CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

Extracting multilayer networks from Sentinel-2 satellite image time series

Roberto Interdonato[1,2]* , Raffaele Gaetano[1,2], Danny Lo Seen[1,2], Mathieu Roche[1,2] and Giuseppe Scarpa[3]

[1]CIRAD, UMR TETIS, Maison de la Télédétection, 500 Rue J.-F. Breton, F-34000 Montpellier, France, [2]UMR TETIS, University of Montpellier, AgroParisTech, CIRAD, CNRS, IRSTEA, F-34000 Montpellier, France (e-mails: raffaele.gaetano@cirad.fr; danny.lo_seen@cirad.fr; mathieu.roche@cirad.fr) and [3]Department of Electrical Engineering and Information Technology (DIETI), University Federico II, 80125 Naples, Italy (e-mail: giscarpa@unina.it)
*Corresponding author. Email: roberto.interdonato@cirad.fr

Action Editor: Hocine Cherifi, University of Burgundy, France

Abstract

Nowadays, modern Earth Observation systems continuously generate huge amounts of data. A notable example is the Sentinel-2 Earth Observation mission, developed by the European Space Agency as part of the Copernicus Programme, which supplies images from the whole planet at high spatial resolution (up to 10 m) with unprecedented revisit time (every 5 days at the equator). In this data-rich scenario, the remote sensing community is showing a growing interest toward modern supervised machine learning techniques (e.g., deep learning) to perform information extraction, often underestimating the need for reference data that this framework implies. Conversely, few attention is being devoted to the use of network analysis techniques, which can provide a set of powerful tools for unsupervised information discovery, subject to the definition of a suitable strategy to build a network-like representation of image data. The aim of this work is to provide clues on how Satellite Image Time Series can be profitably represented using complex network models, by proposing a methodology to build a multilayer network from such data. This is the first work to explore the possibility to exploit this model in the remote sensing domain. An example of community detection over the provided network in a real-case scenario for the mapping of complex land use systems is also presented, to assess the potential of this approach.

Keywords: multilayer network; remote sensing; satellite image time series; Sentinel-2; landscape stratification

1. Introduction

Nowadays, modern Earth Observation (EO) systems include several missions that, through the use of different satellites, continuously monitor our planet by producing huge amounts of heterogeneous data, for example, optical, radar, atmospheric, and so on. A notable example is the Sentinel-2 (S2) EO mission, developed by the European Space Agency (ESA) as part of the Copernicus Programme. This mission supplies optical satellite images at high spatial resolution (up to 10 m) with high temporal revisit period (every 5 days). Thanks to the high revisiting period, such data can be profitably organized in Satellite Image Time Series (SITS), which represent a relevant source of information, expected to support a plethora of land monitoring tasks, such as fire mapping, forest and agricultural monitoring, and climate change analysis. Moreover, since S2 SITS are publicly available, they represent an attractive source for research purposes.

The aim of this work is to show how SITS can be profitably represented using complex network models, by proposing a method to extract a multilayer network from them. The goal is to

explore the possibility to characterize spatio-temporal dynamics of large-scale landscapes by identifying fine-scale geographical objects and modeling larger scale interactions among such objects by means of complex network graphs. The hypothesis is that such complex network models would lead to a spatial representation of satellite data that goes beyond conventional mapping, allowing to model different types of relationships between objects, for example, taking into account, together with the classical set of image properties (radiometry, textures, etc.), the contextual information deriving from the spatial organization of the entities in a landscape.

The need for a richer spatial representation of satellite image data had already led to the introduction, in recent years, of the GEOBIA (GEographic Object-Based Image Analysis) paradigm (Blaschke, 2010), with the aim to encourage the coupling of multiresolution segmentation and object relation models. However, most GEOBIA-based works aim at taking advantage of the aggregating abilities of the object-based approach, focusing on the possibility to extract better characterized entities. Conversely, the possibility of analyzing the spatial relations that may arise from this representation has often been neglected and still represents a challenging task (Chen *et al.*, 2018).

A particular scenario concerns the need to model complex agricultural landscapes by taking into account the existing relations between land units, which often arises in several well-known remote sensing tasks. An example in this direction is the one of land use/land cover (LULC) mapping from remote sensing imagery: while the design of advanced machine learning methods able to fully exploit the expressiveness of SITS for this task is still a current problem (Kolecka *et al.*, 2018; Inglada *et al.*, 2017; Zhang & Du, 2016; Zhu *et al.*, 2017; Interdonato *et al.*, 2019), further complex challenges are arising, such as the "functional" mapping of land use systems (Verburg *et al.*, 2009). The reason for broadening the analysis to the mechanisms behind the whole agricultural system is that, in practical cases, single agricultural plots cannot reflect the full complexity of the land use system to which they belong. For instance, human–environmental relations are a key factor in this scenario that should be effectively managed in order to preserve environmental services (DeFries & Rosenzweig, 2010) and contribute to sustainable solutions (Verburg *et al.*, 2013). In this regard, while LULC mapping can be obtained by classifying satellite images, mapping land use systems requires a broader view, in a landscape agronomy approach (Leenhardt *et al.*, 2010). One of the main limitations of existing remote sensing methods for mapping land use systems (Bégué *et al.*, 2015) is in fact the trouble in overcoming pixel-level and segmentation-based (i.e., object level) approaches, in order to be able to consider an object in its context, that is, the plot within its land use system.

Being able to model an SITS in a multilayer network, thus embedding, in different layers, the different relations between the actors characterizing the complexity of land use systems would certainly be beneficial for these kind of tasks. Furthermore, the availability of a multilayer network model to represent SITS paves the way for the definition of unsupervised remote sensing methods inspired from advanced network analysis techniques. In tasks such as land cover mapping, this would represent an extremely valid support to drive *data hungry* supervised techniques and limit their needs, in a context where collecting and labeling ground truth data is particularly costly, since it can require several days of field work for each study site.

In this exploratory work, we propose a completely unsupervised methodology to extract a multilayer network from a Satellite Image Time Series. We rely on standard segmentation techniques to build a nodeset from the SITS, while layer memberships and edge weights are inferred based on the temporal radiometric profile of each segment. Considered the unsupervised context, we model the association of nodes to the different functional classes with a certain level of uncertainty, by producing fuzzy layer memberships (obtained by the use of fuzzy clustering techniques). Note that, as we investigate our methodology using only Sentinel-2 SITS for sake of simplicity and easiness of reproducibility, the proposed approach stands general enough to deal with any kind of remote sensing imagery, eventually in a multisensor/multisource approach.

As a first example of a practical use of the proposed multilayer network model, we also present an experimental case study concerning an unsupervised landscape stratification task via multilayer community detection. The case study will be centered on the Koumbia area in Burkina Faso, an agricultural zone in the west of the country which is particularly challenging for our methodology due to its fragmented and continuously changing landscape.

The rest of the work is structured as follows: Section 2 discusses related work, Section 3 describes in detail the proposed approach, Section 4 introduces an experimental case study, and Section 5 concludes the work.

2. Related work

As we already pointed out, the use of network models in the remote sensing domain has often been neglected. Nevertheless, we can identify two main lines of research regarding the use of such models in relation with remote sensing, and more generally geospatial data analysis: models based on *interaction graphs* (i.e., focusing on human–environment relationship) and models relying on *object-based image analysis* (OBIA) (i.e., aimed at improving image-based landscape analysis). In the following, we will elaborate more on these two topics.

2.1 Interaction graphs

Generally speaking, landscape changes that can be observed in SITS result from many interacting processes. These processes are often dynamic, spatially distributed, operating at several time and space scales and can involve human activities (Schmidt-Laine & Pave, 2002). Modeling is a preferred approach to study these complex processes of human–environmental interaction, and research is conducted according to three main approaches, each with a large scientific community: systems dynamics (e.g., STELLA (Costanza, 1987)), cellular automata (e.g., SLEUTH (Clarke *et al.*, 1997)), and multiagent systems (e.g., CORMAS (Bousquet *et al.*, 1998)). While these approaches allow to model the main mechanisms occurring in a landscape, modeling the dynamic spatial structures that influence such mechanisms remains challenging. In fact, most of these approaches both do not account for the spatial dimension or limit its representation to points on a fixed-scale grid, limiting the potential in coupling the resulting models with Geographic Information Systems (GIS). With the aim to overcome these limitations, an approach based on the concept of interaction graphs has been developed, namely, Ocelet (Degenne *et al.*, 2009, 2010; Degenne & Lo Seen, 2016). A system represented by interaction graphs can thus evolve when the interaction functions (i.e., edges) are activated and modify the connected entities (i.e., nodes). The entities located in a geographical space can be represented by geometries (points, polygon lines) and/or cells (square, hexagon, or triangle). Between these entities, spatial, functional, hierarchical, and even social relationships can be defined, in the generic form of an interaction graph. Moreover, using this approach, it is possible to associate the time profiles extracted from SITS with objects extracted from very high-resolution (VHR) satellite images and to discriminate among them by taking into account *a priori* knowledge from different domains (Jahel *et al.*, 2017, 2018).

2.2 Object-based image analysis

Regarding OBIA approaches, Guttler *et al.* (2017) investigate the possibility of representing LANDSAT SITS in graph form, proposing a new method to automate the process of detecting and extracting spatio-temporal information. This approach uses traditional OBIA image processing to identify separate sets of objects for each timestamp. Then, a graph-based approach is used to detect coherent areas in space and connect objects belonging to different timestamps. For each detected entity, this method constructs a representation in the form of an evolution graph that summarizes

the evolution of the spatio-temporal phenomenon. The same evolution graph structure is used in Khiali *et al.* (2018), in order to perform a clustering task on LANDSAT SITS. Although these works have the merit to propose a first combination between network analysis and remote sensing, the proposed techniques are tailored on a specific fine-scale task, namely, the identification of landscape elements based on their "local" spatial dynamics, and can hardly be generalized to model larger scale interactions among lands units. This also concerns the possibility to handle a more heterogeneous set of data (multisensor, multisource). In this sense, a more relevant work is presented in Gaetano *et al.* (2009). It proposes an application of the theoretical model developed in Scarpa *et al.* (2009) for the unsupervised land cover classification of complex landscapes from very high-resolution (VHR) remote sensing imagery. The rationale of this methodology is to gather information on heterogeneous land cover classes through the identification of "groups" of geographic objects that share both individual and contextual properties (e.g., trees in a forest vs. trees in urban areas). Eventually, this approach allows for the reconstruction of complex areas through the merging of spatially entangled, statistically coherent clusters of landscape elements. Although its underlying modeling strategy is solely based on spatial dynamics through the use of a single VHR image, the described methodology is prone to the integration of SITS data to account for the temporal dimension, as we will show in the following.

2.3 Targeted contribution

The goal of this work is to investigate a flexible and general approach for the construction of a complex multilayer network starting from a possibly heterogeneous set of geospatial data. Our proposal explicitly builds upon the work carried out in Gaetano *et al.* (2009). To provide a first insight, we rely on the basic idea to exploit two disjoint datasets to, on one hand, provide a spatial delimitation of individual landscape elements and, on the other hand, characterize each of these elements by means of some functional properties. Eventually, a network of geographic objects will be derived using both spatial adjacencies among image objects and suitable links between disjoint, possibly distant elements sharing a well-defined "role" in the landscape (e.g., exhibiting a close behavior in time). The descriptive potential of the provided network will be assessed by trying to identify compact areas corresponding to some semantically coherent land subsystem by means of a multilayer community detection technique. Without loss of generality and to limit the complexity of this exercise, both datasets will be derived by a unique Sentinel-2 time series, tailored on a specific case study that will be described in Section 5.

3. Proposed approach

In this section, we will describe in detail the proposed methodology to extract a multilayer network from an SITS. In the following problem formulation, an SITS may be regarded as a collection of remote sensing images, possibly at multiple modes (optical, radar, etc.) and spatial resolutions, acquired multiple times over the same area and within a fixed temporal range. Even if this formulation is intended to be generic, to fix the ideas, from now on the reader may refer to a specific SITS configuration, composed by a single time series of Sentinel-2 images. Each image of the time series includes only bands at 10-m spatial resolutions, in the visible (red, green, blue) and near-infrared wavelenghts.

Concerning the multilayer network model, we refer here to the one described in Kivela *et al.* (2014):

Definition 1 (Multilayer network model). *Let $\mathscr{L} = \{L_1, \ldots, L_\ell\}$ be a set of layers. Each layer corresponds to a given type of entity relation or edge label. Consider a set $\mathscr{V}$ of entities (e.g., users), then for each choice of entity in $\mathscr{V}$ and layer in $\mathscr{L}$, we need to indicate whether the entity is present in that layer. We denote with $V_{\mathscr{L}} \subseteq \mathscr{V} \times \mathscr{L}$ the set containing the entity–layer combinations in which*
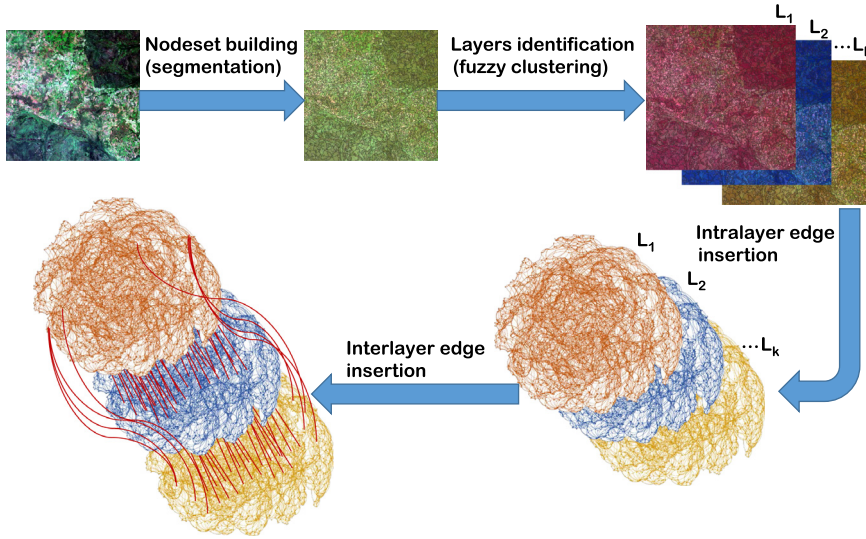
Figure 1. Visual representation of the main steps of the proposed methodology.

an entity is present in the corresponding layer. The set $E_{\mathscr{L}} \subseteq V_{\mathscr{L}} \times V_{\mathscr{L}}$ contains the undirected links between such entity–layer pairs. We hence denote with $G_{\mathscr{L}} = (V_{\mathscr{L}}, E_{\mathscr{L}}, \mathscr{V}, \mathscr{L})$ the multilayer network graph with set of nodes $\mathscr{V}$.

For every layer $L_i \in \mathscr{L}$, let $V_{L_i} = \{v \in \mathscr{V} \mid (v, L_i) \in V_{\mathscr{L}}\} \subseteq \mathscr{V}$ be the set of nodes in the graph of L_i, and $E_{L_i} \subseteq V_{L_i} \times V_{L_i}$ be the set of edges in L_i. To simplify notations, we will also refer to V_{L_i} and E_{L_i} as V_i and E_i, respectively.

Given these assumptions, our problem statement can be formalized as follows:

Definition 2 (SITS to Multilayer-Network Problem). *Given an SITS containing a set $S = \{s_1, ..., s_n\}$ of n satellite images, obtain a multilayer network $G = (V, E, L, w)$, where V is a set of nodes, L is a set of layers, $E \subseteq (V \times L) \times (V \times L)$ is a set of edges, and $w : E \Rightarrow \mathbb{R}$ is an edge weighting function.*

We will now describe in detail the main steps of the proposed methodology, structured as solutions to four different subproblems: *Nodeset building*, *Layers identification*, *Intralayer edges insertion*, and *Interlayer edges insertion*. A visual representation of these main steps is given in Figure 1.

3.1 Nodeset building

The first step of our procedure consists in the identification of a set of nodes for the multilayer network, which corresponds to specific portions of the scene:

Definition 3 (Nodeset Building Problem). *Given an SITS $S = \{s_0, ..., s_n\}$, select a subset of images $S^* \subseteq S$, operate a suitable transform to obtain a composite image $I^* = f(S^*)$, and then obtain a partition $seg(I^*)$ consisting in m distinct segments. Each segment will then correspond to a node in G, such that $m = |V|$.*

In summary, the above definition simply provides a formalization of a generic *image segmentation* workflow as advocated in the OBIA framework (Blaschke, 2010). Upstream, the transform $f(\cdot)$ represents the selection and preprocessing of the pertinent sources (for spatial resolution, acquisition times and sensor modes) aimed at producing a suitable final input for the image segmentation process (e.g., band selection, multitemporal cloud-free composites, image fusion, etc.).

The segmentation operator *seg*(·) can be seen as the process providing, for a given input image, a *partition* into multiple nonoverlapping sets of adjacent pixels (objects, or segments) covering the whole scene, each one being homogeneous with respect to a given, suitably defined homogeneity criterion. In our proposal, segments resulting from this step will constitute the complete nodeset of the multilayer network. For this reason, hereinafter, we will use the terms nodes and segments interchangeably where the context is clear. Please note that in this work we will identify a single nodeset for all the SITS, under the hypothesis that the land units that characterize a given scene do not change significantly in the time interval covered by the SITS.

An intuitive criterion for selecting S^*, which will be followed in this work, is to pick a single image at a given date in which landscape elements at the finest scale can be identified with better precision. As an example, for the characterization of agricultural systems, one typically expects the object layer to fit the crop scale, and an acquisition at the peak of the growing season is typically selected (Lebourgeois *et al.*, 2017), that is, when landscape heterogeneity is at its maximum. However, even using less *ad hoc* criteria (e.g., selecting an image at random) may have a minor impact on the number and the shape of segments if suitable segmentation techniques are used that allow to control the scale and shape of the segmentation, like in Baatz & Schäpe (2000).

While is certainly possible to define different solutions to the Nodeset Building Problem (e.g., by performing a different segmentation for each image in the SITS), we believe that relying on nodes that correspond to the result of a single segmentation is a more robust choice: indeed, image segmentation is a ill-posed, error prone task in image processing, and the joint use of multiple segmentations is likely to increase the uncertainty on the spatial delineation of landscape elements, which must be unique and nonoverlapping to fit our network model. In general cases, the number of segments cannot be known *a priori* (i.e., it depends on the characteristics of a specific image/scene), but it is always possible to tune the parameters to obtain segmentations at different granularities that will correspond to different orders of magnitude for the obtained segment/node set. For typical cases, we believe that a set of nodes with cardinality in an order ranging from 10^3 to 10^5 represents the ideal trade-off between expressiveness and computational tractability of the network (cf. Section 4).

3.2 Layers identification

Once obtained a nodeset for our multilayer network, the next step of the process is to identify a proper set of layers and, consequently, node memberships to these layers:

Definition 4 (Layers identification problem). *Given a set of nodes V, obtain a nonexclusive clustering of V in k subsets $V_1, ..., V_k$, which will correspond to the nodesets of k different layers $L_1, ..., L_k$.*

In our approach, the identification of the different layers of the network basically consists in grouping landscape elements into subsets each corresponding to a unique "functional" class, in an appropriate land use sense. To clarify this concept, let us leverage a simple landscape model in which, at a given spatial scale, objects have a certain semantically precise identity. Referring to the SITS time span, these may be static (building, roads, rocks, permanent vegetation, etc.) as well as dynamic elements (agricultural crops, water basins, temporary meadows, deciduous forests, etc.), each with a given role in the landscape. If elements in the first group might ideally be identified using a single acquisition, in general, one need to recur to information about how landscape elements evolve in time, in other words, about *how surfaces work*, to provide a reliable grouping (e.g., crops vs. temporary meadows, agricultural water tanks vs. natural basins, buildings vs. greenhouses, etc.).

This justifies our choice of relying on SITS to provide a deeper insight on landscape spatio-temporal dynamics. Therefore, in order to identify the different layers of our network, our choice is to gather information on some temporal profiles that can reliably characterize each segment.

More precisely, the idea is to cluster the segments in subsets containing entities that exhibit a similar temporal behavior according to the SITS (e.g., some radiometric features evolve in a similar way through time), even though they may be belong to different, spatially distant land subsystems in the scene. Note that, in general, once the objects of the scene are identified, such characterization could be obtained using any kind of geospatial data collection (images, maps, GIS) providing a full spatio-temporal coverage of the area under analysis.

Practically speaking, we decide to derive a single radiometric index from each image of the time series S, resorting to a multitemporal pixel-level descriptor that can be spatially averaged over each segment. In this work, we will use the Normalized Difference Vegetation Index (NDVI) as a proper radiometric index (Rouse *et al.*, 1974), a measure widely used to estimate the presence of live green vegetation over the land surface, which can also indirectly provide information about the presence of water or concrete surfaces. While the methodology is designed to be flexible in terms of descriptors, in this case, we preferred to opt for a descriptor based on a standard index, which is also compatible with our case study referring to a tropical agricultural landscape (cf. Section 4). Summarizing, the behavioral descriptor of each segment $v_i \in V$ will be an array of $n = |S|$ elements, containing at position j the average NDVI of v_i at the timestamp s_j.

Once obtained the NDVI array of each segment, the aim is to obtain k subsets $V_1, ..., V_k$, which will correspond to the nodesets of k different layers $L_1, ..., L_k$. Given the unsupervised context, we prefer to model the association of nodes to the different functional classes (i.e., clusters) with a certain level of uncertainty. For this reason, we resort to fuzzy clustering, which allows to model the probability $0 \leq \text{prob}(v_i, L_j) \leq 1$ of each segment/node $v_i \in V$ to belong to layer/functional class L_j. Following this approach, a node $v_i \in V$ will belong to the subset V_j (i.e., it will be present in a layer L_j) if $\text{prob}(v_i, L_j) > \frac{1}{k}$, that is, if its probability to be in that cluster is higher than the one inferred by a uniform random distribution of nodes into the k clusters. Note that, according to this rule, the subsets $V_1, ..., V_k$ are likely to be overlapping, since each node/segment may in general be associated to multiple clusters with sufficiently high probability, resorting to a multilayer network compatible with the model introduced in Definition 1. While any clustering algorithm can be applied at this step, in this work, we resort to *fuzzy c-means* clustering (Ross, 2009), which is known to be particularly suitable for time series data.

3.3 Intralayer edges insertion

Next step to solve the *SITS to Multilayer-Network Problem* is how to define the connectivity between nodes in each layer. We can formally define this problem as follows:

Definition 5 (Intralayer edges insertion problem). *Given a set of nodes V and a set of layers L, identify and weight a proper set of intralayer edges $E_{intra} \in E$, connecting nodes belonging to the same layer, that is, $i = j \: \forall \: (u, L_i, v, L_j) \in E_{intra}$.*

Since segments belonging to a specific cluster may not be adjacent, a straightforward solution to link nodes in a single layer would be to resort to the fully connected graph, since all nodes conceptually belong to the same functional class. However, the objective is here to identify connected areas corresponding to uniform land subsystems, so we here chose to create links only among "close" objects in each layer, leaving the description of wide-range spatial interactions to indirect connections in the intralayer graphs. By the way, this choice is directly inspired to the fundamental principle known as *Tobler's first law of geography*, stating that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970).

To this aim, the basic idea is to recur to the strategy developed in Gaetano *et al.* (2010) in a hard clustering framework, in which for each nonoverlapping subset $V_i \in V$, nodes belonging to the same layer are connected by undirected edges based on the spatial position of the corresponding segments in the original scene, according to the pseudo-nearest neighbor scheme. More in details, for each cluster, segments are projected on the lattice of the segmentation image I^* and labelled

independently according to the original segmentation, then all labels are propagated through the background (i.e., the space covered from segments in the other clusters) with a uniform speed law until all the lattice is filled. Links are then set by computing the Region Adjacency Graph (RAG) over the propagated label image, a RAG being the graph built by placing an edge between each couple of adjacent segments.

In our novel fuzzy clustering framework, this solution would also be viable, since it is still possible to consider, for each layer, the whole set of nodes belonging to it and apply the aforementioned method to set intralayer edges. However, if links are created based on the sole spatial proximity of the corresponding segment, such a solution could prevent the direct connection between nodes belonging to the layer with high probability because of the presence of low-probability segments in-between, hence losing a major structural information. In the case of a highly uncertain clustering, one may even resort to per-layer graphs very similar to the full-image RAG, making the network heavily redundant.

For this reason, in order to create edges by giving priority to nodes with stronger cluster memberships, we opt for an iterative application of the basic method, based on the following two steps:

- Given a layer i, for each segment $v \in V_i$, compute its membership rank $R_i(v)$ to the corresponding cluster (i.e., $R_i(v) = r$ if V_i is the rth most probable cluster);
- For each $r \in 1, \ldots, k$, select only the nodes/segments having $R_i(v) \leq r$, and apply the label propagation strategy to this subset to create edges, skipping already existing connections.

As regards edge weighting, recall that our strategy is to leverage spatial interactions among landscape elements of different functional classes to identify coherent subsystems. To this aim, we refer again to the solution proposed in Gaetano *et al.* (2009), which basically tends to reinforce connections among objects of a single cluster sharing a similar spatial context. Even if the reference work had a more fundamental macro-textural approach, this choice remains appropriate here since in our case the spatial context identifies the local environment in which each element evolves. Matter of facts, this is a key information when aiming at the emergence of object communities related to a unique subsystem.

Therefore, borrowing the definition of spatial context of a segment proposed in Gaetano *et al.* (2009), we will define a spatial context for each node/segment based on its adjacencies (at pixel level) with segments belonging to other layers, taking into account the membership probabilities obtained at clustering time by the *fuzzy c-means*.

In order to obtain such spatial context, we perform a pixel-wise analysis of adjacencies for each segment. Given the set of directions $d \in \{N; NE; E; SE; S; SW; W; NW\}$ (i.e., the possible eight pixels adjacent to a given one), for each node $u \in V$, we first define $\text{pix}_d(u)$ as the set of pixels that are adjacent to pixels of u in the direction d, including pixels in u itself. Then, we compute the transition probabilities toward each subset V_i as:

$$p_d(u, V_i) = \frac{\text{adj}_d(u, V_i)}{|u|}, \tag{1}$$

where $|u|$ is the number of pixels in the segment corresponding to node u, and $\text{adj}_d(u, V_i)$ is the weighted count of pixels in $\text{pix}_d(u)$, with weights equal to the probability of each pixel's reference segment to belong to the subset V_i :

$$\text{adj}_d(u, V_i) = \sum_{q \in \text{pix}_d(u)} \text{prob}(v_q, L_i), \tag{2}$$

Summarizing, we can say that $p_d(u, V_i)$ is a "transition probability" representing the likelihood of finding a pixel belonging to a node $v \in V_i$ when moving from a location belonging to node u in the direction d.

Following this methodology, which basically generalizes the formulation proposed in Gaetano *et al.* (2010) to the case where a fuzzy clustering strategy is used, for each $v \in V$, its spatial context will be described by a matrix P of size $k \times 8$, containing a row for each layer, and a column for each direction d. Once obtained the spatial context P for each node/segment, edges between two nodes in the same layer $(u, v) \in V_i$ are obtained based on the cosine similarity between the transition probability array obtained by flattening the P matrices:

$$w(u, v) = 1 - cos(P_{flat}(u), P_{flat}(v)), \qquad (3)$$

where $P_{flat}(u)$ and $P_{flat}(v)$ are the arrays of length $k \cdot 8$ obtained by flattening the respective P matrices (i.e., horizontally concatenating its rows), and $cos()$ is the cosine distance between such arrays. The weight $w(u, v)$ can thus be interpreted as the likelihood of two segments to belong to the same land subsystem.

3.4 Interlayer edges insertion

As the last step of our procedure, we need to identify a proper set of interlayer edges, that is, edges connecting nodes belonging to different layers:

Definition 6 (Interlayer edges insertion). *Given a set of nodes V and a set of layers L, identify and weight a proper set of intralayer edges $E_{inter} \in E$, connecting nodes belonging to different layers, that is, $i \neq j \, \forall \, (u, L_i, v, L_j) \in E_{inter}$.*

In our network modeling strategy, the interest of interlayer edges is evidently to densify connections among nodes belonging to different functional classes that simultaneously appear on the same portion of land, eventually prioritizing small-scale interactions. To this aim, resorting to the RAG of the original segmentation is a quite straightforward choice, provided that only the links between nodes belonging to different clusters are retained. Therefore, once defined $R = RAG(seg(I^*))$ as the RAG built upon the same segmentation used to build the nodeset V (cf. Section 3.1), we import to the multilayer network all the edges from R connecting segments that correspond to nodes belonging to different layers, that is

$$\forall(u, v) \in R \Rightarrow (u, L_i, v, L_j) \in E_{inter} \, s.t. \, u \in V_i, v \in V_j, i \neq j \qquad (4)$$

As regards interlayer edge weighting, we define two alternative weighting schemes:

- **Fixed weights**, that is, all interlayer edges will have the same weight w_{inter}. For this scheme, the idea is to retain control on the impact of spatial proximity in the community detection step. Reserving a deeper insight on this aspect to future research, here, we identified a good heuristic for w_{inter} as half the average of intralayer weights, that is

$$w_{inter} = \frac{\sum_{\forall e \in E_{intra}} w(e)}{2|E_{intra}|} \qquad (5)$$

- **Context-based weights**, that is, the weight between two layers will be weighted based on the cosine similarity between their global contexts, as done for intralayer ones (cf. Equations 1– 3). The global context of a layer is computed as the one of a segment, but considering all the pixels of all the segments in that layer. This strategy is more appropriate under the hypothesis that a significant number of functional classes belong exclusively to one subsystem.

In the following, we will consider fixed interlayer weights as the default option where not explicitly indicated.

4. Case study: Unsupervised agricultural landscape atratification via multilayer community detection

In order to show a practical use of the SITS-based multilayer network introduced in Section 3, in this section we present an experimental case study concerning the stratification of an agricultural landscape into different subsystems.

The case study is centered on an agricultural landscape around the commune of Koumbia, Province of Tuy, Burkina Faso, located in the west of the country. This area is particularly challenging for our methodology due to its fragmented landscape, characterized by the small size of the agricultural plots, the diversity of their spatial arrangements, their entanglement with areas not belonging to the annual cropland such as pastures and fallows, and the continuous changes that may happen both during the growing season (e.g., different growth stages) and from one season to another (e.g., rotation, plot redistribution).

Our specific objective is to leverage the proposed multilayer network model to automatically provide a stratification of the landscape into different subsystems, focusing in particular on the identification of areas dominated by cropping practices. As hinted in Section 1, our goal here is to show in what measure modeling an SITS into a multilayer network enables the definition of unsupervised network analysis techniques that allow to partition the landscape into functionally homogeneous land use subsystems. As pointed out in Bellón *et al.* (2017), this kind of maps is gaining importance in support of agricultural land use system monitoring, to both optimize field data collection strategies for land use/land cover mapping and enable and support more qualitative analyses (e.g., socioeconomic, demographic, etc.). Moreover, with our network-based approach, the identification of a certain land subsystem will eventually be accompanied to a suitable structural description in terms of components of the system (landscape elements), along with a representation of their functional properties and spatial interactions. For this exploratory work, the objective is to highlight the potential of the approach, so the analysis will be limited to measurements on the structure of the network and communities, as well as to visual inspection of the provided stratification maps.

4.1 Experimental setting

The multilayer network for the case study was obtained using a time series of 23 Sentinel-2 images acquired between January 2016 and December 2016. The sole cloud-free image from August 3, 2016 (Figure 2(a)) was selected as S^*. Preprocessing has been limited to the extraction of the visible (red, green, blue) and near-infrared bands (all at 10-m spatial resolution) to perform image segmentation (cf. Section 3.1). Each segment is therefore attributed using the average NDVI per timestamp as described in Section 3.3. In the following, we will consider a value of $k = 10$ for the layer identification process as the default option where not explicitly indicated. This choice is not arbitrary but is related to the expected number of different temporal vegetation dynamics expected to be observed in the area during the selected time span (cropping season), which was estimated using expert knowledge on the study site. Looking at the Sentinel-2 image in Figure 2(a), it can be seen how a vast and diversified cropland area lies along the main diagonal of the scene, while two natural areas covered with mostly forest and shrubland occupy the rest of the scene (i.e., top-right and bottom-left corners).

For the purpose of community detection over the Koumbia multilayer network, we here use the multilayer version of the state-of-the-art *Infomap* algorithm (De Domenico *et al.*, 2015; Edler *et al.*, 2017) . We chose *Infomap* among other alternatives for several reasons: (i) it is fully compatible with our multilayer network model (i.e., categorical/unordered layers, presence of interlayer edges, weighted undirected edges), (ii) it allows to easily take into account edge weights (i.e., through the use of weighted random walks on the network), and (iii) it is extremely efficient (i.e., being able to run in a few seconds even on large networks).

Table 1. Structural characteristics of the multilayer network obtained on the *Koumbia* area for varying *k*, using fixed interlayer weights (cf. Equation 5). The value of *k* is equivalent to the number of layers |*L*|.

|*L*|		V		E_{intra}			E_{inter}					
	tot	mean ± sd	E	tot	$\overline{w}$	$\overline{deg}$	tot	$\overline{w}$	$\overline{deg}$	D_{intra}	D_{glob}	AL
2	2,246	1123 ± 179.0	9,799	6,011	0.97	5.35	3,788	0.48	3.37	0.005	0.007	1
5	2,246	755.2 ± 218.7	35,788	12,870	0.88	11.46	22,918	0.44	20.41	0.008	0.023	1.7
10	2,246	719.1 ± 242.9	126,517	27,433	0.84	24.43	99,084	0.42	88.23	0.009	0.043	3.2
15	2,246	706.06 ± 239.4	268,825	43,285	0.82	38.54	225,540	0.41	200.84	0.01	0.064	4.7
20	2,246	703.45 ± 273.1	466,222	59,968	0.81	53.4	406,254	0.41	361.76	0.01	0.082	6.3

The python code of the proposed methodology is publicly available online[1]. As regards the c-means clustering phase, we used the python implementation provided by Scikit-Fuzzy[2].

4.2 Structure of the multilayer network

Even though we believe that, in practical cases, a value of $k = 10$ is the most appropriate one for the *Koumbia* study site (cf. Section 4.1), we extracted multilayer networks for varying values of *k* in order to study the impact of this parameter (which controls the number of clusters in the fuzzy clustering phase and, consequently, the number of layers in the network) on the process. Table 1 reports on the structural characteristics of multilayer networks obtained on *Koumbia* for $k = 2, 5, 10, 15, 20$. In the table, $\overline{w}$ refers to average (intralayer or interlayer) edge weight, $\overline{deg}$ to the (intralayer or interlayer) average degree, D_{intra} and D_{glob} report the density of the network (calculated, respectively, as the average intralayer density and as the global multilayer one), and *AL* is the average number of layers per node (i.e., the average number of layers in which a node is present).

It can be noted how the value of *AL* is directly proportional to *k* but, for $k > 2$, the average number of nodes in each layer remains rather stable, with an average value always around 700 nodes per layer. This also implies that the density of the networks also increases with *k*, with the increase being more evident when taking into account only the intralayer edges (i.e., D_{intra} for $k > 10$ is one order of magnitude higher than the one for $k \leq 10$). Another characteristic that remains rather stable is the average weight of intralayer edges (with an exception represented by the average weight of E_{intra} for $k = 2$, which is close to 1). This implies that the fuzzy clustering solutions are coherent when varying the value of *k*, that is, even if the number and population of the clusters varies, the similarity between the spatial context of two given node/segments remains stable. By consequence, the average weight of interlayer edges will be stable too, since it is calculated based on the intralayer ones using the heuristic in Equation 5. As a last remark, it can be observed how the average interlayer degree is always higher than the intralayer one: this is indeed largely expected, since the segments adjacent to a given one in the *RAG* (cf. Section 3.4) may be present in different layers, thus resulting in more than one interlayer edge for each couple of adjacent segments (i.e., since the corresponding nodes are likely to be present in several layers).

4.3 Experimental results

In this section, we present the results of the experiments regarding our unsupervised agricultural landscape stratification task via multilayer community detection.

Figure 2(a) and (b) report the reference image of the *Koumbia* study site used for segmentation (in RGB composition, acquired on August 3, 2016 over the study site) along with a reference mapping of the cultivated surfaces. This cropland map has been obtained using a supervised classification algorithm using the methodology presented in Lebourgeois *et al.* (2017) starting from a
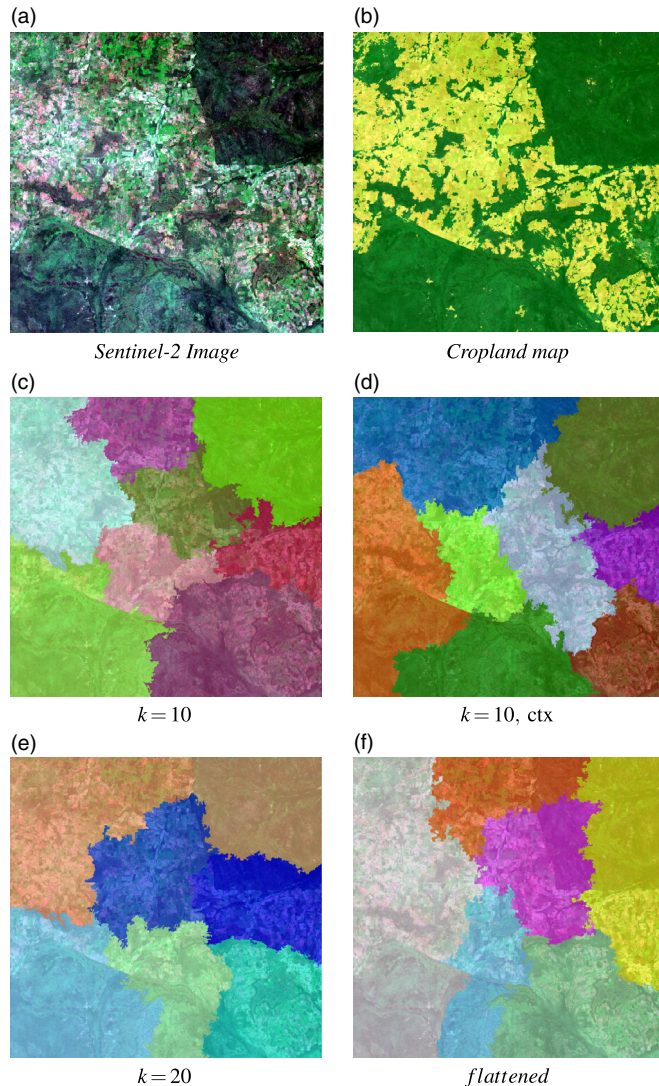
Figure 2. Landscape stratification maps on *Koumbia* (cf. Sentinel-2 image (a) and the corresponding cropland map in (b)), by running *Infomap* on the network with $k = 10$ with fixed (c) and contextual (d) interlayer weights, $k = 20$ with fixed interlayer weights (e) and on the flattened network (f). Different colors correspond to different communities.

larger remote sensing dataset including very high spatial resolution imagery (SPOT6/7) and decametric optical image time series (Sentinel-2 and Landsat 8) and has an estimated overall accuracy of 95%. For this area, noncrop class only contains natural vegetation.

In order to provide a quantitative evaluation of the obtained community solutions, we will use an overlap measure O, aimed at quantifying which percentage of each community lays on a single subsystem (i.e., natural areas or cropland) when considering the cropland map in Figure 2(b). More precisely, for each community, once calculated the percentage (i.e., number of segments over the total ones in the community) which lay on each subsystem, the value of O for that community is given by the maximum percentage. Table 2 reports on the results of the multilayer community detection on the *Koumbia* study site, by running *Infomap* on multilayer networks obtained for $5 \le k \le 20$ using fixed interlayer weights, using contextual interlayer weights (for $k = 10$, ctx), and on the flattened single-layer network. For each solution, we report the number of obtained communities, the multilayer modularity $\mathcal{Q}$ (calculated as proposed in Mucha *et al.* (2010)), and the maximum, minimum, and mean value of O. It can be noted how the number of communities for this set of solutions varies in a limited range, with the solution referring to

Table 2. Results of the multilayer community detection on the *Koumbia* study site by running *Infomap* on: multilayer networks obtained for $5 \leq k \leq 20$ using fixed interlayer weights, using contextual interlayer weights (for $k = 10$, ctx), and on the flattened single-layer network.

Network	#communities	$\mathscr{Q}$	max_O	min_O	$mean_O \pm sd$
$k = 2$	9	0.39	0.99	0.50	0.67 ± 0.15
$k = 5$	6	0.38	0.80	0.55	0.67 ± 0.084
$k = 10$	8	0.26	0.99	0.55	0.72 ± 0.15
$k = 15$	8	0.20	0.99	0.56	0.70 ± 0.14
$k = 20$	7	0.17	0.99	0.62	0.75 ± 0.13
$k = 10$, ctx	8	0.26	0.99	0.51	0.70 ± 0.17
flattened	6	0.43	0.84	0.54	0.66 ± 0.10

$k = 5$ and to the flattened network producing the lowest number of communities (6), and the solution corresponding to $k = 2$ producing the highest one (9). While the higher values of multilayer modularity are obtained by the flattened network (0.43), followed by $k = 5$ (0.39) and $k = 2$ (0.38), qualitative analysis based on overlap values and on the landscape stratification maps (cf. Section 4.3.1) will show how such solutions do not actually correspond to the best ones for the targeted task. Interestingly, the solutions corresponding to $k = 5$ and to the flattened network, even though showing a relatively high modularity, are the only ones which are not able to produce communities that are centered on a single subsystem, that is, they are the only ones with a max_O considerably lower than 0.99. As regards average performances, the best values of $mean_O$ correspond to $k = 10$ (0.72) and $k = 20$ (0.75): qualitative analysis of the landscape stratification maps will in fact confirm the solution corresponding to $k = 10$ as the best one. In order to get an insight in such overlap values, Figure 3 shows the bar charts reporting the percentage of overlap of each community on the two land subsystems (i.e., Cropland and Natural Areas) for solutions obtained with $k = 10$ (a), $k = 10$ with contextual interlayer weights (b), $k = 20$ (c) and on the flattened single-layer network (d). It can be observed how the solutions corresponding to $k = 10$ and $k = 10$, ctx (Figures 3(a) and (b)) contain two communities that are almost completely centered on the *Natural areas* subsystem (i.e., communities 2 and 7 for the former, and communities 6 and 7 for the latter); while the *Cropland* subsystem seems to be more difficult to separate, in both cases several communities can be observed which are strongly centered on this subsystem (e.g., communities 0 and 4 in both charts). Similar observations can be drawn for $k = 20$ (Figure 3(c)), where communities 4 and 6 are well centered on *Natural areas*, and community 0 has a significant overlap with the *Cropland* subsystem. Conversely, the solution obtained on the flattened network (Figure 3(d)) seems to correspond to communities that are not well separated between the two subsystem (with an exception for community 3 which has a significant overlap with *Natural areas*).

4.3.1 Landscape stratification maps

Figures 2(c)–(f) show the results of the *Infomap* community detection over the provided multilayer network in different configuration settings: Figure 2(c) shows the 10-layer network with fixed interlayer weights, Figure 2(d) shows the 10-layer network with contextual interlayer weights, Figure 2(e) shows the 20-layer network with fixed interlayer weights, and Figure 2(f) shows the flattened single-layer network.

The solution obtained on the 10-layer network with fixed interlayer weights (Figure 2(c)) includes eight communities (corresponding to eight random colors in the figure). It can be seen how the natural areas subsystems (the two green communities on the top-right and bottom-left corners) have been mostly separated from the cropland subsystem, which in turn corresponds to six different communities. Summarizing, we can say that the identification of the three main disjoint subsystems has been performed with good precision, with the exception of some borderline
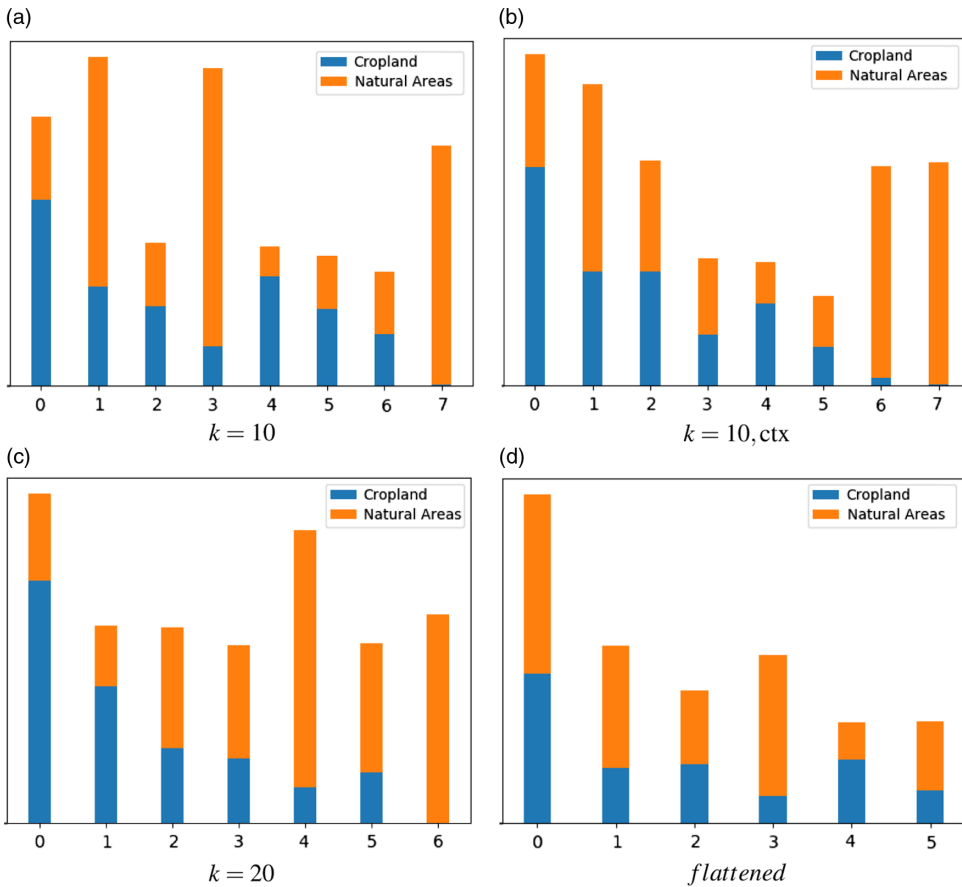
Figure 3. Bar charts reporting the percentage of overlap of each community on the two land cover classes (i.e., Cropland and Natural Areas) for solutions obtained with $k = 10$ (a), $k = 10$ with contextual interlayer weights (b), $k = 20$ (c) and on the flattened single-layer network (d). Size of the bars is proportional to the spatial extent of each community.

forest areas that have been erroneously included in cropland communities (i.e., the portion of the bottom-left forest included in the rightmost violet community). On the other hand, cropland area has been partitioned in six different communities. Conceptually, all these communities can be seen as belonging to a unique crop subsystem, hence in a strict land use sense this may be interpreted as a overpartitioning. However, borders among detected communities seem to grossly follow the transitions among slightly different landscapes, for example, in terms of type of soil (calcareous as in the center-left community in light green), crop vegetation density (center-top community in dark red), or the presence of dense noncrop vegetation (rightmost community in red).

As regards the results obtained on the multilayer network with contextual interlayer weights (Figure 2(d)), the provided map is still mostly consistent with the boundaries among the two, subs, similar in scale with respect to the result in Figure 2(c), but a larger error appears on the orange community. This motivates us, for future works, to investigate more on contextual weights, improving the solution proposed in this work (an interlayer weight for each couple of layer, based on a global context, cf. Section 3.4), in order to provide interlayer contextual weights based on the comparison between portions of each couple of layers (e.g., connected components). Regarding the result obtained starting from a 20-layer networks, here the results are the most unsatisfactory, with five out of the six detected communities spanning over different subsystems. This is mostly

due to the fact that, when raising the number of initial cluster, uncertainty on the definition of layer increases, and spatial adjacencies (i.e., interlayer edges) become dominant.

In order to show the importance of using a multilayer network model, we also show the community solution obtained on the flattened single-layer network (Figure 2(f)), that is, the network obtained merging all the k layers in a single one (which can also be seen as the original RAG obtained on s^*, cf. Section 3.4). It is easy to see how in this case the community structure, which includes seven communities, does not reflect at all the separation between cropland and noncropland areas, confirming how the expressiveness of a multilayer model can be beneficial in this kind of task.

Summarizing, the case study confirmed the benefits coming from the multilayer network model with respect to the application of an unsupervised landscape stratification technique. However, the multilayer network with context-based interlayer weights did not perform better than the one with fixed weights as expected, which underline the need for a further investigation on the definition of a suitable interlayer weighting strategy.

5. Conclusions

In this work, we proposed a completely unsupervised methodology to build a Multilayer Network from an SITS. Standard segmentation techniques are used to build a nodeset from the SITS, then layer memberships and edge weights are inferred based on the temporal radiometric profile of each segment. To best of our knowledge, this is the first work to explore the possibility to exploit this model in the remote sensing domain. In order to test the potential of the proposed model, we present an experimental case study concerning an unsupervised landscape stratification task via multilayer community detection on a tropical agricultural landscape (i.e., the Koumbia area in Burkina Faso). The experimental results confirmed how the use of a multilayer network model can be beneficial in this domain, paving the way for the application of unsupervised network analysis techniques.

Given the exploratory nature of this work, there is plenty of room for future work, aiming at improving and extending the proposed method. A first planned improvement regards the definition of sliding-window-based approaches that would allow for a richer representation of the spatial context of each segment. Another improvement concerns interlayer edge weighting, through the exploration of contextual weighting schemes able to take into account local portions (e.g., connected components) of each couple of layers. As regards the unsupervised landscape stratification task, the definition of a multilayer community detection algorithm specifically adapted to this task may be the object of a future work.

From a broader perspective, the expressive power of the model could be enhanced by coupling the layers issued by spatial characteristics of the SITS described in this work, with layers based on interaction graphs (cf. Section 2.1). This would result in a multilayer network able to represent, in a single integrated object, spatial characteristics and dynamics of the landscape together with the evolving dynamics of the corresponding land use system.

Conflict of interest. The authors have nothing to disclose.

Notes

1 https://gitlab.irstea.fr/raffaele.gaetano/geo2net
2 https://scikit-fuzzy.readthedocs.io

References

Baatz, M., & Schäpe, A. (2000). Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Angewandte geographische informationsverarbeitung xii*, *58*, 12–23.

Bégué, A., Arvor, D., Lelong, C., Vintrou, E., & Simoes, M. (2015). Agricultural systems studies using remote sensing. In P. S. Thenkabail (Ed.), *Land resources monitoring, modeling, and mapping with remote sensing (remote sensing handbook, 2)* (pp. 113–130). Boca Raton: CRC Press.

Bellón, B., Bégué, A., Lo Seen, D., de Almeida, C., & Simões, M. (2017). A remote sensing approach for regional-scale mapping of agricultural land-use systems based on NDVI time series. *Remote Sensing*, *9*(6), 600.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16.

Bousquet, F., Bakam, I., Proton, H., & Le Page, C. (1998). Cormas: Common-Pool Resources and Multi-agent Systems. *Proceedings of the International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE)*, Vol. 2, Lecture Notes in Computer Science 1416, 826–837.

Chen, G., Weng, Q., Hay, G. J., & He, Y. (2018). Geographic object-based analysis (geobia): emerging trends and future opportunities. *Giscience & Remote Sensing*, *55*(2), 159–182.

Clarke, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design*, *24*(2), 247–261.

Costanza, R. (1987). Simulation modeling on the macintosh using stella. *Bioscience*, *37*(2), 129–132.

De Domenico, M., Lancichinetti, A., Arenas, A., & Rosvall, M. (2015). Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys. Rev. X*, *5*(March), 011027.

DeFries, R., & Rosenzweig, C. (2010). Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences*.

Degenne, P., & Lo Seen, D. (2016). Ocelet: Simulating processes of landscape changes using interaction graphs. *Softwarex*, *5*, 89–95.

Degenne, P., Lo Seen, D., Parigot, D., Forax, R., Tran, A., Ait Lahcen, A., Curé, O., & Jeansoulin, R. (2009). Design of a Domain Specific Language for modelling processes in landscapes. *Ecological Modelling*, *220*(24), 3527–3535.

Degenne, P., Ait Lahcen, A., Curé, O., Forax, R., Parigot, D., & Lo Seen, D. (2010, July). Modelling the environment using graphs with behaviour: Do you speak Ocelet? *iEMSs 2010, 8 pages*.

Edler, D., Bohlin, L., & Rosvall, M. (2017). Mapping higher-order network flows in memory and multilayer networks with infomap. *Corr.*, abs/1706.04792.

Gaetano, R., Scarpa, G., & Sziranyi, T. (2010). Graph-based analysis of Textured Images for Hierarchical Segmentation. In *British Machine Vision Conference, BMVC 2010*, *47*(7–2): 2129–2141.

Gaetano, R., Scarpa, G., & Poggi, G. (2009). Hierarchical texture-based segmentation of multiresolution remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*.

Guttler, F., Ienco, D., Nin, J., Teisseire, M., & Poncelet, P. (2017). A graph-based approach to detect spatiotemporal dynamics in satellite image time series. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D. & Rodes, I. (2017). Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, *9*(1), 95.

Interdonato, R., Ienco, D., Gaetano, R., & Ose, K. (2019). Duplo: A dual view point deep learning architecture for time series classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *149*, 91–104.

Jahel, C., Baron, C., Vall, E., Karambiri, M., Castets, M., Coulibaly, K., Bégué, A., & Seen, D. L. (2017). Spatial modelling of agro-ecosystem dynamics across scales: A case in the cotton region of west-burkina faso. *Agricultural Systems*, *157*, 303–315.

Jahel, C., Vall, E., Rodriguez, Z., Bégué, A, Baron, C., Augusseau, X., & Seen, D. L. (2018). Analysing plausible futures from past patterns of land change in west burkina faso. *Land Use Policy*, *71*, 60–74.

Khiali, L., Ienco, D., & Teisseire, M. (2018). Object-oriented satellite image time series analysis using a graph-based representation. *Ecological Informatics*, *43*, 52–64.

Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Mutilayer networks. *Journal of Complex Networks*, *2*(3), 203–271.

Kolecka, N., Ginzler, C., Pazur, R., Price, B., & Verburg, P. H. (2018). Regional scale mapping of grassland mowing frequency with sentinel-2 time series. *Remote Sensing*, *10*(8), 1221.

Lebourgeois, V., Dupuy, S., Vintrou, E., Ameline, M., Butler, S., & Bégué, A. (2017). A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated sentinel-2 time series, VHRS and DEM). *Remote Sensing*, *9*(3), 259.

Leenhardt, D., Angevin, F., Biarnès, A., Colbach, N., & Mignolet, C. (2010). Describing and locating cropping systems on a regional scale. A review. *Agronomy for Sustainable Development*, *30*(1), 131–138.

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, *328*(5980), 876–878.

Ross, T. (2009). Fuzzy logic with engineering applications. *Fuzzy Logic with Engineering Applications* (3rd ed.), 01.

Rouse, Jr., J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring Vegetation Systems in the Great Plains with Erts. *Nasa Special Publication*, *351*, 309.

Scarpa, G., Gaetano, R., Haindl, M., & Zerubia, J. (2009). Hierarchical multiple Markov chain model for unsupervised texture segmentation. *IEEE Transactions on Image Processing*, *18*(8), 1830–1843.

Schmidt-Laine, C, & Pave, A. (2002). Environment: Modelling and models to understand, to manage and to decide in an interdisciplinary context. *Natures Sciences Societes*, *10*(01), 5–25.

Tobler, W. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, *46*(2), 234–240.

Verburg, P H., van de Steeg, J., Veldkamp, A., & Willemen, L. (2009). From land cover change to land function dynamics: A major challenge to improve land characterization. *Journal of Environmental Management*, *90*(3), 1327–1335.

Verburg, P. H., Mertz, O., Erb, K.-H., Haberl, H., & Wu, W. (2013). Land system change and food security: Towards multi-scale land system solutions. *Current Opinion in Environmental Sustainability*, *5*(5), 494–502. Human settlements and industrial systems.

Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, *4*, 22–40.

Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36.