A New Approach Based on Computer Vision and Collaborative Social Networking for Environmental Preservation:

Theory, Tools and Results of Italian ACI Project

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

Roberto Albanesi
Thea s.r.l.
Pavia, Italy
ufficiostampa@albanesi.it

Abstract—**The ACI project (from the Italian acronym Antropentropia Comuni Italiani, i.e., Italian Municipality Anthropentropy) is a collaborative social media based project whose main goal is to measure the impact of invasive presence of human settlements on the environment. The novelty of this approach is to propose a new indicator, which considers not only the area of regions which are occupied by human settlements (buildings, roads, recreational areas), but also their relative shape and contiguity, in order to include in the metric the negative effect of territory fragmentation. Land use and fragmentation is expressed by the Anthropentropy Factor (AF); it is computed by classical computer vision operators (morphological dilation). The paper describes the theory of the new metric, its application to the Italian territory and the involvement of social network and crowdsourcing for generating open data on the current state of land use. One of the aims of this contribution is to show how well known computer vision algorithms, applied to open data collected by the User Generated Content paradigm, can be useful to map the state of degradation of the Italian territory.**

Keywords- Land use, territory fragmentation, anthropentropy crowdsourcing, social network, UGC, morphological operators.

I. INTRODUCTION

There are plenty of aspects of the effects of human activities on environment, and one of the most important is preserving land use form the threat of an inappropriate urban expansion; this is a great challenge of European Community in the field of environmental protection, as documented by several studies and data sheets: the European Environment Agency stated [1] that annual land take in 36 European countries was 111 788 ha/year in the period 2000-2006. In the most urbanized countries, the annual land take increased by 9 % in the last years. One of the consequences of land use is soil sealing, a modification of the soil which prevents it from performing vital functions for the ecosystems. The phenomenon of soil sealing is particularly evident near urban and rural areas and along main road axes. Urbanization is not the only cause of these negative effects: intensive farming and touristic activities (especially for coastlines) contribute to a relevant increase of land use, as reported by the European Environment Agency.

In the case of Italy, the situation is dramatic: land take and soil sealing assessments have been reported by ISPRA [2], showing a rapid medium growth of the sealed surface areas of 6.3% for the period 1956-2006. The main effort of researchers is twofold: to search new indicators to express the land use, and to recall or update data on the current state of the soil. In this paper, both these problems are addressed.

In Section II, the motivations of our approach are explained, in particular with reference to the current state of the art, and the definition of the Anthropentropy Factor (AF) is given. In Section III, the proposed metric is explained and discussed, and the experimental results are reported for some significant cases of Italian territory. In Section IV, conclusions and considerations about future work are given.

A. Related work

In literature, very few contributions are related to new indicators for land use and their automatic or semi-automatic computation. For example, indicators of land quality [3] and bio indicators, such as populations of ants [4] and bryophytes [5], have been investigated. The main limitation of these indicators is that they take into consideration not only land use, but also other aspects of environmental preservation, such as agro-biodiversity, water quality and land contamination/pollution. In fact, these complex indicators consider not only soil use and degradation, but also combined aspects of water, terrain, and biotic resources. Sometimes, these indicators show a week correlation with land use, as they are more sensitive to other features of ecosystem degradation; for example, the maps of bio indicator lichens [6], which seem very similar to our anthropentropy maps (see Section III), are very sensitive to pollution and presence and directions of winds, rather than urbanization; even if the negative effects of the two facts can be correlated, it is difficult to find indicators which isolate the dependence on each single aspect of environmental degradation.

The second problem is to compute land use in a completely automatic way, for example by using image processing primitives: several contributions refer to the use of remote sensing [7], but the most significant results are obtained after a post processing of supervised classification, which requires prior knowledge of the ground cover in the

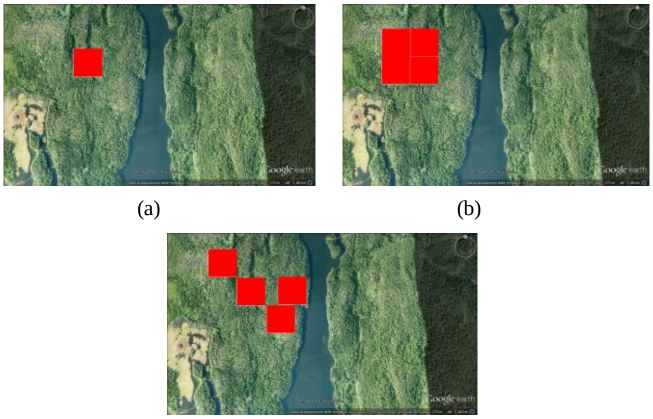
ISBN: 978-1-61208-283-7

study area. Besides, the two main problems of these systems are the coarse map resolution and a limited number of classes. A manual supervised classification is the most effective method, as in Corine Cover Project [8], with 400 classes of land cover, whose data have been used in our approach.

In conclusion, contributions in literature underline that the problems of finding significant indicators for land use and their supervised or unsupervised computation are still an open challenge, with dramatic consequences on environment preservation.

II. A NEW INDICATOR FOR LAND USE

In this research, the first step to measure the impact of human settlement on environment is the computation of a new two-dimensional global metric. The metric is based on the concept of *anthropentropy*: this term is a neologism [9] derived from the Greek term *Anthropos* (*Άνθρωπος*) = man, and *entropy*. As entropy is the well-known measure of disorder of a system, the term anthropentropy refers to the "disorder" introduced in a virgin environment by the presence/penetration of human beings. The basic idea is to associate to a limited geographic area a new indicator, called *Anthropentropy Factor* (*AF* in the following), which expresses the degree of penetration of human settlement in the environment. This is not only based on classical well-known indicators, such as the percentage of soil occupied by human activities and urban expansion, but it takes into consideration also the *shape* of the areas subtracted to nature. For this reason, we have chosen the term two-dimensional metric, where the two dimensions refer to a two-dimensional plan description of the areas of human settlements and their relative shapes and spatial relationships, in particular their contiguity. This choice is supported by the results of several studies in ecology: many researches highlight that the shape of the land is important for wild animal species that inhabit it and that the fragmentation of the territory contributes greatly to limit some environmental key treats, for example biodiversity. In fact, the UN Convention on Biological Diversity [10] considers fragmentation as a major threat to habitats and species populations, because it has a direct impact (due to proximity and disturbance) in creating barriers to the wandering and spreading of animals. One of the most important causes of fragmentation of territory is a disordered urban expansion where the incremental areas (due to new human settlements) are distributed in the territory in such a way that new settlements are not contiguous to existing ones. This is potentially a great drawback, because it increases greatly fragmentation. To illustrate this concept with a simple schematic example, suppose that the initial situation of an area of 16 square km is that of a single urban settlement of square shape of 1 km (Fig. 1a): as a consequence of the urban expansion, this area gets wider, by rising the measure of the side to 2 km (Fig. 2b). The increase of the use of soil passes from 0.0625 to 0.25. Now let us suppose instead of switching to a new situation in which there are three other new settlement squares of 1 km each (Fig. 1c). The new occupation of the land use remains the same (0.25), because cases (b) and (c) have the same relative land take increment (+25%), but the situation, regarding the fragmentation of territory, is clearly more dramatic. Despite the calculated indicators are the same, the first scenario (b) is definitely preferable than the second (c) from the ecological point of view. A simple qualitative example is the construction of a road in the countryside: the impact on wildlife is not just due to the simple loss of land, most likely because the animals of the area will not only reduced their amplitude territory occupied by concrete and asphalt, but it is slightly larger, due to other aspects of environmental pollution (noise, vibration, impossibility to cross the road).


(a) (b)

(c)

Figure 1. The sprawl of urban settlements causes a loss of continuity and consequent territory fragmentation. (a) original situation; (b) contiguous incremental areas (c) non contiguous incremental areas.

For this reason, by considering the increment of urbanized areas, it is important to notice that the relative numeric increment of land use is not a sufficient indicator to take into account. In [11] the problem of land fragmentation is analyzed in full details, but with two great limitations: data refers to 2000 and have no longer updated and (2) the fragmentation taken into consideration is the only one induced by great communication roads. However, if you want to measure the quality of life of individuals, the use of such macroscopic level data may not be sufficient. In order to overcome the limitations of the classical approaches of previous studies, and to improve our knowledge about environment preservation, it is mandatory (a) to update data to a more recent situation and (b) to take into consideration several urban settlements, both from the point of view of considering all their possible typologies, and also their spatial distribution, in order to minimize fragmentation. How you can put all this logic in an indicator that: a) takes into account the relative positions of the urbanized areas and b) is easy to compute and to be updated? The solution proposed here is the metric based on anthropentropy and the consequent definition and computation of indicator *AF* on a given geographic area in a collaborative environment, using the paradigm of User-Generated Content (UGC). The metric and the AF definitions answer challenge (a), while to answer the challenge (b), the second novelty of our research is adopted: *social crowdsourcing*, by involving in the project a social networking community in order to generate open data.

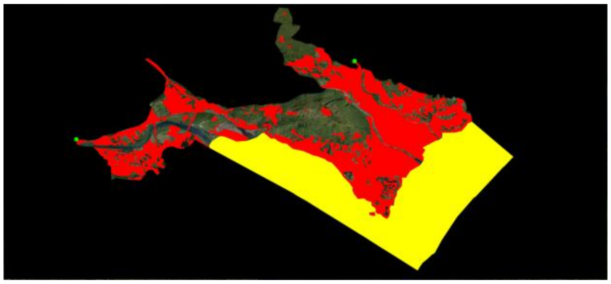


Figure 2. Example of anthropentropy mapping for municipality of Verbania (Italy): in red the anthropic places, in yellow the neutral area.

A. *The AF definition*

The solution consists in considering an urbanized area and then "enlarge" of a security zone by expanding it shape in the two dimensions and, consequently, by increasing its size. The extension has the purpose of taking into account the negative effects (due to noise and pollution) in the immediate vicinity of human urbanized areas; moreover, if the urbanized areas are sufficiently close each other, the enlargement causes an effect of dilation (where contiguous "holes" can be filled), which defines wider areas of individual settlements. The indicator must take into account these enlarged areas, instead of the original ones.

In order to investigate the human presence as the occupation of land by human resources that are not compatible with nature, the whole set of *anthropic places* are defined by four typologies (in the following, referred as *List 1*): *buildings of any kind* (housing, workplaces, schools, hospitals, if there is the actual presence of human activity, also in occasional mode, for example a settlement of voltaic solar panels), *paved roads, railways,* and *places of intensive agriculture* (such as greenhouses, nurseries, where the human presence is fairly continuous and has effectively ousted the wild).

In order to compute the *AF*, some considerations and definitions are necessary. Let us consider a generic geographic region bounded by recognized borders (e.g., a municipality, a county or a state) and let define S the area (in squared kilometers) of the region (assuming you have an appropriate two-dimensional scaled map and the numeric value of S). Within the region, we proceed to the identification of all the sub-regions occupied by the anthropic places of *List 1*. Sub-regions can be disjoint. Each sub-region contains at least one of the anthropic places listed in *List 1*. We define a *neutral sub-region* as a part of territory containing at least these two kinds of elements: (a) stretches of inland water (lakes, lagoons etc.) extending more than two squared kilometers (this choice is to be consistent with the coastline, which is normally considered as administrative boundary of the coastal municipalities) and (b) the land located more than 3,000 m above sea level. For the given region, the union of all the neutral sub-regions (if present), correspond to the *Neutral Zone*. Let define *NA* as the area (in squared kilometers) of the Neutral Zone. Obviously, not all the geographic region necessary contain a neutral zone, in that case NA is set to 0. In Fig. 2, an example of anthropentropy mapping of the municipality of Verbania is shown. The portion of the lake is the yellow area, the red area are the anthropic places (for clarity, the territory outside the municipality is blackened). Each area occupied by anthropic places is enlarged (along the two Cartesian dimensions X and Y) with a factor of "charging" of 50 meters, to give rise to anthropic sub-regions; for example, suppose a generic sub-region is a square of side 1000 meters, then the correspoding anthropic sub-region will be a square of side 1100 meters. The choice of a "buffering" of 50 meters is based on the principle of reciprocity: in Italian laws, most particularly intrusive activities, e.g., hunting, shall be allowed only at a distance greater than 50 m from urban settlements. This distance then seems a good compromise between the two extremes of a too restrictive and a too permissive limit. We define the union of all the anthrophic sub-regions as *Death Zone* of the region. Let define *DA* as the area (in squared kilometers) of the Death Zone. We define the *Anthropentropic Factor AF* as the ratio:

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

The *AF* is a fractional number between 0 (completely uninhabited regions without human active settlements) and 1 (fully populated regions). In (1) the special case of $NA = S$ is not considered, as it would mean that the entire geographic area is occupied by water or it is located above 3,000 m above the sea, thus it is not suitable to urban settlements and the computation of FA is meaningless.

B. *Critical issues in the AF computation and their solutions*

The main purpose of the ACI project is to map the complete Italian territory in such a way that, for each municipality, the AF factor (1) is computed according to the procedure described in Section II.A. The first problem is the availability of a description of land use in terms of anthropic places (see *List 1*), possibly with an accuracy of scale comparable with the size of the dilation (50 m) accomplished by the algorithm. At this purpose, the ACI project refer to the data-set created by the Corine Land Cover (CLC) project [8]; it was born in Europe specifically for the detection and monitoring of the characteristics of cover and land use, especially for environmental protection. The first realization of the CLC data-set refer to 1990 (CLC90) and subsequent updates refer to the year 2000 through the project Image & Corine Land Cover 2000. In 2000, 33 countries adhered to the project, including Italy.

Unfortunately, the first critical issue of the project is that Corine data set is not available for all the Italian territory, but only for 7 regions (over a total of 20 regions), namely for the 40,9% of municipalities (3311 over 8092). For the remaining municipalities, data of land use are not available at all, or they are not available free of charges. To fill the gap of data, an UGC approach has been use (see Section II.C). As our project stresses the importance of open data for society growth, we had to face this problem and we solved it by proposing a solution based on User Generated Content and social networking. Therefore, the results of the ACI project derive from two distinct procedures: (a) the AF mapping based on Corine Land cover data set (due to the avalibility of Corine data set, this part can be developed for a total percentage of 40.9% of Italian municipalities), and (b) the AF mapping based on UGC and social networking open data.

In Section III, results for the two cases are described. First, the procedure for UCG and open data generation of our collaborative project has to be described.

C. UGC and collaborative social networking for AF computation

The ACI project can count on a social network consisting of 7000 users [12]. All of them are potentially involved in the project of UGC. Collaborative users are asked to generate a map (UGC map) on Google Earth environment in which the limits of anthropic sub-regions and neutral zones (if any) are plotted. Obviously, each user will produce the map for the municipality in which he/she actually lives: in this way the knowledge of one's own territory, provided by everyday user experience, is coded in the map. A procedure for the generation of the UGC map of the territory is given to the users of the social network. The procedure must have the following characteristics: (a) it requires a *low level of expertise* in informatics: the only required skills are to be able to open files and do some basic operations such as selecting and tracking area boundaries filled with different colors (red for anthropic sub-regions and yellow for neutral zones), and (b) it is based on *open software* and *open data* (Google Earth maps and Gimp). The user is given a detailed procedure to download the Google Earth map and to fill areas with the proper set up of Gimp tools (types, fillings, colors). In Fig. 3a, an example of UGC map for the municipality of San Martino Siccomario (Pavia, Italy) is shown: red areas refer to the anthropic places, in black the limits of the municipality are shown. After the user has generated the UGC map of his/her own municipality, we received the map and processed it according to a simple computer vision algorithm, implemented in Matlab. It is composed of four steps. The first one is a calibration step, which is necessary to discover the scale of the UGC map. At this purpose a couple of points are select on Google earth original map and their distance is measures in Google Earth with the Ruler tool. The same two reference points are selected in matlab interactive environment and the Euclidean distance is computed (see Fig. 3b, the green line is drawn for euclidean distance and computation of pixel dimensions). This gives the effective area represented by one pixel, which is assumed to be squared. In the second step, the red areas are enlarged by the well-known morphological operator of dilation [13], in X and Y dimensions, assuming a radius of dilation of 50 meters (see Fig. 3c). This step creates the Death Zone. By comparing Fig. 3b and Fig. 3c, it can be noticed that areas are enlarged and, in some points, "holes" are partially or completely filled (further examples are reported in Experimental results, see Fig. 6). In the third step, the areas of dilated zones and neutral zone (if any) are computed by simply counting the number of pixels of the Death Zone and of the Neutral zone (if any) and by multiplying these numbers for the effective area of a pixel computed in the calibration step. This step gives rise to the values of DA and NA in (1*).* In the last step, the area of the entire municipality is read form official on line Istat database [14] and AF factor is computed according to (1).

D. Discussion on applicability, merits and limitations

The first important question is: is AF a significant factor? This indicator gives, in an absolute scale (from 0 to 1), the degree of land use and it is very intuitive to understand. Consider the *paradox of cement*: if every Italian citizen occupied with his/her business (home, work, communication routes) an area of 70 X 70 meters, the whole Italian peninsula (including mountains!) would be populated and the factor AF would be equal to 1 everywhere. The distance from value 1 indicates how far we are from this tragic limit situation. The second important question is: in which aspects can the AF help in environment preservation? As the metric is based on an objective, scientific indicator and it is very simple to understand, it can be a very powerful tool to help legislators and local administrators to evaluate the present situation and to impose limits to urbanization.

Some limitations of the approach have to be underlined, especially in the part of UGC data set, in particular: (a) the precision of Death Zone and Neutral Zone definition varies from map to map, as it is related to the manual skill of selection by mouse of the user (this precision cannot be estimated in advance in a quantitative way), and (b) the knowledge of the land use of the territory varies from map to map, because it is related to the awareness of land use of the user. For the first limitation, several tests have been done to assure that the precision of manual selection by the mouse of the two reference points does not affect the computation of FA. Even with a simulated error of selection up to five pixels, the variation of the indicator AF is less than 0.01%. For the second limitation, a preliminary check on the quality of the UGC map is performed manually, by a strict cooperation with the social network users, in order to verify the level of their expertise. Obviously, the UGC values of AF cannot replace the same precision the computation based on Corine data set, but as the main goal is not replacing but filling the gap of open data, this method can be a valid solution: besides, the method based on UGC maps have three main advantages. First of all, it increases the awareness of the citizens on the state of the environment where they lives. Map reading and, above all, participation to complete them makes us more aware of the environmental richness that we are losing. And it is, after all, the ethical purpose of ACI project, to show that crowdsourcing and social technologies can help to improve society. Besides, the procedure for AF computation can be used as a simulation to study future urban expansions in a *scenario* of *What if?*. This suggests a possible very useful application of our system: to help local government in planning the annual Territory Government Plan where new urban expansions are decided. Finally, it can react rapidly to the local changes for land use, as the main source of knowledge is the citizen (a typical bottom-up knowledge). For what concern the comparison to a fully automatic image processing approach, as we pointed out in

Section I.A (Related work), a fully automatic method to recognize anthropic places is not possible, because general image processing characteristics, such as multispectral information, color and texture are not sufficient to code in a map the level of knowledge necessary to differentiate a nursery for intensive agriculture from a field of sorghum (extensive agriculture), a golf course from a simple lawn, or a complex of abandoned and inhabited houses from an habited one.

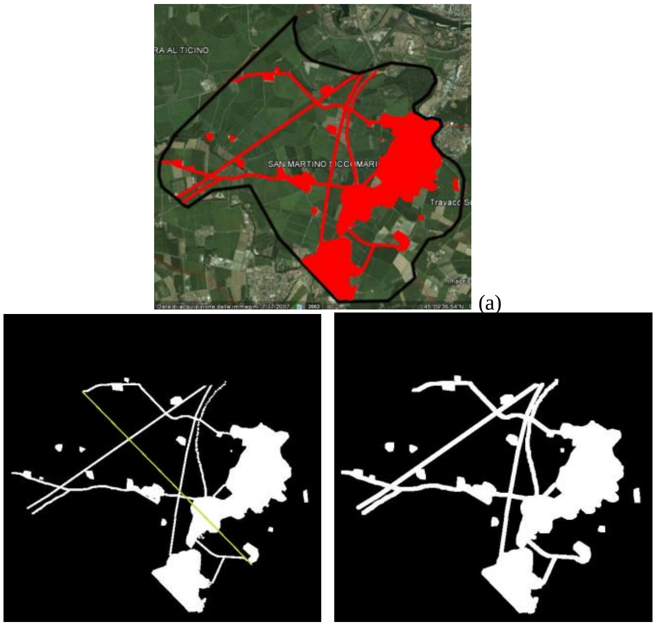
III. THE METRIC AND EXPERIMENTAL RESULTS

A. *The metric*

After the computation of AF index, a simple numeric (and the corresponding visual representation) metric is proposed here: if AF is between 0 and 0.2, the area is considered a very low “anthrophized area” (the ideal condition for the environment). This condition is represented visually on the map by coloring the area into varying levels of green. If AF is between 0.2 and 0.4, the area is considered with a limited but worrying level of anthropentropy. This condition is represented visually on the map by coloring the area into varying levels of yellow. This type of area have to be monitored in time to control its evolution potentially to undesired higher level of anthropentropy. If AF is between 0.4 and 0.6, the area is considered with a serious level of anthropentropy. This condition is represented visually on the map by coloring the area into varying levels of red. In these areas, the presence of humans greatly impact negatively on environment. If AF is between 0.6 and 1, the area is considered with a very serious level of anthropentropy, in an irreversible environmental degradation. This condition is represented visually on the map by coloring the area into varying levels from violet to black. In these areas, the presence of humans has completely compromised the situation. Following the metric and the visual convention on visual mapping, for a given area of the Italian territory (a group of municipalities, i.e., region), it is now possible to generate an *Anthropentropy Map*. In Section III.B, significant results are shown and commented.

B. *Experimental results*

In Fig. 4, an example of Anthropentropy Map for region Trentino (Italy) is shown. The most compromised areas are near big cities, and this is, to a certain extent, a predictable result, however, the situation is quite good because no black areas are present even on the cities, and red areas are limited both in extension and dispersion. This can be quantitatively expressed by the statistics: the minimum value of AF is 0.01, the average is 0.14, and the standard deviation is 0.11. From the anthropentropy map, a low level of fragmentation of green territory is evident. In Fig. 5, the most compromised situation is shown. The black spot refers to the high urbanized area of Milan. The map shows clearly however that the worrying situation is also far off the big cities, in countryside, due to local uncontrolled urban settlement growth. In Fig. 6, some examples of UCG maps are presented, and the corresponding Death Zone computed by the computer vision algorithm based on morphological dilation. The maps refer to the municipalities of Erice (Sicily) and Scarperia (Tuscany), with AF of 0.214 and 0.1637, respectively.



(a)

(b) (c)

Figure 3. AF computation for the municipality of San Martino (Pavia, Italy): (a) original UGC map (b) calibration step (c) Death Zone.

IV CONCLUSIONS AND FUTURE WORK

In this paper, an innovative indicator to evaluate land use is proposed, basing upon the concept of anthropentropy. According to the definition, a method to compute the Antropentropy Factor for a given territory is described, both on open data and on collaborative UGC maps. On the time of writing of the present paper, the ACI project covers a percentage of generation of Anthropentropy Maps for the whole set of Italian municipality (8092) equal to 36,7%, and new results are constantly being collected, both from the social collaborative approach and on new, recently available, open data. Future work includes the mapping of the entire Italian territory and the presentation of results to local and governmental administrations, in order to put in practice an innovative policy of environment preservation based upon a sustainable land use and to measure with some metric the efficacy of the method on real cases of environmental degradation.

ACKNOWLEDGMENT

The authors greatly appreciate for his collaboration Antonio di Gennaro [15], for AF and map generation from Corine data set.

REFERENCES

[1] Overview of exiting Policies to reduce and mitigate soil sealing in the EU and Member States, http://ec.europa.eu/environment/soil/pdf/sealing/4.%20Overview%20of%20existing%20Policies%20EU%20and%20MS%20(A-H).pdf [retrieved: May, 2013].

[2] Land Use, State and impact (Italy). http://www.eea.europa.eu/soer/countries/it/land-use-state-and-impacts-italy [retrieved: May, 2013].

[3] J. Dumanski and C. Pieri, "Land quality indicators: research plan," Agriculture, Ecosystems & Environment 81.2, 2000, pp. 93-102.

[4] J. D. Majer, “Ants: Bio-indicators of minesite rehabilitation, land-use, and land conservation,” Environmental Management, Vol. 7, n. 4, Springer Verlag, 1983, pp. 375-383.

[5] H.G. Zechmeistera, I. Schmitzbergera, B. Steurerb, J. Peterseila, and T. Wrbkaa, “The influence of land-use practices and economics on plant species richness in meadows,” Biological Conservation, n. 114, Elsevier, 2003, pp. 165–177.

[6] http://www.arpa.veneto.it/temi-ambientali/aria/file-e-allegati/docup-2007/09-Isocrono.pdf [retrieved: May, 2013]

[7] J. Rogan and D.M.Chen, “Remote sensing technology for mapping and monitoring land-cover and land-use change,” Progress in Planning, n. 61, Elsevier, 2004, pp. 301–325.

[8] Corine Land Cover Project: http://www.eea.europa.eu/publications/COR0-landcover, [retrieved: May, 2013]

[9] ACI project: http://www.albanesi.it/Ambiente/ACI.htm [retrieved: May, 2013]

[10] The UN convention on biological diversity, http://enrin.grida.no/biodiv/biodiv/cbd/eeafoll.pdf [retrieved: May, 2013]

[11] Fragmentation of ecosystems and habitats by transport infrastructure, http://www.eea.europa.eu/data-and-maps/indicators/fragmentation-of-land-and-forests/eu-ac-fragmentation. [retrieved: May, 2013]

[12] Facebook page: https://www.facebook.com/RobertoAlbanesi, [retrieved: May, 2013]

[13] R. C. Gonzales and R. E. Woods, Digital Image Processing, Pearson Prentice Hall, 2008, Chapter 9, “Morphological Image Processing”.

[14] ISTAT: www.comuni-italiani.it [retrieved: May, 2013]

[15] Linkedin page: http://www.linkedin.com/in/adigennaro [retrieved: May, 2013].

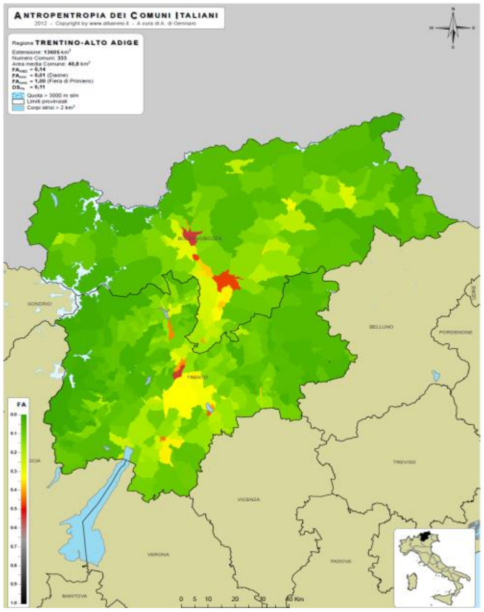


Figure 4. Anthropentropy Map of RegionTrentino, Italy (land area: 13605 squared kilometers).

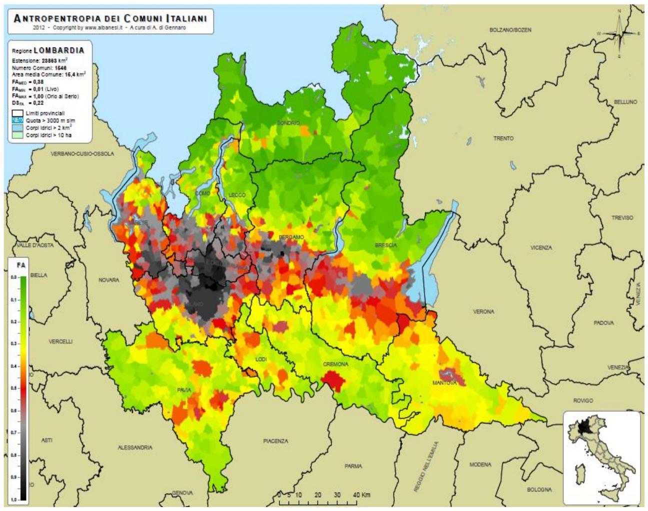


Figure 5. Anthropentropy Map of Region Lombardia, Italy (land area: 23863 squared kilometers).

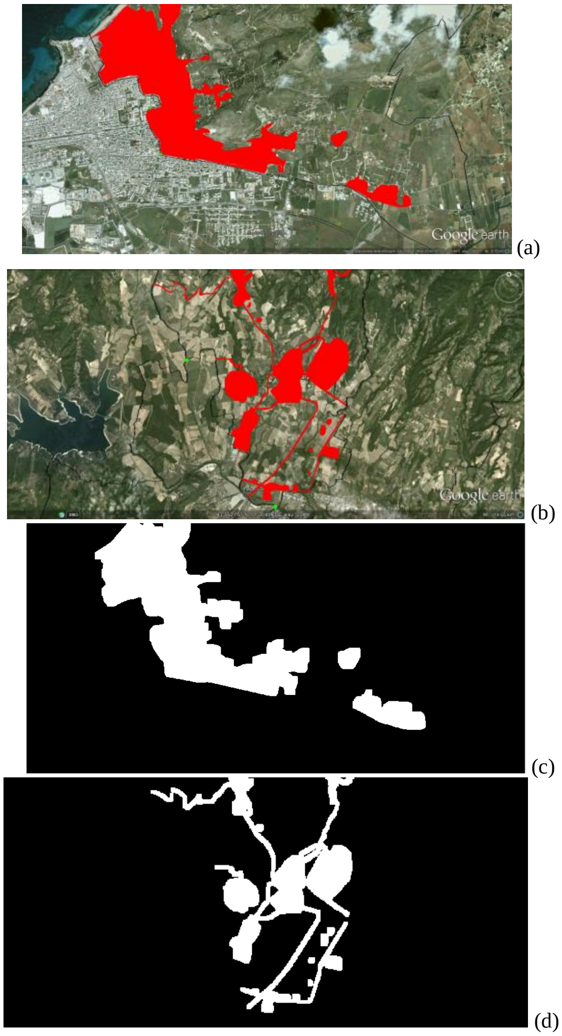


Figure 6. Example of UGC maps for the municipalities of (a) Erice (Trapani, Sicily) and (b) Scarperia (Florence, Tuscany) and their corresponding Death Zones (c and d).