



The 7th International Conference on Ambient Systems, Networks and Technologies  
(ANT 2016)

# Topological and Conceptual Complex Network Models for Environmental Planning

Gianni Fenu<sup>a,\*</sup>, Pier Luigi Pau<sup>a</sup>

<sup>a</sup>*Department of Computer Science, University of Cagliari, Via Ospedale 72, 09124 Cagliari, Italy*

---

## Abstract

The growing importance of environmental planning has encouraged researchers to apply complex network analysis on topological models of environmental networks. Relevant features of current green infrastructure can be derived with common and ad hoc techniques, but results tend to expose only a limited view, whether by geographical areas or by species. In this paper, the possibility to extend complex network analysis to conceptual models with a higher degree of abstraction is explored.

© 2016 The Authors. Published by Elsevier B.V. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer-review under responsibility of the Conference Program Chairs.

doi: 10.1016/j.procs.2016.04.107.

*Keywords:* Complex networks; Environmental networks; Knowledge discovery

---

## 1. Introduction

Environmental planning for landscape development is growing to be a very important topic, as the preservation of wildlife even in an increasingly urbanized society has become a priority for many countries. In the European Union, policies concerning the environment and the protection of biodiversity have converged into the definition of a wide environmental network denominated “Natura 2000”, consisting of nature protection areas designated by the EU and at national level. To understand the implications of the creation of an environmental network, it is sensible to study these networks using complex network analysis (CNA). Applications of these techniques have been successful in a varied amount of fields<sup>1</sup>, including power grids and social networks, often uncovering properties of real networks that were previously thought to be properly modeled with random graphs<sup>2</sup>. CNA is beginning to be applied on topological models of environmental networks, but in many applications, graph models need not be tightly coupled with an underlying geographical structure; rather, graph models can be derived for conceptual entities. Particularly striking is the case for bioinformatics, with biological meanings being found for many mathematical indices<sup>3</sup>.

In this paper, current applications of CNA on environmental networks are reviewed, and possibilities are explored to expand its use to increase the understanding of their features, working at different degrees of abstraction. First, in Section 2, the concept of environmental network is introduced, and its implementations are discussed, providing a

---

\*Corresponding author. Tel.: +39-070-675-8759; fax: +39-070-675-8779.

E-mail address: [fenu@unica.it](mailto:fenu@unica.it)

view of administrative details as well. Section 3 describes the activity of data collection on current nature protection areas that make up the Natura 2000 network. In Section 4, a framework for the analysis of data is established. In Section 5, a case study is discussed, involving reports filed for protection areas in Italy, and more specifically in Sardinia. Lastly, in Section 6, conclusions are drawn and opportunities for future work are outlined.

## 2. Natura 2000 Basics

Current society is characterized by an increasing degree of urbanization, as most human activities are now focused on cities, as opposed to rural areas. Human activities require contiguous infrastructure for connections between cities, for transportation of people and goods, and for the continued provision of many communication services (telephone, Internet, etc.). As increasing amounts of land, including coasts, become part of urbanized systems, the environmental balance is at a peril of being destroyed, with the elimination of habitats from land and sea. Concern for the protection of the environment, particularly endangered species, has given rise to the creation of nature protection areas.

However, since the contiguity of land for human activities has been held as a priority, these areas have long existed only in the form of isolated regions, resulting in a fragmentation of habitats. This is a major limitation, as it has been shown that factors such as an insufficient size of preservation areas, or an excessive distance between areas, can hamper their effectiveness<sup>4</sup>. In addition to this, even when the size of areas is sufficient for the survival of species to be preserved, biodiversity can be negatively affected, due to the frequent cases of inbreeding.

### 2.1. Elements of an environmental network

To address the fragmentation problem, in more recent years, the concept of isolated protection areas has been replaced with a reticular model<sup>5</sup>, in which sites are connected with one another by elements of “green infrastructure”, or other steps are taken to enable a conglomerate of sites to act as a network, in which each site, rather than being setup only to protect very specific categories of wildlife, has a role in contributing to larger preservation goals.

Where network behavior can not emerge from conservation areas by themselves, a commonly implemented way to provide a connection between nature protection areas is the creation of a ‘habitat corridor’ (also known as ‘wildlife corridor’ or ‘green corridor’), defined as a linear strip of habitat connecting two or more larger patches of habitat, surrounded by a dissimilar matrix<sup>6</sup>. Corridors are intended to act as a path for the migration of some animal species, whether ordinarily or in the event of catastrophes hitting one of the larger areas; furthermore, corridors can provide a stable habitat for plants, insects and other smaller animals.

A habitat corridor ought to be tailored to the needs of the species intended to use it<sup>7</sup>. In different cases, it may be more sensible to form either a contiguous strip of land, or a set of disconnected patches (referred to as ‘stepping stones’), arranged to form a line. The usefulness of corridors as a way to improve the conservation of species has often been put into question<sup>8</sup>, and there is still no widespread consensus regarding the long-term effects of habitat corridors. However, it has been shown in a number of studies that their implementation has had a good effect for the biodiversity of some categories of species at least in the short-term<sup>9</sup>.

In the European Union, nature and biodiversity policies have converged in the establishment of a network of nature protection areas, denominated “Natura 2000”. Sites are designated as follows:

- Special Protection Areas (SPA), which are designated by member states according to the EU Birds Directive (2009/147/EC);
- Special Areas of Conservation (SAC), designated by member states according to the EU Habitats Directive (92/43/EEC).

A Special Area of Conservation is generally designated in a two-step process: first, a site is proposed by a member state to become a Site of Community Importance (SCI); once it has been approved as such by the EU, the member state can designate it as a SAC. It is possible for the boundaries of SPAs to overlap with those of SACs and SCIs, and a site can be designated to be a SPA as well as a SAC or SCI at the same time. Natura 2000 sites may include areas dedicated to human activities, as well as privately owned land. In these cases, the site should not be treated as a strictly enforced reserve where human activity is forbidden, but rather as an area where the EU seeks a sustainable

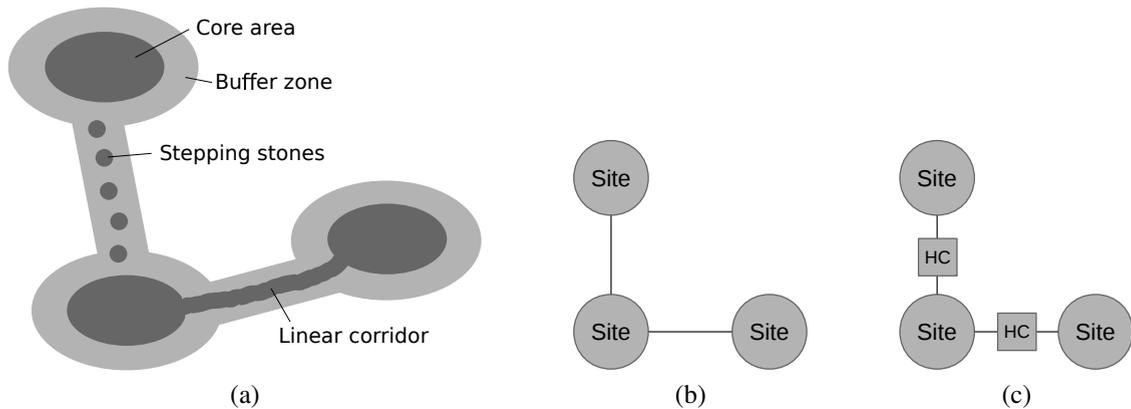


Fig. 1. Topological graph models. (a) Sample configuration of three protected areas and two habitat corridors: a contiguous one, and a set of stepping stones. A buffer zone may be present, acting as a transitional area to keep protected sites at a distance from heavily urbanized areas. (b) A graph model representing this sample configuration, with habitat corridors modeled as edges. (c) A graph model representing the same configuration, with habitat corridors modeled as nodes, and assuming connectivity between protected zones and corridors.

management of human activities and natural resources. Consistently with the concept that a network behavior can emerge in different ways, connection elements are not as thoroughly regulated as sites are.

## 2.2. Environmental complex network analysis

To assess properties of how Natura 2000 sites can act as a network, it can be relevant to build a topological graph model of the network, and apply complex network analysis techniques to extrapolate its advanced properties.

At first glance, a graph model of an environmental network should be rather straightforward to build; graphs are made up of two basic elements, nodes and edges, which generally represent entities and links between entities, respectively. In this context, nodes will represent sites, and edges will be used to represent the existing or potential connections between them (see Figure 1a and 1b). In spite of this apparent simplicity, some nontrivial design choices are necessary – for instance, concerning scale, granularity of data, or the selection of species of interest.

Applications of complex network analysis on topological models for environmental networks are often focused on evaluating the efficiency of current infrastructure with respect to a definite goal. Studies have been performed on different scales and goals, ranging from conservation of bird species at a continental level<sup>10</sup> to dispersal of plants in a peri-urban setting<sup>11</sup>. These topological graph models are generally referred to a single species or a very limited set of species. Their nodes represent sites or patches hosting suitable habitats, and edges are drawn between a pair of nodes if specimens of the species being considered can move between them. This can translate to a suitable habitat corridor being available or, like in the case for birds, the distance between nodes being within their dispersal distance. Note that, if contiguous corridors are not needed, corridors in general can be treated as additional habitat patches, i.e. nodes or sets thereof (a simple example is in Figure 1c).

The graph resulting from a topological model can be analyzed to extract network properties, which may not be obvious from its geographical map. The most crucial observation is whether the graph is connected (i.e. a path exists between any two nodes), or made up of several connected components. If a graph is instantiated for a single species, the latter case is a sign of the existence of separate metapopulations<sup>12</sup>. Node degree, i.e. the number of nodes adjacent to a given one, and betweenness centrality, which expresses a proportion of shortest paths that include a specific network element (a node or an edge), are used to rank network elements by importance. In fact, if nodes are removed from a network – for example, to represent natural disasters – a larger number of shortest paths is affected if those nodes have a high betweenness centrality index, and this is often linked to a stronger impact on network connectivity<sup>13</sup>. Another useful indicator is clustering coefficient, which expresses the proportion of neighbors of a node that are connected to one another, evaluated as a local or global feature. If evaluated locally, nodes with a low clustering coefficient can be associated with portions of the network where alternative paths are scarce or even absent.

### 2.3. Administration of Natura 2000 sites

Although the Natura 2000 network is defined at EU level, the activities of land management and planning are left to the competent administrative offices at national or local level. In Italy, regional administrations are in charge of integrated plans, as well as management plans for specific Natura 2000 sites. Specific competences can be delegated by the regional administration to local offices (Province governments or municipalities). A management plan consists of a document outlining the characterization of a site from multiple points of view (among which: extension and boundaries, subdivision into areas, climate, presence of pollution, habitats, social and economic aspects, landscape), providing a list of possible threats to the continued presence of protected species and habitats, and finally the proposed actions for their conservation, including an assessment of priority for each proposal.

Decisions regarding management plans are determined by technical, regulatory and political aspects; software tools may provide valuable assistance in evaluating technical aspects, among which are topological and geographical constraints derived from the features of the region, biological features of the target species, and budget constraints. As seen previously, complex network analysis is useful to extract features of topological models, even with the limitations derived from building graph instances referred to a small number of species or situations.

## 3. Data Collection and Conceptual Analysis

Reports on each Natura 2000 site (or proposed site) are periodically filed, with the purpose of collecting data on the habitats and species found within the area. Information is gathered on-site by researchers in accordance with a Standard Data Form, and written to a data base. The current version of the Standard Data Form as of this writing was released with Commission Implementing Decision 2011/484/EU, replacing the version previously released with EU Commission Decision 97/266/EC. The form must be filled for a set species of interest as defined in Article 4 of Directive 2009/147/EC, and listed in Annex II to Directive 92/43/EEC.

The Standard Data Form consists of seven sections:

1. Site Information: general information on the site under analysis, such as an identification code, classification (as SPA, SCI, SAC), dates of first and latest update.
2. Site Location: coordinates on Earth (the coordinates of a centroid are given), extension, percentage of marine area, administrative region, etc.
3. Ecological Information: data on each habitat found within the site, with an evaluation of quality on the site (state of conservation, etc.); data on each Annex II species found within the site, with an evaluation (size of population, whether a settlement is permanent, etc.); data on other species considered relevant by reporters.
4. Site Description: percentage of extension cover of each habitat, notes on quality and importance, threats, activities, etc.
5. Site Protection Status: optional references to local designations and relations to other sites.
6. Site Management: administrative references to the bodies in charge of site management.
7. Map of the Site: a graphical map in PDF format is optionally attached.

It follows from item 4 in the list that Natura 2000 sites are made up of patches of different habitats, and each patch may be characterized by the presence of different species. However, item 3 shows that, as part of this data collection activity, habitats and species are linked to the site for which the report is filed, but they are not linked to one another (see the relevant excerpt of the resulting conceptual Entity-Relationship diagram, shown in figure 2). Aspects such as the knowledge of which habitat is ideal for each species are not directly stored, and are supposed to be determined by expert knowledge and external documents.

## 4. A Framework for the Analysis of Species Relations

Complex network analysis has proven to be a successful tool for knowledge discovery in other fields, such as bioinformatics and social networks, even in the absence of a topological model. As such, it can be used to provide

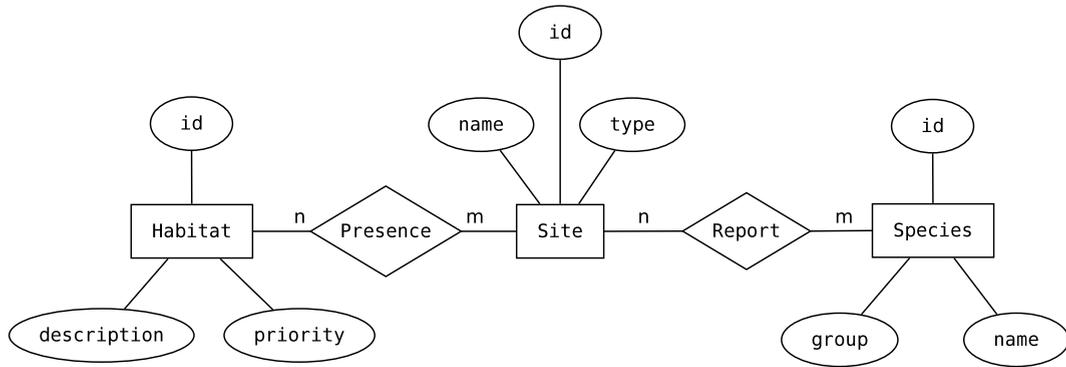


Fig. 2. Simplified Entity-Relationship Diagram used for reporting the presence of habitats and species in each Natura 2000 site. Attributes not shown encompass the evaluation of the presence of each species in a site, e.g. depending on their number and whether they reside permanently or only in certain seasons. Note that habitats and species are not put into a relationship in this data model.

insight into inter-species relations, in order to tackle the problem of choosing species that may represent an interesting object of study with a topological graph approach.

Entities from the data base in figure 2 can be used to build a conceptual network, in which nodes represent species, and edges represent interactions or affinities between species. Several conceptual networks can be built from data made available on the species reported to be found in Natura 2000 sites, varying the criteria used to determine the set of edges, and analyzed with techniques derived from complex network theory. In a simple scenario, links are formed between pairs of species that have been reported to be found in a minimum number of Natura 2000 sites; the resulting graph can be referred to as a shared-report graph. Basic examples of site configurations and resulting graphs can be seen in Figure 3, and used as a frame of reference to interpret results from complex network analysis.

If an area consists of a single site with  $n$  species (or any number of identical instances thereof), the shared-report graph is a complete graph of order  $n$  (Figure 3a shows an example for  $n = 4$ ). As a consequence, for each site in an area under analysis, the subgraph corresponding to the species found in that site will be a complete graph. The  $ABC$  triangle in Figure 3b and 3c is an example thereof.

At each end of the spectrum of possibilities, the ecological network in an area can be made up of a single connected site – corresponding to a complete graph – or a number of sites, each hosting a single species of interest, resulting in a graph with no edges.

In real-world scenarios, sets of species may appear multiple times, including as subsets in sites with many species (see Figure 3d). The resulting graph can be treated as a weighted graph (the number next to each edge represents its strength); alternately, a threshold can be applied, to consider only edges with a certain strength or above. In this example, if a threshold of 2 is chosen, the resulting unweighted graph will consist of three connected components, node  $E$  being an isolated node, and the other components being the pairs  $\{A, B\}$  and  $\{C, D\}$ .

## 5. Analysis Setup and Results

In this case study, a number of shared-report graph instances are built using data on the 93 sites designated as SCI in Sardinia, one of the two main insular regions in Italy (at the time of writing, no site in Sardinia has yet completed its process to be designated as a SAC). Except in one experiment, only the species listed in Annex II were considered. This is because the data on other species of interest is provided by individual reporters without any kind of coordination or normalization, and actually contain some spurious data, due for example to typing errors or the use of aliases for species names. Lastly, an additional experiment is run on data inclusive of the 2314 sites designated as a SCI or a SAC in Italy. Complex network analysis was performed using Cytoscape<sup>14</sup> version 3.2.1.

The thresholding method explained in Section 4 is used to build several unweighted graph instances and compare results. Table 1 summarizes each experiment setup. In the leftmost set of columns, the area (Sardinia or Italy), set of species (Listed in Annex II or all species), and threshold are reported. The graph corresponding to this setup is

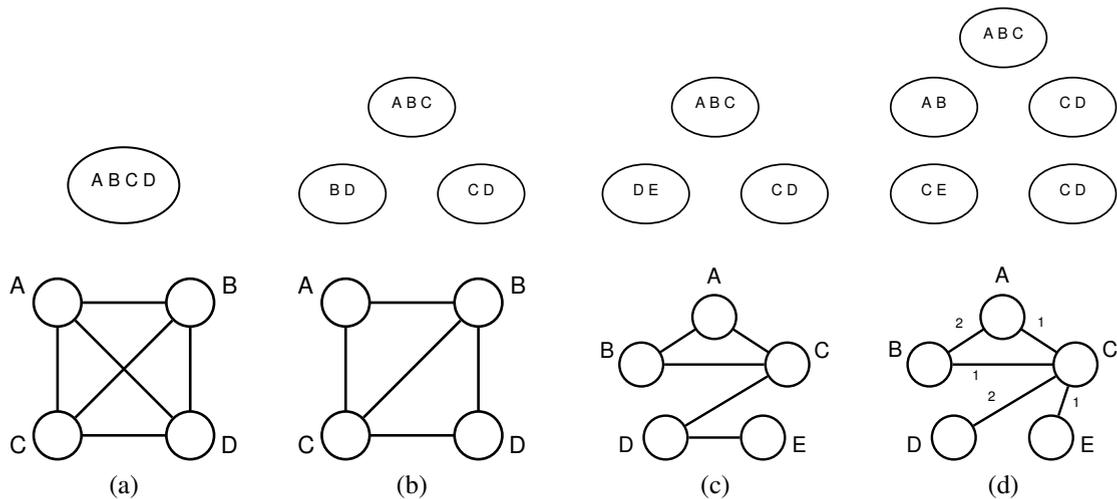


Fig. 3. In the top row, sample sites represented as sets of reported species. In the bottom row, shared-report graphs, where edges are drawn for each pair of species found in the same site. (a) Any number of species in a single site, four in this example, corresponds to a complete graph. (b) When not every species is found in any single site, the corresponding graph is not complete. (c) An example with five species, with species E being reported only with D on a single site; species E is a leaf node in the corresponding graph. (d) An example with sets of species appearing multiple times ( $CD$  twice;  $AB$  also twice, once as a subset in  $ABC$ ). The resulting graph can be treated as a weighted graph (the number next to each edge represents its strength); alternately, a threshold can be applied, to consider only edges with a certain strength or above. In this example, choosing 2 as a threshold cuts out E from the resulting unweighted graph.

built, determining its basic properties, such as the number of nodes and edges, reported in the two following columns. Lastly, complex network analysis is applied on the graph; results are summarized in the rightmost columns.

Table 1. Results of complex network analysis on preliminary models of species relations. Only areas designated as SCI or SAC are considered. See text for full discussion of parameters.

Criteria			Resulting network		Analysis results			
Area	Species	Minimum sites	Nodes	Edges	Isolated nodes	Density	Diameter	Characteristic Path Length
Sardinia	Listed	1	351	44 967	1	0.732	2	1.263
Sardinia	Listed	2	351	34 482	40	0.561	2	1.285
Sardinia	Listed	3	351	28 902	61	0.471	2	1.310
Sardinia	Listed	4	351	24 711	79	0.402	2	1.330
Sardinia	Listed	5	351	21 779	87	0.355	2	1.373
Sardinia	Listed	6	351	19 330	95	0.315	3	1.408
Sardinia	Listed	7	351	17 020	102	0.277	3	1.450
Sardinia	Listed	8	351	15 114	117	0.246	3	1.447
Sardinia	Listed	9	351	13 534	128	0.220	3	1.463
Sardinia	Listed	10	351	12 106	140	0.197	4	1.463
Sardinia	All	1	860	129 953	23	0.352	3	1.629
Italy	Listed	1	853	169 338	8	0.466	3	1.542

Going into further detail, the following is an explanation of the criteria:

- **Area:** the geographical region taken as reference. Data is to be extracted considering the set of Natura 2000 sites in the chosen region. In this study, the focus is on the whole Italian territory, or specifically on Sardinia.

- **Species:** the choice to consider only the species of interest listed in Annex II (“Listed”), or every species included by reporters (“All”). The data on the species of interest includes an evaluation in the form of structured data. Each species is represented by a node in the conceptual network.
- **Minimum sites:** the minimum number of Natura 2000 sites where a pair of species must have been reported, in order to add a link between the corresponding node pair.

The number of nodes is determined from the first two values, whereas the number of edges is determined from all three. In this study, the chosen criteria have always resulted in a network with one large connected component and a number of isolated nodes (i.e. nodes with no incident edges). Generalizing, it is possible for a network to be made up of multiple connected components of more than one node, but this hasn’t happened in these experiments.

Complex network analysis is performed on each instance, treated as an unweighted graph, extracting the following basic features:

- **Isolated nodes:** the number of nodes with no incident edges.
- **Density:** a value between 0 and 1, corresponding to the ratio of existing edges to possible edges. A value of 0 means that there are no edges, and thus every node is isolated. A value of 1 means that the network is fully connected, i.e. there is an edge incident to every pair of nodes.
- **Diameter:** the maximum shortest path length on the graph, considering all possible node pairs.
- **Characteristic Path Length:** the average shortest path length from an arbitrary source node to an arbitrary target node, considering all possible node pairs.

A striking result is that, as criteria are tightened to reduce the number of edges and isolated nodes appear in larger numbers, the diameter of the connected component is always very low. This can be interpreted as the list of reported species being made up of a ‘core’ of species commonly found together in many sites, and a number of species found in a limited number of sites. The fact that diameter is low even in the experiment run on the entirety of Italian land suggests that a wide set of species may indeed adapt to life in large portions of the national territory.

Lastly, a weighted shared-report graph was built using the data concerning sites in Sardinia. Edge weights were assigned according to the number of sites, as explained in Section 4. This instance was analyzed using Cytoscape with modifications to analyze weighted networks. As algorithms for computing betweenness centrality in weighted networks treat edge weights as a distance, weights were assigned based on the reciprocal of the number of sites, to represent the fact that stronger links are to be considered preferential in path construction.

Table 2 reports the top 10 species by betweenness centrality index, divided by the number of node pairs, expressed by  $(N - 1)(N - 2)/2$ , where  $N$  is the number of nodes, to normalize to a scale of 0 to 1.

Table 2. Ranking of species reported to be found in SCIs in Sardinia, by normalized betweenness centrality calculated on the weighted network.

Name	Type	Sites	Betweenness
Hyla sarda	Amphibian	71	0.12483
Turdus merula	Bird	71	0.07566
Upupa epops	Bird	60	0.05980
Falco tinnunculus	Bird	64	0.04169
Larus cachinnans	Bird	59	0.03935
Sylvia melanocephala	Bird	70	0.03734
Bufo viridis	Amphibian	61	0.03206
Carduelis chloris	Bird	66	0.02921
Carduelis carduelis	Bird	67	0.02921
Circus aeruginosus	Bird	52	0.02920

A higher value of the betweenness centrality index means that the node is included in a larger proportion of shortest paths between different species; in a network with a low unweighted diameter such as the one being analyzed, this can be interpreted in the sense that a high-betweenness node shares reports more often with both elements in pair of

species not sharing a report with one another, which can be attributed to higher adaptability or greater tendency to migrate, according to known features of the species. This measure can therefore be considered for the identification of cases with a greater potential of interest for further study, and the possibility to extract this measure using software tools indicates their potential to assist in further research.

## 6. Conclusions and Future Work

In this paper, the possibilities to apply complex network analysis in the field of environmental network have been explored, citing examples of topological analysis and reporting on an experiment involving a conceptual model based on the contextual reporting of species presence. As topological analysis is often focused on a limited number of sites or species, the latter activity can be of assistance in identifying those cases, which hold a greater potential of interest for study. In future work, results of conceptual network analysis will be used to draw guidelines to conduct further topological analysis; furthermore, the definition of more refined criteria to build conceptual networks will be considered.

## Acknowledgements

This essay is written within the Research Program “Natura 2000: Assessment of management plans and definition of ecological corridors as a complex network”, funded by the Autonomous Region of Sardinia (Legge Regionale 7/2007) for the period 2015-2018, under the provisions of the Call for the presentation of “Projects related to fundamental or basic research” in year 2013, implemented at the Department of Mathematics and Computer Science of the University of Cagliari, Italy. Pier Luigi Pau gratefully acknowledges Sardinia Regional Government for the financial support of his PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007-2013 - Axis IV Human Resources, Objective I.3, Line of Activity I.3.1.).

## References

1. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D. U. Hwang, Complex networks: Structure and dynamics, *Physics Reports* 424 (4–5) (2006) 175–308. doi:10.1016/j.physrep.2005.10.009.
2. R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of Modern Physics* 74 (1) (2002) 47–97. doi:10.1103/RevModPhys.74.47.
3. A. Lesne, Complex Networks: from Graph Theory to Biology, *Letters in Mathematical Physics* 78 (3) (2006) 235–262. doi:10.1007/s11005-006-0123-1.
4. J. M. Diamond, The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves, *Biological Conservation* 7 (2) (1975) 129–146. doi:10.1016/0006-3207(75)90052-X.
5. R. H. G. Jongman, Nature conservation planning in Europe: developing ecological networks, *Landscape and Urban Planning* 32 (3) (1995) 169–183. doi:10.1016/0169-2046(95)00197-0.
6. P. Beier, R. F. Noss, Do Habitat Corridors Provide Connectivity?, *Conservation Biology* 12 (6) (1998) 1241–1252. doi:10.1111/j.1523-1739.1998.98036.x.
7. A. M. Fleury, R. D. Brown, A framework for the design of wildlife conservation corridors With specific application to southwestern Ontario, *Landscape and Urban Planning* 37 (3–4) (1997) 163–186. doi:10.1016/S0169-2046(97)80002-3.
8. D. K. Rosenberg, B. R. Noon, E. C. Meslow, Biological Corridors: Form, Function, and Efficacy, *BioScience* 47 (10) (1997) 677–687. doi:10.2307/1313208.
9. L. Gilbert-Norton, R. Wilson, J. R. Stevens, K. H. Beard, A meta-analytic review of corridor effectiveness, *Conservation Biology: The Journal of the Society for Conservation Biology* 24 (3) (2010) 660–668. doi:10.1111/j.1523-1739.2010.01450.x.
10. A. D. Mazaris, A. D. Papanikolaou, M. Barbet-Massin, A. S. Kallimanis, F. Jiguet, D. S. Schmeller, J. D. Pantis, Evaluating the Connectivity of a Protected Areas’ Network under the Prism of Global Change: The Efficiency of the European Natura 2000 Network for Four Birds of Prey, *PLoS ONE* 8 (3) (2013) e59640. doi:10.1371/journal.pone.0059640.
11. A. De Montis, S. Caschili, M. Mulas, G. Modica, A. Ganciu, A. Bardi, A. Ledda, L. Dessena, L. Laudari, C. R. Fichera, Urban–rural ecological networks for landscape planning, *Land Use Policy* 50 (2016) 312–327. doi:10.1016/j.landusepol.2015.10.004.
12. E. S. Minor, D. L. Urban, A Graph-Theory Framework for Evaluating Landscape Connectivity and Conservation Planning, *Conservation Biology* 22 (2) (2008) 297–307. doi:10.1111/j.1523-1739.2007.00871.x.
13. Ö. Bodin, J. Norberg, A Network Approach for Analyzing Spatially Structured Populations in Fragmented Landscape, *Landscape Ecology* 22 (1) (2006) 31–44. doi:10.1007/s10980-006-9015-0.
14. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Research* 13 (11) (2003) 2498–2504. doi:10.1101/gr.1239303.