Mapping excellence in the geography of the sciences:

An approach made possible by using Scopus data

Lutz Bornmann$, Loet Leydesdorff§, Christiane Walch-Solimena$, Christoph Ettl$

$ Max Planck Society, Hofgartenstr. 8, D-80539 Munich, Germany.
§ Amsterdam School of Communications Research, University of Amsterdam, Kloveniersburgwal 48, NL-1012 CX Amsterdam, The Netherlands.

Corresponding author: Lutz Bornmann, bornmann@gv.mpg.de

Abstract

As research becomes an ever more globalized activity, there is growing interest in national and international comparisons of standards and quality in different countries and regions. A sign for this trend is the increasing interest in rankings of universities according to their research performance, both inside but also outside the scientific environment. New methods presented in this paper, enable us to map centers of excellence around the world using programs that are freely available. Based on Scopus data, field-specific excellence can be identified and agglomerated in regions and cities where recently highly-cited papers were published. Differences in performance rates can be flagged on the map using colors and sizes of the marks.

Key words

Scientific excellence; Highly-cited papers; Geographic mapping; Spatial scientometrics

1   Introduction

Bibliometric measures have led to an increased interest among science policy-makers for the identification of centers of excellence in scientific research (see here Danell, 2011; Frenken, Hardeman, & Hoekman, 2009). How can excellence be evaluated in geographic regions and cities? In this study, we attend to this interest and describe new methods to analyze the geographic distribution of scientific excellence. These methods allow, among other things, to visualize regions (and cities within them) that are characterized by a high number of authors having published highly-cited papers. Differently from institutional rankings published thus far e.g. by SCImago Reseach Group (2010) or by Noyons et al. (2003a, 2003b), our methods allow for the spatial identification of scientific excellence by having online access to the Scopus (Elsevier, Amsterdam, the Netherlands) Web interface. We use programs that are freely available on the Internet and easy to handle. The goal of our methods is to produce regional maps showing where excellent papers have emerged and where these papers have frequently occurred. Frenken et al. (2009) suggested to group these methods for mapping the geography of science under the heading of spatial scientometrics.

The geographic mapping of scientific papers can be distinguished from its cognitive mapping (see here Börner, Chen, & Boyack, 2003). In a literature overview, Frenken et al. (2009) identified some descriptive studies which investigated differences among regions or countries, respectively, in terms of their publication output and citations. Most of these studies investigated the spatial distribution of published papers in terms of output (publications) and not impact (using citations). An early study of Matthiessen and Schwarz (1999) researched the scientific output of large European cities. As a result it was found that the north-western part of Europe showed the highest concentration of research output. The urban regions of London, Paris, Moscow and the

Dutch agglomeration (Amsterdam-Hague-Rotterdam-Utrecht) form a "super-league" in Western Europe (to use a terminology from sports).

Only a few studies focused on citation impact (instead of or in addition to publication output) (Frenken, et al., 2009). Batty (2003), for example, analyzed data from the ISI's HighlyCited database (http://www.isihighlycited.com) which consisted of approximately the top 100 most cited individuals in fourteen scientific fields. Based on their first study (Matthiessen & Schwarz, 1999) which focused on publication output of large cities in Europe, Matthiessen, Schwarz, and Find (2002) followed up with an analysis of the strength, interrelations, and nodality of global research centers. In a recently published paper Matthiessen, Schwarz, and Find (2010) asked whether the next generation of hot spots in research of certain disciplines can be predicted. Focusing on the largest cities of the world and the geographic regions around these cities, these authors found "in accordance with other presented findings an extremely high and growing degree of concentration in the large centres, in combination with a series of new centres of concentration" (Matthiessen, et al., 2010, p. 1895).

Spatial bibliometrics has attracted a lot of attention in the most recent past. In *Nature News* Van Noorden (2010) discussed urban regions producing the best research and whether their success could be replicated elsewhere (Florida, 2002; Saxenian, 1996). Living Science (http://www.livingscience.ethz.ch/), created by Luis Bettencourt (Los Alamos National Laboratory in New Mexico) and collaborators under Dirk Helbing at the Swiss Federal Institute of Technology, Zurich, track where arXiv papers are published in real time. In general it pays off for the sciences within a country to identify (by visualization methods) and expand regional centers of excellence (with specific financial support). As a rule, there is a high probability of co-operation between scientists working at a short physical distance. Synergies between ideas and

direct face-to-face communication between scientists are major factors of productivity (Matthiessen, et al., 2002; Wagner, 2008).

2   Methods

Percentile citation impact classes are well suited for identifying the highly cited papers in a scientific field (Bornmann & Mutz, 2011). However, in evaluative bibliometrics there is uncertainty regarding to what percentile rank a paper would have to be considered as highly cited. According to Tijssen, Visser, and van Leeuwen (2002) and Tijssen and van Leeuwen (2006), highly cited papers are those among the top 10% of the most cited papers – that is, papers in or greater than the 90th percentile of a field (see also Lewison, Thornicroft, Szmukler, & Tansella, 2007). In the Essential Science Indicators (ESI) Thomson Reuters classifies as highly cited papers those that belong to the top 1% of papers worldwide (papers in or greater than the 99th percentile), taking into account the field and year of publication. According to the National Science Board (2010) top-performance or highly cited papers are those in the top 1%, too. In the present study we follow the classification of Thomson Reuters and the National Science Board (2010), since the methods proposed here can process no more than 2000 papers due to a systems limit of the Scopus data base. Scopus is used as data base for this study since this is currently the only database in which one can select all papers published in a broader field (e.g., physics & astronomy). A study using the top 10% of the most highly cited papers in a scientific field would exceed this limit in most of the fields. In this study, we focus on the top 1% of papers published in 2007 with a fixed citation window of three years (from 2007 up to the date of research, at the end of 2010).

In the following, the procedure to map the excellent papers (more precisely: the cities of the authors having published the top 1% most highly cited papers) in a certain field is described.

The procedure will be explained for the field of "neuroscience". With the search string "subjarea(neur) and pubyear is 2007 and doctype(ar)" in the advanced search field of Scopus all papers with the document type "article" are retrieved which were published in 2007 within the Scopus journal set of "neuroscience." At December 1, 2010 this search resulted in 40,086 papers (see Table 1) The search was restricted to articles (as document types) since (1) the method proposed here is intended to identify excellence at the research front and (2) different document types have different expected citation rates, possibly resulting in non-comparable datasets.

Table 1. Data used for the maps

Field	Published articles in 2007	Citation limit for the top 1% most highly cited articles	Number of articles belonging to the top 1% most highly cited articles
Neuroscience	40,086	62	405
Physics and astronomy	146,029	40	1496
Social sciences	76,441	22	758
Nature and *Science* articles (Scopus)	1,596		
Nature and *Science* articles (Web of Science)	1,604		

By sorting the search results by citation counts in decreasing order (citation window: from 2007 to the date of search), the 1% of papers at the top of the Scopus list can be marked. At the date of search, 405 papers with at least 62 citations each (gathered between 2007 and the date of search) were marked as the list of the top 1% neuroscience papers. Once these papers (n=405)[1] have been added to a temporary list in Scopus, the area "Refine Results" on the screen shows the affiliations (research institutions) of the authors in descending order. At the top of the list are

---

[1] One percent of 40,086 is 401. However, we included ranks which were tied at this 1% level and thus retrieved 405 records.

those institutions with the highest number of occurrences among the affiliations of the top cited papers.

The selected documents (in the example n=405 papers) are exported by choosing the export format "Comma separated file, .csv (e.g. Excel)" and the output "Complete format." The download in the .csv format must be preprocessed with the program scop2isi.exe. This and the programs mentioned below including the respective user instructions can be downloaded from http://www.leydesdorff.net/maps/ (Leydesdorff & Persson, 2010). The programs and the .csv file must be stored in the same folder. The resulting output file of the program "isi.txt" has to be renamed into "data.txt." This file has to be stored in the same folder as the program cities1.exe. After running, cities1.exe will prompt the user with four questions: with the first and second questions one can set a threshold in terms of a minimal percentage of the total set of city-names in the data or set a minimum number of occurrences. These default answers to the questions ("0") can be followed. The third and fourth questions enable the user to obtain a cosine-normalized data matrix and to generate network data. Both questions should be answered with "N" (meaning: no).

The program cities1.exe creates among other files cities.txt. This file contains all city entries from the top 1% papers downloaded from Scopus (see here Costas & Iribarren-Maestro, 2007). If there is more than one co-author of a publication with an identical address, this leads to a single address (or a single city occurrence) in cities.txt. If the scientists are affiliated with different departments within the same institution, this leads to two addresses or two city occurrences, respectively. Occasionally it happens by using cities1.exe that some erroneous city entries appear at the beginning of cities.txt (for neuroscience there were 18 errors in the file). These entries start with a comma in the line (e.g. ", Mersey Community Forest"). The errors result from technical inconsistencies in the data formats of Scopus. The erroneous entries in

cities.txt have to be deleted, but the number of lines with erroneous entries must be noted. The (corrected) content of cities.txt can then be copied-and-pasted into the GPS encoder at http://www.gpsvisualizer.com/geocoder/. Since no more than 1000 entries can be processed by the encoder, more than 1000 entries in cities.txt must be entered into the encoder in subsequent steps.

After saving the results in the output window of the geo-encoder as a DOS text file (e.g. geo.txt) this data serves as input for cities2.exe. In case of erroneous entries in cities.txt, as many dummy lines with "0,0," ",-," should be added as erroneous lines which had been removed in cities.txt, after the line "latitude, longitude, name, desc, color" in geo.txt. If geo.txt contains all entries from cities.txt with the additional geo data, the program cities2.exe can be used. First of all the program prompts for the name of this output file (here: geo.txt). cities2.exe produces a number of output files in various formats within the folder. If cities2.exe is finished cities3.exe can be used as a final step. This program does not need any user input; it aggregates similar city names with somewhat different geocodes because of different institutional addresses. Furthermore, this program—specifically developed for this project—makes sure that within each map, the circles are coloured according to the number of authors of excellent papers at single cities. The colorization supports the visualisation of the different numbers by different circle radii. With this feature the viewer of a map may realise faster those cities with the highest (and lowest) numbers. We used the percentile rank approach proposed by Bornmann and Mutz (2011) to colorize the circles.

The <u>percentiles</u> were computed as follows: First, the numbers of authors (of excellent papers) X_i for the ith city within n cities (of one map) were ranked in decreasing order

$$X_1 \geq X_2 \geq ... \geq X_n,$$

where X_1 (X_n) denotes the number of papers associated with city names or, in other words, with the largest (lowest) number of author addresses. Secondly, each city is assigned a percentile rank based on this distribution. If, for example, a single city acquires 50 papers whereas 90% of the other cities have 49 papers or less, then this particular city would be in the 90th percentile.

All cities (circles on a single map) are categorized into six percentile rank classes and coloured accordingly:

Blue circle: bottom 50th (cities with a percentile less than the 50th percentile),

Cyan circle: 50th – 75th (cities within the [50th; 75th[percentile interval),

Orange circle: 75th – 90th (cities within the [75th; 90th[percentile interval),

Pink circle: 90th – 95th (cities within the [90th; 95th[percentile interval),

Fuchsia circle: 95th – 99th (cities within the [95th; 99th[percentile interval),

Red circle: top 1% (cities with a percentile equal to or greater than the 99th percentile).

"ucities.txt," that is, the output file of cities3.exe, can be uploaded into the GPS Visualizer at http://www.gpsvisualizer.com/map_input?form=data. The Web page offers a number of parameters that can be set to visualize the data in ucities.txt. The following parameters should be changed: (a) "waypoints" into "default;" (b) "colorize using this field" into "custom field" and choose "color" in this field; (c) "resize using this field" into "custom field" and (d) in "custom resizing field" "n" is written.

When the GPS data has been processed, the Google map is displayed in a small frame, but it is also temporarily available to view on the full screen. The map shows the regional distribution

of the authors of highly cited papers. The background map's opacity can be adjusted or another

layout as available in Google or Yahoo! can be chosen. With the instruments visualized on the

left side of the map it is possible to zoom into the map. At the beginning, the global map is

shown. For the maps presented in the following we zoomed into Europe in order to generate

comparable maps for different publication sets. Similarly, other regional foci can be chosen. To

determine the number of papers for a specific city, one can click on the respective city. Maps

generated in this way can be copied to other programs (like Microsoft Word) by using programs

utilized for screen shots (e.g., Hardcopy). If one uses the download instead of the view command

shown in the Google Maps output page, a html-coded page is saved that includes the data of

ucities.txt. Opening this page within a browser will regenerate the respective Google Map.

There are several problems inherent to the approach proposed here. The user should

always be aware of these limitations when the approach is applied:

1) City name variants (e.g., Zurich and Zrich) in ucities.txt may result in circles that are

   positioned on one another although they should be combined in one bigger circle.

   cities3.exe is intended as an error-correction mechanism to reduce this problem;

   however, it cannot completely be solved using Scopus data. A data cleaning procedure

   is not provided in our approach automatically, but this can be done by the user

   manually.

2) The methods as described above do not allow for the identification of research

   institutions on the map where the authors of the excellent papers are located.[2]

3) As described above the use of cities1.exe with data.txt (created with Scopus data)

   results in a relatively small number of erroneous entries in the output file cities.txt.

   These entries must be deleted before further processing and are therefore missing in the

---

[2] Leydesdorff & Persson (2010) offer a set of similar routines for institutional addresses, but the equivalent of cities3.exe was not yet developed.

final visualization. Particularly, Scopus data from the social sciences and arts & humanities are affected by such erroneous entries.

4) Since it can be assumed for certain fields (e.g., the life sciences) that authors at specific positions in the list of authors (e.g., the first or the last authors) have made particularly significant contributions to a publication it would be interesting to include in the analysis a restricted set of authors (e.g. only the first authors). However, this is not possible with our approach.

5) If there are a high number of publications visualized on the map for one single city two effects could be responsible: (a) Many scientists located in that city (i.e., scientists at different institutions or departments within one institution) produced at least one excellent paper or (b) one or only a few scientists located in this city produced many influential papers. With our approach – assuming cities as units of analysis – we are not able to distinguish between these two interpretations, but this could be done by studying the Scopus search results after the data has been refined to institutions.

6) Because of a systems limit in Scopus only 2,000 papers can be retrieved and downloaded from the Web interface. Thus, fields with more than 2,000 papers among the top 1% – that is, above a total of 200,000 – cannot be visualized using the approach presented here. One may in this case wish to set higher thresholds and study, for example, the 1‰ most highly cited papers.

3   Results

To demonstrate the proposed method we produced three field-specific maps. Figure 1 shows the location of authors in Europe having published highly cited papers in neuroscience. The map is based on the top 1% of articles published in 2007 in a journal of the Scopus journal

set neuroscience (see above and Table 1). On the map it is visible in which regions the authors are located (the circles with different colors on the map) and the frequency of occurrences (author addresses of the papers) per location: (1) the radius of the circles is proportionate to the logarithm of the frequency; (2) the colors of the circles accord with the percentile rank classes of the cities indicated.
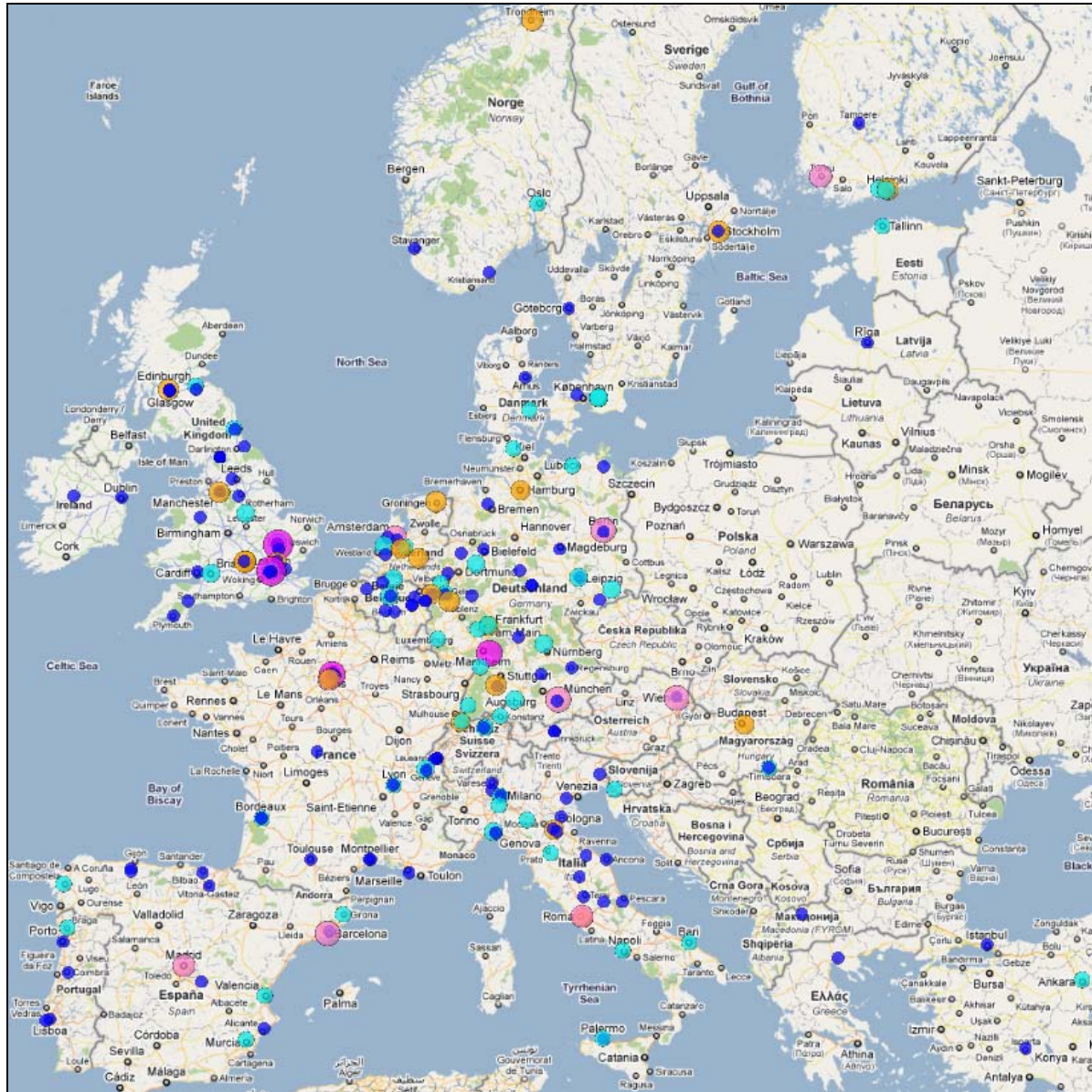

Figure 1. Locations of authors in Europe having published highly cited neuroscience papers in 2007 (this figure appears in colour on the Web (PDF and HTML of this paper), but is not reproduced in colour in the printed version);
http://www.leydesdorff.net/mapping_excellence/figure1.htm

11

For example, when one clicks on the blue circle for Aarhus in Denmark the occurrence of one single author address (or paper) is indicated; this city is classified in this map as belonging to the bottom 50th. For Vienna in Austria, the larger pink circle reflects the publication of 12 papers. Given this number Vienna belongs to the top 10% of the cities worldwide (cities with a percentile equal to or greater than the 90th percentile) locating the most authors of excellent neuroscience papers. If one focuses on geographic regions with a high concentration of excellent papers, these regions are as a rule situated around larger cities (e.g, Vienna, Paris, Munich or London). According to Matthiessen et al. (2010) "cities are almost always also centres for a hinterland, which they more or less dominate" (p. 1880). On the map, a couple of regions with a higher density of circles around a centre are visible; two regions are very salient (1) London – Cambridge – Oxford and (2) Amsterdam – Nijmegen – Rotterdam. A more in-depth search (see the description in the method section) in Scopus shows that the European institution with the largest numbers of authors is located within the first region: the UCL Institute of Neurology in London.
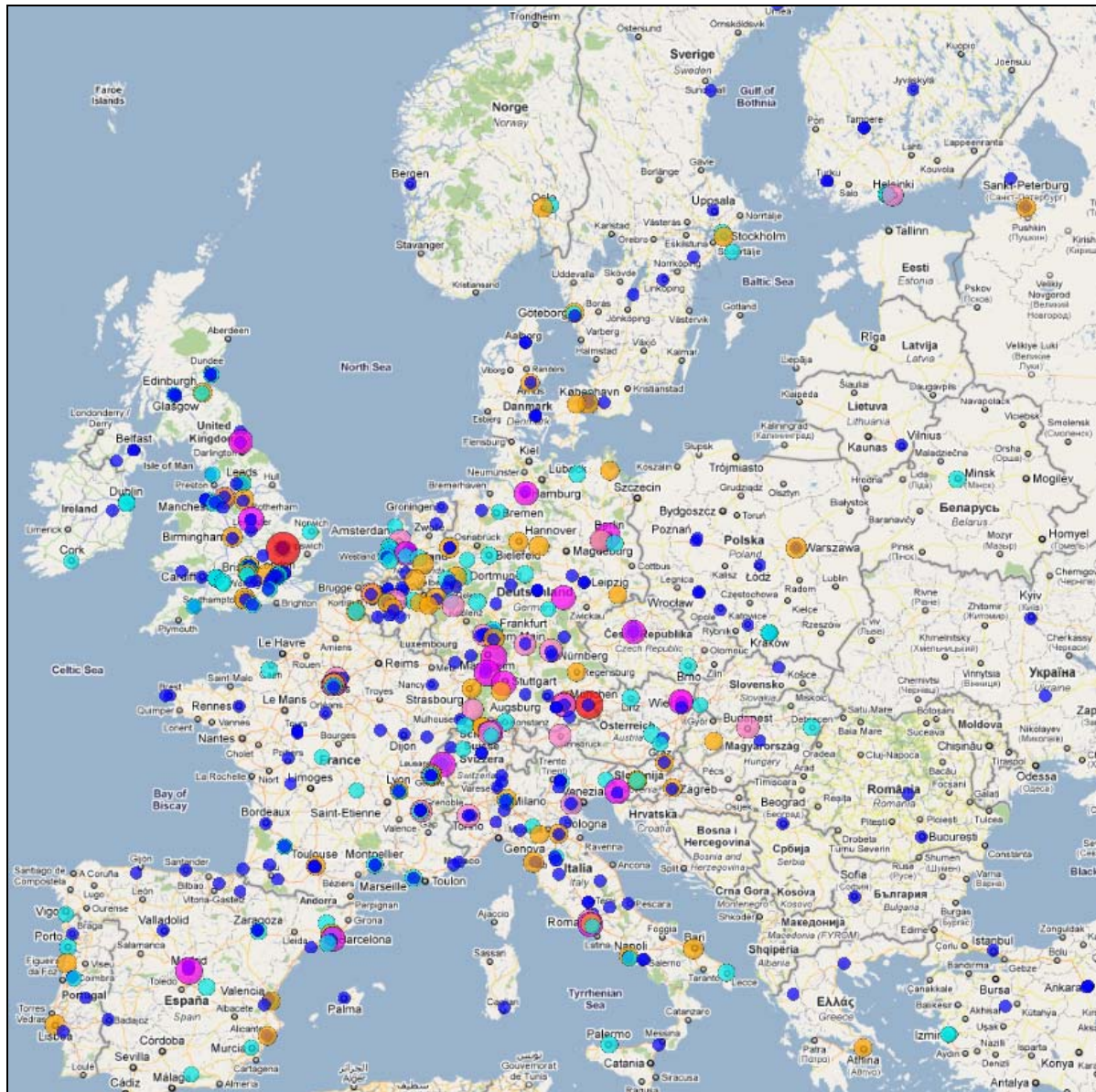
Figure 2. Locations of authors in Europe having published highly cited physics and astronomy papers in 2007 (this figure appears in colour on the Web (PDF and HTML of this paper), but is not reproduced in colour in the printed version);
http://www.leydesdorff.net/mapping_excellence/figure2.htm

Figure 2 shows the corresponding map for physics and astronomy (see Table 1). All in all

146,029 articles have been published in 2007 in this journal set worldwide; the top 1% are those

1496 papers which received at least 40 citations each between 2007 and the date of research

(December 1, 2010). Here again, we see a higher density of circles around London – Cambridge –

Oxford with a red circle for Cambridge. Cambridge belongs to the top 1% of the cities worldwide with the highest numbers of authors of excellent papers. Our results for the two fields neuroscience and physics & astronomy with Cambridge as one of the few European top cities in science worldwide is in agreement with the findings presented by Van Noorden (2010) over all fields.

Since physics & astronomy and neuroscience are natural science fields it is interesting to see how the European map changes in the case of social sciences. Figure 3 reveals the map for the authors of 758 top cited papers (selected from 76,441 articles published in 2007) which received at least 22 citations each between 2007 and the date of research (November 26, 2010) (see Table 1). It is clearly visible that German cities are significantly less frequently locations of highly cited authors in social sciences than in neuroscience and physics & astronomy. For Europe, there are two red circles on the map: Cambridge and London are among the top 1% of the cities worldwide with the highest numbers of excellent papers. Switzerland (Zürich), the Netherlands (Amsterdam and Rotterdam), and Spain (Madrid) can be identified in the next rank (top 5%) using this map.
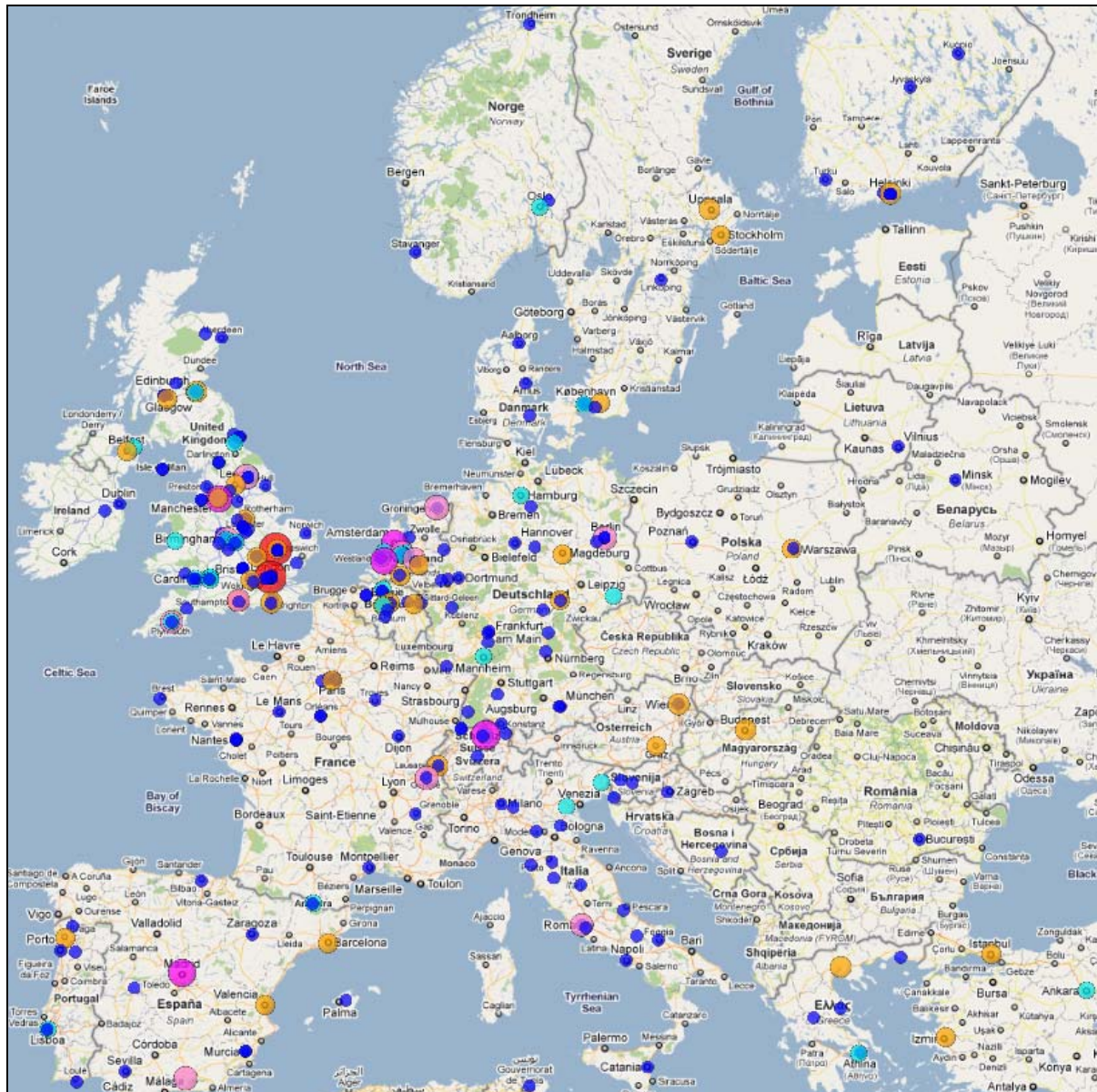
Figure 3. Locations of authors in Europe having published highly cited social sciences papers in 2007 (this figure appears in colour on the Web (PDF and HTML of this paper), but is not reproduced in colour in the printed version);
http://www.leydesdorff.net/mapping_excellence/figure3.htm

As a final step in the analyses for this paper we left the focus on specific fields and present

the results of a fields-overlapping analysis of excellent papers. For this, we downloaded from

Scopus the bibliographic data of all articles published in 2007 in the high-impact journals *Nature*

and *Science* (see Table 1). This analysis is inspired by the maps of Luis Bettencourt and Jasleen

Kaur (Indiana University in Bloomington) published on

www.nature.com/news/specials/cities/best-cities.html. The authors analyzed city addresses

appearing in *Science*, *Nature*, and *Proceedings of the National Academy of Sciences* in 1989,

1999, and 2009.

We downloaded the bibliographic data of all 1,596 articles published in 2007 (date of

research: October 27, 2010). Since the analysis of the data with cities1.exe produced some

erroneous entries (n=36) in cities.txt (see the description in the methods section), we repeated the

same analysis with Web of Science (WoS, Thomson Reuters) data. The WoS data did not

produce any error in cities1.exe. From WoS we downloaded 1,604 *Nature* and *Science* articles

that have been published in 2007 (date of research in WoS: October 27, 2010). The difference

between the WoS and Scopus in the numbers of articles can probably be accounted for by the fact

that publications are (sometimes) differently categorized as a certain document type (here: article)
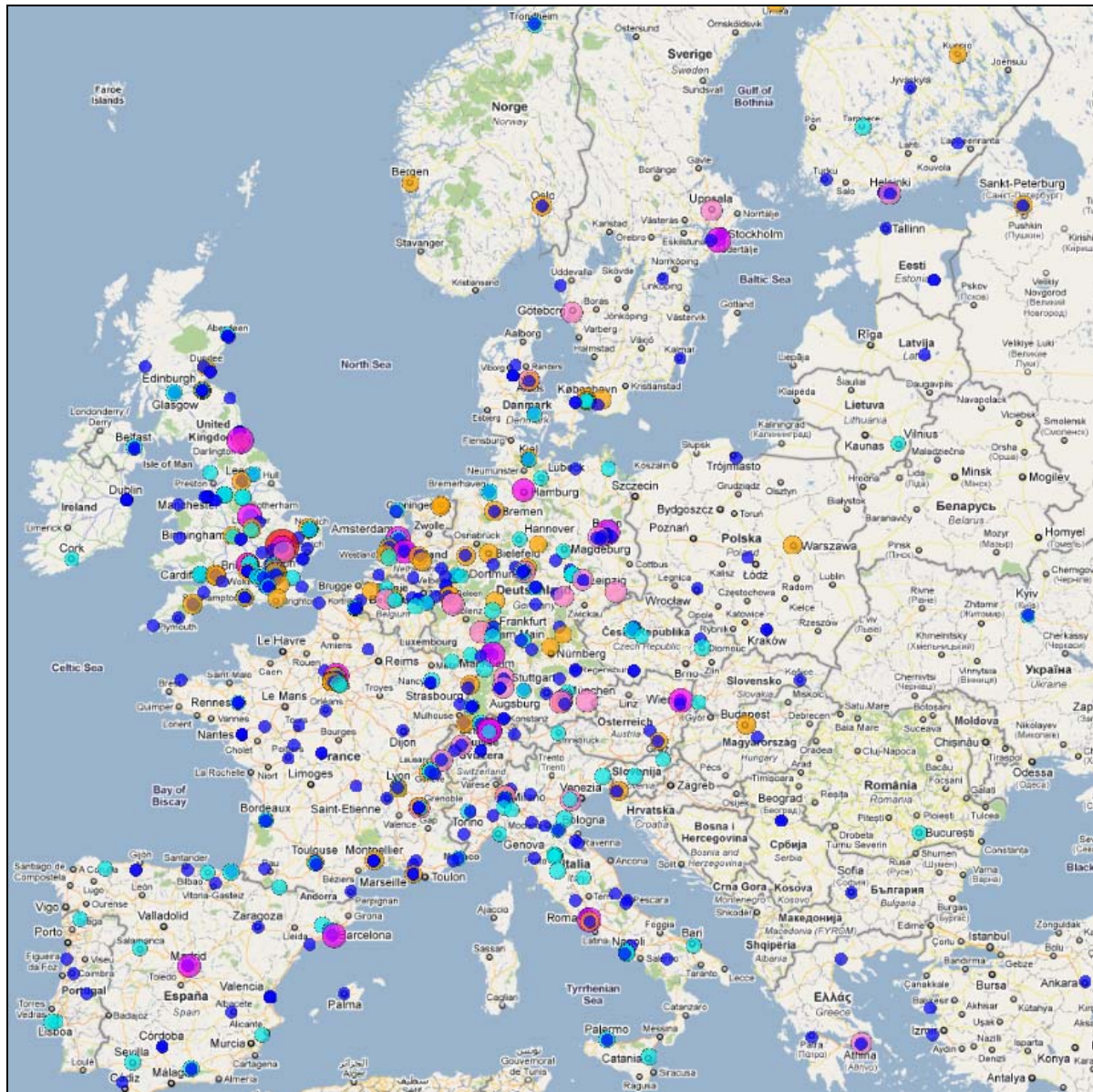
in both data bases (Leydesdorff & Opthof, 2011).

Figure 4. Locations of authors in Europe having published *Nature* or *Science* articles in 2007
(searched in Scopus) (this figure appears in colour on the Web (PDF and HTML of this paper),
but is not reproduced in colour in the printed version);
http://www.leydesdorff.net/mapping_excellence/figure4.htm

Figure 5. Locations of authors in Europe having published *Nature* or *Science* articles in 2007 (searched in Web of Science) (this figure appears in colour on the Web (PDF and HTML of this paper), but is not reproduced in colour in the printed version); http://www.leydesdorff.net/mapping_excellence/figure5.htm

The maps based on the Scopus and WoS data are presented in Figure 4 and Figure 5, respectively. The results are very similar. In agreement to the field-specific maps presented above, higher densities of circles in the neighborhoods of (1) London – Cambridge – Oxford and (2) Amsterdam – Nijmegen – Rotterdam are visible in both figures. The differences between the

maps based on the two data bases are especially due to technical inconsistencies in the data formats of Scopus but not in Web of Science. As a rule this leads to more and other colored circles around one city (see, e.g., Berlin) using the Scopus set than on the WoS based generated map.

4   Discussion

As research becomes ever more a globalized activity, there is growing interest in national and international comparisons of standards and quality in different countries and regions. A sign for this trend is increasing interest in rankings of universities according to their research prowess, both inside but also outside the scientific environment. The methods presented in this paper allow for an analysis revealing centers of excellence around the world using programs that are freely available. Based on Scopus data, field-specific excellence can be identified in regions and cities where recently highly-cited papers were published.

As we could show the spatial concentration of scientific activity in Europe is remarkable. All analyses point out, e.g., that the region London – Cambridge – Oxford is characterized by a high number of excellent research output – measured by top-cited papers or papers published in *Nature* or *Science*. Against the backdrop of this finding and the results of other studies (Frenken, et al., 2009) we propose to introduce a name for this spatial concentration phenomenon in science: the reverse N-effect. The formulation of the N-effect goes back to Garcia and Tor (2009). Their social-psychological research revealed that more competitors doing the same task produce less competition because of motivation-losses. In science, we can assume a reverse N-effect: More competitors (here: prolific scientists) working within the same region produce better results. As this paper shows, the better result may consist of a higher output of excellent papers; the findings of Lee, Brownstein, Mills, and Kohane (2010) provide evidence for the role of the

reverse N-effect as a predictor of the citation impact of collaborations. Due to a systematic spatial bias in the interaction prolific researchers thus favor cooperation with other prolific researchers located in physical closeness.

According to Frenken et al. (2009) "at least three mechanisms may explain why interactions in science are spatially biased towards physically proximate actors. First, serendipitous encounters are more likely when two actors are in close vicinity of each other. Second, the need for face-to-face interaction when engaging in interactions comes at a cost, which increases as a function of travel time. Third, 'the rules of the game' that matter for scientific knowledge production (e.g. funding, labour market regimes, intellectual property right regimes, languages) are spatially differentiated and constrain interaction between institutional frameworks, in particular, between nation-states" (p. 224).

With regards to the reverse N-effect in science it would be interesting to study whether an accumulation effect over time is visible at certain regions worldwide. Do regions where more excellent papers are published produce increasingly excellent papers? The literature research of Frenken et al. (2009) found only two such studies. Both studies point out agglomeration advantages to be present though the evidence is preliminary.

Despite the advantages of the approach proposed in this study to map excellence in science, we recognize the limitations inherent to bibliometric data. (1) Publications are among several types of scientific activities. (2) It is not guaranteed that the addresses listed on the publication reflect the locations where the reported research was conducted. (3) The handling of multiple authorships is different in scientific fields. (4) No standard technique exists for the subject classification of articles (see here Bornmann & Daniel, 2008; Bornmann, Mutz, Neuhaus, & Daniel, 2008; Leydesdorff & Rafols, 2009).

References

Batty, M. (2003). The geography of scientific citation. *Environment and Planning A, 35*, 761-770.

Börner, K., Chen, C. M., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*, 179-255.

Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation, 64*(1), 45-80. doi: 10.1108/00220410810844150.

Bornmann, L., & Mutz, R. (2011). Further steps towards an ideal method of measuring citation performance: the avoidance of citation (ratio) averages in field-normalization. *Journal of Inormetrics, 5*(1), 228-230.

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Use of citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics, 8*, 93-102. doi: 10.3354/esep00084.

Costas, R., & Iribarren-Maestro, I. (2007). Variations in content and format of ISI databases in their different versions: The case of the Science Citation index in CD-ROM and the web of science. *Scientometrics, 72*(2), 167-183. doi: 10.1007/s11192-007-1589-z.

Danell, R. (2011). Can the quality of scientific work be predicted using information on the author's track record? *Journal of the American Society for Information Science and Technology, 62*(1), 50-60. doi: 10.1002/asi.21454.

Florida, R. (2002). Bohemia and economic geography. *Journal of Economic Geography, 2*(1), 55-71.

Frenken, K., Hardeman, S., & Hoekman, J. (2009). Spatial scientometrics: towards a cumulative research program. *Journal of Informetrics, 3*(3), 222-232. doi: 10.1016/j.joi.2009.03.005.

Garcia, S. M., & Tor, A. (2009). The n-effect: more competitors, less competition. *Psychological Science, 20*(7), 871-877.

Lee, K., Brownstein, J. S., Mills, R. G., & Kohane, I. S. (2010). Does collocation inform the impact of collaboration? *PLoS ONE, 5*(12), e14279.

Lewison, G., Thornicroft, G., Szmukler, G., & Tansella, M. (2007). Fair assessment of the merits of psychiatric research. *British Journal of Psychiatry, 190*, 314-318. doi: 10.1192/bjp.bp.106.024919.

Leydesdorff, L., & Opthof, T. (2011). Scopus' SNIP indicator: reply to Moed. *Journal of the American Society for Information Science and Technology, 62*(1), 214-215. doi: 10.1002/asi.21405.

Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: distribution patterns and networks of relations among cities and institutes. *Journal of the American Society for Information Science and Technology, 61*(8), 1622-1634. doi: Doi 10.1002/Asi.21347.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology, 60*(2), 348-362. doi: 10.1002/asi.20967.

Matthiessen, C. W., & Schwarz, A. W. (1999). Scientific centres in Europe: an analysis of research strength and patterns of specialisation based on bibliometric indicators. *Urban Studies, 36*(3), 453-477.

Matthiessen, C. W., Schwarz, A. W., & Find, S. (2002). The top-level global research system, 1997-99: centres, networks and nodality. An analysis based on bibliometric indicators. *Urban Studies, 39*(5-6), 903-927. doi: 10.1080/00420980220128372.

Matthiessen, C. W., Schwarz, A. W., & Find, S. (2010). World cities of scientific knowledge: systems, networks and potential dynamics. An analysis based on bibliometric indicators. *Urban Studies, 47*(9), 1879-1897. doi: 10.1177/0042098010372683.

National Science Board. (2010). Science and engineering indicators 2010, appendix tables. Arlington, VA, USA: National Science Foundation (NSB 10-01).

Noyons, E. C. M., Buter, R. K., van Raan, A. F. J., Schmoch, U., Heinze, T., Hinze, S., & Rangnow, R. (2003a). Mapping excellence in science and technology across Europe: life sciences. Leiden, the Netherlands: Centre for Science and Technology Studies (CWTS), Leiden University.

Noyons, E. C. M., Buter, R. K., van Raan, A. F. J., Schmoch, U., Heinze, T., Hinze, S., & Rangnow, R. (2003b). Mapping excellence in science and technology across Europe: nanoscience and nanotechnology. Leiden, the Netherlands: Centre for Science and Technology Studies (CWTS), Leiden University.

Saxenian, A. (1996). *Regional advantage: culture and competition in Silicon Valley and Route 128*. Cambridge, MA, USA: Harvard University Press.

SCImago Reseach Group. (2010). SIR World Report 2010 :: Global Ranking. Granada, Spain: University of Granada.

Tijssen, R., & van Leeuwen, T. (2006). Centres of research excellence and science indicators. Can 'excellence' be captured in numbers? In W. Glänzel (Ed.), *Ninth International Conference on Science and Technology Indicators* (pp. 146-147). Leuven, Belgium: Katholieke Universiteit Leuven.

Tijssen, R., Visser, M., & van Leeuwen, T. (2002). Benchmarking international scientific excellence: are highly cited research papers an appropriate frame of reference? *Scientometrics, 54*(3), 381-397.

van Noorden, R. (2010). Cities: building the best cities for science. Which urban regions produce the best research - and can their success be replicated? *Nature, 467*, 906-908. doi: 10.1038/467906a.

Wagner, C. S. (2008). *The new invisible college*. Washington, DC, USA: Brookings Press.