Journal of Informetrics 10 (2016) 1025–1036

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

The world network of scientific collaborations between cities: domestic or international dynamics?

Marion Maisonobe*, Denis Eckert, Michel Grossetti, Laurent Jégou, Béatrice Milard

LISST, CNRS-Université de Toulouse, France

ARTICLE INFO

Article history:
Received 14 March 2016
Accepted 16 June 2016
Available online 16 September 2016

Keywords:
Scientific publications
Co-authorship
Research collaboration
Network of cities
Geography of sciences

ABSTRACT

An earlier publication (Grossetti et al., 2014) has established that we are attending a decreasing concentration of scientific activities within "world-cities". Given that more and more cities and countries are contributing to the world production of knowledge, this article analyses the evolution of the world collaboration network both at the domestic and international levels during the 2000s. Using data from the *Science Citation Index Expanded*, scientific authors' addresses are geo-localized and grouped by urban areas. Our data suggests that interurban collaborations within countries increased together with international linkages. In most countries, domestic collaborations increased faster than international collaborations. Even among the top collaborating cities, sometimes referred to as "world cities", the share of domestic collaborations has gained momentum. Our results suggest that, contrary to common beliefs about the globalization process, national systems of research have been strengthening during the 2000s.

1. Introduction

The global "growth of science", world-wide access to transportation, information, and communication technologies, as well as collaborative research policies, have encouraged international scientific cooperation. Together with the continuing spatial diffusion of scientific activities at the world level (Grossetti, Eckert, Gingras, Jégou, & Larivière, 2014; Inhaber, 1977), the increase of scientific collaboration is often described as one of the main features of globalization (Royal Society, 2011; Schott, 1993; Sexton 2012; Wagner, 2008). However, certain scholars using measures based on scientific publications show that the lion's share of scientific collaboration has remained domestic, that is to say intra-national (Frame & Carpenter, 1979; Georghiou, 1998; Hennemann, Rybski, & Liefner, 2012). Measuring the growth of scientific collaboration both within and across countries during the 2000s, and taking into account the share of intercity co-authorships, this article provides new evidence regarding the evolution of the world collaboration network. Showing that science is performed from a growing number of connected places, our work confirms the necessity to adopt a comprehensive approach of scientific activity.

Previous works used to focus only on the top publishing or cited urban areas in the world (Bornmann, Leydesdorff, Walch-Solimena, & Ettl, 2011; Matthiessen, Schwarz, & Find, 2010). They also used to limit their scope to certain macro regions such as Europe (Hoekman, Frenken, & Oort, 2009; Zitt, Barré, Sigogneau, & Laville, 1999). Here, we perform a spatial analysis

* Corresponding author at: LISST, CNRS-Université de Toulouse, Maison de la Recherche, 5 allées A. Machado, 31058 Toulouse Cedex 9, France.
E-mail address: marion.maisonobe@univ-tlse2.fr (M. Maisonobe).

http://dx.doi.org/10.1016/j.joi.2016.06.002
1751-1577/

of the co-authored articles, reviews, and letters extracted from the *Science Citation Index Expanded* with the much sharper spatial resolution of the urban level. We want to answer the question: what are the territorial dynamics underlying the world growth of interurban collaborations during the last decade?

After a review of existing literature on scientific collaboration focusing on the spatial dimension of world science, we present our spatial bibliometrics method based on the geocoding of publications and their assignment to urban areas. At the global scale, even if the overall production was still furnished by groups from one city in 2007, we demonstrate that there has been an increase in collaboration, both at the domestic and international scales. Country-to-country as well as city-to-city differences are discussed. It appears that the more developing is a country, the more scientists located in this country have favored domestic collaborations. To finish, we demonstrate that national systems have remained highly structuring, even when considering only the position of the top world cities.

2. The spatial stakes of scientific collaboration

Exiting literature on the geography of scientific collaborations has provided three different approaches within the spatial scientometrics framework (Frenken, Hoekman et al., 2009): the reasons for scientific collaboration, the proximity and size effects on the propensity to collaborate, and longitudinal tendencies of the world collaboration network.

Observing the increasing number of authors per publication, several scientists have tried to provide explanations. Following Gingras (2002) and Katz and Martin (1997), certain key results deserve to be highlighted. Alongside their historical bibliometric analysis, Beaver and Rosen demonstrated that co-authorship practices improved the world scientific productivity and visibility of the French elite (Beaver & Rosen, 1978, 1979). Focusing on international collaborations alone, Frame and Carpenter (1979) were the first to discuss country-to-country differences at a global scale and during the contemporary era. Their lead was followed by several authors, including Luukkonen, Persson, & Siverstsen (1992), who distinguished between internal (scientific) and external factors; the latter consisting of political incentives and the role of cheaper access to transportation and electronic communication. Apart from these "structural" reasons, Melin carried out a qualitative study showing that, above all, there are individual reasons for collaboration (Melin, 2000). Focusing on scientists' careers, it has also been shown that collaborations are sustained by interpersonal relationships (Cabanac, Hubert, & Milard, 2015). Bozeman and Corley (2004) have highlighted the role of research policies on collaborative practices and identified that research grants have a positive effect on more distant collaborations, even if "most researchers tend to work with the people in their own work group".

At the global scale, there are two ways of analyzing the spatial determinants of scientific collaborations: the first is investigating the role of invariant factors such as geographical distance and scientific weight on the propensity to collaborate, the second is considering geo-historical factors leading to special affinities between territories.

To our knowledge, the first scientists who adapted the gravity model to co-authorship data in order to identify the spatial constraints for scientific collaborations were members of the "Swedish Regional Science Mafia" (Andersson & Persson, 1993). The explanatory variables they identified for scientific collaborations were, in order of importance: scientific size (the publication weight per country), travel time, language similarity, and political unionization. One year later and independently, Katz (1994) was the first to measure the negative impact of geographical distance on university–university co-authorship links within (but not across) several countries. In his study, *Ceteris Paribus*, he showed that "the frequency of research collaboration between domestic universities in the United Kingdom, Canada and Australia decreases exponentially with the distance separating the research partners".

During the 2000s, efforts have been made to enrich the quantitative analysis of scientific collaborations' spatial determinants. In particular, institutional effects were taken into account together with those of geographical distance within a proximity framework. Thus, using a pseudo-regression model, Nagpaul (2003) separately considered the geographical, the thematic, and the socio-economic proximities between countries to explain the international collaboration rate per country. In 2009, Frenken et al. showed at several geographical scales that the "death of distance" theory (Morgan, 2001) did not hold true for scientific collaboration practices. Using gravity equations on three datasets, they demonstrated that in addition to scientific outputs (size effects), both the geographic distance (in kilometers or travel time) and the institutional proximity (boundary effects) are significant to explain the intensity of scientific collaboration measured during the 2000s: first between 36 countries in the world, second between 1316 regions in Europe, and third between 40 regions in the Netherlands (Frenken, Hardeman, & Hoekman, 2009; Frenken, Hoekman et al., 2009). Further, they found that the effect of distances has increased while that of boundaries has decreased between European regions suggesting a better integration of nations within Europe (Hoekman, Frenken, & Tijssen, 2010).

Since the 2010s, much progress has been made in processing the spatial information of bibliometric data at a higher level of resolution (Leydesdorff & Persson, 2010). Performing a bibliometric analysis of urban production has been worthwhile to find that the world scientific production is realized by an increasing number of cities (Grossetti et al., 2014). What is happening is that the previous monopoly of capital cities or historical university/research centers is, little by little, diminishing in almost every country in the world. Also exploiting geolocalization tools, Tijssen et al. have proven that the mean kilometric collaboration distance has increased globally during the 2000s whereas the share of international collaborations has leveled off (Tijssen, Waltman, & van Eck, 2012).

Drawing upon this last family of collaboration studies, the multi-level analysis approach we propose can be used to describe the evolution of the world collaboration network at an unprecedented level of geographical resolution: the urban area level (urban agglomerations, that is to say perimeters merging cities with their suburban areas). For the first time, we can distinguish globally between two kinds of interurban collaborations, domestic collaborations and international collaborations and wonder whether or not the growth of scientific collaborations between cities has been influenced by national dynamics and politics. This question is critical in the field of science studies since the national level of organization is challenged by scholars focusing on the globalization process (Wagner, 2008). Instead of viewing the scientific world as a "global village" or an "invisible college" of scientists, we consider the scientific world to be spatially organized and structured (following Hennemann et al., 2012). In this perspective, the individual unit of analysis is less consistent than the city unit of analysis to study collaboration dynamics at the world level.

3. Materials and methods

In order to measure the scientific activities of various cities, the more robust approach is to process the institutional affiliations (addresses) contained in bibliometric data.

3.1. The geocoding process and the building of scientific agglomerations

The quality of automatic geocoding tools (Google, Yahoo!, Bing, etc.) is actually widely divergent when used on a world-wide and spread over several decades dataset, such as the one we have used: the *Science Citation Index Expanded*. In this bibliometric database, the authors' addresses are decomposed in several fields, of which we selected three: locality, province, and country. Our target scale for geocoding was the locality level. Error control and correction was quite a long procedure, helped by the development of a user-friendly online visualization tool shared among all project participants. For instance, we detected an erroneous location in Southern Germany. The "Garching" text string, which refers to a suburb of Munich harboring a huge science & technological park, was first located by the geocoding tool in remote rural Bavarian area. A comparable problem was visually detected near Chicago (Argonne). A data-quality index was constructed by country, indicating the zones where expert verification was needed. The quality of the geocoding improved step by step. After more than a year of work, with the help of geospatial analysts and cartographers working in fields such as sociology and geography of science, we obtained a fine-tuned/high-resolution spatial database of scientific production over the last decade.

This granularity is itself a source of problems when attempting a comparative approach at the global level. The characteristics of postal addresses, the geographical variability of postal reference systems, and the great diversity of administrative geographical segmentation, prevent any direct comparison between the 18 650 distinct "scientific localities". Our group addressed this problem by merging localities into urban areas. The goal was to build spatially comparable geographical entities at the global level. Once all articles, reviews and letters were extracted from the *Science Citation Index* from two time series (1999–2001; 2006–2008) and geocoded, 10 730 urban perimeters were delineated and used as elementary analysis units to measure scientific activity.

The method we used to build those entities is described in more details in Grossetti et al. (2014). Our aim was to produce universal criteria, and not divisions corresponding to a juxtaposition of national criteria (for example, using Metropolitan Statistical Areas (MSAs) for the USA, urban areas in France, etc., and then comparing the results). For the purposes of this article, it is important to remember that we used a two-step method. First, the city perimeters defined around the 500 top publishing localities were obtained using a supervised procedure based on population density (highly fine-tuned raster data) and scientific production volumes. Second, for smaller urban entities, a simpler criterion of distance (a threshold of 40 km between two localities producing publications) allowed us to identify 'smaller' scientific cities. As a proxy for daily flows, physical distance was the more reliable definitional criterion we could use given the huge disparity of spatial information at the world scale of analysis. Our wish to use a consistent methodology across all cities in the world led us not to retain sophisticated methods that require detailed data on functional activities or commuting (e.g. Makse, Andrade, Batty, Shlomo, & Eugene, 1998). By the means of our method, the Parisian agglomeration includes suburban localities such as Gif-sur-Yvette, Villejuif, l'Université de Versailles-Saint-Quentin-en-Yvelines. Once the urban areas were delineated taking into account the localization of scientific activity, a final step was required: the selection of a counting method.

3.2. The counting method

To measure scientific activities it is first necessary to select a counting unit and a counting system. Second, it is advised to adopt a smoothing method based on averages in order to level short-term fluctuations, in particular spikes related to the variable periodicity of issues in scientific journals (Hennemann & Liefner, 2015). The method to be chosen varies according to the stakes of the research, the analytical scale, the dataset. Here, our dataset is one of the most reliable to obtain a global overview of the world science. The *Science Citation Index Expanded* (SCIe) covers international journals from at least 7 scientific fields (engineering, physics, biology, mathematics, chemistry, medicine, and science of the universe). It should be noted, however, that its coverage is biased toward Anglo-American journals as well as biomedical literature.

The counting unit we considered is the urban area instead of the address. This choice means that we are performing a "whole" instead of a "complete" count (Gauffriau, Larsen, Maye, Roulin-Perriard, & Ins, 2008). As a result, the participation of cities does not depend on the precise number of addresses per city per publication. Doing so, we are simplifying the intra-local information to focus only on the interurban activities. This counting method is more robust at the world level since the number of addresses per city partly depends on the degree of administrative fragmentation which is a country-dependent factor (Eckert, Baron, & Jegou, 2013).

Before applying the whole-normalized counting method on co-authorship data (from 4.2 to 4.3), we consider the geography of the entire scientific production (4.1). In Section 4.1, all scientific publications (even single-authored papers) are divided by types (local, intra-national, international). The aim of 4.1 is to assess the growing share of the interurban co-authored publications among the whole scientific production.

From Sections 4.2 to 4.3, the analysis focuses on interurban collaborations. In order to study collaboration networks and avoid double counts, co-authored publications are normalized using fractional counts. Fractional counting is the best way to avoid double-counts that would disrupt a multiscalar analysis. When studying a co-authorship network, the whole-normalized counting method evaluates each publication according to the number of urban areas it comes from.

For instance, if a publication is signed by scientists located in Paris, Villejuif, Toulouse, and London; only 3 urban areas are counted: Paris, Toulouse and London, since Villejuif is part of the Paris agglomeration. The collaboration volume between all pairs of cities will account for 1/3. In this instance, the share of domestic collaboration is 1/3, and the international share accounts for 2/3. By means of this fractioning method, it is possible to create any sum without losing the real number of co-authored publications issued at the world level.

The volumes of collaborations considered from Section 4.2 to Section 4.3 are based on this counting method. In Section 4.2, we sum interurban co-authorships to study the evolution of the domestic share of collaborations. In Section 4.3, a network analysis is performed on the collaboration network.

Before analyzing the data, a last methodological operation is required: time normalization. In order to smooth minute annual fluctuations in scientific activity, a normalized or moving average is computed over a span of three years. In order to compute a 3-year moving average for the year 2007, the proper formula is: $X_{2007} = (x_{2006} + x_{2007} + x_{2008})/3$.

To sum up, the results presented in the following sections are based on the whole scientific publications set (articles, reviews and letters only) indexed in the SCIe for 1999, 2000, 2001 and 2006, 2007, 2008. In order to compare the state of the world scientific system in 2000 and 2007, spatial analyses were used, which required the processing of over 8 million institutional addresses.

It would probably be wise to qualify our results as they reflect general trends in publication practices as well as the changing management of the bibliometric index performed by the Web of Science's administrators. Indeed, from 2006 to 2008, the database coverage has been enlarged in order to account for the scientific production of a larger number of countries. As a result, more national-oriented journals have been indexed in the SCIe. According to some bibliometricians, this evolution is unfortunate because it changes the philosophy of the source, which was primarily limited to international and mainstream scientific journals with the highest scientific quality and impact (Larsen & Von Ins, 2010; Zitt, Perrot, & Barré, 1998; Zitt & Bassecoulard, 1999). Since many national-oriented journals (mostly American) were actually already indexed in the database before the 2006–2008 expansion, one option is to focus on a smaller set of journals. This is what Leydesdorff et al. (2014) have done in their recent study on the highest cited publications. Focusing on this set, they observed that the weight of China is far less spectacular and that the highest cited publications are more often "international".

However, it is important to notice that even by adopting a restrictive approach, the current scientific dynamism of China and other developing countries do not disappear. Moreover, in comparing the scientific production share by country in 2007 using both the WoS coverage of 2000 and 2007, we have found no significant changes, apart for the share of Russia and the Netherlands which becomes higher with the 2007 perimeter (from 1.2% to 1.75%) and the share of the United States which becomes lower (from 26.2% to 25.6%). As shown by this test, when considering the production indexed in the WoS, the geographical structure which can be observed and the trends which can be measured are not significantly influenced by the change of perimeter (the 2006–2008 WoS expansion). The impacts of this change are more obvious at the level of certain disciplines such as social sciences and humanities that are not included in our analysis. Despite the WoS expansion, the SCI perimeter has remained a significant mirror of scientific cooperation dynamics.

4. Results and discussion

4.1. The growth of multi-city publications

During the last decade, how many publications came from a single city? How many came from several cities in the same country? How many came from several cities in several countries? In Table 1, designed to answer those questions, publications can only belong to one category at a time. When a publication is co-signed by groups from several cities, should one of the cities be located in a different country than the others, the publication is registered in the last category: "several cities, several countries".

At the world level, Table 1 shows a very unambiguous global trend: single-city publications decreased between 2000 and 2007 to the benefit of multi-city publications. The 4-point loss in the share of single-city publications benefited all other types of publications relatively equally: publications from several cities in the same country, and from several cities in different countries. Thus, there is no sign that the number of researchers who tend to seek collaborations abroad has increased over the past decade. The growth of international collaborations was part of the overall growth of intercity collaborations, whether or not they were in the same country. This latest growth was even slightly more pronounced within countries than between

Table 1
The growth of interurban collaborations between 2000 and 2007.

Share of scientific publications signed from (%):	2000[a]	2007[a]
One city	68.8	64.9
Several cities in the same country	15.3	17.3
Several cities, several countries	16.0	17.8
Total%	100.0	100.0
Total number of articles, reviews and letters	763 203	1 117 566

Source: SCI Expanded (articles, reviews, letters), 3-year moving average.
[a] Normalized counting (WNC), 3-year moving average. Source: SCI Exp.

them. Nevertheless, the share of intra-urban publications was still larger than that of interurban publications in 2007, which is in line with the scientific literature dealing with proximity effects.

Focusing on country-to-country differences in the next section demonstrates that even though all countries saw their share of single-city publications decline, the level of benefit for domestic and international collaborations varied substantially from one country to another. In order to go further in the analysis of interurban collaborations, we perform an analysis based on normalized data (fractional counts).

4.2. Variations in the evolution of international and domestic interurban collaborations

The following results are obtained by focusing on interurban collaborations only, that is to say on 35% of the total production in 2007 (386 255 publications are the result of interurban collaborations in 2007). In Table 2, the number of interurban collaborations is split between domestic and international. This table confirms that only a few countries internationalized between 2000 and 2007.

In most countries, the domestic share of interurban collaborations has increased faster than the international share. The reinforcement of intra-national collaborations is most obvious in emerging countries, more exactly countries which have been through major reconfigurations and upheavals during the last several decades (China, Brazil, India, Taiwan, Poland, Turkey, Greece, Czech Republic, Iran). Our hypothesis is that the reinforcement of national systems of research within these countries has led to more scientific autonomy for each. In other words, scientists in these countries depend less and less on international groups to publish in highly visible journals. Indeed, with the deconcentration process studied by Grossetti et al. (2014), opportunities for intra-national collaborations have been opened up.

Only six countries have followed a clear internationalization trend. In particular, we notice that the scientific autarky of the United States has declined during the last decade, following the internationalization trend identified from the 1980s by Frame and Narin (1988). Drawing upon Zitt et al. (1998) and Zitt and Bassecoulard (1999), we can consider this trend both as the result of the growing number of countries and cities involved in the world scientific enterprise and of the transition toward a trans-national model of communication (more and more Anglo-American authors willing to publish with foreign scientists in scientific journals indexed in the SCI during the 1990s). We observe that most countries where a trend toward the internationalization of scientific collaborations can be clearly evidenced are English speaking (Japan and Sweden excepted): United States, United Kingdom, New Zealand, and South Africa.

Russia is the only country where scientific production remained stable between 2000 and 2007 instead of growing (from 26 197 to 26 946 publications, only 749 or 2.6% more publications). This stagnation is an effect of the Russian scientific system's crisis since the collapse of USSR (Wilson & Markusova, 2004). In spite of this very weak progression, the evolution of the collaborative pattern of Russian publications is in line with the global pattern. In this country, the decline of single-city publications has led to an increase of multi-city publications whether intra-national or international.

Table 2 suggests that the integration of new production spaces into the scientific system had two consequences:

- older spaces (traditionally the most visible) have developed their international linkages. However further research is required to determine to what extent they tend to collaborate more with former partners across borders or to make contact with new ones;
- new production spaces or developing ones have reinforced their internal cohesion.

Our interpretation is that the growth of interurban publications has helped reinforce the internal cohesion of national scientific systems in emerging countries, whereas it has served international development in historically central countries, without diminishing their national cohesion. Indeed, the share of intra-national publications is on the rise in all countries.

Similar results are found if we focus on the top publishing cities in the world (Table 3). Few of these cities developed more international links between 2000 and 2007 whereas the majority saw an increase in their domestic share of collaborations. It appears that all "elite" cities follow their national trend (p-value ≤ 0.001); which suggests that even if a city belongs to the 'rich club' (according to scientific production), it does not mean the city is impervious to national logics (Table 3).

Table 2
Evolution of the domestic and international share of scientific collaborations at the country level.

35 most co-authoring countries (2007[a])	% of the country in the whole interurban co-authorship (2007[a])	Dynamic of the domestic interurban co-authorship (2000–2007)[b]		Dynamic of the international interurban co-authorship (2000–2007)[b]	
Czech Republic	0.5	**1.45**	↗↗	**0.85**	↘↘
Poland	1.0	**1.38**	↗↗	**0.85**	↘↘
Portugal	0.6	**1.36**	↗↗	**0.85**	↘↘
Greece	0.6	**1.35**	↗↗	**0.83**	↘↘
Hungary	0.4	**1.33**	↗↗	**0.90**	↘↘
Turkey	1.0	**1.30**	↗↗	**0.63**	↘↘
Iran	0.5	**1.28**	↗↗	**0.73**	↘↘
Brazil	2.0	**1.25**	↗↗	**0.69**	↘↘
Russia	1.4	**1.23**	↗↗	**0.90**	↘↘
Mexico	0.6	**1.20**	↗↗	**0.89**	↘↘
Israel	0.7	**1.18**	↗↗	**0.92**	↘↘
Belgium	1.1	**1.14**	↗↗	**0.94**	↘↘
Taiwan	1.5	**1.13**	↗↗	**0.74**	↘↘
India	1.7	**1.12**	↗↗	**0.84**	↘↘
China	7.3	**1.10**	↗↗	**0.82**	↘↘
Spain	2.6	**1.08**	↗	**0.94**	↘
Argentina	0.4	**1.07**	↗	**0.96**	↘
Denmark	0.8	**1.07**	↗	**0.97**	↘
Austria	0.7	**1.06**	↗	**0.99**	↘
Netherlands	2.1	**1.04**	↗	**0.97**	↘
Australia	2.0	**1.02**	↗	**0.99**	↘
South-Korea	2.1	**1.02**	↗	**0.97**	↘
Finland	0.8	**1.02**	↗	**0.98**	↘
Switzerland	1.4	**1.01**	–	**1.00**	–
Norway	0.6	**1.01**	–	**0.99**	–
Germany	6.8	**1.01**	–	**0.99**	–
Italy	4.0	**1.01**	–	**0.99**	–
Canada	3.3	**1.00**	–	**1.00**	–
France	5.0	**1.00**	–	**1.00**	–
United States	26.7	**0.99**	↘	**1.03**	↗
Japan	6.5	**0.98**	↘	**1.06**	↗
Sweden	1.4	**0.97**	↘	**1.02**	↗
United Kingdom	6.1	**0.90**	↘	**1.10**	↗↗
South Africa	0.4	**0.89**	↘↘	**1.07**	↗↗
New Zealand	0.4	**0.86**	↘↘	**1.08**	↗↗
Total of 35 countries	95.0	**1.03**	↗	**0.97**	↘
World total	100 (386 255)	**1.02**	↗	**0.97**	↘

Key: The% of domestic interurban co-authorship of Czech Republic have been multiplied by 1.45 between 2000 and 2007.
[a] Normalized counting (*WNC*), 3-years moving average. Source: SCIExp (articles, reviews, letters).
[b] % co-authorship 2007/% co-authorship 2000.

'World' cities characterized by an unambiguous trend toward nationalization can be found in countries where the scientific production has been deconcentrating the most over the last decade, in particular: China, South Korea, Spain and Russia.

Moscow is one of the most nationalizing 'world' cities together with Taipei (+11 points). Moscow's domestic share of external collaborations registered a 6-point increase between 2000 and 2007 at the expense of international collaborations. Contrary to Taipei, where half of the collaborations were already intra-national in 2000, Moscow used to account for the weakest domestic share of scientific collaborations in 2000 (23%). In spite of its huge size, Russia's scientific production was still very centralized at the beginning of the decade (Milard & Grossetti, 2006). This progression is critical but needs to be qualified because of the very low production growth in Russia between 2000 and 2007. Taipei is in a very different context than Moscow. According to us, the increase of Taipei's intra-national collaborations depends less on the island's higher scientific cohesion than on its growing isolation imposed by the dramatic development of continental China.

To summarize, in most countries, and even among the top publishing cities, the growth of domestic interurban collaborations has exceeded the growth of international linkages. In other words, there has not been any sizable and unilateral trend toward internationalization at the global level during the last decade. In addition, while the United-States are internationalizing, it should be highlighted that American cities have not been internationalizing at the same rate. Boston and Baltimore have developed their international links faster than Washington and San Francisco.

We expect those observations to have structural effects on the world collaboration network.

Table 3
Evolution of the national share of scientific collaborations at the city level.

The top publishing cities in 2007[a]	Dynamic of the domestic co-authorship (2000–2007)[a]		Dynamic of the international co-authorship (2000–2007)[a]	
Moscow	**1.27**	↗↗	**0.92**	↘
Taipei	**1.21**	↗↗	**0.76**	↘
Hong Kong	**1.20**	↗↗	**0.81**	↘
Beijing	**1.14**	↗↗	**0.81**	↘
Melbourne	**1.11**	↗↗	**0.94**	↘
Madrid	**1.11**	↗↗	**0.94**	↘
Munich	**1.07**	↗	**0.95**	↘
Barcelona	**1.07**	↗	**0.97**	↘
Shanghai	**1.05**	↗	**0.93**	↘
Montreal	**1.04**	↗	**0.98**	↘
Roma	**1.03**	↗	**0.97**	↘
Toronto	**1.02**	↗	**0.99**	↘
Seoul	**1.01**	–	**0.99**	–
Paris	**1.01**	–	**1.00**	–
San Francisco Bay	**1.00**	–	**1.01**	–
Berlin	**0.99**	–	**1.00**	–
Washington-Bethesda	**0.99**	–	**1.01**	–
Durham Research Triangle	**0.99**	–	**1.02**	–
Philadelphia	**0.99**	–	**1.02**	–
Kyoto-Osaka	**0.99**	–	**1.02**	–
Chicago	**0.99**	–	**1.02**	–
Milan-Pavia	**0.99**	–	**1.01**	–
New York	**0.98**	↘	**1.03**	↗
Tokyo	**0.98**	↘	**1.04**	↗
Sydney	**0.98**	↘	**1.01**	↗
Los Angeles	**0.98**	↘	**1.04**	↗
Baltimore	**0.98**	↘	**1.06**	↗
Boston	**0.95**	↘	**1.08**	↗
London	**0.86**	↘↘	**1.09**	↗↗

Key: The% of domestic interurban co-authorship of Moscow have been multiplied by 1.27 between 2000 and 2007.
[a] Fractional counting (WNC), 3-years moving average. Source: SCIExp (articles, reviews, letters).

Table 4
Evolution of the network's global indicators.

The 500 top publishing cities	2000[a]	2007[a]	evolution (%)
Number of cities	500	500	–
Number of links	74 567	92 519	24.1
Links values	178 886	277 658	55.2
Density	0.6	0.75	25.0

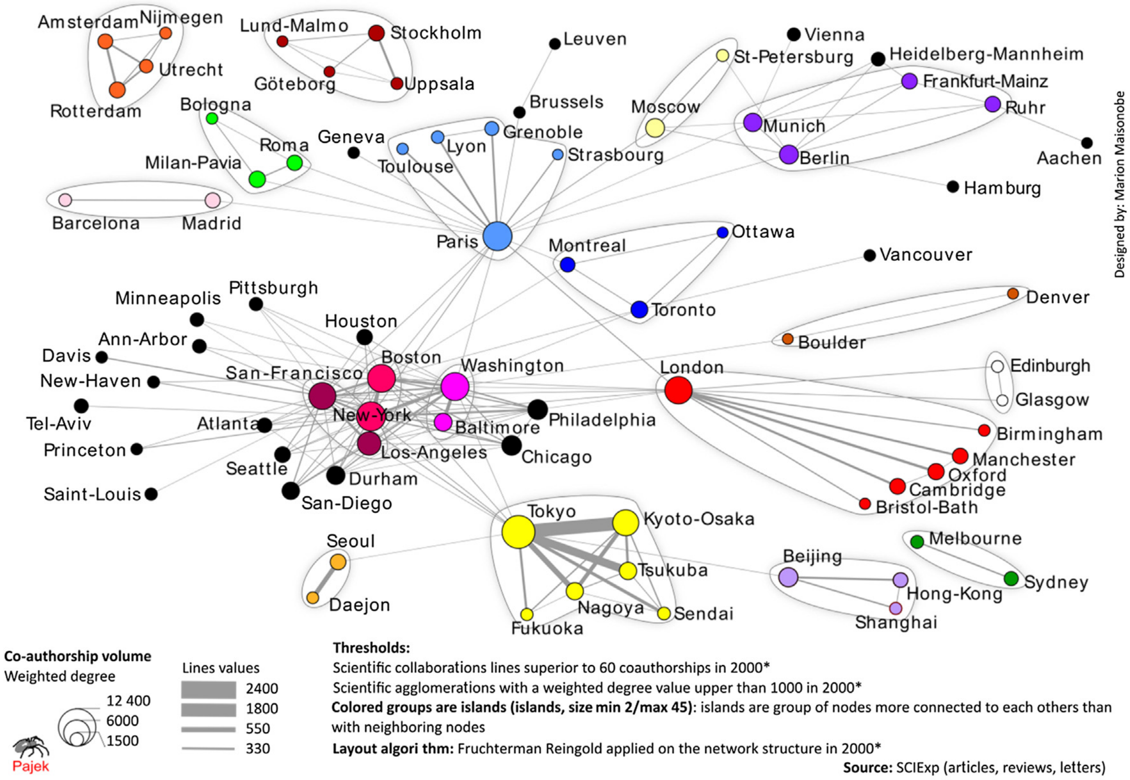
[a] 3-year moving average. Source: SCI Exp (articles, reviews, letters).

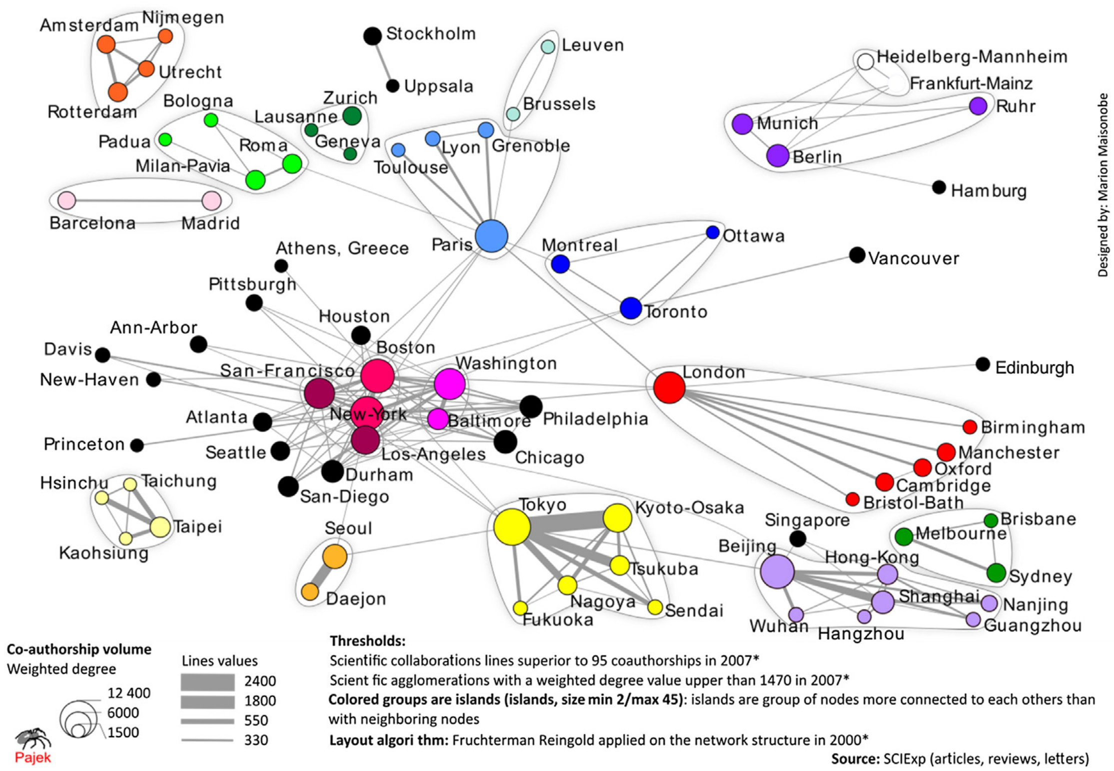
4.3. The evolution of the world collaboration network: densification, deconcentration and strength of domestic links

In order to assess changes in the structure of the world collaboration network, we focus on the co-authorship matrix of the 500 top publishing cities in 2007.[1] These cities are from 59 different countries and are responsible for 87% of world publications in 2007 (960 880 publications). In analyzing the same set of cities in 2000 and 2007, we can track the evolution of the collaboration network not being influenced by the entrance of new production centers during the period. Table 4 shows that there has been a densification process: the isolated or weakly connected cities in 2000 are much more integrated in 2007.

The global connectivity of the network (the density and the degree indicators) increased by 25% between 2000 and 2007. Three quarts of potential linkages were realized in 2007. In 2007, almost every city was related to the others by at least one co-authorship link. This result suggests that the network has densified at the global level. In addition, Table 5 shows that the participation share of the top collaborating cities has decreased within the total collaboration volume.

The 100 top collaborating cities were still involved in the majority of collaborations in 2007 (they account for 57.38% of all collaborations) but there has been a clear, though slow, trend toward the deconcentration of collaborations between 2000 and 2007. This trend suggests that the densification process did not reinforce the top cities' centrality. By focusing on the more central cities of the network, we want to examine the impact of this deconcentration process on their connectivity.





Figs. 1 and 2. World network of interurban scientific collaborations in 2000 and 2007 (Whole Normalized Counting, 3- years moving average).

Table 5
The share of city participation in the interurban network.

Share of co-authorship among the 500 top publishing cities (%)	Share of co-authorship	
	2000[a]	2007[a]
Top 100 cities	59.42	57.38
From the top 100 to the top 200 cities	19.51	19.83
From the top 200 to the top 300 cities	10.90	11.55
From the top 300 to the top 400 cities	6.93	7.15
From the top 400 to the top 500 cities	3.23	4.09
Total%	100.00	100.00

[a] Normalized counting (WNC), 3-year moving average. Source: SCI Exp (articles, reviews, letters).

In focusing on the most collaborative cities and on the strongest links between them, we obtain Figs. 1 and 2. In each figure, the visible cities are those which, ranked in descending order, are involved in 55% of all interurban collaborations, and the visible links are those which, ranked in descending order, account for 20% of all interurban collaborations. Using a similar criterion for both figures ensures their comparability. The fact that an almost identical number of cities (about 80) is obtained in both figures suggests stability.[2] However, since the visible cities are not the same in 2000 and 2007, the stability hypothesis needs to be qualified.

Between 2000 and 2007, the cities belonging to the top of the hierarchy and their links are slightly uneven. The most striking change is the entry of Chinese cities at the expense of Swedish, Russian and American cities (Colorado). Considering the size of each node accounting for the total number of collaborations per city, we noticed that the Chinese cities and more generally Asian cities have experienced the strongest growth in scientific collaborations. Indeed, the world scientific collaboration network is first of all expanding in favor of previously peripheral countries (hence the entrance of Chinese cities in the top of the hierarchy).

Interestingly, the figures show that the trend toward nationalization previously identified in some countries is still obvious when focusing on the top interurban collaborations. Indeed, the strongest links in 2007 are more often intra-national than in 2000. Thus, whereas there were four strong components (distinct groups of cities) in 2000 (the main one + Netherlands, Sweden and Australia), there are six of them in 2007 (the previous ones + Switzerland, Spain and Taiwan). Overall, we notice that national systems are reinforcing.

The fact that only a few “world” cities share very strong international links is in line with many studies on world cities' networks (Matthiessen et al., 2010). However, contrary to most authors in this field of study (the world city consortium GaWC around Peter Taylor), we oppose the idea of a trend toward an increasing concentration of scientific activities within world cities at the expense of smaller cities. Our results suggest that scientific collaborations between cities are strengthening in descending order from the regional to the global level, while passing by the national level.

Although Boston, Paris, New York, London and Beijing are well connected in the global network, their international links are weaker than their domestic linkages. Focusing on Tokyo, New York, London (the 3 “global cities”, analyzed by Saskia Sassen in 1991), as well as on Paris and Beijing, Table 6 shows that, with only a few exceptions, scientists in world cities collaborated more with their national counterparts in 2007.

The network analysis confirms that the growing integration of national systems should not be ignored to understand the geography of scientific activity at the city level. However, national systems are not the only organizations with structural effects on scientific collaboration networks. On Figs. 1 and 2, the color of the nodes depends on the “islands” they belong to according to their collaboration patterns. An “island” is a group of cities that share relationships whose values are higher than the strongest bond they have developed outside their group. In order to detect these groups, we have used a clustering method implemented in the Pajek software (used and defined by Batagelj, Kejzar, & Korenjak-Cerne, 2006 to analyze patent data in the United States). Although, in most cases, islands are entire countries, there are certain instances where islands are sub-national groups. Actually, sub-national groups have been detected in the United States, Germany and the United Kingdom. Not surprisingly, these countries have a more federal organization and the sub-national islands we found within them group together cities belonging to the same province (e.g. California and Scotland).

While the growth of interurban scientific collaborations occurred mainly within national frames between 2000 and 2007, the world-wide network expanded globally at the same time. While the strongest links remained intra-national, scientific relationships developed simultaneously between all the cities of the world system.

[1] This matrix has been computed on the basis of collaborative articles, reviews, and letters indexed in the SCI Exp in 1999–2001 and 2006–2008. We have chosen to stop at 500 cities because these cities figured as major publishing places in both the entire Web of Science catalogue and the SCI Expanded.

[2] *nota bene*: Large research centers such as Helsinki, Copenhagen or Oslo are not represented since they only match one of the two criteria we used for the visualization. Indeed, they are well connected cities (a high number of collaborations) but they don't share any collaboration link superior to the selected threshold.

Table 6
The top 15 scientific partners of 5 World cities (Tokyo, New York, Paris, London, Beijing) in 2007.

First 15 partners of 5 world cities in decreasing order by co-authorship volumes, 2007[a]				
Tokyo	New York	Paris	London	Beijing
Kyoto-Osaka	Boston	Lyon	Cambridge	Shanghai
Tsukuba	Washington	Grenoble	Oxford-Didcot	Hong-Kong
Nagoya	San Francisco	Toulouse	Manchester	Changchun
Sendai	Philadelphia	Marseille-Aix	Bristol-Bath	Wuhan
Fukuoka	Los Angeles	Montpellier	Birmingham	Lanzhou
Sapporo	Chicago	Bordeaux	**Paris**	Nanjing
Shizuoka	Baltimore	Lille	Southampton	Guangzhou
Hiroshima	Durham	**London**	**New York**	Tianjin
Maebashi	Princeton	Strasbourg	Sheffield	Shenyang
Okayama	New Haven	Rennes	Glasgow	Dalian
Tokai	Houston	Nancy	**Boston**	Xian
Kanazawa	San Diego	Nice	Edimbourg	Hefei
Utsunomiya	Brooklyn	**New York**	Leicester	Chengdu
Beijing	Atlanta	Nantes	Leeds	Kunming
Niigata	**London**	**Boston**	Nottingham	Jinan

In bold, foreign cities.
[a] Whole-Normalized Counting (WNC), 3-year moving average 2006–2008. Source: SCIExp (articles, reviews and letters).

5. Conclusion

The data analysis of scientific publications between 2000 and 2007, focusing on the development of interurban collaborations, highlights several global trends. The majority of publications were still produced by a single city, but this proportion declined everywhere in the world. There was an overall increase in the number of publications produced by several cities, within a single country or several. Interestingly, we have noticed that cities located in scientific emerging countries tended to favor domestic interurban co-authorships whereas cities located in more traditionally English-speaking countries internationalized. Actually, in most countries, there has been a general increase of all kinds of collaborations. As a result, the global interurban network of collaborations has densified between 2000 and 2007.

Among the top cities, intra-national links reinforced. Among the top collaborating cities, Asian cities superseded certain Eastern and Northern European cities. All of these results suggest that the worldwide growth of scientific collaborations, referred to as “globalization”, did not develop at the expense of national systems of science.

According to us, these results are the consequence of the decentralization process which took place in almost every country following the devolution of higher education services. This devolution process has favored the growth of research by hiring scientists in new scientific centers. Little by little, their collaborative methods tend to become identical to that of the top cities’ scientists. As we saw, the national level has been an essential component for the development and integration of scientific cities. This level has remained equally structuring for more traditional territories where national and international collaborations have complemented each other. Our results highlight the role of national systems of cities in collaboration dynamics. Instead of focusing on the role of the biggest cities as international hubs, we think that scientific policies, notably in Europe, should take into account this multi-level development of scientific collaborations.

Our contribution is part of an ongoing research process. Some limitations should be discussed since they can open up lines for future research. First, our study takes into account all publications without considering the number of citations received. It will certainly be necessary to repeat the analysis by integrating this variable to see if the trends are the same if we examine the most cited publications. Second, our data end in 2008 and it will be necessary to continue the analysis with more recent data, which means to repeat the geocoding and checking the stability of scientific agglomerations. By this means, it will be possible to test the hypothesis according to which collaborative practices in traditional and new centers of scientific activity have continued to converge after 2007, both at the country level and at the urban area level. Third, our network analysis focuses on the question of the hierarchy of the centrality and overall density. It needs to be continued by focusing on other structuring levels than the national level (regional and macro-regional levels) and by focusing on various disciplines. Indeed, it will be useful to search for more complex structures in the global graph of co-signatures of articles.

Author contributions

Marion Maisonobe: Conceived and designed the analysis, collected the data, contributed data or analysis tools, performed the analysis and wrote the paper.

Denis Eckert: Conceived and designed the analysis, collected the data and contributed data or analysis tools.

Michel Grossetti: Conceived and designed the analysis, contributed data or analysis tools, performed the analysis and wrote the paper.

Laurent Jégou: Conceived and designed the analysis, collected the data and contributed data or analysis tools.
Béatrice Milard: Conceived and designed the analysis, contributed data or analysis tools and wrote the paper.

Acknowledgements

This work received support from the French Research Agency (L'Agence Nationale de la Recherche) through the GEOSCIENCE project and the Excellence laboratory "Structuring of social worlds" (ANR-11-LABX-0066) funded by the French Research Agency (Agence Nationale de la Recherche) as part of the program n° ANR-11-IDEX-0002-02. We are grateful to Drs Y. Gingras, V. Larivière, M. Zitt, A. Shubert, and L. Leydesdorff for helpful discussions. We would like to thank the fmr group for their support in network analysis, OST Paris and Montréal for providing bibliometric data and advises.

References

Andersson, A. E., & Persson, O. (1993). Networking scientists. *The Annals of Regional Science*, *27*(1), 11–21.
Batagelj, V., Kejzar, N., & Korenjak-Cerne, S. (2006). Analyzing the structure of U.S. patents network. In *Data science and classification*. Ljubjana: Springer., p. 358.
Beaver, D., & Rosen, R. (1978). Studies in scientific collaboration: part I. The professional origins of scientific co-authorship. *Scientometrics*, *1*(1), 65–84.
Beaver, D., & Rosen, R. (1979). Studies in scientific collaboration: part II. Scientific co-authorship, research productivity and visibility in the French scientific elite, 1799–1830. *Scientometrics*, *1*(2), 133–149. http://dx.doi.org/10.1007/BF02016966
Bornmann, L., Leydesdorff, L., Walch-Solimena, C., & Ettl, C. (2011). Mapping excellence in the geography of science: an approach based on Scopus data. *Journal of Informetrics*, *5*(4), 537–546. http://dx.doi.org/10.1016/j.joi.2011.05.005
Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*, *33*(4), 599–616. http://dx.doi.org/10.1016/j.respol.2004.01.008
Cabanac, G., Gilles, H., & Milard, B. (2015). Academic careers in computer science: continuance and transience of lifetime co-authorships. *Scientometrics*, *102*(1), 135–150.
Eckert, D., Baron, M., & Jegou, L. (2013). Les villes et la science: apports de la spatialisation des données bibliométriques mondiales. *M@ppemonde*, *110*(2) http://mappemonde.mgm.fr/num38/articles/art13201.pdf
Frame, J. D., & Carpenter, M. P. (1979). International research collaboration. *Social Studies of Science*, *9*(4), 481–497.
Frame, J. D., & Narin, F. (1988). The national self-preoccupation of american scientists – an empirical view. *Research Policy*, *17*(4), 203–212. http://dx.doi.org/10.1016/0048-7333(88)90049-2
Frenken, K., Hardeman, S., & Hoekman, J. (2009). Spatial scientometrics: towards a cumulative research program. *Journal of Informetrics*, *3*(3), 222–232. http://dx.doi.org/10.1016/j.joi.2009.03.005
Frenken, K., Hoekman, J., Kok, S., Ponds, R., Van Oort, F., & van Vliet, J. (2009). Death of distance in science? A gravity approach to research collaboration. In A. Pyka, & A. Scharnhorst (Eds.), *Innovation networks new approaches in modeling and analyzing* (pp. 43–57). Berlin-Heidelberg: Springer-Verlag.
Gauffriau, M., Larsen, P., Maye, I., Roulin-Perriard, A., & Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, *77*(1), 147–176. http://dx.doi.org/10.1007/s11192-007-1934-2
Georghiou, L. (1998). Global cooperation in research. *Research Policy*, *27*(6), 611–626. http://dx.doi.org/10.1016/S0048-7333(98)00054-7
Gingras, Y. (2002). Les formes spécifiques de l'internationalité du champ scientifique. *Actes de la recherche en sciences sociales*, *2*(141), 31–45.
Grossetti, M., Eckert, D., Gingras, Y., Jégou, L., Larivière, V., & Milard, B. (2014). Cities and the geographical deconcentration of scientific activity: a multilevel analysis of publications (1987–2007). *Urban Studies*, *51*(10), 2219–2234. http://dx.doi.org/10.1177/0042098013506047
Hennemann, S., & Liefner, I. (2015). Global science collaboration. In *Handbook of global science technology and innovation*. pp. 347–367. Wiley–Blackwell.
Hennemann, S., Rybski, D., & Liefner, I. (2012). The myth of global science collaboration—collaboration patterns in epistemic communities. *Journal of Informetrics*, *6*(2), 217–225. http://dx.doi.org/10.1016/j.joi.2011.12.002
Hoekman, J., Frenken, K., & Oort, F. (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, *43*(3), 721–738. http://dx.doi.org/10.1007/s00168-008-0252-9
Hoekman, J., Frenken, K., & Tijssen, R. J. W. (2010). Research collaboration at a distance: changing spatial patterns of scientific collaboration within Europe. *Research Policy*, *39*(5), 662–673. http://dx.doi.org/10.1016/j.respol.2010.01.012
Inhaber, H. (1977). Changes in centralization of science. *Research Policy*, *6*, 178–193.
Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*(1), 1–18. http://dx.doi.org/10.1016/S0048-7333(96)00917-1
Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, *31*(1), 31–43. http://dx.doi.org/10.1007/BF02018100
Larsen, P. O., & Von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics*, *84*(3), 575–603. http://dx.doi.org/10.1007/s11192-010-0202-z
Leydesdorff, L., & Persson, O. (2010). Mapping the Geography of Science: Distribution Patterns and Networks of Relations among Cities and Institutes.
Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2014). The European Union, China, and the United States in the top-1% and top-10% layers of most-frequently cited publications: competition and collaborations. *Journal of Informetrics*, *8*, 606–617.
Luukkonen, T., Persson, O., & Siverstsen, G. (1992). Understanding patterns of international scientific collaboration. *Science Technology & Human Values*, *17*(1), 101–126.
Makse, H. A., Andrade, J. S., Batty, M., Shlomo, H., & Eugene, S. H. (1998). Modelling urban growth patterns with correlated percolation. *Physical Review E*, *58*(6), 7054.
Matthiessen, C. W., Schwarz, A. W., & Find, S. (2010). World cities of scientific knowledge: systems networks and potential dynamics. an analysis based on bibliometric indicators. *Urban Studies*, *9*(47), 1879–1897. http://dx.doi.org/10.1177/0042098010372683
Melin, G. (2000). Pragmatism and self-organization: research collaboration on the individual level. *Research Policy*, *29*(1), 31–40. http://dx.doi.org/10.1016/S0048-7333(99)00031-1
Milard, B., & Grossetti, M. (2006). L'évolution de la recherche scientifique dans les régions de Russie: déclin ou déconcentration? *M@ppemonde*, *1*(81) http://mappemonde.mgm.fr/num9/articles/art06105.html
Morgan, K. (2001). The exaggerated death of geography: localised learning, innovation and uneven development. *Technology*, *20*, 23.
Nagpaul, P. S. (2003). Exploring a pseudo-regression model of transnational cooperation in science. *Scientometrics*, *56*(3), 403–416.
Royal Society (Great Britain). (2011). *Knowledge, Networks and Nations Global Scientific Collaboration in the 21 st Century*. London: The Royal Society.
Sassen, S. (1991). *La ville globale: New York, Londres, Tokyo*. Paris: Descartes & Cie. 1996. D.-A. Canal, Trad.
Schott, T. (1993). World science: globalization of institutions and participation? *Science Technology & Human Values*, *18*(2), 196–208.
Sexton, J. (2012). State of the world's science. *Scientific American*, *307*(4), 36–40.
Tijssen, R. J., Waltman, L., & van Eck, N. J. (2012). *Research collaboration and the expanding science grid: measuring globalization processes worldwide*. , arXiv preprint arXiv:1203.4194. http://arxiv.org/abs/1203.4194
Wagner, C. S. (2008). *The new invisible college: science for development*. Brookings Institution Press.
Wilson, C. S., & Markusova, V. A. (2004). Changes in the scientific output of Russia from 1980 to 2000, as reflected in the Science Citation Index, in relation to national politico-economic changes. *Scientometrics*, *59*(3), 345–389. http://dx.doi.org/10.1023/B:SCIE.0000018539.93567.d8

Zitt, M., & Bassecoulard, E. (1999). Internationalization of communication: a view on the evolution of scientific journals. *Scientometrics*, *46*(3), 669–685.
Zitt, M., Perrot, F., & Barré, R. (1998). The transition from national « to transnational « model and related measures of countries» performance. *Journal of the American Society for Information Science*, *49*(1), 30–42.
Zitt, M., Barré, R., Sigogneau, A., & Laville, F. (1999). Territorial concentration and evolution of science and technology activities in the European Union: a descriptive analysis. *Research Policy*, *28*(5), 545–562. http://dx.doi.org/10.1016/S0048-7333(99)00012-8