Scientometrics, Vol. 28. No. 1 (1993) 15-36

THE MEASUREMENT OF INTERNATIONAL SCIENTIFIC COLLABORATION

TERTTU LUUKKONEN,* R. J. W. TIJSSEN,** O. PERSSON,*** G. SIVERTSEN****

**The Academy of Finland, P. O. Box 57, SF-00551 Helsinki (Finland)*
***Centre for Science and Technology Studies, University of Leiden, P. O. Box 9555, 2300 RB Leiden (The Netherlands)*
****Inforsk Group, University of Umeå, S-901 87 Umeå (Sweden)*
*****Institute for Studies in Research and Higher Education, Munthesgate 29, N-0260 Oslo 2 (Norway)*

(Received February 15, 1993)

A growing science policy interest in international scientific collaboration has brought about a multitude of studies which attempt to measure the extent of international scientific collaboration between countries and to explore intercountry collaborative networks. This paper attempts to clarify the methodology that is being used or can be used for this purpose and discusses the adequacy of the methods. The paper concludes that, in an analysis of collaborative links, it is essential to use both absolute and relative measures. The latter normalize differences in country size. Each yields a different type of information. Absolute measures yield an answer to questions such as which countries are central in the international network of science, whether collaborative links reveal a centre – periphery relationship, and which countries are the most important collaborative partners of another country. Relative measures provide answers to questions of the intensity of collaborative links.

Introduction

There has been an increasing interest in international scientific collaboration (ISC) in recent years. ISC is perceived to enhance the scientific capabilities of a country and is considered useful for cost sharing. Experimental work in areas which use expensive equipment needs such a vast amount of resources that work could not be carried out with the support of only one country. An example of such an area is research in experimental high energy physics carried out at the CERN Laboratories.

International collaboration in science is not recent; it started as early as the nineteenth century,[1] but has increased rapidly during the past decades. Between the early 70s and the early 80s, the proportion that internationally co-authored papers are of all papers – one indication of collaboration – doubled,[2] and there is reason to believe that the trend will continue. For example in Europe, this is due to a strong emphasis on collaboration under the auspices of the European Communities.

Elsevier, Amsterdam - Oxford - New York - Tokyo
Akadémiai Kiadó

International collaboration is motivated by many intra-scientific factors, such as a desire to increase knowledge, exchange skills and data, and enhance professional advancement,[3] but it has obtained an additional impetus from facilitated and much less expensive communication (for example, travel, fax and electronic mail) and government support for international scientific contacts and facilities. Intergovernmental science programmes such as the European Communities Science programme have further enhanced international collaboration.

Increased investments in international facilities, joint programmes and greater support for travel abroad have prompted an interest among research managers and policy makers to measure the extent and patterns of ISC and the efficiency of the programmes which are expected to enhance it.[4] Information on ISC can also be used for other purposes, for example, to point out areas of a country's research efforts which are of interest to other countries.[5] Further, information on the intensity of international collaboration between countries has even been used to indicate extra-scientific links between countries.[6]

In order to respond to the demand for information, a host of bibliometric studies using quantitative data based on scientific publications have been carried out in the past few years.[7-14] They have been aimed at measuring the extent of international collaboration in specific research fields or networks of collaboration among countries. A major problem with these studies is a methodological variance which, in some cases, results in conflicting findings. The construction of valid measures of the intensity of collaboration is rendered difficult by vast differences in the volume of research activities in different countries. A country with a large research effort produces a large number of articles and has far more collaborative articles in a literature database than a smaller country. Yet, it may turn out to have less collaboration than a smaller country if the number of collaborative articles is compared to all its articles in the field in question. Moreover, when collaboration between countries of different sizes is studied, the differences in the volume of single and multi-authored articles pose problems. The absolute number of collaborative articles between a small and a large country automatically makes the proportion of collaborative papers for the small country larger than that for the large one, and makes it appear that collaboration is more important for the small country (see section two).

This question has either been overlooked[15] or varying solutions have been offered.[16-19] This paper attempts to clarify the state of methodology in the field. It examines different bibliometric methods to measure collaboration and discusses their adequacy for various purposes.

Bibliometric approaches to the study of international collaboration

Bibliometric studies of collaboration mostly use quantitative data on internationally coauthored research articles provided by the *Science Citation Index* (SCI) and the *Social Science Citation Index* (SSCI). These sources of data provide a complete list of authors and their corporate addresses. The Computer Horizons Inc. (CHI) has processed the 'raw' SCI data for the years 1981–86 by country and scientific field using an eight field classification,[20] and provides such data in a mathematical matrix. Otherwise, the frequency data have to be extracted through online searches or using the CD-Rom version of the (S)SCI data. The collection of data from these sources requires extensive data handling and can be cumbersome. Moreover, the analysis is error-prone due to omissions and differences in spelling of addresses and author names.

The data yield a matrix of inter-institutional collaboration where the number of coauthorships between each country pair is plotted. For example, if a paper has authors from three countries, A, B, and C, six coauthorships (A–B, A–C, B–A, B–C, C–A, C–B) will appear in the matrix. If the number of coauthors from any given country exceeds one, inter-country collaboration is registered only once.

In Table 1 we give an example of a collaboration matrix between a few countries using the CHI data. The simplest way to measure the pattern of collaboration of, for example, Japan and Sweden, is to calculate the percentage of collaborations with authors giving a Japanese and Swedish address (N = 189). This amounts to 1.73% of the total number of Swedish collaborations, but only 1.63% of the Japanese collaborations. Generally, the more countries differ in scientific 'size', the larger the difference in relative terms (e.g. collaborations between a US and Norwegian author make up 22.9% of all the Norwegian coauthorships, but only 1.1% of the US coauthorships).

Table 1

Country-to-country collaborations in 1981–1986, all fields combined: observed frequencies. The matrix is based on information on 30 countries (the largest countries in the world in terms of scientific papers), but for illustrative purposes, only a small selection of countries is displayed

	FIN	JPN	NLD	NOR	SWE	USA		All country-to-coutry collab. (30 largest countries)
FIN	—	50	107	107	549	856	....	3548
JPN	50	—	194	25	189	6082	....	11568
NLD	107	194	—	99	314	2620	....	10440
NOR	107	25	99	—	32	833	....	3645
SWE	549	189	314	632	—	2925	....	10897
USA	856	6082	2620	833	2925	—	....	75665
...	...	....	....	...	....	....	....	.....
All country-to-country collaborations (30 largest countries)	3548	11568	10440	3645	10897	75665	.....	327232

Source: SCI, via Computer Horizons Inc.

The country codes used in this table are as follows: FIN = Finland, JPN = Japan, NLD = The Netherlands, NOR = Norway, SWE = Sweden, USA = USA.

The total number of collaborations Sweden contributed to the matrix of the 30 largest countries was 10 897 and that of Japan 11 568. This means that 3.37% of the collaborations in the matrix were Sweden's and 3.52% Japan's. If Sweden's collaborations were randomly distributed between Japan and the rest of the countries, 3.52% of them would be with Japan. However, Sweden's 'observed' collaborations with Japan were slightly less frequent that 'expected' (1.73%).

The number of collaborations a specific country contributes to the matrix is a function both of its size in terms of scientific papers and its propensity to collaborate internationally. Country sizes differ tremendously, and the differences between their 'observed' and 'expected' collaborations with other countries are generally much larger than in the above example. If we draw far reaching conclusions on the basis of 'observed' values alone, we can easily make mistakes.

We take an example of over- and misinterpretation of coauthorship data from a recent OECD report.[21] it uses data on international coauthorship collaboration. The report states:

> "The collaborative linkages between the two countries (Japan and Sweden) suggest that each focuses on the scientific strengths of the other. Japanese co-operation with Sweden is highest in clinical medicine and biomedicine, which are strong fields of research for the latter country, whereas Swedish collaboration is highest with Japan in chemistry and engineering & technology, which are fields of comparative Japanese strengths. Conversely, there is little collaboration in mathematics and earth & space sciences which are fields where neither country has exceptional research strengths" (p 75).

The conclusions draw on a study by *Miquel* and *Okubo*.[22] The report shows the percentage distribution of Sweden's observed coauthorships with Japan and vice versa. Since Miquel's and Okubo's data are the same as ours, we reproduced the figures on which the above statement is based (Figs 1–2). We give, additionally, information about Sweden's and Japan's shares of all international collaborations in each field in a thirty country matrix, that is the 'expected' distribution of collaborations.

The larger share of clinical and biomedicine in Japan's distribution of collaborations with Sweden is understandable given Sweden's high propensity to collaborate internationally in these fields (Fig. 1). The discrepancies between Japan's 'observed' and 'expected' rates of collaboration with Sweden are greater in this field than the other fields: that is Japan collaborates with Sweden in these fields considerably less frequently than expected. Japan's overall propensity to collaborate internationally is very small.[23] Swedish collaboration with Japan in chemistry, technology and engineering reflects Japanese priorities in international collaboration; Japan collaborates more in technology and engineering than in other fields. The 'observed' rate of Sweden's collaboration with Japan is, nevertheless, far below the 'expected' (Fig. 2).

There are two additional problems in the interpretation of the data in the OECD report. The data are based on a collaboration matrix which does not give any reason to interpret the findings as an indication of the 'focus' of one country toward another or an indication of countries' 'strengths'. 'Focus' implies some kind of causal relationship or direction of the link. Without resorting to additional data such as interview studies, we do not know how collaboration developed and who sought whose company, or whether collaboration is more important to one partner country rather than the other.

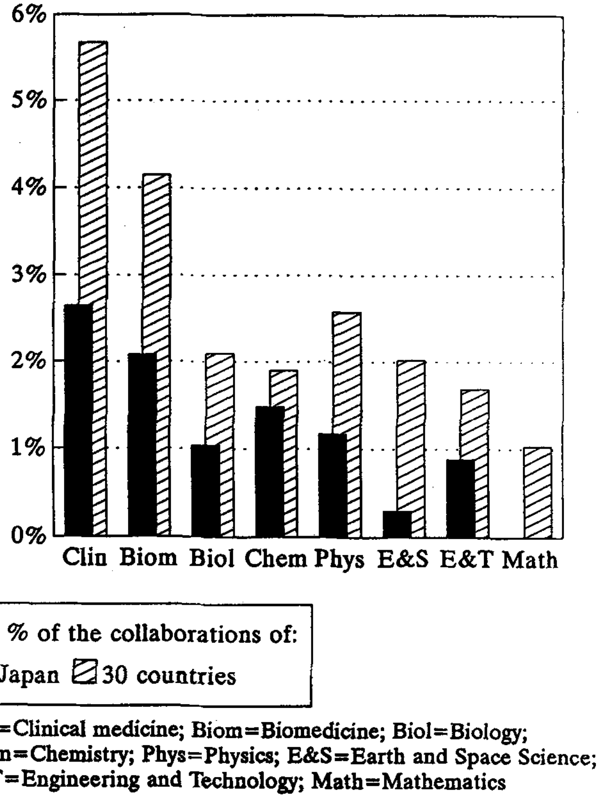

Fig. 1. Collaborations with Sweden as % of the collaborations in the respective field in a matrix of 30 countries

'Strength' has many meanings. The only sense in which the word could be used in this context is *activity* in international collaboration. We can see that Sweden is more active in medical than other fields in terms of international collaboration. The data do not support assumptions of relative strength in performance in terms of citations, quality or article production. It may be misleading to use such a many-faceted word in the interpretation of the data.

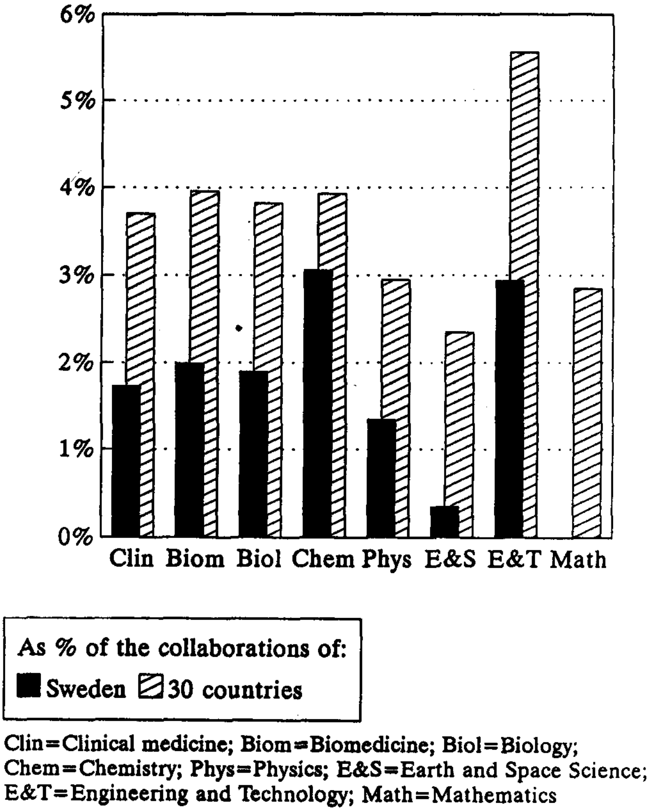

Fig. 2. Collaborations with Japan as % of the collaborations in the respective field in a matrix of 30 countries

If we want to go beyond absolute differences in country sizes and estimate 'propensities' or 'intensities' of collaboration, we have to develop measures which take size into account. Our previous example provides one way to do so when we deal with two-country comparisons. The task is more complicated when we want to develop 'maps of collaboration' which illustrate the structure of collaborations within a set of countries.

Different methods for measuring collaborative networks

Bibliometric studies of ISC face, firstly, the problem of how to attribute papers to the participating countries. We can either attribute papers to all countries that provide authors (i.e. integer counting) or divide the credit among the countries according to a fractional counting procedure. The pros and cons of both for the UK are discussed by e.g. *Anderson* et al.[24] Since our analysis focuses on collaborative links between pairs of countries and not on countries' total output of scientific papers, we will use the former method.

Studies of ISC through quantitative analysis of internationally co-authored papers usually use a data matrix where the rows and columns represent the countries in the same order of appearance. Table 1 is an example of such a matrix. Observe that this example refers to all scientific fields combined. A closer examination of ISC naturally requires a breakdown of the data into scientific disciplines or subfields, which may reveal significant variation in bilateral collaborative relations across the fields.

The ISC relation between a pair of countries can also be coded by a single measure of relatedness instead of using dual percentages. There are several statistical measures of association. The question is which measure to choose? The answer depends on the aspect of the data on which we want to focus. The main consideration is whether and, to what extent, data pertaining to ISC relations with other countries should be incorporated. We distinguish two principal classes of measures: bilateral similarity measures, and multilateral (pattern) similarity measures.

The former can be used if we are primarily interested in comparisons of links between separate pairs of countries. These measures are derived from the observed number of mutual ISC papers C_{xy} of the countries x and y, weighted in one way or another by their total number of ISC papers C_x and C_y. Among the more well-known measures used in science studies are[25]:

Salton's measure[26]:

$$S_{xy} = C_{xy}/\sqrt{C_x \times C_y} \tag{1}$$

and

Jaccard's measure[27]:

$$J_{xy} = C_{xy}/(C_x + C_y - C_{xy}) \tag{2}$$

The maximum and minimum values of both measures are 1 and 0. There are no ready made rules of interpretation for judging the strength of the link.

Table 2 lists the relative ISC values according to those measures. When compared with the data in Table 1 the order of the strengths of links is preserved for most countries. Notable exceptions are the Nordic countries which show a relative increase in the strength of their mutual ISC links. By and large, the Jaccard measure is preferable to Salton's measure since the latter underestimates the collaboration of smaller countries with larger countries; the ISC link between Finland and the USA is a case in point.

Table 2
Measures of relatedness derived from ISC papers in 1981–1986, all fields combined: Salton (in plain print), and Jaccard (in italics). The measures are based on information on the 30 largest countries (in terms of scientific papers), but for illustrative purposes, only a small selection is illustrated

	FIN	JPN	NLD	NOR	SWE	USA	.....
FIN	---	*0.003*	*0.008*	*0.015*	*0.040*	*0.011*	.....
JPN	0.008	---	*0.009*	*0.002*	*0.008*	*0.075*	.....
NLD	0.018	0.018	---	*0.007*	*0.015*	*0.031*	.....
NOR	0.030	0.004	0.016	---	*0.045*	*0.011*	.....
SWE	0.088	0.017	0.029	0.100	---	*0.035*	.....
USA	0.052	0.206	0.093	0.050	0.102	---	.....
....	....	....	....	....	....	....	---

ISC between pairs of countries does not occur in isolation of their collaborative relations with other countries. It is, therefore, important to relate the relative strength of collaborative relations between a pair of countries to their relations with other countries. In most cases, these links are taken into account by comparing the observed and expected numbers of ISC papers. Computation of the expected number of ISC papers according to the *independence model* is based on the assumption that ISC is dependent only on the collaborative efforts of both countries relative to the global ISC activity, that is:

$$E_{xy} = (C_x \times C_y)/T \tag{3}$$

where the statistically expected number of papers E_{xy} is equal to the product of the respective row and column totals (the ISC sizes of both countries) divided by the grand total of the matrix.[28]

Obviously, these expected values are sensitive to the set of countries included in the analysis. It is equally important to note that the non-existent diagonal values in Table 1 (the so-called structural zero's) present a computational drawback, since application of the independence model requires a complete matrix. This problem is solved by application of the *quasi-independence model*,[29] which takes the absence of the diagonal values into account when computing expected values for the off-diagonal elements.

A relative measure of ISC relatedness is expressed by the ratio of observed and expected values. When the ratio exceeds 1, the link is stronger than expected.

When applied to the matrix comprising 30 countries, including all the leading scientific nations (of which part is displayed in Table 1), the expected number of Swedish-Japanese papers is equal to 364 papers. Hence, according to this model, the observed number of ISC papers between Japan and Sweden is only half the expected. By contrast, collaborative relations between Sweden and Finland produce a total of 549 ISC papers, which is more than five times the expected rate. The ratios of observed and expected values for each of the countries used as an example above are listed in Table 3.

Table 3

Ratio (o/e) of observed and expected ISC papers according to the quasi-independence model. The ratios are calculated on the basis of a matrix of the 30 largest countries (in terms of scientific papers), but for illustrative purposes, only a small selection of countries is displayed

	FIN	JPN	NLD	NOR	SWE	USA	....
FIN	---						
JPN	0.43	---					
NLD	1.03	0.56	---				
NOR	3.00	0.21	0.92	---			
SWE	5.03	0.52	0.96	5.64	---		
USA	0.84	1.78	0.85	0.79	0.91	---	
...	....	....	....	....	....	....	....

It is noticeable that the observed value of Finnish collaboration with the USA is below that expected (below 1). The Salton's measure for Finnish-US collaboration in Table 2 is higher than that for Finnish-Norwegian collaboration while the ratio of observed/expected values of the Finnish-Norwegian collaboration in Table 3 is 3, much higher than the value for Finnish-US collaboration. Compared with this method, the Salton and Jaccard formulas produce small values and underestimate the strength of the links between small countries or a pair of countries of which one is small.

An analysis of the ISC profiles of countries yields information as to how they differ with respect to their world-wide ISC distribution.

A common measure of similarity which can also be applied to the ISC patterns of countries x and y, across a set of other countries ($i = 1,\ldots, N$), is Pearson's product-moment correlation, as follows:

$$R_{xy} = \sum_{i=1}^{N} (Z_{x_i} Z_{y_i})/N \tag{4}$$

where Z_{x_i} represents the standardised scores of C_{x_i} (i.e. the mean value of $C_{x_i} = 0$ and the standard deviation of $C_{x_i} = 1$).

This measure gauges the similarity of the pattern of collaboration of two countries. The ISC pattern of the USA differs significantly from that of the other countries, which is understandable considering that the collaborative relations of the USA are relatively evenly distributed across all the other countries (Table 4). Interestingly, there is no similarity whatsoever between the US and Japanese profiles of ISC relationships; the latter emphasizes collaboration with the USA, Western European countries, India and the People's Republic of China. Table 4, however, shows that Japan and Sweden have a high correlation between their profiles which is due mainly to their frequent collaboration with highly industrialized Western countries. They have differences, too; Japan has many collaborations with countries such as the former Eastern bloc countries, the People's Republic of China, India and Canada, while Sweden collaborates relatively more frequently with its Nordic neighbours.

Table 4
Similarity of ISC patterns across all (N = 30) countries (Pearson's correlation). The correlations have been calculated from a matrix of the 30 largest countries (in terms of scientific papers), but for illustrative purposes, only a small selection of countries is displayed

	FIN	JPN	NLD	NOR	SWE	USA	....
FIN	---						
JPN	0.82	---					
NLD	0.83	0.87	---				
NOR	0.96	0.74	0.80	---			
SWE	0.79	0.90	0.89	0.78	---		
USA	0.20	0.03	0.36	0.26	0.15	---	....
...	...	...	...	...	...	...	....

Different methods for mapping collaborative networks

Introduction

The use of tabular data on ISC papers for deriving indicators of scientific collaboration, presented in the previous section, suffices as a point-of-entry for a quantitative analysis of the collaborative links between two countries. However, a more sophisticated analysis of the ISC data is required if one wants to obtain an insight into regularities in the relations between a larger number of countries, for example, to identify 'clusters' of countries with relatively strong ISC links or countries with similar multilateral ISC patterns. Graphical representations ('maps') can provide an overview of such regularities, constructed in an *ad hoc* fashion[30] or by applying sophisticated data-analytical techniques.

Multidimensional scaling (MDS) is one type of multivariate data analysis which can be used to construct maps of networks of ISC relations. Countries are grouped according to a pre-determined, formal decision-rule, which reflects the strength of their relationships.[31] MDS transforms the matrix of ISC frequency data into a set of distances between countries embedded in a high-dimensional Euclidean space where each country is represented by a point. When projected onto the map of the first two dimensions of this space, countries with a close relationship are, on the whole, located near each other whereas unrelated countries are situated a long distance from each other.

In most cases, a two-dimensional space suffices to depict the relevant features of the relational structure. Two-dimensional MDS maps, nonetheless, may suffer from a lack of fit in which case the configuration on the map can be used only for exploratory purposes, for instance, to obtain a rough idea of the structure of the ISC network. On the whole, we should refrain from drawing definitive conclusions from MDS maps, unless other relevant information is added. We must also keep in mind that the location of countries on the map is not fixed, but is (in part) determined by the set of countries included in the analysis. Adding or deleting countries alters the location of countries on the MDS map.

MDS maps are not only well suited for descriptive and exploratory purposes, but they can also provide a spatial framework for adding data which uncover complex features of ISC networks. In the next section, we will illustrate this point by adding connective lines indicating the strengths of bilateral ISC links.

Application: world-wide ISC (1981 – 1986)

In this section, we elaborate on different types of maps and consider their value for addressing different questions regarding ISC, of potential relevance to science policy studies.

Which countries are the major nodes in the global ISC network?

This primarily involves mapping based on the absolute numbers of bilateral ISC papers. An MDS map reflecting the absolute magnitude of ISC links is displayed in Fig. 3.[32]

The map consists of two interdependent layers: (1) the configuration of points on the map, and (2) a network of connecting lines. The first feature results from the analysis of all ISC linkages, while the second feature portrays the (second) strongest bilateral links. Basically, the configuration shows a centre-periphery structure where the USA, Canada and Western European countries are located in the centre. The clustering of peripheral countries depends on a mixture of cultural, political factors (e.g. Eastern Europe), geographical factors (e.g. Japan and China) and language (e.g. Spain and Argentine). *Gómez* and *Méndez*[33] discussed whether such (semi-) peripheral countries profit from the global network and found that the answer varied to some extent with the country and the field of science in question.

There are only a couple of countries which have their strongest ISC link with a country other than the USA; the abundance of grey lines clearly illustrates this. The lines also stress the important role of Great Britain, Germany, France and Sweden as secondary nodes in the global ISC network.

The asymmetrical character of the links should be read, for example, as follows: the largest number of international collaborations of the (former) German Democratic Republic, (former) Czechoslovakia and Bulgaria are with the (former) USSR, while the latter has its largest number of collaborations with the German Democratic Republic. The USA is the second most important partner of the USSR. Most of the US international collaborations are with authors from Great Britain and the latter collaborate second most with authors from Canada.

It is interesting to see that Hungary's ISC links with the USA are stronger than its ties with the USSR. Moreover, the position of Austria and Poland between the Eastern and Western European countries and Poland's links with the USA and the Federal Republic of Germany indicate their intermediary role between Western European countries and those of the former Eastern bloc.

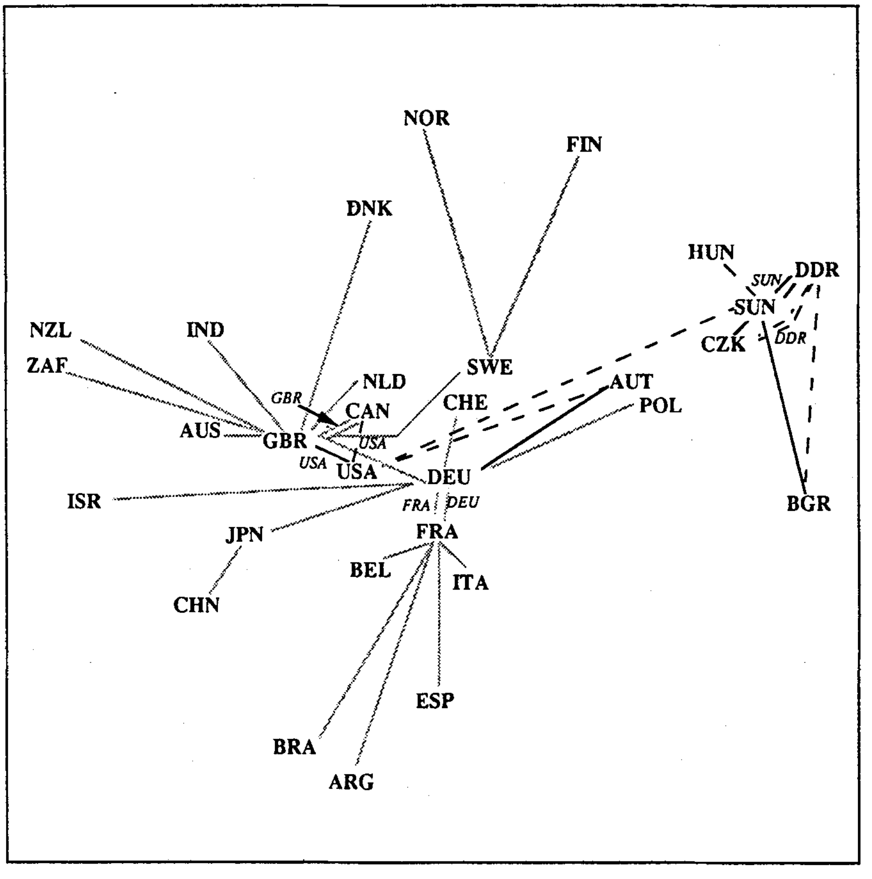

Fig. 3. Network map of international co-authorship relationships amongst 30 countries. Based on observed frequencies. Derived with multidimensional scaling. Fit: 67% (RSQ = 0.67).
Country codes: ARG = Argentina; AUS = Australia; AUT = Austria; BEL = Belgium; BGR = Bulgaria; BRA = Brazil; CAN = Canada; CHE = Switzerland; CHN = People's Republic of China; CZK = Czechoslovakia; DDR = German Democratic Republic; DEU = Federal Republic of Germany; DNK = Denmark; ESP = Spain; FIN = Finland; FRA = France; GBR = Great Britain; HUN = Hungary; IND = India; ISR = Israel; ITA = Italy; JPN = Japan; NLD = the Netherlands; NOR = Norway; NZL = New Zealand; POL = Poland; SUN = USSR; SWE = Sweden; USA = USA; ZAF = South Africa.
Connective lines: ——— Strongest link with a country other than the USA
– – – – Second strongest link – second to link with a country other than the USA
············ Second strongest link – second to link with the USA
Missing lines refer to links with the USA. Ambiguties are avoided by labeling selected lines with the respective country involved.

Does accounting for scale differences between countries affect the ISC structure significantly?

Here the analysis focuses on the relative magnitude of bilateral ISC papers. The MDS map of the relative ISC values according to the Salton's measure is displayed in Fig. 4. Adjusting to differences of scale induces a more complex network which is marked by an increasing diversity of collaborative links between countries. The much smaller number of grey lines reflects a less dominant position of the USA and stronger links between European countries. Basically the structure is, nevertheless, similar to the configuration in Fig. 3 with the exception of one significant new feature, that is, the switching of positions of the Anglo-saxon cluster (GBR, CAN, IND, AUS, NZL and ZAF) and the cluster of Romance language countries (FRA, ITA, BEL, ESP, BRA and ARG).

Examination of the data reveals a multitude of small changes of which the increased relative strength of the Italian-Swedish collaboration seems to be the most pronounced one. Although the adjustment for scale differences uncovers several hitherto hidden strong links between countries (e.g. Belgium and the Netherlands), the core position of the USA and the number of the dashed lines indicates that its dominant position is preserved in large parts of the network.

Will relevant features of the ISC network alter when the global ISC magnitude is incorporated in the analysis?

The analysis incorporates not only the relative bilateral strengths, but also takes into account their strengths relative to world-wide ISC. MDS is applied to the chi-square values of relatedness,

$$X^2_{xy} = (C_{xy} - E_{xy})^2 / E_{xy} \tag{5}$$

which is formally equivalent to applying a special member of the MDS family that is mostly referred to as correspondence analysis. The special version of correspondence analysis is based on the (quasi-)indepedence model[34] and yields the map exhibited in Fig. 5.

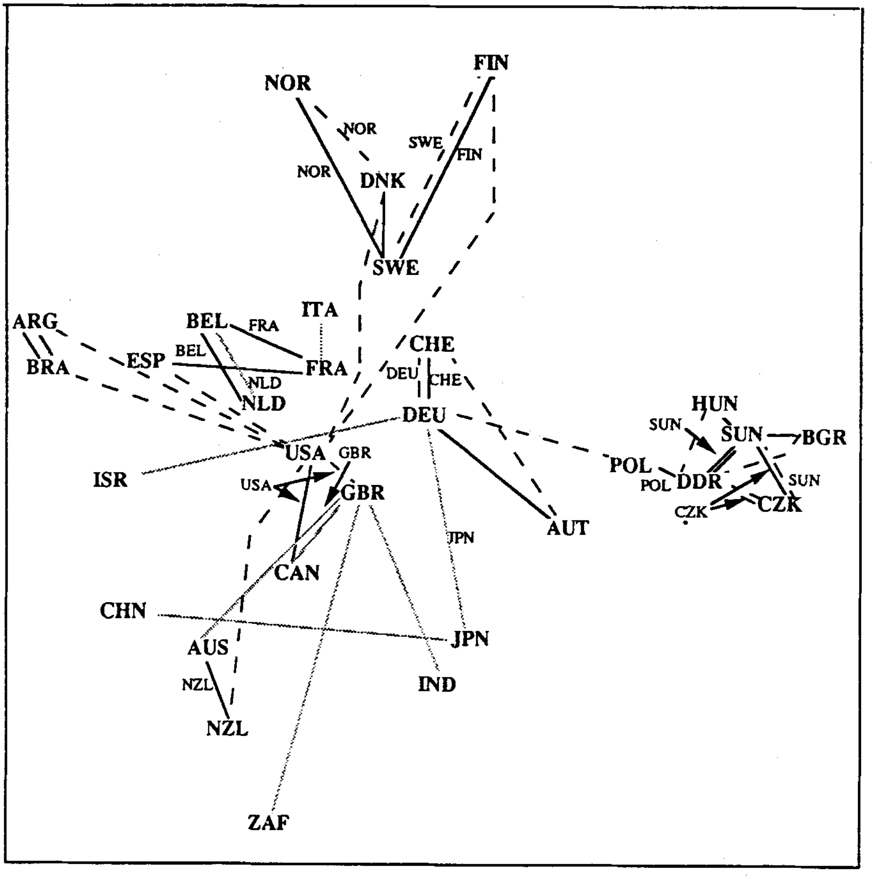

Fig. 4. Network map of international co-authorship relationships amongst 30 countries. Based on Salton's index. Derived with multidimensional scaling. Fit: 68% (RSQ = 0.68).
Connective lines: ——— Strongest link with a country other than the USA
– – – – Second strongest link – second to link with a country other than the USA
.... Second strongest link – second to link with the USA
Missing lines refer to links with the USA. Ambiguties are avoided by labeling selected lines with the respective country involved.

A basic structure arises with features resembling the configuration exhibited in Fig. 3. However, the Nordic countries are now quite peripheral due to their strong intra-Nordic ISC relations. Another noticeable change is the more peripheral position of the USA, compared with the Western European countries (which is no doubt partly due to the emphasis on Europe in this selection of countries).

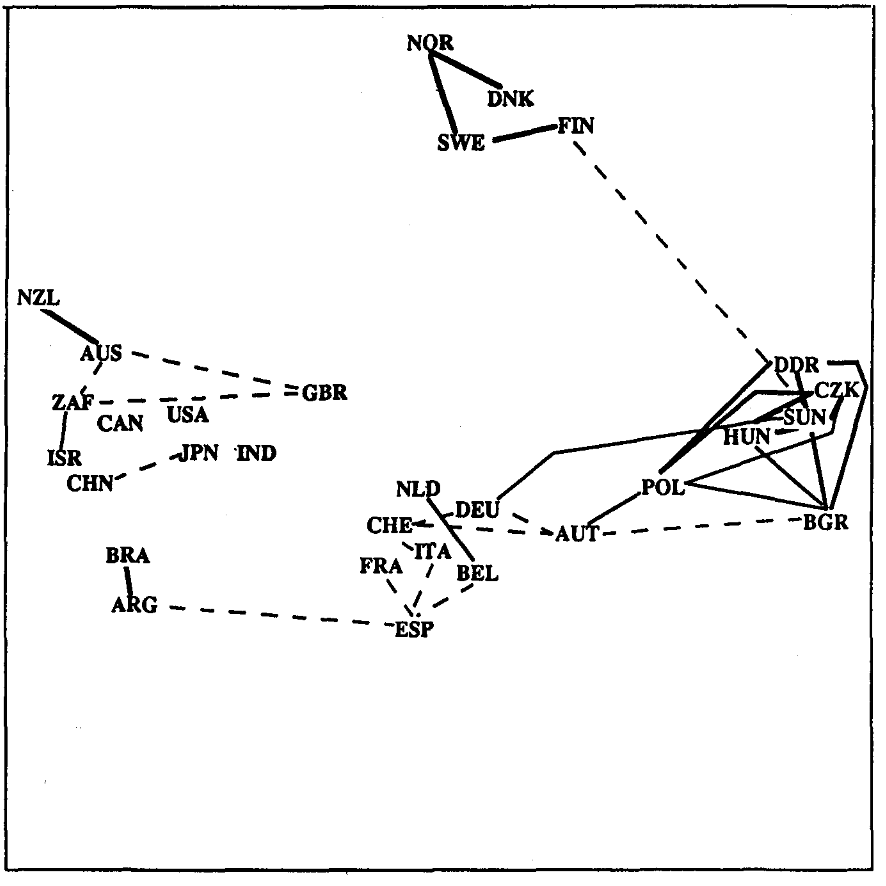

Fig. 5. Network map of international co-authorship relationships amongst 30 countries. Derived with correspondence analysis based on a quasi-independence model. Fit: 51.4%.
Connective lines: (thick line) observed value more than 5 times the expected value
(thin line) between 3 and 5 times
(dashed line) between 2 and 3 times

The structure of links shows a further emphasis on relatively strong bilateral links, in this case, between the Eastern and Central European countries in particular. There is a striking lack of countries which have collaborative ties with the USA that are stronger than twice the expected number. This actually confirms the core position of the USA; its ISC distribution is spread across all the other countries approximately in proportion to their size (in terms of collaborative papers). There are, nonetheless, nine countries with a value higher than 1, with the maximum being 1.9 between the USA and Israel.

Important information about the relative linkages between countries can be inferred by juxtaposing information in Figs 3–5. Consider, for example, Sweden and Finland. Figure 3 indicates that each country has most international collaboration with the USA. Sweden's second most important partner country is Denmark, while Finnish scientists collaborate second most with Swedish scientists. However, the size-independent relations depicted in Fig. 4 show a different picture: Finland is Sweden's most important collaborative partner whereas the USA remains Finland's most important partner country. Finally, when adjusting to the expected world-wide level of collaboration, shown in Fig. 5, yet another picture arises: Finland's link with the USA appears to be more or less as expected in view of the number of internationally coauthored papers of both countries, whereas the number of collaborations between Finland and the former USSR exceeds twice the expected number.

Are there significant differences if the analysis focuses on multilateral ISC patterns instead of bilateral links?

The answer is: yes. Focusing on multilateral patterns yields a different, albeit not completely different, configuration. The MDS analysis based on the Pearson correlations between the ISC profiles of the 30 countries results in the map shown in Fig. 6. The main East-West distinction is still found on the horizontal axis. On the vertical axis the USA has shifted to a location opposed to the former Soviet Union due to the difference between the USA's dispersed ISC profile versus the Soviet profile which concentrates on the Eastern European countries. The countries with significant ISC ties with both the East and the West – the intermediaries, Austria, Poland, and Hungary – are located in between. Adding a third dimension to this map does not differentiate the cluster of countries in the centre of the map. So, it seems that in terms of their ISC profiles, these countries have more or less the same ISC distribution across other countries – a finding corroborated by the high correlations exhibited in Table 3.

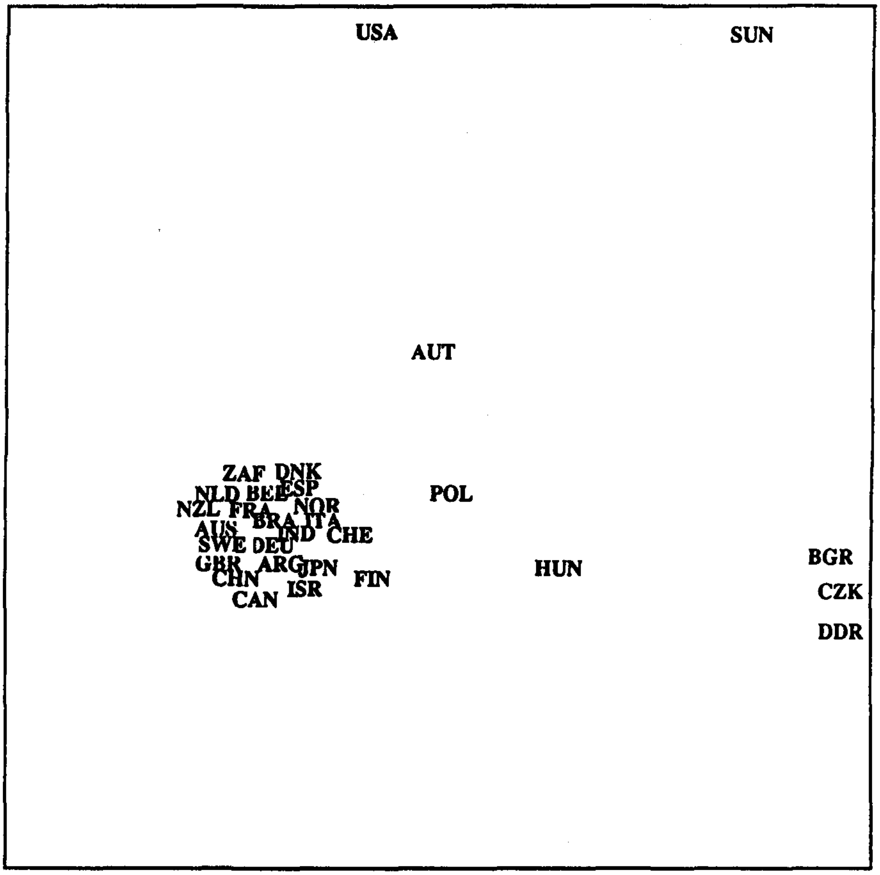

Fig. 6. Map of international co-authorship relationships amongst 30 countries. Derived with multidimensional scaling applied to Pearson correlations. Fit: 97.0%

Conclusions

The choice of measures for gauging the strength of ISC relationships clearly affects the findings. In an analysis of collaborative links, it is essential to use both absolute and relative measures. Absolute measures may yield an answer to questions such as which countries are central in the international network of science, whether collaborative links reveal a centre – periphery relationship, or in absolute size, which countries are the most important collaborative partners of another country. It is essential to recognize that absolute measures are affected by both country size and collaborative activities. Larger countries generally feature as dominant in such absolute comparisons.

If we want to study the relative strength of links between countries we have a choice of several methods which attempt to normalize size.

Bilateral similarity measures such as Salton's and Jaccard measures pay attention to two countries at a time. They change the picture provided by absolute measures, but not very much. They may incorporate a bias against small countries and tend to underestimate the links between them.

Multilateral similarity measures which take a global network into account, such as Goodman's quasi-independence model, calculate the ratio between the observed and expected volume of links. The calculations are sensitive to the countries included in the data set. The model is based on the assumption that the expected number of ISC links between pairs of countries occurs in proportion to the countries' share in the global ISC network. This measure gives equal importance to small and large countries.

Multidimensional scaling methods may be used to construct maps of networks of ISC relations for descriptive and exploratory purposes. Both absolute and relative measures, referred to above, may provide input data for such maps. The strengths of the maps based on different types of data are to a large extent similar to the strengths and weaknesses of such data in a tabular form. Nevertheless, two-dimensional maps may yield a limited picture of the dimensions underlying the data, since international collaborative links are influenced by a multitude of interdependent factors. Two-dimensional maps, therefore, serve to illustrate and draw attention to the most striking factors and can provide a starting point for more in-depth analyses of the collaborative relations between any set of countries. It turned out, nevertheless, in this study, that adding further dimensions produced a similar basic configuration and increased the fit only to a negligible degree.

*

We are grateful to Helmut A. *Abt* and Phoebe A. *Isard* for their useful comments on this paper.

Notes and references

1. D. deB. Beaver, R. Rosen, Studies in scientific collaboration. Part I. The professional origins of scientific co-authorship, *Scientometrics*, 1 (1978) 65–84.
2. T. Luukkonen, O. Persson, G. Sivertsen, Understanding patterns of international scientific collaboration, *Science, Technology, & Human Values*, 17 (1992) 101–126.
3. D. deB. Beaver, R. Rosen, Studies in scientific collaboration. Part II. Scientific co-authorship, research productivity and visibility in the French scientific elite, 1799–1830, *Scientometrics*, 1 (1979) 133–149.
4. E.g. G. Lewison, P. Cunningham, Bibliometric studies for the evaluation of transnational research, *Scientometrics*, 21 (1991) 223–244; The EC SCIENCE/Stimulation Programme. *Scientometrics*, 21 (1991) 325–342.
5. J.-F. Miquel, Y. Okubo, Indicators to measure internationalization of science. A paper presented at a conference organised by the OECD, "Consequences of the Technology Economy Programme for the Development of Indicators", Paris, 2nd–5th July 1990.
6. R. de Bruin, R. R. Braam, H. F. Moed, Bibliometric lines in the sand, *Nature*, 349 (1991) 559–562.
7. R. J. W. Tijssen, H. F. Moed, Literature-based statistical analyses of international scientific co-operation. An exploratory case study of the Netherlands. In: A. F. J. van Raan, A. J. Nederhof, H. F. Moed (Eds), *Science and Technology Indicators. Their Use in Science Policy and Their Role in Science Studies*, Leiden, DSWO Press, 1989, p. 129–145.
8. F. Narin, E. S. Whitlow, Measurement of scientific cooperation and coauthorship in CEC-related areas of science. Vol 1. Commission of the European Communities (EUR 12900 EN), 1990.
9. A. Schubert, T. Braun, International collaboration in the sciences, 1981–1985, *Scientometrics*, 19 (1990) 3–10.
10. J.-F. Miquel, Y. Okubo, *op. cit.* note 5.
11. G. Lewison, P. Cunningham, Bibliometric studies for the evaluation of trans-national research, *Scientometrics*, 21 (1991) 223–244.
12. H. F. Moed, R. E. de Bruin, A. J. Nederhof, R. J. W. Tijssen, International scientific co-operation and awareness within the European Community: Problems and perspectives, *Scientometrics*, 21 (1991) 291–311.
13. F. Narin, K. Stevens, E. S. Whitlow, Scientific co-operation in Europe and the citation of multinationally authored papers, *Scientometrics*, 21 (1991) 313–323.
14. T. Luukkonen, O. Persson, G. Sivertsen, *op. cit.*, note 2.
15. J.-F. Miquel, Y. Okubo, *op. cit.*, note 5.
16. A. Schubert, T. Braun, *op. cit.*, note 9.
17. R. J. W. Tijssen, H. F. Moed, *op. cit.*, note 7.
18. T. Luukkonen, O. Persson, G. Sivertsen, *op. cit.*, note 2.
19. F. Narin, E. S. Whitlow, *op. cit.*, note 8.
20. Earth and space science, mathematics, physics, biomedicine, biology, chemistry, engineering and technology, clinical medicine.
21. Science and Technology Policy – Review and Outlook 1991. OECD: Paris 1992.
22. J.-F. Miquel, Y. Okubo, *op. cit.*, note 5.
23. T. Luukkonen, O. Persson, G. Sivertsen, *op. cit.*, note 2.
24. J. Anderson, P. M. D. Collins, J. Irvine, P. A. Isard, B. R. Martin, F. Narin, K. Stevens, On-line approaches to measuring national scientific output: A cautionary tale, *Science and Public Policy*, 15 (1988) 153–161.
25. Our choice of the measures presented here has been guided by the following methodological and practical considerations: appropriateness, accessibility, and illustrative power.
26. G. Salton, D. Bergmark, A citation study of computer science literature, *IEEE Transactions on Professional Communication PC–22*, 3 (1979) 393–440.

27. P. H. A. SNEATH, R. R. SOKAL, *Numerical Taxonomy*, Freeman, San Francisco, 1973.
28. E.g. T. LUUKKONEN, O. PERSSON, G. SIVERTSEN, *op. cit.*, note 2.
29. L. A. GOODMAN, The analysis of cross-classified data: independence, quasi-independence, and interactions in contingency tables with or without missing entries, *Journal of American Statistical Association*, 63 (1968) 1091 – 1131.
30. E.g. A. SCHUBERT, T. BRAUN, *op. cit.*, note 9; R. DE BRUIN, R. R. BRAAM, H. F. MOED, *op. cit.*, note 6.
31. R. J. W. TIJSSEN, *Cartography of science: scientometric mapping with multidimensional scaling methods*, Leiden, DSWO Press, 1992.
32. Based on nonmetric MDS with the ALSCAL programme supported by the statistical package SPSS. The squared correlation between the actual and fitted distances in a two-dimensional space – the RSQ measure – is equal to 0.67, which means that 67% of the variance is accounted for by these dimensions jointly. Empirically, this proved to be the most parsimonious representation, that is, additional dimensions slightly increase the fit, but hardly affect the structure of the configuration.
33. I. GÓMEZ, A. MÉNDEZ, Are peripheral countries profiting from scientific cooperation networks? In: P. WEINGART, R. SEHRINGER, M. WINTERHAGER (Eds), *Representations of Science and Technology*, Leiden, DWSO Press, 1992, 112 – 123.
34. R. J. W. TIJSSEN, J. DE LEEUW, A. F. J. VAN RAAN, Quasi-correspondence analysis on scientometric transaction matrices, *Scientometrics*, 11 (1987) 351 – 366, describes an alternative method for doing a quasi-correspondence analysis through the application of a standard correspondence analysis to a pre-processed data matrix.