Int J Appl Earth Obs Geoinformation 72 (2018) 26–33

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

Designing a field sampling plan for landscape-pest ecological studies using VHR optical imagery

V. Soti[a,b,c,*], C. Lelong[d], F-R. Goebel[a,b], T. Brévault[a,b,e]

[a] CIRAD, UPR AIDA, F-34398 Montpellier, France
[b] AIDA, Université Montpellier, CIRAD, F-34398 Montpellier, France
[c] CSE, Rue Léon Gontran Damas, Fann Résidence, BP 15532 Dakar, Senegal
[d] Cirad UMR TETIS, Maison de la télédétection, 500 rue JF Breton, 34093 Montpellier, France
[e] Biopass, Isra-Ucad-IRD, Dakar, Senegal

ARTICLE INFO

Keywords:
Heliocheilus albipunctella
Millet
Senegal
Pest control
Pleiades images
Sampling design

ABSTRACT

The objective of this study was to develop an easily replicable sampling methodology using very high spatial resolution (VHSR) optical imagery to study the effect of landscape composition on crop pest incidence and biological control. The methodology was developed for the millet head miner (MHM), *Heliocheilus albipunctella* (de Joannis) (Lepidoptera: Noctuidae), a key pest of millet in Senegal (West Africa). The sampling plan was developed according to two main hypotheses: (i) pest incidence increases with millet abundance in the landscape, and (ii) biological control increases with the abundance of semi-natural habitats in the landscape. VHSR satellite imagery (< 1 m) provided from a *Pléiades* sensor was used to map and to quantify the landscape elements. Covering a square region of 20 × 20 km, a hierarchical, broad-scale land cover map focusing on crop (millet and peanut crops) and tree (tree vegetation) categories was produced and validated with ground truth data. Then, the landscape variables (tree density index and millet crop density index) were calculated based on a regular grid of 100 ha for each cell size covering the study area; the variables were then split into three density classes (low-medium-high) representative of the full landscape heterogeneity and combined into nine landscape patterns. Finally, according to sampling capacity, track accessibility, and statistical constraints, 45 field sites, including five replicates for each landscape pattern, were validated and selected for pest monitoring.

1. Introduction

The spatial distribution and dynamics of crop pest populations and their trophic interactions with primary resources and natural enemies often depend on ecological processes occurring at scales larger than the single crop plot (e.g., Kareiva and Wennergren, 1995; Ricklefs and Schluter, 1993; Tscharntke et al., 2007, 2005; Kareiva and Wennergren, 1995; Ricklefs and Schluter, 1993; Tscharntke et al., 2007, 2005). Landscape composition can affect pest abundance directly by hindering its dispersal, mortality or reproduction or indirectly by fostering its natural enemies. Many studies in recent years, as reviewed by Bianchi et al. (2006a); Chaplin-Kramer et al. (2011), and Veres et al. (2013), have shown that the landscape complexity and particularly higher proportions of semi-natural areas exhibited lower pest abundance or higher pest control in fields. It is therefore important to finely characterize landscape features to better understand how they can affect the spatial dynamics of crop pest populations and their natural enemies (Forman, 1995; Gustafson, 1998; Tischendorf and Fahrig, 2000; Turner et al., 2001). The calculation of landscape indices derived from thematic maps are generally used to quantify the landscape patterns and to test the relationships between landscape properties (composition and/or structure) and the distribution of insect populations (e.g., Forman, 1995; Gustafson, 1998; Tischendorf and Fahrig, 2000; Turner et al., 2001; Carrière et al., 2012). However, sampling strategies and underlying ecological assumptions are rarely well argued. A common statement in landscape-pest studies is that to be efficient, a sampling strategy must be based on environmental gradients that are believed to exercise primary control over the distribution of the crop pest populations and their natural enemies, and the sampling sites must be independently and identically distributed. In the majority of such studies, the environmental gradients are rarely fully considered over the entire study area, and the number of sites and their replicates are often limited because of technical and cost constraints (Veres et al., 2013). However, the proportion of the host-crop and non-crop habitat areas are mostly retained as explanatory landscape variables, and the technologies currently available based on earth observation data could help to fill these

* Corresponding author at: Cirad, UPR AïDA, Carabe Team, TA-B 102/02 Avenue Agropolis, 34398 Montpellier, France.
E-mail address: valerie.soti@cirad.fr (V. Soti).

https://doi.org/10.1016/j.jag.2018.05.016
Received 20 December 2017; Received in revised form 29 May 2018; Accepted 29 May 2018
Available online 05 June 2018
0303-2434/

limitations. To identify arable fields from natural and semi-natural vegetation, remote sensing data are particularly useful since they provide a synoptic view and deliver information over large areas at a high level of detail (Nagendra, 2001). There is a large variety of sensors with a wide range of spatial and spectral resolutions. In landscape-pest studies, remote sensing imagery has rarely been used because the spatial resolution was insufficient to identify habitat biodiversity in highly fragmented areas and because the cost remained too high. However, available sensors with a sub-meter resolution (< 0.5 m) such as Quick bird or WorldView-2 sensors have shown promising results for natural vegetation identification and in particular for tree species at the crown scale (Cho et al., 2015; Immitzer et al., 2012; Karlson et al., 2014; Tooke et al., 2009). This development opens new perspectives in the ecological domain for the identification of species or groups of tree species that could promote crop pest natural regulation.

The aim of the paper is to study how very high spatial resolution (VHSR) optical imagery can help in performing a stratified sampling method that is easily replicable for landscape-pest studies and uses to pest biocontrol models. More specifically, the proposed sampling strategy aims to provide a representative data set including two main statistical constraints: to sample along a relevant gradient of the landscape patterns according to landscape-pest hypotheses, within a highly constrained framework in terms of the number of sampling sites; and to maximize the probability of observing the variability of pest incidence, natural regulation, and insect diversity, especially for natural enemies. In this way, the sampling plan would serve both for field data collection (pest, natural enemies, disease) and as a second step to calculate the pertinent landscape variables around each preselected sampling points that could be tested statistically in landscape pest models.

Therefore, the sampling methodology was developed from the millet head miner (MHM), *Heliocheilus albipunctella*, (de Joannis) (Lepidoptera: Noctuidae), as a case study. This insect species is a key pest of pearl millet in West Africa (Ajayi, 1980; Guevremont, 1982; Ndoye, 1979), causing yield losses up to 85% (Krall et al., 1995; Youm and Owusu, 1998). This study was carried out over an area in the Senegalese Peanut Basin where pest regulation relies only on the action of natural enemies, and the sampling plan was developed according to two main hypotheses that are often tested in landscape-pest studies: (i) pest incidence increases with millet abundance as the “host crop” in the landscape, and (ii) biological control increases with the abundance of semi-natural habitats in the landscape. Very high-resolution remote sensing data were chosen to map the landscape elements because of the small size of the objects that structure the landscape elements, focused on crops (millet and peanut crops) and trees (tree vegetation) as well as the heterogeneity of their spatial distribution. Then, we derived landscape variables to perform the sampling design.

2. Methods

2.1. Study area

The study was carry in the Bambey agroforestry parklands (14° 43′ 42″ N, 16° 33′ 98″ E) located in the Peanut Basin, which is the most important area for staple crop production in Senegal (Fig. 1). Covering an area of approximately 20 × 20 km, this study site was selected because of the spatial heterogeneity of semi-natural vegetation patterns (Fig. 1) and the absence of any insecticidal treatments, which could disturb the biological control of the MHM by natural enemies. Characterized by a semi-arid climate with only one short rainy season from July to October and average annual rainfall varying from 400 up to 600 mm (Badiane et al., 2000), the landscape is characterized by a mosaic of arable land under *Faidherbia albida* parklands. The agricultural system is based on smallholder farming (approximately 0.25 ha) of staple crops dominated by millet and nuts covering 52% and 32%, respectively, of the Bambey area in 2014 (DAPSA, 2014). Livestock farmers and growers live together on the same space.

The *F. albida* parkland is primarily composed of centuries-old trees that are mostly isolated and regularly distributed. To a lesser extent, other tree species are presents in the study area, including *Guiera senegalensis*, *Balanites aegyptica*, *Adansonia digitata, Tamarindus indica* and *Acacia seyal*. The spatial distribution of the trees is characterized by a high spatial heterogeneity, with the highest densities in the North-West and the South-East of the study area.

2.2. Environmental data

A *Pléiades* satellite image was acquired on January 16th, 2013, at a ground resolution of 0.5 × 0.5 m in the panchromatic mode and 2 × 2 m in the multispectral mode, with blue (B), green (G), red (R) and near infrared (NIR) bands. The acquisition date was chosen during the middle of the dry season when most tree crowns were leafy and crops had just been harvested. At this time, millet and peanut fields are characterized by the presence of crop residue and bare soil, respectively. In February 2013, a land-cover field survey was conducted in the study area (Fig. 1). The monitored sites were previously selected using a stratified equal sampling procedure based on a previously acquired *Pléiades* image. Therefore, using the ArcGIS software, the study area was split into 9 tiles, and the sandy tracks were manually drawn. To facilitate the field navigation, we chose sampling sites along the sandy tracks. Twenty sampling sites per image tile, focusing on trees and crops, were preselected via image interpretation of the *Pléiades* scene, and the sites were integrated into a GIS database. Then, the sampling points were exported into a GPS (global positioning system) device, allowing easy navigation of each sampling point. In the field, the vegetation information was collected on the preselected focus point but also within their vicinity, with the aim to collect as much information as possible regarding the vegetation type. The vegetation, including crops and trees, was described on a total of 420 sites with approximately 50 sampling sites per image tile. All this information was geolocalized using a GPS and integrated into a geographic information system (GIS) database.

2.3. Image processing for land cover mapping

A common pre-processing procedure was applied to the *Pléiades* satellite image, including orthorectification and conversion of digital numbers to the top of the atmospheric reflectance. Then, we chose an object-oriented approach and defined a mapping processing chain including different levels of segmentations at various scales and a hierarchical classification (Blaschke, 2010). While it is essential to capture small landscape elements, since the image resolution was smaller than or similar in size to the objects of interest (namely, trees and small patches of land), there was a large variability of intra-class spectral signatures and other per-pixel indicators (Blaschke et al., 2014). Thus, pixel-based classifications would lead to a speckled result far from a mapping product. In addition, strong post-processing smoothing would also decrease the potential accuracy of the classification. For the calibration of the method and learning the algorithms, we split the initial 420-sites of the ground truth data set, leaving 164 sites for the evaluation step and addressing only the 256 sites remaining. A radiometric and textural analysis was then performed on these 256 learning sites to identify the most relevant descriptors of the land covers, with a special focus on major crops (millet and peanut) and the most abundant tree species. First, 36 potentially explicative radiometric variables proposed in the eCognition Developer® software (Baatz and Schäpe, 2000) were derived, including, for example, the normalized difference vegetation index (NDVI) and the normalized difference water index (NDWI), which are useful to discriminate vegetation from bare soil and to separate different classes of vegetation (Pettorelli et al., 2005; Tucker, 1979). The soil-adjusted vegetation index (SAVI, Huete (1988)) and the brightness index are known to minimize the influences of the soil brightness from the spectral vegetation (Pouget et al., 1990). Then, 64

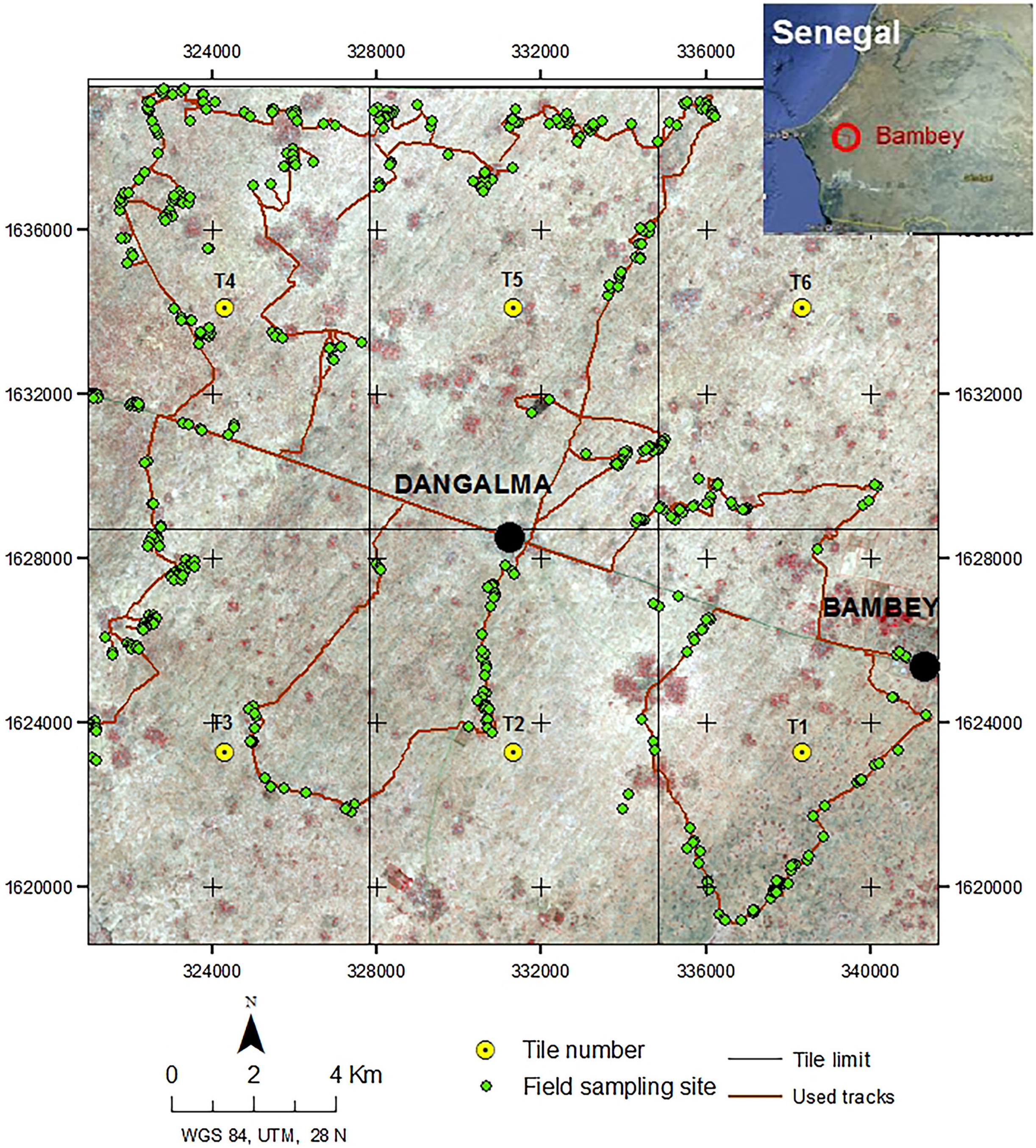

Fig. 1. Location of the study area divided in 6 tiles with the 497 ground field sites used for the image processing steps. Infrared coloured *Pléiades* image from 16 January 2013, ©CNES 2013, distribution Airbus DS/ *Pléiades* Image/ISIS programme.

textural indices were calculated, corresponding to the 8 different indices derived with the ENVI® software (Germain and Baylou, 1997; Haralick, 1979) for 8 different neighbourhood sizes (3, 7, 11, 15, 19, 21, 31 and 35 pixels) to cover the various scales of landscape element textures. The predictive relevance of these indices was then assessed using a principal component analysis (PCA) with R software. The individual factor map, for instance, was analysed to decipher which classes of land cover could be correctly discriminated from the other classes, and in which principal axis. The eigenvalues of each principal axis allowed the determination of which variables contributed to the information deciphered on each principal axis. The variables factor map was used to verify that the corresponding indicators are not correlated. Finally, the associated p values were used to test whether the discrimination potential of each selected variable was overestimated. At the end of this step, eighteen descriptors only were retained: 12 radiometric indices (the mean of each of the blue, green, red, and near-infrared bands, the mean of the panchromatic band, blue ratio, green ratio, infrared ratio, brightness, brilliance index, NDVI, and NDWI) and 6 textural indices (the variance and homogeneity in a neighbourhood of both 15 and 21 pixels, the mean and contrast in a neighbourhood of 7 pixels). Using the 'multi-resolution segmentation' algorithm in the eCognition Developer® (Baatz and Schäpe, 2000), several scale and shape parameters were tested to find the best compromise to encompass the large range of patch sizes characterizing the tree vegetation and field crops present in the study area. Three segmentation levels were found as appropriate to detect the 16 targeted land use classes, which are composed of height prevalent tree species (*Faidherbia albida, Balanites aegyptica, Adansonia digitata, Acacia seyal, Guiera senegalensis,* *Mangifera indica, Azadirachta indica,* and *Tamarindus indica*), the two major crops (millet and peanut), market gardens, fallow land, bare soil, water, roads, and buildings. The first level (L1) included 3 general land cover classes: roads, urban, and rural areas. The second level (L2) was based on the segmentation at the sub-field scale, allowing the differentiation of cropped areas from natural vegetation, and even millet and peanut fields from other crops (such as fallows). The third level (L3) was aimed at separating tree vegetation from any other land cover type, with a specific focus on detecting small vegetation patches such as individual trees. Once each data set was partitioned, a classification ruleset was developed. For each class, the radiometric descriptors were based on the spectral analysis, and the textural descriptors were defined by expert knowledge using fuzzy rules. These allowed class membership functions to be established. The membership functions were then applied at each image segmentation level to finally produce two hierarchical maps resulting from the respective L2 and L3 image segmentations: a general map where trees were regrouped into one single class named “tree vegetation”, and a detail map with the height of prevalent tree species of the study area. Finally, the classification accuracy was evaluated using the ground truth data that were not used in the classification process (corresponding to 164 sites). In the error matrix, the allocated land-cover class of the validation objects was compared to the observed land-cover class, and the quality of the classification was measured through the overall accuracy coefficient and the Kappa index (Congalton, 1991). The overall accuracy is a measure of how many ground truth pixels were correctly classified. The Kappa index represents the proportion of agreement obtained after removing the proportion of agreement that could be expected to occur by chance (Foody, 1992). The latter returned values range from 0 for poor agreement between the predicted and observed values to 1 for perfect agreement (Cohen, 1960).

2.4. Landscape metrics

According to our hypotheses (see introduction), we quantified two landscape variables from the land cover map obtained at a fine scale: the natural and semi natural habitats, mainly trees, and the millet crop proportion area, the host crop habitat. In a first step, using the ESRI ArcGIS ™ software (Redlands, CA, USA), the complete image was divided into a regular hexagonal grid where each cell was 100 ha (Fig. 1). As recommended by Birch et al. (2007) hexagons are preferable in ecological applications when the analysis includes aspects of connectivity or movement paths. In our study case, the hexagonal area was chosen in relation to the maximum displacement capacity of the MHM natural enemies. According to expert knowledge, this distance was estimated for natural enemies to be approximately 1 km from the field sampling assuming a short-distance dispersal of noctuid moths when resources are locally available.

In a second step, the hexagonal grid was used to cut the land use map into 491 landscape units, which allowed the calculation of each cell’s two landscape variables, the tree density index (TDI) and the millet crop density index (MCDI). These indices supposedly influenced the incidence of *H. albipunctella* and the biological control by natural enemies. The first variable concerns the natural vegetation and is mainly composed of isolated trees whose presence and density are supposed to favour the natural regulation of the MHM. The natural vegetation can provide food and alternative prey but also shelter from natural enemies (Elliott et al., 1999; Geiger et al., 2009; Gentry, 1988; Lys and Nentwig, 1994; Marino and Landis, 1996; Otieno et al., 2011). We hypothesized that the natural pest control increases with the abundance of natural habitats, which are mainly trees in the traditional parkland agroforestry in the landscape (‘natural enemies’ hypothesis). Then, we calculated the TDI (tree density index), which is the relative proportion of tree vegetation cover:

$$TDI_i = \begin{cases} \dfrac{TA_i}{B_i - TA_i} & \text{if } \; TA_i \leq \dfrac{B_i}{2} \\ 1 & \textit{otherwise} \end{cases} \tag{1}$$

where TA_i is the surface area of tree vegetation within grid cell i and the cell area B_i. Once calculated for each hexagon of the grid, a tree density map was generated that covered the entire study area.

The second landscape variable is concerned with the millet crop density, and supposedly influences the MHM incidence as a "source area" for the MHM, which feeds on the millet panicle for its development cycle. It is also well-known that the MHM is a univoltine noctuid species undergoing its pupal diapause in the soil during the dry season from October to August. In the same way, we hypothesized that the MHM incidence will be higher in fields with a landscape dominated by millet crops than in other fields. Therefore, the millet crop density index (MCDI) represents the relative proportion of the millet area, and it was calculated for each grid cell of the land cover map:

$$MCDI_i = \begin{cases} \dfrac{MA_i}{B_i - MA_i} & \text{if} MA_i \leq \dfrac{B_i}{2} \\ 1 & \textit{otherwise} \end{cases} \tag{2}$$

where MA_i is the area of the millet crop within a grid cell i and cell area B_i. Once calculated for each hexagon of the grid, a millet crop density map was generated that covered the entire study area.

2.5. Sampling design

Due to time and resource constraints, the number of field sites was fixed to a total of 45. Then, to consider the gradient of the landscape variable combination densities and minimize the statistical replications, we decided to keep at least 9 landscape patterns with 5 replicas in each landscape pattern. Therefore, from the TDI and MCDI density maps, we first used an agglomerative hierarchical clustering algorithm and, more specifically, the complete-linkage clustering algorithm to split both the landscape variables into three optimal classes: low, medium, and high densities. From the resultant tree dendogram, this method allows users to specify the number of clusters, k, to be generated. In our case, we initially chose 4 clusters and then used the cut-tree function based on the higher relative loss of the inertia criteria to select 3 fin. l homogenous classes of densities. The Euclidean distance was calculated to measure the dissimilarity between each pair of observations and the Ward’s minimum variance method for the cluster dendogram partitioning. Therefore, we used the R freeware Team (2016) and “Fastcluster,” “cluster,” “devtools” and “JLutils” packages. In a second step, all three density classes of both landscape indices were combined in pairs (3 × 3 density classes) and allowed to produce a grid map composed of nine landscape patterns. Then, in a third step, 45 sites representing 5 replications in each landscape pattern were randomly selected from the resulting cluster. To do this step, we used the random polygon selection tools in the ArcGIS software, which allowed specifying the number of replications for each landscape pattern class. Finally, two sampling plans were randomly generated for each of the 45 sites. The second plan was used when the first had field access constraints that were too strong. The 45 resulting sampling sites were finally validated on the field in terms of their landscape pattern combination. Then, observations of the MHM incidence, parasitism rate, crop damage, and natural enemy diversities were carried out on the 45 selected fields during the following growing season.

3. Results

3.1. Land cover map

The land-cover map validation of the general map (L2) showed a good match between the predicted and observed classes with a global

Int J Appl Earth Obs Geoinformation 72 (2018) 26–33

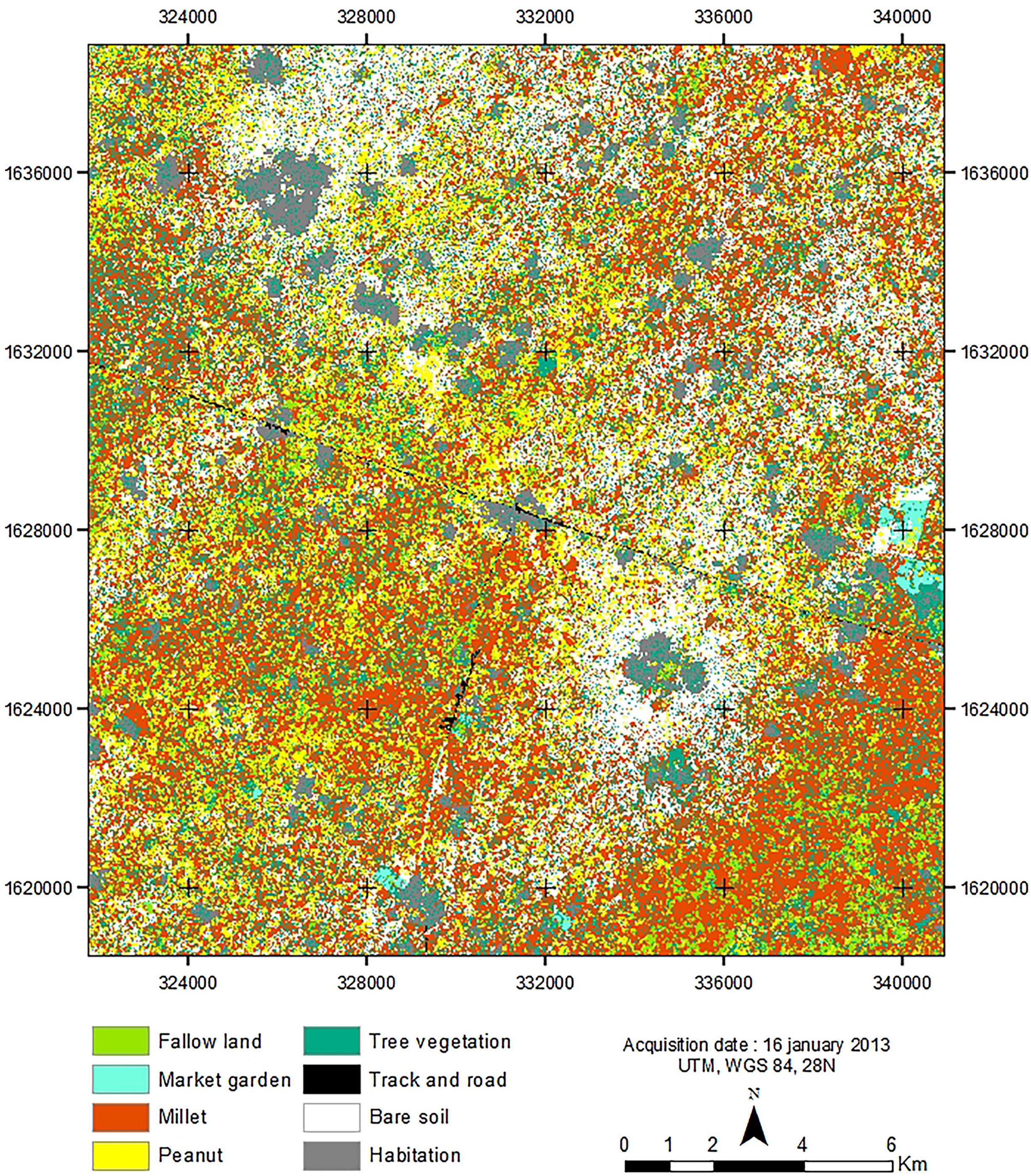

Fig. 2. Final land cover map resulted from the object-based classification of the second level of the image segmentation allowing differentiating cropped species from natural vegetation (*Pléiades* Image, ©CNES 2013, distribution Airbus DS/ Spot Image/ISIS programme).

accuracy rate of 71% and a Kappa index of 0.57. The best discrimination results were for the "bare soil" (100%), "tree vegetation" (85%), and "roads" (83%), followed by "urban areas" (70%), "millet fields" (55%) and "peanut fields" (46%). Regarding the map at a finer scale (L3), the identification of tree species was poor, with a global accuracy rate of 38% and a Kappa index of 0.31. Specifically, the global accuracy rate was 81% for *Guiera senegalensis*, 38% for *Balanites aegyptica*, 27% for *Adansonia digitata*, 20% for *Acacia seyal*, 17% for *Tamarindus indica*, and 6% for *Faidherbia albida.* Such results showed that the *Pléiades* image was not appropriate for consistent discrimination of tree species in our case study. Accordingly, the tree species were not taken into account in the sampling strategy, and the second classification level (Fig. 2) that grouped all trees in a general class named "tree vegetation" was used for the landscape metrics calculation.

3.2. Landscape metrics

According to our sampling capacity, the hierarchical clustering algorithm allowed splitting the two maps representing the TDI and the MCDI into three classes: low, medium and high densities. With 55% of the surface area, the medium TDI was dominant in the study area. The high TDI at 22% was found primarily in the north-west region of the study area, and the low density (23%) was located primarily in the centre and the north of the study area. Regarding the MCDI, a higher

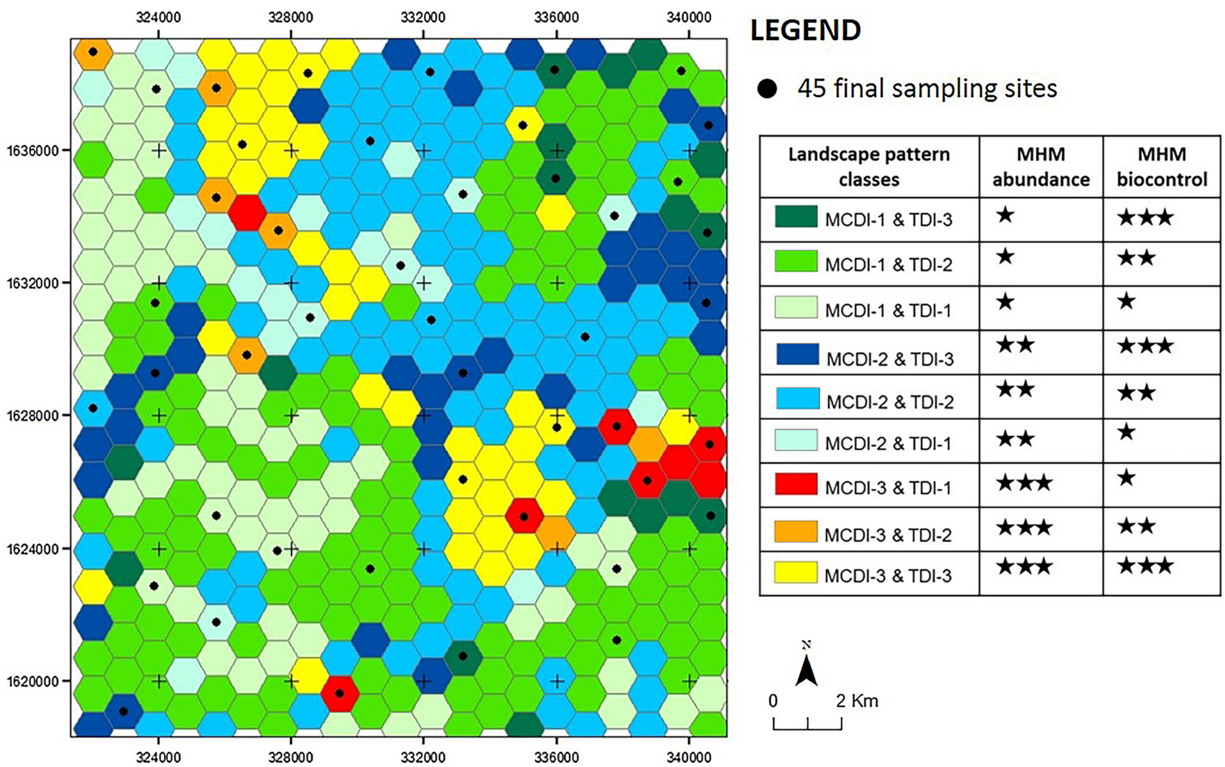

Fig. 3. Result of the final stratification design. The colour indicates the 9 landscape patterns resulting from the combination of the density indices MCDI and TDI (1: high density; 2: medium density; 3: low density). Stars and their numbers (1 to 3) indicate the expected level of MHM abundance and biocontrol based on the landscape pattern assumptions.

density of the millet crop prevailed in the south region of the study area (48%). Conversely, the presence of the millet was lower in the north, and north-east regions of the study area had a medium crop density at 38%. Areas where the millet density was low represented only 13% of the study area.

3.3. Sampling design

A sampling stratification map composed of 9 landscape patterns was obtained following the combination of the TDI and MCDI class density pairs (Fig. 3). The pattern that combined a high millet density with a medium tree vegetation density was the most common in the study area (30%), followed by the medium millet density with a medium tree vegetation density (24%). Combinations with the low millet density were less frequent in the study zone, but they were sufficient to meet the five sites for each landscape pattern as a statistical requirement.

The geographic coordinates of the centre of each cluster were calculated using the ArcGIS software, exported onto a GPS and validated in the field based on the landscape features and track accessibility. Once validated in the field, the 45 sampling sites were selected for further entomological, agronomic and landscape field surveys.

4. Discussion

The land-cover map at the second level (L2) derived from the *Pléiades* imagery allowed a good discrimination of the natural vegetation dominated by trees, bare soil, tracks or roads and urban areas (Kappa = 0.70). The satellite image was acquired in January during the middle of the dry season. During this season, the majority of the trees were well identified, because they were leafy (85% well recognized). The millet and peanut crops were classified based on the presence of post-harvest residues from the millet plots on the ground. An earlier acquisition, in November or December, would have been more favourable for millet identification, as millet residues would be less dry and, therefore, more easily detectable with remote sensing. This result shows that spectral and spatial properties of a *Pléiades* image are still not sufficient to well discriminate the different types of crops in our study zone characterized by small plots of approximately 0.25 ha. Indeed, numerous studies carried out with medium-resolution images made the same conclusions, incriminating the low resolution of the images and the fragmented landscape characterized by small sizes fields and resulting in high regional variability in terms of agricultural systems and practices (Vintrou et al., 2012). To overcome this problem, multitemporal and multispectral remote sensing imagery with a coarse resolution (i.e., NOAA-AVHRR, SPOT-VEGETATION and TERRA-MODIS) have been widely used for crop identification in recent years (Hountondji et al., 2006; Justice et al., 1985; Smith et al., 2003; Vintrou et al., 2012) and have reduced the confusion between different types of land cover (Husak et al., 2008). However, these studies remain unable to determine the crop type, especially in arid areas such as in the Sahel belt. Regarding tree species identification, the results showed some limitations with a low efficiency using the *Pléiades* images in the detection of isolated trees species. However, recent studies in West Africa using very high-resolution images WorldView-2 (8 spectral bands and 46 cm spatial resolution) showed better results in tree species identification (Karlson et al., 2016). Two factors could explain this difference. The first is the profile of the Bambey agroforestry parklands where the tree density is much lower and the tree species are limited. The second factor relates to the methodology based on the use of satellite imagery time series data, which facilitates tree discrimination based on differences between their individual phenological cycles. According to these statements, the use of satellite time series data could improve the discrimination of crops and trees in our case study. First, to more accurately identify the two major crops, we suggest acquiring an additional satellite image just before the harvesting period, which generally occurs in October. Nevertheless, the presence of clouds during this period could hamper the image acquisition. Second, to further separate the tree species, we propose two different approaches: (1) using satellite images with richer spectral bands and a high revisit time, such as the WorldView-3 sensor, which has a panchromatic band at 0.31 × 0.31 m and eight multispectral bands at 1.24 × 1.24 m, during the dry season when trees are leafy; or (2) acquiring one *Pléiades* image per month during the dry season in order to map all the tree species as a function of the evolution of the leaf density.

We confirm that the remote sensing data, through the characterization of land cover/use, is very useful to quantify the landscape elements in terms of their composition and spatial distribution. The approaches offer a quasi-continuous and complete land cover with a gradient of variability, which could be quantitatively estimated and located in space. In landscape ecology, this allows the study of the effect of landscape on the presence or abundance of animal species. In our study, we were able to quantify the trees and millet crops. However, the selection of the landscape indices is limited by various constraints: the precision of the object image classification and the number of fields/observation sites monitored, which was limited in our case to 45 due to time and resource constraints. It would have been interesting to consider the sampling plan in the tree species diversity as the third variable to identify species promoting the diversity and abundance of natural enemies, thus enhancing natural regulation (Bianchi et al., 2006b). However, it is recommended to limit the number of landscape variables to consider because of their biological interpretation issues. Nevertheless, we plan to determine the tree diversity richness around each sampling site in a second step of this research project.

The resulting sampling strategy is based on environmental data for MHM populations and their natural enemies as a function of the habitats. Combining the two landscape indices calculated from a land cover map, we obtained a sampling plan offering nine pattern modalities with five or more possible replications for each. Generally, most studies dealing with the effects of landscape on the pest incidence concern a limited number of monitored sites in relatively small areas (Elliott et al., 2002a,b; Elliott et al., 1999; Holland and Fahrig, 2000; Marino and Landis, 1996; Prasifka et al., 2004; Purtauf et al., 2005). Here, we propose a sampling design mixing a grid and an equal stratified sampling method that is considered the most robust and accurate for habitat suitability modelling (Hirzel and Guisan, 2002). This systematic method is simple to run and has the merit to consider all the possible combinations of the environmental variables, which can play a key role in the distribution of pests and their natural enemies. In the case of studies dealing with more than two landscape metrics, which result in more than 200 hundred possible combinations, a non-supervised clustering method is generally used (Danz et al., 2005; Roux et al., 2013) to reduce the strata for the selection of the final samples. Nevertheless, the quality of the implementation of such a sampling design will depend on two critical requirements. First, it is necessary to obtain a fine land cover map covering all the study area. The use of very high-resolution satellite image, such as the *Pléiades* sensor, is useful to inventory land cover exhaustively over a large area. Additionally, knowledge on insect biology and ecology is essential because it allows better targeting of relevant landscape variables and thus limits the number of required landscape patterns. The selection of a suitable sample design method ensures that the samples for which time and money has been invested for collecting can support the desired inferences (Elliott and Décamps, 1973). Knowledge on the dispersal capacity of the targeted species is also important to consider in the sampling plan (Gilchrist and Meats, 2012) to fix the size and the distance of each observation site. Unfortunately, the latter is rarely known, making it necessary to choose an important distance between the sampling sites to limit any risk of spatial autocorrelation for the variable of interest (Dale and Fortin, 2014; Pasher et al., 2013). In our case study, based on the knowledge on the MHM bioecology and their natural enemies, the minimum distance between each of the 45 sampling sites was fixed at 2 km. This distance is sufficiently long to allow a multi-scale analysis to be performed in the future to identify the spatial units at which landscape management resulting from the research could be developed and implemented (e.g., Dunford and Freemark, 2005; Eigenbrod et al., 2008; Holland et al., 2004).

5. Conclusion

This work proposed a methodology to design a sampling plan for a wide range of empirical landscape ecology studies. From the case study of the millet head miner, *H. albipunctella* in Senegal, we demonstrated the value of very high spatial resolution satellite imagery (< 1 m) to extract landscape features and, more specifically, tree vegetation as a potential proxy of the MHM incidence and its regulation by natural enemies. This work could be easily reproduced to develop systematic sampling strategies at a landscape scale for studies addressing pest incidence and the conservation of biological control. To improve the sampling design, we suggest considering tree species, thus performing tree identification, using satellite images with more spectral bands, such as the WordView-3 sensor (8 bands and a spatial resolution of 30 cm in panchromatic).

Authors' contributions

V. Soti, R. Goebel and T. Brévault conceived the study, and participated in its design and coordination. V. Soti supervised the overall project and drafted the manuscript. C. Lelong contributed to the image processing. All authors wrote, read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This study was funded by the SAFSE project (CIRAD fund AI-210723, Monptellier, France) and TRECS project (CNES fund DAR 4800000915, Toulouse, France). The satellite imagery was obtained through the ISIS French program and processed by Coline Billand (UFR FAM, Dept SIAFEE, Agroparitech) who prepared the landscape map. We want to thank Emile Faye (DAPSA-Senegal) and Bamba Diop (CSE) for their respective contributions to improve this paper. We also want to thank Cheikh Thiaw from the ISRA-CNRA Institute of Bambey and Assize Touré from the CSE (Centre de Suivi Ecologique), the Ecological Monitoring Center of Dakar who have both facilitated the landscape survey on the field. We also want to thank Danny Lo Seen, Annelise Tran and Agnes Begue from Cirad in Montpellier-France for their English review contributions and Diam Abdoul Sow and Mor Fall as guides and botanical experts for field environmental surveys.

References

Ajayi, O., 1980. Insect pests of millet in Nigeria. Samaru Miscellaneous Paper N°97. pp.10, ref.11.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. Angewandte Geographische Informationsverarbeitung XII 58, 12–23.

Badiane, A., Khouma, M., SENE, M., 2000. Région de Diourbel: Gestion des eaux. Drylands Research Working Paper Drylands Research 14. pp. 15.

Bianchi, F., Booij, C., Tscharntke, T., 2006a. Sustainable pest regulation in agricultural landscapes: a review on landscape composition, biodiversity and natural pest control. Proc. R. Soc. Lond. B: Biol. Sci. 273, 1715–1727.

Bianchi, F.J., Booij, C., Tscharntke, T., 2006b. Sustainable pest regulation in agricultural landscapes: a review on landscape composition, biodiversity and natural pest control. Proc. R. Soc. Lond. B: Biol. Sci. 273, 1715–1727.

Birch, C.P., Oom, S.P., Beecham, J.A., 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecol. Modell. 206, 347–359.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. Remote Sens. 65, 2–16.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R.Q., van der Meer, F., van der Werff, H., van Coillie, F., 2014. Geographic object-based image analysis–towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191.

Carrière, Y., Goodell, P.B., Ellers-Kirk, C., Larocque, G., Dutilleul, P., Naranjo, S.E., Ellsworth, P.C., 2012. Effects of local and landscape factors on population dynamics of a cotton pest. PLoS One 7, e39862.

Chaplin-Kramer, R., O'Rourke, M.E., Blitzer, E.J., Kremen, C., 2011. A meta-analysis of crop pest and natural enemy response to landscape complexity. Ecol. Lett. 14,

922–932.
Cho, M.A., Malahlela, O., Ramoelo, A., 2015. Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation implications: Dukuduku forest patch as case study. Int. J. Appl. Earth Obs. Geoinf. 38, 349–357.
Cohen, J.A., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46.
Congalton, R., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Rentote Sens. Environ. 37, 35–46.
Dale, M.R., Fortin, M.-J., 2014. Spatial Analysis: A Guide for Ecologists. Cambridge University Press.
Danz, N.P., Regal, R.R., Niemi, G.J., Brady, V.J., Hollenhorst, T., Johnson, L.B., Host, G.E., Hanowski, J.M., Johnston, C.A., Brown, T., 2005. Environmentally stratified sampling design for the development of Great Lakes environmental indicators. Environ. Monit. Assess. 102, 41–65.
DAPSA, 2014 Direction de l'Analyse, de la Prévision et des Statistiques Agricoles, 2013. Rapport de présentation des résultats définitifs de l'enquête agricole 2013–2014.
Dunford, W., Freemark, K., 2005. Matrix matters: effects of surrounding land uses on forest birds near Ottawa, Canada. Landsc. Ecol. 20, 497–511.
Eigenbrod, F., Hecnar, S.J., Fahrig, L., 2008. The relative effects of road traffic and forest cover on anuran populations. Biol. Conserv. 141, 35–46.
Elliott, J.-M., Décamps, H., 1973. Guide pour l'analyse statistique des échantillons d'invertébrés benthiques, Annales de Limnologie. Station Biologique du lac d'Orédon 79–120.
Elliott, N.C., Kieckhefer, R.W., Lee, J.-H., French, B.W., 1999. Influence of within-field and landscape factors on aphid predator populations in wheat. Landsc. Ecol. 14, 239–252.
Elliott, N., Kieckhefer, R., Beck, D., 2002a. Effect of aphids and the surrounding landscape on the abundance of Coccinellidae in cornfields. Biol. Control 24, 214–220.
Elliott, N., Kieckhefer, R., Michels, G., Giles, K., 2002b. Predator abundance in alfalfa fields in relation to aphids, within-field vegetation, and landscape matrix. Environ. Entomol. 31, 253–260.
Foody, G., 1992. On the compensation for chance agreement in image classification accuracy assessment. Photogramm. Eng. Remote Sens. 58, 1459–1460.
Forman, R.T., 1995. Some general principles of landscape and regional ecology. Landsc. Ecol. 10, 133–142.
Geiger, F., Wäckers, F.L., Bianchi, F.J.J.A., 2009. Hibernation of predatory arthropods in semi-natural habitats. BioControl 54, 529–535.
Gentry, A.H., 1988. Tree species richness of upper Amazonian forests. Proc. Natl. Acad. Sci. 85, 156–159.
Germain, C., Baylou, P., 1997. Caractérisation multi-échelle de l'anisotropie d'une texture, 16° Colloque sur le traitement du signal et des images, FRA, 1997. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.
Gilchrist, A., Meats, A., 2012. Factors affecting the dispersal of large-scale releases of the Queensland fruit fly, Bactrocera tryoni. J. Appl. Entomol. 136, 252–262.
Guevremont, H., 1982. Recherches sur l'entomofaune du mil, Rapport annuel de Recherches pour 1982. CNRA, Tarna, Niger p. 69.
Gustafson, E.J., 1998. Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1, 143–156.
Haralick, R.M., 1979. Statistical and structural approaches to texture. Proc. IEEE 67, 786–804.
Hirzel, A., Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability modelling. Ecol. Modell. 157, 331–341.
Holland, J., Fahrig, L., 2000. Effect of woody borders on insect density and diversity in crop fields: a landscape-scale analysis. Agric. Ecosyst. Environ. 78, 115–122.
Holland, J.D., Bert, D.G., Fahrig, L., 2004. Determining the spatial scale of species' response to habitat. BioScience 54, 227–233.
Hountondji, Y.C., Sokpon, N., Ozer, P., 2006. Analysis of the vegetation trends using low resolution remote sensing data in Burkina Faso (1982–1999) for the monitoring of desertification. Int. J. Remote Sens. 27, 871–884.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.
Husak, G., Marshall, M., Michaelsen, J., Pedreros, D., Funk, C., Galu, G., 2008. Crop area estimation using high and medium resolution satellite imagery in areas with complex topography. J. Geophys. Res.: Atmos. 113.
Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sens. 4, 2661–2693.
Justice, C.O., Townshend, J.R.G., Holben, B.N., Tucker, C.J., 1985. Analysis of the phenology of global vegetation using meteorological satellite data. Int. J. Remote Sens. 6, 1271–1318.
Kareiva, P., Wennergren, U., 1995. Connecting landscape patterns to ecosystem and population processes. Nature 373, 299.
Karlson, M., Reese, H., Ostwald, M., 2014. Tree crown mapping in managed woodlands (parklands) of semi-arid West Africa using WorldView-2 imagery and geographic object based image analysis. Sensors 14, 22643–22669.
Karlson, M., Ostwald, M., Reese, H., Bazié, H.R., Tankoano, B., 2016. Assessing the potential of multi-seasonal WorldView-2 imagery for mapping West African agroforestry tree species. Int. J. Appl. Earth Obs. Geoinf. 50, 80–88.
Krall, S., Youm, O., Kogo, S., 1995. Panicle insect pest damage and yield loss in pearl millet. In: Proceeding of an International Consultative Workshop on Panicle Insect Pest of Sorghum and Millet. ICRISAT Sahelian Centre, Niamey, Niger. pp. 135–145.
Lys, J.-A., Nentwig, W., 1994. Improvement of the overwintering sites for Carabidae, Staphylinidae and Araneae by strip-management in a cereal field. Pedobiologia (Germany).
Marino, P.C., Landis, D.A., 1996. Effect of landscape structure on parasitoid diversity and parasitism in agroecosystems. Ecol. Appl. 6, 276–284.
Nagendra, H., 2001. Using remote sensing to assess biodiversity. Int. J. Remote Sens. 22, 2377–2400.
Ndoye, M., 1979. L'entomofaune nuisible du mil a chandelle (Pennisetum typhoides) au Sé né gal'. Compte rendus des travaux: Congré s sur la Lutte Contre les Insectes en Milieu Tropical. pp. 13–16.
Otieno, M., Woodcock, B.A., Wilby, A., Vogiatzakis, I.N., Mauchline, A.L., Gikungu, M.W., Potts, S.G., 2011. Local management and landscape drivers of pollination and biological control services in a Kenyan agro-ecosystem. Biol. Conserv. 144, 2424–2431.
Pasher, J., Mitchell, S.W., King, D.J., Fahrig, L., Smith, A.C., Lindsay, K.E., 2013. Optimizing landscape selection for estimating relative effects of landscape variables on ecological responses. Landsc. Ecol. 28, 371–383.
Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol. Evol. 20, 503–510.
Pouget, M., Madeira, J., Le Floch, E., Kamal, S., 1990. Caracteristiques spectrales des surfaces sableuses de la region cotiere nord-ouest de l'Egypte: application aux donnees satellitaires SPOT.
Prasifka, J.R., Heinz, K.M., Minzenmayer, R.R., 2004. Relationships of landscape, prey and agronomic variables to the abundance of generalist predators in cotton (Gossypium hirsutum) fields. Landsc. Ecol. 19, 709–717.
Purtauf, T., Roschewitz, I., Dauber, J., Thies, C., Tscharntke, T., Wolters, V., 2005. Landscape context of organic and conventional farms: influences on carabid beetle diversity. Agric. Ecosyst. Environ. 108, 165–174.
Ricklefs, R.E., Schluter, D., 1993. Species diversity: regional and historical influences. Species Diversity in Ecological Communities. pp. 350–363.
Roux, E., Gaborit, P., Romaña, C.A., Girod, R., Dessay, N., Dusfour, I., 2013. Objective sampling design in a highly heterogeneous landscape - characterizing environmental determinants of malaria vector distribution in French Guiana, in the Amazonian region. BMC Ecol. 13, 45.
Smith, J.H., Stehman, S.V., Wickham, J.D., Yang, L., 2003. Effects of landscape characteristics on land-cover class accuracy. Remote Sens. Environ. 84, 342–349.
Team, R.C., 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
Tischendorf, L., Fahrig, L., 2000. How should we measure landscape connectivity? Landsc. Ecol. 15, 633–641.
Tooke, T.R., Coops, N.C., Goodwin, N.R., Voogt, J.A., 2009. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. Remote Sens. Environ. 113, 398–407.
Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I., Thies, C., 2005. Landscape perspectives on agricultural intensification and biodiversity–ecosystem service management. Ecol. Lett. 8, 857–874.
Tscharntke, T., Bommarco, R., Clough, Y., Crist, T.O., Kleijn, D., Rand, T.A., Tylianakis, J.M., van Nouhuys, S., Vidal, S., 2007. Conservation biological control and enemy diversity on a landscape scale. Biol. Control 43, 294–309.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150.
Turner, M.G., Gardner, R.H., O'neill, R.V., 2001. Landscape Ecology in Theory and Practice. Springer.
Veres, A., Petit, S., Conord, C., Lavigne, C., 2013. Does landscape composition affect pest abundance and their control by natural enemies? A review. Agric. Ecosyst. Environ. 166, 110–117.
Vintrou, E., Desbrosse, A., Bégué, A., Traoré, S., Baron, C., Seen, D.L., 2012. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. Int. J. Appl. Earth Obs. Geoinf. 14, 83–93.
Youm, O., Owusu, E., 1998. Assessment of yield loss due to the millet head miner, Heliocheilus albipunctella (Lepidoptera: Noctuidae) using a damage rating scale and regression analysis in Niger. Int. J. Pest Manage. 44, 119–121.