

A structural analysis of publication profiles for the classification of European research institutes

BART THIJIS,^a WOLFGANG GLÄNZEL^{a,b}

^a *Katholieke Universiteit Leuven, Steunpunt O&O Indicatoren, Leuven (Belgium)*

^b *Hungarian Academy of Sciences, Institute for Research Policy Studies, Budapest (Hungary)*

In the present study we propose a solution for a common problem in benchmarking tasks at institutional level. The usage of bibliometric indicators, even after standardisation, cannot disguise that comparing institutes remains often like comparing apples with pears. We developed a model to assign institutes to one of 8 different groups based on their research profile. Each group has a different focus: 1. Biology, 2. Agricultural Sciences, 3. Multidisciplinary, 4. Geo & Space Sciences, 5. Technical and natural Sciences, 6. Chemistry, 7. General and Research Medicine, 8. Specialised Medicine. Two applications of this methodology are described. In the first application we compare the composition of clusters at national level with the national research profiles. This gives a deeper insight in the national research landscape. In a second application we look at the dynamics of institutes by comparing their subject clustering at two different points in time.

Introduction

The comparative analysis of the research performance of universities and non-university research institutes is often faced with the problem of their different research profiles. Publication activity and citation impact of different universities are – without standardisation and normalisation – not immediately comparable with each other. Moreover, even normalisation of indicators cannot disguise that comparing, e.g., a medical school with a business school still remains an exercise of ‘comparing apples

Received February 12, 2007

Address for correspondence:

BART THIJIS

Steunpunt O&O Indicatoren, KU Leuven, Dekenstraat 2, B-3000 Leuven, Belgium

E-mail: Bart.Thijs@econ.kuleuven.be

0138–9130/US \$ 20.00

Copyright © 2007 *Akadémiai Kiadó, Budapest*

All rights reserved

with pears'. This is one of the most serious common problems in benchmarking tasks at the institutional level. Another important aspect in macro and meso studies is that the research profiles of even multidisciplinary institutions do usually not represent the national situation. To overcome these problems, we develop a method to cluster research institutes in order to obtain groups of 'likewise' institutes in terms of their publication profile. Benchmarking can thus take place within the same group. The size of groups can also serve for the direct cross-country comparisons of national publication profiles giving a deeper insight into structures than the national total. Moreover, dynamic analysis sheds light on changes of the composition and size of profile clusters as well as on possible changes in the profile of individual institutions. This paper presents the methodology for clustering institutes and describes the creation of a predictive model that can be used for the assignment of institutes to one of the different groups. In the last section two different applications of this grouping are shown. However, let us start with the presentation of the data sources and the processing.

Data sources and processing

Data were extracted from the yearly updates of the *Web of Science* database of Thomson-ISI (Philadelphia, PA, USA). Only papers of the document type article, letter, note and review indexed in the 1992 to 2005 volumes have been selected. This data has undergone a detailed cleaning and then processed to bibliometric indicators.

Publications were assigned to countries and institutions according to the address in the by-line of the paper. A 3-step assignment procedure was developed. For each country under study, a list of distinct names of institutes as occurring in the extracted addresses is compiled. This list contains thus all possible synonyms and spelling variance/errors of research institutes. Secondly, each entry in this list with a number of publications above a certain threshold was assigned, if possible, to a unique, known institute. Finally, this thesaurus is matched with all addresses in a paper's by-line. This procedure has been applied to all publications with at least one European address. The papers were assigned to corresponding research institutes.

Subject classification of the publications was based on the field assignment of journals according to sixteen major fields of science developed in Leuven and Budapest [GLÄNZEL & SCHUBERT, 2003]. These fields are Agriculture & Environment, Biology (Organismic & Supraorganismic Level), Biosciences (General, Cellular & Subcellular Biology, Genetics), Biomedical Research, Clinical and Experimental Medicine I (General & Internal Medicine), Clinical and Experimental Medicine II (Non-Internal Medicine Specialties), Neuroscience & Behaviour, Chemistry, Physics, Geosciences & Space Sciences, Engineering, Mathematics, Social Sciences I (General, Regional & Community Issues), Social Sciences II (Economical & Political Issues) and Arts & Humanities.

Citations received by these papers have been determined for three-year citation window beginning with the publication year on the basis of an item-by-item procedure using special identification-keys (so-called cluster-keys) made up of bibliographic data elements. Citation data up to 2005 has been extracted from the WoS.

Methodology

In order to map the European institutional landscape we have applied an advanced version of a method developed by THUIS & GLÄNZEL [2006]. This methodology consists of three steps.

The first step in our procedure was the breakdown of the individual institutes' publication output into research fields and to construct their research profile. Unlike in the analysis of individual universities described by VAN RAAN [2004] where a 'spectral analysis' of the output based on ISI Subject Categories is applied, we do not need such fine-grained subject structures for the joint cluster analysis of the European institutional publication profiles, and have therefore used the 16 major fields of the sciences (13 fields), social sciences (2 fields) and humanities (1 field). This research profile can be seen as a vector holding the share of each of the 16 fields in the total set of publications of the respective institute. This data is standardised and takes only values between 0 and 1. This means that the number of papers an institute produces within a certain time frame has no effect on their profile. However this does not apply for institutes with a very low publication activity as their share in separate fields comes close to 0 or 1. In order to keep the influence of these small institutes within reasonable limits, small institutes with publication output beneath a given threshold are removed.

In a next step these research profiles are used as an input for a cluster analysis. Several clustering algorithms are utilised to test the stability of the proposed cluster solution. Different stopping rules and procedures suggested by MILLIGAN & COOPER [1985] are applied to determine the number of clusters.

After having obtained these groups, a predictive model is created using discriminant analysis. This model enables us to predict group membership of research institutes based on their research profile.

Research profiles

A set of 15 European countries was selected (EU15 without Greece but including Switzerland). All publications indexed between 2001 and 2003 with at least one address in one of these countries have undergone the assignment procedure described above. Table 1 shows the percentage of addresses that could be uniquely assigned to a European research institute for each country. This resulted in a set of 2775 institutes with at least one publication. For each of the institutes the total number of publications

over the three-year period was counted and the research profile was calculated. Figure 1 shows the distribution of the number of papers per institute. Only institutes with at least 20 publications were selected in this step. In all, 1767 institutes, that is, 63.7% of all institutes with publication output recorded in the database have met this criterion.

Table 1. Percentage of uniquely assigned addresses and number of institutes used in the analysis

Country	Assigned addresses	Number of institutes
Austria	87.8%	71
Belgium	98.7%	585
Denmark	92.8%	77
Finland	95.5%	91
France	71.9%	249
Germany	75.9%	206
Ireland	93.0%	58
Italy	88.0%	291
Luxemburg	66.0%	19
Netherlands	88.4%	149
Portugal	90.1%	55
Spain	83.9%	266
Sweden	93.1%	124
Switzerland	85.2%	78
UK	87.9%	456

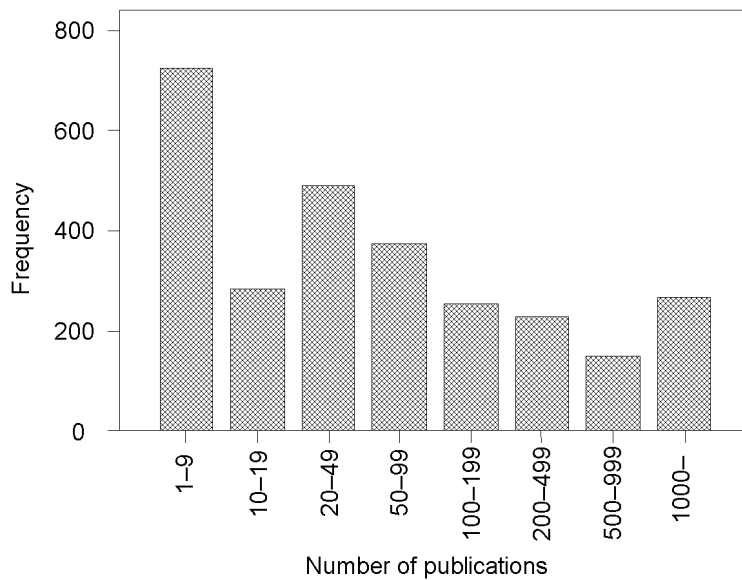


Figure 1. Distribution of papers over institutes in the 15 European countries

The clustering procedure

Cluster analysis, using the Ward algorithm with Squared Euclidean Distances, results in characteristic groups of universities and other academic institutions. In our study we have applied two different stopping rules to determine the number of clusters, particularly, the pseudo-F index according to CALINSKI & HARABASZ [1974] and $Je(2)/Je(1)$ index introduced by DUDA & HART [1973]. Large values of these indexes indicate distinct clustering. Figure 2 shows the results for different number of clusters (Duda-Hart's $Je(2)/Je(1)$ index is multiplied by 100). These results are not supportive for one particular number of clusters. The two-cluster solution suggested by both the $Je(2)/Je(1)$ and the pseudo-F index only separates medical institutes for non-medical ones. This rough classification has certainly a sense but proved not to be useful for the purpose described in the outset. The Duda-Hart $Je(2)/Je(1)$ suggests eight clusters as the second optimum solution.

These eight clusters can be characterised by the fields in which the institutes are specialized: Biology (BIO), Agriculture (AGR), Geo- and Space sciences (GSS), Technical and Natural sciences (TNS), Chemistry (CHE), General and Research Medicine (GRM), Specialised Medicine (SPM) and a last cluster with institutes having a Multidisciplinary profile (MDS).

Table 2 presents the average research profile for each of the eight clusters. Activity higher than 15% is highlighted. These profiles show a distinct and clear specialization within each individual cluster. Clusters 4 and 6 (GSS and CHE, respectively) are characterized by an extremely high degree of specialization; almost 90% of their research activities are devoted to one single field each. Even the two medical clusters (GRM and SRM) show the same high specialization which is, however, split up and distributed over the two subject fields in 'clinical and experimental medicine'; but the 2/3–1/3 proportion of these distributions are both contrary and complementary. The composition of the other clusters is more multidisciplinary. The TNS cluster (# 5) comprises the natural and technical sciences, the clusters BIO and AGR show considerable activity in the fields 'agriculture' and 'biology', interesting enough, almost mirroring the same contrary and complementary picture of GRM/SRM found in the life sciences. Finally, the third cluster (MDS) has been found a truly multidisciplinary cluster with activity in all science fields and less skewed publication distributions over fields; no field has a higher share than 20% here.

Social sciences and humanities make up only a very small share of the activity of the European institutions in all clusters.

The number of institutes assigned to the clusters is presented in Table 3. The clusters with multidisciplinary institutes and institutes with specialized medicine comprise jointly about 57% of all institutes. The clusters with Biology, Technical and Natural Sciences and the General and Research Medicine still hold a reasonable share

of institutes while the three small clusters, the Geo- and Space Sciences, Agriculture and Chemistry each hold about 3% of all institutes. The existence of one large cluster is often an undesired effect of the chosen linkage method (Ward) but in this case inspection of the data is clearly supportive for one larger multidisciplinary group.

Table 2. Research profile per cluster as expressed by the subject representation of their research output

Field	BIO	AGR	MDS	GSS	TNS	CHE	GRM	SPM
Agriculture (A)	20.0%	66.7%	6.0%	6.2%	4.9%	4.6%	0.2%	0.6%
Arts & Humanities (U)	0.0%	0.0%	0.3%	0.0%	0.2%	0.1%	0.1%	0.1%
Biology (Z)	57.2%	25.1%	9.6%	1.8%	2.3%	1.6%	4.1%	5.2%
Biomedical Research (R)	7.2%	5.2%	11.2%	0.1%	1.5%	1.2%	9.3%	9.8%
Biosciences (B)	14.6%	5.2%	15.5%	0.5%	2.6%	2.1%	7.0%	5.9%
Chemistry (C)	5.5%	14.3%	16.8%	2.9%	27.8%	86.6%	0.6%	1.1%
Engineering (E)	1.4%	5.1%	9.0%	4.0%	35.0%	7.3%	0.6%	0.7%
General & Internal medicine (I)	9.0%	1.2%	12.7%	0.2%	0.8%	0.1%	64.8%	31.2%
Geo- & Space science (G)	7.1%	5.5%	4.5%	87.4%	7.7%	1.3%	0.0%	0.1%
Mathematics (H)	0.8%	0.7%	5.7%	0.5%	6.9%	0.7%	0.1%	0.2%
Neuroscience & Behavior (N)	1.8%	0.5%	9.1%	0.0%	1.1%	0.1%	2.6%	4.8%
Non-internal medicine specialties (M)	7.7%	4.0%	16.8%	0.3%	2.5%	0.9%	33.5%	61.9%
Physics (P)	1.4%	3.1%	10.5%	4.7%	37.0%	10.8%	0.4%	0.6%
Social Sciences I (S)	0.6%	0.3%	1.7%	0.1%	0.4%	0.0%	0.7%	1.5%
Social Sciences II (O)	0.3%	0.5%	1.3%	0.0%	0.3%	0.2%	0.0%	0.1%

Table 3. Number of research institutes in each of the eight clusters

Cluster	Code	Counts	Share
Cluster 1 (Biology)	BIO	146	8.3%
Cluster 2 (Agriculture)	AGR	59	3.3%
Cluster 3 (Multidisciplinary)	MDS	550	31.1%
Cluster 4 (Geo- & Space Science)	GSS	57	3.2%
Cluster 5 (Technical & Natural)	TNS	261	14.8%
Cluster 6 (Chemistry)	CHE	51	2.9%
Cluster 7 (General & Research Medicine)	GRM	182	10.3%
Cluster 8 (Specialised Medicine)	SPM	461	26.1%
Total		1767	100.0%

Some of the most typical members of each group are listed here: 'Wageningen University and Research Center' with about 40% of all publications assigned to our subject field "biology", the 'Danish Institute of Agricultural Sciences' is with 72% in agriculture a true representative of its group. In the third cluster most of the large European universities with many specialties are grouped. Examples are: 'Catholic University of Louvain (K.U.Leuven)' or 'LMU Munich'. Obviously, most of the national research councils are included in this cluster as well. In the fourth cluster (Geo & Space Sciences) we can find the Italian 'National Institute for Astrophysics (INAF)' and the Spanish 'Astrophysics Institute of the Canary Islands (IAC)'. The French institute 'CEA (Commissariat à l'Énergie Atomique)' is one of the typical members of

the group specialized in Technical and Natural Sciences, others are 'Delft University of Technology' and 'IMEC'. In the Chemical group we can find 'BASF AG', 'Institut Francais du Petrole' as well as the Dutch company 'DSM'. Our grouping resulted in two clusters with a main focus on medical sciences. In the first medical group the focus is more on general and research medicine. General hospitals make up a large part of this group, e.g. 'Niguarda CA Hospital of Milan'. Other institutes in this group are the 'European Institute of Oncology' or 'Netherlands Cancer Institute and Antoni van Leeuwenhoek Hospital'. In the last cluster we find several universities like the 'Erasmus University Rotterdam' or 'Medical University of Luebeck'. We can also find specialized institutes like the 'National institute for the rest and care of the elderly' or 'Nuffield Orthopaedic Centre' among the institutions with specialized medical profile.

Predictive model

In a last step a predictive model is created for the assignment of other research institutes or other research profiles to one of the 8 groups. The 1767 institutes used in the cluster analysis are used as the training set. The resulting model can then be applied to the other institutes. For the creation of this model we use discriminant analysis. This technique uses linear functions (latent variables) of the predictive variables in the dataset, in this case the vector representing the research profile. Such a function or canonical root classifies cases into one of two groups. By adding a function to the analysis it is possible to distinguish between one more group. This means that for the classification into 8 separate groups we'll need 7 different linear functions. Discriminant analysis can be disturbed by out-liners but these cases were mostly removed by excluding institutes with less than 20 publications.

The statistics for these 7 discriminant functions are presented in Table 4. The *eigenvalue* indicates the importance of this function for the classification of cases into the given groups. The '*% of variance*' statistic is the share that each function has in the total of the explained variance. In this case, the first three functions account for nearly 80% of all explained variables. The *canonical correlation*, finally, gives information on the association of the grouping by the discriminant function and the dependent variable. Each of these functions also have a significant Wilks' Lambda value which indicates that different groups have indeed different, discriminating mean values on this function.

Based on these functions a classification model is constructed using Fisher's coefficients. This results in eight different linear functions, each assigned to one of the different groups. For each observation the resulting value for each function can be calculated. The observation is assigned to the group connected to the function with the highest value. The exact coefficients of each classification function can be found in Appendix 1.

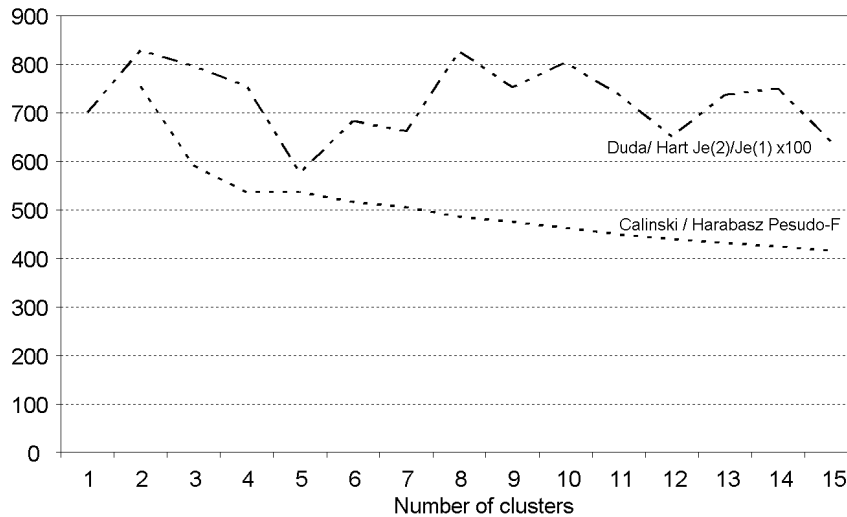


Figure 2. Two different stopping rules applied to the classification of European research institutions

Table 4. Eigenvalues and statistical functions of the discriminant analysis

Function	Eigenvalue	% of variance	Cumulative %	Canonical correlation
1	7.572(a)	36.4	36.4	0.940
2	4.573(a)	22.0	58.4	0.906
3	4.290(a)	20.6	79.0	0.901
4	1.320(a)	6.3	85.3	0.754
5	1.179(a)	5.7	91.0	0.736
6	1.171(a)	5.6	96.6	0.734
7	0.700(a)	3.4	100.0	0.642

The main outcome of the discriminant analysis can be summarized as follows. The constructed model was used to reclassify the 1767 institutes into the 8 different groups. 93% of all institutes were assigned correctly. This ratio, the significant values of Wilks' Lambda and the Canonical Correlations substantiate the predictive power of the model and justify the use of this model as a predictive tool.

Application

In this section we show two different applications of the proposed methodology. First we use the grouping to get a deeper insight in the national research characteristics by comparing national profiles to group sizes within each country. Next we show the

application of this method to shed light on dynamics and changes of research profiles of institutes. For these applications the predictive model can be applied to all institutes in the dataset, including the small institutes with less than 20 publications. This means that we have a set of 2775 European research institutes assigned to one of these eight groups. Table 5 gives the distribution of all institutes over these groups. The shares differ from Table 3.

Table 5. The distribution of all institutes over clusters

Group	Code	Share
Group 1 (Biology)	BIO	7.7%
Group 2 (Agriculture)	AGR	5.4%
Group 3 (Multidisciplinary)	MDS	26.1%
Group 4 (Geo- & Space Science)	GSS	3.2%
Group 5 (Technical & Natural)	TNS	14.1%
Group 6 (Chemistry)	CHE	6.1%
Group 7 (General & Internal Med.)	GRM	12.3%
Group 8 (Non-internal Med. Spec.)	SPM	25.1%

National comparison

To gain more insight into the research landscape of several European countries a comparison is made between the national profile and the distribution of institutes over the 8 different groups. We use the *Activity Index* (see, [FRAME, 1977]) to indicate whether a certain field has a lower or higher share in the country's total than it has in the European total. Thus in our analysis the European profile is used as reference instead of the world profile. This indicator is also used to compare the distribution of institutes over the eight groups within a specific country with the overall shares within Europe.

A value of 1.0 means that the share of this field/group in one country is equal to reference standard, that is, the corresponding share in Europe. A value less/greater than 1.0 indicates a lower/higher than average share representation. A value of 0 indicates a total absence of the field or group.

Figures 3a and 3b present national publication profiles and cluster representation for the European countries. The EU15 reference standard is indicated as a solid line, and has the form of a regular dodecagon (publication profile) and octagon (cluster representation), respectively. The cross-European comparison of *cluster size* and *composition* reflects national research preferences which are expected to by and large correspond to the national publication profiles (see, for instance, [GLÄNZEL, 2001]). Nonetheless, the extremely high cluster representation in several countries is worth mentioning (CHE in Belgium, BIO in Portugal, GRM in Italy, GSS in Spain, etc.), although some of these phenomena such as the presence of earth and space science institutes in Spain are not at all striking unexpectedly.

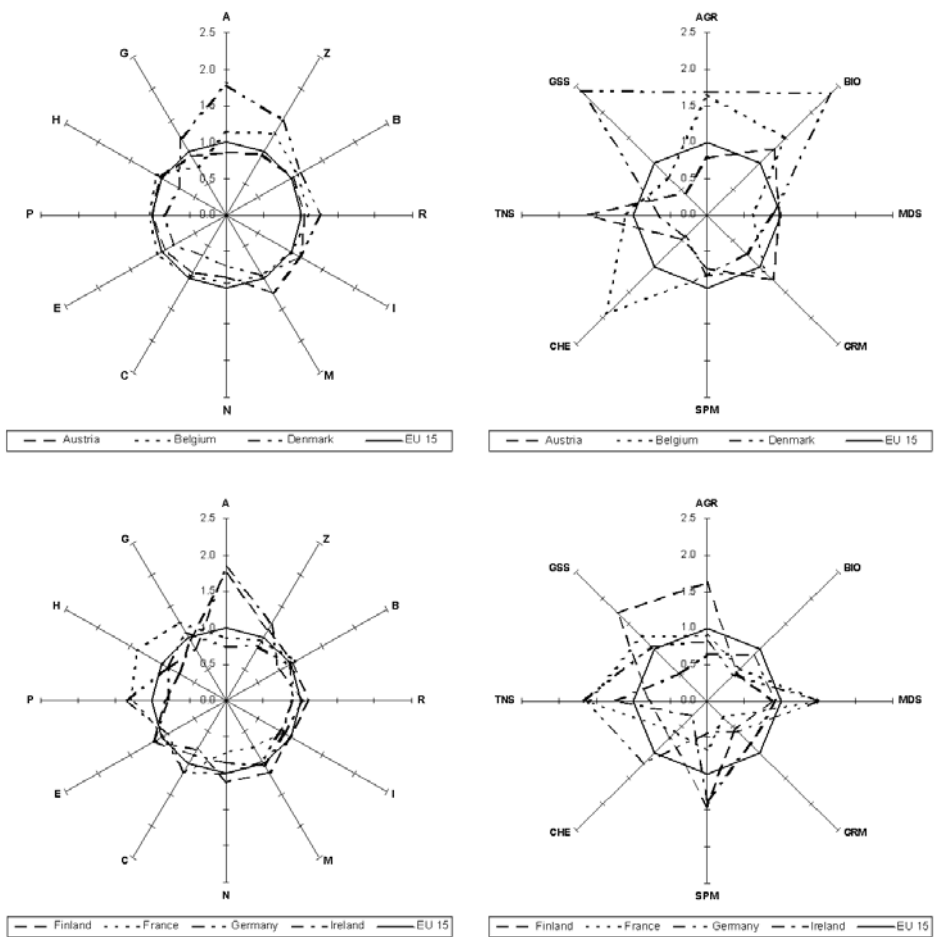


Figure 3. a) National publication profile (left) and cluster representation (right) for the European countries (first group)

It is also interesting to see that in some countries there is a discrepancy between the national profile and the cluster profile especially for some specific fields. In Belgium there is a high share of chemical institutes while the number of publications in the national total is what we expected for the European reference. The situation in Portugal is the complete opposite. The share of chemical papers is way above the standard while the number of chemical institutes is rather low. A further investigation of the data showed that groups that are over represented in Portugal (BIO, MDS and TNS) produce most of these chemical papers.

In Spain we find a striking discrepancy between the share of papers in geo and spaces sciences and the number of institutes. The share of institutes really peaks while the share of publications remains around the reference of Europe.

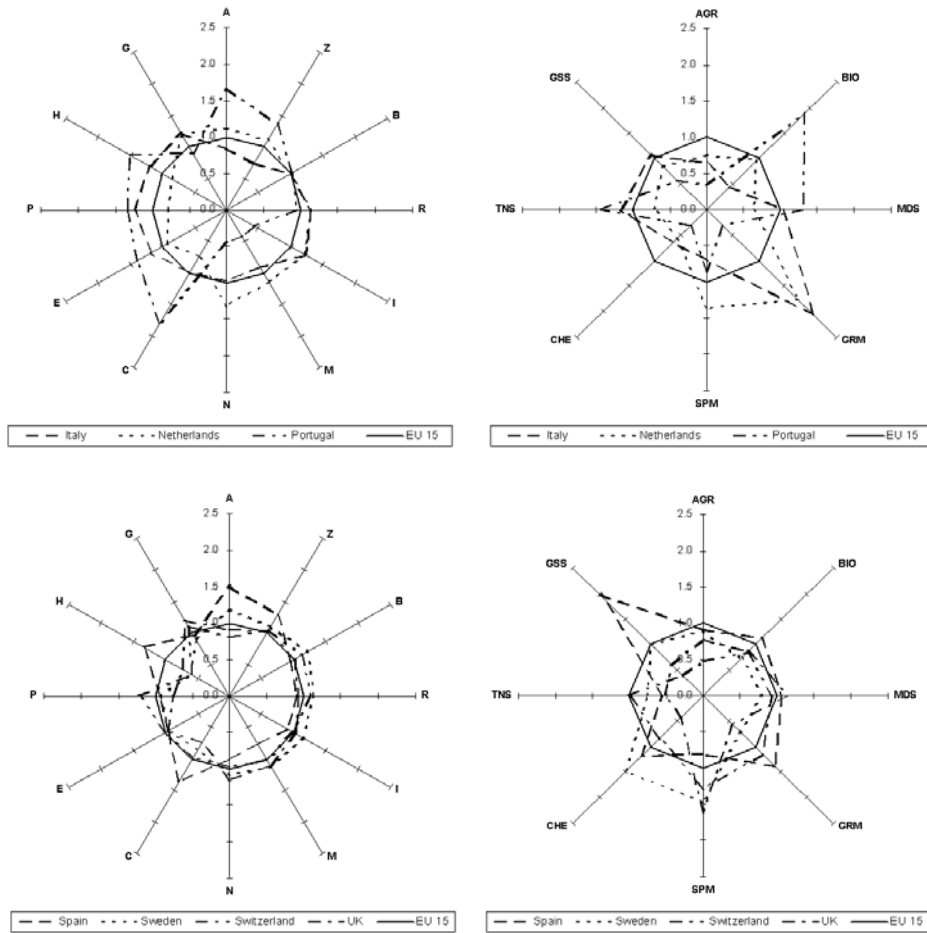


Figure 3. b) National publication profile (left) and cluster representation (right) for the European countries (second group)

Profile dynamics

Our proposed grouping also allows a dynamic look at the institutional subject clustering. In order to study the evolution of both clusters and institutional profiles, we compare the classification of institutes based on their research profile of 1992–1994

with their corresponding classification one decade later, based on 2001–2003 publications. Table 6 presents the evolution within each group.

Table 6. Cluster dynamics based on changes of cluster representation and composition in time (in column percentage)

		Group membership 1992–1994							
		BIO	AGR	MDS	GSS	TNS	CHE	GRM	SPM
2001–2003	BIO	70.2%	14.7%	2.0%	1.5%	0.3%	2.3%	2.0%	1.0%
	AGR	13.0%	70.5%	0.6%	0.0%	1.3%	1.5%	0.0%	0.8%
	MDS	9.9%	7.4%	78.1%	1.5%	10.4%	16.0%	4.3%	6.4%
	GSS	0.6%	1.1%	0.6%	89.6%	1.3%	0.0%	0.0%	0.0%
	TNS	2.5%	2.1%	4.7%	7.5%	81.4%	13.0%	0.0%	0.0%
	CHE	0.0%	1.1%	2.4%	0.0%	3.9%	66.4%	0.3%	0.0%
	GRM	0.6%	2.1%	3.6%	0.0%	0.0%	0.0%	55.5%	11.8%
	SPM	3.1%	1.1%	7.9%	0.0%	1.3%	0.8%	37.9%	80.0%

Overall shows this table a relatively high stability of the institutes. According to our expectations, the main diagonal is dominant. In the first five clusters and the last one, around 3/4 of all institutes or more remain in their own clusters. Here migration is relatively low. Only the clusters BIO and AGR form a certain exception; migration between these clusters is quite considerable. Since these clusters have been found ‘complementary’, this inter-cluster migration can be considered a normal phenomenon. The same applies to clusters GRM and SPM. However, the evolution from natural and technical sciences (CHE and TSN) towards multidisciplinary is in any case worth mentioning.

Conclusions

The main findings of this study are that it is possible to group research institute on the basis of their research profile and that this clustering proved to yield stable and valid results. From the statistical viewpoint it shows that a valid classification model can be constructed to assign each institute to one of eight groups. This first study aimed at laying the groundwork for various possible applications which promise a much deeper insight in the institutional structure of the national research landscape.

The first potential application is creating the methodological framework for cross-institutional comparative studies. The evaluative analysis of institutional research performance, that is, of publication activity and citation impact should preferably be conducted on basis of institutes with likewise research profiles; otherwise, outcomes of cross-institutional studies might be biased or distorted, even if a breakdown by subject disciplines is used.

The second important application is the dynamic look at institutional publication profiles in the context of national research policies, as has, for instance, been done in our meso study on Brazilian science [LETA & AL., 2006]. The analysis of cluster dynamics from the sectoral perspective, for instance, in the context of the Triple Helix model seems also to be a promising extension of this field of applications.

The structural analysis of collaboration between European research institutes based on their research profiles forms the topic of a third application. The investigation of inter-cluster collaboration vs. intra cluster co-operation is already the task of an ongoing project of the authors.

References

- CALINSKI, T., HARABASZ, J. (1974), A dendrite method for cluster analysis. *Communications in Statistics*, 3 : 1–27.
- DUDA, R. O., HART, P. E. (1973), *Pattern Classification and Scene Analysis*. Wiley, New York.
- FRAME, J. D. (1977), Mainstream research in Latin America and the Caribbean, *Interciencia*, 2 : 143–148.
- GLÄNZEL, W. (2001), National characteristics in international scientific co-authorship, *Scientometrics*, 51 (1) : 69–115.
- GLÄNZEL, W., SCHUBERT, A. (2003), A new classification scheme of science fields and subfields designed for scientometric evaluation purposes, *Scientometrics*, 56 (3) : 357–367.
- LETA, J., GLÄNZEL, W., THIJIS, B. (2006), Science in Brazil. Part 2: Sectoral and institutional research profiles. *Scientometrics*, 67 (1) : 87–105.
- MILLIGAN, G. W., COOPER, M. C. (1985), An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50 (2) : 159–179.
- THIJIS, B., GLÄNZEL, W. (2006), The influence of author self-citations on bibliometric meso-indicators. The case of European universities. *Scientometrics*, 66 (1) : 71–80.
- VAN RAAN, A. F. J. (2004), Measuring Science, In: H. F. M. MOED, W. GLÄNZEL, U. SCHMOCH (Eds), *Handbook of Quantitative science and Technology Research. The Use of Publication and Patent Statistics in Studies on S&T Systems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 19–50.

Appendix 1
Classification function coefficients

Field	Clusters							
	1	2	3	4	5	6	7	8
A	80.392	157.137	63.721	50.982	60.717	42.628	62.327	58.544
B	76.428	65.537	86.906	79.840	83.739	65.842	64.559	68.854
C	68.028	64.657	82.179	71.591	93.754	136.314	68.402	67.736
E	63.922	59.828	75.611	66.336	107.189	73.140	63.524	62.255
G	93.895	79.930	95.353	265.486	101.958	79.934	85.021	84.046
H	101.009	90.576	115.203	92.096	111.865	89.592	102.683	104.077
I	77.836	74.906	79.347	73.811	76.422	67.554	122.417	89.208
M	76.790	72.208	80.497	74.774	77.901	68.249	85.606	105.088
N	77.809	75.691	87.126	72.872	83.133	73.691	62.352	54.114
O	9.170	-2.635	2.380	20.746	-7.836	19.198	-1.667	-3.380
P	78.023	81.745	86.944	76.454	112.264	68.354	74.318	73.943
R	51.762	56.758	55.786	55.339	52.288	28.461	41.944	50.260
S	47.896	23.438	49.032	-4.143	25.437	31.409	55.385	43.914
U	59.711	60.442	63.865	84.365	61.233	101.459	80.461	88.880
Z	130.174	81.579	73.427	59.012	66.220	57.191	66.856	67.872
Constant	-69.985	-81.618	-53.802	-127.005	-68.180	-72.677	-63.923	-56.998

Fisher's linear discriminant functions