A New Biodiversity Composite Indicator Based on Anthropentropy and Forest Quality Assessment

Framework, Theory, and Case Studies of Italian Territory

Silvia Assini
Dept. of Earth and Environment Sciences
University of Pavia
Pavia, Italy
silviapaola.assini@unipv.it

Maria Grazia Albanesi
Dept. of Electrical, Computer and Biomedical Engineering
University of Pavia
Pavia, Italy
mariagrazia.albanesi@unipv.it

***Abstract*— This paper describes a new environmental composite indicator, which is based on two previously defined single indicators, related to land use (the Anthropentropy Factor) and to quality of forests (the Forest Status Quality Indicator). The framework for the definition of the composite indicator is an innovative formalization of the multidisciplinary approach, which connects knowledge and expertise of two different scientific fields: vegetation science and computer science. The proposed method for the indicator computation combines the classical algorithms of computer vision, to process data from Geographic Information Systems, and the phytosociological approach, to assess the floristic composition of the forests. The goal is to build a deep knowledge about the impact of land use and forest quality, at a landscape level, on biodiversity conservation, by studying the impact of anthropic activities, both inside (urban and rural areas) and outside (forests) the areas occupied by human activities. The knowledge is expressed by a single composite indicator and its assessment can be used for environmental preservation policy actions, to guide local government decisions for a biodiversity conservation in the landscape. The new indicator and the methodological approach is validated by presenting experimental results on two case studies in the North-West of Italy.**

Keywords - biodiversity; land use; environmental indicator; forest status quality; Anthropentropy Factor; composite indicator.

I. INTRODUCTION

This paper describes the continuation and the extension of the research project reported in [1]: theory have been improved and the number of case studies has been increased, in order to validate the original ideas and to enlarge the scope of applications of the proposed approach. The target is still the same: to propose and measure indicators related to some specific aspects of biodiversity, within a given territory. Biological diversity means the variability among living organisms from all sources including, *inter alia,* terrestrial, marine, and other aquatic ecosystems and the ecological complexes of which they are part: this includes diversity within species, between species and of ecosystems [2]. Actually, the habitat loss and fragmentation due to agriculture and urbanization, the introduction of alien species and the climatic changes are among the most important causes of biodiversity loss. Thus, the recent Strategic Plan for Biodiversity 2011-2020 and the Aichi Biodiversity Targets [3] include, among the others, two important goals: to reduce the direct pressures on biodiversity and to improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity. The vision of this Strategic Plan is a world of “living in harmony with nature,” where “by 2050, biodiversity is valued, conserved, restored and wisely used, maintaining ecosystem services, sustaining a healthy planet and delivering benefits essential for all people.” The goal of this work is not only to give a quantitative description of the biodiversity, but also to identify the sub-areas, within the territory under investigation, where the application of policy actions for its conservation is a more serious and pressing issue. In particular, two specific aspects of environmental preservation have been considered in [1]: land use and forest quality. In order to motivate this fundamental choice, it is important to understand the relationships between these two aspects and biodiversity conservation.

Land use can be defined in different ways, and for this reason its meaning is often a source of misunderstanding. In our research, we adopt its broadest sense: the definition and classification, within a limited territory, of the areas of *anthropic places*, i.e., places occupied by human activities (for every-day life, economic and productive activities), and of the areas of wild nature. According to this definition, the territory under investigation is *de facto* partitioned into two classes: the areas of anthropic activities, where the human presence is fairly continuous and has ousted the wild, and the “wild areas”, where the opposite occurs. A correlated term is *land take*: it expresses the variation of the land use over time, (e.g., one year for *annual land take*).

The limitation of land use against the threat of an exceeding, out of control, urban sprawl and the mitigation of the loss of wild habitats are very important for biodiversity conservation. This is evidenced by the fact that the European Environment Agency considers land use and biodiversity in the same target and objective for the policy actions of 2010-2050 decades [4]. The relationship between land use and biodiversity conservation can be motivated also referring to the well-known DPSIR framework [5], which describes the impact of human activities on the environment as a chain of causes-effects. The chain connects five rings: the Driving forces (D), the Pressures (P), the State (S), the Impacts (I), and Responses (R). If we use this framework to investigate

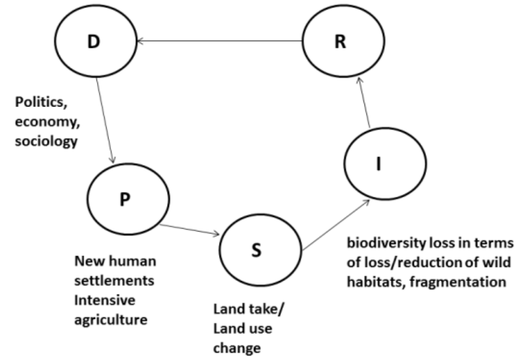


Figure 1. The DPSIR framework and the relationships between land take and biodiversity loss (D: driving forces, P: pressures, S: state, I: impacts, R: responses).

the possible relationships between land use changes and biodiversity loss (see Fig. 1), we can identify examples in which the driving forces, which are the starting point of the chain, are the same for both the phenomena, or, at least, are closely related. For example, social motivations (D), such as demographic expansions, generate pressures on the environment in order to increase the areas occupied by new settlements, roads, and services (P). As a consequence, the environment enters in a new state S, where land use changes over time, with an increase of anthropic areas, at the expense of natural, wild areas. In turn, this state change has two possible negative impacts on biodiversity: (a) the loss of the territory, both for vegetal and animal species, and (b) the territory fragmentation, which is considered as one of the most dangerous threats to biodiversity [6], [7]. Demographic expansion is one of the most simple and intuitive examples of pressures that can act as a strong motivation for both land use changes/land take and biodiversity loss. There are plenty of examples that lead to the loss of wild natural areas: urban and rural expansion, new roads and communication lines, settlements for industries, tourism and services, intensive farming. Moreover, processes that cause land use change are specific for the different parts of Europe [8]: forest management affects the Boreal and Alpine regions, abandonment and intensification are mainly encountered in the Mediterranean areas, urbanization and drainage are typical of the Continental and Atlantic regions. Concerning Italy, further studies [9], [10], [11] identified a distinction between the planar belt, where urbanization and agricultural intensification are the main pressures of biodiversity loss, and the hilly-montane belt, which is more affected by abandonment and the consequent forest re-colonization.

If we consider the problem of land use from a quantitative point of view, the most complete research regarding Europe is the Corine Land Cover Project [12]; the artificial areas (intended as soil sealed territories) cover only the 4% of the land in Europe, as compared to a 34% of forests. However, this percentage rises to a global value of 51%, if we consider the areas that support all the anthropic activities, the economic growth and food production (agriculture, crops, pasture and semi-natural vegetation). The situation is particularly dramatic if we consider the phenomenon of land take over the past years. For example, in Italy, the most recent report on land take [13] shows an annual value of 7.3%, equal to 21.890 square kilometers, the equivalent of 70 ha/day, or 8 square meters/s.

The second aspect of environmental preservation considered in [1] is the presence, in a given territory, of forests, and their quality, expressed by some quantitative measure. This is consistent with the chosen definition of land use: as the territory is partitioned into two areas, i.e., anthropic places vs. natural, wild areas, we are interested into the study of one aspect (at least) referring to each of the partitions: land use is related to the first one (anthropic places), the presence of forests to the second one (wild areas). We have chosen the forests for their importance and benefits for the human well-being (the so-called ecosystem services), such as flood prevention, erosion control, CO_2 absorption, climate regulation, *refugium* function for wild plants and animals, recreation, science and education.

Moreover, even if this partition is clearly visible on the territory, the two worlds are far to be completely separated. In particular, the quality of the forest can be influenced by human activities, such as pollutions, climate changes, use of pests and the introduction of invasive species. In other term, it is important not only the quantitative presence of forests, but also their quality [14], [15]. Indeed, in the “wild area”, due to the sporadic presence of humans (for example, for agriculture), old-growth forests, with a high degree of biological richness, are sometimes replaced by young forests with poor species.

Also, intensive forests and plantations are far to be “true wildness”, even if they are clearly outside the anthrophized areas. For all these reasons, it is important to give a quantitative measure of the quality of the forests, as the simple measure of the area percentage is a too simplistic parameter to understand the actual contribution the forests give to the biodiversity of the territory.

A. *Novelties of this contribution*

In previous researches, we have proposed new indicators, both for land use measurement [16], [17], i.e., the *Anthropentropy Factor* (*AF*), and for the forest quality assessment [1], i.e., the *Forest Quality Status Indicator* (*FSQ*). In addition, a first attempt to study the correlation between the two indicators [1], with regard to a precise area of the North-West of Italy, the Province of Pavia, has become our first case study. However, it is evident that the distinct computation of the two indicators, although promising and interesting, does not provide an overall view of the impact of the two critical aspects on the biodiversity of a given territory. Therefore, the research continued with the challenging goal to define a new composite indicator, based on both *AF* and *FSQ*, in order to give a more significant measure of the state of the biodiversity inside a given territory, at least for the two major phenomena, i.e., land use and forest quality. This can be considered a relevant improvement on the state of the art; in fact, biodiversity can be measured [18] by considering different parameters (composition, structure or function) and at different levels of biological organization (genetic, population-species, community-ecosystem, landscape). Usually, the produced indicators in literature refer only to *one* of such organization levels. On the contrary, the composite indicator here built relates, at the same time, to *two* organization levels: landscape (considering urban and forest patches) and community-ecosystem level (considering forests and their quality).

Moreover, the main novelties carried out with this new research activity, if compared to the previous one [1], can be summarized as follows:

1. The definition of a new composite indicator, called *Biodiversity Composite Indicator* (*BCI*). It is based on the two "simple" indicators, i.e., *AF* and *FSQ* that become, in turn, the two sets of variables on which the composite indicator is built on. Building a composite indicator is not a simple matter of crunching numbers, but a clear stated and well-defined method has to be followed [19]. For this purpose, it is very important to define a theoretical framework (see point 2).
2. The formalization of multidisciplinary approach, which is the core of the theoretical framework. The goal of the formalization is to describe how the two sciences, vegetation science and computer science, can cooperate and under what constraints.
3. A new case study has been proposed, and the comparison with the first one [1] has been fully investigated, by considering the two indicators *AF* and *FSQ*.
4. The new *BCI* indicator has been calculated for both the two case studies, in order to give a global assessment on the environmental issue of biodiversity conservation for the two target areas.

The paper is organized as follows. Section II describes the multidisciplinary approach and the new conceptual framework for the definition of the composite indicator. Section III summarizes the theory about the two indicators and presents the results for the two case studies. Section IV describes the fundamental steps for the definition of the composite indicator and shows the experimental results on the two case studies on the Italian territory. Conclusion and considerations about future work in Section V close the article.

II. THE MULTIDISCIPINARY APPROACH

One of the most important and distinctive characteristics of our research is its highly multidisciplinary approach, which bridges across two important fields of our modern scientific research: vegetation science and computer science. Both of them have knowledge, tools and paradigms that are able to assess the impact of human activities for a sustainable future. In particular, we have identified three hierarchical levels of "integration": knowledge, methods (or paradigms) and data (see Fig. 2) .

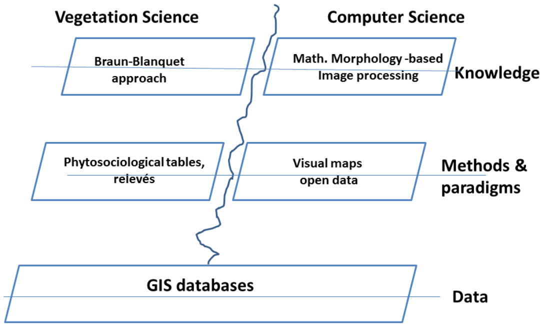


Figure 2. The formalisation of the framework of cooperation between computer science and vegetation science.

At the first level, we find the concepts, approaches and algorithms useful to process data and information at the lower levels. In particular we have, for computer science, the mathematical morphological operators applied in 2D digital images processing of computer vision [20]. They have been used for the computation of the AF indicator. Concerning vegetation science, we use the Braun-Blanquet approach [21], [22], [23], [24] to determine in the FSQ definition the floristic composition of the forests and to assess their quality from the authenticity perspective [14]. Authenticity is a measure of the health and integrity of the ecosystems; it can be assessed considering different aspects, and we have chosen the composition (number of layers, presence and percentage of alien and protected species) and the continuity (in particular, we have considered the areas occupied by the forests.)

At the second level of the hierarchy of our framework, we find the methods and the paradigms used to translate the "knowledge" level in practical tools, in order to process the lower levels (data). For vegetation science, we used the phytosociological tables [25] and, in some cases, information have been integrated by bibliographic references and phytosociological *relevés*, collected in the area where the forest type occurs. For computer science, we have used Quantum GIS [26], an open source Geographic Information System (GIS), and its primitives to compute areas and intersect boundaries in the territories of the case studies.

At the lowest, third level of the hierarchy, we find the raw data, expressed in a visual form (maps of GIS systems). It consists of three databases: the ERSAF (Ente Regionale Servizi Agricoltura e Foreste) database [27], the Corine Land Cover database [12], and the ISTAT (Istituto Nazionale di Statistica) database of Italian administrative boundaries [28]. We consider the portion of these databases, which refers to the same territory, i.e., the region Lombardia in the North-West of Italy. The first two databases are used to represent the presence of forests and the land use classification, respectively. The third one is used to focus the attention on the two case studies, namely two of the twelve provinces of region Lombardia, i.e., the provinces named "Pavia" and "Lodi". Some interesting characteristics of the territory of Lombardia and of the two case studies are reported in Table I. The level of GIS databases is the conceptual bridge, which connects vegetation science and computer science, the channel through which the two disciplines can communicate and exchange information and knowledge. In order to assure that this communication is valid and generate useful and meaningful information, some constraints are to be respected. We have defined five constraints, described as follows.

Format constraint: data are expressed in a common format: in this way, the visual maps of the same territory (geodata) can be processed by GIS tools. To adhere to this constraint, the ERSAF database (raster GIS data) has been vectorized, in order to be consistent to the other vector geodata.

Temporal constraint: data stored in the databases have to be referred to equal or very close temporal periods. In our case, data of Lombardia refer to the period 2007-2011.

Granularity constraint: the different databases has to consider the same data granularity as reference in the different computations. As our databases refer to a geographical territory, adopting the same data granularity means that both the indicator computations refer to the same geographical unit. In our research, we have chosen the municipality as common data granularity. This choice is motivated by the fact that, in Italy, the municipality is the administrative division, which is in charge of adopting local policy on the territory for land use and/or environmental requalification. Therefore, it is important to consider this granularity if we want to use the indicators to support decision-makers.

Precision constraint: in all the indicator computations and geometrical transformations (vectorization, change of coordinates), we have a maximum error less of 50 meters, which is consistent to the definition of the most important aspect of the *AF* indicator computation, the dilation step (see Section III).

Availability constraint: all the GIS databases are public and available according to the open access paradigm. This choice is particularly important in the construction of the composite indicator, as underlined in Section V.

Once we have formalized the framework, we can use it to explain how to create the composite indicator, starting from the two "simple" ones, i.e., *AF* and *FSQ*. The framework and the method here proposed are quite general, and can be applied also adding more than two simple indicators. In the following sections, a brief recall of the theory of the two indicators is summarized. Moreover, data of the two case studies are reported and compared.

III. THE SIMPLE INDICATORS

In this section, the two simple indicators are presented, for land use and forest quality evaluation, respectively. In Section II.A, the flow chart of the algorithm for land use estimation is described, with some fundamental considerations about the computer science paradigms and tools, and their innovative aspects in this research field.

TABLE I. REPRESENTATIVE DATA FOR THE TERRITORY UNDER INVESTIGATION: THE REGION OF LOMBARDIA, NORTH WESTERN ITALY, AND THE TWO CASE STUDIES, PAVIA AND LODI.

Territory	Main characteristics	
Region Lombardia	Area	23.844 square km
	Minimun Altitude above sea level	11 m
	Maximum Altitude above sea level	4.021 m
Case study °1: Province of Pavia	Area	2.965 square km
	Number of municipalities	190
	Average Minimun Altitude above sea level	53 m
	Average Maximum altitude above sea level	951 m
Case study °2: Province of Lodi	Area	782 square km
	Number of municipalities	61
	Average Minimun Altitude above sea level	40 m
	Average Maximum Altitude above sea level	101 m

In Section II.B, the forest quality indicator is defined by following the same theoretical framework.

A. Land use indicator

The Anthropentropy Factor [16], [17] is an environmental indicator of type B, using the standard European Environmental Agency taxonomy [5]. The AF expresses in an absolute, continuous scale, from 0 to 1, the degree of the impact of anthropic human activities due to the land use. Furthermore, it is a performance indicator; in fact, besides its definition, also a metric is given, which describes a table of reference to assess the environmental situation, in an optimum or near-optimum condition.

The metric maps the *AF* values to five intervals (class of land use), where only the first one is very desirable, the second one is near-optimum, until the last one, which refers to the worst situation of irreversible environmental degradation. In Table II, the metric of the *AF* indicator is described. For a reasoned treatment of the metric and its relationship with a possible policy making for a sustainable development, see [17]. Here, we recall the basic definition and concepts that are essential to understand the metric and analogies and differences between this indicator and the *F*orest *S*tatus *Q*uality Indicator (see Section III.B).

Anthropentropy is a neologism, from the ancient Greek term *Anthropos* (*Ἄνθρωπος*) = "man", and *entropy*, that, in turn, derives from the ancient Greek terms *en* (ἐν) = "inner", and *tropé* (τροπή) = "transformation".

TABLE II. THE METRIC ON THE *AF* INDICATOR FOR LAND USE.

Class of land use and map color	**Evaluation of Land Use**	
	Intervals of AF	***Meaning***
1 light green	$0 <= AF <= 0.2$	Very low level of anthropentropy, *ideal* situation for nature and human beings
2 green	$0.2 < AF <= 0.4$	A first worrying level of anthropentropy, but the situation is still *good*
3 yellow	$0.4 < AF <= 0.6$	A serious level of anthropentropy, with a beginning negative impact of anthopization on the environment.
4 red, light violet	$0.6 < AF <= 0.8$	A very serious level of anthropentropy, with a great negative impact of anthopization on the environment.
5 violet, black	$0.8 < AF <= 1$	The worst situation, with an irreversible environmental degradation.

In fact, the *AF* indicator means to express the transformations and the consequent “disorder” introduced in natural ecosystems by the presence and disturbance of human beings. The algorithm for *AF* computation is described in the block-diagram in Fig. 3. The first step of the algorithm is the identification of a delimited part of a geographic territory under consideration and the computation of its *area S*, in squared kilometers. As previously discussed in the general framework, data granularity refer to the Italian municipalities; thus, in our computation, S is the area of a given municipality the indicator refers to.

The second step for the *AF* computation algorithm is the identification of all the land parcels occupied by anthropic places, such as human settlements, factories, roads, and so on (for a complete taxonomy of the anthropic places, see [15]). We call these parcels *anthropic sub-regions*. This is performed by using the Corine Land Cover database [10] and primitives of GIS software to extract, intersect and subtract areas that are identified by the 44 classes of land use of Corine Land Cover project.

The third step takes into consideration the shapes and the contiguity of the anthropic sub-regions. In fact, each sub-region is geometrically enlarged by the morphological operator of dilation [20] (along both the two Cartesian dimensions X and Y) with a factor of "buffering" (radius of the circular dilation) of 50 meters, to give rise to anthropic sub-regions. The choice of a 50 meters distance/limit has been discussed in [16], and it seems a good compromise between a too restrictive and a too permissive limit. After performing the dilation on each of the anthropic sub-regions of the municipality, the union of all of them is taken and it corresponds to the *Death Zone*. Let define *DA* as the area (in square kilometers) of the Death Zone. We think that this step of the algorithm makes the AF indicator a “true” naturalistic evaluation of land use, because it not only compute the percentage of land occupied by human activities, but also takes into consideration the *shape* of the areas subtracted to nature, and their relative positions, thus incorporating the important aspect of land fragmentation and its impact on biodiversity.

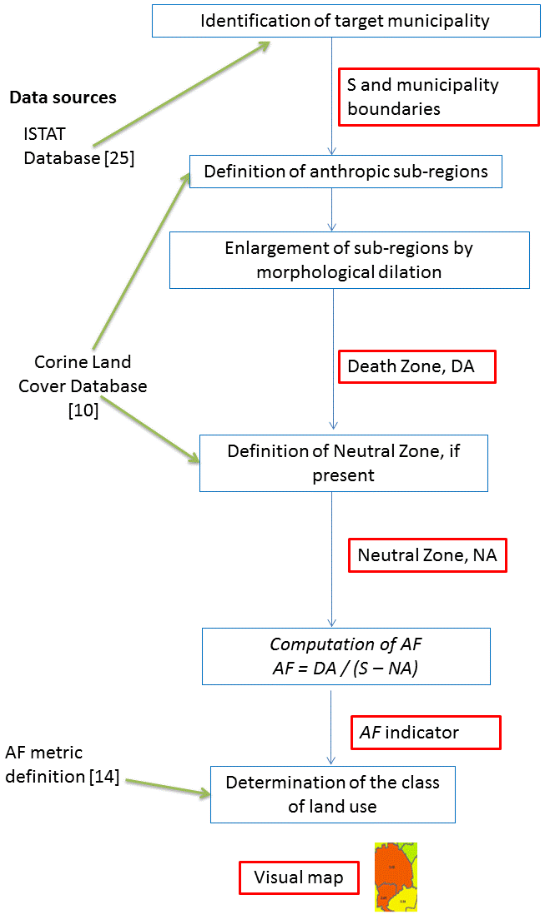


Figure 3. The flow chart of the algorithm for the AF computation: in blue the computational steps, in red the outputs. Shadowed green arrows show where data sources are used.

In order to understand the importance of the dilation, a simple example can be useful: in Fig. 4, the situation for the municipality of Monte Cremasco (Lombardia, latitude 45°22'31"80 N, longitude 09°34'20"64 E) is shown. The anthropic sub-regions are depicted in red. This example is particularly meaningful, as the anthropic sub-regions are scattered in different part of the territory, with a lot of “hole” and unconnected areas, thus the territory has a high degree of fragmentation. In Fig. 5, the result of the dilation on the red areas is shown. The new areas, which have been added by the dilation of the original ones, are shown in violet. The dilation causes the phenomena of filling the little “holes” and connect unconnected regions that are not so far. In fact, the dilation is performed with a radius of 50 meters, but in our example this has relevant effects, because the original anthropic area was highly fragmented. A little hole in an anthropic sub-regions is not a “wild” area, because the influence of anthropic activity on the nature is still very high. By filling holes and enlarging the outside perimeter of the anthropic sub-regions, the algorithm actually rises the areas occupied by humans, by taking into account the bad side effect of fragmentation of the territory on animal and plant species. In this example, the effect of the dilation is a gain of the Death Zone of about 18%, if compared to the case of simply measuring the area of the anthropic sub-regions (without dilation).

The third step of the algorithm excludes the regions where human settlements are not possible, i.e., the part of the municipality, if any, occupied by inland waters (e.g., lakes or lagoons) or lands located more than 3,000 m above sea level. We define all these sub-regions as *neutral sub-regions.* The union of all the neutral sub-regions (if present), corresponds to the *Neutral Zone*. Let define *NA* as the area (in square kilometers) of the Neutral Zone.

After the computation of *S*, *DA*, and *NA*, we can define the *Anthropentropy Factor AF* [16] as the ratio:

$$AF = DA / (S - NA) \qquad (1)$$

We have computed the *AF* indicator for all the municipalities of the two case studies.

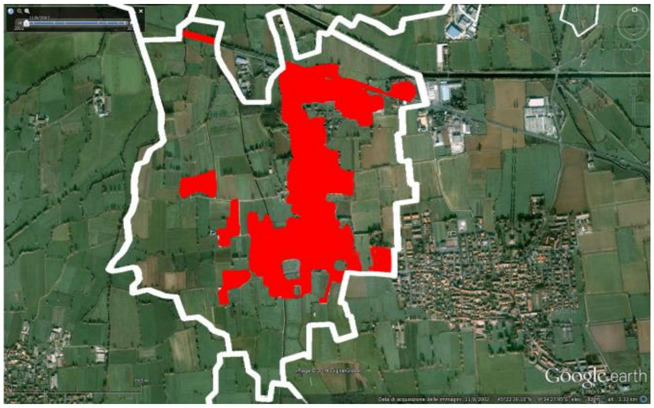


Figure 4. The anthropic sub-regions for the municipality of Monte Cremasco (Lombardia): in red, the anthropic regions are superimposed on the standard Google Earth map of the municipality, in white the boundaries of the municipality.

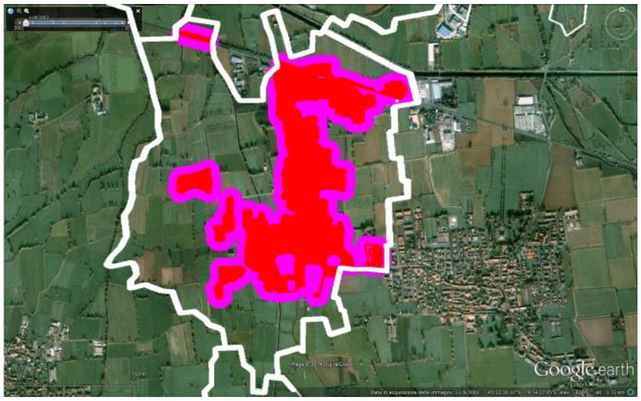


Figure 5. The anthropic sub-regions after the dilation: in red, the original sub-regions, in violet the added area. The Death Zone is the union of red and violet regions.

The first one is the province of Pavia, which is located around its chief town, Pavia (latitude, longitude: 45°11'7"44 N, 09°9'45"00 E), in the North-West of Italy. The province consists of 190 municipalities. In Fig. 6, the map of the area of the province of Pavia is shown: for each municipality, its territory is depicted in a color related to the class of land use, as specified in Table II (from green, yellow, red and black). In Fig. 7, the equivalent map is shown for the second case study, the province of Lodi, which is located around its chief town, Lodi (latitude, longitude: 45°18'52"20 N, 09°30'14"04 E). The province consists of 61 municipalities. The first case study has been already investigated in [1]. We have chosen the province of Lodi, as second case study, because it is close to the province of Pavia along the Po river (the longest and most important Italian river); therefore, it includes a territory very similar, for geomorphology and climate, to a great part of the province of Pavia. However, differences are notable, as the province of Pavia includes also a montane territory in the South part, which is completely missing in the province of Lodi. In the discussion of the results (See Section IV.B), it will be interesting to discover how differences and analogies on the geomorphology and altitude of the municipalities can affect the composite indicator. For all these reasons, we consider the second case study a good term of comparison to the first one.

The novelties of this algorithm, from the computer science side of the multidisciplinary approach, are the application of the mathematical morphology operator of dilation, for the computation of the *Death Zone*, to GIS data, and the application of the constraints of our framework in each of the computational steps of the algorithm. In particular, mathematical morphology have been widely use in computer vision theory as useful spatial data analysis [29]. However, morphological operators have been used on GIS data mainly for preprocessing and filtering in data acquisition phases [30], for extracting simple information on the spatial disposition of primitives (such as roads lines [31]) or to better detect areas that have to be categorized [32]. In our approach, the mathematical operator is used to generate new knowledge about the territory, as it is related to the definition of the concept of *Death Zone*, with its impact on the fragmentation of the territory and, consequently, on the biodiversity. The second aspect is that, without a careful study or precision and granularity of data, the application of mathematical morphology operators on GIS maps cannot give reliable and significant results. Therefore, the *joint use* of the operator dilation with the constraints of our framework (in particular, granularity and precision constraints in computational steps of Fig. 3) is a new and distinctive contribution of algorithms and tools of computer science, in particular in the computational sustainability field [33].

Even if the *AF* indicator is able to take into considerations quantitative extensions, shapes and relative positions of the anthrophized areas of a territory, it does not give any hints on the state of the green areas *outside* the anthropic areas, which is the goal of the second indicator here described.

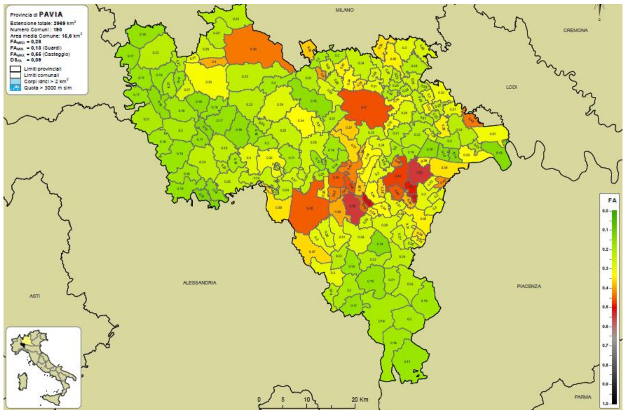


Figure 6. The visual map for the *AF* land use indicator for the municipalities of the first case study (Pavia Province, Lombardia, North Western Italy); the meaning of the colors is explained in Table II.

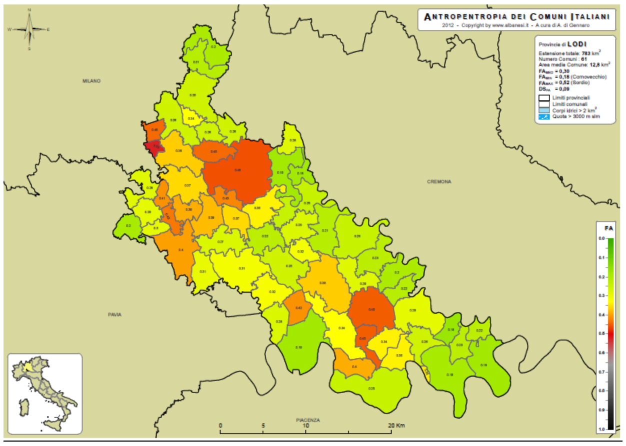


Figure 7. The visual map for the *AF* land use indicator for the municipalities of the second case study (Province of Lodi, Lombardia, North Western Italy)); the meaning of the colors is explained in Table II.

B. *Forest quality indicator*

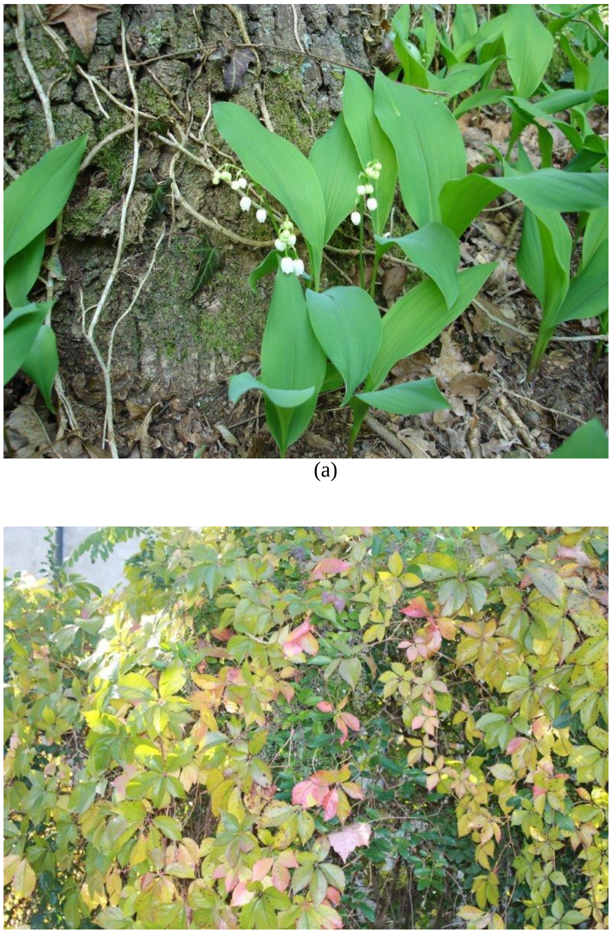
The assessment of forest quality differs according to the different components that can be evaluated (ecological, social and/or economic components associated to forests). In many assessment systems, environment has been relegated to a relatively unimportant element, if compared with other issues such as economic importance, although there are now also some specialized indicator sets relating to the environment, such as WWF's Living Planet Index [34]. Other examples include: the IUCN well-being index [35], that divides indicators into two classes, the first relating to human well-being (socio-economic) and the second to the environment (ecological, environmental services etc.) and the Montreal Process criteria and indicators [34], for temperate and boreal species outside Europe, which uses seven criteria (and 67 indicators) including the conservation of biological diversity. For our purposes, the *FSQ* indicator [1] expresses the forest quality status as the value of its ecological components, with particularly reference to the biodiversity conservation. We have chosen the following components:

- the number of forest layers: more layers correspond to higher biodiversity;
- the presence of protected species according to the Lombardia regional law [L.R. 10/2008]: more protected species mean higher and better biodiversity;
- the presence of alien species: a lower number of alien species mean higher and better biodiversity.

For the target area of region Lombardia, in Figs. 8a and 8b, examples of protected and alien species are shown, respectively. We have imposed some limitation on the GIS data, following two important constraints:

- In the *FSQ* computation, only natural forests have been considered, i.e., plantations were excluded.
- Only forests occurring on areas greater than 10.000 square meters have been considered. In fact, floristic richness, in forest patches smaller than 1 ha, is generally very low [36].

Following the theoretical framework, in particular the granularity constraint, also the *FSQ* indicator is computed for each municipality of the target territories. We define a set of sub-regions occupied by natural forest F_i (i = 1, 2, n). Each of F_i may have one or more occurrences, denoted by the index k, in the territory (k = 1, 2, max(i)). Each k-th occurrence is characterizes by: (a) an area A^k_i, expressed in square meters, for i= 1,2,,…n and k = 1, 2,… max(i) and (b) a type of T_i, derived from the GIS ERSAF Database "Map of the Forest Types of Lombardia" [27], which classifies forests on the basis of their physiognomy (dominant woody species) and the ecological characteristics of the site where they occur (geological substrate, type of soil, etc.) [37]. As for the *AF* computation, we take into consideration the same two target territories: the province of Pavia and the province of Lodi. In the first one, there are 66 different forest types, but only 32 of them have occurrences whose areas are greater than 10.000 square meters. Therefore, for the *FSQ* computation of the province of Pavia, *n* = 32. In the second case, the province of Lodi, only 11 forest types survive the area constraints. Therefore, for the province of Lodi, *n* = 11. As described before, the province of Pavia includes a portion of montane territory, which is not present in the province of Lodi. For this reason, the province of Pavia is characterized by a higher number of forest types. In Table III, a list of the types T_i and the relative reference *syntaxa* is provided, referring to the types, which are present in *both* the two provinces. The two provinces have ten forest types in common, therefore, the T_i are listed for i=1, 2,. 10. In Table IV, the forest types for the province of Pavia are listed. The province of Pavia has twelve forest types, which are not present in the province of Lodi, therefore, the T_i are listed for i=11, 12,. 32. There is only one forest type in the province of Lodi, which is not in the province of Pavia (labeled T_{33}, see Table V). The Type Lab field in the tables is a data label, which refers to the database [27] used as input source. Moreover, the difference between the two case studies is related not only to the number of forest types, but also to the



(a)

(b)

Figure 8. Examples of protected (a) and alien (b) species, according to the regional law of Lombardia: (a) *Convallaria majalis* (b) *Parthenocissus quinquefolia*. Photos have been acquired during the relevés in the target territories.

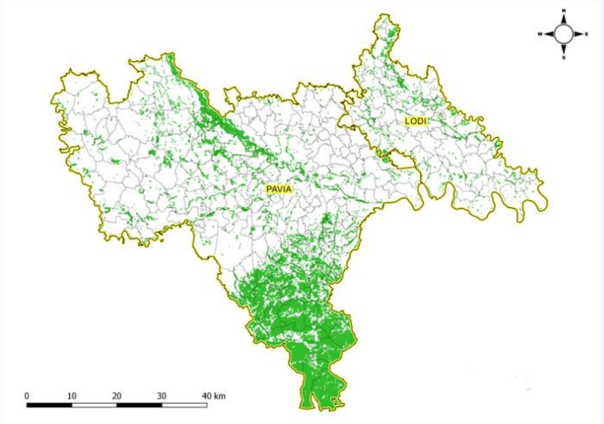


Figure 9. The two provinces of Pavia and Lodi: the areas of the forests are depicted in green.

TABLE III. FOREST TYPES IN COMMON BETWEEN THE TWO CASE STUDIES, THE PROVINCE OF PAVIA AND THE PROVINCE OF LODI.

Type Lab[a]	**Description of forest types T_i and relative reference *syntaxa***
1	T_1: Oak-Hornbeam wood of the lowlands *Syntaxa: Polygonato multiflori-Quercetum roboris* subass. *carpinetosum* and *anemonetosum* Sartori 1984; *Quercus robur*, *Carpinus betulus* and *Physospermum cornubiense* community; *Quercus robur*, *Carpinus betulus* and *Holcus mollis* community
12	T_2: Oak wood of inland sand dunes *Syntaxa: Quercus robur* community
13	T_3: Oak wood of stony river beds *Syntaxa: Quercus robur* and *Brachypodium rupestre* community
14-15	T_4, T_5: Oak-Elm wood (also including the Black Alder variant) *Syntaxa: Polygonato multiflori-Quercetum roboris* subass. *ulmetosum* Sartori 1984
173	T_6: Typical Black Alder wood *Syntaxa: Osmundo regalis-Alnetum glutinosae* Vanden Berghen 1971; *Carici elongatae-Alnetum glutinosae* W. Koch 1926 et R. Tx. 1931; *Carici acutiformis-Alnetum glutinosae* Scamoni 1935
177	T_7: Willow wood of bank *Syntaxa: Salix alba* community*; Salicetum albae* Issler 1926
180	T_8: *Salix cinerea* wood *Syntaxa: Salicetum cinereae* Zolyomi 1931
188	T_9: Pure *Robinia pseudoacacia* wood *Syntaxa: Robinia pseudoacacia* community
189	T_{10}: Mixed *Robinia pseudoacacia* wood *Syntaxa:Robinia pseudoacacia, Quercus robur* and *Ulmus minor* community

a. According to ERSAF database [27]

areas occupied by the forests, which is significantly lower in Lodi. This is evident by observing Fig. 9, where the two neighbor provinces are shown, with the forest areas depicted in green (without any differentiation among forest types): the province of Pavia has a considerably presence of forest in the South, where the altitude is higher. On the contrary, the province of Lodi has forests only near the boundaries, as the rest of the planar belt is occupied mainly by agricultural crops.

For each forest T_i of Tables III, IV, and V, we found the correspondence with one or more phytosociological tables [25]. When this correspondence was not reported by the above mentioned authors, we used other bibliographic references or phytosociological *relevés* collected in the area the forest type occurs.

For each forest type T_i, we defined a set of the following indicator components (s_i, a_i, p_i):

- Stratification (number of layers) of a forest type i (s_i): this component analyzes the quality of the forest structure. The tree and the herb layers are always present in a forest. The shrub layers (high-shrub and/or low-shrub layers) were considered valuable if their total cover were > of 10% of the sampled forest area (indicated in the phytosociological tables) or at least one species presented an abundance value equal to 2.
- Percentage frequency of alien species (a_i) in the corresponding phytosociological table/s. When more phytosociological tables described a forest type T_i, a

mean value between the percentages of each table was calculated.

- Percentage frequency of protected species (p_i) in the corresponding phytosociological table/s. When more phytosociological tables described a forest type T_i, a mean value between the percentages of each table was calculated.

The three components can assume only discrete values, from 0 to 3, according to an *if – then – else* algorithm described in the following of this paragraph.

TABLE IV. FOREST TYPES THAT ARE PRESENT ONLY IN THE PROVINCE OF PAVIA.

Type Lab[a]	**Description of forest types T_i and relative reference *syntaxa***
20, 23	T_{11}, T_{12}: *Quercus pubescens* wood of the carbonatic substrates (also including the Chestnut variant) *Syntaxa: Quercus pubescens, Euphorbia cyparissias* and *Epipactis helleborine* community
26, 27	T_{13}, T_{14}: *Quercus petraea* wood of the carbonatic substrates and mesic soils (also including the Chestnut variant) *Syntaxa: Physospermo cornubiensis-Quercetum petraeae* Oberd. et Hofm. 1967
28	T_{15}: *Quercus cerris* wood *Syntaxa: Quercus cerris, Cruciata glabra* and *Anemone trifolia* community
45, 48, 49, 50, 57	T_{16}, T_{17}, T_{18}, T_{19}, and T_{20}: Chestnut wood on drift; Chestnut wood of the carbonatic substrates (mesic soils, meso-xeric soils, xeric soils); Chestnut wood of the siliceous substrates and mesic soils *Syntaxa: Physospermo cornubiensis-Quercetum petraeae* Oberd.et Hofm. 1967*; Castanea sativa* and *Corylus avellana* community
63, 64, 65	T_{21}, T_{22}, and T_{23} *Ostrya carpinifolia* and *Fraxinus ornus* wood (of layer, of cliff, typical) *Syntaxa: Knautio drymeiae-Ostryetum* Mondino et al. 1993
84	T_{24}: Birch wood *Syntaxa: Betula pendula* community
88	T_{25}: Primitive Beech wood *Syntaxa: Trochiscantho-Fagetum* Gentile 1974; *Fagus sylvatica* and *Acer opulifolium* community
89, 96, 97, 105	T_{26}, T_{27}, T_{28}, and T_{29}: Beech wood of the carbonatic substrates (high-montane, montane, montane of xeric soils, submontane) *Syntaxa: Trochiscantho-Fagetum* Gentile 1974; *Fagus sylvatica* and *Acer opulifolium* community
99	T_{30}: Beech wood of the siliceous substrates *Syntaxa: Trochiscantho-Fagetum* Gentile 1974; *Fagus sylvatica* and *Acer opulifolium* community
172	T_{31}: Black Alder wood of gulley *Syntaxa: Alnus glutinosa*, *Populus alba* and *Ulmus minor* community
183	T_{32}: White Poplar formation *Syntaxa: Populus alba* community

a. According to ERSAF database [27]

TABLE V. THE UNIQUE FOREST TYPE THAT IS PRESENT ONLY IN THE PROVINCE OF LODI.

Type Lab[a]	**Description of forest types T_i and relative reference *syntaxa***
5	T_{33}: Oak-Hornbeam of the hills *Syntaxa*: *Castanea sativa*, *Carpinus betulus* and *Quercus petraea* community

a. According to ERSAF database [27]

While the definition of quality of stratification is independent on the altitude of the forest, the definition of values related to the percentages of alien and protected species is different, according to the altitude, because usually the impact of human activities decreases with the altitude. Thus, naturalness is higher in the montane belt than in planar belt. We differentiate between forest types belonging to the class “high hilly and montane” (altitude > = 500 m) and forest types belonging to the class “planar and low hilly” (altitude < 500 m). The three components (s_i, a_i, p_i) are defined according to an empirical *if – then – else* algorithm:

```
If the number of layers = 2, then $s_i = 1$
    Else if number of layers = 3, then $s_i = 2$
    Else if number of layers = 4, then $s_i = 3$
For altitude <500 m:
If the percentage of alien species is > 40 then $a_i = 0$
        Else if alien species range is (15- 40] then $a_i = 1$
        Else if alien species range is (5- 15] then $a_i = 2$
        Else if alien species range is [0- 5] then $a_i = 3$
If percentage of protected species range is (0.5-3] then $p_i = 1$
        Else if protected species range is (3- 6.5] then $p_i = 2$
        Else if protected species range is > 6.5 then $p_i = 3$
For altitude > = 500 m:
If the percentage of alien species is > 10 then $a_i = 0$
        Else if alien species range is (5-10] then $a_i = 1$
        Else if alien species range is (2-5] then $a_i = 2$
        Else if alien species range is [0- 2] then $a_i = 3$
If percentage of protected species range is (0.5-5] then $p_i = 1$
        Else if protected species range is (5- 10] then $p_i = 2$
        Else if protected species range is > 10 then $p_i = 3$
```

For each of the forest type *i* of Tables III, IV, and V, we computed the relative value set of (s_i, a_i, p_i), according to the *if – then – else* algorithm and the phytosociological tables and/or relevés: the complete value set is reported in Table VI.

After determining the values of the set of components for stratification, alien and protected species, it is now possible to define the *F*orest *S*tatus *Q*uality Indicator (in the following, *FSQ*) of a given territory of a municipality as

$$FSQ = \Sigma_i \;\; \Sigma_k \, (s_i + a_i + p_i) * A^k_i / S \qquad (2)$$

where *i* is one of the significant forest type (*significant* means that at least one occurrence of the forest has A^k_i >= 10.000 square meters) that is present in the territory under investigation (Tables III and IV for Pavia, Tables III and V for Lodi), A^k_i is the area of the *k-th* occurrence of forest type *i*, and *S* is the area of the municipality. The number of occurrences may vary, from a minimum of 1 to a maximum, which depends on the forest type. The *FSQ* definition is the weighted values of the components, where the weights are the ratios between the areas of the forests and the area of the territory under investigation. The wider is the area occupied by a forest, the higher is its contribution to the global quality of the territory. Besides, its contribution is weighted by the values of the components (stratification, alien, and protected species) as described in the *if – then – else* algorithm.

TABLE VI. THE VALUE SET OF COMPONENTS FOR STRATIFICATION, ALIEN AND PROTECTED SPECIES, FOR EACH FOREST TYPE OF BOTH THE CASE STUDIES.

Type Lab[a]	**Components (s_i, a_i, p_i)**
1	3,2,3
5	3,2,3
12	2,2,1
13	3,3,3
14-15	3,2,2
20, 23	3,3,1
26, 27	2,3,3
28	3,3,2
45, 48, 49, 50, 57	2,3,3
63, 64, 65	3,3,2
84	1,3,0
88	3,3,3
89, 96, 97, 105	3,3,3
99	3,3,3
172	3,3,1
173	2,3,2
177	1,1,0
180	2,2,0
183	3,1,0
188	2,1,0
189	3,2,0

TABLE VII. THE METRIC ON THE *FSQ* INDICATOR FOR FOREST QUALITY.

Class of forest quality	**Evaluation of Forest quality and policy**	
	Intervals of FSQ	***Suggested policy***
1 Unsatisfactory	0 <= *FSQ* <= 0.9	Very low level forest quality. A high-impact policy of restoration and/or requalitication of forest is mandatory.
2 Satisfactory but improvable	0.9 < *FSQ* <= 1.8	Sufficient forest quality but improvable. A policy for forest biodiversity conservation is preferable.
3 Good	1.8 < *FSQ* <= 3.6	Good forest quality, the first level of satisfactory situation. A policy for the conservation of existing forests is suggested.
4 Optimum	3.6 < *FSQ* <= 4.5	The optimun situation, with a high quality of forests. A policy for the conservation of existing forests is suggested. Anyway, if shrublands and grasslands are scarse or absent, a policy for their biodiversity conservation has to be considered.
5 Overbalanced	*FSQ* > 4.5	The overbalanced situation, forests have overcome other ecosystems. A policy for shrubland and grassland biodiversity conservation is highly suggested.

The summation in (2) is for all the forest types of the territory under investigation, and for all the occurrences of the forests. The *FSQ* value can range from 0 (no forests are present in the territory with at least one occurrence of A^1_i> 10.000) to a maximum of 9, which refers the "perfect", quite unrealistic, situation of forests of very high quality (set of components (s_i, a_i, p_i) =(3,3,3)), which occupy the entire territory of the municipality $(\Sigma_\iota\ \Sigma_\kappa\ A^k_i = S)$. By using an approach similar to the *AF* metric, we have defined a set of ranges for the *FSQ* indicator. In Table VII, the metric for the *FSQ* indicator and the suggested policy actions are shown.

C. Results for the two indicators

By referring to the general framework of the multidisciplinary approach, the different knowledge of vegetation science and computer science has been combined, by following the stated rules (format, temporal, granularity and precision constraints), on the GIS databases, and the results are the values of *AF* and *FSQ*, computed for all the municipalities of the province of Pavia and Lodi, according to eqs. (1) and (2), respectively. In Figs. 10 and 11, the plot of *AF* and *FSQ* are reported for Pavia and Lodi, respectively. On the X-axis, the municipalities are listed according the alphabetical order on their names, and each of them is labeled by a numerical value (from 1 to 190 for Pavia, from 1 to 61 for Lodi), to increase readability. On the Y-axis, the values of the two indicators are plotted. By comparing the two provinces, it is clear that the *AF* values are comparable for the two cases: on the contrary, the *FSQ* values are considerably lower for the second case (Lodi, see Fig. 11, blue lines) than for the first one (Pavia, see Fig. 10, blue line). The *FSQ* indicator is even considerably lower than the *AF* indicator for the second case (the red line is over the blue line, for most of the cases), and this is more surprising, if we consider that the two indicators have different scales (from 0 to 1 for *AF*, from 0 to 9 for *FSQ*). This is evident also from the scatter plot, where the relations *FSQ* vs. *AF* are depicted (see Figs. 12 and 13, for Pavia and Lodi, respectively). From the scatter plot, we see a similar dependency between the two indicators in the two provinces, with most of the municipality with the *FSQ* values agglomerated around the Y-axis (*FSQ* = 0), but with different performance in terms of class of metrics. In fact, we can see that the *AF* indicator shows a comparable land use level for both the two provinces: almost all the *AF* values are below the first worrying level of *AF* = 0.4 (classes 1 and 2 of land use, see Table II) and few are in class 3, no one in classes 4 and 5. However, performance in terms of forest quality are very different. The *FSQ* values for Pavia cover all the classes (see Table VII), while for Lodi, the *FSQ* values are all in the first, unsatisfactory class (*FSQ* < = 0.9). This means that in the second case study, not only forests are less present in the landscape of Lodi (as it can be infer by simply analyzing the cumulative GIS image of the forests, see Fig. 9), but also that their quality is not so high to compensate the quantitative negative situation. The dispersion plots (Fig. 13 and Fig. 14) show that the two indicators are quite independent and this is a positive result, this means that the two indicators are related to independent and different pressures on the environment: land use and forest quality. This is also confirmed by the correlation between the *AF* and *FSQ*, which is very low: the correlation coefficients are equal to -0,197651 and -0,205527, for Pavia and Lodi, respectively. The analysis of the dispersion plots also reveals that all the municipalities with serious levels of *AF* (> 0.4) have very low levels of forest quality (*FSQ* < 0.9), for both the provinces. This underlines a worrying trend to neglect the ecological compensations to mitigate the impact for increasing urbanization. This suggest the fact that in both the two provinces environment and biodiversity loss are scarcely considered in the land use policies.

It is evident that the two indicators, *AF* and *FSQ*, express different pressure on the biodiversity, and that it is not simple to get an overall view of the situation of a territory, if we consider the two environmental indicators separately. This is due, not only to the fact that they are assessed according to two different metrics, but also because it is difficult to compare different municipalities by using a couple of un-normalized values, instead of only one. For this reason, a further step in this research has been taken in the direction of building a composite indicator, described in Section IV.

IV. THE COMPOSITE INDICATOR

After the definition of the two simple indicators and their computation on the case studies, it is now possible to build a unique, composite indicator, with the aim to give an overall description of the two aspects of environmental preservation: land use and forest quality.

A. *Theory of composing AF and FSQ indicators*

Composite indicators are increasing recognized to be very relevant in policy action assessment and communication to citizens about social, economic and environmental issues. They have undisputed advantages, against known or otherwise controllable disadvantages. In fact, a composite indicator is able to summarize, in a more compact and powerful way, multiple concepts, related to different “single” indicators. Moreover, it reduces the set of considered data and for this reason it is easier to be interpreted and to be communicated to the citizens. The last aspect is particularly important in environmental issues, where communication of the state of the territory is relevant, as it is the first step to raise awareness on the environment preservation. The main disadvantages of composite indicators are related to the accuracy of their definition: as data and dimensions are reduced, relevant information may be missed, if the construction process lacks of statistical or conceptual knowledge, or if it is not transparent and fully described, in terms of data selection and applied algorithms. Concerning data selection, we use the availability constraint of our framework, which is implemented by the open data paradigm: all the databases we have used are fully available on Internet and they can be processed by open source GIS software [26]. Concerning the algorithm for the composite indicator computation, in our research we follow a rigorous method [19]. The method consists of several steps, which can be summarized in the following paragraphs.

Theoretical framework: in this step, the basis for the selection of the single indicators are to be settled. This step generally involves knowledge of experts and stakeholders of the target issue. As already discussed in Section II, the formalization of the multidisciplinary approach individuates vegetation science and computer science as the fundamental disciplines, and a set of *constraints* to assure a robust communication between the two sciences.

Data selection: this step has been already described, as it consists of the choice of the databases and the computation of *AF* and *FSQ* indicators. The formulas for the definitions of the two environmental indicators (see (1) and (2)) and the data set of the two case studies (see Section IV) are the outputs of this step.

Imputation of missing data: this step is relevant whenever the collected data are not complete, and missing data has to be replace in some way. In our case, Corine Land Cover [12] and ERSAF [27] databases are very detailed, and the problem of missing data is quite irrelevant, as the classification of land use and the forest areas are quite enough for the granularity and the precision adopted in the framework. There is only one aspect that can be reported to the problem of missing data: in the ERSAF databases, some typologies of forest cannot be assigned with the component of stratification, alien and protected species, because their description is too vague. They are two typologies of forest labeled as “unclassified forest areas” (ERSAF [27] Type Lab Fields 900 and -100). However, these data occupy an area of 3.6% of the entire region Lombardia, therefore, it is reasonable to omit these data and to adopt a Complete Cases approach [19], where data are simply discarded, as their irrelevance on the entire set of data. As a rule of thumb, if a variable has less of 5% of missing data [38], the cases can be omitted and the Complete Case Analysis is a simply but robust choice.

Statistical analysis: in Table VIII, the main statistical data measured on *AF* and *FSQ* indicators are provided for the two case studies.

Normalization: this step is fundamental to compose indicators, which are expressed in different scales or measure units. In our case, *AF* range of definition is [0-1], while *FSQ* range is [0-9]. Moreover, the two indicators are *discordant* [39]: in fact, low values of *AF* expressed a positive assessment on the environmental issue (See Table II), while, on the contrary, low level of *FSQ* expresses a negative assessment (See Table V). In the general theory, a composite indicator *X* can be written in the form:

$$X = \mathrm{F}\,[\mathrm{N}_1(\mathrm{x}_1), \mathrm{N}_2(\mathrm{x}_2), \ldots, \mathrm{N}_j(\mathrm{x}_j)] \qquad (3)$$

where F is a function of aggregation, x_j are the single j-th indicator, N_j is a normalization function. In our case, $j = 2$ and $x_1 = AF^m$ and $x_2 = FSQ^m$, where the variable *AF* and *FSQ* can assume *m* distinct values (*reference data-set*), corresponding to all the municipalities of our case studies, namely $m = 190 + 61 = 251$. There are plenty of possibility in choosing the normalization and aggregation functions [39]. In our study, we start with the simplest linear normalization, which reports the range of indicators from 0 to 1:

$$x_j - \min(x_j) / [\max(x_j) - \min(x_j)] \qquad (4)$$

For the case of *AF* indicator the normalization is the concatenation of two linear transformations: the first one considers the complement to 1, (1 – *AF*), in order to make the two indicators concordant, and the second is the linear transformation of (4). Therefore, we have:

$$N_1\,(AF^m) = [(1 - AF^m) - \min(1 - AF^m)\,]/\,[\max(1 - AF^m) - \min(1 - AF^m)] \qquad (5)$$

$$N_2\,(FSQ^m) = [FSQ^m - \min(FSQ^m)\,]/\,[\max(FSQ^m) - \min(FSQ^m)] \qquad (6)$$

Aggregation and Weighting functions: after normalization, the choice of the aggregation function is the final step for the definition of the composite indicator. It allows to merge the information of the two normalized indicators to give, as output, a unique value for each item of the reference data-set. In our case, we have chosen the simplest solution of the linear aggregation with equal weights (0.5), in order to give the same importance to the two phenomena, i.e., land use and forest quality. This operative choice is the most common in the case of a limited number of indicators with a low degree of correlation [19], as in our study.

Therefore, we can define the Biodiversity Composite Indicator (*BCI*) as:

$$BCI^m = 0.5 * N_1\,(AF^m) + 0.5 * N_2\,(FSQ^m) \qquad (7)$$

where N_1 and N_2 are the normalized functions defined in (5) and (6), respectively, and the index *m* cover all the municipalities of both the provinces, i.e., $m = 1,2,\ldots 251$.

B. Results for composite indicator

The Biodiversity Composite Indicator gives a value, for each municipality, in the range of [0-1], where 0 means the worst situation and 1 the best one. Obviously, in the composite indicator, we lose the distinction of what factor influences the result (if land use or forest quality), but we have an overall, absolute assessment of how the two aspects are combined in the impact on biodiversity.

In Fig. 14 and Fig. 15, the *BCI* values for the two provinces of Pavia and Lodi are shown, respectively. On the X-axis, the municipalities are listed according the alphabetical order of their name, and each of them is labeled by a numerical value (from 1 to 190 for Pavia, from 1 to 61 for Lodi), to increase readability. On the Y-axis, the values of the Biodiversity Composite Indicator are plotted. If we define a metric for the *BCI* similar to the two metrics for the simple indicators (see Table II and Table VII), we can settle five classes of situations (see Table IX), of increasing performance in term of global impact on biodiversity of both the two aspects (land use and forest quality).

TABLE VIII. REPRESENTATIVE STATISTICAL DATA FOR THE TWO INDICATORS FOR THE TWO CASE STUDIES, PAVIA AND LODI.

Territory	**Main characteristics**	
	AF indicator	***FSQ indicator***
Range of indicator	[0-1]	[0-9]
Case study °1: Province of Pavia		
Number of computed values	190	190
Minimum value	0.133	0
Maximum value	0.548	5.34978
Average value	0.2757	0.4949
Standard Deviation	0.091	1.01
Case study °2: Province of Lodi		
Number of computed values	61	61
Minimum value	0.175	0
Maximum value	0.524	0.58
Average Value	0.305	0.0832
Standard Deviation	0.088	0.162

By comparing the values of *BCI* for the two case studies (Fig. 14 and Fig. 15), we can infer that the composite indicator expresses a more serious situation for the province of Lodi, than for Pavia. In fact, in the second case study, none of the municipalities shows a value greater than 0.6 of the composite indicator, while in the first case study, few but existing cases refer to good or excellent situation (*BCI* > 0.6, classes 4 and 5).

In order to compare directly the two case studies, we can report on the same plot the percentage of municipalities that fall in each of the class of impact on biodiversity (see Fig. 16). Also, in this analysis, the first case study (Pavia) outperforms the second one (Lodi). In fact, the percentages of the municipalities belonging to the classes 1 and 2 (see Table IX), which refer to the dramatic and serious impact on biodiversity, are always higher in the second case study (Lodi) than in the first one (Pavia).

The last issue we want to discuss refers to the question if the altitude may influence the *BCI* performance. The results from the experiments seem to indicate that the altitude influences the *BCI* values, only if its range includes a significant percentage of municipalities of "high" altitude (previously defined as altitude > 500 m). In fact, in the first case study, the correlation between the *BCI* and altitude is enough good (0.584), while in the second case the correlation is weak and, even, negative (-0.196). This is visually confirmed by comparing Fig. 17 and Fig. 18. In these two plots, the *BCI* values are reported for the two provinces, as a function of the average altitude of the municipality the *BCI* value refers to. In the first case study, it is true that in the higher classes of our composite indicator (classes 4 and 5) we find *only* montane municipalities.

TABLE IX. THE METRIC ON THE *BCI* INDICATOR.

Evaluation of the Biodiversity Composite Indicator		
Class of impact on biodiversity	***Intervals of BCI***	***Meaning***
1	0 <= *BCI* <= 0.2	The worst situation, with a dramatic impact on biodiversity.
2	0.2 < *BCI* <= 0.4	A very serious level of pressures on biodiversity.
3	0.4 < *BCI* <= 0.6	A first worrying impact on biodiversity of the compound effect of land use and forest quality.
4	0.6 < *BCI* <= 0.8	A good situation, with a satisfactory impact of land use and forest degradation on biodiversity.
5	0.8 < *BCI* <= 1	The excellent situation, with a very low impact of land use and forest degradation on biodiversity.

In the second case study, where the altitude range is very limited (40-101 m), the dependency between altitude and *BCI* is not significant at all. A better idea of the correlation between altitude and *BCI* could be obtained considering the whole region of Lombardia, where the altitude range is considerably wide (11-4021 m).

V. CONCLUSION AND FUTURE WORK

In this paper, we presented a new composite indicator, which can express the pressures on biodiversity, by considering two distinct phenomena: land use and forest quality status. In order to define the composite indicator, a conceptual framework has been proposed, starting from the multidisciplinary approach used in the research.

The framework states clearly the constraints, the methodologies and knowledge used by the two scientific fields involved in the research: vegetation science and computer science. We defined the composite indicator, based upon two simple indicators, the *Anthropentropy Factor* and the *Forest Status Quality Indicator*, and we analyzed two case studies, for a total number of 251 municipalities of the North-West Italian territory, i.e., region of Lombardia.

Current and future developments of this work include the computation of the composite indicator for the entire region of Lombardia. It is an ambitious goal, as the region ranks first in Italy for the population and the number of local municipalities (1530), second for population density and the fourth for area. Moreover, the region is also the most invaded by non-native species, which represent 16.9% of the total vascular flora [40]. The region Lombardia is a complex territory with very different geological, geomorphological, and climatic, bioclimatic, and phytogeographical characteristics. As a consequence, a high floristic and vegetation richness is present: the vascular flora includes 3220 entities [40], while the forest vegetation includes 174 forest types [27], [37]. For all these reasons, the computation of the Biodiversity Composite Indicator on the whole region can give interesting hints and priorities on the biodiversity conservation in different environmental conditions and contexts.

ACKNOWLEDGMENT

The authors greatly appreciate Antonio di Gennaro for his collaboration, and for importing and pre-processing of data from GIS databases.

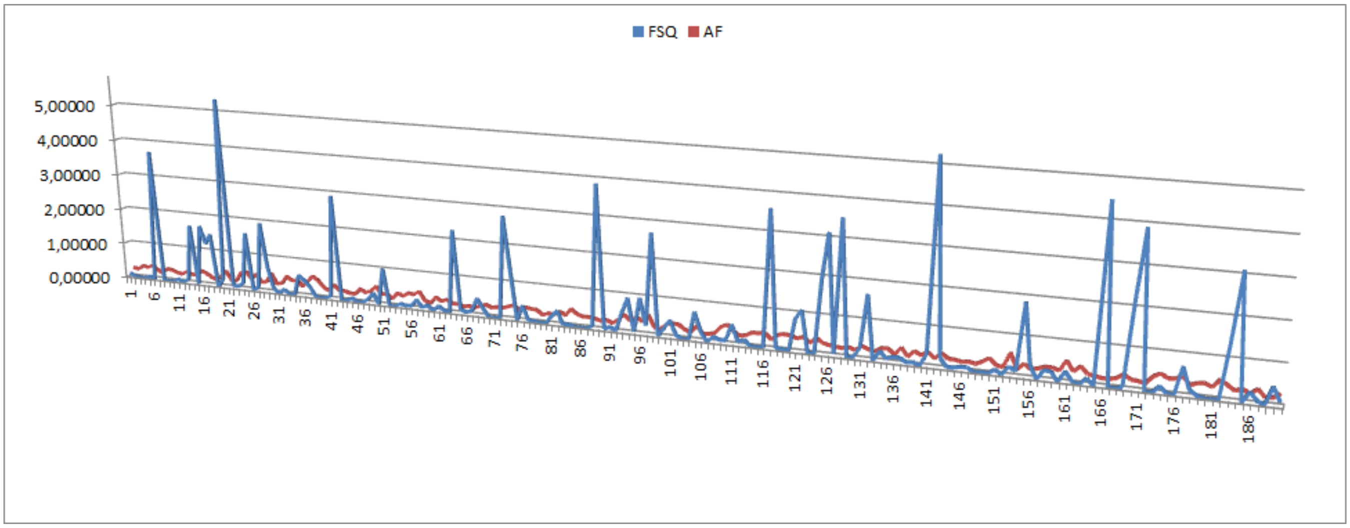


Figure 10. The two indicators, Anthropentropy Factor (*AF*) and Forest Status Quality (*FSQ*), for all the 190 municipalities of the province of Pavia.

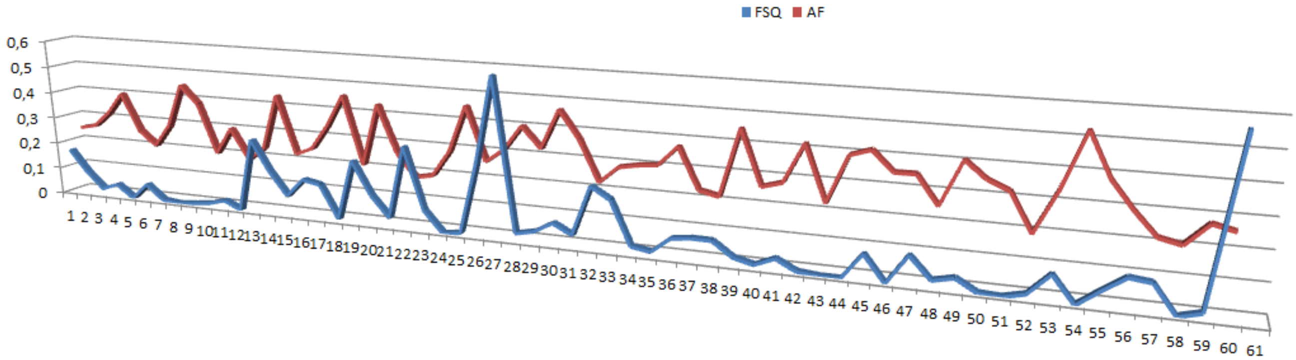


Figure 11. The two indicators, Anthropentropy Factor (*AF*) and Forest Status Quality (*FSQ*), for all the 61 municipalities of the province of Lodi.

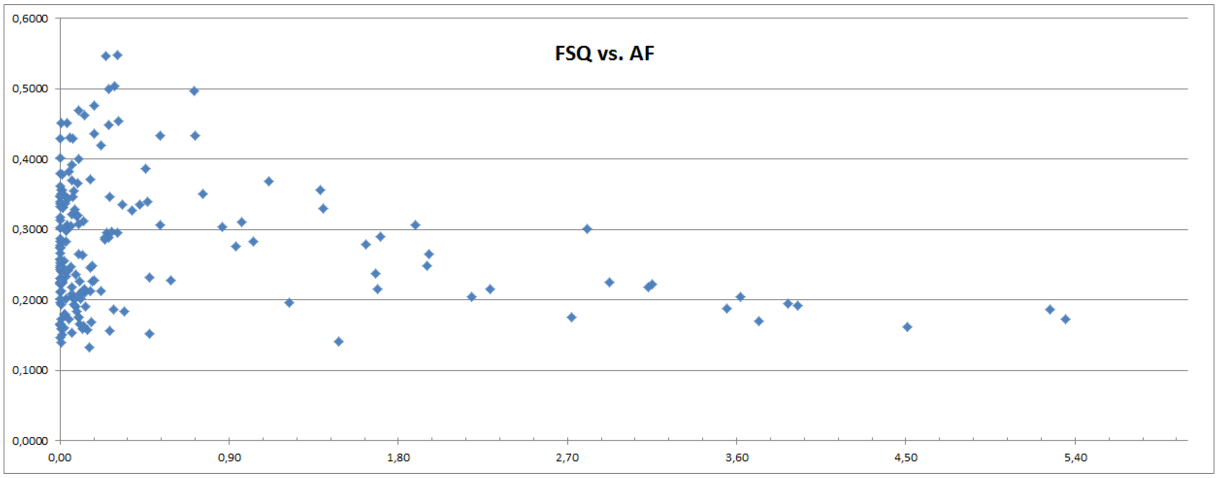


Figure 12. The relationship between the two indicators for land use and forest quality: dispersion plot of *FSQ* (on the X axis) vs. *AF* (on the Y-axis), for all the municipalities of the province of Pavia.

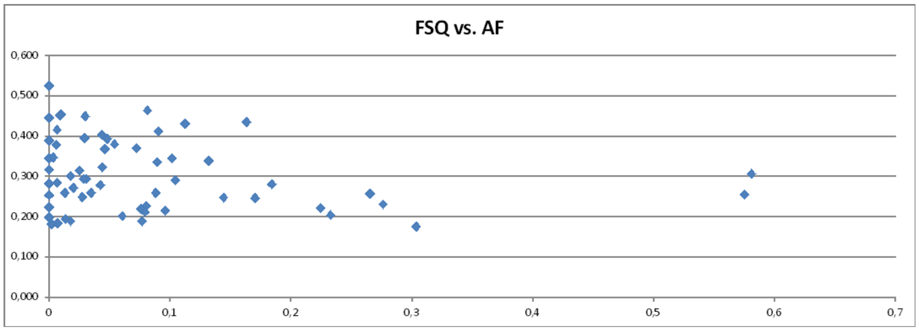


Figure 13. The relationship between the two indicators for land use and forest quality: dispersion plot of *FSQ* (on the X axis) vs. *AF* (on the Y-axis), for all the municipalities of the province of Lodi.

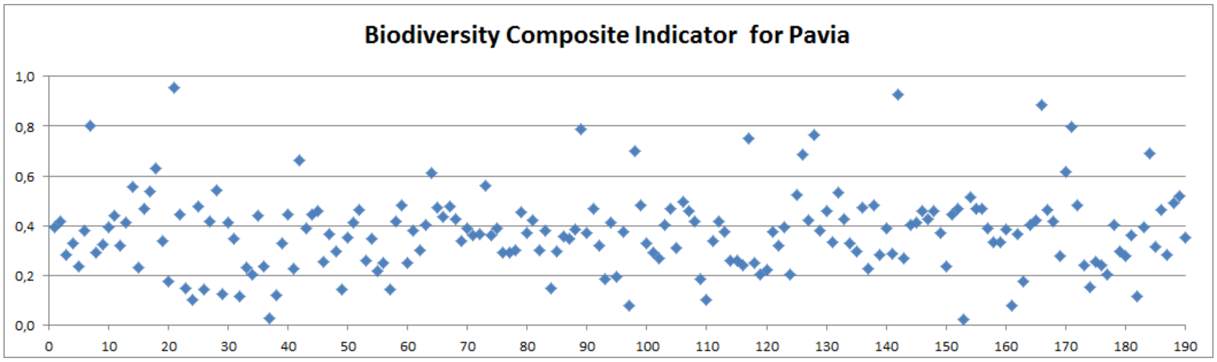


Figure 14. The Biodiversity Composite Indicator for all the municipalities of the first case study, the province of Pavia.

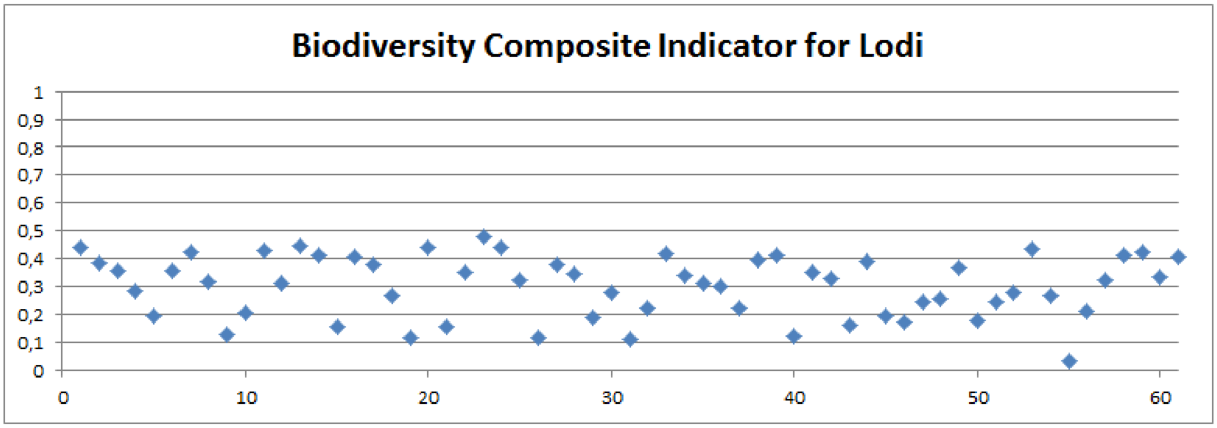


Figure 15. The Biodiversity Composite Indicator for all the municipalities of the second case study, the province of Lodi.

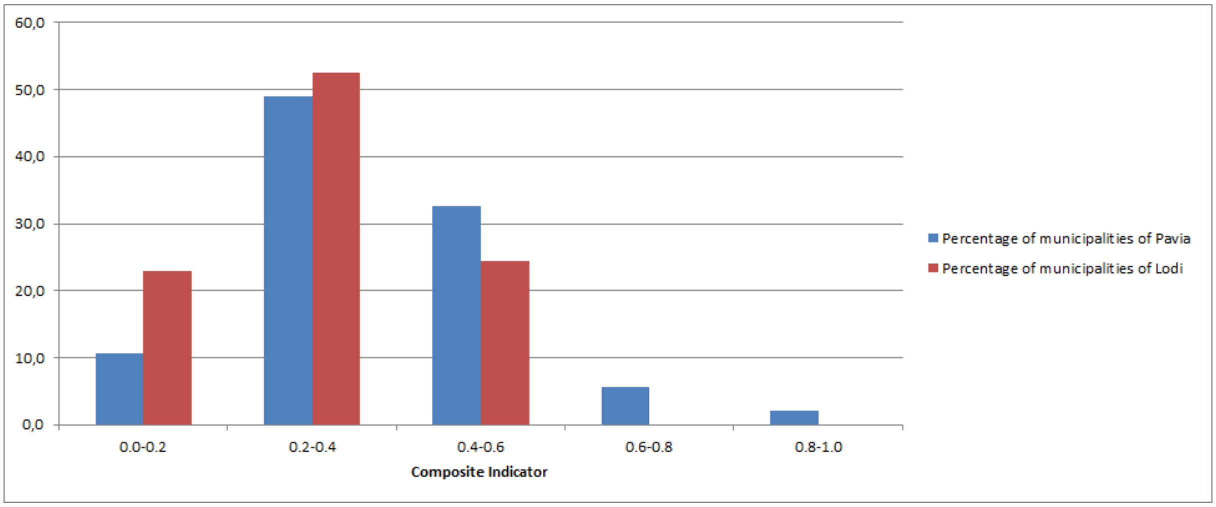


Figure 16. The percentage of municipalities falling into the five classes of pressure impact expresses by the sub-ranges of the Biodiversity Composite Indicator, for the two case studies (Pavia and Lodi).

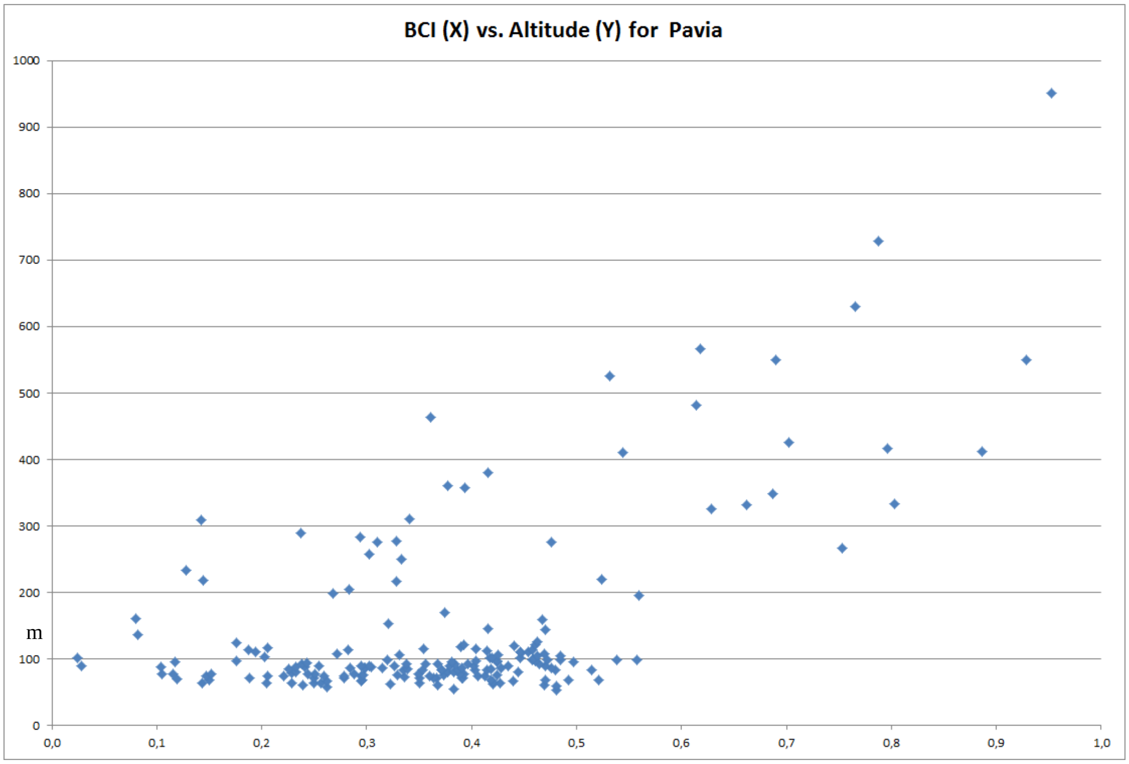


Figure 17. The relationship between the *BCI* and the average altitude (in meters) of the municipalities of Pavia.

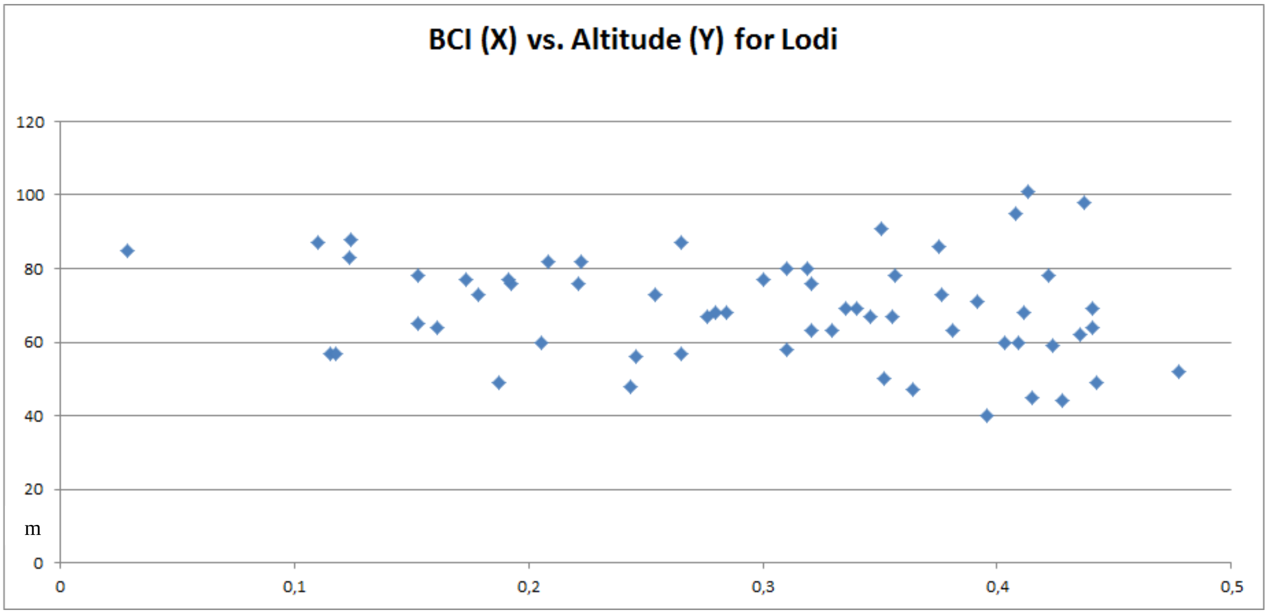


Figure 18. The relationship between the *BCI* and the average altitude (in meters) of the municipalities of Lodi.

REFERENCES

[1] S. Assini and M. G. Albanesi, "A Naturalistic Indicator of the Forest Quality and its Relationship with the Land Use Anthropentropy Factor," Proceedings of the Sixth International Conference on Bioenvironment, Biodiversity and Renewable Energies (BIONATURE 2015), Rome (Italy), 24-29 May 2015, pp. 27-33, Copyright (©) IARIA, 2015, ISBN 978-1-61208-410-7.

[2] United Nations, "Convention on Biological Diversity," 1992, https://www.cbd.int/convention/text/ [retrieved: November, 2015].

[3] United Nations, "Strategic Plan for Biodiversity 2011-2020 and the Aichi Biodiversity Targets," 2010. https://www.cbd.int/decision/cop/?id=12268 [retrieved: November, 2015].

[4] European Environment Agency, "Towards a green economy in Europe - EU environmental policy targets and objectives 2010–2050," EEA Report n. 8/2013, http://www.eea.europa.eu/publications/towards-a-green-economy-in-europe

[5] E. Smeets and R. Weterings, "Environmental indicators: typology and overview," European Environment Agency Technical Report n. 25/1999 http://www.eea.europa.eu/publications/TEC25 [retrieved: November, 2015].

[6] The UN convention on biological diversity, http://enrin.grida.no/biodiv/biodiv/cbd/eeafoll.pdf [retrieved: November, 2014].

[7] European Environment Agency, "Landscape fragmentation in Europe," Joint EEA-FOEN report n. 2/2011, ISSN 1725-9177, http://www.eea.europa.eu/publications/landscape-fragmentation-in-europe [retrieved: November, 2015].

[8] F. Gerard et al., "Land cover change in Europe between 1950 and 2000 determined employing aerial photography," Progress in Physical Geography, vol. 34 (2), pp. 183-205, 2010.

[9] A. Falcucci, L. Maiorno, and L. Boitani, "Changes in land-use/land-cover patterns in Italy and their implications for biodiversity conservation," Landscape Ecology, vol. 22, pp. 617-631, 2007.

[10] S. Assini, F. Filipponi, and F. Zucca, "Land cover changes in abandoned agricultural land in the Northern Apennine (Italy) between 1954 and 2008: spatio-temporal dynamics," Plant Biosystems, in press, doi: 10.1080/11263504.2014.983202.

[11] S. Assini, F. Filipponi, M. Brusoni, and F. Zucca, "Vegetation of abandoned areas in the northern Apennines (Italy): phytosociological aspects and biodiversity analysis," Acta Botanica Gallica: Botany Letters, vol. 161 (4), pp. 379-393, 2014, doi: 10.1080/12538078.2014.948065.

[12] Analysis of changes in European land cover from 2000 to 2006, http://www.eea.europa.eu/data-and-maps/figures/land-cover-2006-and-changes-1 [retrieved: November, 2015].

[13] ISPRA, "Land use in Italy, Report n. 195/2014," Transl. "Il consumo di suolo in Italia, Edizione 2014, http://www.isprambiente.gov.it/files/pubblicazioni/rapporti/Rapporto_Consumo_di_Suolo_in_Italia_2014.pdf [retrieved: November, 2015].

[14] N. Dudley, R. Schlaepter, W. Jackson, J.P. Jeanrenaud, and S. Stolton, Forest Quality, Assessing Forests at a landscape Scale. Eartscan, London, Sterling VA, 2006.

[15] N. Dudley, Forests In Trouble: A Review of the Status of Temperate Forests World-wide, World Wildlife Fund, London, 1992.

[16] M. G. Albanesi and R. Albanesi, "A new approach based on computer vision and collaborative social networking for environmental preservation: theory, tools and results of Italian ACI project," Proc. of The Eight International Multi-Conference on Computing in the Global Information Technology ICCGI2013, Nice (France), IARIA ed, 21-26 July 2013, pp. 16-21, ISBN 978-1-61208-283-7.

[17] M. G. Albanesi and R. Albanesi, "A Decision-making Support System for Land Use Estimation Based on a New Anthropentropy Predictive Model for Environmental Preservation – Theory, Model and Web–based Implementation," International Journal On Advances in Intelligent Systems, IARIA ed, v 7 n 1&2 2014, pp. 85-102.

[18] R. F. Noss, "Indicators for Monitoring Biodiversity: A Hierarchical Approach," Conservation Biology, Vol. 4, n. 4, pp. 355-364

[19] M. Nardo, M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, and E. Giovannini, Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Statistics Working Papers, No. 2005/03, OECD Publishing, Paris, France, http://www.oecd-ilibrary.org/economics/handbook-on-constructing-composite-indicators_533411815016 [retrieved: November, 2015].

[20] R. C. Gonzales and R. E. Woods, Digital image processing, Pearson Prentice Hall, 2008, Chapter 9, "Morphological image processing".

[21] J. Braun-Blanquet, Plant Sociology. New York: Mac Grew Hill Book Company, 1931.

[22] E. Biondi, E. 2011. "Phytosociology today: Methodological and conceptual evolution," Plant Biosystems, vol. 145 (suppl. 1), pp. 19-29, 2001.

[23] C. Blasi, E. Biondi, and J. Izco, "100 years of plant sociology: A celebration," Plant Biosystems, vol. 145 (Supplement), pp. 1-3, 2011.

[24] R. Pott, "Phytosociology: A modern geobotanical method," Plant Biosystems, vol. 145 (suppl. 1), pp. 9-18, 2011.

[25] C. Andreis and F. Sartori, eds, "Forest vegetation of Lombardia. Phytosociological classification," Transl. "Vegetazione forestale della Lombardia. Inquadramento fitosociologico," Archivio Geobotanico, vol. 12-13, pp. 1-215, 2011.

[26] QuantumGIS software, http://www.qgis.org/en/site/ [retrieved: November, 2015].

[27] Regional Organization for Services in Agricolture and Forests, Transl. Ente Regionale Servizi Agricoltura e Foreste, "Map of Forest Types of Lombardia", Transl. "Carta dei tipi reali forestali della Lombardia," http://www.geoportale.regione.lombardia.it/galleria-mappe, page 3 [retrieved: November, 2015].

[28] ISTAT shapefile for administrative boundaries: http://www.istat.it/it/archivio/24580 [retrieved: November, 2015].

[29] R. M. Haralick, R. Sternberg Stanley, and Z. Xinhua, "Image Analysis Using Mathematical Morphology," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-9, n.4, 1987, pp.532-550.

[30] M. M. Ansoult and P. J. Soille, "Mathematical Morphology: a Tool for Automated GIS Data Acquisition from Scanned Thematic Maps," Photogrammetric Engineering and Remote Sensing, vol. 56, n. 9, 1990, pp. 1263-1271.

[31] C. Zhu, W. Shi, M. Pesaresi, L. Liu, X. Chen, and B. King, “The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics,” International Journal of Remote Sensing, vol. 26, 2005, pp. 5493-5508.

[32] F. Rottensteiner and C. Briese, “A new method for building extraction in urban areas from high-resolution LIDAR data,” International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, Vol. 34 (3/A), 2002, 295-301.

[33] C. P. Gomes, "Computational sustainability: Computational methods for a sustainable environment, economy, and society,” The Bridge, Vol. 39, N. 4, 2009, 5-13.

[34] J. Loh, Living Planet Index 2003. WWF and the UNEP-World Conservation Monitoring Centre,Gland, Switzerland and Cambridge, UK, 2003.

[35] R. Prescott-Allen, The Well-Being of Nations: A Country-by-Country Index of Quality of Life and the Environment. Island Press, Covelo, California and Washington DC, 2001.

[36] P. Digiovinazzo, G.F. Ficetola, L. Bottoni, C. Andreis, and E. Padoa-Schioppa, “Ecological thresholds in herb communities for the management of suburban fragmented forests,” Forest Ecology and Management, vol. 259, pp. 343-349, 2010.

[37] R. Del Favero, “Forest types of Lombardia”, Transl. “I tipi forestali nella Regione Lombardia,” CIERRE Edizioni, 2002.

[38] R.J.A Little and D.B. Rubin, Statistical Analysis with Missing Data, Wiley Interscience, J. Wiley & Sons, Hoboken, New Jersey, 2002.

[39] F. Aiello and M. Attanasio, "How to transform a batch of simple indicators to make up a unique one." Proc, of XLII Meeting of Italian Society of Statistics, CLEUP (Padova), 327-338, 2004.

[40] L. Celesti-Grapow et al., “Non-native flora of Italy. Species distribution and threads,” Plant Biosystems, vol. 144 (1), pp. 12-28, 2010.