SCIENTOPOGRAPHY

World maps and charts based on scientometric indicators.

A bird's eye view on the metropolis and beyond

A. Schubert and T. Braun
Information and Scientometrics Research Unit (ISSRU),
Library of the Hungarian Academy of Sciences,
Budapest (Hungary)

Introduction

Topography is defined in the dictionaries as "the science of drawing on maps and charts or otherwise representing the features of a region" or, in a broader sense the "study or description of a region system or entity, showing specific relations of component parts as to shape, size, position, etc."

Scientopography is defined here as the topography of scientometric features (i.e., publication and citation based indicators) of geographical or geopolitical regions. Examples of scientopographical representations include, among others, *proportional maps, proximity diagrams* and *relational charts.*

The topographical approach seems to be particularly suitable to define and redefine such concepts as *center, periphery, metropolis,* etc. Topographic relations are not merely formally defined but, by their very nature, visually represented, thereby they may have a very strong heurisitic power.

Exactly constructed maps are justly thought to reflect the objective facts and nothing but the objective facts. The question, however, arises *which version of the facts* is to be considered (1). Maps are overly sensitive to the *point of view* both in the literal and the figurative sense of the word. By taking a plain photographic view of Europe from 200 km above France and Hungary, we shall have an idea what differences a relatively small deviation in the viewpoint can make. What appears to be central and fundamental in one picture turns to be peripheral and insignificant in the other. This kind of relativity is a feature that must always be taken into account while considering and evaluating topographical information.

Maps as presentation tools of statistical data

Maps are considered nowadays the most effective technique for portraying spatial relationships among statistical data (2, 3). Yet, it was not until the 17th century that the combination of cartographic and statistical skills came together to construct the first *data maps* (usually referred to in cartography as *thematic maps*) (4). One of these first data maps was Edmund Halley's chart showing trade winds and monsoons on a world map (5).

The most widespread techniques of thematic cartography are *choroplet maps*, where different areas on the map are shaded differently according to the value of certain statistical variables, and *combined maps*, where certain well-known statistical chart forms (spot, bar, pie, etc.) are simply superimposed on a geographic map. A combination of both methods is demonstrated in Figure 1. Here shading denotes whether the given country belongs to the European Community; the number of German co-authored papers are shown by a bar diagram.

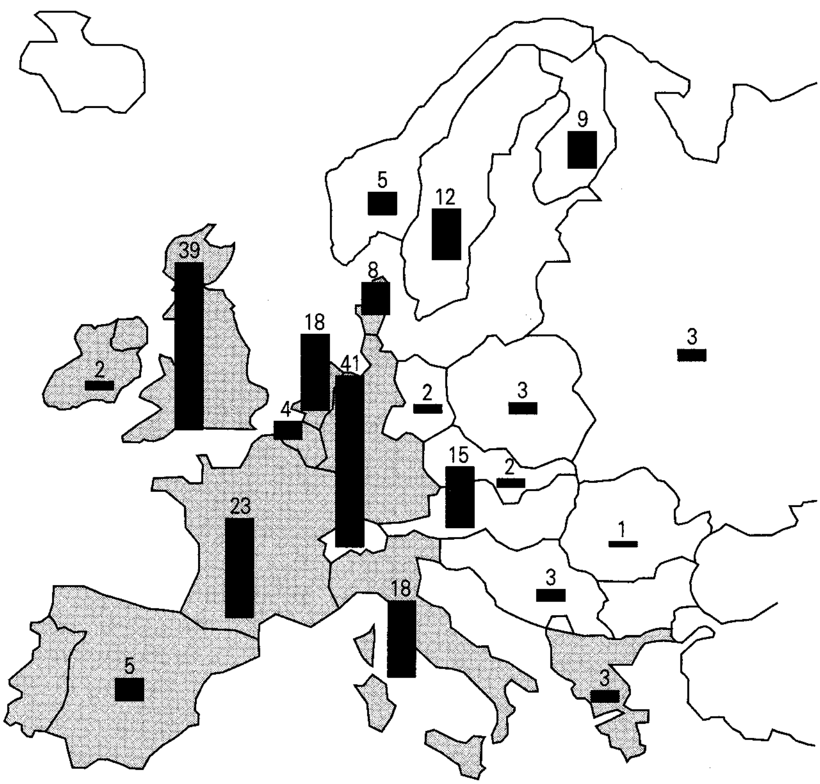


Figure 1. Geographical distribution of highly cited papers among German-European co-operations (6)

Proportional maps

In *proportional maps*, the relative position of the entities concerned (viz., countries) is more or less "natural", while their size (area) is proportional to certain scientometric extension (publication output, citation rate, etc.) (7). Thereby, the relative weights of the countries are easily visualized and, assuming an implicit knowledge of the "natural" proportions, significant deviations can readily be pinpointed. By complementing the proportional maps with proper shading or colouring (or both), multidimensional representation of scientometric indicators becomes possible (8).

A collection of non-scientometric (mainly socio-economic) proportional maps can be found in Ref. (9). Its only science related map is Map 20 on Science Power: its schematic version is shown in Figure 2.

Changes over about a decade can be followed by comparing Figure 2 with similar maps compiled for the period 1981-85 (Figure 3 **a-d** [7]).

Although there are some changes in the details worth noting (for instance the appearance of P. R. China on the map), the basic proportions seem to be unaltered. This stability suggests a certain kind of objectivity of the map. It should, however, also be remembered that both of these maps are viewed from the same specified point of view: that of the Philadelphia based Science Citation Index database, even if the technique of data processing was somewhat different in both cases. A very intriguing question is how such maps taken from various angles would compare. The primary step would be to compile suitable databases representing alternative (preferably "non-metropolitan") viewpoints. This direction seems to be a worthwhile topic for future studies.

Proximity diagrams

Proximity diagrams require the definition of a suitable proximity (similarity) measure. Typical examples are the various co-occurrence (co-authorship, co-citation, co-word, etc.) maps, where the entities considered are placed the closer to each other on the map the greater the similarity between them. For obtaining an exact map (with exactly

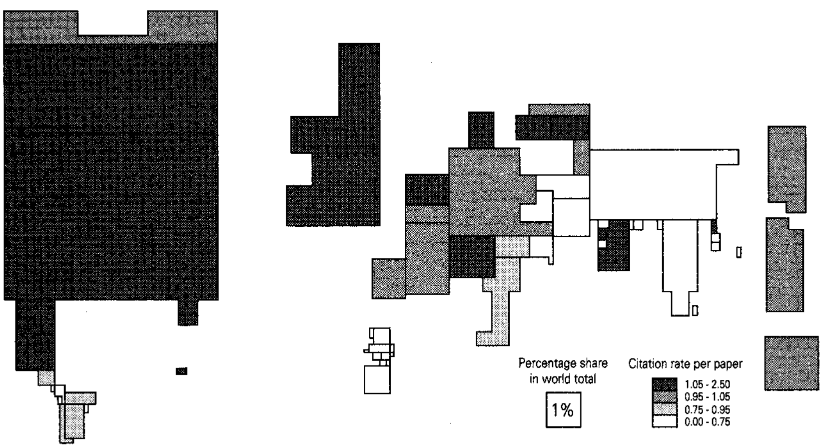


Figure 2. World share in scientific publications (1973) and average number of citations (1973-78)

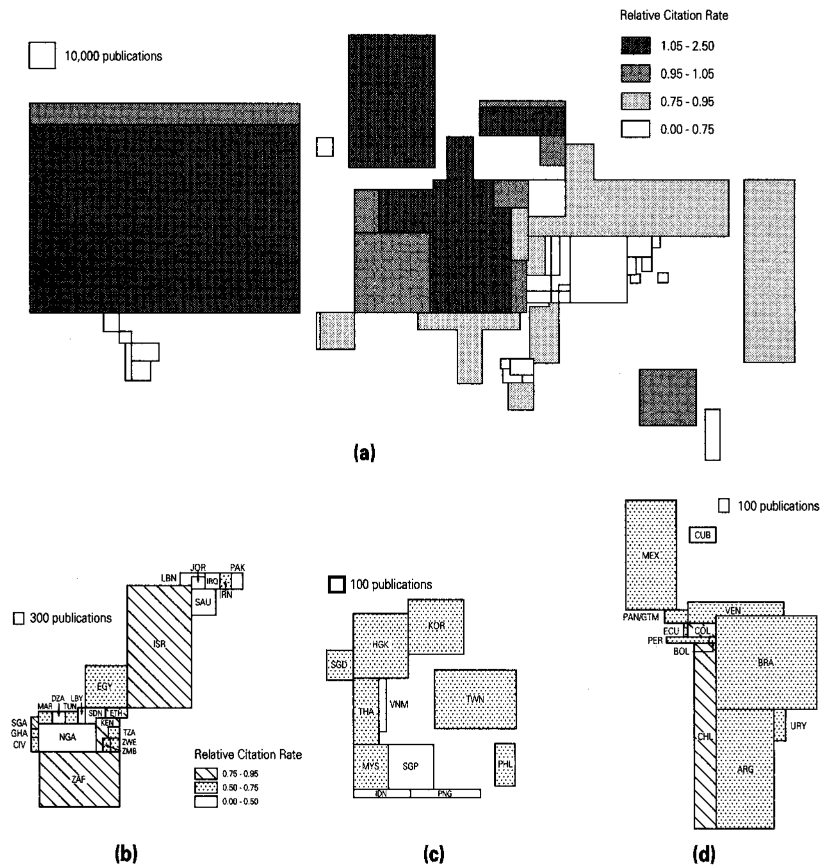


Figure 3. Proportional maps of the World (a), Africa & Near East (b), Far East (c), and Latin America (d) based on publication activity and relative citation rates, 1981-85

determined positions on the plane or space), the existence of a proper distance function (fulfilling reflexivity, symmetry and the triangle inequality) is needed, which is rarely the case. More frequently, the positions are only partially determined by the chosen "metric", and the final pattern may contain arbitrary elements (e.g., the original geographic order can partially be retained).

The map shown in Figure 4 is a more detailed version of that published in (10). Co-authorship links between countries are defined by Salton's measure:

$$r(i,k) = n(i,k)/sqr(n(i)n(k)),$$

where n(i) denotes the total number of publications of the i-th country, n(i,k) denotes the number of publications co-authored by scientists from the i-th and the k-th country. All links greater than 0.5 % are indicated on the map.

Hardly anywhere else than on this map can one see so strikingly the formation of a real *Megapolis* of the eight leading Western countries (USA, UK, Italy, France, Germany, Switzerland, the Netherlands and Belgium) as well as the relation of the rest of the

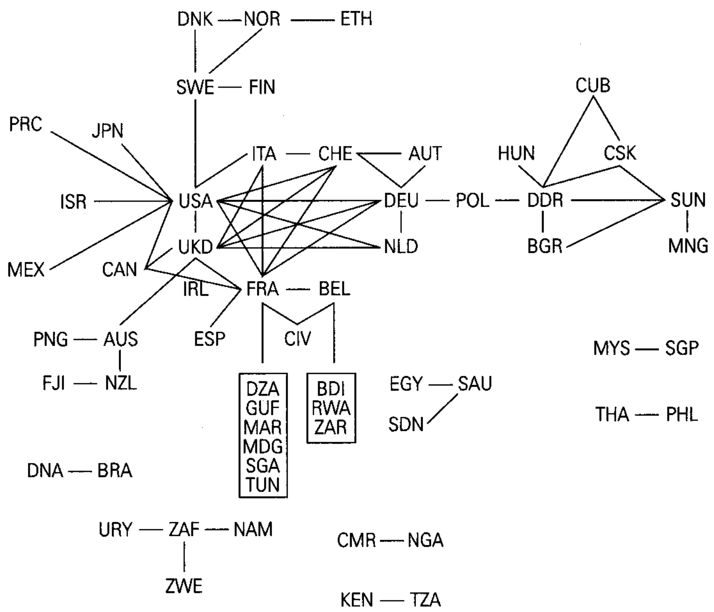
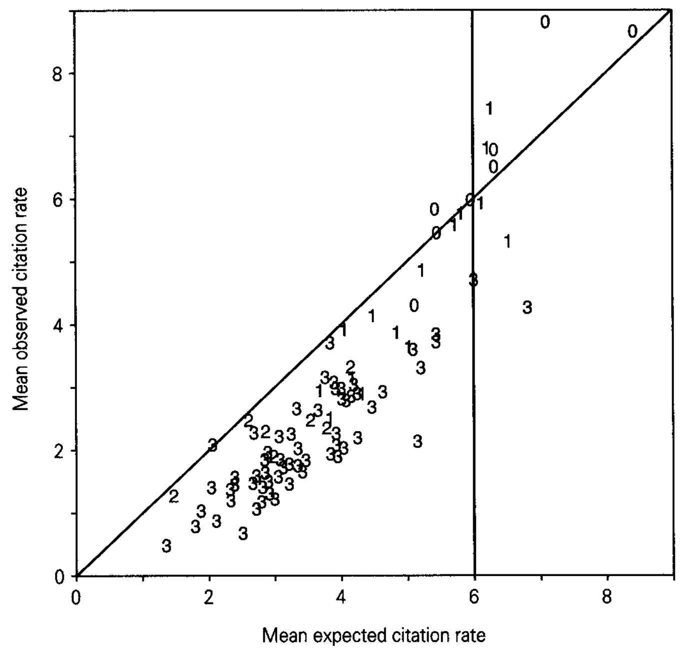


Figure 4. World map of international cooperative links in the sciences, 1983-1985



0: countries of the Megapolis (cf. Figure 5), 1: developed countries, 2: countries of the Soviet Bloc, 3: Third World countries

Figure 5. Relational chart of observed (vertical axis) against expected (horizontal axis) citation rates. Source and citation periods: 1980-89.

world to this metropolis and to each other. Beside a number of more or less obvious facts (e.g., cluster formation on purely geographical grounds - but why in one case and why not in other? -, the links of the former colonies to the mother countries, etc.) light is shed also on some more delicate relations (however evident it seems, the situation of the East European bloc – we are in the mid-1980s! – is worth a closer look).

A question of special interest would be to follow the time trends of cooperation maps, with particular attention to the changes in the turn of the decade from the 1980s to the 1990s.

Relational charts

Relational charts have been introduced to correlate expected and observed citation rates (12). Expected citation rates are those calculated on the basis of the average citation rates of journals (kind of "impact factors"), observed citation rates are those actually received. Position of a country on this kind of charts reflects simultaneously its publication strategy or ambition, as well as its relative success or failure in meeting the envisaged standards.

A typical relational chart can be seen in Figure 6.

A large collection of such relational charts concerning various fields and subfields of science for the period 1981-85 can be found in Ref. (13).

The diagonal of the chart separates lower-than-expected from higher-than-expected citation rates. The vertical line separates lower-than-world-average from higher-than-world-average expected citation rates. The horizontal position on the chart thus signifies the publication strategy or, if you like, the ambitions of a country; the vertical position measures the fulfillment of the expectation, i.e. the extent of the actual success. The four zones delimited by the straight lines are usually occupied by similar groups of countries, largely independently of the time periods considered or even of the subject fields under study. The most prestigious North-East zone (publication in highly cited journals and reaching an even higher citation rate) is typically populated by several *Megapolis* countries plus maybe some from the Scandinavian region. A position in the South-East zone is characteristic of countries with ambitious, "citation-conscious" publication strategy but also with a lack of matching influence. The most typical member of this area is Israel, occasionally accompanied by countries from any of the four groups denoted by 0 to 3. The South-West zone is always the most crowded as another consequence of the well-known "skew distribution" laws of scientometrics. In accordance with the definite and characteristic upward bending shape of the graph, its lowest end closely approaches the diagonal, and this is the home of what used to be the Soviet Union and several of its close associates. This position simply means that these countries published almost exclusively in their own low-impact journals, where they automatically reached the standards set by themselves.

Most conspicuously, the North-West zone is, as a rule, empty. This is a spectacular refutation of the often stressed belief that the cheapest way of reaching higher-than-expected citation rates is to publish only in low-impact journals, where even a few citations might count "higher than expected". In fact it might work in one or two individual cases, but as a strategy, it is perfectly wrong. In the long run, even in mediocre journals, significantly more citations are attracted by authors, institutes, etc. earning credit by regularly publishing in leading, prestigious journals than by those using the mediocre ones exclusively.

Conclusions

Scientopography does not mean a brand new approach to problems of science studies. However, by combining the well-established techniques of scientometrics and topography, it appears to provide rather suitable tools in representing and analysing quantitative data relating to the scientific activity, performance and impact of countries, regions, etc. Particularly, it may help to define and characterize the relation of metropolitan and provincial areas in the world of science.

The few examples given in this presentation taught, among others, that:

1) In spite of its apparent objectivity, topography (whether traditional or "sciento-") can never be free of distorting factors originating from the choice of the point of view and/or the mapping technique.
2) Proximity in science can be explained only partially by factors external to science. In science, being an elitist endeavour by its very nature, strong links between countries can be set by a few active individuals (in extreme cases, even by just one).
3) In the dimensions of ambition and success, countries have almost as characteristic and stable positions as if their longitudes and latitudes were concerned. The standards are set by a chosen (or, rather, naturally selected) few; no other strategy than complete assimilation seems to hold out any promises (the predominance of English language is but a superficial phenomenon).

REFERENCES

1) *Szilard:* "I am going to write down all that is going on these days in the project. I am just going to write down the facts -not for anyone to read, just for God".
Bethe: "Don't you think God knows the facts?"
Szilard: "Maybe he does, but not this version of the facts". WEART, S.R., WEISS-SZILARD G. (Eds), *Leo Szilard: His Version of the Facts. Selected Recollections and Correspondence,* MIT Press, Cambridge, MA, 1978.

2) SCHMID, C.F., SCHMID, S.E., *Handbook of Graphical Presentation,* Wiley, New York, 2nd ed., 1979.

3) SCHMID, C.F., Statistical Graphics. Design Principles and Practices, Wiley, New York, 1983.

4) TUFTE, E.R., *The Visual Display of Quantitative Information,* Graphics Press, Cheshire, CT, 1983.

5) HALLEY, E., An Historical Account of the Trade Winds and Monsoons, Observable in the Seas Between and Near the Tropicks; With an Attempt to Assign the Phisical Cause of Said Winds, *Philosophical Transactions,* 183: 153-168 (1686).

6) GLÄNZEL, W., CZERWON, H.-J., What Are Highly Cited Publications? A Method Applied to German Scientific Papers, 1980-1989, *Research Evaluation,* 2(3): 135-141 (1992).

7) BRAUN, T., SCHUBERT, A., A Topographical Approach to World Publication Output and Performance in the Sciences, 1981-1985, *Scientometrics,* 19(3-4): 159-165 (1990).

8) SCHUBERT, A., BRAUN, T., Three Scientometric Etudes on Developing Countries as a Tribute to Michael Moravcsik, *Scientometrics,* 23(1): 3-19 (1992).

9) KIDRON, M., SEGAL, R., The New State of the World Atlas, Simon & Schuster, New York, 1984.

10) SCHUBERT, A., BRAUN, T., International Collaboration in the Sciences, 1981-1985, *Scientometrics,* 23(1): 3-19 (1992).

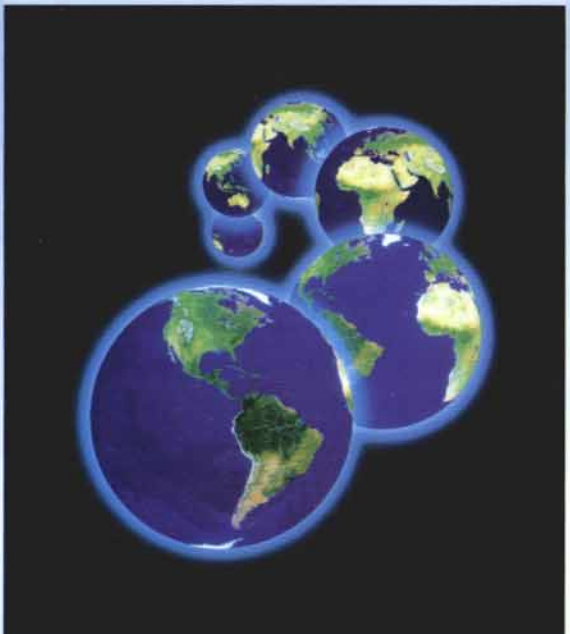
12) SCHUBERT, A., BRAUN, T., Relative Indicators and Relational Charts for Comparative Assessment of Publication Output and Citation Impact, *Scientometrics,* 9(5-6): 281-291 (1986)

13) SCHUBERT, A., GLÄNZEL, W., BRAUN, T., Scientometric Datafiles. A Comprehensive Set of Indicators on 2649 Journals and 96 Countries in All Major Science Fields and Subfields, 1981-1985, *Scientometrics,* 16(1-6): 3-478 (1989)

▪

LES SCIENCES HORS D'OCCIDENT AU XX[E] SIÈCLE

SÉRIE SOUS LA DIRECTION DE ROLAND WAAST



VOLUME 6

LES SCIENCES AU SUD
ÉTAT DES LIEUX

ROLAND WAAST
ÉDITEUR SCIENTIFIQUE

ORSTOM éditions

LES SCIENCES HORS D'OCCIDENT AU XX^{E} SIÈCLE

20TH CENTURY SCIENCES: BEYOND THE METROPOLIS

SÉRIE SOUS LA DIRECTION DE ROLAND WAAST

VOLUME 6

LES SCIENCES AU SUD
ÉTAT DES LIEUX

SCIENCES IN THE SOUTH
CURRENT ISSUES

ROLAND WAAST
ÉDITEUR SCIENTIFIQUE

ORSTOM Éditions
INSTITUT FRANÇAIS DE RECHERCHE SCIENTIFIQUE POUR LE DÉVELOPPEMENT EN COOPÉRATION
PARIS 1996