Journal of Informetrics 11 (2017) 976–988

Contents lists available at ScienceDirect

Journal of Informetrics

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

Review

A spatial scientometric analysis of the publication output of cities worldwide

György Csomós

University of Debrecen, Department of Civil Engineering, 2-4 Ótemető u., Debrecen 4028, Hungary

ARTICLE INFO

Article history:
Received 4 April 2017
Received in revised form 16 August 2017
Accepted 17 August 2017

Keywords:
Scientific output
Journal articles
Spatial scientometrics
Geovisualisation
Scopus

ABSTRACT

In tandem with the rapid globalisation of science, spatial scientometrics has become an important research sub-field in scientometric studies. Recently, numerous spatial scientometric contributions have focused on the examination of cities' scientific output by using various scientometric indicators. In this paper, I analyse cities' scientific output worldwide in terms of the number of journal articles indexed by the Scopus database, in the period from 1986 to 2015. Furthermore, I examine which countries are the most important collaborators of cities. Finally, I identify the most productive disciplines in each city. I use GPS Visualizer to illustrate the scientometric data of nearly 2200 cities on maps. Results show that cities with the highest scientific output are mostly located in developed countries and China. Between 1986 and 2015, the greatest number of scientific articles were created in Beijing. The international hegemony of the United States in science has been described by many studies, and is also reinforced by the fact that the United States is the most important collaborator to more than 75% of all cities. Medicine is the most productive discipline in two-thirds of cities. Furthermore, cities having the highest scientific output in specific disciplines show well-defined geographical patterns.

1. Introduction

The University of Oxford (England) is one of the most prestigious universities in the world, with a remarkable scientific output[1] (Bonaccorsi, Haddawy, Cicero, & Hassan, 2017; Lin, Huang, & Chen, 2013). It occupies the top position in nearly every university ranking (for example, it is ranked sixth in QS World University Rankings[2] 2016–2017). Tianjin University, the first modern university in both Tianjin and China, occupies a mid-level position in worldwide university rankings (for example, it is only in the 481–490th position in QS World University Rankings 2016–2017), and its scientific output is much smaller than that of the University of Oxford. The University of Oxford and Tianjin University are clearly not on the same tier in terms of scientific output. However, in this spatial scientometric study, I aim to measure and geovisualise scientometric data of cities, not of organisations. The main goal of this paper is to examine cities' scientific output, analyse the characteristics of their international collaboration, and present which scientific disciplines are the most productive ones in each city. Using the aforementioned example as an illustration, I aim to compare the scientific output of the cities of Oxford and Tianjin; i.e.,

E-mail address: csomos@eng.unideb.hu

[1] Both the definition and the measurement method of scientific output have key importance in this analysis, and will be presented thoroughly in "Data and Methods".

[2] QS World University Rankings 2016–2017: https://www.topuniversities.com/university-rankings/world-university-rankings/2016.

http://dx.doi.org/10.1016/j.joi.2017.08.006
1751-1577/

the scientific output of a mid-sized city with fewer than 170,000 inhabitants against the scientific output of a megacity with more than 15 million people.

According to Frenken, Hardeman, and Hoekman (2009), the first studies discussing the spatial distribution of science were published in the 1970s; however, spatial scientometrics has only recently begun to attract more attention (Bornmann & Waltman, 2011). The geospatial measurement of cities' (entire urban regions') scientific output using various scientific indicators appears in many studies. Matthiessen and Schwarz (1999) examine the scientific output of European urban regions in terms of the number of papers in the Science Citation Index (SCI). Zhou, Thijs, and Glänzel (2009a) analyse the scientific output of Chinese (provincial-level) administrative regions in terms of the number of scientific publications collected in the SCI Expanded database. In an article published in *Nature News*, van Noorden (2010) discusses which urban regions produce the best research in the world, and whether their success can be replicated. Bornmann, Leydesdorff, Walch-Solimena, and Ettl (2011) and Bornmann and Waltman (2011) classify cities and urban regions as "centers of excellence in scientific research", based on the total number of excellent papers (top 1% most highly cited papers). Bornmann and Leydesdorff (2011), and Bornmann and Leydesdorff (2012) examine whether there is a relationship between the total scientific output of cities and the number of highly cited papers published in those cities. Csomós and Tóth (2016) and Csomós (2017) explore the global position of cities in terms of corporate research and development, based on the number of scientific publications created by corporate researchers and engineers.

These studies are limited by their focus on specific geographic regions (e.g., Europe, China, etc.) or specific research areas (e.g., neuroscience, physics and astronomy, etc.). In this paper, I examine cities' scientific output based on Scopus data, in a search for answers to three research questions:

- Which cities in the world have the highest scientific output, and how has this output changed over time?
- For a given city, which countries are its most important collaborators?
- Which disciplines are the most productive in each city?

The structure of the paper is as follows: In Section 2, I describe the data collection process and methodology. In Section 3, a parallel is made between the results of the study and the answers to the research questions. In Section 4, I present the limitations of the analysis; and finally, in Section 5, I present the conclusions.

2. Data and methodology

2.1. Data collection

Before presenting the results, I will address two important issues that need to be clarified: First, how can scientific output of cities be measured, and second, which cities should be examined in terms of scientific output?

2.1.1. Determination of the source and type of scientific publications

In this analysis, cities' scientific output is obtained by measuring the total number of publications written by authors who are affiliated with a professional organization (e.g., universities, firms, hospitals, governmental and non-governmental institutions, etc.) in that city. Scientific publication data is indexed by several databases. To achieve the main goals of this spatial scientometric analysis, Web of Science and Scopus are chosen as the most relevant databases containing scientific publication data. Both databases include the authors' affiliations (i.e., names, addresses) and sum up the total number of articles created by authors affiliated with specific organisations in the same city. Furthermore, both databases assign a publication to the cities in which they were produced, rather than the cities in which the affiliated organisations' headquarters are located.[3] This is an important distinction, because some large organisations (universities, multinational corporations, national or international research institutions) house their semi-autonomous but scientifically productive units (university hospitals, corporate research centres) in one city and their command and control centres in another city (corporate headquarters, universities' main campuses, etc.). For example, IBM, a corporation that is among the world's leaders in scientific output, is headquartered in Armonk, New York, but only 17% of the corporation's total number of publications were created here (i.e., the authors have indicated this location as the institutional address on their publication). Scopus assigns 18 additional affiliations to IBM, including the Thomas J. Watson Research Center in Yorktown Heights, New York, where 47% of IBM's publications were created. As a matter of fact, both Armonk and Yorktown Heights are located within the New York-Newark metropolitan area, whereas the IBM Almaden Research Center, which has roughly the same number of publications as the Armonk headquarters, is located in San Jose, California. Other scientifically productive IBM subsidiaries are in Ruschlikon, (Switzerland), Bangalore (India), Haifa (Israel), Tokyo, and Beijing. Obviously, the number of scientific publications created

[3] It should be noted that both Web of Science and Scopus present the addresses (i.e., the name of cities) reported by the authors of a publication. In most cases, these are the addresses at which a publication was produced. However, it is generally at the authors' discretion to choose what address they wish to indicate on a publication, and if they choose to report the address of their organization's headquarters, then this will be the information that is presented in Web of Science and Scopus.