The scientific production of Italian agricultural engineers: a bibliometric network analysis concerning the scientific sector *AGR/10 Rural buildings and agro-forestry territory*

Andrea De Montis,[1,2] Amedeo Ganciu,[3] Fabio Recanatesi,[4] Antonio Ledda,[1,2] Vittorio Serra,[1,2] Mario Barra,[1] Stefano De Montis[1]

[1]*Department of Agricultural Sciences, University of Sassari;* [2]*Department of Civil and Environmental Engineering and Architecture, University of Cagliari;* [3]*Department of Architecture and Design, La Sapienza University, Rome;* [4]*Department of Agricultural and Forestry Sciences, Tuscia University, Viterbo, Italy*

Abstract

According to a worldwide well-known attitude, also in Italy, the assessment of scientific production in the last decades has been progressively based on the analysis of the impact through bibliometric variables. Various data sets, such as Scopus by Elsevier and Web of Science by Thomson Reuters, are designed and maintained to index a steadily increasing range of essays: mostly journal articles, book chapters, and conference proceedings. The indexing relays on the capacity to evaluate and update specific impact measures by keeping track of the citations representing the relations between the essays. The related opportunity to interpret bibliographic systems as research and development (R&D) networks attracted the interest of scientists operating, beyond the field of bibliometric analysis, in the realm of social networking. Network analysis belongs to mechanical statistics and is able to make sense of interconnected systems including very large sets of nodes and links. In this paper, we present a network approach to the review of the scientific production in the time period January, 2003-June, 2016 of Italian agricultural engineers, namely scientists belonging to the Italian ministerial scientific disciplinary sector *AGR/10 - rural buildings and agro-forestry territory*. Starting from 238 articles indexed in the Web of Knowledge database and published by 87 AGR/10 scholars, we apply four network analysis approaches to the study of the citations among articles, the most influential journals and topics, the co-authorship, the most favourite keywords with their evolution in time, and the communities' pattern. We discover that Italian agricultural engineers are interlaced in a sparse network with a still limited tendency toward citing each other and are inclined to team up in established research groups based on a single university. As for the dualism between rural buildings and territory, we document on a relevant expansion of the issues related to landscape analysis and planning and a continuous renewal of studies concerning the relation between rural buildings and biomass and energy management. We advance that Italian agricultural engineers are not confronted anymore with two monolithic macro themes, *i.e.* building design and landscape analysis and planning. Instead, the complexity and interplay between these two domains has dramatically increased in a somehow diverging universe of even more specialised and trans-scale topics.

Introduction

According to a worldwide long-standing practice and to recent legislative acts and institutional policies, also in Italy the assessment of scientific production in the last decades has been progressively based on the analysis of the impact through bibliometric variables. Italian academic scientific production is structured by the Ministry of University and Research (MIUR) in 14 scientific areas grouping sectors that cover the whole panorama of knowledge, from physics and mathematics to social and political sciences (MIUR, 2016). The majority of scientific sectors (SS) are bibliometric, *i.e.* their evaluation is based on the quantitative analysis of the performances recorded in selected research indexing browsers. The National Agency for Research Evaluation (Italian acronym ANVUR) has chosen Scopus by Elsevier (www.scopus.com) and Web of Science by Thomson Reuters (http://ipscience.thomsonreuters.com/product/web-of-science/). These datasets are designed and maintained to index mostly journal articles, book chapters, and conference proceedings. Measures quantify the impact of scientists and articles by keeping track of the citations received by the essays. In this paper, by Italian agricultural engineers we refer to scientists working in the academia and belonging to the SS named *Rural buildings and agro-forestry territory* and heretofore referred to simply with its code *AGR/10*, as it is the tenth sector included in the scientific area *Agricultural and veterinary sciences*. The SS AGR/10 has been recently inserted in a macro sector (MS) named *Agricultural, forest and biosys-*

tems' engineering (code 07/C1), with the other two SSs AGR/08 *Agricultural hydraulics and agro-forestry civil works* and AGR/09 *Agricultural mechanics*. According to the MIUR's description, scientists belonging to the SS AGR/10 focus on two broad themes: i) rural buildings analysis and design for a wide range of destinations including livestock farming, food production, forestry management, and rural housing; and ii) rural territory and landscape analysis and planning with the evaluation of planned transformations, survey and representation of the environment and the landscape, and rural landscape recovery and planning. For further details, we refer the reader to the description released by MIUR (2016). The SS AGR/10 has sometimes changed denomination evolving in the last seventy years from the pioneering studies on livestock farming, rural building design, and agricultural topography to the recent researches concerning rural-urban landscapes, green areas and greenways design, precision livestock farming, and biomass and energy management. In the last twenty years, a major shift affected the official denomination of the entire MS 07/C1 implying an enlargement of the engineering applications' range from mere agricultural to *lato sensu* biological systems. The Society for the Engineering of Agriculture, Food, and Biological Systems (ASABE), proposes a unifying definition of Agricultural and Biological Engineering, which *is the discipline of engineering that applies engineering principles and the fundamental concepts of biology to agricultural and biological systems and tools […]* (Aguado *et al.*, 2011). The concept has been interpreted by Italian colleagues during the ninth Conference of the Association of the Italian Agricultural Engineers: *biological or biosystems engineering can be considered a scientific field, which comprehends the application of engineering principles to processes structuring territorial phenomena to study, model, elaborate, and valorise biological systems for developing optimally agriculture, food production, land use, and the environment* (Santini, 2009). Under these definitions, which have not been adopted worldwide yet, international organisations, such as the Thematic Network on Education & Research in Biosystems Engineering in Europe (ERABEE), have been working for the homogenisation of teaching curricula in the field of biological/biosystem engineering (Comparetti *et al.*, 2009; Aguado *et al.*, 2011).

According to a relatively recent determination, currently the SS AGR/10 (as well as the whole MS 07/C1) is bibliometric, *i.e.* the assessment of the performance of its scientists takes into account the quantitative measures cited above. The system of essays and citations between them can be regarded as a bibliographic network and in general included in the so-called research and development (R&D) networks. In addition, the system of acquaintances between co-authors presents similarities to many social networks (Newman, 2001). The paradigm of complex network analysis (CNA) allows the interpretation of the intertwined sets of nodes and arcs representing bibliographic features (Garfield *et al.*, 1964). As recent studies demonstrate (Jin *et al.*, 2011; Caschili *et al.*, 2014), bibliographic network analysis consists of a semi-automatic tool able to characterise a given research field by detecting the most influential authors, the pattern of collaboration in research groups, the hottest issues, and the polarisation of relevant communities of essays over certain topics. The aim of this paper is the investigation of the scientific production of the SS AGR/10 and indexed in the Web of Science database. We start from the selection of 238 articles published between January, 2003 and June, 2016. Our argument unfolds according to the following Research Questions (RQs). RQ1 attains the investigation of the most influential essays and the patterns of their citations. RQ2 regards the most prominent authors and the nature of their collab-

oration groups. RQ3 refers to the clarification of the hottest research topics and their evolution in time. The contents of this paper are structured as follows. In the next section, we present a state of the art summary on the principles and tools adopted to study bibliographic systems as networks. In the third section, we present the network representations selected in this paper and clarify the meaning and algorithms for quantifying centrality and detecting communities in a network. In section four, we apply network analysis to the characterisation of the scientific output of Italian scholars belonging to the ministerial SS AGR/10. In the fifth section, we discuss the results obtained with reference to the RQs proposed above and, in the sixth and concluding one, we summarise the main messages of this paper and present critical observation on possible limitations of our investigation.

Bibliographic analysis: a state of the art summary

The introduction of automatic systems supporting bibliographic analysis in scientific research is due to Garfield *et al.* (1964). The successive development and availability of greater computing power has allowed the application of these analyses to many other fields with a variety of scopes and levels of detail (Ding, 2011). Caschili *et al.* (2014) referred to scientific essays concerning Strategic Environmental Assessment (SEA), Jin *et al.* (2011) introduced, under the concept of R&D network, more complex studies of the cross-fertilising influences between different research sectors (Newman, 2001), and the economic development connected to scientific discovery due to the so called *absorptive capacity* (Calero-Medina and Noyons, 2008). The success of an essay is usually meant as reputation and influence on the production of ideas in a given scientific field, and is positively correlated with the number of citations it receives (Chen, 2005). Some authors are more influential than others (Eom and Fortunato, 2011), while some colleagues, limitedly cited, exert a transitory influence and are rapidly marginalised from the scientific debate (Price, 1965). The probability of success depends on the achievement of relevant results within the research domain, such as the discovery of the Higgs boson awarded in 2013 the Nobel prize in physics, or external events with extraordinary geographic, social, and economic impact, such as new epidemic events for medicine (Chen, 2005).

Materials and methods

The study of the production in a given SS can be directed to the reconstruction and characterisation of a social system - a particular R&D or, namely, a bibliographic network - where essays are interlaced by the citations among each other. According to Garfield *et al.* (1964) and Jin *et al.* (2011), the analysis of the citations in a bibliographic network and the interpretation of the dominant trajectories of research can be developed by means of the application of graph theory. In the last twenty years, graph theory has been adopted extensively to a variety of fields and types of structures in the domain of complex network analysis (Albert and Barabási, 2002; Barthélemy, 2011). A graph, heretofore referred to as network, is a system including nodes (or vertices) and arcs (or edges). In a bibliographic network, nodes usually stand for essays, keywords, and authors, and arcs for relations whatsoever, such as the presence of citations. A network G can be formalised by an ordered triple including the sets N(G) of nodes, E(G) of edges, and Ψ of attributes attached to the edges. It can be also represented through a mathematical notation, a graphic framework, and a geometric adjacency matrix A(G)-based description. A network is called

topological, if it reports just on the relational characteristics, direct, when the relations show a preference or sense, weighted, if it considers the system of properties of the relations, and spatial, if it takes into account the geographical attributes of its elements, and bipartite, when it divides heterogeneous nodes in subsets of homogeneous nodes (Comincioli, 2004).

Network analysis implies the study of local characteristics, such as the centrality of a given node, or the investigation of global features, such as the detection of properties of emerging communities of nodes and edges. Centrality is a concept of paramount importance in network analysis and describes the topological prominence of a node in terms of number of edges converging to it. According to CNA, centrality is measured by the degree k_i of a certain node i and, in an undirected network, can be obtained as follows (Caldarelli, 2007):

$$k_i = \sum_{j=1,n} a_{ij} \qquad (1)$$

where a_{ij} is an element of the adjacency matrix A(G). In directed graphs, the in-degree (k_i^{in}) and out-degree (k_i^{out}) describe the number of arcs that, respectively, enter and exit from the node i and can be calculated as follows (Caldarelli, 2007):

$$k_i^{in} = \sum_{j=1,n} a_{ji} \qquad (2)$$

$$k_i^{out} = \sum_{j=1,n} a_{ij} \qquad (3)$$

Total degree is a measure of the global prominence of a node and is obtained by the sum of the in-degree and out-degree, according to the equation:

$$k_i^{tot} = k_i^{in} + k_i^{out} \qquad (4)$$

In Figure 1, simple frameworks illustrate the meaning of these centrality measures in the case of the reference articles (nodes in blue) quoted in a scientific essay (in red). The in-degree records how many times an article i quotes past essays j, while the out-degree measures the number of times a given essay j is quoted in the article i. Total degree is a measure that describes in a unique

number the topologic overall centrality of red and blue articles (Appendix Figure 1). With respect to community detection in networks, the pattern of citations leads to various composition of articles. Some articles may be more densely connected each other with respect to other sets of essays. Community detection in networks has been applied to many scientific sectors, such as social sciences (Newman and Girvan, 2004), and bibliometric analysis to investigate the tendency of authors to cluster around certain research topics (Caschili *et al.,* 2014). Community detection tools have been studied in the last years (Fortunato, 2010) and can be classified in three main families: divisive methods (Girvan and Newman, 2002), spectral analysis (Kleinberg, 1999), and optimisation methods (Newman, 2006). In this paper, we draw our method from the last family of tools, where the best community structure corresponds to the maximisation of an objective function indicated as modularity Q and obeying to the following equation:

$$Q = \frac{1}{2m} \sum \left(a_{ij} - \frac{k_i k_j}{2m} \right) * \delta(C_i, C_j) \qquad (5)$$

where m stands for the total number of edges and δ is a parameter equal to 1, if two nodes whatsoever i and j belong to the same community and to zero elsewhere. Modularity ranges from zero to 1 depending on the quality of the separation between communities. According to Fortunato and Barthélemy (2007), modularity is quite sensitive to the maximisation algorithm adopted and the resolution parameter γ. We adopt the Louvain algorithm (Blondel *et al.*, 2008), since it is very fast also for large networks including thousands of nodes and is able to disentangle the complexity of emergent nested communities. A finer analysis of the topics covered in an article network can be performed through a specialisation index adopted by Caschili *et al.* (2014), according to the following equation:

$$S_y^{\varphi} = \frac{ST_y^{\varphi} * CT_y^{\varphi}}{\sum_y \left(ST_y^{\varphi} * CT_y^{\varphi} \right)} \qquad (6)$$

where φ stands for the community, y for the category name used by the Journal of Citation Reports to characterise the journals' focus on particular scientific issues, *ST* for the number of essays labelled with a certain category name y, and *CT* for the number of citations received by those essays. The specialisation index is normalised and expressed in percentage.

With reference to the RQs of our work, we have analysed the bibliographic network of the SS AGR/10 by adopting four types of network representations, as reported in Table 1.

Network representation A is topologic and direct and is ideal

Table 1. Research questions and corresponding network representations developed to investigate the bibliographic network of the SS AGR/10.

RQs	Network representation			
	Code	Elements	Type	Analysis
RQ1, RQ3	A	Articles, citations	Topologic, direct	Centrality and communities
RQ2	B	Co-authors	Weighted, undirect	Centrality
RQ3	C	Keywords	Weighted, undirect	Centrality and communities
RQ3	D	Keywords, year	Weighted, direct, bipartite	Centrality

for characterising the system of essays (nodes) and citations (edges). It implies the analysis of the centrality and the community detection that are useful to satisfy RQ$_1$ and RQ$_2$ concerning the prominence of essays and the role of citations with respect to local and global properties of the network. Network representation B is weighted and undirected and fits the analysis of the centrality of colleagues that have co-authored one or more (depending on the weight) papers. This representation supports arguments that respond to RQ$_2$, with respect to the leadership and the establishment and consolidation of research groups. Network representation C is similar to the previous case and is suitable for the establishment of relations under relevant keywords. It involves the analysis of centrality and of communities and provides results responding to RQ$_3$, as for the understanding of the hottest topics developed by scientists. The last network representation D is weighted, directed and bipartite, as two sets of elements constitute the focus of the analysis and are related through directed arcs featured by a weight meaning the frequency of a given keyword. This representation implies a centrality analysis able to respond to RQ$_3$, by describing the evolution of the most favourite topics in a certain scientific domain.

A bibliographic network analysis of the Italian SS AGR/10

In this section, we apply the bibliographic network analysis to the study of the most recent trends of the SS *AGR/10*. We select 238 articles listed in the Thompson Reuters © Web of Knowledge database and published in the period from January, 2003 to June, 2016 by 87 AGR/10 scholars. In Table 2, we sketch a synthesis of the composition of the colleagues considered.

The group includes 60 still employed and 27 retired scholars at the moment of the writing of this paper. Information has been drawn from the official MIUR website (http://cercauniversita. cineca.it/ php5/docenti/cerca.php). As for the retired colleagues, we have considered the scientific contribution of the single scientist to the development of the SS. In this respect, we have taken into account the affiliation to the Italian Society of Agricultural Engineers (http://www.aiia.it/).

In the Web of Knowledge database, each article is associated to an. ISI file, which conveys in tags all the relevant information including authors, publication year, title, name of the journal, keywords, and references. In Appendix Figure 2, we report on example of .ISI file. We processed the dataset through Sci2Tool (https://sci2.cns.iu.edu), free software powered by Cyberinfrastructure for Network Science Centre at Indiana University. Adopting Sci2Tool, we applied network analysis to map the set of .ISI files into the network representations introduced in Table 1. Then, we performed centrality analysis and community detection through the tool software Gephi (https://gephi.org/). The remainder of this section includes four subsections that illustrate the development and interpretation of the results obtained through the network representations illustrated in Table 1.

The network of articles

We performed network representation A by considering a topological and direct network, where the nodes represent articles and the edges citations between them (Figure 1, top).

The network includes 7925 articles (nodes) cross-referenced through 9257 citations (oriented arcs) with an average density equal to 1.17 citations per article. The analysis of the distribution of the centrality values is illustrated in the bottom of Figure 1. The in-degree, standing for the number of citations received by an arti-

cle, ranges between 1 and 9. One hundred and forty-five (corresponding to 1.83% of the entire set of) articles have never been quoted in ISI journals, while producing 5349 citations (corresponding to 58%). On the other side, 6799 articles (86%) are quoted just one time while they do not cite any articles: they are *passive* items listed in the reference section and included by *active* nodes that quoted at least one other article. According to the analysis of the interplay between in-degree and out-degree distribution, the class of not quoting articles ($k_i^{in}=0$) exerts a fairly high influence on the specific scientific panorama, as each article receives seven citations (Table 3).

The articles develop on a variety of topics including network

Table 2. SS AGR/10 scientists selected in this study.

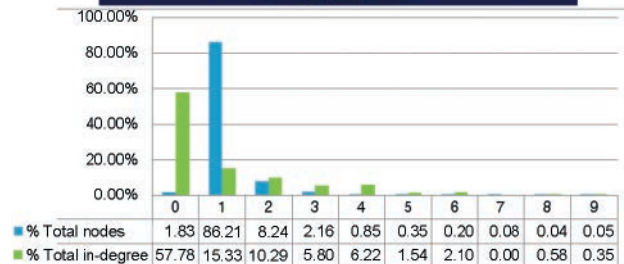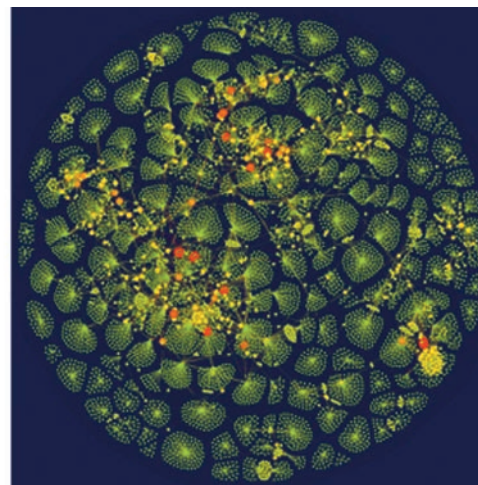| Current status	Role	Number
Employed	Full professors	11
	Full professors (now in other SS)	2
	Associate professors	18
	Assistant professors	29
	Total employed	60
Retired	Full professors	17
	Associate professors	9
	Assistant professors	1
	Total retired	27
Total		87

SS, scientific sectors.

Figure 1. On top: network representation A of the articles: node dimension and colour (from green to red) is proportional to the number of citations received (out-degree). On the bottom: analysis of the total in-degree distribution.

analysis, chemical and thermal properties of building materials, multicriteria evaluation, and the relation between geo-science and computer science. As for the contribution of Italy, we considered articles with at least 5 citations and verified that a fairly large number of them (20 out of 57) have been co-authored by Italian colleagues (Figure 2).

As for the analysis of the specialisation in certain topics introduced in section 3, we directed our analysis to selected (576) arti-

cles quoted at least two times and unambiguously identified by the attribution of the digital object identifier (DOI). The analysis reveals that the scientific production is published in 126 journals classified in 85 category names. In Table 4, we characterise the most favourite journals by AGR/10 colleagues: dominant issues are related to agricultural engineering with a specialisation for environmental and geographical sciences, rural building design, and energy technologies applied to agriculture. The most favourite

Table 3. Six influential papers ($k_i^{out}=7$) in the article network.

First author	Title	Journal/Publisher	Year
Barabási, L.A.	Emergence of scaling in random networks	Science	1999
Hengl, T.	Finding the right pixel size	Computers & Geosciences	2006
Papadakis, G.	Review Paper (SE-Structures and Environment): radiometric and thermal properties of, and testing methods for, greenhouse covering materials	Journal of Agricultural Engineering Research	2000
Saaty, T.L.	The analytic hierarchy process	McGraw-Hill, New York	1980
Scarascia-Mugnozza, G.	Effects of solar radiation on the radiometric properties of biodegradable films for agricultural applications	Biosystems Engineering	2004
Watts, D.J.	Collective dynamics of 'small-world' networks	Nature	1998
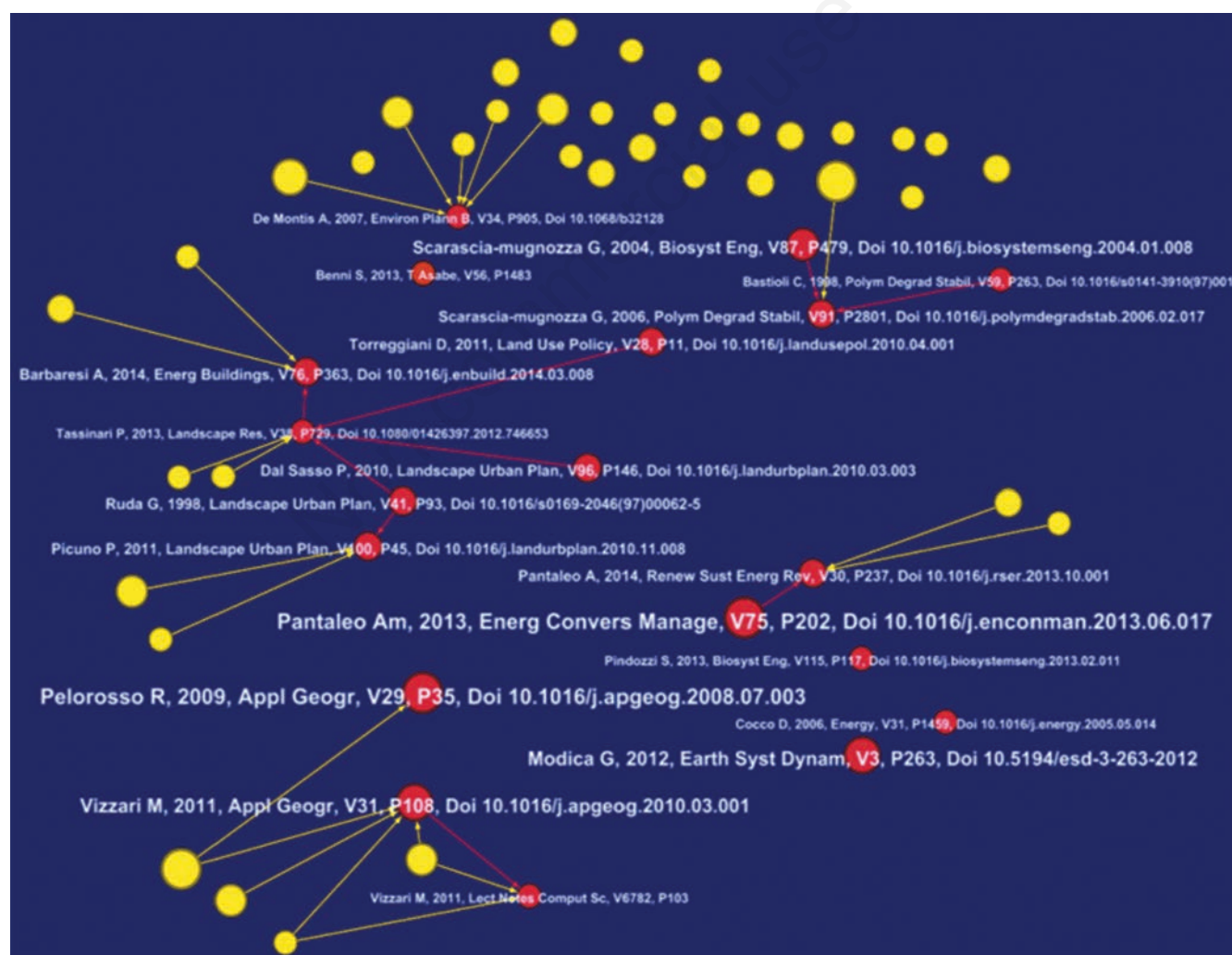
Figure 2. The article network is pruned by deleting the articles quoted less than 5 times: in red the essays co-authored by Italian colleagues.

journals are *Landscape and Urban Planning*, *Biosystem Engineering*, *Landscape Ecology*, and *Energy and Buildings*.

Category names are very complex and have been grouped in 13 macro sectors (for the correspondence between macro sectors and category names, see Table 5).

According to Figure 3, in the article network of the SS AGR/10 the leading specialisations concerns environmental sciences, geography, energy and fuels, and agriculture. The specialisation index can be applied to the study of the patterns of articles belonging to sub-networks, *i.e.* communities, of similar articles and extracted as follows.

The application of the algorithm for the communities' detection implies the selection of a proper value of modularity Q, depending on the resolution parameter γ. In Appendix Figure 3, we analyse the number of communities *versus* Q. We decide to compromise at a resolution parameter γ equal to 40, by choosing a community distribution including 41 clusters at the still reasonable modularity value Q equal to 0.717.

In Table 6 and Figure 4 (on the top), we characterise and illustrate the relevant communities detected in the article network. Communities 25 and 30 are quite large, as they include more than 65% of nodes and edges counted in the entire bibliographic network.

The five communities described in Table 6 absorb all the articles quoted at least two times and labelled with the DOI. We direct a fine inspection of the specialisation of these communities. The remaining 36 communities are quite small, each including on average 30 nodes.

These communities often show only one article that cites all the remaining nodes, which do not quote any node in the same or in other communities. This phenomenon leads to the emergence of many disconnected communities and provide the reason why in the article network it is so difficult to aggregate clusters. In Figure 4 (on the bottom), we analyse the patterns of specialisation of the five most influential communities in the article network. The

largest community 25 includes essays concerning mostly environmental sciences, geography, and energy and fuels. Community 30 comprehends articles treating mainly agriculture, material science and engineering. Community 33 is almost entirely focused on energy and fuels, while community 17 and 31 on agriculture. These last communities host a smaller number of articles and are quite polarised on a unique set of arguments.

The network of co-authors

In this section, we apply network representation B by characterising a weighted and undirected graph, where the nodes stand for authors, edges for the co-authorship, and the weight for the frequency of the collaboration, *i.e.* the number of papers they have
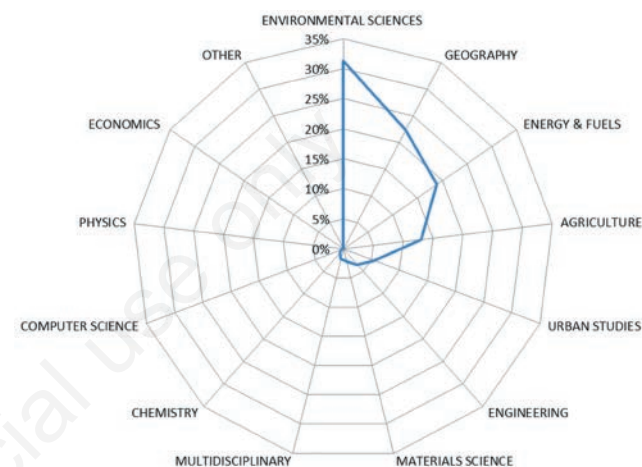
Figure 3. Specialisation pattern of the entire article network, with respect to the macro sectors: the prominence of environmental sciences emerges.

Table 4. Characterisation of the most favourite (with more than 10 articles in the article network) journals by AGR/10 colleagues.

Journal	Number of articles	Out-degree	JCR category names
Landscape and Urban Planning	52	147	Ecology; Geography, Physical
Biosystem Engineering	28	81	Agricultural Engineering; Multidisciplinary
Landscape Ecology	19	44	Ecology; Geography, Physical; Geosciences, Multidisciplinary
Energy and Buildings	18	46	Construction & Building Technology; Energy & Fuels; Engineering, Civil
Remote Sensing of Environment	17	41	Environmental Sciences; Remote Sensing; Imaging Science & Photographic Technology
Applied Energy	15	38	Energy & Fuels; Engineering; Chemical
Applied Animal Behaviour and Science	14	31	Agriculture; Dairy & Animal Science; Behavioral Sciences; Veterinary Sciences
Journal of Agricultural Engineering Research	14	43	Agricultural Engineering; Multidisciplinary
Biomass and Bioenergy	12	28	Agricultural Engineering; Biotechnology & Applied Microbiology; Energy & Fuels
Energy Conversion and Management	11	34	Thermodynamics; Energy & Fuels; Mechanics; Physics, Nuclear
Journal of Polymers and the Environment	11	39	Engineering, Environmental; Polymer Science
Applied Thermal Engineering	10	30	Thermodynamics; Energy & Fuels; Engineering, Mechanical; Mechanics
Computers and Electronics in Agriculture	10	21	Environmental Studies; Geography
Energy	10	31	Thermodynamics; Energy & Fuels
Land Use Policy	10	26	Environmental Studies
Renewable and Sustainable Energy Reviews	10	26	Green & Sustainable Science & Technology; Energy & Fuels

Table 5. Clustering of category names in macro sectors.

Macro sector	JCR category name
Agriculture	Agricultural Engineering
	Agriculture
	Agriculture, Dairy & Animal Science
	Agriculture, Multidisciplinary
	Agronomy
	Behavioral Sciences
	Dairy & Animal Science
	Food Science & Technology
	Forestry
	Horticulture
	Plant Sciences
	Respiratory System
	Veterinary Sciences
Chemistry	Biochemistry & Molecular Biology
	Biotechnology & Applied Microbiology
	Chemical
	Chemistry Organic
	Chemistry, Analytical
	Chemistry, Applied
	Chemistry, Multidisciplinary
	Chemistry, Physical
	Electrochemistry
	Engineering, Chemical
	Geochemistry & Geophysics
	Spectroscopy
Computer Science	Computer Science, Artificial Intelligence
	Computer Science, Information Systems
	Computer Science, Interdisciplinary Applications
	Engineering, Electrical & Electronic
Economics	Economics
Energy & Fuels	Construction & Building Technology
	Energy & Fuels
	Green & Sustainable Science & Technology
	Physics, Nuclear
	Thermodynamics
Engineering	Acoustics
	Engineering
	Engineering, Biomedical
	Engineering, Civil
	Engineering, Environmental
	Engineering, Mechanical
	Engineering, Multidisciplinary
	Mechanics
Environmental Sciences	Biodiversity Conservation
	Ecology
	Environmental Sciences
	Environmental Studies
	Evolutionary Biology
	Genetics & Heredity
	Limnology
	Mathematical & Computational Biology
	Meteorology & Atmospheric
	Mineralogy
	Reproductive Biology
	Soil Science
	Water Resources
Geography	Geography
	Geography, Physical
	Geosciences, Multidisciplinary
	Imaging Science & Photographic Technology
	Remote Sensing

To be continued to the next column.
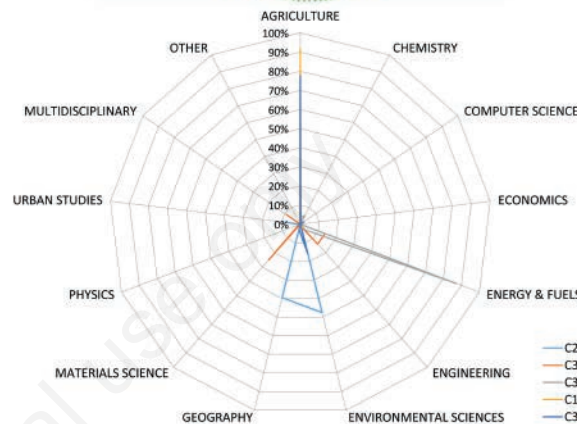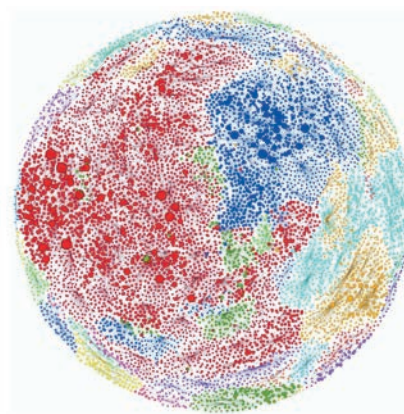
Figure 4. Top: Visual representation of the communities' pattern in the article network: nodes are coloured in red for community 25, in blue for community 30; bottom: analysis of the specialisation in communities 17, 25, 30, 31, and 33.

Table 5. Continued from previous column.

| Macro sector	JCR category name
Materials Science	Materials Science, Characterisation & Testing
	Materials Science, Coatings & Films
	Materials Science, Multidisciplinary
	Polymer Science
Multidisciplinary	Mathematics, Interdisciplinary Applications
	Multidisciplinary
	Multidisciplinary Sciences
	Physics, Multidisciplinary
Physics	Physics, Applied
	Physics, Atomic, Molecular & Chemical
	Physics, Fluids & Plasmas
	Physics, Mathematical
Urban Studies	Planning & Development
	Public, Environmental & Occupational Health
	Transportation
	Transportation Science & Technology
	Urban Studies
Other	Anthropology
	Operations Research & Management Science
	Pharmacology & Pharmacy
	Psychology, Mathematical
	Social Sciences, Mathematical Methods
	Sociology
	Statistics & Probability
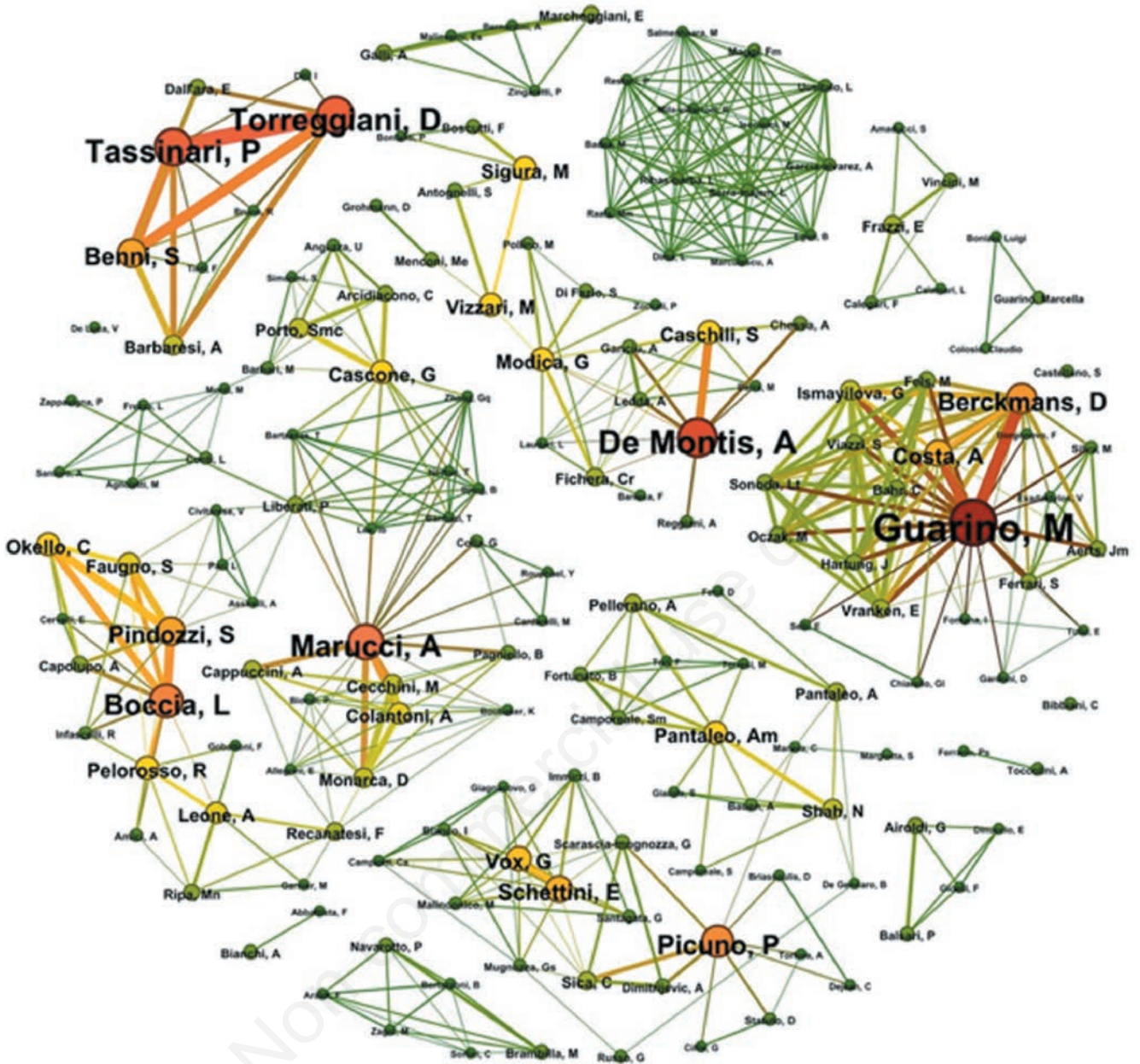
Figure 5. Network of co-authors for the scientific sector AGR/10. Node dimension and colour (green to red) depends on the number of ISI paper co-authored. For ease of reading, authors with less than 4 papers are not represented. Arc thickness and colour (green to red) vary, according to the number of papers written by the authors corresponding to the end nodes.

Table 6. Topological analysis of the communities in the article network.

Community	Number of articles		Total out-degree		Total in-degree	
	Complete	Selected	Complete	Selected	Complete	Selected
25	3652	325	4427	874	4435	1244
30	1474	110	1734	302	1740	466
33	647	71	807	205	803	231
17	598	46	545	103	640	85
31	406	24	453	54	446	67

written together. The network includes 512 authors (nodes), who have collaborated 2868 times in the writing of 1096 essays.

Figure 5 confirms that AGR/10 scientists collaborate mainly on the basis of institutional prevalent relations determined by the affiliation to a given academic institution. The patterns of the network clearly indicate a polarisation on some leading scholars, who co-authored a large number of essays.

The analysis of the distribution of the scientific production (Appendix Figure 4) reveals that the majority of colleagues (458 corresponding to 90% of the entire set of co-authors) have co-authored a maximum of four papers. The same analysis signals that very productive (*i.e.*, with more than ten papers) scholars are quite rare (thirteen people).
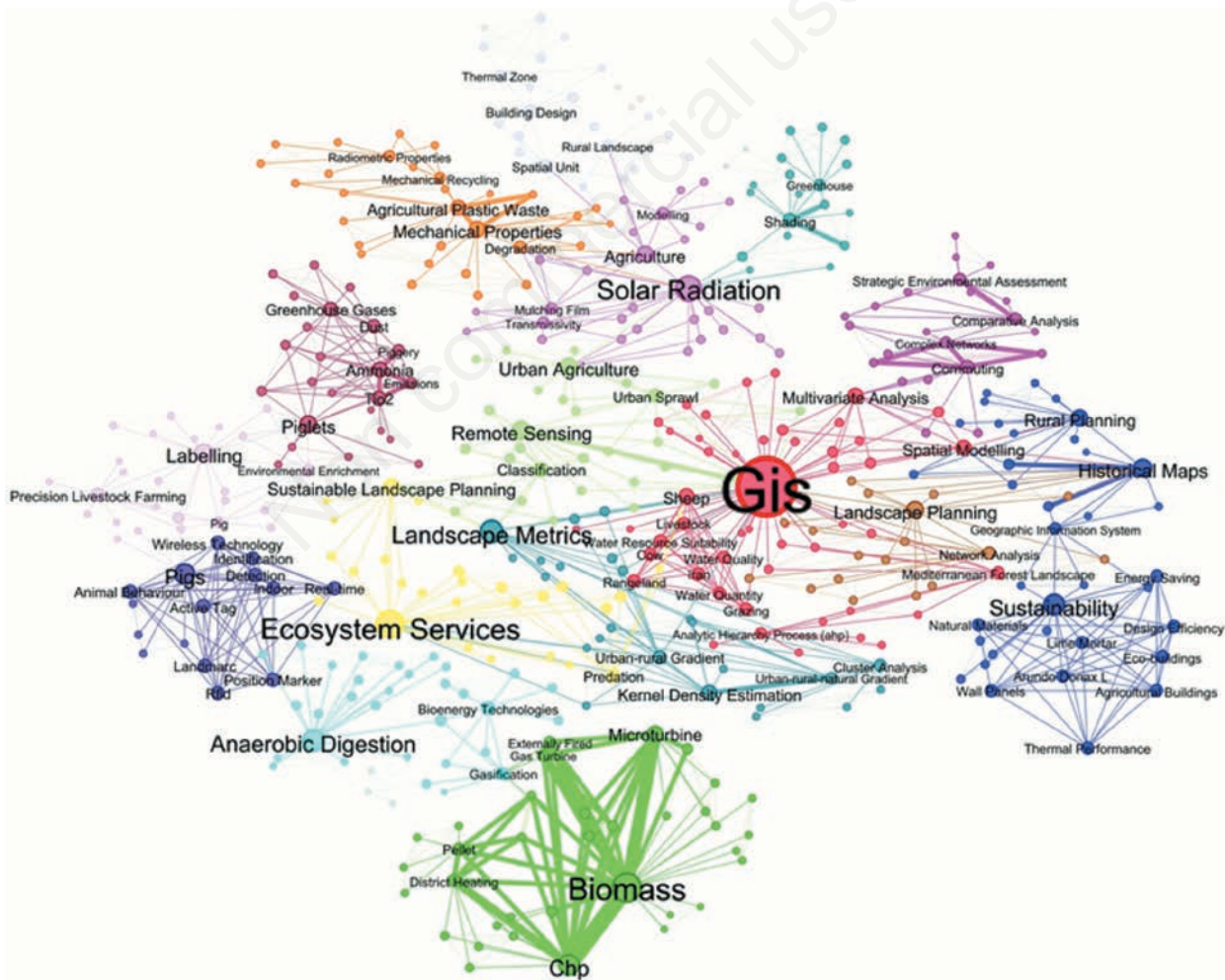
Network of keywords

In this subsection, we develop on the network representation C, which proposes a weighted and undirected graph, where nodes stand for keywords, edges for the co-appearance in the same essay, and the weight for frequency of appearance. The network of keywords includes 841 nodes and 2165 arcs and is represented in Figure 6.

Centrality analysis reveals that the co-authorship network

Table 7. Ranking of the most quoted keyword by frequency of appearance.

Keywords	Frequency
Gis	59
Biomass	27
Ecosystem Services	25
Solar Radiation	23
Landscape Metrics	20
Anaerobic Digestion	19
Sustainability	18
Chip	18
Pigs	17
Italy	17
Historical Maps	15
Remote Sensing	15
Labelling	15
Landscape Planning	14
Piglets	14
Urban Agriculture	14

Figure 6. The network of keywords of the scientific sector AGR/10. Node dimension is proportional to the frequency of each keyword, while arc thickness to the frequency of the co-presence in the same article.

polarises around some hot keywords that characterise the research interests of AGR/10 colleagues. The most prominent concepts confirm the preponderance of issues connected to rural landscape and environment analysis and planning: ecosystem services, sustainability, landscape metrics, and GIS. Other prominent keywords correspond to clusters of issues that are more isolated than the previous cases: biomass, solar radiation, and anaerobic digestion. In Table 7, we report the most frequent keywords that characterise the production in the SS AGR/10: a very prominent role is played by GIS, which appears with a frequency more than two times larger than the one of the second keyword in rank (biomass).

Community detection analysis yields 85 communities, with a resolution parameter γ equal to 1 and a very good modularity value (Q=0.934). In this case, no nested communities were discovered. In Appendix Figure 5, we report on the characterisation of community pattern. The largest communities 10 and 50 comprehend arguments close to the keywords *GIS* and *ecosystem services*. Communities 52 and 17 refer to the keywords *biomass* and *anaerobic digestion*, while communities 68 and 55 to *solar radiation* and *agricultural plastic*.

Network keyword-year

This sub section reports on the application of the last network representation D, by developing on a weighted, directed and bipartite graph, where two sets of nodes correspond to keywords and year of appearance and the weight to the frequency of a given keyword. We have taken into account some of the most favourite keywords selected in the previous section and used in the time period 2006-June, 2016. Figure 7 illustrates the keyword-year network.

The most quoted keyword *GIS* has been used throughout the whole period, with the exception of the years 2007, 2008, and 2010. Other keywords, such as *landscape metrics* and *ecosystem services* have come to the fore only recently. Other traditional keywords, such as *biomass* and *agricultural plastic waste* have been subject to a new interest from 2011 on.

Discussion

In this section, we discuss the results obtained in this paper with respect to the possibility to characterise the scientifics production and collaboration between Italian scientists belonging to the SS AGR/10 *rural buildings and agro-forestry territory* and according to the dataset maintained by Web of Science. We have studied the SS AGR/10 through a package of four network representations, in order to argument with reference to three RQs, as described in Table 1. Network representation A provides us with insights about the most influential articles and research issues (RQ1 and RQ3). The network of articles is quite sparse, as it shows a relatively low density of citations per article (1.17). Six papers are very influential and include one article first-authored by an AGR/10 colleague. These articles develop on many topics: network analysis, chemical and thermal properties, building materials, multicriteria evaluation, geo-science and computer science. The most favourite five journals are *Landscape and Urban Panning*, *Biosystem Engineering*, *Landscape Ecology*, *Energy and Buildings*, and *Remote Sensing of Environment* and reflect a quite balanced interest of AGR/10 colleagues for both: i) rural environmental and landscape analysis planning, and ii) rural buildings analysis and design. This fully reflects what indicated in the ministerial description of the SS AGR/10 (MIUR, 2016). A fine inspection of the specialisation reveals that four macro sectors -

following the ranking, *environmental sciences*, *geography*, *energy and fuels*, and *agriculture* - characterises 85% of the articles cited at least two times and uniquely classified with the DOI. Community detection of the article network indicates the emergence of five relevant communities. The remaining 36 communities include a small number of nodes often connected to a single citing node and thus are disconnected to each other. The largest two communities present a specialisation mostly in *environmental sciences*, *geography*, and *energy and fuels* and in *agriculture*, *material science*, and *engineering*. The other three are much more polarised on a unique macro sector. Network representation B supports arguments that attain the role of AGR/10 colleagues and research groups (RQ2). The analysis of the co-author network clearly reveals that AGR/10 scientists collaborate mainly with colleagues belonging to the home university. On the other hand, a slight attitude to collaborate in inter-academic groups can be detected. Network representation C offers other results responding to RQ3, as for the understanding of the hottest topics developed by scientists. The characterisation of the network of keywords clearly confirms the duopoly territory-buildings, where preponderant broad issues, such as GIS, landscape, and sustainability are accompanied by important concepts, such as biomass, solar radiation, and
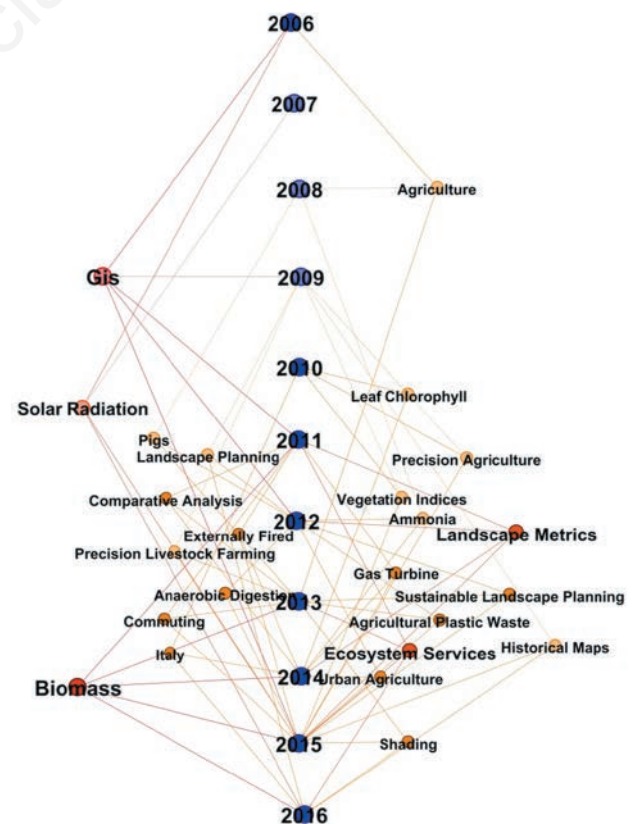

Figure 7. Bipartite network of keywords and year of appearance in article published by AGR/10 colleagues from 2006 to June, 2016.

anaerobic digestion. As for RQ$_3$, network representation D complements these results with a dynamic analysis of the temporal occurrence of the keywords. The most quoted keyword is *GIS*, while other territorial concepts (ecosystem services) have emerged recently. More stable is the presence of building design related keywords: the interest for biomass and agriculture plastic waste has revamped since 2011.

Conclusions

In this section, we address the concluding remarks and argue on some limitations that open future research alleys. In this paper, we have demonstrated how complex network analysis provide researchers with simple yet powerful tools for inspecting the structure of scientific production in a given field and sector of science. We have developed on a semi-automatic analysis of the Italian SS AGR/10 *rural buildings and agro-forestry territory* by mapping to networks and characterising the data base maintained by Web of Science on the articles published from January, 2003 to June, 2016. Different representations deploy many types of networks and two principal analyses, concerning centrality and community detection. They have been key to the clarification of the RQs proposed and concerning the leading scholars, the structure of the research groups, the specialisation in scientific fields, and the dynamics of research interests developed by AGR/10 colleagues. As an overall unifying argument, we found confirming quantitative evidence of the increasing scientific interest for issues related to the territory and a revamping new attention for arguments traditionally related to rural buildings. This ultimately demonstrates a leading role of the ministerial description of the SS AGR/10 (MIUR, 2016), in an always-changing interpretation of the well-known dualism between rural buildings and territory.

Yet, we are aware that our approach presents some limitations and may be integrated as follows. First, we directed our interest to the recent production of AGR/10 scholars recorded only in the Web of Science database. This source is representative of only a part of a wider output, which is indexed in other sites, such as primarily Scopus by Elsevier, an important source for Italian scientists, and Google Scholar citation (https://scholar.google. com/intl/en/scholar/citations.html), a broader system embracing also grey scientific literature. Our analysis is also limited, it does not include essays, which - although not indexed - have influenced decades of research in the SS AGR/10. In this respect, our analysis does not aim at proposing a retrospective summary of the seminal contribution of the pioneers and founding fathers of the SS AGR/10. Another important limitation of this study attains the exclusive interest for the output of Italian - *i.e.* belonging to the SS AGR/10 - agricultural engineers. Agricultural engineering worldwide encompasses a variety of themes correspondingly related to rural buildings and agro-forestry territory. Italian contribution to the field is important and emerges in this paper in some respects. A complete analysis of the main research issues covered by international colleagues would imply a much wider and demanding selection of authors worldwide in both academic and professional contexts.

References

Aguado P., Ayuga F., Briassoulis D., Panagakis P., Febo P., Comparetti A., Scarascia-Mugnozza G., O'Donnel C.,

Navickas K., Fehrmann J. 2011. The transition from agricultural to Biosystems Engineering University Studies in Europe. Proceedings if the IMSCI 2011 - 5th International Multi-Conference on Society, Cybernetics and Informatics, 1:1-6.

Albert R., Barabási A.L. 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74:47-97.

Barabási A.L., Albert R. 1999. Emergence of scaling in random networks. Science 286:509-12.

Barthélemy M. 2011. Spatial networks. Phys. Rep. 499:1-101.

Blondel V.D., Guillaume J.L., Lambiotte R., Lefebvre E. 2008. Fast unfolding of communities in large networks. J. Stat. Mech-Theory E. 10:P10008.

Caldarelli G. 2007. Scale-free networks: complex webs in nature and technology. Oxford University Press, Oxford, UK.

Calero-Medina C., Noyons E. 2008. Combining mapping and citation network analysis for a better understanding of the scientific development: the case of the absorptive capacity field. J. Informetr. 2:272-9.

Caschili S., De Montis A., Ganciu A., Ledda A., Barra M. 2014. The strategic environment assessment bibliographic network: a quantitative literature review analysis. Environ. Impact Asses. Rev. 47:14-28.

Chen C. 2005. Visualising critical trails of scientific knowledge. Presented at the Joint 4S/ISSI Panel on Mapping Science. Annual Meeting of the Society for Social Studies of Science (4S), Pasadena, CA, USA.

Comincioli V. 2004. Modelli matematici. Available from: http://www.multimediacampus.it/download/biomatematica2.pdf

Comparetti A., Febo P., Orlando S., Scarascia Mugnozza G. 2009. Dall'ingegneria agraria all'ingegneria dei biosistemi: didattica e ricerca in Europa e negli USA. Proceedings of the IX National Conference of the Italian Association of Agricultural Engineers, Ischia, Italy.

Ding Y. 2011. Scientific collaboration and endorsement: network analysis of co-authorship and citation networks. J. Informetr. 5:187-203.

Eom Y.-H., Fortunato S. 2011. Characterising and modeling citation dynamics. PLoS One 6:e24926.

Fortunato S. 2010. Community detection in graphs. Phys. Rep. 486:75-174.

Fortunato S., Barthelemy M. 2007. Resolution limit in community detection. Proc. Natl. Acad. Sci. 104:36-41.

Garfield E., Sher I.H., Torpie R.J. 1964. The use of citation data in writing the history of science. Institute for Scientific Information Inc., Philadelphia, PA, USA.

Girvan M., Newman M.E. 2002. Community structure in social and biological networks. Proc. Natl. Acad. Sci. 99:7821-26.

Hengl T. 2006. Finding the right pixel size. Comput. Geosci. 32(9):1283-98.

Jin J.H., Park S.C., Pyon C.U. 2011. Finding research trend of convergence technology based on Korean R&D network. Expert Syst. Appl. 38:15159-71.

Kleinberg J.M. 1999. Hubs, authorities, and communities. ACM Comput. Surv. 31:5.

MIUR, 2016. Settori scientifico disciplinari [scientific sectors]. http://www.miur.it/0002Univer/0021Offert/0092Settor/index_cf2.htm Accessed: May 2016.

Newman M.E.J. 2001. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. 98:404-9.

Newman M.E.J. 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. 103:8577-82.

Newman M.E.J., Girvan M. 2004. Finding and evaluating community structure in networks. Phys. Rev. E 69:026113-27.

Papadakis G., Briassoulis D., Scarascia Mugnozza G., Vox G., Feuilloley P., Stoffers J.A. 2000. Review Paper (SE-Structures and Environment): radiometric and thermal properties of, and testing methods for, greenhouse covering materials. J. Agr. Eng. Res. 77:7-38.

Price D.J.D.S. 1965. Networks of scientific papers. Science 149:510-5.

Saaty T.L. 1980. The analytic hierarchy process. McGraw-Hill, New York, NY, USA.

Santini A. 2009. Presentazione. Proceedings of the IX National Conference of the Italian Association of Agricultural Engineers, Ischia, Italy.

Scarascia-Mugnozza G., Schettini E., Vox G. 2004. Effects of solar radiation on the radiometric properties of biodegradable films for agricultural applications. Biosyst. Eng. 87:479-87.

Watts D.J., Strogatz S.H. 1998. Collective dynamics of 'small-world' networks. Nature 393:440-2.