SOCIAL NETWORK ANALYSIS IN GEOSCIENCES: SCIENTIFIC COLLABORATION BETWEEN PERIGLACIAL SCHOLARS IN THE IBERIAN PENINSULA (2000-2017)

C. GARCÍA-HERNÁNDEZ[1]*, J. RUIZ-FERNÁNDEZ[1], E. SERRANO-CAÑADAS[2]

[1]Department of Geography, University of Oviedo, Oviedo, Spain.

[2]Department of Geography, University of Valladolid, Valladolid, Spain.

ABSTRACT. *This research examines the collaboration relationships between those authors affiliated to Iberian institutions and involved in periglacial studies between 2000 and 2017. The data has been extracted from the Google Scholar database and processed using Social Network Analysis (SNA) techniques. A total of 230 scientific publications have been selected, involving 198 researchers of which 113 belong to Spanish institutions, 38 to Portuguese institutions and 47 to foreign institutions. The work carried out by the Iberian community (assembled through national groups linked to the International Permafrost Association) has considerably improved its possibilities of dissemination and citation in this period: There has been an increase in the number of publications per year, in the number of authors per publication, and in the number of publications with international collaboration. Nowadays there is a group of very productive and well-connected authors, who play an important role as intermediaries, and future prospects are good. However, 50% of the authors have a low capacity for interaction, and there is a low presence of women in the network: These imbalances constitute the main challenges to be faced by the Iberian community in the next decade. Beyond the Iberian community of periglacial scholars, this study provides an example of methodology to be applied in other co-authorship networks in subdisciplines of the Earth Sciences, and its results are discussed in this sense. The results of this research offer valuable information for the management of scientific collaboration programs, the selection of representatives and to weigh the possibilities and needs of the Iberian network. But, especially, it presents a socioscientific framework of reference for early career scientists and makes them aware of the need to integrate themselves into a booming scientific community.*

Análisis de redes sociales en Ciencias de la Tierra: colaboración científica en estudios de periglaciarismo en la Península Ibérica (2000-2017)

RESUMEN. *Esta investigación examina las relaciones de colaboración entre los autores afiliados a instituciones ibéricas e involucrados en estudios de periglaciarismo entre 2000 y 2017. Los datos se extrajeron de la base de datos de Google Scholar y se procesaron utilizando técnicas de análisis de redes sociales (SNA). Se han seleccionado un total de 230 publicaciones científicas, con 198 investigadores, de los cuales 113 pertenecen a instituciones españolas, 38 a instituciones portuguesas y 47 a instituciones extranjeras. El trabajo realizado por la comunidad ibérica (reunido a través de grupos nacionales vinculados a la Asociación Internacional de Permafrost) ha mejorado considerablemente sus posibilidades de difusión y citación en este período: ha aumentado el número de publicaciones por año, en número de autores por publicación, y el número de publicaciones con colaboración internacional. Hoy en día hay un grupo de autores muy productivos y bien conectados, que desempeñan un papel importante como intermediarios, y las perspectivas de futuro son buenas. Sin embargo, el 50% de los autores tienen poca capacidad de interacción y la presencia de mujeres en la red es baja: estos desequilibrios deben constituir los principales desafíos a los que se enfrentará la comunidad ibérica en la próxima década. Por otra parte, este estudio proporciona un ejemplo de metodología para ser aplicada en otras redes de coautoría en subdisciplinas de Ciencias de la Tierra, y sus resultados se discuten en este sentido. Los resultados de esta investigación ofrecen información valiosa para la gestión de programas de colaboración científica, la selección de representantes y para sopesar las posibilidades y necesidades de la red ibérica. Pero, especialmente, presenta un marco sociocientífico de referencia para aquellas personas que inician su carrera investigadora, haciéndolas conscientes de la necesidad de integrarse en una comunidad científica en auge.*

*Corresponding author: Cristina García-Hernández, Department of Geography, University of Oviedo, C/ Amparo Pedregal s/n, 33011 Oviedo, Spain. E-mail address: garciahcristina@uniovi.es

1. Introduction

Social Network Analysis (SNA) is a method that enables us to interpret patterns of behaviour within a network that can be identified through "graph theory". A graph is formed by means of a set of components (nodes) joined by lines (edges); its construction

allows us to identify and analyse the structures that emerge from the relationships established between the nodes. In words of Newman (2001a:1) a social network is "A set of people or groups each of which has connections of some kind to some or all of the others. In the language of social network analysis, the people or groups are called actors and the connections ties. Both actors and ties can be defined in different ways depending on the questions of interest. An actor might be a single person, a team, or a company. A tie might be a friendship between two people, a collaboration or common member between two teams, or a business relationship between companies".

SNA fits into the research tradition of the social sciences since the 1950s, and of interest are both the study of human interaction patterns and the implications of these networks' structures in the way information is distributed (Newman, 2001b). These types of studies have numerous applications, including the analysis of family networks, commerce, correspondence and affiliation, among many others (Wasserman and Faust, 1994). The development of the Internet, with the creation of a virtual space where relations multiply at a speed that is unprecedented in history, has enabled a much greater development of these types of studies in recent years (Scott, 2012).

SNA can also be applied to research on scientific communication. The first studies, focused mainly on the dynamics of citation and co-citation, date back to the 1960s, when the work of Price (1965, 1966) and Crane (1969) pioneered this field. However, studies focused on co-authorship networks have undergone further development, essentially since the 2000s (Newman, 2000, 2001a, 2004; Barabási *et al*., 2002). These are works whose interest lies in their ability to explain how knowledge flows and how information exchanges occur, to identify the existence of groups and determine the degree of interaction and influence of each of the elements that form part of a network. In the Iberian context, SNA is a technique that is increasingly applied in the study of co-authorship networks, although it is much more developed in Spain (e.g. Molina *et al*., 2002; Valderrama-Zurián *et al*., 2007; Ávila-Toscano *et al*., 2014, among others) than in Portugal, for reasons related to the late start of university studies on sociology compared to Spain (Martins, 2011; Varanda *et al*., 2012).

In the Iberian Peninsula, several bibliometric studies have been published regarding the forms of production in geosciences (López-Ontiveros and Molinero, 1995; García-Ruiz, 1999, 2008; Gutiérrez *et al*., 2013; Lasanta and García-Ruiz, 2013), as well as studies reflecting the production related to the permafrost and periglacial studies, both in Spain and in the Iberian Peninsula as a whole (Gómez-Ortiz and González-Martín, 1988; Gómez-Ortiz and Palacios, 1995; Peña-Monné, 1998; Gómez-Ortiz *et al*., 2001; Gómez-Ortiz *et al*., 2003; Gómez-Ortiz and Vieira, 2006; Serrano, 2016). Nevertheless, up to now no SNA have been carried out regarding scientific production related to these disciplines in the Iberian Peninsula and, although international bibliometric studies have been carried out (e.g. Dorn, 2002), there are no SNA in the field of geomorphology to date.

Geomorphology is a field of interdisciplinary study in which different specialists converge and collaborate on a regular basis. Scientific collaboration is a growing trend globally, and it is considered to increase both the quality (Lee and Bozeman, 2005) and

the impact of research: The best connected scholars usually have a better citation-based performance (Abbasi *et al*., 2011), but some of the main benefits of collaboration are for groups and institutions, and not only for individuals.

In this study we examine networks in which the actors are Iberian scientists researching on permafrost and periglacial processes, and the ties between them are their scientific collaborations in the period 2000-2017. These scientists are largely connected to the Spanish and Portuguese groups of the International Permafrost Association (IPA). The aims of this research are to:

i. Determine the degree of interaction of the authors and institutions involved.

ii. Identify the best-connected members of the network: those who collaborate the most, who have the greatest capacity for intermediation and are closer to the other members.

iii. Identify the existence of clusters or groups of authors and institutions.

iv. Assess the way that collaborations in the Iberian network evolve, both between Spain and Portugal and with other countries.

v. Discuss the interest and limitations of Social Network Analysis in other scientific communities involved in geomorphological studies.

2. Methodology

To carry out this study, a list was created containing authors belonging to Spanish and Portuguese institutions that participated in publications in the field of periglacial studies between 2000 and 2017. The data was manually extracted from the Google Scholar database, using keywords related to this field of study, as "periglacial", "permafrost", "active layer", "rock glacier", "ice cave" and "soil thermal regime" (terms in Spanish, Portuguese and English). Only articles in scientific journals, books and book chapters were taken into account. To locate publications in which the subject was related to periglacial studies, despite the absence of keywords in the title, the *Google Scholar Citations* profile of the authors involved was also examined to carry out a second selection of publications. Subsequently, the data was refined by manually eliminating duplicate references and homogenising signatures. The same methodology was applied to the analysis of collaboration between institutions.

The data was processed using Node XL network analysis software, which made it possible to identify network structures by analysing all combinations of pairs of authors in each publication. This software also enabled us to create graphical output for some of the analyses by means of graphs (graphical representations made up of nodes and edges, which enable a system made up of interconnected elements to be represented), produced according to the Fucherteman-Reingold algorithm. In SNA the main interest of the research is not so much in the actors that make up a network as in the connections that are established between them, although knowing certain individual characteristics may be useful in order to understand the general structure of the relationships. To analyse these relationships, we used centrality measures, degree centrality, closeness centrality

and betweenness centrality. Degree centrality indicates the number of co-authorships that are established, taking a certain factor as a reference (in this case, the number of collaborations established by each researcher, institution or country). Closeness centrality takes into account the distance that is established between each node (researcher, institution or country) with the others: for example, maximum closeness centrality would be established between a node that was directly connected to all the other nodes of the network. Betweenness centrality measures the capacity of an element to facilitate connection between others, controlling information flows (this is determined by counting the number of times that a node facilitates the most direct connection between two others). On the other hand, the existence of groups or clusters was identified according to the distance between their members (authors or institutions), placing the threshold at two maximum points of distance. Lastly, the clustering coefficient shows the extent to which the collaborations of each node occur within a given group.

3. Results

A total of 230 scientific publications involving 198 researchers were analysed: 113 belong to Spanish institutions, 38 to Portuguese institutions, and 47 to institutions from other countries (Fig. 1). Half of the publications involved three or fewer authors, and only 17% were signed by more than five authors. Figure 2 shows the evolution of the average number of authors and institutions involved in the production of each publication. From an average of 3.7 authors and 2.1 institutions in 2000, these figures rose to an average of 4.8 authors and 2.9 institutions in 2016 (an increase of over 30% in authors and over 35% in institutions). This figure also shows an upward trend in the number of publications per year, from 9 a year in 2000 to 35 in 2016.
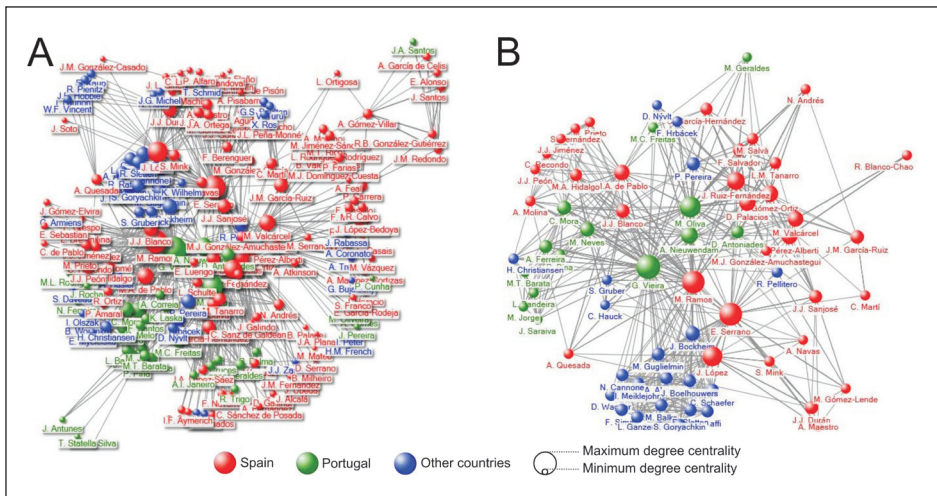

Figure 1. Graph showing the authors and their collaborative relationships between 2000 and 2017, according to the country of origin of their institution: (a) All nodes (authors) visible (b) Authors visible if degree centrality is above 10.
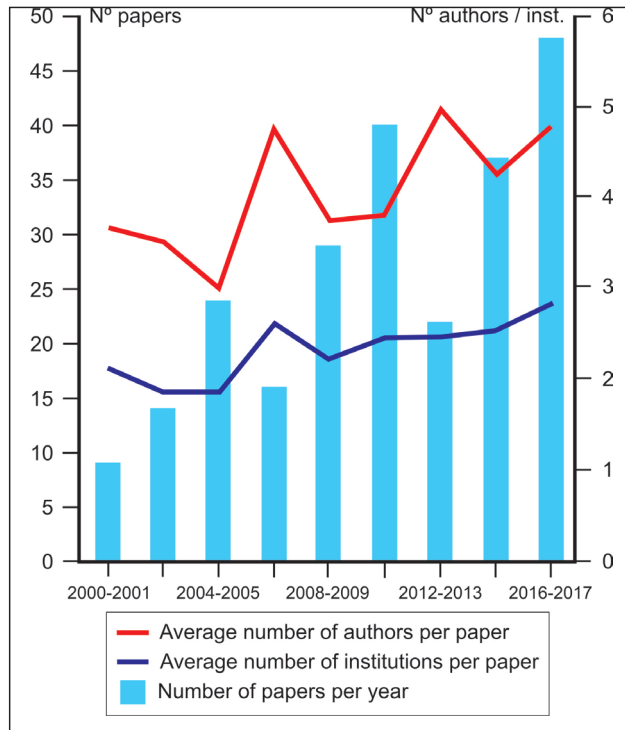
Figure 2. Evolution of the number of publications and the average number of institutions and authors per article.

The most prolific author published 67 articles, although a large number of members of the network (50%) only participated in one. Because of this, although the average is 2.3 articles per author, if we only take into account the researchers who published more than one article the average is 4.72, while among those who worked on more than two articles it rises to 11.85 publications per author.

There was a positive evolution in the collaborations established between authors from Spanish and Portuguese institutions: from an average of 2.3 publications per year over the period 2000-2007 to an average of 14 publications over the period 2008-2016. The same occurred regarding the contributions of authors from the Iberian Peninsula with authors from other countries: since 2012 there was a gradual increase from 4 annual publications to 10 in 2016 (Fig. 3).

Regarding the level of cooperative interaction of the authors and institutions, Tables 1 and 2 show the main general indicators of the Iberian network of periglacial studies. The maximum number of collaborations achieved by one author was 75, well above the average and median for the remaining authors (Table 1). The authors' degree centrality is also shown in Figures 1a and 1b. For the proximity there are only three levels which are represented as minimum, maximum and average/median. Regarding
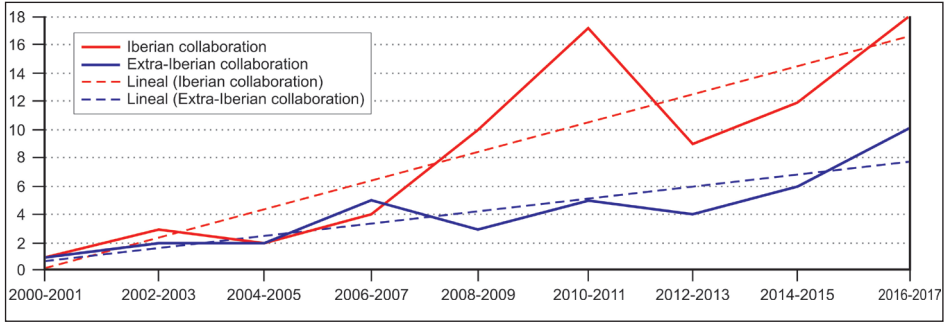
Figure 3. Evolution of the number of collaborations between Spanish and Portuguese authors (Iberian) and authors from other contries (Extra-Iberian).

betweenness centrality (i.e., the number of the most direct paths between nodes, which pass through a certain node), again there are large differences between authors, with maximums far above the average which, in turn, is far from the median (Table 1; Fig. 4a). The same pattern is repeated for institutions (Table 2; Fig. 4b). Tables 3 and 4 show the main authors and institutions that achieved high levels of interaction in this scientific network.

Table 1. Descriptive statistics for authors.

Indicator	Minimum	Maximum	Average	Median
Degree centrality	1.000	75.000	10.899	7.500
Closeness centrality	0.001	0.003	0.002	0.002
Betweenness centrality	0.000	4774.655	161.404	0.000
Clustering coefficient	0.000	1.000	0.822	1.000

Table 2. Descriptive statistics for institutions.

Indicator	Minimum	Maximum	Average	Median
Degree centrality	1.000	42.000	9.706	6.000
Closeness centrality	0.005	0.011	0.007	0.007
Betweenness	0.000	713.255	40.368	0.000
Clustering coefficient	0.000	1.000	0.799	1.000
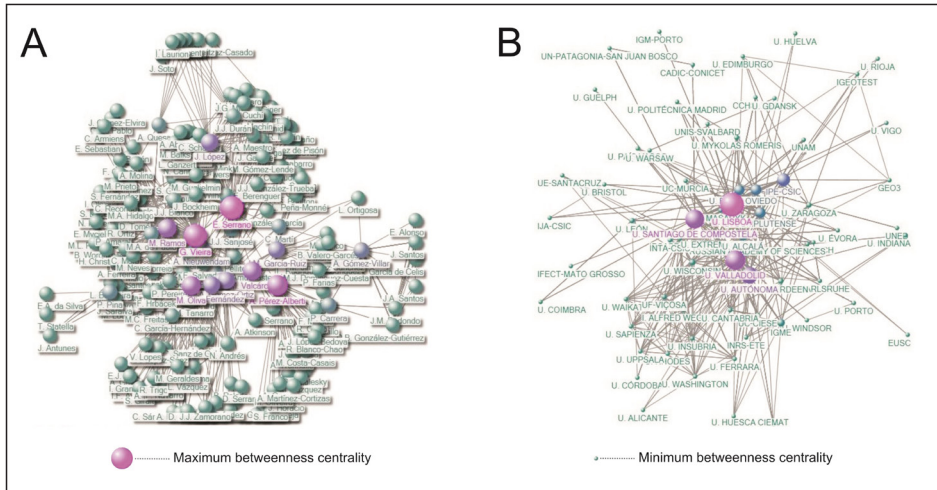
Figure 4. Graph showing betweenness centrality: (a) Author network (b) Institution network.
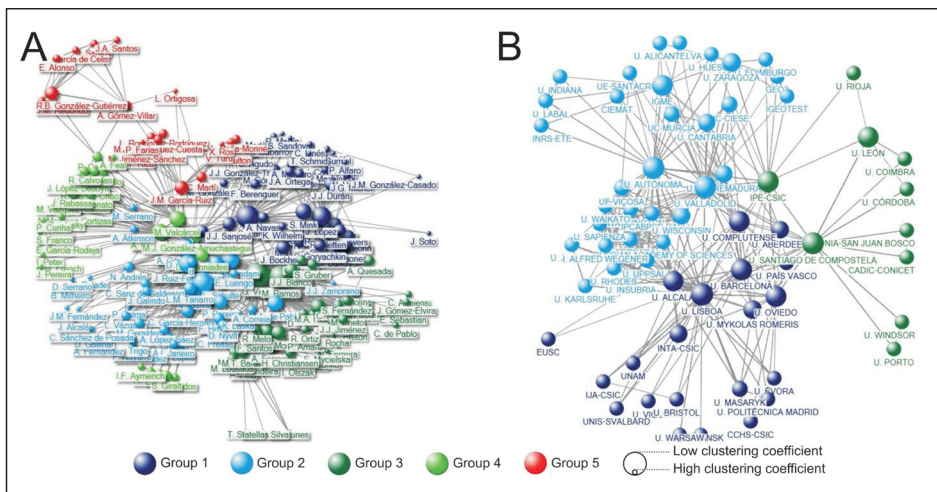

Figure 5. Graphs showing the intensity of collaborations in the groups: (a) Authors and (b) Institutions. Node size becomes larger as the grouping coefficient gets smaller, highlighting the authors or institutions that have a greater capacity to interact with elements of groups different from the one with which they preferentially collaborate.

Table 3. The 20 best-positioned authors according to the structural indicators of the collaboration network.

Author	Degree centrality	Author	Between-ness centrality	Author	Closeness centrality	Author	Cluster-ing coeffi-cient
G. Vieira	75	E. Serrano	4774.655	G. Vieira	0.003	E. Serrano	0.167
E. Serrano	67	G. Vieira	4260.718	E. Serrano	0.003	G. Vieira	0.172
M. Ramos	64	A. Pérez-Alberti	3136.503	M. Oliva	0.003	A. Pérez-Alberti	0.199
M. Oliva	56	M. Oliva	2339.941	J. Ruiz-Fernández	0.003	M. Oliva	0.201
J. López	50	M. Valcárcel	2115.525	A. Nieuwendam	0.003	M. Ramos	0.201
J. Ruiz-Fernández	39	M. Ramos	1942.058	A. Gómez-Ortiz	0.003	J. López	0.228
A. Nieuwendam	37	A. Gómez-Ortiz	1419.985	D. Palacios	0.003	J. Ruiz-Fernández	0.270
A. Gómez-Ortiz	35	J. López	1405.146	M. Ramos	0.003	M. Valcárcel	0.274
D. Palacios	34	J. Ruiz-Fernández	1404.892	A. Pérez-Alberti	0.003	A. Gómez-Ortiz	0.274
A. Pérez-Alberti	33	D. Palacios	1198.985	M. Valcárcel	0.003	J.M. García-Ruiz	0.277
M. Valcárcel	30	J.M. García-Ruiz	1172.765	J. López	0.002	D. Palacios	0.301
M.A. de Pablo	30	A. Nieuwendam	848.014	J. Bockheim	0.002	A. Nieuwendam	0.308
J. Bockheim	25	A. Gómez-Villar	768.264	J.J. Blanco	0.002	M.A. de Pablo	0.338
C. Mora	25	P. Carrera	471.352	J.M. García-Ruiz	0.002	J.J. Sanjosé	0.376
M. Neves	25	P. Pina	465.957	P. Pereira	0.002	J.J. Durán	0.386
J.M. García-Ruiz	23	C. Martí	402.769	M.A. de Pablo	0.002	C. García-Hernández	0.399
J.J. Blanco	22	D. Antoniades	378.789	J.J. Sanjosé	0.002	A. Gómez-Villar	0.400
J.J. Sanjosé	21	M.A. de Pablo	352.468	C. Mora	0.002	A. Ferreira	0.410
J.J. Durán	21	J.J. Sanjosé	284.836	M. Neves	0.002	P. Pereira	0.419
P. Pereira	21	A. Quesada	239.715	A. Ferreira	0.002	M. Neves	0.460

Table 4. The 20 best-positioned institutions according to the structural indicators of the collaboration network.

Institution	Degree centrality	Institution	Betweenness centrality	Institution	Closeness centrality	Institution	Clustering coefficient
U. Lisboa	42	U. Lisboa	713.255	U. Lisboa	0.011	U. Lisboa	0.224
U. Valladolid	40	U. Valladolid	468.476	U. Valladolid	0.010	U. Santiago de Compostela	0.258
U. Autónoma	36	U. Santiago de Compostela	447.892	U. Autónoma	0.009	U. Valladolid	0.277

Institution	Degree centrality	Institution	Betweenness centrality	Institution	Closeness centrality	Institution	Clustering coefficient
U. Alcalá	27	U. Autónoma	321.031	U. Complutense	0.009	U. Autónoma	0.292
U. Complutense	21	IPE-CSIC	201.871	IPE-CSIC	0.009	IPE-CSIC	0.346
U. Wisconsin	20	U. Oviedo	132.424	U. Alcalá	0.008	U. Barcelona	0.346
U. Barcelona	20	U. Complutense	121.842	U. Barcelona	0.008	U. Oviedo	0.383
IPE-CSIC	19	U. Barcelona	113.148	U. Santiago de Compostela	0.008	U. León	0.400
UF-Viçosa	18	U. Alcalá	86.843	U. Oviedo	0.008	IGME	0.407
U. Santiago de Compostela	18	IGME	51.624	U. Wisconsin	0.008	U. Com-plutense	0.409
U. Oviedo	18	U. Wisconsin	20.336	UF-Viçosa	0.008	U. Alcalá	0.457
IP-CABPSS-RAS	17	U. León	13.186	U. Aberdeen	0.008	U. Zaragoza	0.611
I. Alfred We-gener	17	U. Aberdeen	10.985	IP-CABPSS-RAS	0.008	U. Mykolas Romeris	0.639
U. Rhodes	15	U. Mykolas Romeris	9.253	I. Alfred Wegener	0.008	INTA-CSIC	0.700
U. Uppsala	15	U. Zaragoza	8.892	U. Rhodes	0.008	U. Wisconsin	0.745
U. Waikato	15	U. Extremadura	8.212	U. Uppsala	0.008	U. Aberdeen	0.750
U. Insubria	15	UF-Viçosa	5.754	U. Waikato	0.008	U. Ex-tremadura	0.750
U. Ferrara	15	INTA-CSIC	4.400	U. Insubria	0.008	U. Cantabria	0.750
U. Sapienza	15	U. Cantabria	3.593	U. Ferrara	0.008	UC-Ciese	0.867
U. Washington	15	UC-Murcia	0.993	U. Sapienza	0.008	UF-Viçosa	0.900

There are a total of 2158 collaborative relationships (in pairs), with an average of 10.8 collaborations per author. The maximum distance between two people in the network is six points and the average distance is 2.62. By applying a maximum threshold of two separation points, we identified a total of five clusters or groups of authors (Fig. 5a) with 38.6 authors on average. The largest group is made up of 48 authors (1.7 points of separation, 60% belonging to Spanish institutions, and 40% from other countries); the second has 46 members (1.9 points of separation, 67% Spain, 24% Portugal and 9% other countries); the third has 45 (1.9 points, 38% Spain, 44% Portugal and 18% other countries), the fourth has 32 (2 points, 63% Spain, 13% Portugal and 25% other countries) and the fifth has 22 (1.9 points, 82% Spain, 5% Portugal and 14% other countries). A total of five authors did not meet the criteria for inclusion in any of the groups. In the largest group there were 384 collaborations (8 per author), in the second 523 (11 per author), 420 in the third group (9 per author), 154 in the fourth (4.8 per author) and 141 in the smallest group (6.4 per author).

The average cohesion index of the clusters is high (0.22), with the group consisting of 32 authors having the highest cohesion index (0.18) and the group consisting of 22 authors having the lowest cohesion index (0.31). This same operation was carried out

with the network of institutions (68 in total), where we found three groups (Fig. 5b) with a maximum separation threshold of 3 points, an average distance of 1.7 and an average cohesion index of 0.26. In total, 660 collaborative relationships between institutions were found. The largest group is made up of 33 institutions (242 collaborations within the cluster), the second group comprises 24 other institutions (241 collaborations) and the third 11 institutions (25).

4. Discussion

4.1. *Factors explaining the increase in the number of publications and collaborations since 2000*

4.1.1. *A trend that fits in with the epistemological culture of periglacial geomorphology*

A rising trend is observed in the number of annual publications, showing the state of maturity of this field of research in the Iberian Peninsula. The number of authors per article has also increased significantly, a phenomenon that in recent years has affected production in most scientific areas (Bordons and Gómez, 2000; Larivière *et al*., 2016). In the case that concerns us, this could be due to the increasing complexity of the discipline, which requires interdisciplinary collaboration and in-depth knowledge of the abundant scientific production that is generated. However, recent studies maintain that the collaborative trends of the different disciplines reflect to a large extent their own epistemological culture (Wagner *et al*., 2017). Each scientific field represents a different mode of organization in relation to cooperation, and the predominant motivations may be different. Beyond the exchange of knowledge and ideas, many other reasons can promote scientific collaborations; for example learn new methods and become familiar with new technologies (Haythornthwaite, 2006). Although there have not been specific studies on the field of geomorphology, in other sub-fields of geosciences such as soil science, seismology and geophysics, the need to share resources (equipment, samples) explains to a large extent the increase in the number of collaborations (Wagner and Leydesdorf, 2003; Wagner, 2005). In fact, the peculiarities of periglacial studies, which usually require carrying out field work in remote areas and applying techniques and instruments that are progressively more complex and expensive, may also make it interesting to have collaboration between professionals, or even between different research groups with access to samples or direct experience in certain geographical areas.

4.1.2. *The importance of the creation of the Iberian sections of the IPA*

In 1994, the IPA-Spain group was created and integrated as an associate member of the IPA, thus initiating a process of collaboration between institutions and researchers, which continues to the present day. Its beginning should be placed at the "Permafrost and Climate in Europe" (PACE) European Project, and the promotion of IPA-Spain by a nucleus of active emerging groups (University of Barcelona, Complutense University of Madrid, University of Santiago de Compostela and University of Zaragoza) that organised studies on cold regions and permafrost in Spain. The incorporation of the Portuguese community in the IPA (IPA-Portugal) in 2005 resulted in joint meetings

(Iberian Meetings of the IPA) since 2007, just when Iberian collaboration increased (Fig. 3). So far, eight biannual Iberian meetings were held (Table 5). These meetings have encouraged the development of research secondments and joint projects, so that relationships between researchers in the field of permafrost and periglacial studies have grown and consolidated. This, in turn, benefits the group of researchers and those who are beginning their research career.

Table 5. Activities of the IPA regarding the establishment of scientific networks.

Meeting	Year	City	Hosting group	Iberian Attendees
I IPA-Spain	1994	Madrid	UCM*	~15
II IPA-Spain	1995	Santiago Compostela	USC	~20
III IPA-Spain	1997	Andorra	UB	45
IV IPA-Spain	1999	Albarracín	Unizar*	45
V IPA-Spain	2001	Potes	UVa*	31
I EUCOP	2001	Roma	IPA-USR	5
VI IPA-Spain	2003	San Ildefonso-La Granja	UCM-UCLM	30
II EUCOP	2005	Potsdam	IPA-AWI	4
I IPA-Iberian	2007	Guarda	UL*	40
II IPA-Iberian	2009	Sigüenza	UAH*	45
III EUCOP	2010	Svalbard	IPA-UNIS	8
III IPA-Iberian	2011	Ancares	USC*	35
IV IPA-Iberian	2013	Nuria	UB	35
IV EUCOP	2014	Evora	IPA-UL-UE	17
V IPA-Iberian	2015	Valladolid	UVa*	38
VI IPA-Iberian	2017	Mieres	UNIOVI*	55

*Attendance of members of the IPA Executive Committee.

4.2. Factors promoting international cooperation

4.2.1. General distribution of the network according to nationality

The international character of the network in general, and of most of the clusters, is remarkable. Only two of the five clusters (the smallest ones) show a low level of international cooperation. And, both the first and the fourth groups include authors from non-Iberian institutions. The fourth group is the one that best reflects the general distribution of the network, with a significant number of authors from Spanish institutions

and a similar proportion of authors connected to Portuguese and foreign institutions. The high number of scientists from Spanish institutions is justified by the many academic institutions in this country. In Portugal, according to data from the *Agência de Avaliação e Acreditação do Ensino Superior*, there were 41 university centres in 2016, while in Spain there were 83, according to data from the Spanish Ministry of Education, Culture and Sports.

Despite this difference, the Portuguese authors' capacity for interaction and production is noteworthy. They are mostly concentrated in the University of Lisbon, the most productive and best-connected institution in the entire Iberian network. Bibliometric studies related to other disciplines indicate a high interactive capacity at an international level of the Portuguese scientific community when compared to its Spanish counterpart (Filippo and Levin, 2015). In Spain there is a more equitable distribution both of productivity and of interactions between institutions. The universities of Valladolid, Alcalá, Barcelona, the Autonomous University of Madrid, Complutense and Santiago de Compostela, are in the top positions along with the Instituto Pirenaico de Ecología (IPE), a research centre belonging to the Spanish National Research Council (CSIC).

4.2.2. *Internal and external factors promoting internationalisation of research*

Internationalisation of research (considered as collaboration with foreign research groups) has been increasing globally but, in Spain, the increase in scientific collaboration has been especially remarkable since the 1990`s (Wagner and Leydesdorff, 2005). International scientific collaboration is particularly advantageous in some fields of science: both in Portugal and especially in Spain, the number of citations in the articles of Earth Sciences is higher when they are internationally co-authored (Glänzel *et al*., 1999). In the case of Iberian periglacial scholars this process was one of the objectives of the PACE Project, in which the scholars promoting IPA-Spain collaborated. The previous existence of an internal structure, made up of national groups linked to the IPA, stimulated the progressive increase in Iberian attendees to European and international meetings (EUCOP and ICOP), especially since 2008 (Table 5), and culminated with the celebration of the EUCOP IV in Évora (Portugal). This favoured the initiation of new contacts and inclusion in international networks (PAG, APPEWG-IPA, ANTPAS, PLPCWG-IPA, PYRN) that allowed previously established networks to be consolidated and expanded in the second decade of the 21st century. In these networks, emerging researchers from previous phases who are now established and active are taking a leading role. This allows us to be optimistic about the continuity of Iberian research on cold regions and permafrost.

In addition, this internal structure has enabled to take advantage of the international context and conditions, such as the International Polar Year (IPY) or the promotion of international networks for the control of parameters related to Climate Change, such as the Global Terrestrial Network for Permafrost (GTN-P) and the Circumpolar Active Layer Monitoring System (CALM), which have favoured the development of personal and institutional relationships. As an example, a turning point for the Iberian collaborations occurred in 2007, when the first IPA-Iberian meeting was organised and the first IPY was held.

4.3. Strengths and imbalances of the Iberian network

4.3.1. Concentration of interaction capacity and under-representation of women

Among the centrality measures employed, degree and betweenness centrality are shown to be the most significant, with considerable differences among the authors. Specifically, the fact that in both indicators the average and the median are so far apart, and the maximums soar in relation to the average, indicates that a significant part of the actors in the network have a low capacity for interaction and intermediation, which are both concentrated in a small group of authors and institutions detailed in Tables 3 and 4. Given the interaction capacity of these actors, it is not surprising that they coincide with researchers who have a lower clustering coefficient, demonstrating their ability to interact beyond the limits of their usual collaboration group. The same phenomenon also occurs between institutions. Actually, the fact that some collaboration are timely might not be so bad, since weak ties, combined with strong ones, are also beneficial in a network: while the first ones ensure a fluid communication and the possibility of sharing data and resources, the second guarantees the inclusion of new perspectives and experiences (Wagner, 2005).

On the other hand, the under-representation of women is notable, as they correspond to only 15% of the researchers who make up the Iberian network. Moreover, only 4 women appear among the 20 best-positioned authors according to any of the structural indicators (there are no women among the top 10), and only one woman (a Portuguese researcher), appears among the 20 best-positioned in terms of the number of collaborations established (Table 1).

In this sense, the Iberian network is no exception: several studies show woman's underrepresentation in geoscience academia, with 40% of Bachelor's Degrees awarded to females in 2010 while women still make up at around 20% of Tenured and Tenure-Track Faculty (Ceci *et al.*, 2014). In Spain, the difference between male and female geoscientists is more noticeable in each step that involves their research career (Alonso-Zarza *et al.*, 2008, Pereira and Díaz, 2016). In fact, women represent 50% of the people that occupy the positions of Assistant Professor in Spanish University, but only 20% of those of Full Professor (Ministry of Education, Culture and Sport, 2016). There are many ways the heterogeneity of a network can benefit those who are part of it, particularly exchanges between people with different skills, information and experiences, thus generating more creative environments and fostering the resolution of problems (Reagans and Zuckerman, 2001). Therefore, a scientific network needs a balanced presence of men and women to produce better science (Holmes *et al.*, 2008). However, in the absence of new analyses that could be applied to collaboration networks on periglacial studies at European and international levels, we can assert that the results of the Iberian network for the period 2000-2017 do not show a very promising panorama in this sense.

Several studies show that women are attracted to those projects that seek to improve society (Paderewsky-Rodríguez *et al.*, 2017). Thus, those actions conducted to show the main goals of periglacial science as "socially significant" could make those

studies more interesting for women, and some seminars or events could be conducted in this sense. Other initiatives, as enhancing the visibility of women in leading positions (Glass and Cook, 2016), have been identified as essential in addressing the underrepresentation of women in science, while mentoring programs appear to be quite successful in the short term (Kosoko-Lasaki *et al*., 2006). Finally, new studies highlighting the main barriers that women have to face in the development of their academic/scientific carrier in Earth and Planetary Sciences, should be conducted in order to help the whole scientific community (men and women) to confront these unbalances together.

4.4. A small-world network with strong links between its best-connected members

The existence of few clusters (five clusters of authors and three of institutions) indicates that there is a high level of cooperation and cohesion between groups. Moreover, the clusters have a high number of members (38.6 authors and 22.5 institutions per cluster). In this scientific network, 100% of the authors are connected, either directly or through intermediaries. Authors such as Newman (2004) suggest that in a scientific collaboration network there is a "giant component" that includes a proportion of, approximately, 90% of the authors. So, in this case, we can consider that we are dealing with an extremely well-connected network. This idea is reinforced by the fact that the average distance between authors is 2.62 points, well below the average distance observed in large networks, which is six degrees of separation (Newman, 2001). This scientific collaboration network meets most of the requirements of networks known as "small world" (Newman, 2000), which implies minimal separation between the elements that make it up, as well as a high clustering coefficient (Watts and Strogatz, 1998).

There is, therefore, a group of especially well-connected authors who play an important role as intermediaries, ensuring the fastest connection among other members and constituting key factors in avoiding the existence of so-called "structural holes" in the network (Burt, 1995). Authors such as Goh *et al*. (2003), argue that it is unusual for the best-connected members of a network to collaborate with each other. However, in the Iberian network dedicated to the study of permafrost and periglacial studies, these scientists have actively collaborated during the last years (Fig. 1), showing the existence of strong ties between them, which improves scientific productivity and increases the probability of citation (Katz and Martin, 1997; Bordons and Gómez, 2000; Larivière *et al*., 2016). Since co-authorship networks are determined by dynamic processes that contribute to their emergence and development, they are prototypes of a complex network in continuous evolution. Thus, scientific networks usually go through a starting point, a stage of development in which active groups arise driven by highly productive members, a stage of maturity and, finally, a decline (Price, 1966; Barabási *et al*., 2002). Everything indicates that the scientific network we are dealing with, which represents the Iberian community related to permafrost and periglacial studies, has definitely gone through the first stage and is immersed in the second.

4.5. SNA in geosciences: What for?

4.5.1. The interest of these kind of studies

As a precedent, this research has allowed to verify how the organization of scientific events (e.g. ICOP, EUCOP, IPA-Iberian meetings), the development of policies and specific programs for the promotion of certain scientific sub-fields (e.g. PACE, IPY) and the creation of institutions supporting them (e.g. IPA, IPA-Spain, IPA-Portugal) can promote scientific collaboration, especially at the international level. Moreover, we observed the existing feedback between collaboration and productivity and detected imbalances (e.g. few women) in the configuration of the network. Finally, we examined the way in which certain external and internal factors (e.g. of political, institutional and sociocultural nature) can influence its configuration. The information provided here can be used by current members of the scientific network to make decisions, enabling them to assess new possibilities for interaction, or the need to involve new members. But this information can also be useful to avoid the isolation of those who initiate research in this field, since it offers a socioscientific frame of reference that facilitates their orientation and makes them aware of the need to be part of a scientific community in full development (e.g. identifying the most active groups and scholars). Lastly, this information could also be of help to those responsible for the design and financing of interdisciplinary or inter-institutional collaboration programs, or for the selection of representatives in scientific committees. Similar studies in other sub-disciplines of earth sciences will make it possible to establish comparisons with larger or smaller scientific communities, with their own social characteristics and epistemological traditions.

4.5.2. Limitations of this study and possible alternatives

The most significant limitations of this study have to do with the data source, because it is not exhaustive. Therefore, it is likely that some publications, especially those considered less relevant and which were not included in Google Scholar profiles, were not recorded. However, using sources that do not take into account non-indexed publications (e.g. Scopus repository) would leave out an important part of the Iberian scientific production of this period. It is also possible that the identification of publications from keywords resulted in some gaps, since the scanning was done manually. In this sense, data mining techniques, which have been successfully applied in other scientific contexts (e.g. Bose and Mahapatra, 2001; Ngai *et al.*, 2009), could be applied here: Specifically, it seems advisable to use techniques such as web scraping (Vargiu and Urru, 2012) and text mining (Grupta and Lehal, 2009) to extract and detect information automatically, thus allowing for broader and more exhaustive scanning.

On the other hand, the quality of the images is not a point that we can celebrate about the use of graph layout in NodeXL Basic, a free and open-source software package for Microsoft Excel. Despite the advantage of the free access, the export resolution of the images (96 dpi), was quite poor. Moreover, NodeXL graphs do not include a legend. Thus, we used Corel Draw X7 in order to add a legend to each image and, finally, we resized and resampled the images with another specific software (Irfan View), improving

their final resolution. Some of these problems could be solved using the fully featured version of NodeXL, which offers advanced features for graphs and visualization.

5. Conclusions

This study enabled us to identify authors, institutions and groups of researchers involved in developing the Iberian Peninsula's scientific network that studies permafrost and periglacial environments. The evolution of collaborations in the period 2000-2017 was positive, resulting in a considerable increase both in the annual production of publications and in the number of authors per publication. The authors are currently more diverse in terms of institutions and countries of origin, although the number of researchers is higher in Spain (113), where there are a larger number of research centres and where the creation of the national IPA research group occurred 11 years earlier than in Portugal. Even so, the Portuguese authors (38) show a high capacity for interaction, and the University of Lisbon is the most productive and best-connected institution in the entire Iberian network.

The remaining authors in the network (47) belong to non-Iberian institutions. We observed a very positive evolution in the number of international collaborations since 2000, including not only co-authorships between Spain and Portugal (which rose from 2.3 to 14 publications per year) but also between Iberian researchers and those from countries like Argentina, Italy, the United Kingdom, Switzerland, Bulgaria and the Czech Republic.

The results of this analysis indicate that the work carried out by the Iberian community regarding permafrost and periglacial studies has considerably improved its potential for dissemination and citation compared to the year 2000. The creation of the national groups associated with the IPA has made it possible to hold meetings and training activities, and to take advantage of various international circumstances. All the indicators that were assessed indicate that the Iberian network has passed the starting stage and is in full development. A group of authors, some of whom emerged in the 2000s and now well established, form a stable and active core which is necessary in order to incorporate new members and allow the network to go on evolving. Given the strength of interpersonal and inter-institutional ties, as well as the international connections we observed, future prospects are good.

However, certain imbalances in the development of the network seem evident: 50% of the authors have a low capacity for interaction and intermediation, which is concentrated in a small group of institutions. On the other hand, women represent only 15% of the researchers who make up the Iberian network. These imbalances should be the main challenges to be addressed by the Iberian community in the next decade.

The interest of these studies justifies carrying out future research related to the networks of authors and institutions involved in other Earth Sciences subdisciplines, both at regional and global level. However, especially if more extensive searches are intended (i.e. continental or global research, or research concerning several subdisciplines), it would be advisable to automate the process of searching and extracting data by using web scraping and text mining techniques. These improvements would enable a broader and more detailed analysis of the evolution of the scientific network's structure.

Acknowledgements

References

Abbasi, A., Altmann, J., Hossain, L. 2011. Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics* 5 (4), 594-607. https://doi.org/10.1016/j.joi.2011.05.007.

Alonso-Zarza, A.M., Álvarez-Marrón, J., Calonge, A., Díaz, C., Díez-Balda, M.A., Gil-Peña, I., Gómez, M.I. 2008. Mujeres y geología en España. In: *Mesa redonda*. *VII Congreso Geológico de España*. Las Palmas de Gran Canaria. *Geotemas*, 589.

Ávila-Toscano, J.H., Marenco-Escuderos, A., Madariaga Orozco, C. 2014. Indicadores bibliométricos, redes de coautorías y colaboración institucional en revistas colombianas de psicología. *Avances en Psicología Latinoamericana* 32 (1), 167-182. https://doi.org/10.12804/apl32.1.2014.12.

Barabási, A.L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T. 2002. Evolution of the social network of scientific collaborations. *Physica A*. 311, 590-614. https://doi.org/10.1016/S0378-4371(02)00736-7.

Bordons, M., Gómez, I. 2000. Collaboration networks in science. In: B. Cronin, H.B. Atkins (Eds.). *The Web of Knowledge: a festschrift in honor of Eugene Garfield*. Medford, Information Today, pp. 197-213.

Bose, I., Mahapatra, R.K. 2001. Business data mining - a machine learning perspective. *Information & Management* 39 (3), 211-225. https://doi.org/10.1016/S0378-7206(01)00091-X.

Ceci, S.J., Ginther, D.K., Kahn, S., Williams, W.M. 2014. Women in academic science: A changing landscape. *Psychological Science in the Public Interest* 15 (3), 75-141. https://doi.org/10.1177/1529100614541236.

Burt, R.S. 1995. *Structural holes: the social structure of competition*. Harvard University Press, Cambridge.

Crane, D. 1969. Social structure in a group of scientists: a test of the "invisible college" hypothesis. *American Sociological Review* 34, 335-353. https://doi.org/10.2307/2092499.

Filippo, D., Levin, L. 2015. Estudio de la producción Ibérica en Biblioteconomía y Documentación a través del análisis de Comunidades Bibliográficas. In: *Desafíos y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC 2015 (Madrid, 16 y 17 de noviembre de 2015)*. Universidad Complutense de Madrid, Madrid.

García-Ruiz, J.M. 1999. *La producción científica de la Geomorfología española y su impacto, a través de las publicaciones periódicas*. Instituto Pirenaico de Ecología, Zaragoza.

García-Ruiz, J.M. 2008. Las citas internacionales de las revistas españolas de Geografía. *Boletín de la Asociación de Geógrafos Españoles* 46, 207-225. http://age.ieg.csic.es/boletin/46/11-LAS%20CITAS.pdf.

Glänzel, W., Schubert, A., Czerwon, H.J. 1999. A bibliometric analysis of international scientific cooperation of the European Union (1985-1995). *Scientometrics* 45 (2), 185-202. https://doi.org/10.1007/BF02458432.

Glass, C., Cook, A. 2016. Leading at the top: Understanding women's challenges above the glass ceiling. *The Leadership Quarterly* 27 (1), 51-63. https://doi.org/10.1016/j.leaqua.2015.09.003.

Goh, K.I., Oh, E., Kahng, B., Kim, D. 2003. Betweenness centrality correlation in social networks. *Physical Review E* 67, 017-101. https://doi.org/10.1103/PhysRevE.67.017101.

Gómez-Ortiz, A., González-Martín, J.A. 1988. Geomorfología periglaciar: desde una visión histórica hacia nuevas perspectivas. In: M. Gutiérrez Elorza, J.L. Peña Monné (Eds.), *Perspectivas en Geomorfología*. SEG, Zaragoza, pp. 79-94.

Gómez-Ortiz, A., Palacios, D. 1995. La investigación española sobre geomorfología de regiones frías. Áreas geográficas, estudios actuales y resultados. *Revista de Geografía* 29 (1), 7-19. https://www.raco.cat/index.php/RevistaGeografia/article/view/50289.

Gómez-Ortiz, A., Martí-Bono, C.E., Salvador, F. 2001. Evolución reciente de la Geomorfología glaciar y periglaciar en España (1980-2000), balance y perspectivas. In: A. Gómez-Ortiz, A. Pérez-González (Eds.), *Evolución reciente de la Geomorfología española (1980-2000)*. Geoforma, Logroño, pp. 139-184.

Gómez-Ortiz, A., Palacios Estremera, D., Salvador Franch, F. 2003. La investigación reciente en Geomorfología Periglaciar en España. La labor de la IPA-España. *Scripta Nova* 7 (155), 133-156. http://www.ub.es/geocrit/sn/sn-155.htm.

Gutiérrez, F., Harvey, A., García-Ruiz, J.M., Silva, P., Cendrero, A. 2013. Geomorphological research in Spain. *Geomorphology* 196, 1-12. https://doi.org/10.1016/j.geomorph.2012.12.014.

Gupta, V., Lehal, G.S. 2009. A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence* 1 (1), 60-76. https://doi.org/10.4304/jetwi.1.1.60-76.

Haythornthwaite, C. 2006. Learning and knowledge networks in interdisciplinary collaborations. *Journal of the American Society for Information Science and Technology* 57 (8), 1079-1092. https://doi.org/10.1002/asi.20371.

Holmes, M.A., O'Connell, S., Frey, C., Ongley, L. 2008. Gender imbalance in US geoscience academia. *Nature Geoscience* 1 (2), 79-82. https://doi.org/10.1038/ngeo113.

Katz, J.S., Martin, B.R. 1997. What is research collaboration? *Research Policy* 26, 1-18. https://doi.org/10.1016/S0048-7333(96)00917-1.

Kosoko-Lasaki, O., Sonnino, R.E., Voytko, M.L. 2006. Mentoring for women and underrepresented minority faculty and students: experience at two institutions of higher education. *Journal of the National Medical Association* 98 (9), 1449. https://pdfs.semanticscholar.org/8cbf/f61aa43d95e4d354f56acccc5f94f286cbb0.pdf.

Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., Sugimoto, C.R. 2016. Contributorship and division of labor in knowledge production. *Social Studies of Science* 46 (3), 417-435. https://doi.org/10.1177/0306312716650046.

Lasanta, T., García-Ruiz, J.M. 2013. La producción de la geografía española a través de las bases de datos de Scopus e ISI Web of Knowledge. In: T. Lasanta-Martínez, J. Martín-Vide (Coords.), *La Investigación Geográfica en España (1990-2012)*. Asociación de Geógrafos Españoles, Instituto Pirenaico de Ecología (CSIC), Centro de Ciencias Humanas y Sociales (CSIC), Madrid, pp. 221-234. http://digital.csic.es/bitstream/10261/88158/1/Publicaciones%20de%20geografos%20espa%C3%B1oles.pdf.

López-Ontiveros, A., Molinero, F. 1995. Evolución reciente de la investigación geográfica española. *Boletín de la Asociación de Geógrafos Españoles* (21-22), 119-132.

Martins, D. 2011. A emergência da análise de redes sociais como campo de pesquisa: perspetiva da análise da produção científica em português e espanhol a partir do Google académico, *Alexandria: Revista de Ciencias de la Informacion*, 5 (8), 17-29. http://revistas.pucp.edu.pe/index.php/alexandria/article/view/214/208.

Molina, J.L., Muñoz-Justicia, J.M., Domènech I-Argemí, M. 2002. Redes de publicaciones científicas: un análisis de la estructura de coautorías. *Redes: Revista Hispana para el Análisis de Redes Sociales* 1, 3-31. https://doi.org/10.5565/rev/redes.29.

Ministerio de Educación, Cultura y Deporte. 2016. *Datos y Cifras del Sistema Universitario Español. Curso 2015-2016*. Available at: https://www.educacionyfp.gob.es (last access: 24/04/2019).

Newman, M.E. 2000. Models of the small world. *Journal of Statistical Physics*, 101, 819-841. https://doi.org/10.1023/A:1026485807148.

Newman, M.E. 2001a. Scientific collaboration networks. I. Network construction and fundamental results, *Physical Review* 64, 01613-1-8. https://doi.org/10.1103/PhysRevE.64.016131.

Newman, M.E. 2001. The structure of scientific collaboration networks. *PNAS* 98, 404-409. https://doi.org/10.1073/pnas.021544898.

Newman, M.E. 2004. Coauthorship networks and patterns of scientific collaboration. *PNAS* 101, 5200-5205. https://doi.org/10.1073/pnas.0307545100.

Ngai, E.W.T., Xiu, L., Chau, D.C.K. 2009. Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications* 3. https://doi.org/10.1016/j.eswa.2008.02.021.

Paderewski-Rodríguez, P., García-Arenas, M.I., Gil-Iranzo, R.M., González, C.S., Ortigosa, E.M., Padilla-Zea, N. 2017. Initiatives and strategies to encourage women into engineering. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* 12 (2), 106-114. https://doi.org/10.1109/RITA.2017.2698719.

Peña-Monné, J.L. 1998. Los estudios sobre procesos de clima frío en España: balance y perspectivas. In: A. Gómez-Ortiz, F. Salvador, L. Schulte, A. García (Eds.), *Procesos biofísicos actuales en medios fríos*. Universitat de Barcelona. Barcelona, pp. 43-54.

Pereira, D., Díaz, C. 2016. Are women properly represented in scientific publication and research? Interim results from a Spanish case study in Earth Sciences. *Episodes: Journal of International Geoscience* 39 (1), 52-58. https://doi.org/10.18814/epiiugs/2016/v39i1/89237.

Price, D.J.S. 1965. Networks of scientific papers. *Science* 149, 510-515. https://doi.org/10.1126/science.149.3683.510.

Price, D.J.S. 1966. Beaver D. Collaboration in an invisible college. *American Psychologist* 21, 1011-1018. https://doi.org/10.1037/h0024051.

Reagans, R., Zuckerman, E.W. 2001. Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams. *Organization Science* 12, 502-517. https://doi.org/10.1287/orsc.12.4.502.10637.

Scott, J. 2012. *Social network analysis*. Sage, London. https://doi.org/10.1007/978-1-4614-1800-9_178.

Serrano, E. 2016. Periglaciarismo y permafrost. *Polígonos. Revista de Geografía* 28, 15-29. https://doi.org/10.18002/pol.v0i28.4283.

Valderrama-Zurián, J.C., González-Alcaide, G., Valderrama-Zurián, F.J., Aleixandre-Benavent, R., Miguel-Dasit, A. 2007. Redes de coautorías y colaboración institucional en Revista Española de Cardiología. *Revista Española de Cardiología* 60 (2), 117-130. https://doi.org/10.1157/13099458.

Varanda, M., Rego, R., Fontes, B., Eichner, K. 2012. A análise de redes sociais no mundo lusófono: contributos para o conhecimento de uma comunidade científica em desenvolvimento. *Redes: Revista Hispana para el Análisis de Redes Sociales* 22, 147-188. https://doi.org/10.5565/rev/redes.434.

Vargiu, E., Urru, M. 2012. Exploiting web scraping in a collaborative filtering-based approach to web advertising. *Artificial Intelligence Research* 2 (1), 44. https://doi.org/10.5430/air.v2n1p44.

Wagner, C.S. 2005. Six case studies of international collaboration in science. *Scientometrics* 62 (1), 3-26. https://doi.org/10.1007/s11192-005-0001-0.

Wagner, C., Leydesdorff, L. 2003. Seismology as a case study of distributed collaboration in science. *Scientometrics* 58, 91-114. https://doi.org/10.1023/A:1025479524390.

Wagner, C.S., Whetsell, T.A., Leydesdorff, L. 2017. Growth of international collaboration in science: revisiting six specialties. *Scientometrics* 110 (3), 1633-1652. https://doi.org/10.1007/s11192-016-2230-9.

Wasserman, S., Faust, K. 1994. *Social network analysis: Methods and applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815478.

Watts, D.J., Strogatz, S.H. 1998. Collective dynamics of 'Small-World' networks. *Nature* 393, 440-442. https://doi.org/10.1038/30918.