EXAMINING THE TREND OF RESEARCH ON CHEMOMETRIC ANALYSIS: A BIBLIOMETRIC REVIEW

Sarifah Fauziah Syed Draman, Nur Izzatul Qayyimah Ahmad, Anis Suffia Suliman, Mohammad Alif Aiman Shamsu Kamal Bahrin, Nur Amyliya Azahari

Faculty of Chemical Engineering

Universiti Teknologi MARA (UiTM) Terengganu Branch, Bukit Besi Campus,

23200 Dungun, Terengganu, Malaysia

Corresponding author: sfauziah@uitm.edu.my

Abstract

Bibliometrics analysis is a technique that refers to the collection, the handling, and the quantitatively assess of bibliographic data, resulting from scientific publications. This bibliometric review aimed to examine the trend of research on chemometric analysis over the past 37 years (1983-2020). Chemometrics is the chemical discipline that uses mathematical, statistical and other methods to design or select optimal measurement procedures and experiments. By analysing chemical data, chemometric analysis can provide the full relevant chemical information needed. Various areas and principles such as multivariate statistics and pattern recognition, time series analyses, information technology, computation, numerical analysis, operations research, decision theory and operations research contribute to chemometrics. This bibliometric review was applied to examine the chemometric analysis research trend between 1983-2020 using the Scopus-indexed documents. There are many research types on the chemometric analysis carried out by the researcher during the study period. This study's finalised 1299 documents were conducted overall for further analysis based on the' title' search results. Standard bibliometric indicators were used in this review analysis. Two software is used (Harzing Perish and VOSviewer) in completing the bibliometric review. This study found 12 types of documents that have been published related to chemometric analysis. 89.92 % of the total publications were from articles. The highest source type is journals (94.53%). The first research on chemometric analysis was published in the year of 1983. Then, the growth of the related publications has risen gradually. However, there is some decrease in publications between the years. English is the top language used in the publications. An analysis by country, China is ranked first in productivity with 225 (13.64%) published documents. Citation analysis identified key authors and documents that have shaped the progress of this review. Author co-citation analysis used to determine the intellectual structure of the chemometric analysis knowledge base. In conclusion, this review can better understand development trends that have emerged in this chemometric analysis over the past 37 years. Furthermore, the review result can also offer a reference for future research.

Keywords: Chemometric Analysis, Chemometric, Bibliometric Review, Scopus, Harzing Perish, VOSviewer

Table of Contents

1 Introduction

There are two roots in bibliometrics, which are' biblio' and' metrics.' The word "biblio" is derived from a book's combination meaning "biblion" in Latin and Greek. For example, the measurement is derived either from the Latin or Greek word "metrics" or "metricus" respectively, each meaning measurement, the term "metrikos" representing the science of measurement (Sengupta, 1992) that cited by Osareh (1996). The application of mathematics and mathematical analysis to books and other media of communications is discussed through bibliometrics. Bibliometrics analysis describes the quantitative treatment of the discourse's characteristics and behaviour documented (Pritchard, 1969) cited by Osareh (996). Bibliometrics analysis aims to shed light on written communication mechanisms, the structure of a discipline, and its growth route. By counting and evaluating the different aspects of written communication, written communication is seen by written communication (Pritchard, 1969:348, 1972:38-39) cited by (Osareh, 1996).

The chemical discipline that uses the mathematical and the statistical methods to plan or select the optimal calculation of the procedures and experiments and to provide complete knowledge with the study of chemical data is the real concept of chemometrics (Swede et al., 1972) that cited by (Otto, 2007). Chemometrics is a discipline sourced the chemistry. The development of the quantitative structure behaviour relationships or the measurement of analytical-chemical data is the standard applications of the chemometric methods. One explanation for the data flood produced by modern analytical instrumentation is that analytical chemists develop chemometric strategies in particular (Otto, 2007). Chemometrics is the study of using the quantitative or statistical techniques to link the observations made on a chemical device or process to the system's state (Lavine, B. and Workman, J., 2008). A chemometric analysis is especially useful for food science, plant extracts, and biological research (Jokić, S. et al., 2012).

In general, the number of citations for chemometrics continues by seeing steady growth amongst all fields. The numbers of publications published can be related to developing new chemometric methods (Brown, S. D. et al., 1996). In this review, the Scopus database is being used as the reference to get the data of chemometric analysis. Scopus is the most extensive base of peer-reviewed scientific literature abstracts and bibliographic references. Over 18,000 titles from 5,000 publishers allowing for multidisciplinary, international integration regularly updated data were then examined for thirty-seven years from 1983 until 2020 related to the chemometric study to explore the research trend. The data on chemometric analysis study is based on the title only. In bibliometric review, essential information, such as the author profile and citation tracker, can provide a complete view of the trends. The index includes the 'H' tool showing the findings of this analysis. In format tables and or graphics, the results were can presented the rankings.

Furthermore, to visualise the results graphically, the VOS viewer software is used. The VOS viewer is software designed by a Leiden University research team that extracts data from databases such as the Web of Science or Scopus and creates graphs describing the result of bibliographic links. Finally, Harzing Publish or Perish is another software that is used for bibliometric analysis result. It is a software program that retrieves academic citations and analyses them (Van Eck and Waltman, 2009). In summary, bibliometric review involves the uses of a set of data sources, then analyses them and provides metrics such as the total number of articles and the total number of citations, the average

number of citations per paper, citations per author, papers per author and citations per year, the h-index of Hirsch and related parameters, and more.

1.1 Problem Statement

The trend for chemometrics research can be seen if there is consistent development in this field. The number of references can estimate distributions related to the advancement of new chemometric strategies. Bibliometric analysis techniques have been steadily gaining popularity as one of the approaches in revealing research trends/patterns (Ahmi et al., 2020). Scopus information base is being used as the reference to get the information of the chemometric analysis. The following research questions (RQ) are some of the standard indicators in the bibliometric review analysis. Then, it will be the focus of this study.

RQ1: What is the current publication trend in the chemometric analysis?

RQ2: Which are the most influential articles on chemometric analysis?

RQ3: Which themes involving chemometric analysis are the most popular among scholars?

RQ4: Who are the most influential authors on chemometric analysis?

RQ5: What is the current state of collaboration involving chemometric analysis?

RQ6: What is the intellectual structure of current research on chemometric analysis?

RQ7: What kinds of issues hamper research on chemometric analysis?

RQ8: What areas involving chemometric analysis need an additional study?

1.2 Objectives

Based on the problem statement, as mentioned in the previous section, the objectives of this study are:

1. To examine the trends of research in chemometric analysis using the bibliometric review technique
2. To describe the trends of chemometric research in standard bibliometric analysis indicators using Microsoft Excel, Harzing Publish or Perish and VOSViewer software.

1.3 Methods

1.3.1 Method to find document and source types, year of publications and authorship and co-authorship analysis.

Data obtained for this study is extracted from the Scopus database through the Ez Access UiTM website. Scopus is the database that provides a broad and different selection of sources for many branches of studies (Hallinger and Kovacevic, 2019; Mongeon and Paul-Hus, 2016; Udomsap and Hallinger, 2020). Scopus was chosen as one of the reliable databases as it has a wide range of publications that have been reviewed and offers bibliographic data for further understanding (Muritala et al., 2020).

This study's search is limited to only the title of the articles of "Chemometric Analysis". The documents under the title of "Chemometric Analysis" are finalised with 1299 documents starting from 1983 until 2020. The years that the research on this title has been about 37 years is more than three decades. The documents' refine values such as the subject areas, languages of the documents, and geographical distribution publications under the title are exported in comma-separated values file (CSV) for further analysis. The data also need to save in Research Information Systems (RIS) file. RIS file is a bibliographic citation file saved in a format developed by the research information system. RIS files provide information such as title, author, publication date, keywords, publisher, issue number, and start and end page.

Harzing Publish or Perish software was used to analyse the data selection by using the Scopus database's exported values to observe the trends and identification of the data. Other than that, VOSViewer is used to show the bibliographic data for the study and to project the bibliometric map networks (van Eck & Waltman, 2014; Mohamed et al., 2020). It is a practical tool to present the distance of main topics to show the relations and build maps based on the co-occurrence from the data obtained (Li et al., 2019).

1.3.2 Method to determine most productive authors in authorship and co-authorship analysis, languages of the documents and subject area.

The data were retrieved from the Scopus database via Ezaccess UiTM Libraries e-Resources website. This study's research trend is based on the title, "Chemometric Analysis" in December 2020. This study only focuses on the title of the articles. Bibliometric analysis was applied to examine the research trend on chemometric analysis over the past 37 years from 1983-2020. Based on the title of article search results, this study finalised 1299 documents obtained for further research. A few tools were used for this bibliometric review in completing the analysis. Scopus database is for data collection. Harzing Publish or Perish (www.harzing.com) software is used to obtain the citations and the metrics such as the total number of citations, average citations per paper, citations per author citations per year, g-index and h-index. Meanwhile, VOSviewer version 1.6.16 (www.vosviewer.com) software was used to create and visualise the bibliometric networks map based on co-authorship, co-occurrence, and citation (Ahmi et al., 2020).

In the Scopus database, we need to export all the refine values in CSV format which is we need to use all the data such as most productive authors, languages of the documents, subject area, and others. Microsoft Excel opened the CSV format that we exported. And also we need to export the data in RIS. Format and CSV Excel format that we need to use the RIS. The structure in the Harzing Publishes or Perish software to get the number of cited publications (NCP), total citations (TC), average citations per publication (C/P), average citations per cited publication, h index and g index. And for the CSV format, we use VOSviewer software to visualise the maps based on the data collected from the Scopus