



HAL
open science

Regional Distribution of Research: The Spatial Polarization in Question

Marion Maisonobe

► **To cite this version:**

Marion Maisonobe. Regional Distribution of Research: The Spatial Polarization in Question. Handbook Bibliometrics, De Gruyter Saur, pp.377-396, 2020, 10.1515/9783110646610-036 . hal-03097835

HAL Id: hal-03097835

<https://hal.science/hal-03097835>

Submitted on 6 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

5.8 Regional Distribution of Research: The Spatial Polarization in Question

Marion Maisonobe

Abstract: The objective of this chapter is threefold: showing how research activity is spatially organized in the contemporary world at different scales, what are the different means and datasets available to measure its spatial distribution, and what can explain the degree of polarization of this activity. The chapter adopts a critical view on the literature that stresses the spatial polarization of this activity as if it was a necessary feature of it. First, the chapter demonstrates that contrary to most expectations, this polarization has globally diminished since the second part of the twentieth century. Second, taking the example of the regional distribution of research in Europe, it insists on the various means and sources used to study this distribution and its evolution through time. To finish, it highlights an overlooked dimension of research activity: the fact that the institutional geography (authors' affiliations) does not reflect where research is actually done, especially in the case of field ground-based research; and it gives some direction for future research.

Keywords: geography of science, spatial scientometrics, regional economics, urban geography, spatial polarization.

Introduction

Scholars from distinct fields with different research questions are working at studying the geography of scientific production. We currently observe three main reasons and three main types of scholars interested to this aim:

- The first reason relates to the research of indicators to monitor scientific activity and to the will to relate scientific policies to the spatial organization of science. It interests in priority scientometricians and more specifically among them spatial scientometricians. Apart from the construction of indicators, this approach targets a better understanding of the “growth of science” as it is measurable through the growing publication of academic papers.
- The second reason relates to the attempt to link indicators of territorial development, innovation, and economic growth to indicators related to research activity (investment in R&D as well as research output). It interests in priority scholars specialized in regional economics that are keen to understand the mechanisms allowing transferring knowledge into innovation that can be useful for civil so-

Marion Maisonobe, UMR Géographie-cités, CNRS – Université Panthéon-Sorbonne – Université de Paris – EHESS, Campus Condorcet, Bâtiment de recherche Sud, 5, cours des Humanités, 93322 Aubervilliers, France, marion.maisonobe@cnrs.fr

- ciety either through the commercialization of new products or through radical changes in people's lives and beliefs.
- The third reason relates to the ambition of understanding spatial mechanisms leading to the contemporary distribution of men and activities and their connections through material infrastructures (roads, railways, etc.) and immaterial exchanges of information and knowledge. The will of understanding certain territorial arrangement in particular also enters into this line of research: spatial dynamics of certain national systems, macro-regions, and linguistic areas. This approach interests in priority scholars belonging to the field of geography, but also historians and sociologists.

Historically, Joseph Ben-David was one of the first scholars who tried to elaborate a theory to explain the changing geography of scientific production by analyzing a move of the world science's center of gravity from Germany to the United-States between the mid-nineteenth century and the early 1970s. Ben-David was not only interested in scientific power, but by the entanglement between systems of higher education and the organization of scientific research, which explains why he sometimes referred to the nations he was focussing on as "centers of learning" or as "scientific centers" (Ben-David, 1991). His work falls within the stream of institutionalist sociology of science, which subsequently has lost ground since the 1970s. Later, most studies focussing on science production at the world level were performed by scientometricians using the Web of Science database to analyze the growth of science. As a result, the precise question of the spatial distribution of scientific production and its global evolution has more to do with spatial scientometrics nowadays, as can be seen from the presence of this chapter hereby.

This chapter begins by a review of the literature on the spatial distribution of contemporary research. It then shows that my group's observations on this point – obtained from analyses conducted on the Web of Science – diverge significantly from several published works on the subject in regional economics and urban geography as well as from political discourses on the spatial organization of research. Beyond my group's observations, the chapter takes the example of the regional distribution of research in Europe to stress the limits of what is measurable with bibliographic sources and of the means used to measure it. To finish, I insist on the advisability of using new sources to study the geography of research, taking into account not only the institutional geography of science but also sources that make it possible to locate the concrete place where research is carried out. To this aim, I give the example of the Research in Svalbard (RIS) database, which allows localising precisely the sites where researchers go to perform their field ground research on this arctic territory.

On the Spatial Polarization of Research

On the Spatial Polarization of Research before the 2000s

Research as we understand it today is an activity that was practised only in a very limited number of places until the mid-twentieth century. As mentioned by Jean Gottmann in the conclusion of his discourse on the “The need of an international policy for the sciences” at the Nobel Symposium in 1973: “the apparent order which existed after 1945, owing to the high concentration of science and technology in a very small number of countries and centres, has now been eroded. If any international policy for science has been operating in the last 20 years, it has certainly been one of geographic decentralization. It has borne results; one consequence of it is precisely the need for more common moorings” (Gottmann, 1973). The spatial diffusion of research is a phenomenon that can be related to an increase in the number of countries developing a system of higher education and investing in science during the second part of the twentieth century. The effects of this phenomenon are also observed within countries by Herbert Inhaber in an article published in 1977 in *Research Policy*, “Changes in centralization of science” (Inhaber, 1977). By retrieving data from the Institute for Scientific Information (ISI) on scientists’ affiliations, Inhaber reckons that “When the relative change in concentration from 1967 to 1972 was considered, almost all of the countries considered had a fairly strong decrease. For most, the change was of the order of 2% a year”. For Inhaber, “this dispersion of scientists may be due to less reliance on a small centralized corps of scientists once a nation reaches a “take-off” point in science”. Together with an increasing spatial dispersion of scientists within countries, referring to 1990’s UNESCO data, Thomas Schott reports an increasing participation in the world of science by a growing number of countries: “In recent decades, expenditure on research and development has increased nearly everywhere in absolute amounts and in most places also as a percentage of the gross national product. Scientists and engineers engaged in research and development have also increased nearly everywhere, both in absolute numbers and as a fraction of the population of their respective countries. All these measures indicate a growth in scientific activity in nearly every country in recent decades” (Schott, 1991). Despite these evolutions, Schott insists on the growing concentration of scientific activity in the wealthiest countries as measured by the growth in the global share of high-standard publications per country (+1.61 for the wealthiest countries against +0.96 for the poorest countries and -2.58 for communist countries between 1973 and 1986). This analysis focusses on the scientific domination of liberal democracies over communist countries following the collapse of the USSR. The classification Schott uses to group countries is very much influenced by this context and helps him to demonstrate the drop of Soviet science, which is of course very significant during the period under study. If we look into more details of the data computed by Schott (Table 1), we can observe that actually, among the “liberal democracies”, it is only Japan and “the rest of Western Europe” that have registered considerable

increases in their global share of publication. When normalizing by countries' population growth, the picture is interestingly different, suggesting a spatial diffusion of the scientific activity within countries, which has been faster within the wealthiest countries than within the poorest countries between 1973 and 1986. In the post-2000s literature on the spatial polarization of research, as shown in the next subsection, the issues of data standardization and aggregation continue to play an important role in defining what is meant by "spatial polarization" and thus explain that even empirical analyses can disagree on the matter.

On the Spatial Polarization of Research after the 2000s

In 2005, confirming the diffusion trends identified earlier, Evan Schofer and John W. Meyer published in the *American Sociological Review* the result of an extensive research on "The Worldwide Expansion of Higher Education in the Twentieth Century" (Schofer and Meyer, 2005). Their research shows that a growing enrolment of the population in higher education and the development of university systems happened in every region of the world during the twentieth century. The authors analyze the institutional processes that led to what they call "a new model of society" including scientization, democratization and the expansion of human rights, the rise of development planning, and the structuration of the world polity. According to them, "the result is a highly expanded, and essentially global, system of higher education". However, if the global growth in the population of students, in the number of universities, and R&D institutions is a shared observation among academics, there is more debate regarding changes in the geography of research output.

Table 1: Changes in Production of Scientific Papers, 1973–1986. After Schott, 1991.

	Percentage of scientific papers*			Ratio of percentage papers to percentage population		
	1973	1986	Change	1973	1986	Change
<i>Scientific Community</i>						
Liberal Democracies:	81.77	83.38	+1.61	4.29	4.90	+0.61
United States	38.23	36.53	-1.70	7.12	7.49	+0.37
Canada	4.39	4.38	-0.01	7.84	8.45	+0.61
Japan	5.26	7.92	+2.66	1.92	3.23	+1.31
United Kingdom	9.21	7.97	-1.24	6.46	6.95	+0.49
West Germany	6.04	5.99	-0.05	3.85	4.87	+1.02
France	5.56	4.99	-0.57	4.21	4.45	+0.24
Rest of Western Europe	9.74	12.17	+2.44	1.67	2.34	+0.67
Australia, New Zealand	2.38	2.41	+0.03	5.75	6.18	+0.43
Israel	0.96	1.02	+0.06	11.53	11.75	+0.22
Communist Countries:	13.18	10.59	-2.58	1.25	1.14	-0.11
Soviet Union	9.00	7.39	-1.60	1.15	1.07	-0.09
Eastern Europe	4.18	3.20	-0.98	1.54	1.38	-0.16

Poor Countries:	5.06	6.02	+0.96	0.07	0.08	+0.01
North Africa, Near East	0.43	0.56	+0.14	0.12	0.13	+0.01
South Africa	0.45	0.53	+0.08	0.73	0.79	+0.06
Rest of Africa	0.35	0.31	-0.04	0.05	0.04	-0.01
India	2.53	2.35	-0.19	0.17	0.15	-0.02
Rest of Asia	0.33	1.07	+0.74	0.01	0.03	+0.02
Latin America	0.97	1.20	+0.23	0.13	0.15	+0.02

*Note: Articles by authors in each country and region in a year as percentage of total papers produced worldwide. Change is difference in that percentage 1973–86.

Sources: Number of papers: National Science Foundations SP # 1 Science Literature Indicators Database, Indicator I_b (1973-journal set), from CHI Research, derived from *Science Citation Index*.

Population: *World Tables Compressed Data Diskettes (DEK)*, *op. cit.*; *Demographic Yearbook 1975* (New York: United Nations, 1979), and *Demographic Yearbook 1987*, *op. cit.*

According to certain scholars, research is an activity that is intrinsically concentrated and tends to become more and more polarized though time. Empirically, this vision is being nurtured by the gap between the distribution of the world population and the distribution of research output and impact. For instance, in the early 2000s, Michael Batty used the ISI's HighlyCited database composed of the top 100 or so cited individuals in 14 scientific fields to compute a level of spatial concentration of these individuals. He comments: "The pattern of concentration that this analysis reveals is remarkable: 1222 scientists work in 429 institutions which are located in 232 places in 27 countries. Almost half these scientists are in 50 institutions in 5 countries, most being in the United States" (Batty, 2003). In 2005, Richard Florida also relies on this source to observe: "Scientific advance is even more concentrated than patent production. Most occurs not just in a handful of countries but in a handful of cities – primarily in the United States and Europe. Chinese and Indian cities do not even register. As far as global innovation is concerned, perhaps a few dozen places worldwide really compete at the cutting edge" (Florida, 2005). In both cases, the spatial distribution of the top cited scientists is analyzed only for a given year, which does not allow identifying a spatial trend. However, Florida links his observation to a dynamic theory. According to him, the concentration of creative talents in a few hotspots able to connect to the global system of cities is intensifying from the 1990s. This theory borrows from regional economics the principles according to which innovating activities tend to cluster in a few centers because of "the powerful productivity advantages, economies of scale, and knowledge spillovers such density brings". In 2010, the bibliometric analyses of Christian W. Matthiessen, Annette W. Schwartz, and Soren Find offered a more precise and longitudinal analysis on the geography of research output encompassing the world's most publishing urban areas (Matthiessen et al., 2010). Instead of a reinforcing domination of the traditional centres, they observe that "Clearly, the position of European and North American research is under challenge from new high-producing centres outside these continents", which suggests a geographical expansion at the world level. However, measuring the evolution of the number of publications retrieved in the Science Citation Index (SCI) between 1996

and 2006, they remark: “Total growth has been 28 per cent and growth in the top 30 cities has been 34 per cent, which demonstrates a concentration process”. In 2013, Michel Grossetti, Denis Eckert, Yves Gingras, Laurent Jégou, Vincent Larivière, and Béatrice Milard explored the same issue but offered an analysis more in line with pre 2000s observations (Grossetti et al., 2013). Instead of focussing on the most publishing places, their analysis encompasses all the publishing places retrieved in the SCI-Expanded. They demonstrate a deconcentration trend in the distribution of the scientific production happening both within most countries and between countries at the world level. First, they show that “the number of countries needed to account for 80 % of world publications was 7 in 1978, 10 in 1988 (including the USSR), 13 in 1998 (with a united Germany and Russia separated from other former USSR countries), and 16 in 2008” (Table 2).

Table 2: Evolution between 1987 and 2007 of the 30 most productive countries in world publications. After Grossetti et al., 2013.

Country	Country/world share in 1987 ^a	Country/world share in 1997 ^a	Country/world share in 2007 ^a	1987 rank	1997 rank	2007 rank
USA	34.3	29.8	24.5	1	1	1
China	0.9	2.4	8.6	18	12	2
Japan	7.3	9.0	7.8	4	2	3
Germany	7.4	7.4	5.9	3	3	4
UK	7.8	7.4	5.4	2	4	5
France	5.3	5.5	4.3	5	5	6
Italy	2.5	3.4	3.5	9	8	7
Canada	4.3	3.7	3.4	7	6	8
India	2.6	2.2	3.0	8	11	9
Spain	1.3	2.3	2.7	13	9	10
South Korea	0.1	1.1	2.7	42	16	11
Russia	5.1	3.5	2.2	6	7	12
Australia	2.1	2.2	2.1	10	10	13
Brazil	0.5	0.9	2.1	27	21	14
The Netherlands	1.8	2.0	1.7	11	13	15
Turkey	0.1	0.5	1.6	44	26	16
Taiwan	0.3	1.0	1.6	33	18	17
Poland	1.0	0.9	1.4	17	20	18
Sweden	1.6	1.6	1.2	12	14	19

Table 2 (Continued)

Country	Country/world share in 1987 ^a	Country/world share in 1997 ^a	Country/world share in 2007 ^a	1987 rank	1997 rank	2007 rank
Switzerland	1.2	1.3	1.2	14	15	20
Belgium	0.9	1.0	0.9	19	17	21
Iran	0.0	0.1	0.8	73	51	22
Israel	1.0	0.9	0.8	16	19	23
Greece	0.3	0.5	0.7	31	28	24
Austria	0.6	0.7	0.6	23	24	25
Denmark	0.8	0.8	0.6	20	22	26
Finland	0.6	0.7	0.6	21	23	27
Mexico	0.2	0.4	0.6	35	32	28
Czech Republic	0.5	0.4	0.5	24	29	29
Singapore	0.1	0.3	0.5	45	35	30

Fractional counting of publications, with 3-year moving average.

Second, they show that: “The proportion of world publications by the 25 cities ranked among the top 30 in both 1997 and 2007 regressed by 3.7 percentage points. Unsurprisingly, the share of newcomers (absent from the top 30 in 1997 and present in 2007) increased by 1.1 percentage points. The share of the five other agglomerations (present in 1997 and absent in 2007) regressed by 0.7 percentage points during the same period”. It thus emerges that a twofold process was at work during the 2000s: the rise of new players in the global game, including China, and a continuation of the spatial deconcentration of scientific production observed within countries since the 1970s. This work therefore calls into question the assimilation of the research activity of an activity that is intrinsically polarized. More recent studies show that this deconcentration has been going on during the 2010s and is also measurable when focussing on the spatial distribution of collaboration (Maisonobe et al., 2016), of citations (Maisonobe et al., 2017), and of the 10% top cited publications in the world (Maisonobe et al., 2018). Despite these global and comprehensive analyses, a series of publications in regional economics and urban geography still consider the spatial concentration of scientific activities as an ongoing or intensifying phenomenon. In the next section, we attempt to explain this apparent contradiction.

How to Measure a (De)concentration Process?

Different Methods and Scales, Different Observations?

As noted by Susanne A. Frick and Andrés Rodríguez-Pose for the measure of urban concentration of the population (Frick and Rodríguez-Pose, 2018), there exists an important variety of methods that can be used to measure a degree of spatial polarization and its evolution. Regarding the spatial concentration of research, the literature relies either on the evolution of the Gini index (Dolores León et al., 2011; Zitt et al., 1999), on the share of the top centers in research intensity (Grossetti et al., 2013; Maisonobe et al., 2018), or on the scaling parameter (called “beta”) in urban scaling laws (Nomaler et al., 2014; Pumain et al., 2006). In addition, we observe a growing number of studies relying on network analysis metrics to measure the spatial distribution of knowledge exchanges and the evolving centrality of certain places in R&D networks (Bergé et al., 2017; Hong, 2008). Both the territorial units of analysis and the scale of analysis vary across studies.

Among all this literature, urban scaling studies are those emphasizing the most the concentrated and metropolitan nature of research activity. Unlike other methods of measuring spatial concentration of an activity, urban scaling laws measure a degree of urban concentration by taking into account the distribution of the activity in relation to the population of the cities in which the activity is distributed (see subsection III.2). In addition to being an index of concentration, the scaling parameter therefore makes it possible to know in which part of the urban hierarchy the activity under study is concentrated. Consequently, this methodology contributes to a better understanding of varying distribution patterns across countries. As shown by Grossetti et al. (2013), the geography of research depends on historical processes and planning choices that have been made throughout history and still have effects nowadays. This may explain why, in urban scaling studies, the degree of adjustment of research activities to the urban hierarchy (R^2 value of the regression model) is generally lower than that of jobs in sales and public administration, or of employment in general. By focussing on England and Wales, Arcaute et al. notice this specificity in the spatial distribution of patents: “In addition to economic hubs, one also encounters knowledge hubs, which also present dragon-king-like qualities and which are not necessarily correlated with size. These hubs are the outcome of path-dependencies that give rise to emergent properties that are not present in all cities as is the case of patents. This is most dramatically demonstrated by the dominance of patent production in Cambridge, UK” (Arcaute et al., 2015). Despite their poorer fit to the urban hierarchy, and despite the variety of spatial boundaries and methods available to measure scaling coefficients, R&D activities generally appear more concentrated in the most populated territories (beta index greater than 1). However, among the authors highlighting these clear superlinear relationships between the distribution of R&D activity and urban population, there are two opposing interpretations. First, Bettencourt et al. rely on Florida’s theory and consider super-linear regimes of scal-

ing as the result of increasing returns to network interactions in the most populated places (Bettencourt et al., 2007). Second, Pumain et al. (2006) offer an evolutionary theory for interpreting urban scaling laws. According to them, human sectors characterized by a super-linear regime are innovative sectors concentrated in the top of the urban hierarchy at the time of their emergence. Their subsequent diffusion in the urban hierarchy transforms the scaling regime into a linear one (when the activity becomes common) and then into a sub-linear one (when it becomes mature).

This second interpretation seems adequate to describe the gradual deconcentration of research described in part II. Universities and research centers used to be rare institutions located in only a few places. With decentralization policies and the massification of students, these institutions have become much more common and located in an increasing number of cities. The spatial diffusion of these institutions has led to the world deconcentration of research results.

Yet, two recent counter examples deserve to be mentioned that highlight the limits of urban scaling to measure the spatial polarization of research activity. First, the study of Arcaute et al. shows that the beta index computed on 2000–2011 patents' distribution is very sensitive to the boundary definition used to delineate cities and that, for England and Wales, its relation with the urban population can be either linear or superlinear. Second, Leitão et al. discuss the classical use of OLS regressions for urban scaling laws, and, using different regression methods on multiple datasets, show that the relation of UK-patents [2000–2011] and OECD-patents [2008] to the population distribution is linear (Leitão et al., 2016). Their result questions the relevance of studies arguing that R&D activity naturally tends to concentrate in the most populated areas.

As the concentration index, the territory, and the method used to measure spatial polarization can lead to various observations, we consider it important to take into account geographical variations and avoid generalizations. Indeed the existing literature suggests that in some countries, R&D activity remains concentrated in the most populated areas, while other countries have a less centralized research system. There is also strong evidence that at the global level, the number of countries and cities involved in R&D activity have increased since the mid-twentieth century. Focusing on the European territory, the following sub-section highlights the sensitivity of observations to methodological choices.

The Regional Distribution of Research in Europe

In this sub-section, we examine the evolution of the geography of research at the European level between 2001 and 2012 using normalized publication outputs from the Web of Science database, as well as demographic, general employment and R&D data from Eurostat. R&D data from Eurostat include R&D expenditure (GERD), R&D employment, and patent applications to the EPO. Our goal is to compare the results obtained using different indicators and different methods. In order to relate

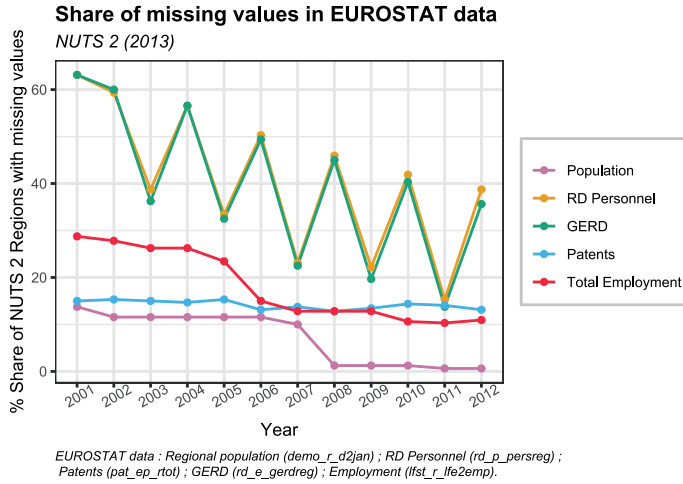
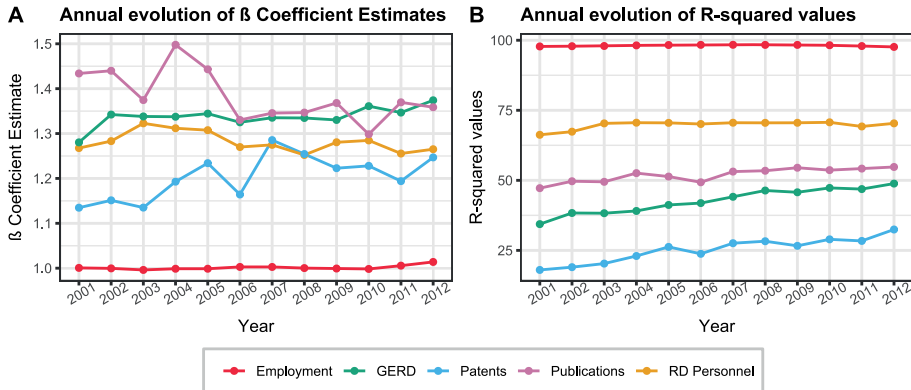


Figure 1: Share of missing values in Eurostat data at the NUTS 2 level.

the publication data to demographic data from Eurostat and to compare their spatial polarization to that of other R&D indicators, we mainly aggregate the publication data at the NUTS 2 level (which corresponds to national sub-regions). As a result, we conduct this analysis at a less precise spatial resolution than that of previous studies on the world geography of science.

We have chosen to start the analysis in 2001 because of the amount of missing data before this year in the Eurostat database and to stop the analysis in 2012 because patent data have stopped being registered after this date. In addition, post-2013 reforms in countries such as France and Poland have modified the administrative level of collect of R&D data, creating a discontinuity in the time series and making it difficult to follow a geographical evolution from 2001 to 2020 at the NUTS 2 level. The frequency of collection of R&D data following the recommendations of the Frascati Manual varies across country. As shown in Figure 1, certain countries such as Germany only collect these data biannually. For this reason, we have applied an interpolation method to estimate the missing values when the number of missing values for one region was strictly inferior to 11.

In order to include all the countries of the EU-28 perimeter together with Iceland, Norway, and Switzerland in the analysis, we have chosen to make other specific adjustments. As reflected by the important share of missing values at the NUTS 2 level (Figure 1), some countries report Eurostat data only at national level, such as Switzerland and Denmark (before 2007). In certain areas, the data are only available at an intermediate level (NUTS1), such as for Lithuania, Slovenia, Belgium, and the Southern Finland region, but also for the regions of Sachsen and Bavaria in Germany, as well as for Scotland and North-West England in the UK. In order to make the London area and the Budapest region comparable with other capital regions, we have also chosen to consider the NUTS1 level for these specific territories.



OLS Regression models: $\log(Y) = a + \beta \log(N) + e$ with $N =$ EUROSTAT Regional population (*demo_r_d2jan*).
 EUROSTAT data (RD Personnel : *rd_p_persreg* ; Patents : *pa_ep_rot* ; GERD : *rd_e_gerdreg* ; Employment : *fst_r_lfe2emp*).
 European perimeter: EU28 + NO, IS and CH. Spatial resolution: NUTS 2 2013 except for: UKI, UKD, UKM, DED, DE2, BE1, BE2, BE3, F11, HU1, LT0, SI0, CH0, DK0.
 For these areas, upper levels are used because of EUROSTAT data availability. NA observations are approximated using an interpolation method.
 Publications data: Web of Science/OST-HCERES (articles, reviews, letters). Fractionation at the level of urban areas (see Maisonneuve et al., 2018)

Figure 2: Annual evolution of the scaling parameters and of R-squared values.

By making these adjustments and by interpolating the data, we succeeded in removing all the missing values and covering 241 territories in our analysis.

This analysis is organized in three steps. First, we monitor the evolution of the scaling coefficient (beta) for all the three R&D indicators (R&D personnel, GERD, and patent applications) as well as for the spatial distribution of publication data. As a reference indicator, we consider the spatial distribution of employment data (all human sectors of activity combined) knowing that employment is expected to be less concentrated than R&D activity (in red on the figures). In the following steps, we compare the results of the first analysis with the results obtained by computing the Gini coefficient and the evolution of the percentage share of the top 30 European regions. To estimate the scaling coefficient, we use a classical Ordinary Least Square regression method between the considered variables and the number of inhabitants by regions (regional population).

As expected, employment exhibits a linear regime of scaling (beta equals to one) whereas R&D activity appears more concentrated in the most populated regions at the European scale (beta superior to one). However, the scaling parameter of R&D personnel and publication outputs has decreased over the period (Figure 2 A). While it appears the scaling parameter of R&D expenditure and patents has increased over the period, this information is subject to caution since the goodness of fit (R^2) is inferior to 50% for both indicators (Figure 2 B), meaning that these indicators poorly fit to population data. When, following the advice of Finance and Swerts regarding the interpretation of scaling parameters (Finance and Swerts, 2020), we consider their 95% confidence intervals (Figure 3) and note that the observed changes over the years have been too limited to be interpreted as changes in spatial polarization.

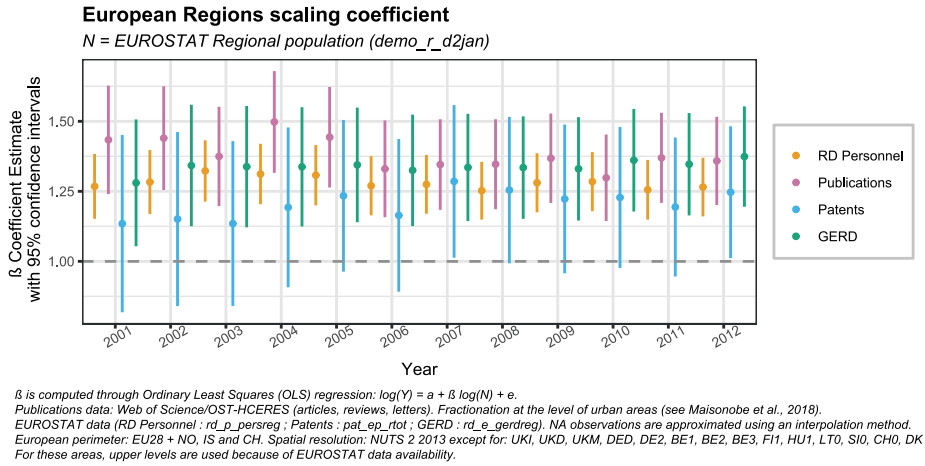
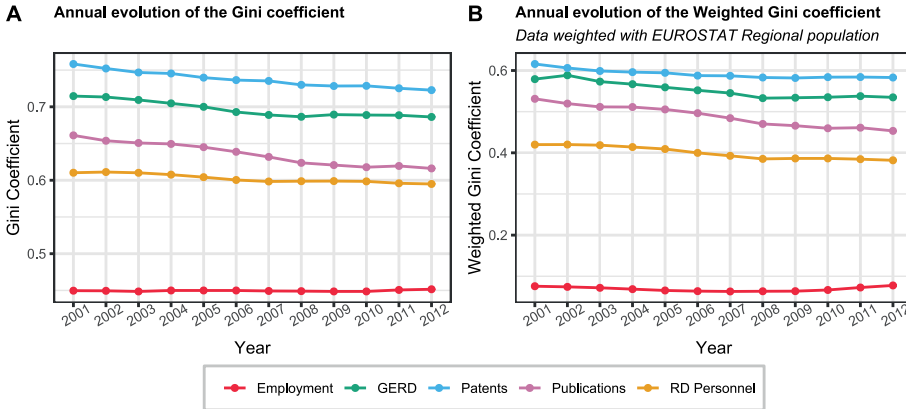


Figure 3: Annual evolution of the scaling parameters and of R-squared values.

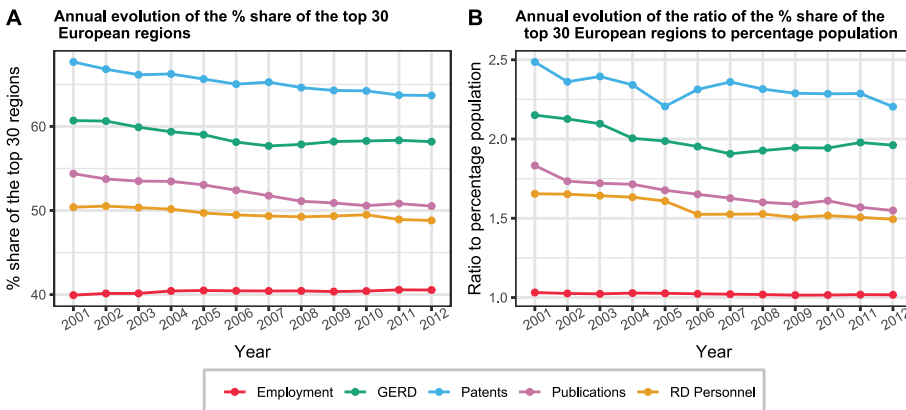
Even if the measured changes are limited, the value of the Gini coefficient (Figure 3 A) as well as the one of the Gini coefficient weighted by the number of inhabitants (Figure 3 B) both exhibit a decrease of concentration for all indicators except the reference indicator (total employment). This analysis suggests a movement of greater convergence of European regions in terms of R&D activity consistent with the one measured by Zitt et al. (1999) for the 1990s and by Dolores Leon et al. (2011) for the 2000s relying on the evolution of the Gini index. While R&D outputs and expenditures still appear more concentrated than R&D personnel this gap appears to be narrowing over the period, especially in terms of publications. After 2008, the implementation of excellence policies have developed across European countries leading to a growing concentration of research funds in certain areas, which could explain that the concentration of R&D expenditures remains stable after this date.

As these policies tend to increase geographical inequalities among scientists, several authors have pointed out their effects in recent years. Indeed, it seems that if a process of dispersal of R&D activity is at work, it is important that the allocation of funding can accompany this process (Aagaard et al., 2019; Langfeldt et al., 2015; Larivière et al., 2010). When analyzing the evolution of the per cent share of the top 30 European regions the gap between the distribution of R&D expenditure and that of R&D personnel appears even wider (Figure 4 A). While the top 30 European regions concentrate about 60% of the total R&D expenditures, they only represent about 50% of the R&D personnel. Considering the ratio between the percentage share in R&D activity and the population share (Figure 4.B) does not make this difference less important.



EUROSTAT data (RD Personnel : rd_p_persreg ; Patents : pat_ep_tot ; GERD : rd_e_gerdreg ; Employment : lst_r_lfe2emp ; Population : demo_r_d2jan). European perimeter: EU28 + NO, IS and CH. Spatial resolution: NUTS 2 2013 except for: UKI, UKD, UKM, DED, DE2, BE1, BE2, BE3, FI1, HU1, LT0, SI0, CH0, DK0. For these areas, upper levels are used because of EUROSTAT data availability. NA observations are approximated using an interpolation method. Publications data: Web of Science/OST-HCERES (articles, reviews, letters). Fractionation at the level of urban areas (see Maisonobe et al., 2018)

Figure 4: Annual evolution of the Gini coefficient and of the Weighted Gini coefficient.



EUROSTAT data (RD Personnel : rd_p_persreg ; Patents : pat_ep_tot ; GERD : rd_e_gerdreg ; Employment : lst_r_lfe2emp). European perimeter: EU28 + NO, IS and CH. Spatial resolution: NUTS 2 2013 except for: UKI, UKD, UKM, DED, DE2, BE1, BE2, BE3, FI1, HU1, LT0, SI0, CH0, DK0. For these areas, upper levels are used because of EUROSTAT data availability. NA observations are approximated using an interpolation method. Publications data: Web of Science/OST-HCERES (articles, reviews, letters). Fractionation at the level of urban areas (see Maisonobe et al., 2018)

Figure 5: Annual evolution of the activity share and ratio of the top 30 European regions.

The results obtained when comparing the Gini coefficient with the evolving share of the top 30 regions are equivalent. Weighting and normalizing these indicators with the regional population does not modify the observations. The evolution of the scaling parameter, on the other hand, exhibits interesting differences. The scaling parameter for patents and R&D expenditures is smaller than that of publication activity, which is the opposite of what we measure with the value of the Gini coefficient and that of the top regions' percentage share. Since the goodness of fit is very low for these two indicators, it is possible that another model will better adjust to their distribution and thus better describe their spatial polarization (Leitão et al., 2016). Be-

yond this, other regression methods could also enable to account for zero values, which would be useful for long-term analyses of R&D activity distribution (Finance and Cottineau, 2018). In addition, the choice of regions as the level of analysis is disputable since the scaling method lends itself to analysis at the city level. Thus, although the scaling approach can be very interesting, especially at the level of a given national system, it requires accurate data at a fine level of analysis and a lot of caution in interpretation.

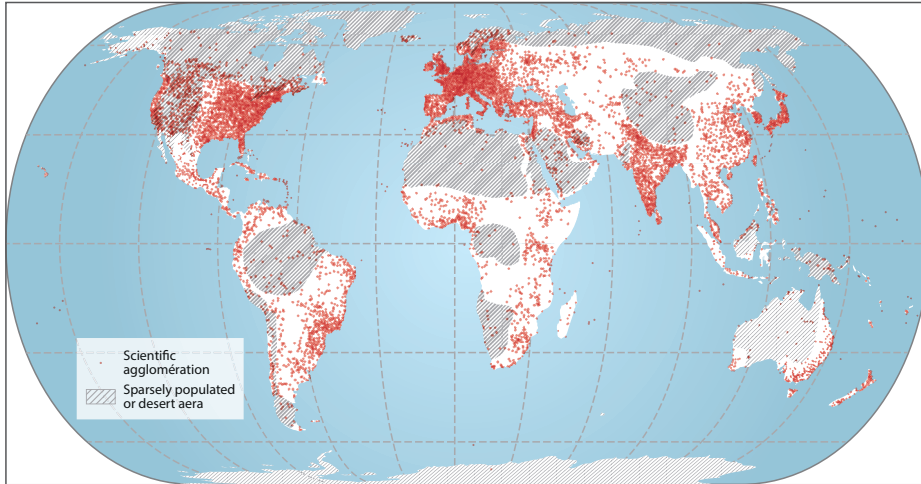
Despite these limitations, the three approaches show that R&D activities are geographically more concentrated than the active population at the level of the EU28. The Gini and activity share approaches indicate a gradual convergence between regions with a lower level of concentration at the end of the period. In particular, the gap between publishing activity and the distribution of research staff tends to narrow. The script and datasets used for this analysis are available at the following repository: <https://framagit.org/MarionMai/the-regional-distribution-of-research-in-europe> (July 15, 2020).

New Prospects for the Geography of Science

Beyond an Institutional Geography

Previous sections have highlighted the existence of several methods and numerous indicators to measure the degree of polarization of research activities. Depending on the territories, data are unevenly available, and, in the case of bibliographic sources such as the Web of Science or Scopus to measure publishing activity, there are representation biases related to the completeness of the databases and the evolution of their coverage over time (Basu, 2010; Zitt et al., 2003). On the administrative data side, although OECD countries are encouraged to follow the recommendations of the Frascati manual, there are significant differences in the pace of collection and in territorial divisions that complicate comparisons between countries. To our view, the geography of science literature should more systematically take into account and discuss these different limitations in the future. To improve the reproducibility of research in this still emerging field, our team shared the dataset of urban areas used to produce analyses at the world level and launched a geospatial application called NETSCITY that can geocode, cluster, and map any bibliographical dataset at the urban area and country level (Maisonobe et al., 2019).

However, both when studying the distribution of scientific personnel and production, the available data only reflect an institutional geography in the sense that these are the reported locations of the activity: the location of the laboratory where the researcher is attached and, in the case of company research, possibly the location of the company's head office. Thus, a field study financed by the Occitania Region in France has shown that researchers working in the Institutes of Technology of the small-sized town of Figeac signed their publications from Toulouse,



Map 1: Places from which at least one publication has been published between 1999 and 2014 (by Laurent Jégou).

where the headquarters of their lab is located (Chauvac and Naverau, 2019). Moreover, an important part of the scientific activity is not conducted within laboratory sites: applied sciences require fieldwork to be performed; history, archaeology, and humanities require the conducting of archival research and/or excavations; and many scientists, whatever their disciplines, praise the virtue of literary retreats and summer schools for writing and exchanging ideas. As a result, the geography obtained through administrative data and authors' affiliations only reflects partially the actual distribution of research.

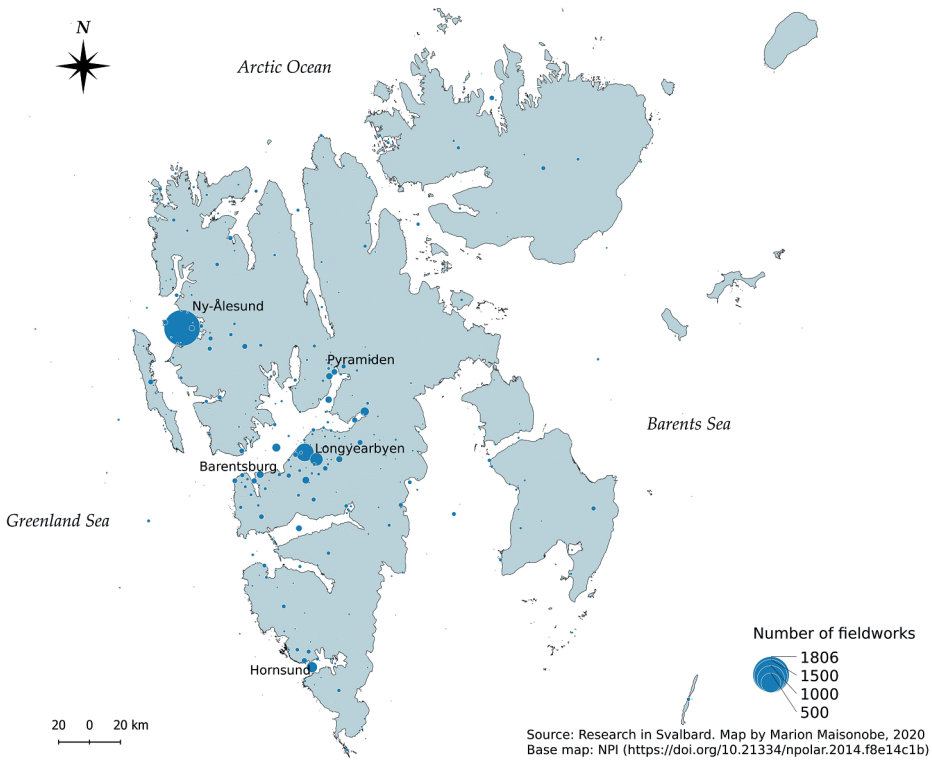
Nevertheless, when mapping all the places where at least one publication registered in the Web of Science has been signed during the 2000s according to authors' affiliations, we can observe that this covers some very remote areas (Map 1). If this is the case, it is because certain scientists sign from the actual place where they have done their research, but this is not often the norm, in particular because of the competition between universities to appear well ranked at world level. Since authors' affiliations might not be the best indicator of where research is actually done, other approaches could be used such as retrieving the name of the places mentioned by the authors in the title, the abstract or even the full text of their publications. Another possibility is to refer to other available sources, such as databases registering the activity of scientists on their field ground. In the next subsection, we give the example of the Research in Svalbard (RIS) database.

The Case of the Research in Svalbard (RIS) Database

Svalbard is a territory located just below the exact location of the North Pole. Under the sovereignty of Norway, this archipelago is home to 13 scientific stations of eleven countries and welcomes every year a considerable number of researchers who carry out measurements and observations (Strouk, 2020). Much of the research carried out in this territory is related to global warming and its effects on fauna, flora, permafrost, and glaciers. As these topics have become more important in recent years, the number of researchers conducting research in Svalbard has tended to increase. In order to be able to control the situation, and to ensure the safety of all, Norway has set up a monitoring system. Researchers coming to do their research at Ny-Ålesund stations must indicate in advance on the Research in Svalbard (RIS) platform the project they are going to carry out there and fill in the GPS coordinates of the specific sites where they will be working, the number of times, and the period they will visit them. Researchers visiting other localities in Svalbard are also invited to detail their project on the platform so that RIS can also be a tool for coordination and possibly convergence between similar projects. However, it is likely that the database obtained does not cover all the projects carried out in the area. Nevertheless, this source appears well built and interesting enough to give an overview of a rarely considered aspect of the geography of research.

The map below highlights the sites visited by all scientists who have recorded a project in the platform before January 2020 and the number of fieldworks carried out and ongoing in these sites. To improve the readability of the map, the points located within a radius of one kilometer from each other have been aggregated. It can be observed that the stations located in Ny-Ålesund, Longyearbyen, Barentsburg, and Hornsund regularly feature in the list of places mentioned because they serve as stopover points for the majority of missions. They can be the place where boats are refuelling and where the scientists rest for a few days before moving on. In addition, a large number of sites declared by scientists are located near these towns. For example, the Kongsfjorden setting with the Kronebreen and Kongsvegen glaciers near Ny-Ålesund on the one hand and the Adventdalen valley near Longyerbyen, which is home to the University of Svalbard, on the other, are two natural areas most visited by scientists who have registered projects on the platform. While it is obvious, for logistic reasons, that the geography of the stations influences the distribution of research carried out in such a remote territory, it is also observed that a very large number of ad hoc missions are carried out at greater distances from the stations as well as at sea (small blue dots on the map). Thus, even in the northernmost part of our world, it is possible to find mechanisms for the polarization of research activities with zones of attraction around research stations and less frequented areas. This being said, the occupation of this territory by researchers of an increasingly diverse number of nationalities is a good indication of the increasingly globalized nature of research and its presence even in very remote lands, at a dis-

tance from major urban centers. Yet, this activity would not have been visible through traditional spatial bibliometrics approaches.



Map 2: Number of fieldworks per site registered in the Research in Svalbard database.

Conclusion

In this chapter, we have returned to the apparent paradox between the literature showing a continuous movement of spatial deconcentration of research activity, between and within countries, since the mid-twentieth century, and the literature stressing the metropolitan and naturally concentrated nature of research activities. We have shown that the latter, inspired by regional economics, was mainly based on observations made at a given date but not on longitudinal analyses. In order to overcome the apparent contradictions between these approaches, we have highlighted the different methods available to measure the spatial concentration of research activities and the different indicators available. An application of these different methods to the case of European data indicates that even at the European level, where research systems are mature, there has continued to be a slight devolution of research activity between 2001 and 2012. Thus, although universities and research

centers are obviously not as dispersed as some community-based services, and thus continue to be more highly developed in some cities, regions, and countries, their polarization is not as great as it once was. Beyond the variety of methods, the question of data accessibility, the level of spatial resolution of the data, and their completeness is addressed. Finally, we demonstrate the opportunity of increasing the number of sources that can inform the geography of research. In particular, it is proposed to go beyond traditional mapping approaches based on affiliation data and public statistics, to work towards investigating not only an institutional geography of research but also a geography of science taking into account the very places where research is carried out, places that can be extremely remote and at a distance from urban centers.

References

- Aagaard, K, Kladakis, A & Nielsen, MW 2019, 'Concentration or dispersal of research funding?', *Quantitative Science Studies*, vol. 1, no. 1, pp. 117–49, doi: https://doi.org/10.1162/qss_a_00002.
- Arcaute, E, Hatna, E, Ferguson, P et al. 2015, 'Constructing cities, deconstructing scaling laws', *Journal of The Royal Society Interface*, vol. 12, no. 102, doi: <https://doi.org/10.1098/rsif.2014.0745>.
- Basu, A 2010, 'Does a country's scientific 'productivity' depend critically on the number of country journals indexed?', *Scientometrics*, vol. 82, no. 3, pp. 507–16, doi: <https://doi.org/10.1007/s11192-010-0186-8>.
- Batty, M 2003, 'The geography of scientific citation', *Environment and Planning A* 35, pp. 761–70, doi: <https://doi.org/10.1068/a3505com>.
- Ben-David, J 1991, *Scientific Growth: Essays on the Social Organization and Ethos of Science*, in G Freudenthal (ed.), Los Angeles: Univ. of California Press.
- Bergé, LR, Wanzenböck, I & Scherngell, T 2017, 'Centrality of regions in R&D networks: a new measurement approach using the concept of bridging paths', *Regional Studies*, vol. 51, no. 8, pp. 1165–78, doi: <https://doi.org/10.1080/00343404.2016.1269885>.
- Bettencourt, LMA, Lobo, J, Helbing, D et al. 2007, 'Growth, innovation, scaling, and the pace of life in cities', *Proceedings of the National Academy of Sciences*, vol. 104, no. 17, p. 7301, doi: <https://doi.org/10.1073/pnas.0610172104>.
- Chauvac, N & Naverau, B 2019, *Localisation Géographique de la Production Scientifique en Occitanie*, Rapport intermédiaire, Toulouse, CCRRDT OCCITANIE.
- Dolores León, M, Fernández, AM & Flores, E 2011, 'Scientific and Technological Knowledge of Universities in the EU-15: Implications for Convergence', *European Planning Studies*, vol. 19, no. 4, pp. 683–703, doi: <https://doi.org/10.1080/09654313.2011.548468>.
- Finance, O & Cottineau, C 2018, 'Are the absent always wrong? Dealing with zero values in urban scaling', *Environment and Planning B: Urban Analytics and City Science*, vol. 46, no. 9, pp. 1663–77, doi: <https://doi.org/10.1177/2399808318785634>.
- Finance, O & Swerts, E 2020, 'Scaling Laws in Urban Geography. Linkages with Urban Theories, Challenges and Limitations', in D Pumain (ed.), *Theories and Models of Urbanization: Geography, Economics and Computing Sciences*, pp. 67–96, Springer International Publishing, Cham, doi: https://doi.org/10.1007/978-3-030-36656-8_5.
- Florida, RL 2005, 'The world is spiky', *The Atlantic*, vol. 296, no. 3, pp. 48–51.

- Frick, SA & Rodríguez-Pose, A 2018, 'Change in urban concentration and economic growth', *World Development*, vol. 105, pp. 156–70, doi: <https://doi.org/10.1016/j.worlddev.2017.12.034>.
- Gottmann, J 1973, 'The need of an international policy for the sciences', in *Nobel Symposium No. 26*, pp. 7–13, Oslo.
- Grossetti, M, Eckert, D, Gingras, Y et al. 2013, 'Cities and the geographical deconcentration of scientific activity: A multilevel analysis of publications (1987–2007)', *Urban Studies*, doi: <https://doi.org/10.1177/0042098013506047>.
- Hong, W 2008, 'Decline of the center: The decentralizing process of knowledge transfer of Chinese universities from 1985 to 2004', *Research Policy*, vol. 37, no. 4, pp. 580–95, doi: <https://doi.org/10.1016/j.respol.2007.12.008>.
- Inhaber, H 1977, 'Changes in centralization of science', *Research Policy*, vol. 6, pp. 178–93, doi: [https://doi.org/10.1016/0048-7333\(77\)90024-5](https://doi.org/10.1016/0048-7333(77)90024-5).
- Langfeldt, L, Benner, M, Siverstsen, G et al. 2015, 'Excellence and growth dynamics: A comparative study of the Matthew effect', *Science and Public Policy*, vol. 42, no. 5, pp. 661–75.
- Larivière, V, Macaluso, B, Archambault, É et al. 2010, 'Which scientific elites? On the concentration of research funds, publications and citations', *Research Evaluation*, vol. 19, no. 1, pp. 45–53, doi: <https://doi.org/10.3152/095820210X492495>.
- Leitão, JC, Miotto, JM, Gerlach, M & Altmann, EG 2016, 'Is this scaling nonlinear?', *Royal Society Open Science*, vol. 3, no. 7, 150649, <https://doi.org/10.1098/rsos.150649>.
- Maisonobe, M, Grossetti, M, Milard, B et al. 2016, 'The global evolution of scientific collaboration networks between cities (1999–2014): multiple scales', *Revue française de sociologie*, vol. 57, no. 3, pp. 417–41, doi: <https://doi.org/10.3917/rfs.573.0417>.
- Maisonobe, M, Grossetti, M, Milard, B et al. 2017, 'The global geography of scientific visibility: a deconcentration process (1999–2011)', *Scientometrics*, vol. 113, no. 1, pp. 479–93.
- Maisonobe, M, Jégou, L & Cabanac, G 2018, 'Peripheral Forces', *Nature Index*, no. 563, S18-S19.
- Maisonobe, M, Jégou, L, Yakimovich, N et al. 2019, 'NETSCITY: a geospatial application to analyse and map world scale production and collaboration data between cities', in *ISSI'19: 17th International Conference on Scientometrics and Informetrics*, Rome.
- Matthiessen, CW, Schwarz, AW & Find, S 2010, 'World Cities of Scientific Knowledge: Systems, Networks and Potential Dynamics. An Analysis Based on Bibliometric Indicators', *Urban Studies*, vol. 47, no. 9, pp. 1879–97, doi: <https://doi.org/10.1177/0042098010372683>.
- Nomaler, Ö, Frenken, K & Heimeriks, G 2014, 'On Scaling of Scientific Knowledge Production in U.S. Metropolitan Areas', *PLOS ONE*, vol. 9, no. 10: e110805, doi: <https://doi.org/10.1371/journal.pone.0110805>.
- Pumain, D, Paulus, F, Vacchiani-Marcuzzo, C et al. 2006, 'An evolutionary theory for interpreting urban scaling laws', *Cybergeo: European Journal of Geography*, no. 343, doi: <https://doi-org.inshs.bib.cnrs.fr/10.4000/cybergeo.2519>.
- Schofer, E & Meyer, JW 2005, 'The Worldwide Expansion of Higher Education in the Twentieth Century', *American Sociological Review*, vol. 70, no. 6, pp. 898–920, doi: <https://doi.org/10.1177/000312240507000602>.
- Schott, T 1991, 'The world scientific community: Globality and globalisation', *Minerva*, vol. 29, no. 4, pp. 440–62.
- Strouk, M, 2020, 'Science en puissances : la recherche scientifique, porte d'entrée vers la gouvernance arctique pour les États observateurs', to be published.
- Zitt, M, Barré, R, Sigogneau, A et al. 1999, 'Territorial concentration and evolution of science and technology activities in the European Union: a descriptive analysis', *Research Policy*, vol. 28, no. 5, pp. 545–62, doi: [https://doi.org/10.1016/S0048-7333\(99\)00012-8](https://doi.org/10.1016/S0048-7333(99)00012-8).

Zitt, M, Ramanana-Rahary, S & Bassecouard, E 2003, 'Correcting glasses help fair comparisons in international science landscape: Country indicators as a function of ISI database delineation', *Scientometrics*, vol. 56, no. 2, pp. 259–82.