the study. In the approach, site-wise transitional matrix were determined using conditional probabilities described in Section 2.5.1. To implement spatial-temporal MRF we set $\eta = 0$ in Equation (2) but, other parameters remain as described in Section 2.6. This eliminates data interaction in I reducing it to a MRF Ising model:

$$P(\mathbf{c}|\mathbf{x}) = \frac{1}{Z(\mathbf{x}^{t0})} \exp \left\{ \sum_{i \in S} A[\lambda 1(c_i^{t0}, \mathbf{x}^{t0})] + \sum_{i \in S} \sum_{j \in N} I[\lambda 2(c_i^{t0}, c_j^{t0})] + \sum_{i \in S} \sum_{t \in T} \sum_{k \in K} TP[\lambda 3(c_i^{t0}, c_k^{t1}, \mathbf{x}^{t0}, \mathbf{x}^{t1})] \right\} \quad (11)$$

where I :

$$I(c_i^{t0}, c_j^{t0}) = \begin{cases} \beta \cdot h(c_i^{t0}, c_j^{t0}) & \text{if } c_i^{t0} = c_j^{t0} \\ h(c_i^{t0}, c_j^{t0}) & \text{if } c_i^{t0} \neq c_j^{t0} \end{cases} \quad (12)$$

3.4.4 Approach 4 is the spatial-temporal CRF approach presented in this paper, where the site-wise transition matrix is computed from conditional probabilities as shown in Section 2.5.1. The approach is an extension of MRF forest change detection method in (Liu et al., 2008) into CRF crop type classification using crop phenology information.

3.5 Results

In conducted experiments, we applied cross-validation based on reference data in Table 2 to evaluate our approach (approach 4) vis-à-vis approaches 1 to 3 in Section 3.4. Overall accuracy (OA), kappa statistic, producer (Prod) and user (User) accuracy measures were computed from error matrices generated by comparing classified pixels against validation set. These accuracy measures are illustrated in Tables 3 to 10.

Tables 3 and 4 illustrate results of approach 1, RF, in both epochs (June and March). Approach 1 considers no spatial-temporal information and thus accuracy values are low. Classification accuracy is particularly low in March with many pixels mislabeled as

shown in Figure 4 compared to June. This is because of different phenological states of crops as demonstrated in Figure 3.

Class	Prod	User	Class	Prod	User
Grain	59.7%	30.5%	Grain	94.6%	72.2%
BL	49.7%	49.2%	BL	95.3%	85.3%
Unp	13.0%	53.8%	Unp	18.8%	57.6%
OA	40.3%		OA	76.6%	
Kappa	10.8%		Kappa	62.9%	

Table 3. Approach 1, March

Table 4. Approach 1, June

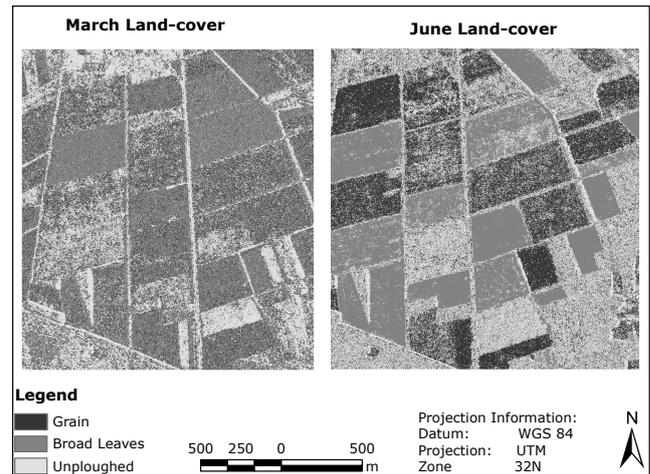


Figure 4. Approach 1, RF, classification used as baseline.

Introduction of spatial interaction by approach 2 using CRF improves discrimination of crops as shown in Tables 5 and 6. The OA improves by 3.2% and 11.7% in March and June respectively.

Class	Prod	User	Class	Prod	User
Grain	78.8%	25.6%	Grain	95.7%	91.0%
BL	50.8%	61.0%	BL	97.5%	89.7%
Unp	14.7%	61.3%	Unp	36.9%	64.0%
OA	43.5%		OA	88.3%	
Kappa	16.5%		Kappa	80.2%	

Table 5. Approach 2, March

Table 6. Approach 2, June

Results of addition of temporal potential using approach 3 in (Liu et al., 2008) are shown in Tables 7 and 8. Approach 3 improved classification OA by 48% and 12.8% in march and june when compared to approach 1 in Tables 5 and 6. In comparison to approach 2 OA improved by 45.1% and 1.1% in march and june respectively. This depicts importance of temporal information.

Class	Prod	User	Class	Prod	User
Grain	98.3%	86.8%	Grain	95.89%	90.92%
BL	95.6%	92.5%	BL	97.01%	92.48%
Unp	39.2%	75.0%	Unp	40.05%	63.25%
OA	88.3%		OA	89.4%	
Kappa	80.4%		Kappa	81.9%	

Table 7. Approach 3, March

Table 8. Approach 3, June

Results of the new spatial-temporal method, approach 4, are illustrated in Tables 9 and 10. The use of conditional probabilities to determine the temporal transition matrix improved classification accuracy significantly. Impact of this approach is evident

compared to CRF mono-temporal classification in Tables 5 and 6 considering classification accuracy improved by 47.3% and 2.9% in March and June respectively. In comparison with approach 3, MRF spatial-temporal classification, approach still improved OA by 2.5% and 1.8% in march and june respectively.

Class	Prod	User	Class	Prod	User
Grain	97.5%	91.0%	Grain	94.2%	94.7%
BL	94.9%	94.9%	BL	96.2%	95.1%
Unp	47.2%	67.1%	Unp	46.7%	47.8%
OA	90.8%		OA	91.2%	
Kappa	84.1%		Kappa	84.5%	

Table 9. Approach 4, March

Table 10. Approach 4, June

A comparison of performance of all methods using OA is depicted in Figure 5. The diagram shows that approach 4 generally performs well in both epochs compared to other approaches.

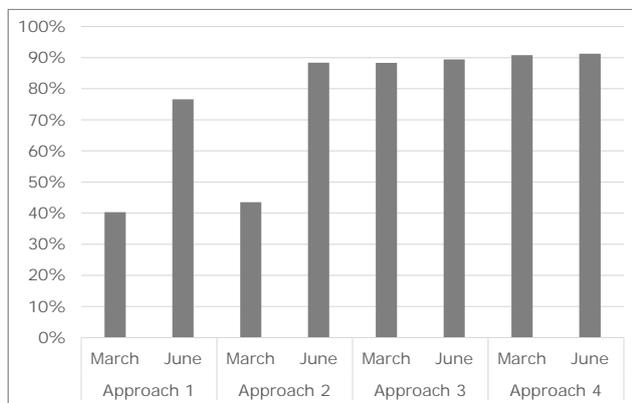


Figure 5. Overall accuracy summary of all approaches.

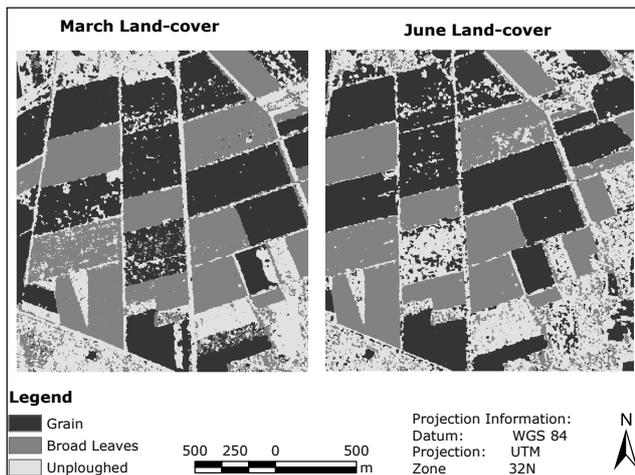


Figure 6. Land-cover maps produced using approach 4.

4. Discussion

The aim of this study was to design a spatial-temporal approach that exploits phenological changes within a vegetation season to improve discrimination of crops. Random forest, approach 1, used in A of CRF considers pixel-wise labelling by voting using decision trees. Results of this approach are characterized by mislabeled pixels in the form of "salt and pepper" as shown in Figure 4 which impacts classification accuracy. This illustrates

significance of context in classification. We used RF as a baseline to incorporate spatial and temporal context using CRF.

As demonstrated in Tables 5 and 6, spatial interaction introduced by CRF significantly improved crop discrimination. The CRF interaction potential models site dependencies in a statistically sound manner in order to allocate a label to a given site. This minimizes "salt and pepper" effect common in per pixel approaches by ensuring that sites are labelled in regard to neighbouring labels and data.

Addition of temporal term further enhanced crop discrimination by exploiting crop phenology. Phenology enhanced class discrimination because different crops vary spectrally in time. For instance, the unploughed group of crops grow throughout the year, Figure 3, and has higher accuracy in June when other crops are in growing stage. Its spectral discrimination reduces later in March when most crops are in preparation stage because they are covered by similar land-cover. We incorporated this knowledge into TP for classification using approaches 3 and 4.

Results from approach 3 and 4 indicate an increase in classification accuracy. The two approaches use site-wise conditional probabilities to compute temporal transition matrix. Conditional probabilities determine probability of inter- and intra-class changes. During a vegetation season, observed spectral changes are only within a class. Therefore, transition from class A to another class B are minimal assuming no natural or artificial crop interference. We use this fact to enhance discrimination of classes that can not be separated easily, for instance in March epoch. However, approach 3 is based on MRF which ignores spatial interactions in data a fact that makes it perform lower than CRF. We therefore implemented the approach in CRF, i.e. approach 4.

Observations from approach 3 led us to design a robust method (approach 4) to incorporate phenology information in all epochs. We use site-wise conditional probability to compute temporal transition matrix. So far our approach demonstrates stable accuracy in considered epochs. This is because conditional probability facilitate bi-directional exchange of temporal information in form of probabilities thereby enhancing accuracy of a class with low discrimination in either epoch. The approach is suitable for seasonal crop classification since uncertainty in a class at a given epoch is resolved in another. For instance, low accuracy in grain and broad leaved classes in March, see Table 3, is improved by our approach as shown in Table 9. Moreover, our temporal term (TP) is data dependent unlike in (Hoberg et al., 2015) where it was determined by empirically.

5. Conclusion and outlook

Our study has demonstrated significance of spatial-temporal information for seasonal crop classification. Spatial interaction significantly enhances spatial dependency minimizing "salt and pepper" effect common in per pixel classifiers. The introduced temporal potential resolved classes with low accuracy in March 2009 using information from June 2009 based on crop phenological changes in time. Moreover, our approach outperforms existing MRF state-of-the art spatial-temporal approach. Follow up studies will consider classification of sub-categories of crops and using images acquired at more than two epochs.

ACKNOWLEDGEMENTS

The project is funded by the German Federal Ministry of Education and Research (Project 50EE1326).

References

- Bargiel, D., Herrmann, S., Lohmann, P. and Sörgel, U., 2010. Land use classification with high-resolution satellite radar for estimating the impacts of land use change on the quality of ecosystem services. In: *ISPRS TC VII Symposium: 100 Years ISPRS*, Vol. XXXVIII, IAPRS, Vienna, Austria, pp. 68–73.
- Breiman, L., 2001. Random forests. *Machine Learning* 45(1), pp. 5–32.
- Buja, K. and Menza, C., 2013. Sampling Design Tool for ArcGIS - Instruction Manual. NOAA, Silver Spring, MD.
- Deutscher Wetterdienst, 2012. Mittelwerte der Temperatur und des Niederschlags bezogen auf den aktuellen Standort. Online: <http://www.dwd.de>.
- Geman, S. and Geman, D., 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6(6), pp. 721–741.
- Gomez-Chova, L., Fernández-Prieto, D., Calpe, J., Soria, E., Vila, J. and Camps-Valls, G., 2006. Urban monitoring using multi-temporal SAR and multi-spectral data. *Pattern Recognition Letters* 27(4), pp. 234–243.
- Goodchild, M. F., 1992. Geographical information science. *International Journal of Geographical Information Systems* 6(1), pp. 31–45.
- Gould, S., Rodgers, J., Cohen, D., Elidan, G. and Koller, D., 2008. Multi-class segmentation with relative location prior. *International Journal of Computer Vision* 80(3), pp. 300–316.
- Haralick, R. M., Shanmugam, K. and Dinstein, I., 1973. Textural Features for Image Classification. *IEEE Transactions on Systems Man and Cybernetics* 3(6), pp. 610–621.
- Hastie, T., Tibshirani, R. and Friedman, J., 2011. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics, 2nd edn, Springer.
- He, W., Jäger, M., Reigber, A. and Hellwich, O., 2008. Building Extraction from Polarimetric SAR Data using Mean Shift and Conditional Random Fields. In: *EUSAR*, Friedrichshafen, Germany, pp. 1–4.
- He, X., Zemel, R. and Carreira-Perpiñán, M., 2004. Multiscale conditional random fields for image labeling. In: *CVPR*, Vol. 2, Washington, DC, USA, pp. 695–702.
- Hoberg, T. and Müller, S., 2011. Multitemporal Crop Type Classification Using Conditional Random Fields and Rapid-Eye Data. In: *ISPRS*, Hannover, Germany.
- Hoberg, T. and Rottensteiner, F., 2010. Classification of settlement areas in remote sensing imagery using conditional random fields. In: *ISPRS TC VII Symposium-100 Years ISPRS*, Vienna, Austria.
- Hoberg, T., Rottensteiner, F. and Heipke, C., 2010. Classification of multitemporal remote sensing data using Conditional Random Fields. In: *IAPR Workshop on Pattern Recognition in Remote Sensing*, pp. 1–4.
- Hoberg, T., Rottensteiner, F. and Heipke, C., 2011. Classification of multitemporal remote sensing data of different resolution using Conditional Random Fields. In: *ICCV*, pp. 235–242.
- Hoberg, T., Rottensteiner, F., Feitosa, R. and Heipke, C., 2015. Conditional Random Fields for Multitemporal and Multiscale Classification of Optical Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 53(2), pp. 659–673.
- Jeon, B. and Landgrebe, D., 1992. Classification with spatio-temporal interpixel class dependency contexts. *IEEE Transactions on Geoscience and Remote Sensing* 30(4), pp. 663–672.
- Jianya, G., Haigang, S., Guorui, M. and Qiming, Z., 2008. A review of multi-temporal remote sensing data change detection algorithms. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVII(B7), pp. 757–762.
- Kenduiwo, B. K., Tolpekin, V. A. and Stein, A., 2014. Detection of built-up area in optical and synthetic aperture radar images using conditional random fields. *Journal of Applied Remote Sensing* 8(1), pp. 083672–1 – 083672–18.
- Korc, F. and Förstner, W., 2008. Interpreting terrestrial images of urban scenes using discriminative random fields. In: *ISPRS*, Beijing, China, pp. 291–296.
- Kosov, S., Kohli, P., Rottensteiner, F. and Heipke, C., 2013. A two-layer conditional random field for the classification of partially occluded objects. *ArXiv e-prints*.
- Kramer, O., 2010. Iterated local search with Powell's method: a memetic algorithm for continuous global optimization. *Memetic Computing* 2(1), pp. 69–83.
- Kumar, S., 2006. Discriminative random fields. *International Journal of Computer Vision* 68(2), pp. 179–201.
- Lafferty, J. D., McCallum, A. and Pereira, F., 2001. Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: *ICML*, Morgan Kaufmann, San Francisco, CA, USA, pp. 282–289.
- Leite, P., Feitosa, R., Formaggio, A., da Costa, G., Pakzad, K. and Sanches, I., 2011. Hidden Markov Models for crop recognition in remote sensing image sequences. *Pattern Recognition Letters* 32(1), pp. 19–26.
- Li, S. Z., 2009. *Markov random field modeling in image analysis*. Advances in pattern recognition, 3rd edn, Springer-Verlag, London.
- Liu, D., K., S., Townshend, J. and Gong, P., 2008. Using local transition probability models in markov random fields for forest change detection. *Remote Sensing of Environment* 112(5), pp. 2222 – 2231.
- Liu, D., Kelly, M. and Gong, P., 2006. A spatial-temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery. *Remote Sensing of Environment* 101(2), pp. 167–180.
- Lu, D. and Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing* 28(5), pp. 823–870.
- Lu, W.-L., Murphy, K., Little, J., Sheffer, A. and Fu, H., 2009. A Hybrid Conditional Random Field for Estimating the Underlying Ground Surface From Airborne LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* 47(8), pp. 2913–2922.

- Melgani, F. and Serpico, S., 2003. A Markov random field approach to spatio-temporal contextual image classification. *IEEE Transactions on Geoscience and Remote Sensing* 41(11), pp. 2478–2487.
- Moser, G. and Serpico, S., 2011. Multitemporal region-based classification of high-resolution images by markov random fields and multiscale segmentation. In: *IGARSS*, Vancouver, Canada, pp. 102–105.
- Niemeyer, J., Rottensteiner, F. and Sörgel, U., 2013. Classification of urban LiDAR data using conditional random field and random forests. In: *JURSE*, pp. 139–142.
- Niemeyer, J., Wegner, J. D., Mallet, C., Rottensteiner, F. and Söergel, U., 2011. Conditional random fields for urban scene classification with full waveform lidar data. In: *Photogrammetric Image Analysis*, Springer, pp. 233–244.
- OpenCV, 2014. Random Trees. <http://docs.opencv.org/modules/ml/doc/ml.html>. [Accessed: 20-11-2014].
- Parikh, D. and Batra, D., 2008. CRFs for image classification. Technical report, Carnegie Mellon University.
- Solberg, A. H. S., Taxt, T. and Jain, A., 1996. A markov random field model for classification of multisource satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing* 34(1), pp. 100–113.
- Swain, P. H., 1978. Bayesian Classification in a Time-Varying Environment. *IEEE Transactions on Systems, Man and Cybernetics* 8(12), pp. 879–883.
- Tso, B. and Mather, P. M., 2009. *Classification methods for remotely sensed data*. 2nd edn, CRC Press, Boca Raton.
- Tupin, F., 2010. Fusion of Optical and SAR Images. In: U. Sörgel (ed.), *Radar Remote Sensing of Urban Areas*, Remote Sensing and Digital Image Processing, Vol. 15, Springer, Netherlands, pp. 133–159.
- Wegner, J., Hänsch, R., Thiele, A. and Sörgel, U., 2011a. Building Detection From One Orthophoto and High-Resolution In-SAR Data Using Conditional Random Fields. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4(1), pp. 83–91.
- Wegner, J., Söergel, U. and Rosenhahn, B., 2011b. Segment-based building detection with conditional random fields. In: *JURSE*, Munich, Germany.
- Zhong, P. and Wang, R., 2007a. A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images. *IEEE Transactions on Geoscience and Remote Sensing* 45(12), pp. 3978–3988.
- Zhong, P. and Wang, R., 2007b. Using combination of statistical models and multilevel structural information for detecting urban areas from a single gray-level image. *IEEE Transactions on Geoscience and Remote Sensing* 45(5), pp. 1469–1482.