Mapping Global Science Using International Co-authorships:
A Comparison of 1990 and 2000

Caroline S. Wagner[*] and Loet Leydesdorff
Science & Technology Dynamics,
Amsterdam School of Communications Research (ASCoR),
University of Amsterdam

Abstract

Using the Science Citation Index (CD-Rom version) for 1990 and 2000, this paper analyses international co-authorships. The paper presents the methodology for identifying these co-authorships. Analysis is presented on observed linkages at the global level and on regional bases. The architecture of the network is further explored using statistical methods and factor analysis to reveal intense relationships as well as the core members of a global network. Findings show that, in the 10 years between 1990 and 2000, the global network has expanded to include more nations and it has become more interconnected. Regional networks show emerging hubs. Within the global network, a core set of countries has expanded from six in 1990 to eight in 2000. Factor analysis suggests that large countries compete with each other for partners in the global network. We discuss implications for public policy of the rise of a global network of scientists operating somewhat independently of national interests.

Keywords: international cooperation in science and technology; global science; networks; co-authorships

1. Introduction

This paper analyses the dynamics of science as a global system by examining international co-authorships. It improves upon earlier analysis by including all the countries of the world, and by applying a range of tools, including social network analysis and factor analysis, to expose the network. We explore whether the network created by international links among scientists might be taking on the features of a global system overlaid upon and somewhat independent of the national systems. We discuss the implications of the emergence of such a system for knowledge creation and public policy.

There has been a rapidly growing literature discussing the increase in international linkages in science. Authors have approached the question from three perspectives: 1) scientometric analysis of the increase in the interconnectedness of scientists (examples include, Glänzel, 2001; Zitt, et al., 2000; Luukkonen, 1993; Okubo et al., 1992); 2) social sciences analysis of collaboration in general (Katz and Martin, 1997; Gibbons et al., 1994) and international linkages in particular (Schott, 1998; Stichweh, 1996); and 3) policy analysis of the implications of linkages for funding and outcomes (Advisory Council of Canada, 2001; Wagner et al., 2000).

Explanations and theories about why international linkages in science are increasing depend upon the perspective of the author and the scope of the data being examined. Table 1 provides a schema of the factors considered and theories advanced to explain the increases in collaboration.

[*] Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands, Wagner@pscw.uva.nl.

Factors….	Internal to Science	External to Science
Relating to the diffusion of scientific capacity	Center-periphery dynamic of lagging countries seeking to cooperate with the leaders (Ben-David, 1971)	Official support for investment in research and development (Wagner et al., 2001)
Relating to the interconnectedness of scientists	Internal disciplinary differentiation of science (Stichweh, 1996) Field-specific characteristics of 'big science' (Galison, 1987)	Historical relationships related to geographic proximity or colonial ties (Zitt et al., 2000) Increase in international trade (Schott, 1991) Growth of information and communications technologies (Gibbons et al., 1994)

Table 1. Explaining the rise in international collaboration in science

The theories of why international collaboration is increasing appear to consider factors both internal and external to science. In addition, they address either the diffusion of scientific capacity or the interconnectedness of researchers, although, at times, the two features are not well delineated. In discussing reasons for the diffusion of scientific capacity, for example, Schott (1998, 1991), following Ben-David (1971) and Shils (1988), sees the progression related to a succession of countries that have acted as "centers" for world science, with countries at the periphery (often smaller countries) trying to emulate the organisation and orientation of scientific work at the center. As they emulate and adapt the practices of the core country, the capacity of the periphery countries grows. Wagner et al. (2001) document and index capacity, and they suggest that the increasing investment in research and development by governments and non-governmental organisations (such as the World Bank)—groups interested in using science as a tool to aid development—has contributed to the diffusion of capacity.

A greater volume of literature addresses the increasing interconnectedness of researchers. Reasons offered for this phenomenon (ones highlighting factors internal to science) include Stichweh's (1996) assertion, following Price (1963) that collaboration arises from the dynamics of internal differentiation of science into specialized disciplines. Galison has suggested that, at least for some sciences, the scale of investment is so large that no single nation will undertake it alone. These field-specific characteristics make some collaborations unavoidable. (Galison, 1987, quoted in Kim, 2002) Factors external to science offered as explanations for collaboration include geographical proximity and historical determinants explored by Zitt et al. (2000), and the ubiquitous mention of the rise of information and communications technologies as influencing the interconnectedness of everyone, not just those in the scientific community. (Gibbons et al., 1994 is just one example)

We are interested in examining both the diffusion of capacity and the interconnectedness of researchers at the network level since we view these as related phenomena. By looking at 1990 and 2000, we sought to test the hypothesis that global science has grown more decentralized and more interconnected as well as more inclusive of new entrants.   Further, we expected that growth in international collaboration documented by others (Glänzel, 2001; Luukkonen, 1992; Schubert, 1990; Georghiou, 1998; Doré, 1998) has resulted in stronger, more robust networks among scientifically active countries. In addition, we expected to find that developing countries have become more active participants within regional networks, with links to more advanced countries within their region (e.g., South Africa) growing more quickly than links among developing countries themselves. In the conclusion, we discuss how our findings can be interpreted in light of the different theories of why international collaboration in science has increased.

2. Time Frame for the Study

The 10-year interval between 1990 and 2000 is a particularly interesting time to examine scientific collaborations. Five events significantly changed the environment for science during this time: 1) the break-up of the Soviet Union and the end of the bi-polar world, 2) the reunification of Germany, 3) the growth of information and communications technologies (specifically, the Internet and the World Wide Web), 4) the rise of the European Union and the growth in its funding for science, and 5) the globalization of industry.[1]   Each of these factors has influenced either the supply of scientific knowledge (the newly independent states have been encouraged to expand their participation in global science, for example) or the demand for science (the local information requirements of global businesses have encouraged investment in research). Each of these factors has also affected the political process that provides funding and support for science in ways too numerous and complex to be described here. However, it is important to set this as a marker, as we will come back to this question in the conclusion.

3. A Confusion of Terms

The terms describing the growth of international collaboration are varied. Unfortunately, at least in the English language, the word "international" as applied to science has at least three different meanings. It can alternatively refer to the political interactions of nations around scientific interests, or to the political origins of research practitioners. A third meaning is "universal" in the sense that knowledge creation in science is independent of the political or personal attributes of those creating and using that knowledge. At times, the term could be inferred to carry all these meanings.

In their book *Denationalising Science*, Crawford, Shinn, and Sörlin (1993) define "international" as "activities involving persons, equipment, or funds from more than one country."   They note that the term implies political interest and organisation. To take one step back, the concept of the "national" connotes political, economic, and geographic congruence. This would certainly carry over to descriptions of a "national science policy," or "national research and development budgets" as these elements of policy are reported by groups such as the Organisation for Economic Cooperation and Development (OECD). The idea of the "international," at least as it first emerged in the 19th century, was that interactions were possible among these political nations, and that groups from within these nations could work together under the political protection and support offered by the nation-state. This concept is descriptive of science as it has been practised over the past 50 years—nationally funded practitioners do seek to work with others under politically determined guidelines—but the application of the term requires a more nuanced understanding.

Crawford, Shinn, and Sörlin (1993) argue for the use of the term "trans-national" rather than "international" in order to remove the political connotations of the latter term. This approach has an intuitive appeal. It does appear that scientists create linkages independent of the political process, relationships that seem to "rise above" nations and create a meta-network or a super system. Scientists, after all, are not necessarily seeking to establish political ties when they collaborate. However, in most cases, funding is still tied to national interests. Thus the term "international" may still be preferable to "trans-national," since it includes within it the political interests as well as the sense of the "universal" nature of scientific knowledge.

The confusion of terms is more than simply a semantic curiosity. The application of the term "international" to the extra-national linkages among scientists masks the rise of an intriguing and somewhat puzzling shift in the organisation of science. To the question of why international science is growing so rapidly, an initial response is sometimes

---

[1] There are a number of different descriptions of the process of globalization of industry. One useful description, directly relevant to this paper, is contained in the Science & Engineering Indicators 2000. (NSB, 2000).

incredulity: "Of course international science is growing: science is international!" By this, people usually mean "universal" in the sense that knowledge is available to anyone regardless of their political identity, gender, race, or other outwardly defining characteristics. However, science taking place collaboratively across geographical distances inherently has more transaction costs than intra-laboratory collaboration. Working across different time zones, as just one example, can affect productivity. Moreover, given the immense importance attached to recognition in science (e.g., access to funds) international collaboration might be considered as a less appealing choice for the most eminent or ambitious scientists. The fact that the opposite appears to be true, and that international links in science proliferate despite these countervailing factors, suggests that the motivating forces behind this shift, whether internal or external to science or both, are very compelling indeed.

Because the term "international science" can be a source of confusion, given the various meanings it carries, we refer to the linkages being created among scientists, the ones that extend beyond or transcend national borders, as "global science." We choose this term because we want to explore the possibility that science is growing increasingly independent of the national political system implied in the term "international." Moreover, this follows the frequent use of the term "global business" to mean business that has an international component over and above national or local interests.

4. Methodology

The place to begin examining global science is by detailing the patterns of linkages among researchers from different nations.   Both quantitative and qualitative approaches can shed light on this question, but this article focuses on the results of quantitative analysis of the co-occurrence of authorships among practitioners in different countries in the two years examined. An effective way to quantify linkages is using co-authorships.

A scientific document is co-authored if it has more than one author, and it is *internationally co-authored* if at least two of these authors list addresses in different countries. Data on co-authored articles can be retrieved from almost any bibliographic database. However, the Science Citation Index (SCI) produced by the Institute for Scientific Information (ISI) is the most reliable source for a comprehensive survey. We drew data from the Science Citation Index CD-Rom[2] from 1990 and 2000. The addresses are generally well standardized and can be compared over time. The addresses make it possible to study co-authorships using *nations* as the unit of investigation.[3] We did not distinguish among types of contributions (reviews, letters, proceedings, and journal articles) because we are seeking social connection to reveal the structure of the network regardless (at this point) of the scientific status of the output.

When using co-authorships as the indicator of international linkages, there are a number of validity issues to consider. Scientific collaboration may lead to a range of outcomes, and a co-authored article is only one of these outputs. Katz and Martin (1997) have detailed a number of issues associated with using co-authorship data as an indicator of collaboration. They note that collaboration does not necessarily lead to co-authored publications. Conversely, co-authorship does not necessarily indicate collaboration: for example, research leaders may list their names on articles without actually contributing to the work. Finally, in many cases, although international collaboration is taking place, co-authors list a common institution in which they are working, thus leading to an

---

[2] The CD-Rom version is preferable to the SCI Extended version found on the Internet. The Extended version is regularly updated, and therefore, it would be very difficult to repeat this analysis, even from day to day. Assuming that the extended index can change, the data set could be altered and therefore the results presented here could not be reproduced.

[3] The United Kingdom is considered here in its component parts. Addresses are provided as England, Scotland, Wales, and Northern Ireland, and each is handled as a separate political unit for the purposes of this analysis.

understating of international linkages. A level of uncertainty is inherent in this data. As Melin and Persson (1996) have suggested, we can hope that "significant scientific collaboration leads to co-authored papers in most cases, the main reason being the priority claims of the scientists involved."

Counting and representing the patterns of linkages can be done in several ways. We agree with the finding of Luukkonen et al. (1993) that the "choice of measures for gauging the strength of the … relationships clearly affects the findings." (p. 33) Accordingly, we examined the SCI data using different methods. Although we used other methods to analyse the data, the findings presented here represent our assessment of the most illustrative ways to analyse and visualize the data when one is seeking to study network dynamics. We concur with Luukkonen et al., (1993) who suggest that it is important to use both absolute and relative measures if one wants to acquire a full picture of collaborative linkages. Absolute measures allow us to see the overall size of the network, as well as which countries are in the center and which are at the periphery of the network. We use these measures to view the network at the global and regional levels.

Relative measures normalize size and therefore provide insight into the strengths of links between countries. We also applied these measures in an effort to reveal the architecture of the network. Relatedness measures such as the Salton Index (i.e., the cosine) illuminate the architecture of the relationship of countries. The Salton Index can be applied using social network analysis, showing intensity of relationships as well as revealing the core group of countries operating at the global level. Multilateral measures take the global network into account based on the assumption that the expected numbers of linkages between countries occur in proportion to a country's share in the global network.

Factor analysis allows us to rotate the correlations and find commonalities and differences among countries that are not obvious from a visual inspection of the data, and which do not become apparent using (hierarchical) relatedness measures. Factor analysis presents clusters of countries with strong relationships as well as intensities of relationships visualized in a multi-dimensional space.

We use these three sets of measures: descriptive statistics of observed values, relatedness measures (illustrated through social network analysis), and factor analysis to assess the absolute and relative positions of countries in the global system. They allow us to gain a "snapshot" of what the global science network looked like in the two years studied. The ten year interval allows us to compare the two years to explore whether the network has changed over time.

Counting co-authorships

We began by collecting into a single data set all papers produced for the relevant year. Table 2 inventories this data set.

Year	Unique documents in SCI	Addresses in the file	Authors for all records	Internationally co-authored records	Addresses, internationally co-authored records	Percent internationally co-authored documents
2000	778,446	1,432,401	3,060,436	121,432	398,503	15.6
1990	590,841	908,783	1,866,821	51,596	147,411	8.7

Table 2. Data used to create international network of co-authorships

The first analytic task is determining how to attribute papers to participating countries. Previous efforts to count co-authorships have used fractional, links, and integer counting in order to compare performance of countries or to examine networks. Fractional counting attributes the numbers proportionally, so that the number of authors on any given paper reduces the share of each participating country.   A second way of counting is to identify the number of *links* represented among the countries involved, with

each bilateral relationship counting as "1." The normalization in terms of number of links is more common in network analysis (Newman, 2000). A third way of counting is integer (Luukkonen 1993) or whole/distinct count (Zitt et al., 2000) that attributes a count of "1" to each occurrence of authorship by a country created by the participation of researchers from that country. Table 3 shows how the different methods of counting affect the totals. Since we are interested in using counts to see how the network is developing between countries, we used integer counting.

During the counting phase, we applied a two-tiered analysis. The data was placed into an asymmetrical matrix containing raw occurrence data that recorded all countries across one axis, and all articles on the other. A count is placed in the corresponding cell created by this matrix. Table 4 provides an illustration. The digit in the cell represents the number of unique addresses for each country appearing in the address line of the respective article. This occurrence table was used to conduct factor analysis in SPSS to identify patterns within the global network.

Country / Method	USA	Switz	Scotland	Germany	France (2)
Integer	1	1	1	1	1
Fractional	0.16	0.16	0.16	0.16	0.32
Links	5	5	5	5	8

Table 3. An example of different methods of counting co-authorships

Table 4. A sample of data in an asymmetrical matrix of the occurrence of authorship in an internationally co-authored paper (SCI 1990)

6

We then converted the occurrence table into a binary matrix of only ones and zeros ("integer counting") in order to construct a symmetrical matrix of countries appearing on both axes, with the co-occurrence of addresses appearing in the corresponding cell. Table 5 provides an illustration. The co-occurrence table was also used to conduct social network analysis.

COUNTRY	AFGHANISTA	ALBANIA	ALGERIA	ANGOLA	ARGENTINA	AUSTRALIA	AUSTRIA	BAHAMAS	BAHRAIN	BANGLADESH	BARBADOS	BELGIUM	BELIZE	BENIN	BERMUDA	BHUTAN	BOLIVIA	BOPHUTHATS	BOTSWANA	BRAZIL	BRUNEI	BULGARIA	BURKINA_FA	BURMA	BURUNDI	CAMBODIA	CAMEROON	CANADA
AFGHANISTAN																												
ALBANIA																												
ALGERIA						3						4										3						
ANGOLA																												
ARGENTINA						2	1					3					2			40		2						26
AUSTRALIA			3		2		16			3		20								10		2	3					179
AUSTRIA					1	16						27								7		4						53
BAHAMAS																												
BAHRAIN																												
BANGLADESH						3						4																5
BARBADOS																												11
BELGIUM			4		3	20	27			4									2	22		5			4			85
BELIZE																												
BENIN																											1	
BERMUDA																												2
BHUTAN																												
BOLIVIA					2																							2
BOPHUTHATSWANA																												
BOTSWANA												2																
BRAZIL					40	10	7					22															1	61
BRUNEI																												
BULGARIA			3		2	2	4					5																8
BURKINA_FASO																												
BURMA					3																							
BURUNDI												4																

Table 5. A sample of data in a symmetrical matrix of co-occurrences of authorship between countries in internationally co-authored papers, 1990

Once we collected all the data in the co-occurrence matrix for 1990 and 2000, we applied two types of analyses. Previous analysts have used a number of different approaches depending upon the question they were bringing to the data. Two measures have been used in scientometric analyses to achieve the weighted results (Luukkonen et al., 1993): 1) bilateral similarity measures, and 2) multilateral (pattern) similarity measures (e.g., Pearson correlations). The first method – bilateral similarity measures – we also found very helpful for our purposes. We used factor analysis for rotating the Pearson correlation matrices.

The bilateral measures weight the data in order to illustrate links between separate pairs of countries.   The resulting figures are derived from the observed number of shared papers C_{xy} of the countries x and y, weighted by the total number of papers, C_x and C_y. Table 6 shows these measures and equation.

Measure	Equation
Salton's measure	$S_{xy} = C_{xy}/\sqrt{C_x \times C_y}$
Jaccard's measure	$J_{xy} = C_{xy}/(C_x + C_y - C_{xy})$ *or* $C_x \cap C_y / C_x \cup C_y$

Table 6. Relatedness measures and equations considered to analyse the data

Luukkonen et al. (1993) state that "The Jaccard measure underestimates the collaboration of smaller countries with larger ones, but the Salton measure underestimates the collaboration of smaller countries with each other."   In a careful analysis, we find that Jaccard's measure does not serve us as well as the Salton Index. Jaccard's Index provides the intersection of the two countries as a percentage of the sum, while the Salton Index provides the intersection as a weighted percentage. But the difference is more than a factor two: whereas the Jaccard Index focuses on strong links in segments of the database (e.g., the strong relations between Croatia and Slovenia), the Salton Index organizes the relations geometrically so that they can be visualized as structural patterns of relations (Hamers et al., 1989 explores this in detail). Unlike the Pearson correlation, however, the Salton Index remains non-parametrical (Ahlgren et al., in preparation; Leydesdorff & Zaal, 1988). Thus, in keeping with Glänzel (2001), we have used the Salton Index as a measure of the networked relatedness of countries.

Social network analysis allowed us to explore and visibly depict the intensity and dynamics of interrelationships among researchers from different countries at the multilateral level. A social network is a collection of people, each of whom is acquainted with some subset of the others. (Newman 2000) We aggregate these social networks at the level of countries. Such a network can be represented as a set of points ("vertices") denoting people or countries, joined in pairs by lines ("edges") denoting acquaintance, or, in this case, co-authorship. Vertices that attract many links are called "hubs."   In order to visualize social networks of co-authorships across countries, we used Ucinet software to analyse the matrix of co-occurrences of authorship shown in Table 5. (For large networks such as this, Ucinet provides the algorithms that find the interrelationships and affiliations among networks.) Bringing the matrix of co-occurrences/co-authorships into Ucinet for analysis, and then into Pajek for the visualization of the networks, we sought to identify the clusters representing collaborations at a global level and within regions.[4]

5. Findings: Comparison of Collaborative Networks

Within each of the three types of analysis applied, the data support our initial expectations: between 1990 and 2000, the global network has expanded (more players are involved), and it has become more interconnected (more links occur between players). The cluster created by scientifically advanced countries has expanded, but some other nations (e.g., the Arab countries in the Middle East) are grouped into otherwise disconnected networks. At the regional level, networks have expanded (more players), become more decentralized (increasing number of hubs), and are more interconnected (more linkages). Clustering retains features related to geographical proximity and historical relationships, but these are no longer the strongest features affecting links. Overall, the data supports the expectation that the science system is further developing

---

[4] For information about the software and its applications, see http://vlado.fmf.uni-lj.si/pub/networks/pajek/ [last visited, December 2002]

into a network structure operating at the global level, one that is taking on dynamic characteristics.

5.a Observed Linkages within Global Science

At the global level, using observations of pairs of linkages, the network of interactions is shown to be very strong and highly interconnected. Figure 1 illustrates the linkages of all the countries of the world in 2000. (At this level, the figure does not differ markedly from one for 1990 so we don't include that figure.) The main point to note is that science is a highly interconnected network, with a dense core and a number of periphery countries. We will take this network apart into its regional and hierarchical components as we progress.
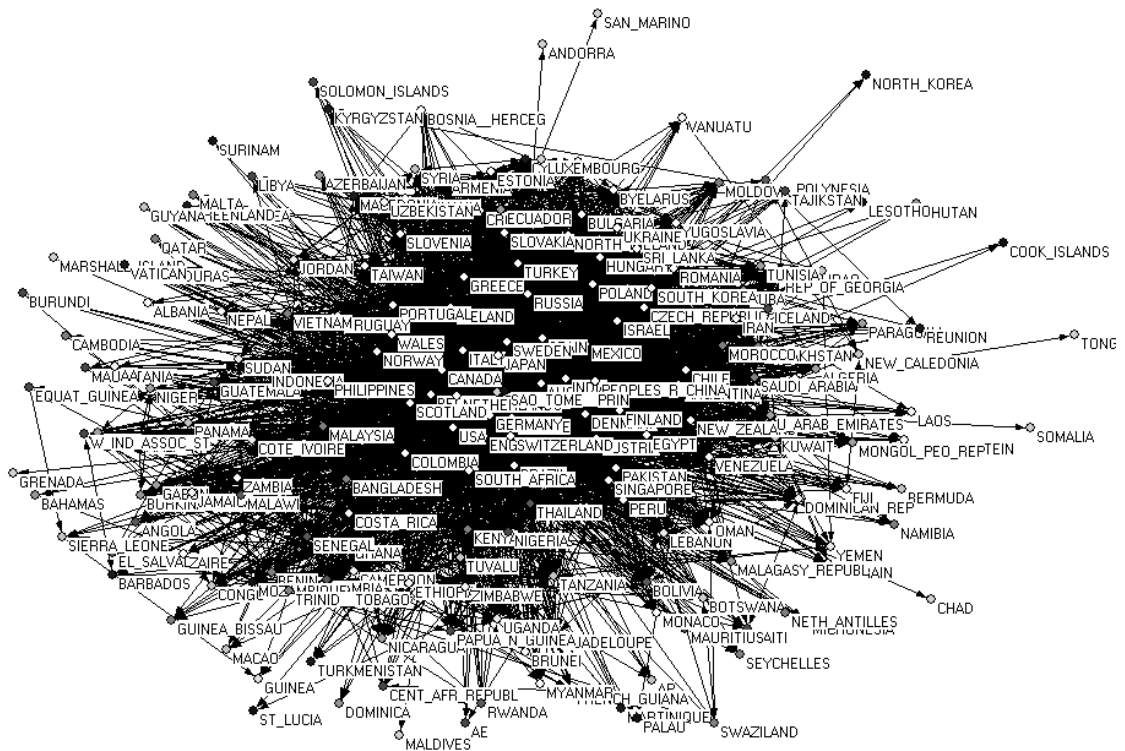
Figure 1. 199 Countries in a network of scientific co-authorship, 2000

Using the Pajek software tool (specifically the algorithm to find a core within a network) we analysed this network for core groups.[5] A large core of cooperating countries —a group that expands from 37 to 54 countries from 1990 to 2000—is shown in Table 7.

---

[5] A subset of vertices is called a *k*-core if every vertex from the subset is connected to at least *k* vertices from the same subset.

37 countries forming core set of international co-authorship relations in 1999	54 countries forming a core set of international co-authorship relations in 2000
	ARGENTINA
AUSTRALIA	AUSTRALIA
AUSTRIA	AUSTRIA
BELGIUM	BELGIUM
BRAZIL	BRAZIL
BULGARIA	BULGARIA
	BYELARUS
CANADA	CANADA
	CHILE
	COLOMBIA
	CROATIA
CZECHOSLOVAKIA	CZECH REPUBLIC
DENMARK	DENMARK
EGYPT	EGYPT
ENGLAND	ENGLAND
	ESTONIA
FINLAND	FINLAND
FRANCE	FRANCE
GERMAN FEDERAL REPUBLIC	GERMANY
GERMAN DEMOCRATIC REPUBLIC	
GREECE	GREECE
HUNGARY	HUNGARY
INDIA	INDIA
IRELAND	IRELAND
ISRAEL	ISRAEL
ITALY	ITALY
JAPAN	JAPAN
	LATVIA
	LITHUANIA
	MEXICO
NETHERLANDS	NETHERLAND
NEW ZEALAND	NEW ZEALAND
NORTH IRELAND	NORTH IRELAND
NORWAY	NORWAY
PEOPLES REPULBIC OF CHINA	PEOPLES REPUBLIC OF CHINA
POLAND	POLAND
PORTUGAL	PORTUGAL
	ROMANIA
USSR	RUSSIA
SCOTLAND	SCOTLAND
	SINGAPORE
	SLOVAKIA
	SLOVENIA
SOUTH AFRICA	SOUTH AFRICA
	SOUTH KOREA
SPAIN	SPAIN
SWEDEN	SWEDEN
SWITZERLAND	SWITZERLAND
	TAIWAN
	TURKEY
	UKRAINE
	URUGUAY
USA	USA
WALES	WALES
YUGOSLAVIA	YUGOSLAVIA

Table 7. The lists of the core set of mutually collaborating countries within the global network in 1990 and 2000

5.b Regional Networks

Studies of network architecture find that real networks are clustered, meaning that they possess local communities in which a higher than average number of people know one another. (Watts and Strogatz, 1998) Zitt et al. (2000) find that geographic links are strong in science. As a result, we decided to observe networks at regional levels to see if decentralization and expansion of the network clusters is occurring within smaller groups than those evidenced at the global level.[6] Our findings support our expectations: the regional networks have expanded (more players are participating) and they are more interconnected (more linkages). In general, the greatest growth in 2000 can be seen among countries that already had a robust network in 1990. Nevertheless, even among countries that had relatively weak or non-existent networks in 1990, there has been a development of the network. The only exception is countries in the region of the Middle East and North Africa, where no growth can be seen in the network, although there are more countries operating within the *core* of that network by 2000. (Figures 9 and 10)

Figures 2 and 4 show the interrelationship of the network of European Union and Accession Countries growing considerably between 1990 and 2000. The density of the network in 2000 becomes so concentrated that it is no longer possible to see lines connecting the most active participants. (If it were part of the European Union, Switzerland would be part of this central group, as well.) Figure 3 shows the 2000 cluster of the European Union, Accession Countries and adds Russia. (See also Figure 15.)
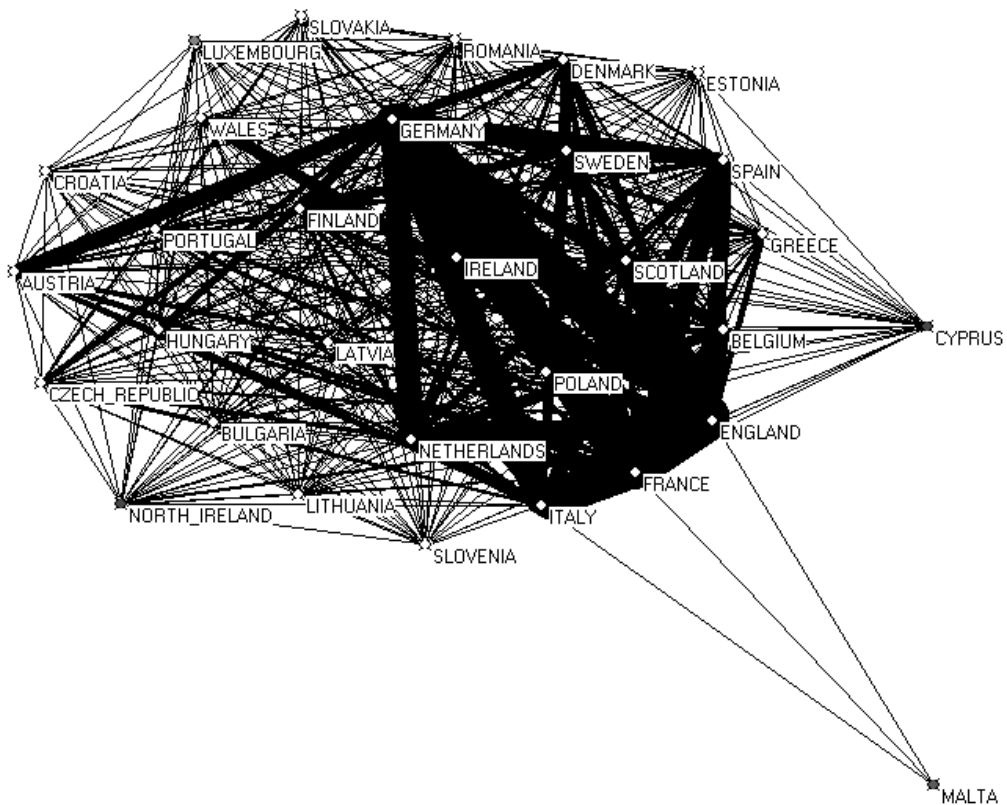
Figure 2. Network of 31 members of the European Union and Accession Countries co-authoring in 2000

---

[6] The figures in this section are derived using Pajek software. The lines are proportional—thicker lines represent a more intense relationship. A core group is always displayed as white vertices. Peripheral network members and non-participating countries appear as shaded dots.
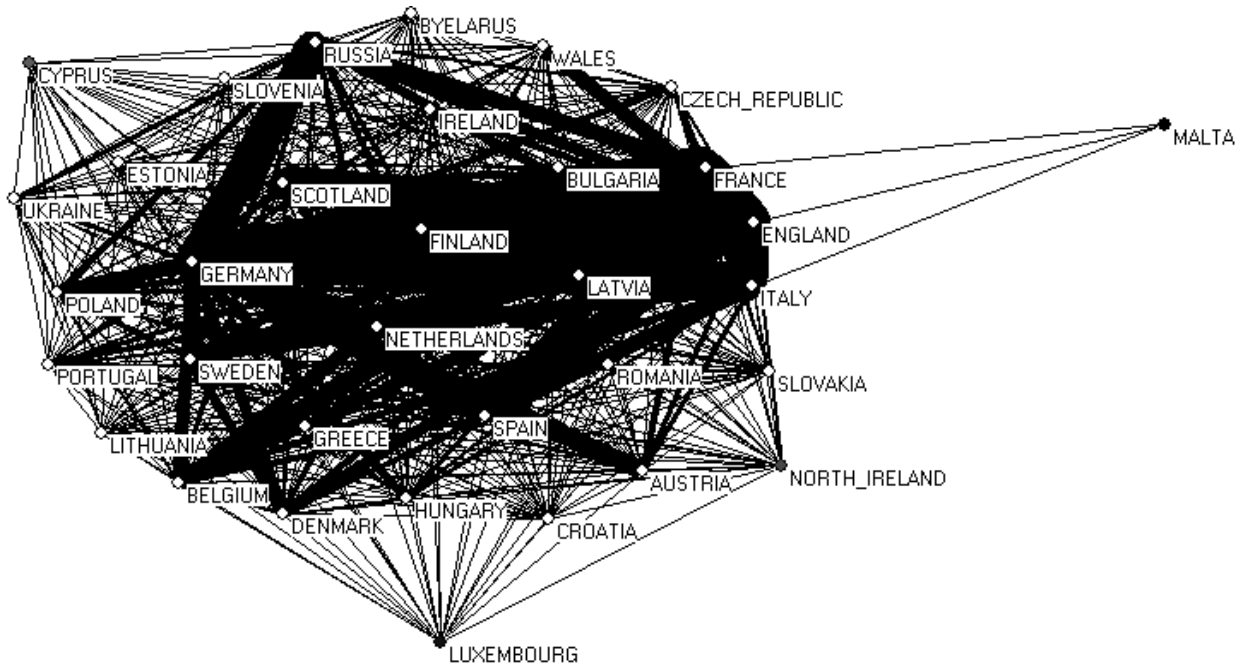
Figure 3. Network of European Union, Accession Countries and Russia, 2000
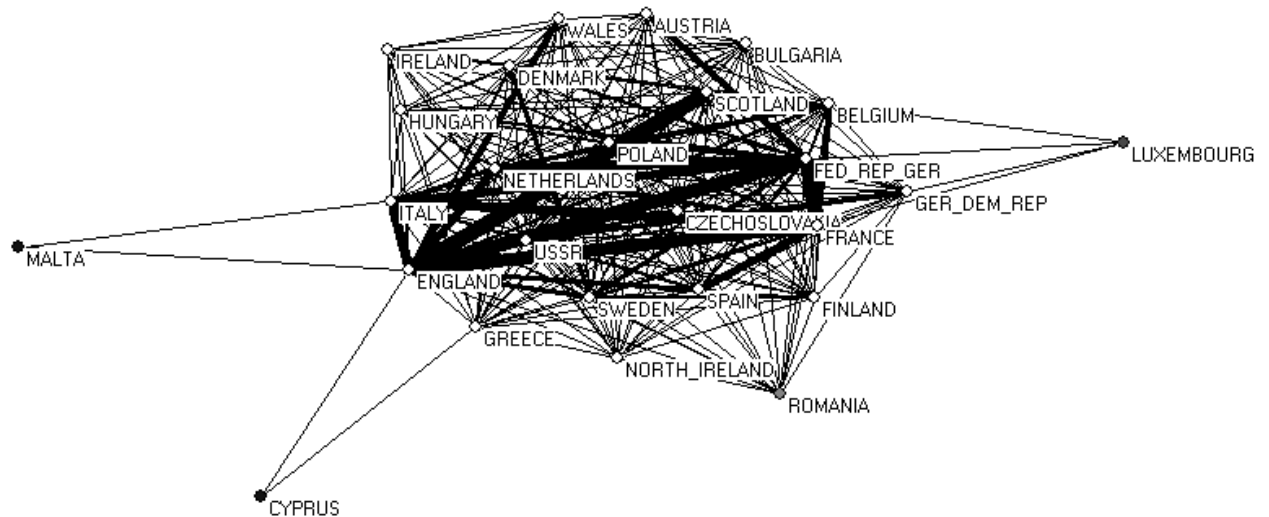
Figure 4. Network of European Union, future Accession Countries, and the USSR co-authoring in 1990

A network of scientists in The Americas also saw some change at the observed level between 1990 and 2000. Latin America has a much greater network of interconnection in 2000 than it had in 1990, both in terms of intensity of connection and in terms of the number of countries joining the network. (Not illustrated here.) Figures 5 and 6 illustrate the connection in the Americas in the two years examined. (See also Figures 16 and 17 for a view of the Americas with the United States and Mexico added into a relational weighting analysis.) The number of links between the U.S.-Canada core and the smallest countries in the network demonstrates the drawing power of that large science system.
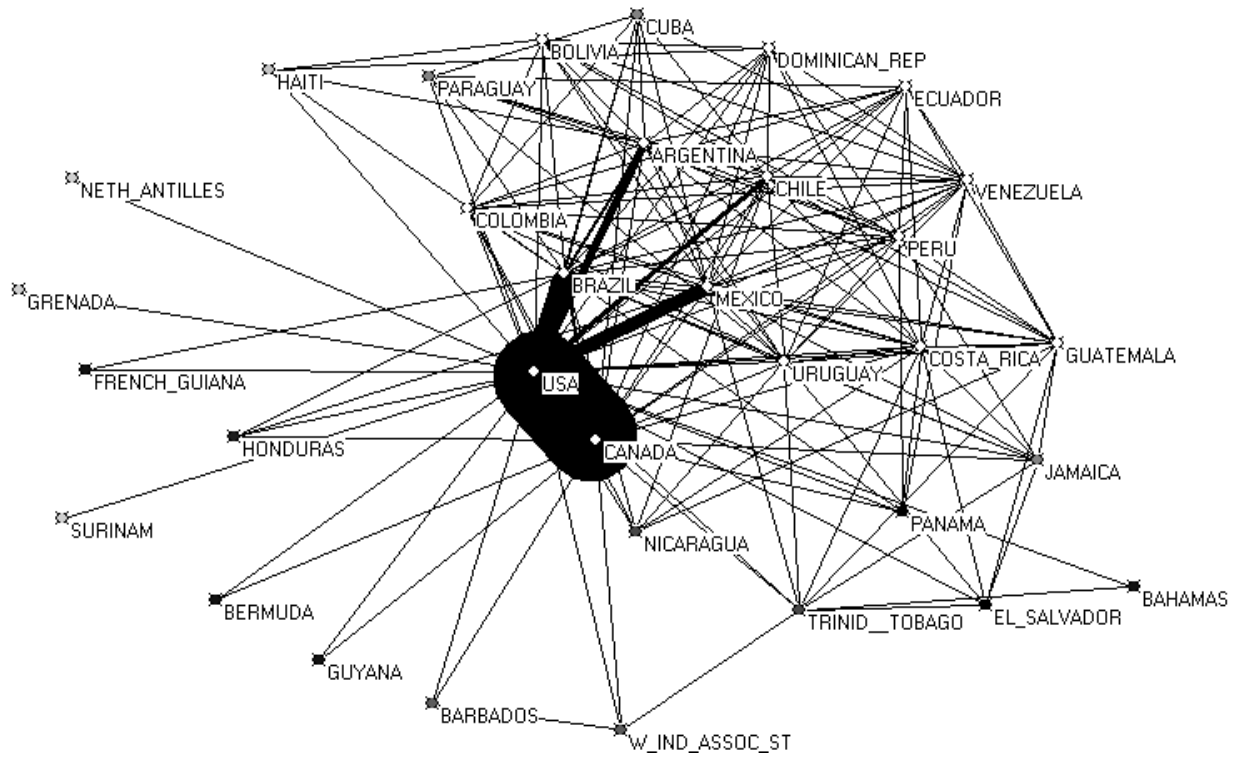
Figure 5. Network of 33 American countries co-authoring in 2000
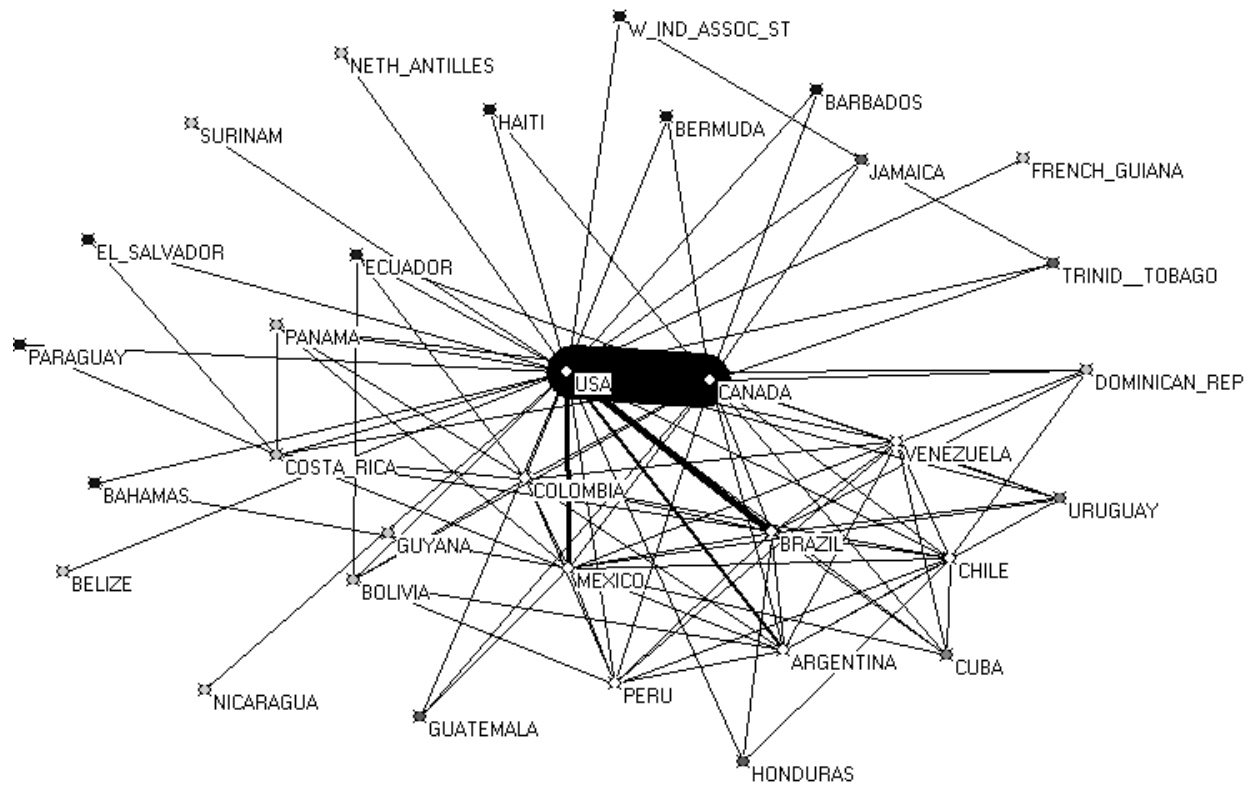
Figure 6. Network of 33 American countries co-authoring in 1990

Similarly, the countries of the North American Free Trade Agreement (NAFTA), the U.S., Canada, and Mexico, which have made an effort to increase collaboration, show a changed pattern between 1990 and 2000. The relationship among the three countries is dominated by the United States in 1990. (Not illustrated separately.) The network relationship shows the intensity of the relationship between the United States and Canada, which may be the strongest bilateral scientific relationship in the world. In addition, the link between the United States and Mexico grows considerably stronger over the decade, although there is little growth between Canada and Mexico.

The network of African co-authorships provides a compelling look at a region that has often been thought to be lagging in terms of participation in global science. Despite the lack of active participation of African scientists in some parts of the scientific world, figures 7 and 8 illustrate an active network in sub-Saharan Africa. Moreover, the network has grown in the decade between 1990 and 2000. More countries are participating in the network, and small hubs appear in 2000. In addition to Cote d'Ivoire—which seems to function as hub for Francophone collaborations—South Africa, Kenya, Nigeria, and Ethiopia have become more central. However, a number of publishing countries remain outside the network (e.g., Chad, Somalia).
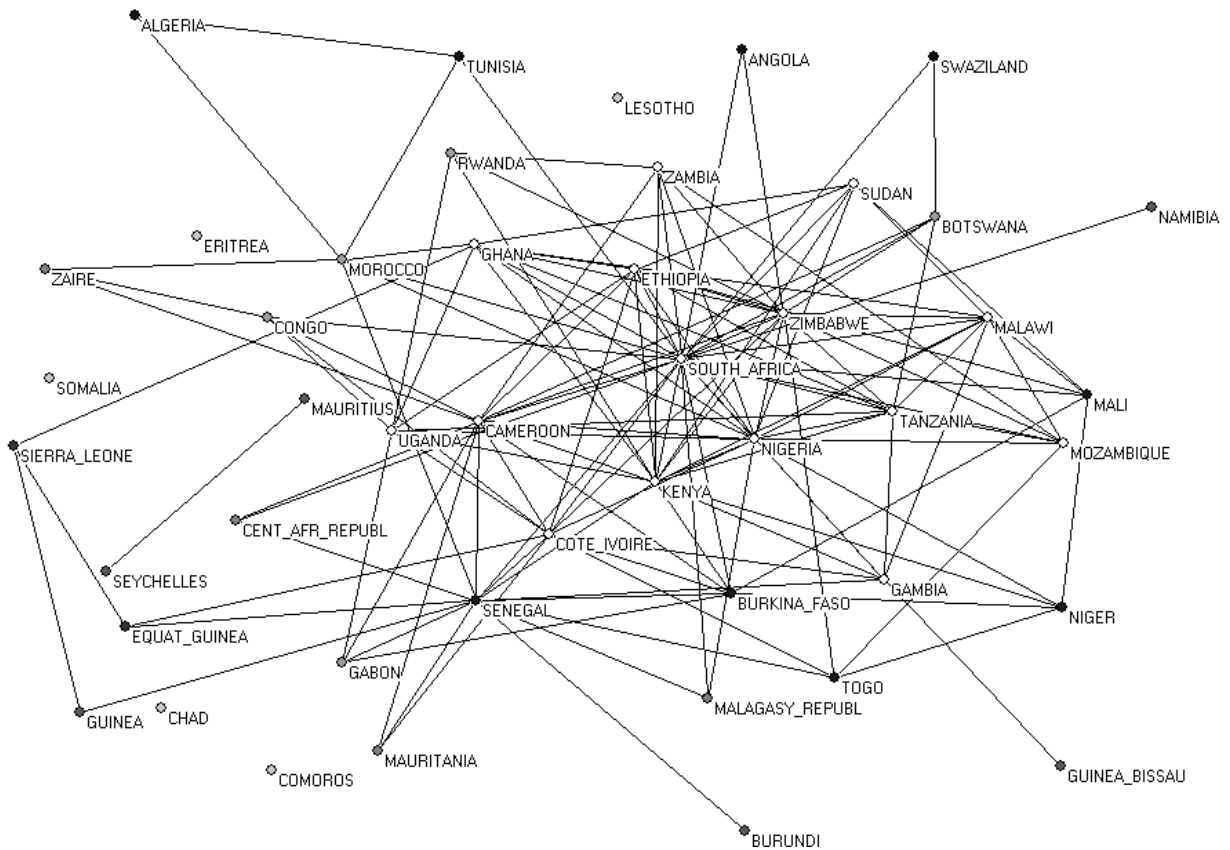
Figure 7. Network of 47 sub-Saharan African countries' co-authorships, 2000
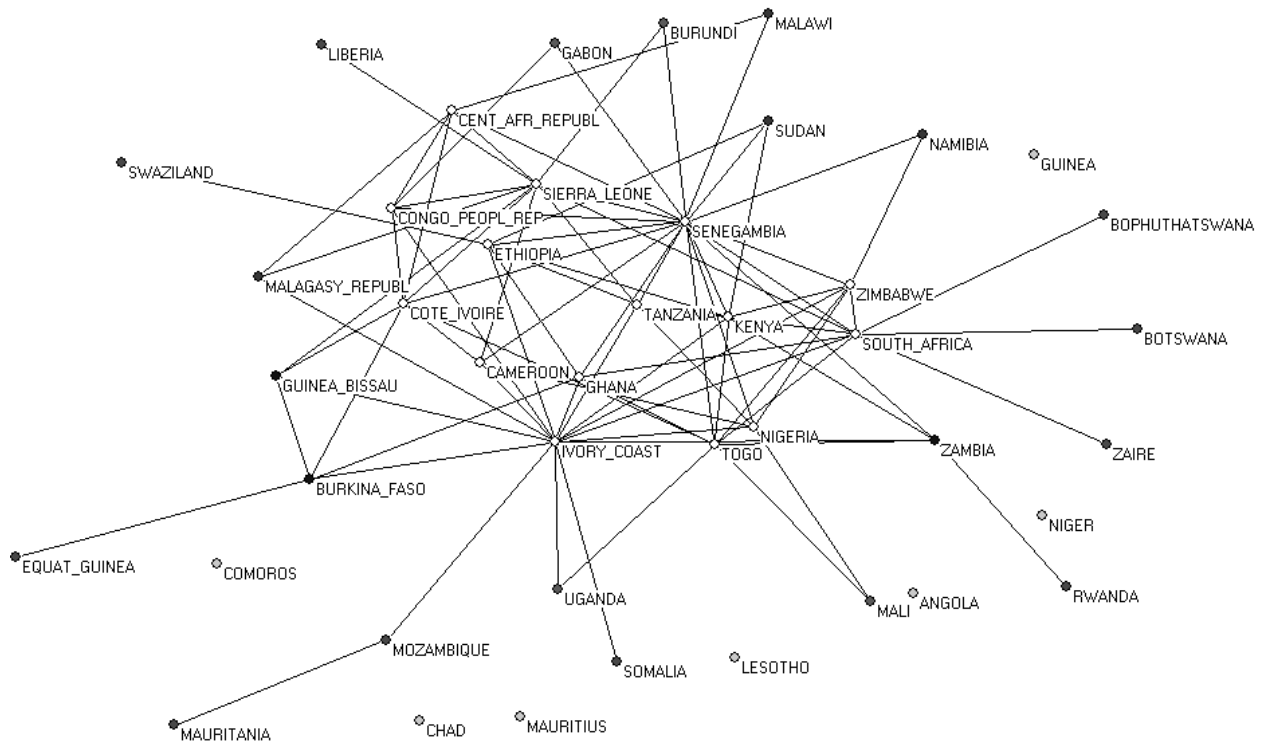
Figure 8. Network of 45 sub-Saharan African countries' co-authorships, 1990
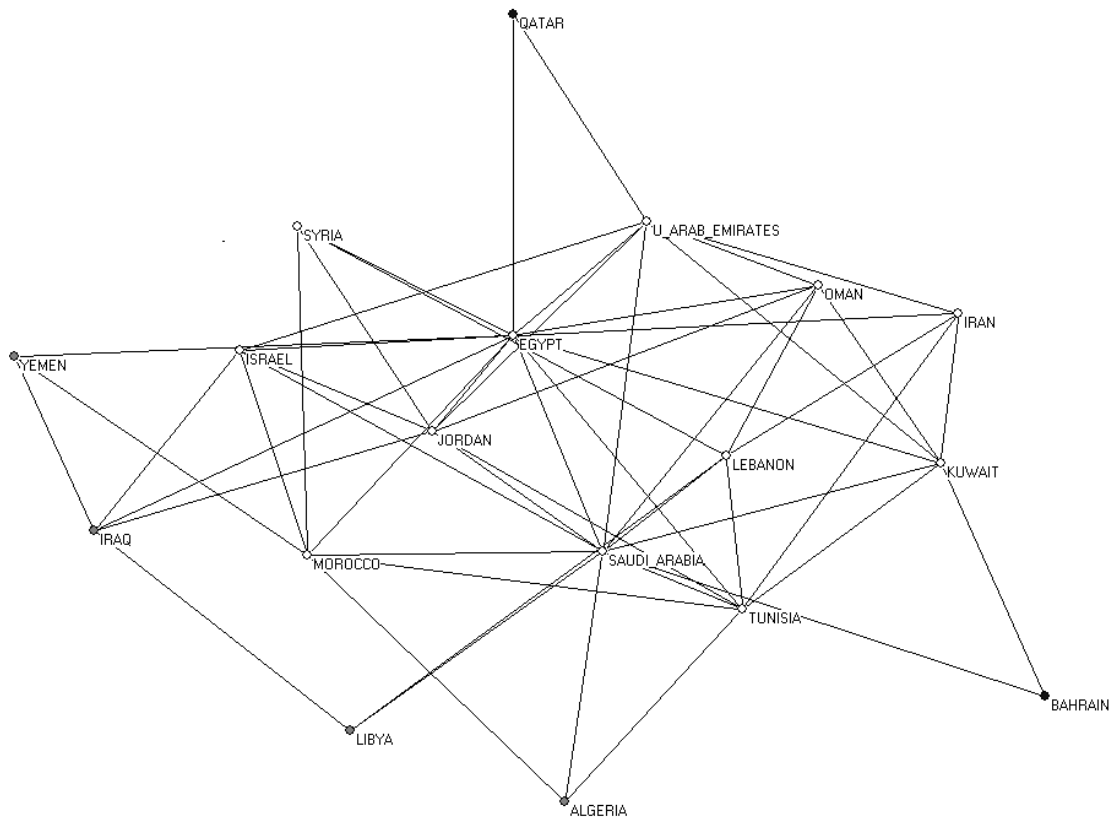
Figure 9. Network of 18 Middle Eastern and North African country co-authorships, 2000
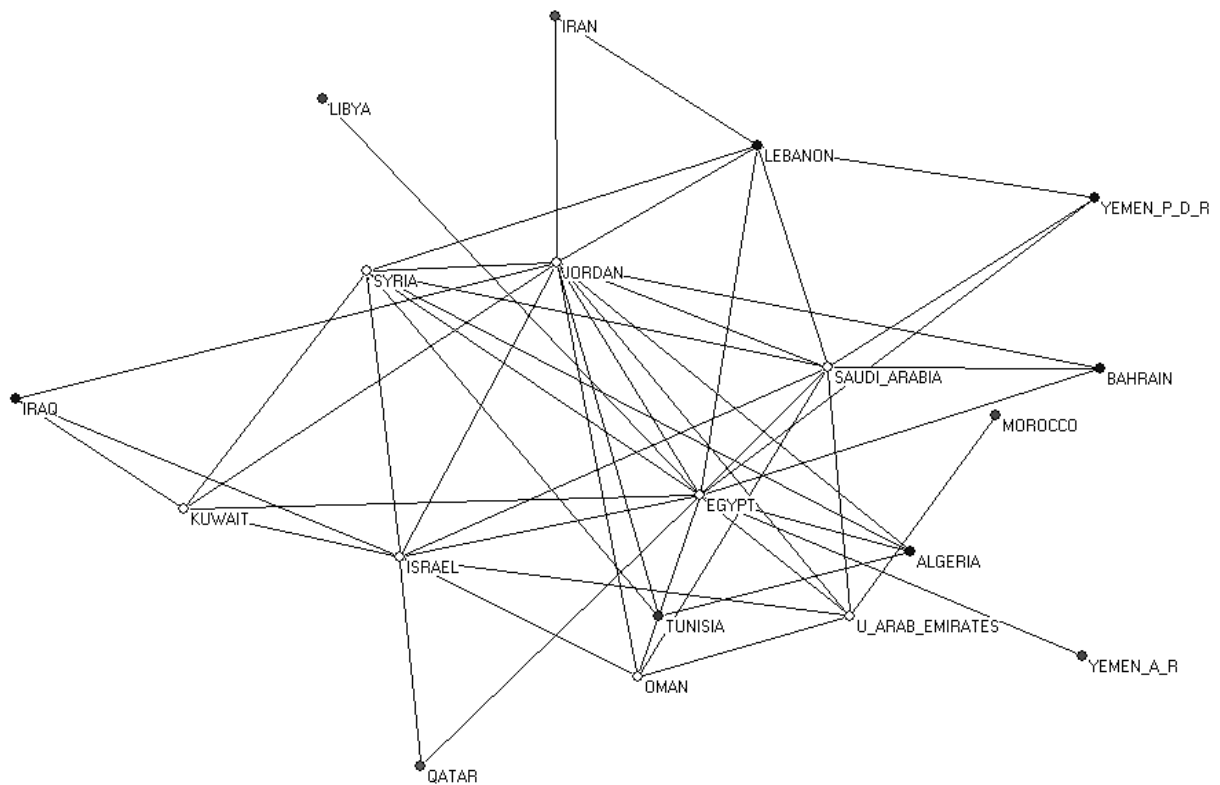
Figure 10. Network of 19 Middle Eastern and North African country co-authorships, 1990

The Asia-Pacific region has also seen an intensification of relationships and some shift in the role of the People's Republic of China and Japan between 1990 and 2000. Figure 12 exhibits a loose network in 1990, but one that had grown significantly by 2000, shown in figure 11, with the People's Republic of China taking a more integrated role as a hub in Asian science. In a pattern similar to Latin America, a number of smaller countries that were unconnected in 1990 have now joined the network, although several very small publishing nations are still unconnected (e.g., Tuvalu, Maldives, Micronesia).
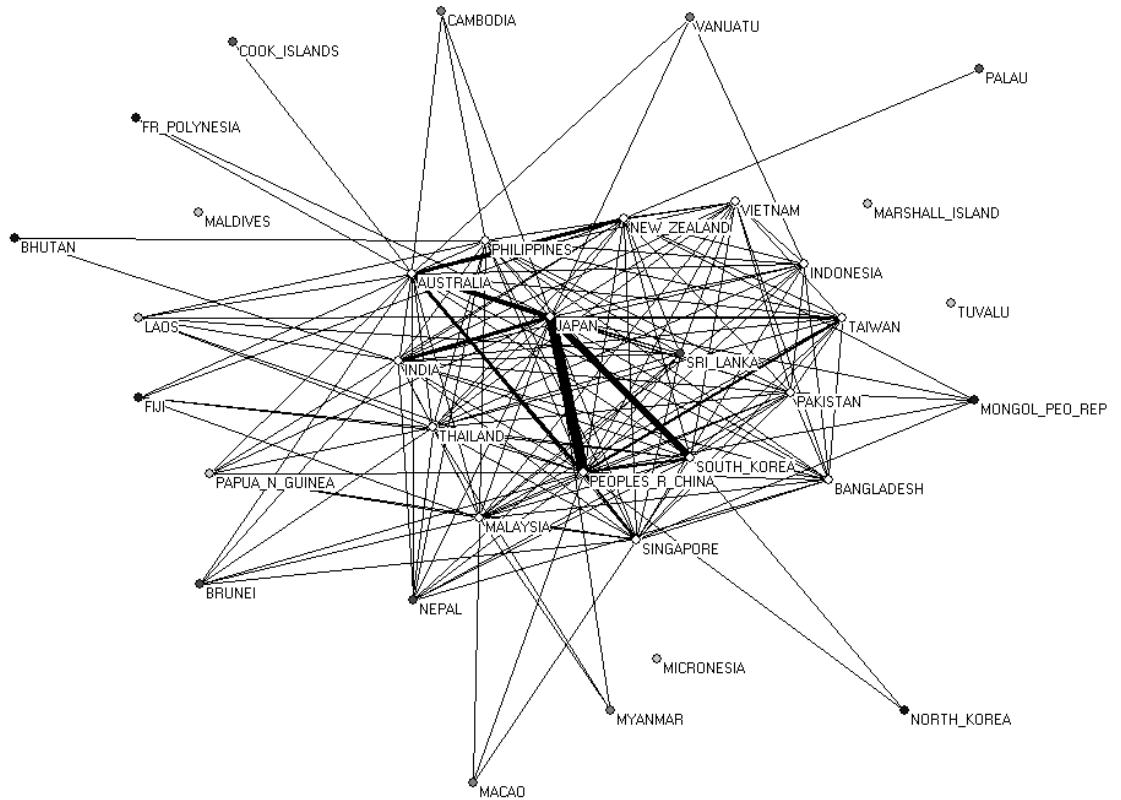
Figure 11. Network of 35 Asia-Pacific countries' co-authorships, 2000
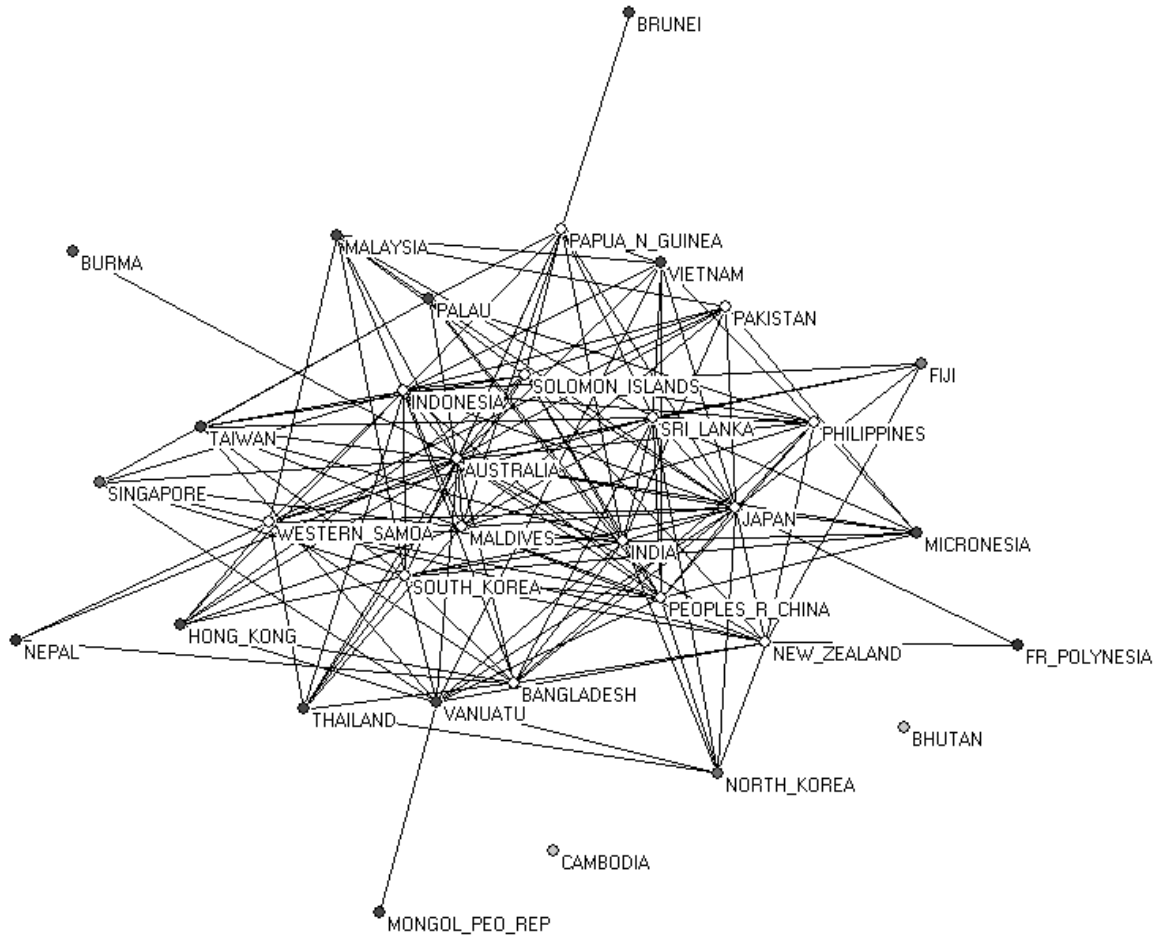
Figure 12. Network of 35 Asia-Pacific countries' co-authorships, 1990

5.c Relational Measures

Relational measures normalize the data and reveal the network architecture within the multilateral relationships. As discussed in the methodology section, we found the cosine or Salton Index to be the most illustrative of the overall hierarchy and structure at the network level, while the Jaccard Index focuses more on outlaying relations forming strong segments in the data.[7] The Salton Index normalizes for the size of the participants while retaining a measure of the volume of the linkages between them. We took the co-occurrence matrix and applied the Salton Index equation to it to create a fully indexed set of data for 1990 and 2000. Then, we exported the data into Ucinet and into Pajek to develop visualizations of the hierarchies and structures within the global network.

The results of the Salton Index (SI) analysis at the level of the core relationships among collaborative countries are revealing. By setting a threshold for relationships, it is possible to see that, even at a low threshold of SI ≥ 0.01, many countries fall out of the network. When the threshold is raised to 0.05, 42.1 percent of all countries fall out of the network in 1990; in 2000, the number is somewhat less at 35.4 percent. We are left with a component of 123 countries that are collaborating; 61 are bi-connected,[8] that is, they are collaborating with more than a single other nation (Table 8).

Year	Number of countries in the database	Countries related to at least one other country	Countries networked with more than a single other country
2000	198	128 (64.6%)	61
1990	178	103 (57.9%)	41

Table 8. Numbers of countries connected to the global network

Figures 13 and 14 show the networks for global linkages in 2000 and 1990 at the level of a Salton Index larger than 0.05. Six core countries relate to each other in a strong core set in 1990: England, the Federal Republic of Germany, France, Italy, The Netherlands, Switzerland, and the United States. In 2000, two additional countries join this group: Belgium and Spain.

Some major changes during the 1990s can be noted. By 2000, the relations of the former colonial powers with their overseas territories have faded away with the exception of the relations within the British Commonwealth. The Francophone nations form the core of an emerging African cluster. South Africa is not part of this cluster: it belongs to the group of advanced nations. The more diffuse picture in 1990 can also be explained by a secondary network centered on the USSR, as a subset within the core set. Note that, through Finland, this sub-network was also related to a Scandinavian cluster. The latter was at that time more oriented towards the U.S. than toward the E.U. member states.

In 2000, figure 13, the Eastern European part of the former Soviet group has now completely merged with the OECD set, particularly through Germany. This set of advanced nations also relates to a South American group, which is also more integrated than it was in 1990. The Caribbean and Central American networks, however, are separately organized in both years. The network of Arab countries has become more structured, but in 2000 this group is no longer related to the main grouping of more advanced countries.

---

[7] Zitt et al. (2000) use "expected/observed" as the normalization. The results of using this measure are very akin to using the Jaccard measure, but with even more emphasis on outliers (cf. Michelet, 1988).

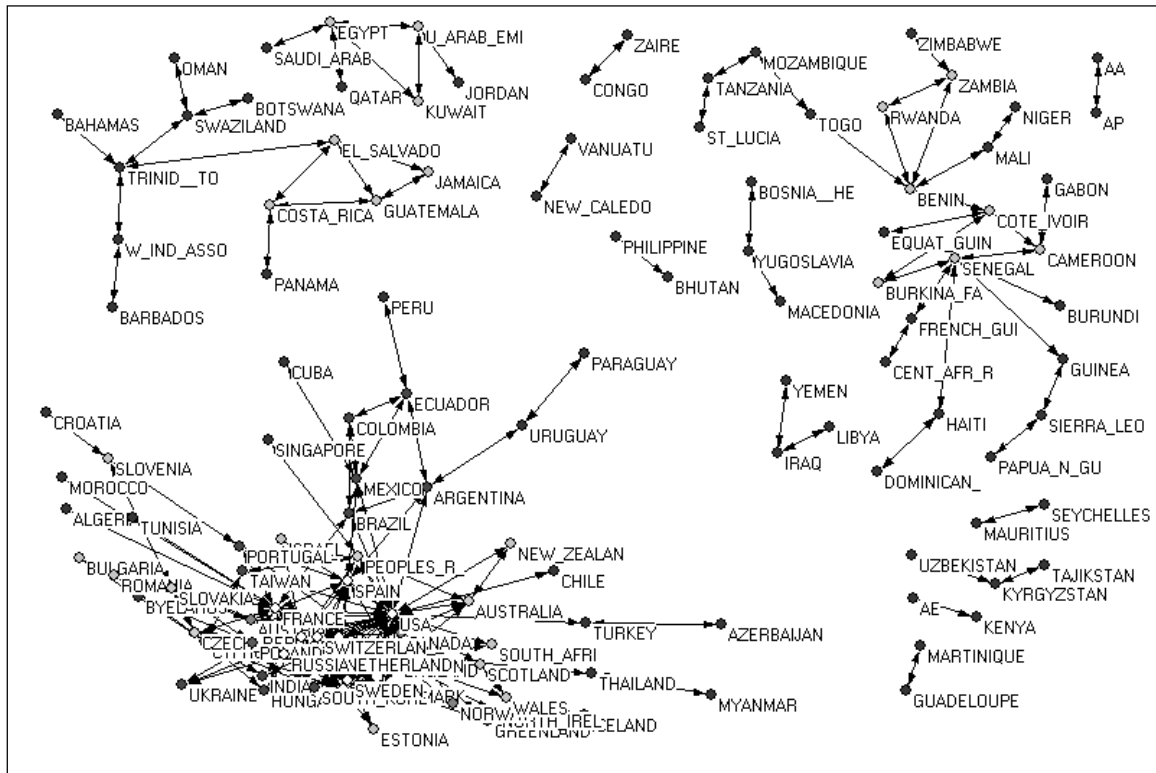[8] Bi-connected components can be considered hierarchically.

Figure 13. 128 countries collaborating in 2000 (Salton Index ≥ 0.05). (The cluster of eight core countries is indicated with white marks.)
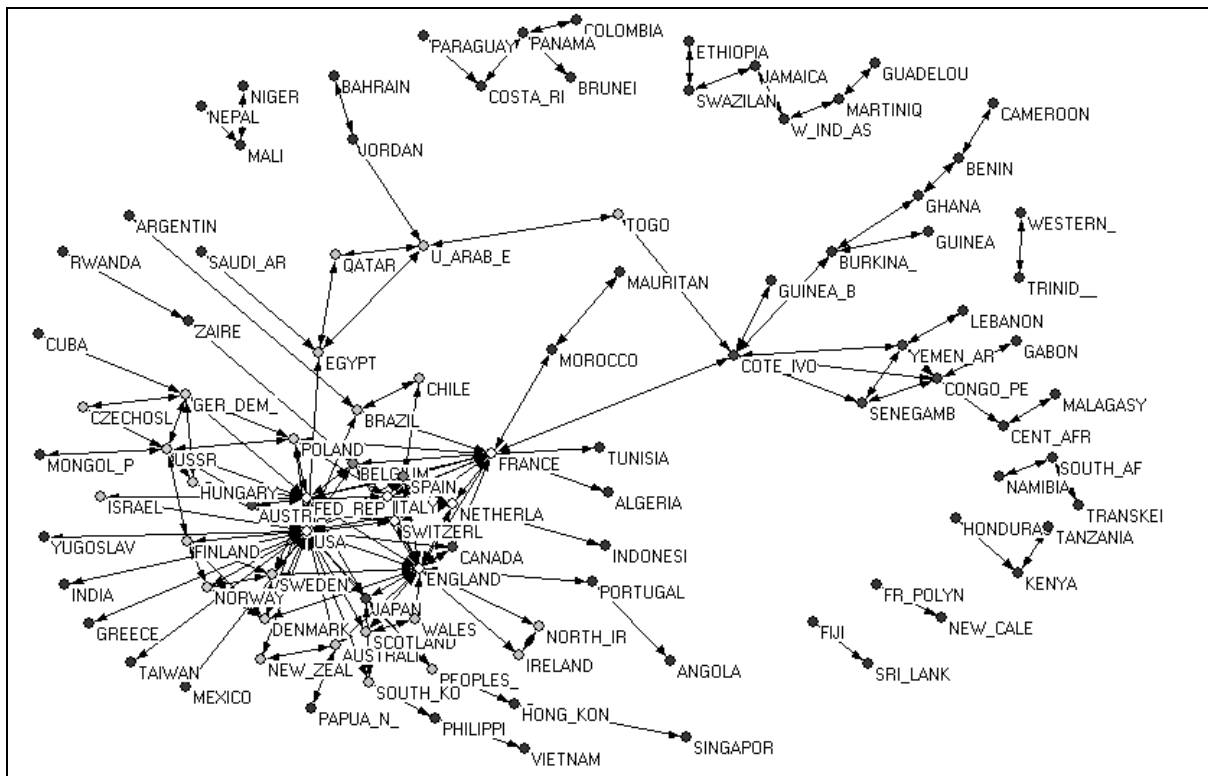

Figure 14. 103 countries collaborating in 1990 (Salton Index ≥ 0.05). (The cluster of six core countries is indicated with white marks.9)

---

[9] These visualizations are based on using the Fruchterman Reingold (1991) algorithm in Pajek (Batagelj and Mrvar, 2000).

In addition to revealing the core group of collaborating countries, the Salton Index allows us to view relationships that are particularly intense. We found, for example, that there are country pairs, usually geographically proximate, that have a very close relationship in science. These geographical groupings may otherwise be isolated.

It is instructive to examine the relational analysis at the regional level to put into perspective the earlier observed networks, figures 2 through 12. The European Union looks considerably different when examined using the Salton Index. Figure 15 exposes the hierarchy of the European Union and the Accession Countries, showing the hubs within this network at the 0.05 level. Note that Denmark and Sweden are part of the core when this is defined at the European level, whereas they were not at the global level.
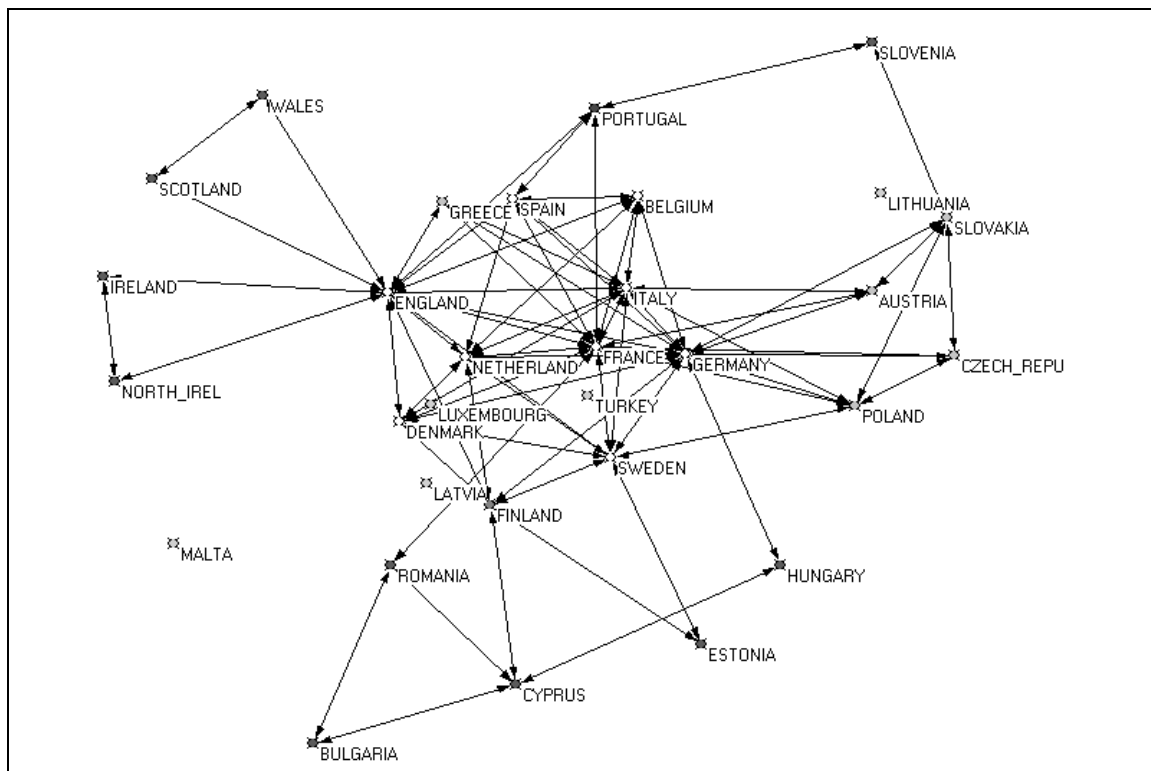
Figure 15. Core and related countries of the European Union and the Accession Countries, Salton Index ≥ 0.05, 2000[10]

Similarly, Figures 16 and 17 expose the hierarchy of weighted relations within Latin America when the United States, Canada, and Mexico are added as member of that network. Unlike the observed figures that show only regional links, when the United States is added into Latin America, it is possible to observe the drawing influence it has on collaborative opportunities within that region. Whereas the relation between Brazil and the U.S. was the single axis of collaboration in 1990, the pattern of collaborations has become more elaborate in 2000.

---

[10] The representations in Figures 15, 16 and 17 are based on the algorithm of Kamada & Kawai (1989) for the visualization. While in the previous analysis the objective was to distinguish among networks, we now focus on the inner structure of specific networks.
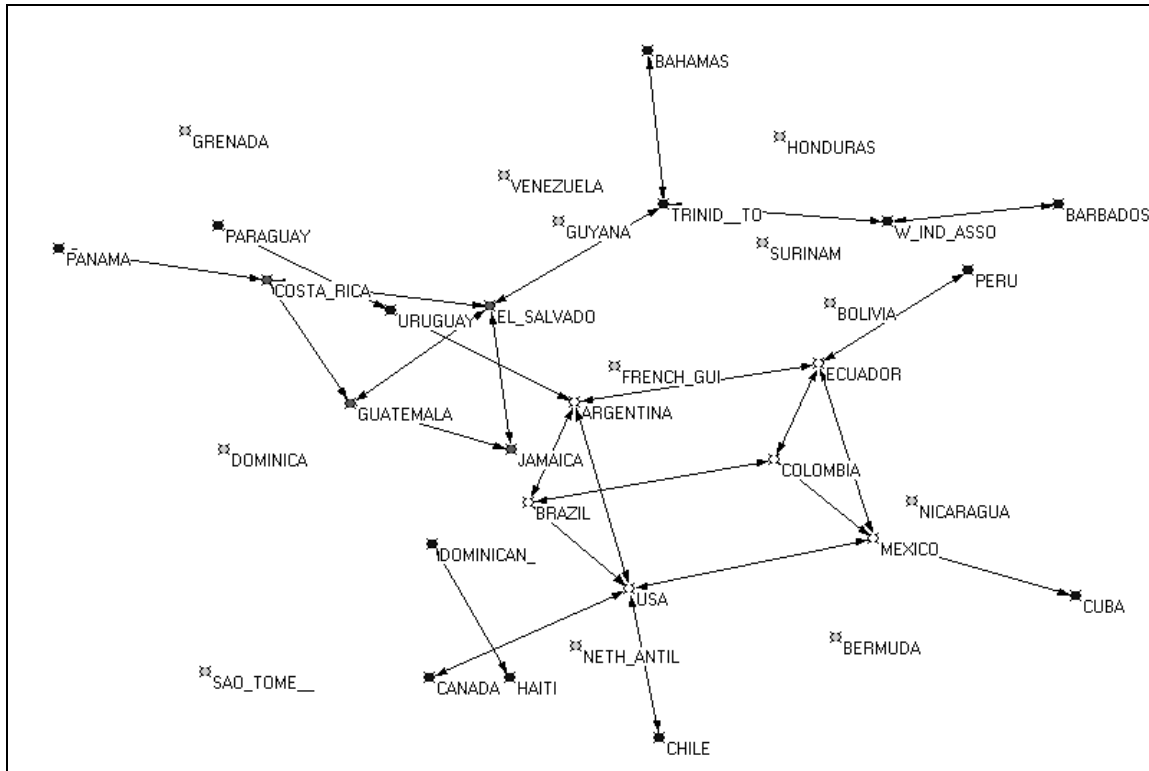
Figure 16.   Core and related countries of Latin America with the U.S., Canada, and Mexico included, Salton Index ≥ 0.05, 2000
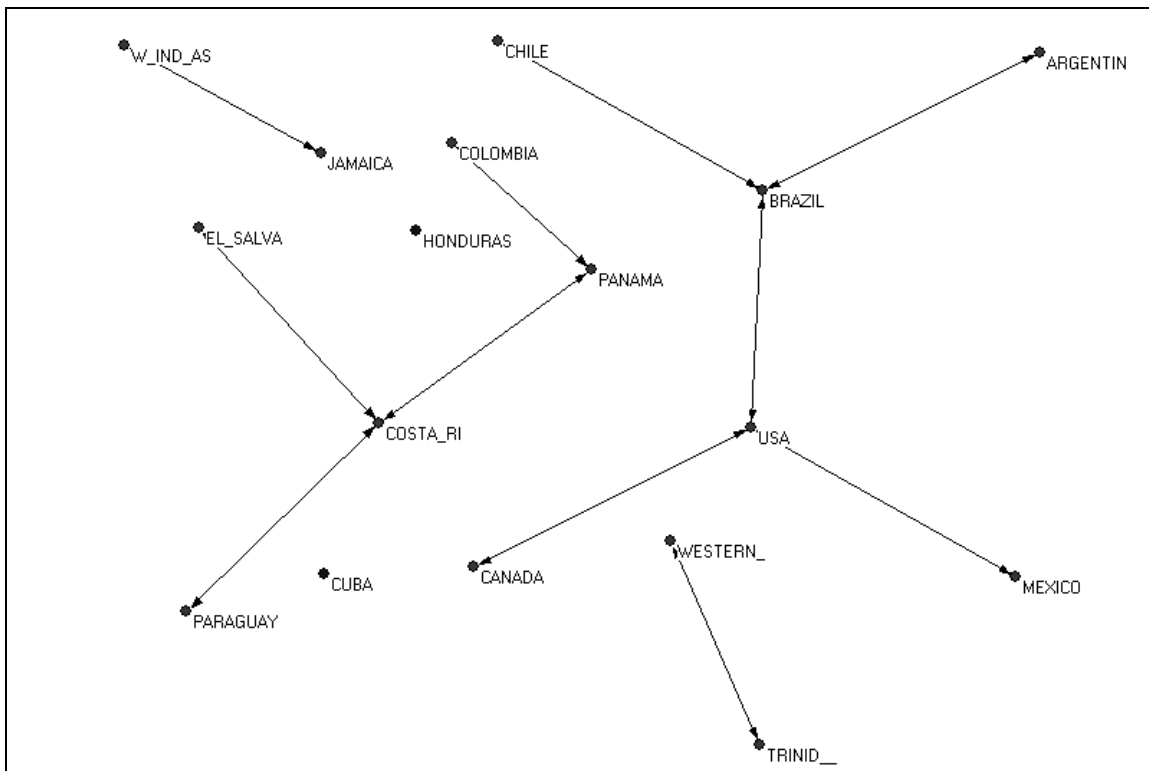

Figure 17. Core and related countries of Latin America with the U.S., Canada, and Mexico included, Salton Index ≥ 0.05,1990

5.d Factor Analysis

We applied factor analysis to the asymmetrical matrix of co-authorships created using the SCI addresses from the 2000 and 1990 CD-Rom data. Factor analysis allows us to find a commonality of relationships among the variables in the data. We factor analysed this matrix forcing different numbers of factors to reveal different information about the structure and architecture of the relationships. In comparison to the observed data presented above, factor analysis enables us to recognize structural properties of correlation and variation that are not observable by inspection of the matrix level.

Factor analysis reveals that countries group in five distinct clusters that reveal geographic proximity or historical linkages. Table 9 shows the breakdown of geographic groupings based on the solution for ten factors. These clusters are:

1) a U.S. dominated cluster representing both proximity in the Pacific region and historical/political ties.

2) A Scandinavian cluster including Sweden, Denmark, Norway, Finland, and Estonia.

3) A continental European cluster that has 4 sub-clusters.

4) A second European cluster including France, Spain, and Belgium that relates to the Francophone and Latin world, including Senegal, Cameroon, Argentina, and Morocco.

5) A British Commonwealth cluster.

	Component 1	2	3	
USA	*-0.593*			
JAPAN	*-0.33*		*-0.124*	
SOUTH_KOREA	*-0.254*	*-0.111*		
CANADA	*-0.242*			*U.S. dominated*
PEOPLES_R_CHINA	*-0.221*			
TAIWAN	*-0.187*	*-0.111*		
ISRAEL	*-0.102*			
INDIA	*-0.101*			
SWEDEN	0.234		-0.254	
DENMARK	0.211		-0.187	
NORWAY	0.206		-0.186	Scandinavian
FINLAND	0.176	-0.107	-0.221	
ESTONIA			-0.132	
PORTUGAL	0.178			**CONTINENTAL**
GREECE	0.17			**EUROPEAN**
GERMANY	0.174	-0.293	-0.106	
ITALY	0.175	-0.139		
SLOVAKIA	0.172	-0.162		
SLOVENIA	0.158	-0.123		**Middle Europe**
AUSTRIA	0.151	-0.132		
POLAND	0.125	-0.213		
CZECH_REPUBLIC	0.148	-0.165		
SWITZERLAND	0.143	-0.153		
RUSSIA		-0.304		
CYPRUS		-0.243		
ROMANIA		-0.215		
BULGARIA		-0.194		Eastern Europe
HUNGARY		-0.156		
UKRAINE		-0.146		
BYELARUS		-0.1		
ENGLAND	**0.298**	**0.51**		
SCOTLAND	**0.198**	**0.364**		
WALES	**0.135**	**0.268**		**European**
IRELAND	**0.12**	**0.149**		
AUSTRALIA		0.224		COMMONWEALTH
NEW_ZEALAND		0.15		
NORTH_IRELAND		0.145		
KENYA		0.124		
SOUTH_AFRICA		0.12		
ZIMBABWE		0.1		
FRANCE	**0.248**	**-0.148**	**0.461**	
SPAIN	**0.156**	**-0.1**	**0.187**	European
BELGIUM	**0.225**		**0.113**	
SENEGAL			0.253	
CAMEROON			0.252	LATIN WORLD
ARGENTINA			0.226	
MOROCCO			0.211	
Cote d'Ivoire, Mexico, Brazil, Colombia, Burkina Fasso, Ecuador, Benin, Tunisia, Cuba, Gabon, Algeria, French Guinea, Chile, Venezuela, Togo, and Niger have factor loadings larger than unity on the Francophone cluster.				

Table 9. Clusters of collaborating countries based on a Factor 2 level analysis

Figures 18 a and b and 19 a and b show the global network as a two and three factor structure for 2000 and 1990, respectively. On the above side of the picture in Figure 18 a, the Anglo-Saxon dimension is spanned between the U.S. and the various parts of the U.K. On the right side of the picture, one can observe the European dimension, gradually moving down from the U.K. to Germany and Russia. The third factor (exhibited in Figure 18 b) opens up a second dimension within the European domain between France on one side and the Scandinavian countries on the other. (Remember that the Scandinavian countries were in 1990 more connected to the U.S. than to the E.U. in the hierarchical core/periphery analysis above using the Salton Index.)

        In these representations, the peripheral countries occupy positions in the middle since individuals and institutions vary in their collaborations with the leading countries that span the network. This global net acts as a superstructure of international collaboration and competition to which the new entrants connect by creating linkages either directly or via intermediate hubs.
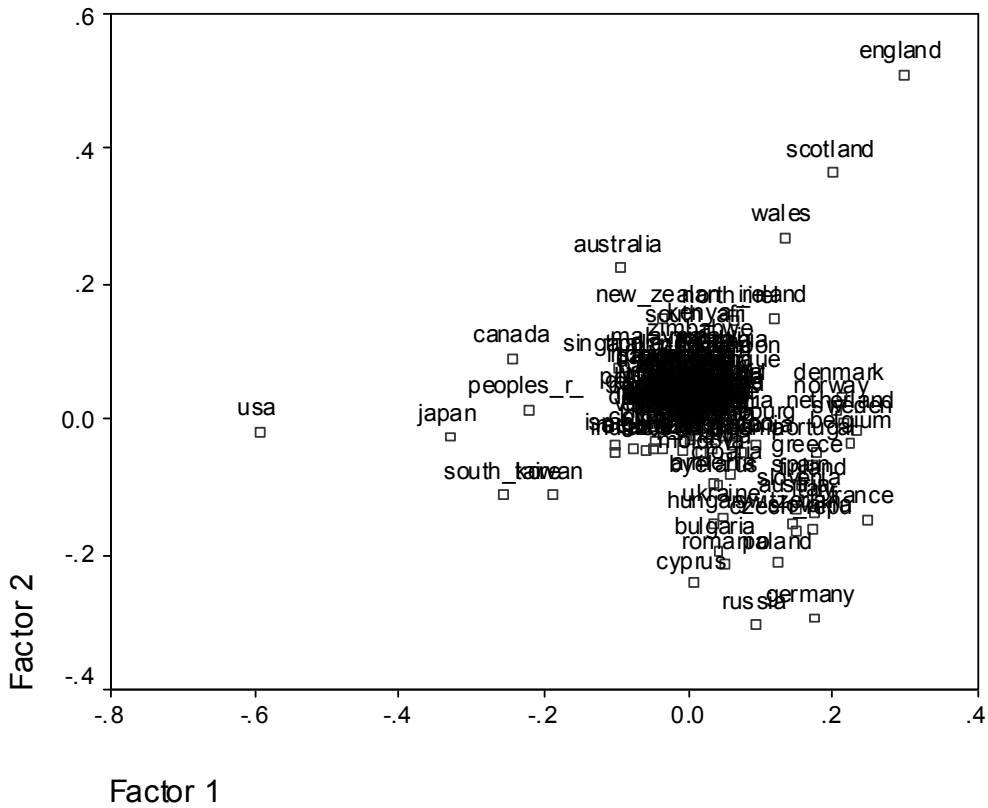
Figure 18 a. Factors One and Two used as dimensions for the mapping the rotated factor solution of the network of international co-authorship relations in 2000
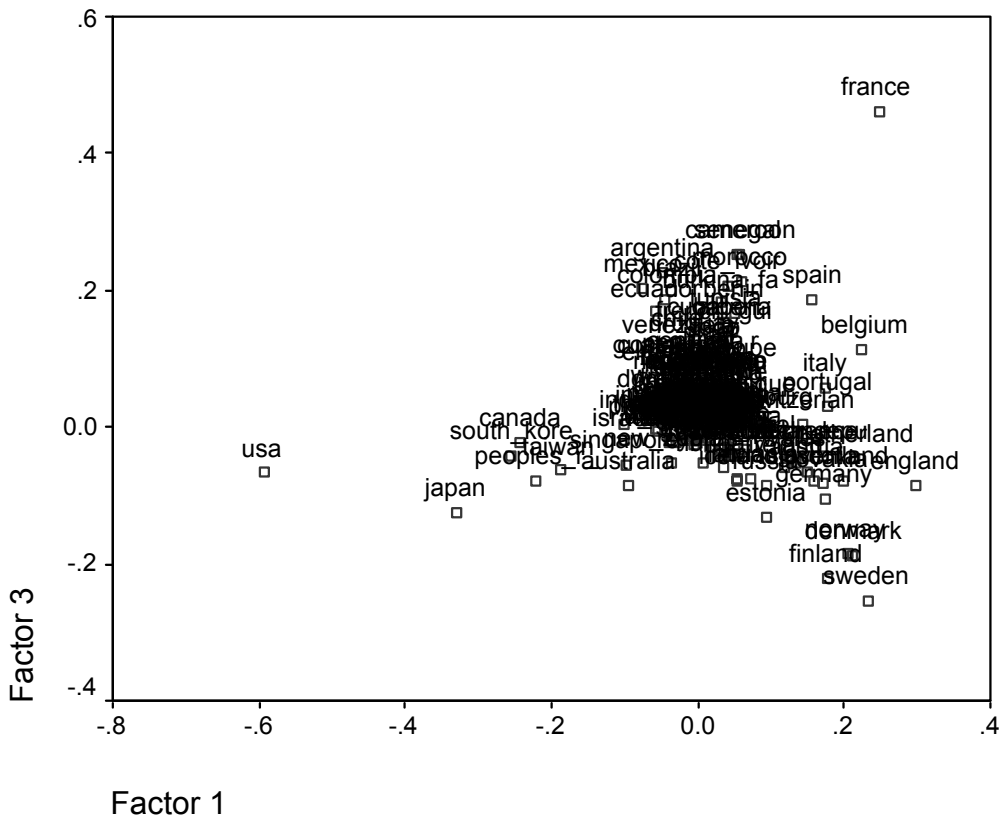
Figure 18 b. Factors One and Three used as two dimensions for the mapping of the rotated factor solution of the network of international co-authorship relations in 2000

Figure 19 a. Factors One and Two used as dimensions for the mapping the rotated factor solution of the network of international co-authorship relations in 1990

Figure 19 b. Factors One and Three used as two dimensions for the mapping of the rotated factor solution of the network of international co-authorship relations in 1990

The disappearance of the Soviet Union has contributed to a further integration in the relevant networks of international co-authorship relations. The *position* of the Soviet Union in the international arena in 1990 was taken over by Germany in 2000, see figures 18 a and 19 a. The European network, however, has become more complex. When using social network analysis for discovering a hierarchy (in terms of center and periphery) we found a strong cluster in terms of relations—including also the U.S. and Switzerland—but the factorial analysis shows that the core network is structurally differentiated and differentiating. The co-authorship relations with authors from Scandinavian countries, for example, have begun to play in the European theatre during this decade.

It is interesting to note that the U.K., the U.S., and Germany are counter-posed against one another in Figure 18 a and b for 2000. This suggests that when researchers are seeking an international partner, they chose among these three countries. In essence, these three large players (plus France as the fourth, and Scandinavia more recently as the fifth) compete with each other for partners in the global network. Partners that cooperate with authors in England are less likely to cooperate with the United States, for example.

The factor analysis reveals that the network is highly structured. Table 10, for example, exhibits the factor loadings on factor ten in a ten factor solution. An Arab cluster can be identified that loads exclusively on this factor with the exception of Malaysia that loads also on another (mainly Asian) factor.

EGYPT	.592
UNITED ARAB EMIRATES	.506
SAUDI ARABIA	.427
KUWAIT	.313
JORDAN	.249
QATAR	.231
OMAN	.147
IRAQ	.121
MALAYSIA	.109

Table 10. Factor loadings on Factor 10 in a ten factor solution

6. Observations and Conclusions

We set out to explore whether the increase in international cooperation, documented in a number of articles, is affecting the organisation of science at the global level. Specifically, we were interested in finding out if participation in global science is expanding to include more countries, and if researchers are more interconnected at the international level. The data suggest that, indeed, between 1990 and 2000, international co-authorship relations not only grew spectacularly in volume, but that this layer of knowledge production has become more pronounced as an interconnected structure. Expansion is shown by the increased number of countries linking into regional networks, providing evidence of a diffusion of scientific capacity to peripheral countries, and interconnection is shown by the growth in linkages between country pairs. The analysis further suggests that the network is becoming more decentralized, with regional "hubs" emerging (e.g., South Africa), and a strong core group of collaborating countries growing from 6 to 8 countries.

The increased volume of internationally co-authored publications seems to have reinforced emerging structures at the global level. The global level can be considered as providing increasingly a system of reference other than the national systems. However, this system is highly structured: The factor analysis reveals that some of the leading countries compete for co-authorship relations with less developed countries. For example,

we show that the U.S. has become an increasing partner for Latin American authors, somewhat to the detriment of the further development of regional relations.

Regional relations have remained strong where they have been based on cultural patterns, such as the Francophone community, and in collaboration between small neighboring countries like Holland and Belgium or Slovenia and Croatia. Increasing relationships are overwhelmingly observable among the member states of the European Union, but these nations are at the same time firmly embedded in a core structure that includes the United States and Switzerland as two non-E.U. partners in the hierarchy. Secondary networks like the one carried by the Soviet Union and its allies have faded away during the 1990s. Although a greater number of countries are connected (see table 7), some independent networks remain disconnected from the core structure in 2000.

In summary, this data show that the center-periphery model of international scientific collaboration can be replaced with a model that accounts for various centers that both collaborate among them and compete with one another for human resources from smaller national systems. The reputationally-controlled reward structure of science (Whitley, 1984) functions at the supra-national level and can thereby reinforce these developments. This supra-national level is internally differentiated, but at the same time hierarchically structured. A core group is both competitive and highly related. At the lowest levels of the hierarchy, segments of the world are not yet connected to the global structure.

7. Policy Implications

Political support is needed to create and maintain research institutions and to fund scientific research and development. As a result, science funding is aligned with political boundaries at the level of the *nation* where support can be garnered and where public budgets are large enough to allow for such investments. Funds are allocated out of public budgets in part based on the expectation of non-scientific benefits such as prestige (Nobel prizes) and economic spill-overs (innovation). (Wagner 2002)

A virtuous circle between political and scientific communities has been reinforced over the past 50 years as science has catalyzed innovation systems within developed economies. Large federal bureaucracies within the core countries have developed to manage the relationship between the political and scientific communities. A secondary set of publicly-funded programs has grown up to disseminate the results of science for economic application, often at the regional and local levels.

The rapid rise of global science has significant implications for this system. First it challenges the capacity of national bureaucracies to manage the funding of science in a way that ensures benefits accrue to taxpayers. (If science takes place on a geographically distributed basis, ensuring spill-overs will become increasingly difficult.) Second, it challenges the planning and allocation system within these bureaucracies, ones that have built up intricate evaluation and accountability measures to assure the political system that value is being created. Third, it challenges the identity of the practitioner within the political-scientific system to find a balance between the unwritten social contract (public funds to create appropriable goods) and the changing nature of knowledge creation (occurring across the globe; cf. Gibbons et al., 1994).

National systems that are structured to appropriate the results of science, with some adaptations, may be fixed to benefit from the increased output of knowledge at the global level. Nevertheless, to the extent that spill-overs occur locally (and there is literature to suggest that this is a significant amount of learning), a region may be poorly served by research that takes place at a distance. The challenge to policymakers and managers of science will be finding links between local needs and the knowledge created within the global network.

Nations that have never developed a national science system face a different challenge: does participation in global science require first building a *national* system? Or, can regions or nations pick and choose when and where to develop capacities in order to take advantage of the global system in some places, and to selectively build local capacities when spill-overs are particularly desirable? It would appear that the emerging dynamic is one operating at the local-global interchange rather than at the national level, perhaps making national entities a burdensome obstacle rather than a facilitator of participation in global science. These questions need further research.

Our findings suggest that global science is taking on the characteristics of a knowledge system with its own internal dynamics. Where previously, cross-border links in science were considered to be the extension of national systems finding opportunities to complement each others' capabilities, the emerging global network appears to be operating independently of and orthogonally to national systems of science. The national systems continue to operate, but perhaps at a lower level of efficiency, with resources being diffused in international networks. In that case, both the political and the scientific management of science will need to better adapt to the network of global science as they consider allocation of resources and appropriation of results.

References

Advisory Council on Science and Technology, 2000, Reaching Out: Canada, International Science and Technology, and the Knowledge-based Economy, the Expert Panel on Canada's Role in International Science and Technology (Industry Canada, Ottawa).

Ahlgren, Per, Bo Jarneving, and Ronald Rousseau (in preparation), Requirement for a co-citation similarity measure, with special reference to Pearson's correlation coefficient.

Batagelj, V., & A. Mrvar, 2000, Drawing genealogies, Connections 21, 47-57.

Beaver, Donald deB., 2000, Reflections on scientific collaboration (and its study): Past, present and prospective. Keynote lecture delivered at the Second Berlin Workshop on Scientometrics and Informetrics/Collaboration in Science and in Technology held in Berlin.

Ben-David, Joseph, 1971, The scientist's role in society: A comparative study (Englewood Cliffs, New Jersey).

Crawford, Elisabeth T., Terry Shinn (editors), 1993, Denationalizing science: The contexts of international scientific practice (Kluwer Academic Publishers, Dordrecht).

Doré, Jean-Christopher, Tiiu Ojasoo, and Yoshiko Okubo, 1996, Correspondence factorial analysis of the publication patterns of 48 countries over the period 1981-1992, Journal of the American Society for Information Science 47 8, 588-602.

Franck, Georg, 2002, The scientific economy of attention: A novel approach to the collective rationality of science, Scientometrics 55 1, 3-26.

Fruchterman, T., & E. Reingold , 1991, Graph drawing by force-directed replacement, Software--Practice Experience 21, 1129-1166.

Galison, Peter, 1987, How experiments end (University of Chicago Press, Chicago).

Georghiou, Luke, 1998, Global cooperation in research, Research Policy 27, 611-626.

Gibbons, M., et al., 1994, The new production of knowledge: The dynamics of science and research in contemporary societies (Sage Publications, London).

Glänzel, Wolfgang, 2001, National characteristics in international scientific co-authorship relations, Scientometrics 51 1, 69-115.

Hamers, 1989, Similarity measures in scientometric research: The Jaccard Index versus Salton's Cosine Formula, Information Processing & Management 25 3, 315-318.

Kamada, T., & S. Kawai, 1989, An algorithm for drawing general undirected graphs, Information Processing Letters 31 1, 7-15.

Katz, Sylvan and B.R. Martin, 1997, What is 'research collaboration?' Research Policy 26. 1, 1-18.

Kim, Ki-Wan, 2002, Global science, local research: The adoption and development of physics in South Korea, EASST conference paper.

Leydesdorff, Loet, 1995, The challenge of scientometrics: The development, measurement, and self-organisation of scientific communications (DSWO Press, Leiden University: Leiden); at http://www.upublish.com/books/leydesdorff-sci.htm

Leydesdorff, Loet, 2001, A sociological theory of communication: The self-organization of the knowledge-based society, (Universal Publishers: Parkland, FL); at http://www.upublish.com/books/leydesdorff.htm

Leydesdorff, Loet, and Rolf Zaal, 1988, Co-words and citations relations between document sets and environments, Informetrics 87/88, Leo Egghe and Ronald Rousseau (editors) (Elsevier: Amsterdam) 105-119.

Luukkonen, Terttu, Olle Persson, and Gunnar Sivertsen, 1992, Understanding patterns of international scientific collaboration, Science, Technology, & Human Values 17 1, 101-126.

Luukkonen, Terttu, R.J.W. Tijssen, O. Persson, G. Sivertsen, 1993, The measurement of international scientific collaboration, Scientometrics 28 1, 15-36.

Melin, G. and O. Persson, 1996, Studying research collaboration using co-authorships, Scientometrics 36 3, 363-377.

Michelet, Bertrand, 1988, L'Analyse des Associations, Thèse de doctorat (Université Paris VII: Paris).

Narin, Francis, 1991, Gobalisation of research, scholarly information, and patents - ten year trends, Proceedings of the North American serials interest group (NASIF) 6th Annual Conference, The Serials Librarian 21, Nos.2-3.

Newman, M.E.J, 2000, The Structure of Scientific Collaboration Networks, Santa Fe Institute Working Paper. Santa Fe Institute webpage publications.

Okubo, Y., J. F. Miquel, L. Frigoletto, and J.C. Doré, 1992, Structure of international collaboration in science: Typology of countries through multivariate techniques using a link indicator, Scientometrics 25 2, 321-351.

Price, D. de S., Little Science, 1963, Big Science…and Beyond, (Columbia University Press, New York).

Schott, Thomas, 1998, Ties between center and periphery in the scientific world system: accumulation of rewards, dominance and self-reliance in the center, Journal of World Systems Research 4, http://csf.colorado.edu/jwsr/.

Schott, Thomas, 1991, The world scientific community: Globality and globalization, Minerva 29, 440-462.

Schubert, A., W. Glänzel, and T. Braun, 1989, Scientometric datafiles. A comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields 1981-1985, Scientometrics 161-6, 3-478.

Shils, E., 1988, Center & periphery: an idea and its career, 1935-1987, 250-282, in L. Greenfield and M. Martin (editors) Center: Ideas and institutions (University of Chicago Press, Chicago).

Stein, Josephine, 1999, External relations in the European Union, the United States and Japan and international research and technological development cooperation, (Policy Research in Engineering Science and Technology, Manchester, United Kingdom).

Stichweh, Rudolf, 1996, Science in the system of world society, Social Science Information 35, 327-340.

Smith, David and J. Sylvan Katz, 2000, Collaborative approaches to research, HEFCE Fund review of research policy and funding, Final Report, (University of Sussex, United Kingdom).

Wagner, Caroline, Science and foreign policy: the elusive partnership, Science and Public Policy, December 2002.

Wagner, Caroline, Irene Brahmakulam, Brian Jackson, Anny Wong, Tatsuro Yoda, 2001, Science and Technology Collaboration: Building Capacity in Developing Countries? (RAND, Santa Monica).

Wagner, Caroline S., Allison Yezril, and Scott Hassell, 2000, International Cooperation in Research and Development: An Update to an Inventory of U.S. Government Spending, (RAND, Santa Monica).

Watts, D.J. and S. H. Strogatz, 1998, Collective dynamics of "small world" networks. Nature 393, 440-442.

Zitt, Michel, E. Bassecoulard, Y. Okubo, 2000, Shadows of the past in international cooperation: Collaboration profiles of the top 5 producers of science, Scientometrics 47, 627-657.

Index of Figures and Tables