Scientometrics (2013) 94:1239–1251
DOI 10.1007/s11192-012-0792-8

Iberian universities: a characterisation from ESI rankings

Tânia F. G. G. Cova · Alberto A. C. C. Pais · Sebastião J. Formosinho

Received: 5 April 2012 / Published online: 14 June 2012

Abstract The access to bibliographic and citation databases allows to evaluate scientific performance, and provides useful means of general characterisation. In this paper we investigate the clustering of Iberian universities, resulting from the similarity in the number and specific nature of the scientific disciplines given by the Essential Science Indicators database. A further refining of the analysis, as provided by PCA, clearly reveals the relationship between the universities and the scientific disciplines in the main groups. Similarity between universities is not dictated only by the number of areas in the ranking, but also stems from the nature of the ranked scientific areas and the specific combination in each university.

Keywords Iberian universities · Ranking areas · Essential science indicators · Principal component analysis

Introduction

In recent decades, the expansion of science and technology has been coupled to an increasing necessity of evaluating scientific productivity in the various disciplines of knowledge. This has made the measure of scientific output of researchers and institutions, an important task for the scientific community. Universities are, in most cases, the backbone of higher education and scientific research (Braun et al. 2007). The evaluation of their performance is of paramount importance because they live in an enourmously competitive marketplace. There are different ways to assess science and information flows. The difficulty is to clearly understand what good science is and to know if the chosen data reflects quality (Bornmann et al. 2012). In fact, the access to bibliographic and citation databases is changing the way scientific performance is evaluated (Moedet al. 1995; Cronin 2001; Adam 2002; Bar-Ilan et al. 2007; Csajbók et al. 2007; Thelwall 2008; Tian et al. 2008; Sicilia et al. 2011). Several different indicators have been used in research evaluation due

T. F. G. G. Cova · A. A. C. C. Pais (✉) · S. J. Formosinho
Department of Chemistry, University of Coimbra, Rua Larga, 3004-535 Coimbra, Portugal
e-mail: pais@qui.uc.pt

to easy access to the relevant databases (Garfield 1995, 2006; Leydesdorff 2005). The majority of these indicators was developed to quantify both the production of researchers, using the total number of published papers and the number of papers published in a period of time, or the impact of their publications, using the total number of citations, the average number of citations per paper, the number and percentage of significant papers published, relative citation rates, or even combining some of those indicators (see e.g., Martin 1996; Bishop et al. 2003; Van Leeuwen et al. 2003; Hirsch 2005; Vinkler 2007; Schreiber 2008; Alonso et al. 2009; Vieira and Gomes 2010; Zhang et al. 2011; Leydesdorff and Rafols 2011; Serenko and Dohan 2011; Aksnes et al. 2012).

In a recent publication, some of us have investigated the way European countries are joined in clusters according to similarity in the profile of their citations, on the basis of the 22 scientific disciplines provided by essential science indicators (ESI) (Almeida et al. 2009). A clear pattern arises, implying that geographical and cultural factors strongly influence the scientific fabric of these countries. Thus, one of the major factors behind Science in Europe is the geographical proximity. Bilateral cooperation between countries does not suffice to account for such similarities.

The cluster that includes Portugal, Czech Republic and Hungary is an exception to the general trend of geographical proximity. However, it reflects the historical role of three cultures in these countries: Christian, Jewish and Muslim. On that perspective, one would expect that Spain would be in this same cluster but, in fact, as far as science is concerned, Spain has at present geographical proximity with the trans-Pyrenees cluster of France, Germany and Switzerland. This could reflect long term policies, planning and investment in science, because Spain clusters with Portugal in other fields, notably in patents profiles (OCDE 2010). This leads us to carry out a comparative study of the Iberian universities on their research contributions to new knowledge embodied in academic publication.

The ESI field rankings for universities provide a useful criterion to quantitatively assess universities on their best performances. In this study, 55 Iberian universities are compared taking into account 22 scientific disciplines, given by the ESI database, available on the Internet (ESI 2010).

This database considers the scientific citations from countries with at least one hundred thousand citations for all these scientific disciplines, in a period of 10 years. The updates of ESI always have a certain differential in relation to the original database (indexed articles of the Thomson Scientific journal), and from which they are drawn up (ISI 2010). The data corresponding to the update of ESI in July 2010 covers the period of 10 years + 4 months, January 1st 2000 to April 30th 2010. Table 1 shows the scientific disciplines being considered, in agreement with the descriptions of this database.

An important issue is the validity of considering citations in scientific disciplines as a relevant indicator for quality Science. We note that, although debatable (Vanclay 2011), citation-metrics have been used and are deemed appropriate (Moed et al. 2012), e.g, to assess journal quality or impact.

Citations correspond to an examination at an international level, which is repeated year after year upon publication of the paper. The obvious alternative, number of papers, corresponds to an examination at an international level, valid only per a certain period. Citations are thus an appropriate indicator to assess the impact of scientific and research institutions including universities, which are the driving force of scientific production in almost every country in the world. Recognizing this fact, ESI considers the scientific areas of each institution (Field Rankings) only if they are located in the top 1 % of worldwide citations in the respective area: the so-called “ranking areas”.

Table 1 Description of the scientific disciplines, using the original notation of the ESI database

Scientific disciplines	
Agricultural Science	*Mathematics*
Biology and Biochemistry	*Microbiology*
Chemistry	*Molecular Biology and Genetics*
Clinical Medicine	*Multidisciplinary*
Computer Science	*Neuroscience and Behavior*
Economics and Business	*Pharmacology and Toxicology*
Engineering	*Physics*
Environment/Ecology	*Plant and Animal Science*
Geosciences	*Psychiatry/Psychology*
Immunology	*Social Sciences: General*
Materials Science	*Space Science*

Data analysis

Data analysis resorted to standard chemometrics techniques, including hierarchical cluster analysis (HCA), with Ward's linkage, for exploratory assessment of the data structure (Ward Jr. 1963) and principal component analysis (PCA), for data visualization and variable selection, on the variance–covariance matrix (Jolliffe 2002). The software code was developed by the authors in Octave (Eaton et al. 2007).

The main goal of clustering algorithms is to group the data into a number of sensible clusters according to their similarities. In the exploratory data analysis, HCA was used, for which a recent review can be found in reference (Almeida et al. 2007). Ward's method is a standard procedure for such analysis, and will be the one selected for the present study. In this method the distance between objects or groups is established based on minimizing the within cluster variance, when the new cluster is constructed (Ward Jr. 1963).

The HCA procedure is graphically represented by a dendrogram, which consists of a cluster structure and illustrates the fusions or divisions made at each successive stage of the analysis. It allows inspecting the overall structure of the data, and estimate the number of clusters.

In turn, PCA allows compressing the data by reducing the number of dimensions, with a minimized loss of information. The most influential variables in the system are highlighted, and underlying factors may be identified. Based on an orthogonal linear transformation, PCA defines a lower dimensionality system, such that the highest variance of the data comes projected on the first principal component (PC1), the second largest on the second coordinate (PC2), and so on (Brereton 2003). PCA requires the solution of an eigenvalue problem, either based on the correlation or variance/covariance matrices of the original variables. In this work we present results based on the variance/covariance approach.

The dataset analyzed includes information on 55 independent universities. In this study, we excluded the "Multidisciplinary" discipline which covers articles published in Nature, Science and Proceedings of the National Academy of Sciences, but 98% of which are distributed by the other 21 areas based on scientific citations (ESI 2010, URL accessed in January and July 2010).

The techniques require a description of the objects, i.e., points in Euclidean space. In this analysis, the universities correspond to the objects. Each university is described by a vector with 21 components, each component given the value one if the area is ranked, and zero otherwise. The total number of ranking areas in one university will be denoted as *Rk*.

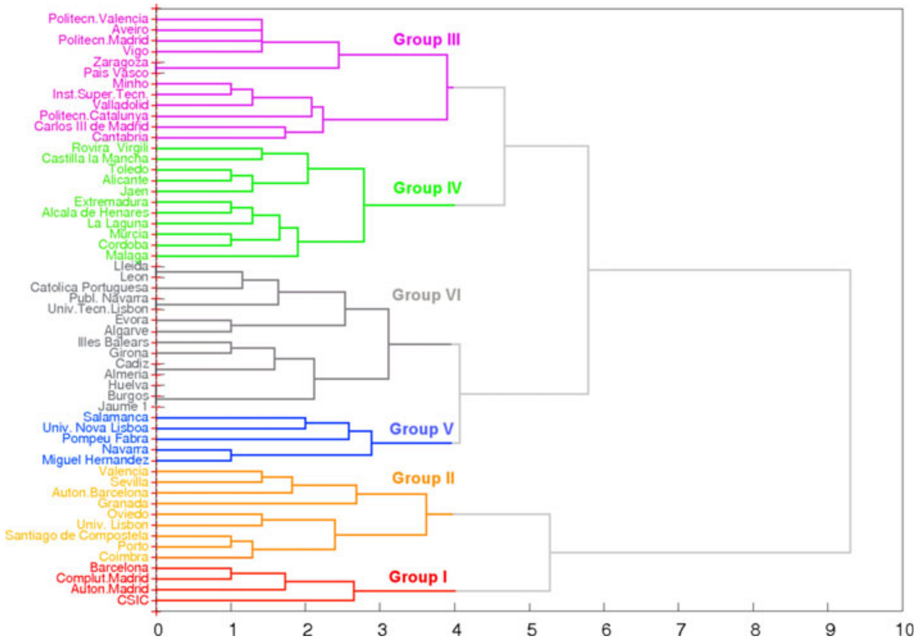

Fig. 1 Similarity among Iberian universities in terms of scientific areas (Field Rankings) contemplated by the ESI database in January 2010. Dendrogram constructed resorting to Ward's method with Euclidean distances, using the 21 scientific areas as variables defining each institution

Relative positioning of the Iberian universities

The dendrogram presented in Fig. 1 provides a very simple two dimensional plot of the data structure indicating the merging universities and the merging distances. It is constructed on the basis of the total existing information for the 55 institutions. From this figure, it is apparent that the data possesses a super-structure in which six groups of institutions are visible, highlighted by the colours. These groups were established based on a distance cut close to $d = 4$.

The "Top Universities" (Group I), whose group includes the *CSIC (Consejo Superior de Investigaciones Scientificas*), and the "Classical Universities" (Group II) can be found in the lower end of the figure. These are very different from the remaining four groups. In the top are apparent the "Polytechnic Universities" (Group III), a group so called because it incorporates all polytechnic universities in Spain, which is not very distant from Group IV. In turn, Groups V and VI are the closer.

Spain possesses the *CSIC* ($Rk = 18$), which is the largest public institution devoted to research in the country and the third at the European level, and three "Top Universities", the *University of Barcelona* ($Rk = 19$), *Complutense de Madrid* ($Rk = 16$) and *Autonoma de Madrid* ($Rk = 17$). In the group of "Classical Universities" there are six universities from which the *Autonoma de Barcelona* shows the highest Rk, with $Rk = 15$, then the university of *Valencia*, with $Rk = 14$. However, clustering does not depend solely on the value of Rk, and is also dependent on the nature of the scientific areas. For example, *Granada* and *Sevilla*, both with $Rk = 12$, and *Oviedo* with $Rk = 7$, the lowest ranking, are in the same

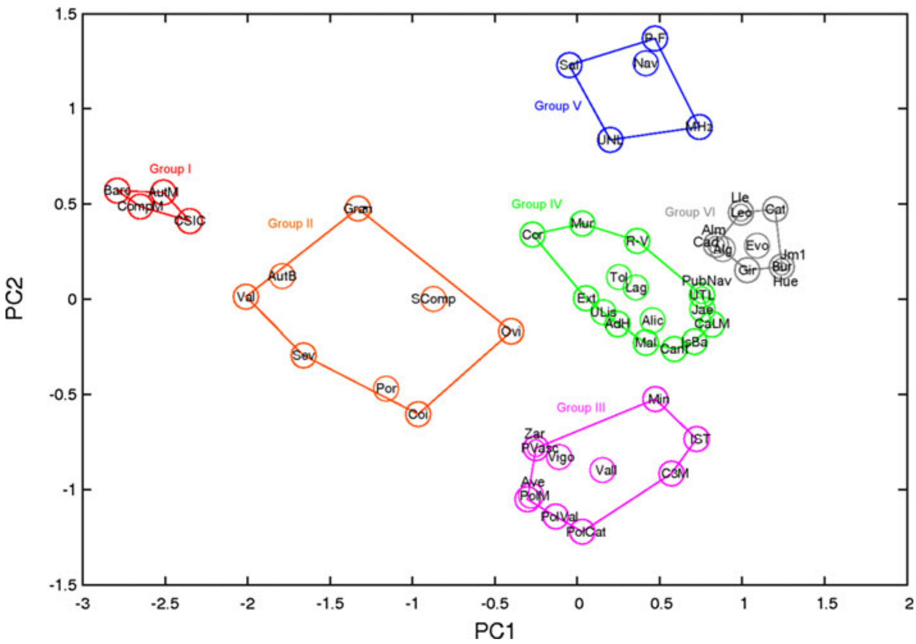

Fig. 2 Representation of the Iberian universities on the two main principal components. Each group is depicted in 2D convex hull form, showing the groups boundaries. The *red*-coloured group refers to the "Top Universities" (Group I), the *orange*-coloured group refers to the "Classical Universities" (Group II), the *purple*-coloured one to the "Polytechnic Universities" (Group III). The remaining three groups (Groups IV, V and VI) correspond to groups of general characteristics with intermediate and low impact. For convenience, the names of institutions have been duly abbreviated (color figure online)

group. The university of *Santiago de Compostela* with a *Rk* = 10 has the same ranking level of the best universities in Portugal, *Porto* and *Coimbra*.

Figure 2 depicts a composed view of the universities represented in the new orthonormal principal component system. This representation, in two dimensions, allows the visual discrimination between the institution classes. As a preliminary PCA result, the data scores representation is in direct agreement with the results obtained via HCA. For clarity, the groups are identified by matching colours in HCA and PCA. The boundaries of these groups do not overlap. Furthermore, this figure shows that the descrimination between the universities of higher and lower impact on research lies essentially on PC1.

In the update of July 2010, the dendrogram suffers some changes in the distribution of the groups, because some of the institutions have changed, by loss or gain of ranking areas. Figure 3 presents the new dendrogram, corresponding to this update. The colours of the original large groups are mantained for easier identification of changes. In this structure, there are some transitions from one group to a neighboring one: for example, the group of "Top Universities" (Group I) finds himself enriched with another institution, the *Univ. Autonoma de Barcelona*, lost by the group of "Classical Universities" (Group II). *Granada, Oviedo* and *Lisbon*, have evolved from the "Classical" into groups of lower impact. However, according to PCA results (see below), *Granada* remains in the orbit of this group, and possesses a high *Rk*.

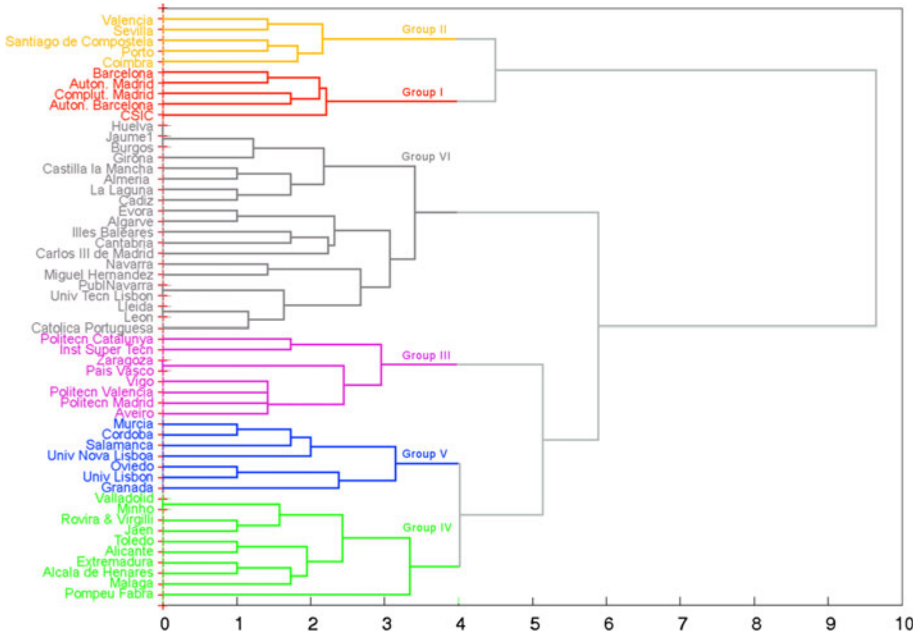

Fig. 3 Dendrogram constructed resorting to Ward's method with euclidean distances, using the 21 scientific areas for the update of July 2010

Discrimination between universities

Let us now start analyzing the PCA results on the update of July 2010. Note that PCA replaces most of exploratory data analysis tasks, and provides a good data visualization, as shown before for the January 2010 results.

Table 2 presents the dependency of the PCA transformed data on each original variable, for the first three principal components. The criterion for selecting a significant load is based on the comparison to the average expected value. Since PC constitutes an orthonormal vectorial base, for a *m* dimensional case we expect an average value of

$$1/\sqrt{m} \tag{1}$$

In summary, the variables of highest weight for the first component (PC1) are *Biology and Biochemistry*, *Physics* and the *Environment/Ecology*, while *Materials Science*, *Engineering* and *Physics* have the highest contribution to the second component (PC2). Those that contribute the most to the third component (PC3) are *Agricultural Science*, *Plant and Animal Science* and *Clinical Medicine*.

Figure 4 depicts the scores of the Iberian universities in two dimensions. The third dimension (PC3) is represented by bars whose amplitude can be read on the PC2 axis. From Fig. 4, it is seen that the six groups spread essentially along PC1, but PC2 also contributes for the discrimination of some of them. It is also seen that the third component displays additional information for each group. The groups of the "Top Universities" (Group I) and "Classical Universities" (Group II) are clearly detached from each other and from the remaining ones. Group III, "Polytechnic", is also clearly separated in the PC1/PC2 plane from every other group. These observations are compatible with those extracted from Fig. 3. Some degree of overlap is found between groups IV and V and between

Table 2 Impact of all original variables on the first three principal components

Scientific disciplines	PC1	PC2	PC3
Agricultural Science	−0.165	0.170	**0.606**
Biology and Biochemistry	**−0.331**	**0.271**	−0.175
Chemistry	−0.159	**−0.220**	**−0.249**
Clinical Medicine	**−0.251**	−0.197	**−0.448**
Computer Science	−0.077	−0.216	0.155
Economics and Business	0.015	0.027	−0.080
Engineering	**−0.234**	**−0.439**	**−0.223**
Environment/Ecology	**−0.305**	−0.084	−0.038
Geosciences	−0.185	0.100	−0.015
Immunology	−0.155	0.092	0.030
Materials Science	**−0.269**	**−0.478**	0.123
Mathematics	**−0.250**	0.028	0.034
Microbiology	**−0.247**	0.146	0.027
Molecular Biology and Genetics	−0.157	**0.218**	−0.032
Neuroscience and Behavior	**−0.228**	**0.287**	−0.029
Pharmacology and Toxicology	**−0.270**	0.037	0.032
Physics	**−0.331**	**−0.338**	0.110
Plant and Animal Science	−0.207	0.075	**0.461**
Psychiatry/Psychology	−0.136	0.076	0.023
Social Sciences: General	−0.211	0.195	−0.104
Space Science	−0.038	0.019	0.010

The most relevant contributions are highlighted in bold

groups IV and VI. Note, however, that if the dominant value of the third principal component is used, the closer groups are IV and V, not far from III in PC1, but the latter differing in both PC2 and PC3, and, finally, the most remote is VI, both in PC1 and PC3. This, again, is compatible with the observations from Fig. 3.

The "Top Universities", with the highest values of *Rk* (between 15 and 19), are depicted on the negative side of PC1. In the positive side are present those universities characterized by small values of *Rk* (for example, *Evora* and *Catolica Portuguesa*). For Group II, the "Classical" universities, *Rk* values vary between 10 and 14. In the "Polytechnic", Group III, the variation is between 5 and 7. These values are easily understandable, since the top and classical universities are active in most areas, while the latter focus essentially on technical disciplines. In Fig. 5 it is noticeable that the distribution of the institutions along the first principal component (PC1), are essentially based in the number of the ranking areas (*Rk* values).

In order to characterize the second component (PC2) we now analyze, separately, three groups that are disposed or have a significant dispersion along this component: Group III formed by the "Polytechnic Universities", with a predominance of the field of Nanosciences, Group V of intermediate/high impact, with a predominance in the Biology field, and Groups IV and VI with intermediate/low and low impact.

Figure 6 further explores the PCA results, displaying the representation, on the basis of histograms, of the average values of the scientific areas of greatest contribution to the

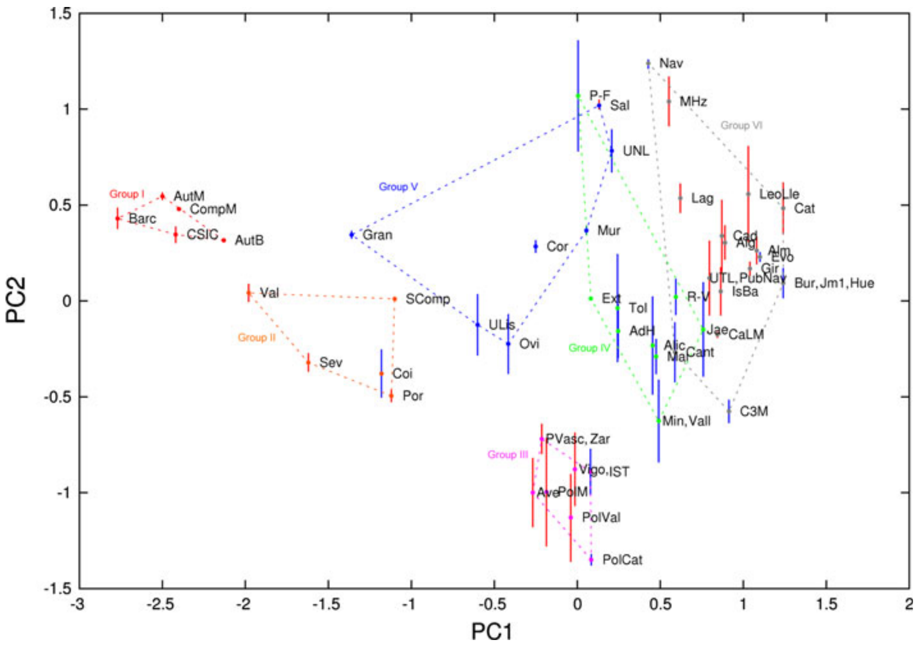

Fig. 4 Scatter plot of the Iberian universities in 2D for July 2010. The positive and negative scores are represented by *bars* in the third dimension. *Red* refers to positive scores, while *blue* corresponds to negative scores. For convenience, we used abbreviations for the names (color figure online)

second component (*Materials Science*, *Engineering*, *Physics*, *Neuroscience and Behavior*, *Biology and Biochemistry*, *Chemistry* and *Molecular Biology and Biochemistry*). This figure reflects the relevance of these areas in the distribution of the institutions in the four groups considered.

We can conclude that the predominant areas in Group III, “Polytechnic Universities”, are the field of Nanosciences, as *Materials Science* and *Engineering*, in addition to *Chemistry* and *Physics*. These areas are less dominant or even absent (case of *Materials Science* and *Physics*) in Group V. In the latter group, we highlight the areas of the field of Biology (*Biology and Biochemistry*, *Neuroscience and Behavior* and *Molecular Biology and Genetics*). Groups IV and VI are characterized by an intermediate dominance of all these areas.

Considering the highest loading in the second component (PC2), *Materials Science*, we can see that the corresponding average influences significantly the position of the four groups on PC2. Consider, as an example, the universities of *Navarra* (Group V) and *Polytechn Catalunya* (Group III), located on opposite sides of the second component (PC2). *Navarra* possesses areas such as *Molecular Biology and Genetics* and *Neuroscience and Behavior*, which are absent in Groups III, IV and VI. On the other hand, at the *Politecn Catalunya*, the areas of *Materials Science*, *Engineering*, *Physics* and *Chemistry* are dominant. These are not present in the former institution. Based on this example, we can conclude that PC2 reflects the influence of the specific nature of the scientific areas in the distribution of groups.

In Fig. 4, and as previously mentioned, the third component is represented by bars of amplitude proportional to the respective value (read on the axis of PC2). Different colours are used as this component is positive or negative. The working hypothesis that we will

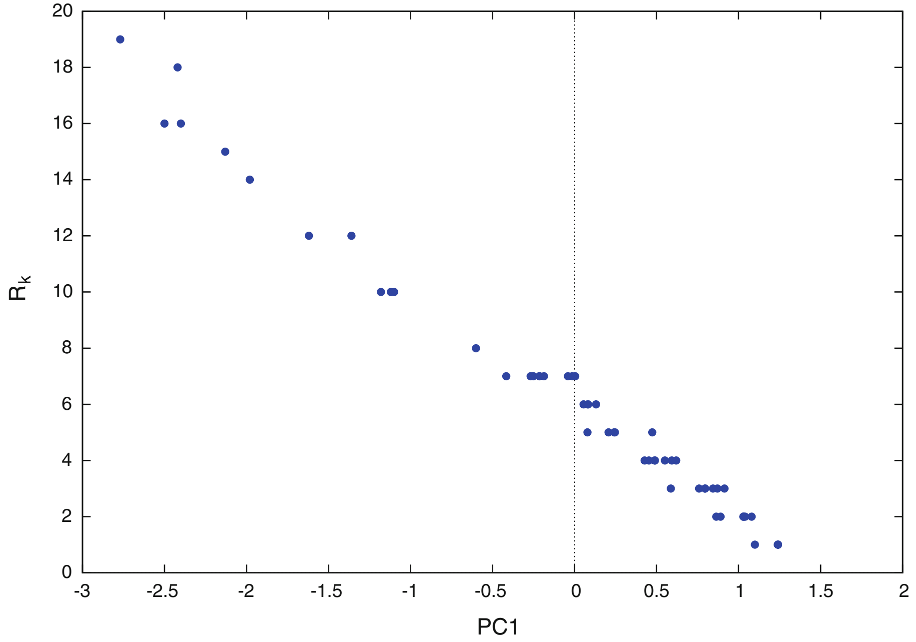

Fig. 5 Distribution of the universities along the first principal component according to the number of the ranking areas (*Rk* values). In the extreme negative of PC1 are located the "Top Universities" with the highest values of *Rk* (*Rk* = 15 to 19). Universities with a very small values of *Rk* are located at the opposite end

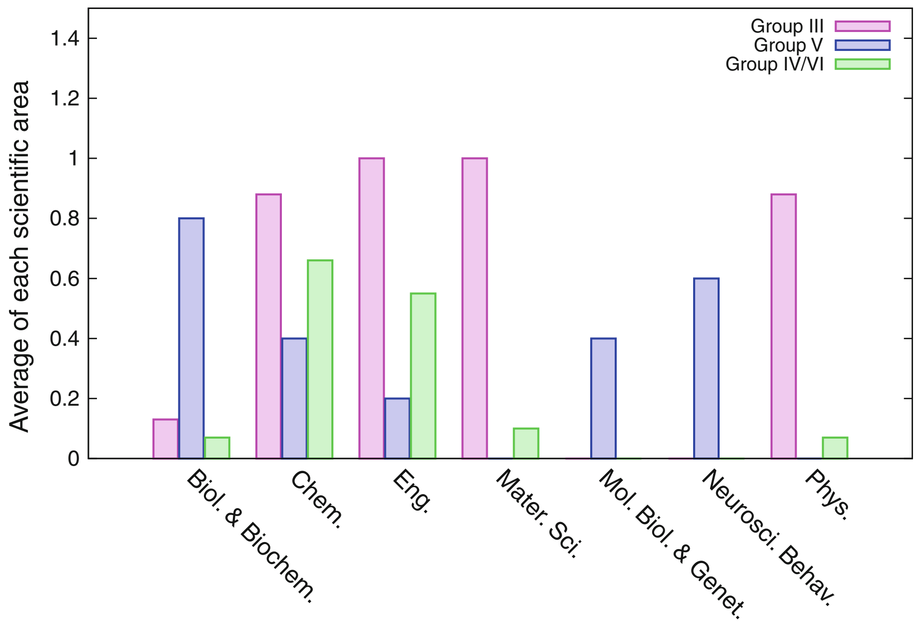

Fig. 6 Relative impact of the scientific areas of greatest contribution to the second component (PC2), based on the average values of each ranking area, for the four groups considered (III, V, IV and VI)

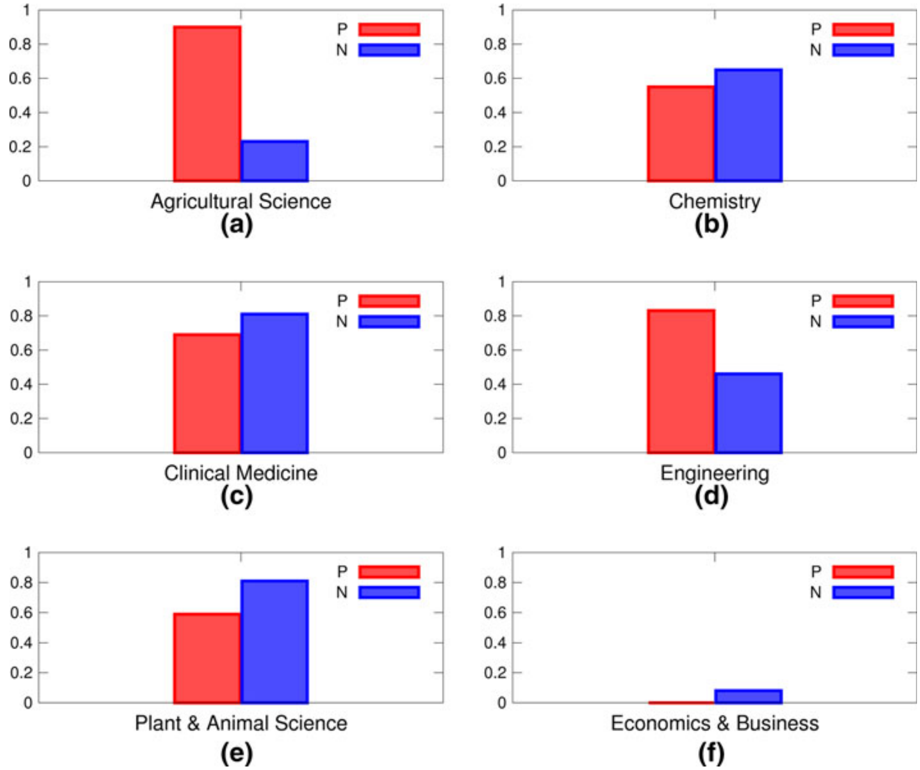

Fig. 7 Relevance of the scientific areas of greatest weight in the third component (PC3) in the discrimination of the institutions. The average of the variables for universities with positive scores are shown in *red*. The *blue bars* represent the average of the variables for universities with negative scores (color figure online)

consider is that the third component is characterized by the combination of certain scientific disciplines. The areas corresponding to *Agricultural Science* and *Plant and Animal Science* when combined with other areas of Nanoscience, such as *Computer Science*, *Engineering*, *Materials Science* and *Physics* leads to positive scores in PC3 and the corresponding bars have a very high amplitude. As an example of this type of combination we can quote universities *Polytechn Madrid, Polytechn Valencia* and *Aveiro*.

The negative scores arise as a result of the combinations of *Biology* and other areas as *Business and Economics*, *Social Sciences* and *Engineering*. An example of such combination stands out, *Pompeu Fabra*. This institution has a negative PC3 and high amplitude, due to the combination of *Business and Economics* and *Social Sciences*. The university of *Pompeu Fabra* is the only Iberian university that has the combination of *Economics and Business* and *Social Science: General* as ranking areas. The scores of the universities of *Navarra*, *Murcia*, *Salamanca*, *Santiago de Compostela* and *Granada* are virtually null in the PC3, given that these institutions present a very balanced distribution of areas in the Biology field.

An effort is now made to assess the relevance of the areas of greatest contribution in PC3 (*Agricultural Science*, *Plant and Animal Science*, *Clinical Medicine*, *Chemistry* and *Engineering*) to discriminate the institutions in the third dimension (positive and negative scores). Figure 7 shows the impact of the areas of greatest contribution to PC3, based on the average values.

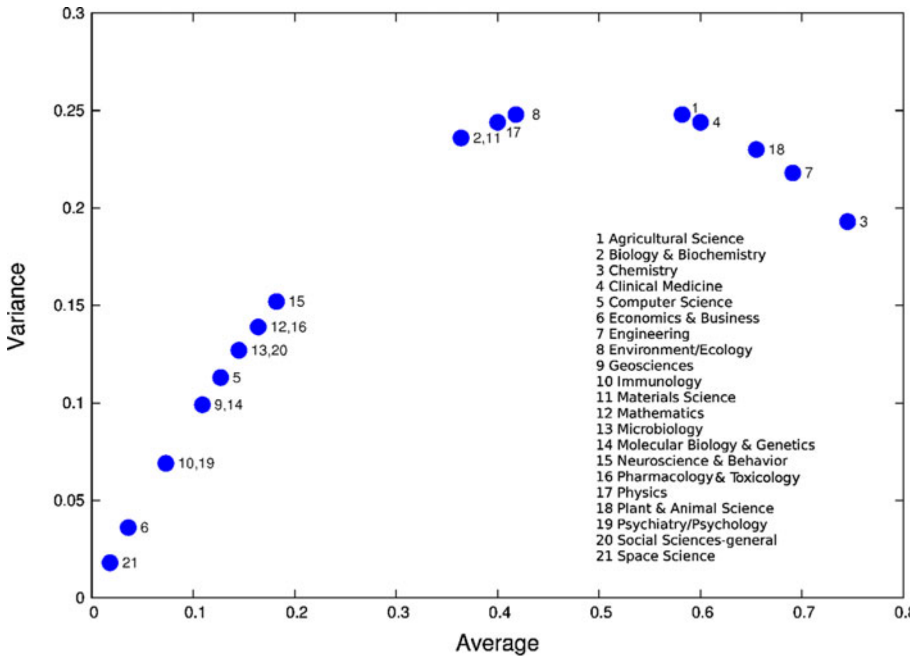

Fig. 8 Variance values as a function of the mean of each variable to the set of institutions under study (ESI 2010)

Panels (a) and (e) support the conclusions drawn previously: the scientific disciplines of *Agricultural Science* and *Plant and Animal Science* have a great influence on the discrimination of the institutions scores in the third dimension.

Figure 8 shows the variance values according to the average of each scientific area. We can consider that the variability is associated with the deviation from the intermediate value of 0.5 (for binary variables, 0 and 1). The closer to the intermediate value, the greater the variability associated to the variables.

The most significant loadings (of high variability) have the average close to 0.5. For example, *Agricultural Science* and *Environment/Ecology* are the scientific disciplines which have the highest variance (0.25), are at the same distance from the mean value (0.58 and 0.42, respectively). They are followed by *Clinical Medicine* and *Physics* with 0.24. On the other hand, the least significant loadings, such as *Space Science* and *Economics and Business* present the values of variance very close to zero (0.02 and 0.04, respectively). Scientific disciplines with an higher variance, i.e., close to an average value of 0.5, populate the higher loadings in the first components and are the main responsables for the descrimination between universities.

Conclusion

The results presented in this paper indicate that scientific disciplines are paramount for the evaluation of universities and their performance. It is recognised that the number of scientific disciplines in the ranking level is a quantitative measure of the performance of each institution in the scientific research. It is also stressed that the specific combination of some scientific disciplines discriminates the groups of universities, with an intermediate impact, in a qualitative way.

References

Adam, D. (2002). Citation analysis: The counting house. *Nature*, *415*, 726–729.

Aksnes, D., Schneider, J., & Gunnarsson, M. (2012). Ranking national research systems by citation indicators a comparative analysis using whole and fractionalised counting methods. *Journal of Informetrics*, *6*(1), 36–43.

Almeida, J., Barbosa, L., Pais, A., & Formosinho, S. (2007). Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. *Chemometrics and Intelligent Laboratory Systems*, *87*(2), 208–217.

Almeida, J., Pais, A., & Formosinho, S. (2009). Science indicators and science patterns in Europe. *Journal of Informetrics*, *3*(2), 134–142.

Alonso, S., Cabrerizo, F., Herrera-Viedma, E., & Herrera, F. (2009). h-index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, *3*(4), 273–289.

Bar-Ilan, J., Levene, M., & Lin A. (2007). Some measures for comparing citation databases. *Journal of Informetrics*, *1*, 26–34.

Bishop, N., Gillet, V., Holliday, J., & Willett, P. (2003). Chemoinformatics research at the university of sheffield: A history and citation analysis. *Journal of Information Science*, *29*(4), 249–267.

Bornmann, L., Schier, H., Marx, W., & Daniel, H. (2012). What factors determine citation counts of publications in chemistry besides their quality? *Journal of Informetrics*, *6*(1), 11–18.

Braun, T., Dióspatonyi, I., Zádor, E., & Zsindely, S. (2007). Journal gatekeepers indicator-based top universities of the world, of europe and of 29 countriesa pilot study. *Scientometrics*, *71*(2), 155–178.

Brereton, R. G. (2003). *Chemometrics: Data analysis for the laboratory and chemical plant*. Chichester: Wiley.

Cronin, B. (2001). Bibliometrics and beyond: Some thoughts on web-based citation analysis. *Journal of Information Science*, *27*(1), 1–7.

Csajbók, E., Berhidi, A., Vasas, L., & Schubert, A. (2007). Hirsch-index for countries based on essential science indicators data. *Scientometrics*, *73*(1), 91–117.

Eaton, J. W., Bateman, D., & Hauberg, S. (2007). *GNU Octave manual: A high-level interactive language for numerical computations*. Network Theory Ltd., version 3 for Octave version 3.2.4.

Essential Science Indicators (2010). The Thomson Corporation. http://scientific.thomson.com/products/esi. Accessed Janurary and July 2010, http://scientific.thomson.com/support/recorded-training/esi/. Accessed July 2010.

Garfield, E. (1995). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, *122*, 108–111.

Garfield, E. (2006). Citation indexes for science. a new dimension in documentation through association of ideas. *International journal of epidemiology, 35*(5), 1123–1127.

Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United states of America*, *102*(46), 165–169.

ISI Web of Knowledge (2010). The Thomson Corporation. http://isiwebofknowledge.com/. Accessed January 2010.

Jolliffe, I. (2002). *Principal component analysis*. (2^{nd} ed.). New York: Springer.

Leydesdorff, L. (2005). Evaluation of research and evolution of science indicators. *Current Science-Bangalore*, *89*(9), 1510–1517.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, *5*(1), 87–100.

Martin, B. (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics*, *36*(3), 343–362.

Moed, H., De Bruin, R., & Van Leeuwen, T. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics*, *33*(3), 381–422.

Moed, H. F., Colledge, L., Reedijk, J., Moya-Anegon, F., Guerrero-Bote, V., Plume, A., & Amin, M. (2012). Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in an informed way. *Scientometrics*. doi:10.1007/s11192-012-0679-8

Organisation for Economic Co-operation and Development (2010). http://stats.oecd.org/Index.aspx?DatasetCode=PATS_IPC. Accessed January 2010.

Schreiber, M. (2008). A modification of the h-index: The hm-index accounts for multi-authored manuscripts. *Journal of Informetrics*, *2*(3), 211–216.

Serenko, A., & Dohan, M. (2011). Comparing the expert survey and citation impact journal ranking methods: Example from the field of artificial intelligence. *Journal of Informetrics*, *5*, 629–648.

Sicilia, M., Sánchez-Alonso, S., & García-Barriocanal, E. (2011). Comparing impact factors from two different citation databases: The case of computer science. *Journal of Informetrics*, *5*, 698–704.

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of information science*, *34*(4), 605–621.

Tian, Y., Wen, C., & Hong, S. (2008). Global scientific production on gis research by bibliometric analysis from 1997 to 2006. *Journal of Informetrics*, *2*(1), 65–74.

Vanclay, J. K. (2011). Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*. doi:10.1007/s11192-011-0561-0

Van Leeuwen, T., Visser, M., Moed, H., Nederhof, T., & Van Raan, A. (2003). The holy grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, *57*(2), 257–280.

Vieira, E., & Gomes, J. (2010). A research impact indicator for institutions. *Journal of Informetrics*, *4*, 581–590.

Vinkler, P. (2007). Eminence of scientists in the light of the h-index and other scientometric indicators. *Journal of Information Science*, *33*(4), 481–491.

Ward, J., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244.

Zhang, L., Thijs, B., & Glänzel, W. (2011). The diffusion of h-related literature. *Journal of Informetrics*, *5*, 583–593.