Chemometrics and Intelligent Laboratory Systems 222 (2022) 104514

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Chemometrics in food science and technology: A bibliometric study

J.L. Aleixandre-Tudo [a,b,*], L. Castello-Cogollos [c,d], J.L. Aleixandre [a], R. Aleixandre-Benavent [d,e]

[a] Instituto de Ingeniería de Alimentos para el Desarrollo (IIAD), Universitat Politecnica de Valencia, Valencia, Spain
[b] South African Grape and Wine Research Institute (SAGWRI), Department of Viticulture and Oenology, Stellenbosch University, Matieland, South Africa
[c] Departamento de Sociología y Antropología Social, Universidad de Valencia, Valencia, Spain
[d] Unidad de Información e Investigación Social y Sanitaria (UISYS-CSIC), Universitat de Valencia, Valencia, Spain
[e] Ingenio (CSIC-Universitat Politecnica de Valencia), Spain

ARTICLE INFO

Keywords:
Chemometrics
Foodstuffs
Spectroscopy
Bibliometrics
PCA
PLS

ABSTRACT

Chemometrics has been defined as the discipline that provides maximum information from chemical data. Food science and technology applications use chemical data and are therefore suitable for chemometrics evaluation. Bibliometric studies provide an enhanced understanding of the progression, research status as well as future trends of a research field. The main aim of the study was therefore, to provide a bibliometric evaluation of the research literature employing chemometrics techniques in food science and technology applications. For this, a search strategy including the single term chemometric* was performed. The metadata obtained from the bibliometric search was subsequently analyzed. Indicators of the scientific productivity and quality such as the number of articles, citations, or funding activity, were obtained in combination with the most relevant keywords, authors, and countries. The progression over time of the bibliometric indicators was also presented and discussed. Chemometrics appeared as a prolific and healthy research field with increased funding received in the last decade. PCA, PLS and DA are still the preferred methods for most applications. A big part of the research is related to the combined use of spectroscopy and chemometrics. Finally, China and Brazil appeared as the leading countries in applying chemometrics to foodstuffs.

1. Introduction

The first reference to the term "chemometrics" appeared in the literature in the 70's as the English translation of the Swedish term "kemometri" as proposed by Svante Wold [1–3]. One of the most accepted definitions of chemometrics states the following "Chemometrics is the chemical discipline that uses mathematical, statistical and other methods employing formal logic to design or select optimal measurement procedures and experiments, and to provide maximum relevant chemical information by analysing chemical data" [4]. As per the definition, it is clear that chemical data analysis is the main purpose of chemometrics [5]. A combination of mathematical and, therefore statistical methods is applied to qualitative and quantitative analysis, process analytical technologies (PAT), involving chemical data, or even design of experiments [3,6]. Moreover, it has been argued that despite the term appearing only in the 70s, chemometrics applications appeared much earlier together with the use of primary univariate statistics in chemical data [1]. Nevertheless, it seems to be consensus now that chemometrics involves the use of non-perfect data sets in a multivariate manner [1,5,7], which again would bring the birth of chemometrics later in time.

If we understand chemometrics as a multivariate exercise, the availability of more advanced computers marked the first explosion of chemometric methods. In the early 70s, intense research led to the development of multivariate approaches to interpret analytical chemical data in areas such as pattern recognition, factor analysis, spectroscopy (specially NIR) or even chromatography [3,8–10]. Despite these early advancements in the field, it seems that the explosion of chemometrics only happened a decade ago. More and more researchers worldwide were in demand of chemometrics packages. The natural progression of the field led to a very intense increase in the number of users along with a decrease in the number of experts chemometricians. Interestingly, this evidence raised the question of whether chemometricians are still needed or whether they will be needed in the near future [1,11].

Brereton et al. (2013) identified the reasons for such concerns. Limited success to establish chemometrics as a solid field of research, the

* Corresponding author. Instituto de Ingeniería de Alimentos para el Desarrollo (IIAD), Universitat Politecnica de Valencia, Valencia, Spain.
E-mail addresses: joaltu@upvnet.upv.es (J.L. Aleixandre-Tudo), lourdes.castello@uv.es (L. Castello-Cogollos), jaleixan@tal.upv.es (J.L. Aleixandre), rafael.aleixandre@uv.es (R. Aleixandre-Benavent).

https://doi.org/10.1016/j.chemolab.2022.104514
Received 18 November 2021; Received in revised form 21 January 2022; Accepted 1 February 2022
Available online 5 February 2022

reluctance of industry to fund long term chemometrics projects, the view of chemometrics as second-class statistics by expert statisticians or finally, the involvement of chemists themselves in chemometrics tasks rather than making use of expert chemometricians [1]. The research systems in the different countries seem also to be playing against chemometrics. Scientists are increasingly judged according to the amount of funding they can raise, which prevents them from spending time developing the core fundamentals of chemometrics and rather spending time in more highly funded applied fields. With already visible signs, the future of chemometrics is pointing towards a rapidly growing community of users with limited knowledge of core chemometrics alongside a decreasing number of leading pure chemometricians. In addition, it is nowadays possible to generate large amounts of data with novel advanced instrumentation such as imaging devices or mass spectrometers. The processing of these large and complex data sets using some of the most relevant machine learning tools appears to be one of the main challenges for the future of chemometrics [1].

As briefly discussed, data used in chemometric applications is typically unstructured, complicated, and messy, which is not the type of data pure statisticians are familiar. This fact is probably the main difference between statistics and chemometrics. Moreover, chemometrics can deal with problems associated with instrumental data such as peak resolution, presence of artefacts or data pre-processing. Data analysis using chemometrics can be divided into exploration (classification or discrimination) and prediction (calibrations) tasks [6]. The most suitable exploration technique is principal component analysis (PCA) [12]., As unsupervised exploratory analysis, classification and discrimination mostly use PCA decomposition with some distance computation. On the other hand, supervised discrimination and classification tasks, make use of projection approaches, i.e. PCA or partial least squares (PLS) [13]. Some of these methods are linear discriminant analysis (LDA), canonical analysis, or soft independent modelling of class analogy (SIMCA) [14, 15]. K-nearest neighbours (KNN) and support vector machines (SVM) are other commonly used methods for discrimination. Moreover, methods for three-dimensional data are becoming more and more popular for applications such as mass spectrometry, fluorescence, or time data. N-way methods, as they are also known, include the Tucker-3 algorithm, the more recent parallel factor analysis (PARAFAC) [16] or Tri-PLS, which provides a tridimensional extension of PLS.

Regrading quantification (calibration) methods, multiple linear regression (MLR), principal component regression (PCR) or PLS are the most common linear methods used. For nonlinear calibrations, PLS for nonlinear data, artificial neural networks (ANN) or the genetic algorithm [17] might be used. In addition, chemometrics data is also characterised by the need for data pre-processing. Mean centring and scaling are widely used in most chemometric applications. Nonlinear data pre-processing using the Kubelka-Munck or the Box-Cox transformations are also frequent for data linearisation. Data smoothing using the Savitzky-Golay method reduces de influence of random noise. Finally, the Fourier transform method was a game-changer for signal processing for spectroscopy applications with wavelet analysis presented as its natural progression for large data sets [18].

A previous bibliometric study showed chemometrics is a fast-growing and highly competitive research area [2]. The study was based on metadata obtained from six of the most relevant journals in chemometrics and reported the collaboration patterns among the main actors in the field [2]. In our study, the main aim was to evaluate the bibliometric findings in chemometrics but applied to the food science and technology subject category. We believe that this study is of special relevance as food science and technology is a field strongly linked with chemistry, analytical chemistry and, an increasing presence of process analytical technologies (PAT) and subsequently to chemometric applications. This bibliometric study provides a better understanding of the main contributors, their interactions and, the current and future main areas of interest in the field. In addition, this study discusses some of the question raised about the future of chemometricians and of chemometrics as a field.

2. Materials and methods

The single term "chemometric*" was used in the search strategy. The term was truncated with an asterisk to include "chemometric" and "chemometrics". The term chemometrics is a plural noun with a singular meaning, whereas chemometric is used as an adjective in chemometric methods [19]. The search was conducted in the topic field, including the title, abstract and keywords. The Web of Science core collection (WOS) database of Clarivate Analytics was used to extract the bibliometric metadata. The search was restricted to articles and reviews published in the Food Science and Technology subject category. Therefore, other record types such as notes, or editorial material, were excluded. No time limitation was considered. The search was conducted from the appearance of the first article in 1979 until 2020. A keyword standardisation task was performed. Synonyms, spelling variations, acronyms and derivation of the same term were standardised.

The scientific performance, as well as the scientific impact of a research topic, can be assessed using bibliometric studies. This study includes several indicators of the scientific impact, such as the number of papers, citations, impact factor of the published papers, and funding received. Moreover, a keyword analysis was also assessed to identify the most relevant topics and their progression over time. Finally, the scientific activity of authors contributing to the field was also evaluated. The Journal Citation (JCR) last available report when the search was conducted was used to extract the indicators of scientific impact.

To visually represent the relationships between qualitative and quantitative data, social network analysis (SNA) is used. A representation of the data using nodes and ties is attempted to measure and map complex relationships among data. Using spheres and connecting lines, SNA provides a visualization of the frequency and relationship between keywords and countries. The frequency of appearance and the co-occurrence between two terms is thus evaluated. The co-occurrence indicates the number of times keywords or countries are repeated in the same record. Network visualization is improved by defining a minimum threshold or number of publications that include the terms. VOSViewer and Pajek software are used for network representation and visualization [20]. Pajek software is based on the Kamada and Kawai (KK) algorithm for network generation. The algorithm considers an attraction-repelling force for each pair of nodes (*i,j*) that are connected. The algorithm finds an optimal proportion of connections with an ideal length in the graph. At the same time, the symmetry between the distances that connect each pair of nodes (*i,j*) is also maintained in the network [21,22]. The ideal distance is defined with the following formula in which L_0 represents the side length of the drawing frame and represents the network diameter, defined as follows:

$$l_{i,j} = \frac{L_0}{\max\limits_{i<j} d_{i,j}} \times d_{i,j}$$

Due to the often impossibility of optimising the above function for some type of networks, the algorithm uses a spring model to minimise the energy function E of the network topology using the below formula:

$$E = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{2} k_{i,j} \left(|p_i - p_j| - l_{i,j} \right)^2$$

In this function: $k_{i,j}$ represents the stiffness between the nodes i and j; $l_{i,j}$ the ideal distance between nodes i and j; and p_i and p_j are the positions of the nodes i and j, respectively. The KK algorithm finds a visual position for each pair of nodes i and j, with a proportional Euclidean distance to $l_{i,j}$. Subsequently the KK algorithm defines a diameter matrix that stores the theoretical distances ($d_{i,j}$) between the nodes. $d_{i,j}$ represents the hop counts between nodes and finds the shortest hop counts for the possible paths between the nodes i and j. For network interpretation, spheres' size represents the number of times an item was observed in the records, whereas the thickness of the lines indicates the intensity of the cooccurrences. Larger spheres indicate a higher frequency of appearance, whereas thicker connecting lines indicate a more intense co-occurrence of items. For this study, the item corresponded to keywords and countries.

3. Results

Articles, citations, and signatures. Table 1 shows the number of articles, citations, the ratio citation per article, the signatures, and the ratio signatures per article. The data is presented in 5-year periods from 1991. The articles published prior to 1991 are all grouped under one period due to the few references present. This first period, therefore, covers from the first record in 1979–1991. Interestingly, the number of published articles doubled every five years through the entire analyzed period to a final 2118 records for the last five years period. An increasing number of citations, in line with the number of publications was also observed, except for the last period. However, considering the recent publication of the articles in the last period and the number of citations received to date (20.645), the observed increasing trend is expected to continue. The ratio citation per article presented a maximum from 1996 to 2010, with the highest ratio observed during the period 2001–2005 with 50.29 citations per article.

Interestingly, a decrease in the ratio was identified for the 2011–2016 period with 27.08 citations per article. The doubling of the publications numbers not being accompanied by the same increase in the citation counts is causing the observed decrease in the citations per article. However, this can be solely due to the time from publication and the data should be revisited in the future to confirm or negate these results. Finally, a steady increase in signatures per article is also worth mentioning. From just above two signatures per article at the beginning of the period (2.44) to more than five as was recently observed (5.32) in the last five years period.

Funding *activity.* Table 2 provides information about the funding activity of the records obtained. The results show the number of papers and citations for the funded and nonfunded publications. Funding information is only available in WOS from 2011, and therefore, only the results of the last decade are included. An increase in the number of funded papers over the last ten years was observed, with a total of 81% of funded papers in 2020. From a proportion 70/30 funded/non-funded to a slightly above 80/20 proportion. In addition, a very similar trend can be observed from the citations received by funded and nonfunded papers. However, this is not standardised with the number of published papers. To account for this, the ratio citations per paper showed that funded papers do not stand out as more frequently cited with almost an equal total citations per paper ratio for nonfunded records over the period under evaluation.

Journals. The most productive journals, citations, citations per article, funded papers, impact factor and subject categories are summarised in Table 3. *Food Chemistry* appears as the leading journal in the number of publications (673), followed by the *Journal of Agricultural and Food Chemistry* (400). *Analytical Methods* (279) and *Food Analytical Methods* (226) also appeared above 200 publications. Five journals were also identified with more than 100 papers. These were *Food Research International*, *LWT-Food Science and Technology*, *Journal of the Science of Food and Agriculture*, *Food Control*, *European Food Research and Technology,* and *AOAC International*. *LWT-Food Science and Technology* (85.7%) and *Analytical Methods* (83.2%) appeared as the two journals with more funded papers. *Food Chemistry* appeared as the journal with the highest number of citations, followed by the *Journal of Agricultural and Food Chemistry,* with also a substantial number of citations compared with the other journals. However, in terms of citations per article this ranking is inverted with the *Journal of Agricultural and Food Chemistry* as the leading journal with 41.67 citations per article. In terms of subject categories, excluding *Food Science and Technology*, *Chemistry applied* and *Chemistry Analytical* were also highly represented. Moreover, areas such as *Agriculture multidisciplinary*, *Nutrition and Dietetics*, *Spectroscopy*, *Toxicology* and *Engineering chemical* were also observed.

Most cited papers. Table 4 shows those papers with the highest number of citations and the standardised citations considering the publication year of the articles. The most cited paper reported a review on the use of nondestructive measurements using NIR spectroscopy for the assessment of fruit and vegetable quality. The paper was published in 2007 from a collaboration by researchers from two institutions, the Catholic University of Leuven in Belgium and Stellenbosch University in South Africa. The article discusses the most widely used spectral pre-processing methods and the most used linear and nonlinear regression techniques used for such applications. Outlier detection techniques and model accuracy and robustness methods are also reviewed. Interestingly, calibration transfer was also discussed in the article to avoid the repetition of expensive and time-consuming calibrations. Finally, the temperature effect on the spectral properties of fruits and vegetables was also discussed, and several options were provided [23].

The second most cited article was published in the *Analytical Methods* journal and reported a tutorial review on principal component analysis (PCA). The article was published from a collaboration between the Universities of Copenhagen and the University of Amsterdam. A small data set was used to illustrate the features of this widely used chemometrics tool [24]. Pure chemometricians wrote the article with the aim of helping readers to understand, use and interpret PCA. Interestingly, when the citations are standardised with the number of years from publication this article stood out with 134.67 citations per year. The third most cited article was reported in *Trends in Food Science and Technology* and presented hyperspectral imaging as an emerging process analytical tool for food quality and safety. This article was published by a collaboration of researchers from three institutions in Ireland. The authors showed the applications and potential of the technique [25]. However, despite its common adoption, the deployment in the food industry as a process control tool seems to be occurring slowly due to the high cost of the equipment and data processing complexity. Other relevant topics identified in the most frequently cited papers presented in Table 4 include spectroscopy applications to assess the quality of different food and beverages such as meat, olive oil, or dairy. The topics fraud, adulteration, authenticity, or geographical location were also identified. Moreover, spectroscopy applications in reflectance mode seem to be a topic of high interest. Most of the identified articles were review studies.

Keywords. Fig. 1 shows the keyword networks per period. In this case, the search period was divided into three periods corresponding to <2001, 2001–2010, and 2011–2020. All three maps show the keyword chemometrics playing a central role in combination with spectroscopy as the second most frequently used keyword with the strongest co-occurrence observed between them. A table with the most frequently used keywords is also provided (Table 5). It is worth mentioning that in the keywords map, the size of the spheres represents the number of times the word was found in co-occurrence with the other related terms, whereas in the table, the frequency reported includes the number of times the keyword was used, as a single term or in co-occurrence with other keywords. For the initial period (<2001), the term infrared was found to co-occur with spectroscopy intensively. Very strong co-occurrence was also observed between chemometrics and infrared with olive oil, indicating an intense research activity for this food product

Table 1
Evolution of scientific indicators in chemometrics during the 1978–2020 period.

Five-years period	Nº of articles	Nº of citations	Citations/ Articles	Signatures	Signatures/ Articles
<1991	9	60	6.67	22	2.44
1991–1995	44	1222	27.77	127	2.89
1996–2000	97	3970	40.93	347	3.58
2001–2005	253	12,723	50.29	1022	4.04
2006–2010	467	21,508	46.06	2048	4.39
2011–2015	1124	30,439	27.08	5406	4.81
2016–2020	2118	20,645	9.75	11,270	5.32
Totals	4112	90,567	22.03	20,242	4.92

Table 2
Evolution of funded and non-funded papers during the 2011–2020 period.

Years	Funded Papers					Papers without funding				
	Papers	% Papers	Citations	% Citations	Citations/papers	Papers	%Papers	Citations	%Citations	Citations/papers
2011	100	69.9%	3425	70.7%	34.25	43	30.1%	1416	29.3%	32.93
2012	145	72.5%	4314	75.1%	29.75	55	27.5%	1432	24.9%	26.04
2013	143	65.9%	3696	69.7%	25.85	74	34.1%	1605	30.3%	21.69
2014	181	70.7%	5069	65.3%	28.01	75	29.3%	2688	34.7%	35.84
2015	235	76.3%	5449	80.2%	23.19	73	23.7%	1345	19.8%	18.42
2016	260	76.2%	4535	77.8%	17.44	81	23.8%	1295	22.2%	15.99
2017	290	74.0%	4241	75.6%	14.62	102	26.0%	1372	24.4%	13.45
2018	325	75.8%	4328	80.4%	13.32	104	24.2%	1053	19.6%	10.13
2019	367	76.9%	2308	78.1%	6.29	110	23.1%	646	21.9%	5.87
2020	388	81.0%	714	82.4%	1.84	91	19.0%	153	17.6%	1.68
Totals	2434	75.1%	38,079	74.5%	15.64	808	24.9%	13,005	25.5%	16.10

during this initial period. In terms of the chemometric techniques employed, PCA, multivariate data analysis (MVDA), PLS and discriminant analysis (DA), in order of appearance in Table 5, appeared highly represented and well connected with chemometrics. Near-infrared (NIR), fluorescence, NMR and Raman with Fourier transform, also in this order in Table 5, are represented in the map, thus showing strong co-occurrence and demonstrating the presence of a wide variety of spectroscopy techniques. The main tasks performed dealt with characterisation, adulteration, and authentication problems, with only the latter two appearing among the twenty most used keywords. Applications for olive oil, fatty acids, phenolic compounds, oils, almonds, beef, fruits, edible oils, and meat were first investigated during this period, with the first three foods showing the highest frequency. Intense associations between fruits and beef with spectroscopy were also observed. Moreover, the terms quality and control were also represented and well connected, indicating the main purposes when applying these techniques. These terms were not represented in Table 5. HPLC appeared as the analytical methods of choice. Finally, intense associations between pattern recognition and metals and sensory analysis and fuzzy methods indicate intense research areas for this period, especially for the later association (Table 1).

In the second period under study, the keywords near-infrared, infrared and Fourier transform showed strong links with chemometrics and spectroscopy and occupy top positions in Table 5. In terms of the chemometric techniques employed, PCA, PLS, MDVA and DA were represented in both Fig. 1 and Table 5. Fig. 1 also shows the presence of the keywords pattern recognition with lower co-occurrence, and for the first time ANN (artificial neural networks). Authentication and adulteration seem to be the main tasks attempted, with the keyword classification observed in the network. Near-infrared, NMR and fluorescence showed intense activity both as single terms and in co-occurrence. Moreover, Raman, and for the first time MIR, were observed in the keyword network, with ATR not linked to this later term. Finally, UV and Visible were also observed in the map, completing the range of spectroscopies used. The term profiling is also observed in the map together with quality and control. Analytical techniques such as MS and SPME were identified with the frequently used HPLC and GC (gas chromatography). The inclusion of GC is associated with volatile compounds. Chemometric applications were also used for geographic origin and botanical origin identification tasks. Olive oil, wine and honey seemed to be leading the research activity during this period in combination with cheese, fish, sugar, or cider applications.

During the last period under study (2011–2020), NIR, Fourier transform, and food were strongly linked with chemometrics. PCA stood out as observed in the previous periods with PLS, MDVA and DA as per Table 5 and PLS-DA as identified in the keywords network within the chemometrics tools used. The most widely represented spectroscopies were NIR (the only one appearing in Table 5) and NMR, Fluorescence, Raman, MIR, UV and Visible. More recent spectroscopy techniques such as induced coupled plasma optical emission spectroscopy (ICP-OES) was observed for the first time in the co-occurrence map, but most importantly, the term imaging was identified in this last period of analysis. Authentication, adulteration, and classification were again the most widely used terms (Table 5) together with quantification, profiling, and fingerprinting. Mass spectrometry appears now in a higher position in Table 5, with HPLC and GC as the most widely used analytical techniques. The keywords quality (Table 5) and control appeared together with terms such as fraud and safety. Interesting terms such as metabolomics, electronic nose and antioxidants were also observed. Finally, only olive oil and phenolic compounds appeared within the 20 most common terms used. Other foodstuffs such as milk, meat, honey, wine, and fatty acids were also observed on the map.

Authors. In the first period under study (<2001) two authors led the publication activity (Table 6). Both authors were based in the United Kingdom, with some papers with shared co-authorship. Wilson R.H. published papers reporting several spectroscopy applications mainly using infrared and Raman spectroscopies to assess a variety of foodstuffs. This author also reported applications for adulteration, authenticity, process control and fermentations. As mentioned above, similar topics were also observed for Kemsley E.K. as some of the papers were written in co-authorship. In addition, this author reported papers comparing PCA and PLS for dimensionality reduction tasks, chemometrics for classification problems, a genetic algorithm approach to calculate canonical variates or avoiding overfitting for high dimensional data with artificial neural networks. These applications could be considered as fundamental chemometrics work. In third place, Aishima T. focused his research using GC and sensory data to assess mainly soya sauce, katsuobushi and tea. The author makes wide use of pattern recognition techniques, and to a lesser extent, spectroscopy applications to evaluate the flavour profile of foodstuffs.

During the following period under evaluation, it is observed that the most prolific authors have on average doubled the number of publications when compared with the previous period, despite a shorter period (10 years) being assessed. The most prolific author was Karoui R., with 24 publications. This author developed his career between France and Belgium and his research dealt with topics such as infrared and fluorescence applications for geographical discrimination of dairy products such as milk or cheese. Other applications include the evaluation of soil texture or the assessment of the botanical origin of honey. Spectroscopy applications to assess fish products or egg freshness were also observed. The second most prolific author (Dofour E.) also worked at a French institution and mainly dealt with fluorescence applications for dairy, cheese, meat, and fish products. Authentication and geographical origin determination were also a topic of interest. In third place, Downey G., an Irish researcher, showed intense research activity in areas such as infrared spectroscopy in a wide variety of foodstuffs and for adulteration or geographical identification tasks. In addition, several papers using hyperspectral imaging were also observed for this author.

For the final period under evaluation (2011–2020), the number of publications was again overall double than those observed for the most

Table 3
Most productive journals, citations, citations per article, impact factor, and subject category (>35 papers).

Journal	Papers	Papers without funding	% Papers without funding	Funded papers	% Funded papers	5-Year Impact Factor	Citations	Citations per paper	Subject Category	Quartile
Food Chemistry	673	204	30.3%	469	69.7%	6.219	22,541	33.49	Chemistry, Applied Nutrition & Dietetics Food Science & Technology	Q1 Q1 Q1
Journal of Agricultural and Food Chemistry	400	235	58.8%	165	41.3%	4.289	16,666	41.67	Agriculture, Multidisciplinary Chemistry, Applied Food Science & Technology	Q1 Q1 Q1
Analytical Methods	279	47	16.8%	232	83.2%	2.296	3614	12.95	Chemistry, Analytical Food Science & Technology Spectroscopy	Q2 Q2 Q2
Food Analytical Methods	226	55	24.3%	171	75.7%	2.551	2589	11.46	Food Science & Technology	Q2
Food Research International	171	38	22.2%	133	77.8%	5.084	4460	26.08	Food Science & Technology	Q1
LWT-Food Science and Technology	154	22	14.3%	132	85.7%	4.385	2489	16.16	Food Science & Technology	Q1
Journal of the Science of Food and Agriculture	139	65	46.8%	74	53.2%	2.945	2243	16.14	Food Science & Technology Chemistry, Applied Agriculture, Multidisciplinary	Q1 Q2 Q1
Food Control	131	32	24.4%	99	75.6%	4.421	2652	20.24	Food Science & Technology	Q1
European Food Research and Technology	108	58	53.7%	50	46.3%	2.341	1536	14.22	Food Science & Technology	Q2
Journal of AOAC International	103	59	57.3%	44	42.7%	1.399	624	6.06	Chemistry, Analytical Food Science & Technology	Q3 Q3
Foods	96	27	28.1%	69	71.9%		300	3.13		
Journal of Food Science	88	35	39.8%	53	60.2%	2.693	1848	21.00	Food Science & Technology	Q2
International Journal of Food Science and Technology	71	28	39.4%	43	60.6%	2.516	961	13.54	Food Science & Technology	Q3
Journal of the American Oil Chemists Society	69	39	56.5%	30	43.5%	1.869	1634	23.68	Food Science & Technology Chemistry, Applied	Q2 Q2
Journal of Food Composition and Analysis	67	14	20.9%	53	79.1%	3.829	1029	15.36	Food Science & Technology Chemistry, Applied	Q1 Q1
Journal of Food Engineering	62	18	29.0%	44	71.0%	4.332	2042	32.94	Engineering, Chemical Food Science & Technology	Q1 Q1
Meat Science	59	24	40.7%	35	59.3%	4.095	2249	38.12	Food Science & Technology	Q1
Food and Bioprocess Technology	56	15	26.8%	41	73.2%	3.666	2021	36.09	Food Science & Technology	Q1
International Journal of Food Properties	51	23	45.1%	28	54.9%	1.774	414	8.12	Food Science & Technology	Q2
Journal of Food Measurement and Characterisation	47	12	25.5%	35	74.5%	1.649	207	4.40	Food Science & Technology	Q3
Journal of Food Science and Technology-Mysore	44	10	22.7%	34	77.3%	2.705	353	8.02	Food Science & Technology	Q2
European Journal of Lipid Science and Technology	44	15	34.1%	29	65.9%	2.302	763	17.34	Food Science & Technology	Q2
Food Additives and Contaminants Part A-Chemistry Analysis Control, Exposure and Risk Assessment	37	15	40.5%	22	59.5%	2.516	466	12.59	Chemistry, Applied Food Science & Technology Toxicology	Q2 Q2 Q3

prolific authors in the ten preceding years. In this case, the author Sun D.W. appeared first with 53 publications. This author, whose research career was developed in Ireland, focused his activity on topics such as modelling freezing, thawing, cooling, and heating of foodstuffs, machine vision and image processing techniques applied to food processes, ultrasound applications as well as the use of Raman and surface-enhanced Raman spectroscopy (SERS), but also vibrational spectroscopy, of food products and processes. The second most prolific author was the Australian researcher Cozzolino D. with a strong focus on infrared spectroscopy applications for a wide variety of foodstuffs.

Table 4
Most cited papers with more than 250 citations.

Authors	Title	Source	Citations	Standardised Citations
Nicolai, Bart M.; Beullens, Katrien; Bobelyn, Els; Peirs, Ann; Saeys, Wouter; Theron, Karen I; et al.	Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review	Postharvest Biology and Technology 2007; 46 (2): 99-118	1174	90.31
Bro, Rasmus; Smilde, Age K.	Principal component analysis	Analytical Methods 2014; 6 (9): 2812-2831	808	134.67
Gowen, A. A.; O'Donnell, C. P.; Cullen, P. J.; Downey, G.; Frias, J. M.	Hyperspectral imaging - an emerging process analytical tool for food quality and safety control	Trends in Food Science & Technology 2007; 18 (12): 590-598	735	56.54
Cen, Haiyan; He, Yong	Theory and application of near infrared reflectance spectroscopy in determination of food quality	Trends in Food Science & Technology 2007; 18 (2): 72-83	532	40.92
Nychas, George-John E.; Skandamis, Panos N.; Tassou, Chrysoula C.; Koutsoumanis, Konstantinos P.	Meat spoilage during distribution	Meat Science 2008; 78 (1–2): 77-89	469	39.08
Reid, Linda M.; O'Donnell, Colm P.; Downey, Gerard	Recent technological advances for the determination of food authenticity	Trends in Food Science & Technology 2006; 17 (7): 344-353	365	26.07
Moore, Jeffrey C.; Spink, John; Lipp, Markus	Development and Application of a Database of Food Ingredient Fraud and Economically Motivated Adulteration from 1980 to 2010	Journal of Food Science 2012; 77 (4): R118-R126	357	44.63
Luykx, Dion M. A. M.; Van Ruth, Saskia M.	An overview of analytical methods for determining the geographical origin of food products	Food Chemistry 2008; 107 (2): 897-911	348	29.00
Huang, Haibo; Yu, Haiyan; Xu, Huirong; Ying, Yibin	Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review	Journal of Food Engineering 2008; 87 (3): 303-313	321	26.75
Llorach, Rafael; Martinez-Sanchez, Ascension; Tomas-Barberan, Francisco A.; Gil, Maria I.; Ferreres, Federico	Characterisation of polyphenols and antioxidant properties of five lettuce varieties and escarole	Food Chemistry 2008; 108 (3): 1028-1038	296	24.67
Yang, H; Irudayaraj, J; Paradkar, MM	Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy	Food Chemistry 2005; 93 (1): 25-32	289	19.27
Aparicio, R; Roda, L; Albi, MA; Gutierrez, F	Effect of various compounds on virgin olive oil stability measured by Rancimat	Journal of Agricultural and Food Chemistry 1999; 47 (10): 4150-4155	285	13.57
Karoui, Romdhane; De Baerdemaeker, Josse	A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products	Food Chemistry 2007; 102 (3): 621-640	262	20.15

Intense work was also observed for the application of spectroscopy and chemometrics for food quality and safety. A high presence of review articles on relevant topics was also identified for this author. The third most prolific author corresponded to van Ruth S.M., based in the United Kingdom, and with intense research in areas such as food fraud and traceability of various foodstuffs with or without the use of spectroscopy applications, especially vibrational spectroscopy, and imaging techniques.

Finally, the appearance of some authors in several of the periods evaluated also deserves consideration. This is the case for Mangas-Alonso J.J., Karoui R. or Cozzolino D., appearing in two periods. Moreover, two authors (Downey G., and Engelsen, SB) were represented in the three evaluated periods, showing high publication intensity. The Spanish researcher Mangas-Alonso J.J. showed a strong focus on applying chemometrics tools during the production process of cider. On the other hand, the research carried out in Denmark by the author Engelsen, S.B. focused on several applications using chromatography, near infrared and NMR metabolomics, and several papers dealing with more fundamental chemometrics and process analytical technologies work.

Countries. Fig. 2 shows the evolution of the most prolific countries per period. The bigger the countries' font size, the more intense publication activity under the period evaluated. Interestingly, Spain appears the most prolific country during the first period under evaluation (<2001) with 28 published papers. United Kingdom (18), Italy (15) and Denmark (12) also showed prolific scientific activity. Countries also represented were the United States (7), France (7) or Greece (5) and Sweden (5). During the second period under study, Spain appeared again as the most prolific country (107 publications) but followed closely by the United States (97) and with less scientific productivity by Italy (69), France (41), China (35), Denmark (32) and United Kingdom (31). Finally, during the third period under study, the ranking was led by China (767) with almost double the publications than Brazil (367). Spain appears in this case in third place with 349 published papers, followed closely by Italy with 323 publications and the United States (293).

The international collaboration among the most prolific countries is shown in Fig. 3. Four publications in collaboration were the threshold applied. The strongest collaboration was found between the United States of America and China with 62 papers in partnership. The second most prolific collaboration was between China and Ireland (35 papers). In third place, China appears again, but in this case, collaborating intensively with Australia with 24 joint papers. From these observations, China is the most prolific country in terms of international collaboration. Interestingly solid collaboration activity was found between Spain and Italy (24), two countries with similar geographic location, climate and therefore production of foods. Intense collaboration was also observed for Ireland, mainly with China and the United States, with this later being also a very active country internationally. Intense collaboration between neighbouring countries was also observed such as for Malaysia and Indonesia, France and Belgium or Spain and Portugal. Strong collaboration activity was also identified for New Zealand, mostly with United States of America and China.

4. Discussion

Bibliometric studies analyse metadata from databases and provide further understanding of the status of a research field. By analyzing the records obtained from the bibliometric search, the relevant information is provided to newcomers and established researchers in the field under study. Moreover, by analysing the existing body of literature, a new perspective of the research status can be provided, with the potential identification of research trends, and existing knowledge gaps. Collaboration patterns among individuals, institutions and countries are also often provided. Bibliometric studies have been reported in a variety of chemometrics related topics, such as an earlier study on the historical,

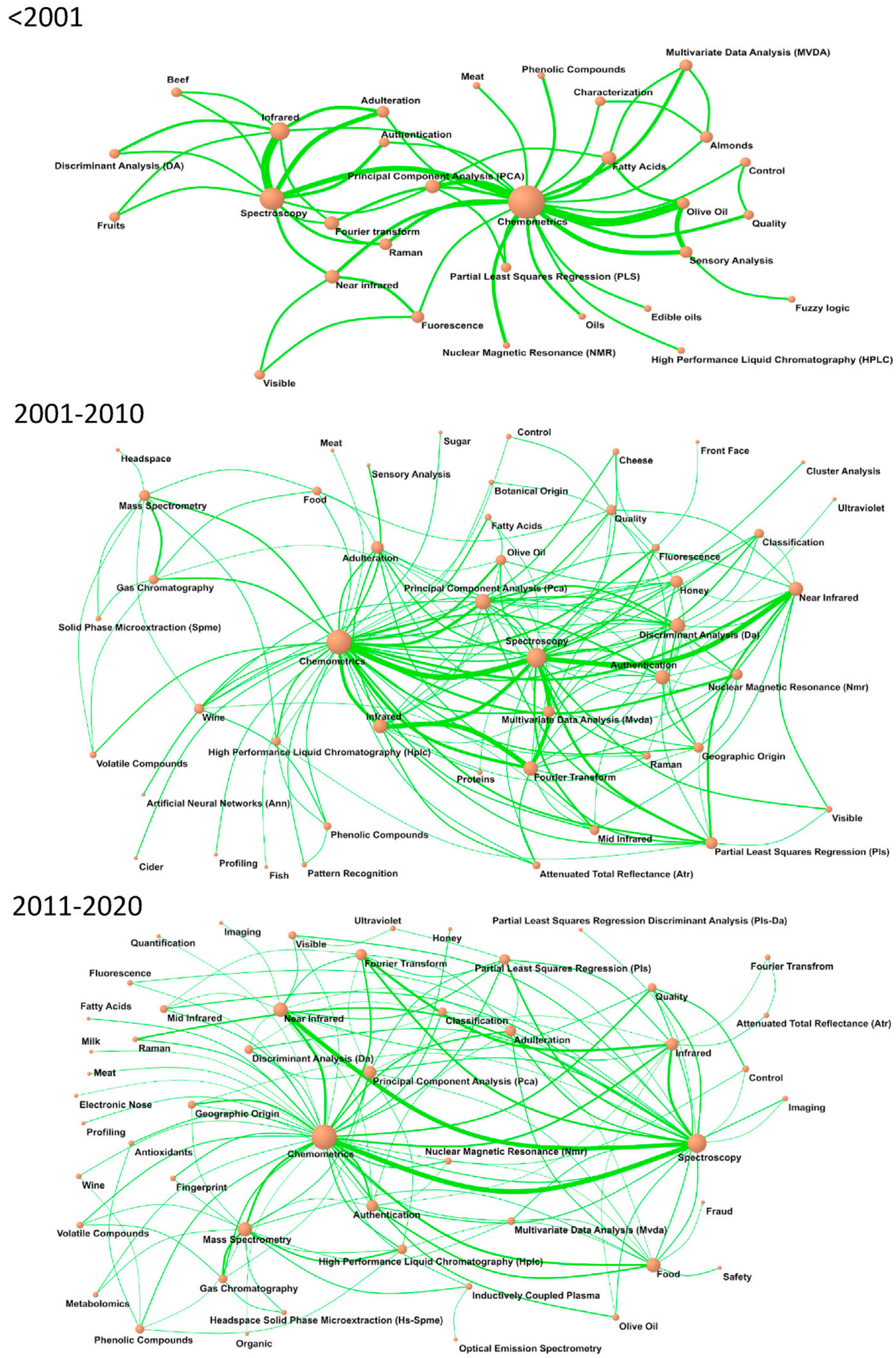


Fig. 1. Fig. 1. Keyword networks of co-occurrence. The figure shows the keyword networks of three periods (i.e. articles published prior to 2001, during 2001–2010 and finally during the 2011–2020 period) to assess the keyword usage evolution over time.

linguistic, and sociological interpretation of the term chemometrics [19], collaboration patterns and networks in chemometrics [2] or a study on the strongly related field of spectroscopy [26].

Over time, the number of papers published showed a proportional increase with double the papers every five years. These results could be associated with the overall increase in published papers worldwide [27]. However, it might also be a sign of the good health of the discipline. Increasing amounts of data are nowadays generated. The food industry is also immersed in the fourth industrial by incorporating the newest technology and implementing process monitoring systems and process analytical technologies (PAT). This scenario requires data handling tools to process complex and large datasets in a multivariate manner. The nature of this data, very often coming from spectroscopic measurements, collected at-line, in-line or on-line [28,29], makes it suitable for chemometrics analysis. Significant variation in the raw materials or data collected from food manufacturing processes seems ideal for chemometric applications. The variable chemical composition of foodstuffs is also opening opportunities for research and development of chemometrics applications. This is clearly represented in recent publications that investigate applications in a wide variety of products and for varying purposes, for example, banana flour adulteration [30] or contamination with arsenic in teas [31].

In addition, the number of citations goes in line with the increase in the number of papers. An average citations/paper of around 45 is

Table 5
Most frequently use keywords for the three periods evaluated.

<2001		2001–2010		2011–2020	
Keywords	Frequency	Keywords	Frequency	Keywords	Frequency
Chemometrics	54	Chemometrics	332	Chemometrics	1232
Spectroscopy	23	Spectroscopy	193	Spectroscopy	741
Olive oil	15	PCA	98	Near infrared	391
Infrared	15	Near infrared	93	PCA	348
PCA	15	Infrared	90	Mass spectrometry	321
MVDA	12	Fourier transform	87	Food	314
Near infrared	10	PLS	78	Infrared	276
Fatty acids	9	MVDA	68	PLS	258
Sensory analysis	9	DA	66	Authentication	254
PLS	8	NMR	58	MVDA	238
DA	7	Olive oil	54	Fourier transform	237
Fluorescence	7	Authentication	51	Quality	219
Adulteration	7	Mass spectrometry	47	Adulteration	216
Quality	7	Geographic origin	43	HPLC	195
NMR	7	HPLC	42	Geographic origin	193
Pattern recognition	6	Wine	42	Gas chromatography	175
Raman	6	Adulteration	40	Olive oil	172
Fourier transform	6	Honey	39	DA	172
Authentication	6	Fluorescence	39	Classification	159
Phenolic compounds	5	Gas chromatography	38	Phenolic compounds	158

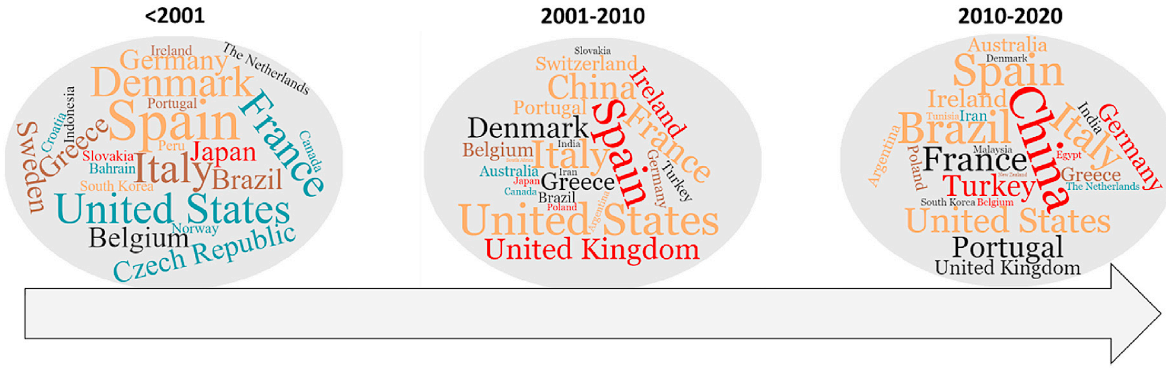
Table 6
Most prolific authors during the period evaluated.

<2001		2001–2010		2011–2020	
Authors	Publications	Authors	Publications	Authors	Publications
Wilson, RH	11	Karoui, R	24	Sun, DW	53
Kemsley, EK	11	Dufour, E	22	Cozzolino, D	40
Aishima, T	9	Downey, G	22	van Ruth, SM	31
Phillips, GO	8	Irudayaraj, J	21	Karabagias, IK	31
Jurasek, P	6	O'Donnell, CP	17	Rohman, A	29
Mangas-Alonso, JJ	6	De Baerdemaeker, J	15	Chen, Q	29
Engelsen, SB	6	Ranalli, A	14	Granato, D	28
Morales, MT	6	Bosset, JO	14	Kontominas, MG	27
Moreno, J	5	Contento, S	13	Karoui, R	27
Defernez, M	5	Cozzolino, D	11	Boyaci, IH	23
Munck, L	5	Engelsen, SB	10	Badeka, AV	23
Downey, G	5	Aparicio-Ruiz, R	10	Kontakos, S	20
Latorre, MJ	4	He, Yong	10	Pu, H	20
Herrero-Latorre, C	4	Garcia-Gonzalez, DL	9	Engelsen, SB	19
Picinelli, A	4	Paradkar, MM	9	Elliott, CT	18
Suarez, B	4	Mangas-Alonso, JJ	9	Farag, MA	18
Berenguer-Navarro, V	4	Arvanitoyannis, IS	9	Godoy Alves Filho, E	17
Blanco-Gomis, D	3	Norgaard, L	9	Valderrama, P	17
Gutierrez, F	3	Lucera, L	9	Sousa de Brito, E	16
Garcia, s	3	Man, YBC	8	Downey, G	16

observed for the period 1996–2010 which can indicate the average number of citations per paper in this field. As mentioned previously, the number of citations of the last decade should not be considered as papers are still getting noticed and cited by the scientific community. Lower ratios were observed in other bibliometric studies, with around 25 citations per paper for studies on renewable energy [32], probiotics [33] or nanotechnology [34], 30 for water use efficiency [35] or a lower ratio of around 20 for a bibliometric study on spectroscopy applications in food science and technology [26]. The citations/paper indicator might be also showing the healthy status of this prolific field of study. The increase in the number of signatures/papers observed overtime is also worth mentioning. This trend can also be a sign of, on the one hand, the participation of expert chemometricians within the food scientists' teams or, on the other hand, the involvement of the food scientists themselves in chemometric tasks. It is the perception of the authors that pure chemometrics papers are generally signed by a lower number of authors compared with food science and technology related papers. The data showed by Li et al. (2019) might confirm the above statement as the number of authors per paper reported in this study, analysing only six journals strongly devoted to chemometrics work, was lower than the one reported here.

The funding activity was also evaluated, with interesting trends observed. Firstly, the fact that most of the chemometrics papers are funded, shows the acceptance and wide presence of this field in food science and technology research. Secondly, the increasing trend in the number of funded papers (by 10% in the last decade) is also showing the vitality of chemometrics applications. However, based on these, it is not possible to ascertain if chemometrics as a field is receiving increased funding. The results showed that works including chemometrics applications are more likely to be funded. On the other hand, the assumption that funded works are more frequently cited, as funded papers should be part of the most competitive and high-quality research projects, does not hold, as was identified here, with an equal number of citations received for both funded, and non-funded papers. It is possible that high quality non-funded work is also produced.

Two journals, i.e. Food Chemistry and Journal of Agricultural and Food Chemistry, appeared as leading journals in food science and technology for chemometrics applications. These journals are both leading



	<2001		2001-2010		2011-2020
papers	country	papers	country	papers	country
28	Spain	107	Spain	767	China
18	United Kingdom	98	United States	367	Brazil
15	Italy	68	Italy	349	Spain
12	Denmark	41	France	323	Italy
7	France	35	China	293	United States
7	United States	32	Denmark	130	France

Fig. 2. Intensity of publication per country over time. The top part of the figure shows the evolution of the most prolific countries using cloud maps with a higher font size indicating a higher number of publications. The table includes the number of publications of the top six prolific countries per period.

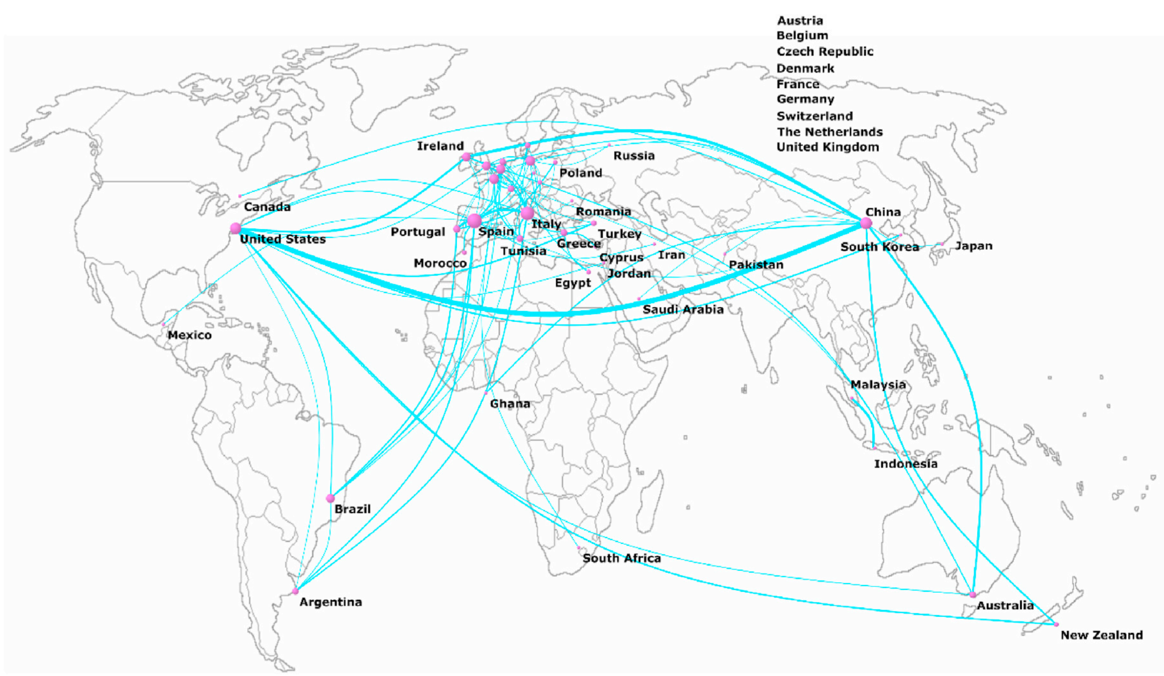


Fig. 3. International collaboration network (threshold >4 in collaboration). Forty countries represented in the map.

journals in terms of the number of publications and impact factor (the former) and citations per paper (the latter). None of the journals strongly devoted to chemometrics work, as reported by Li et al. (2019), was represented among the journals with more than 35 papers. This can be a sign of the widespread use of chemometrics applications within the food science and technology community, which confirms the reasoning exposed by previous authors who predicted a rapid growth in the number of non-pure chemometricians using chemometrics tools [1,11]. The analysis of the subject categories yielded interesting results as chemistry-related subject categories were also highly represented, a fact that refers back to the use of chemical data in chemometrics work [4,5].

The most cited papers with the highest number of citations can be used to indicate the research activity and identify research trends. Several of the studies corresponded to review articles dealing with the current food science and technology applications. The use of imaging techniques, such as hyperspectral imaging, for the evaluation of foodstuffs along the production chain, appeared highly represented and might be nominated as one of the current and future research trends. The data obtained from imaging instruments requires strong chemometrics tools for processing and analysis [36]. The possibility of using images opens numerous opportunities for the evaluation of food products from farming and crop production to processing or even storage and post-harvest of fruits and vegetables to ensure food quality and safety [25,37,38]. Spectroscopy applications are widely represented, showing the strong link between spectroscopic and chemometrics applications for both qualitative and quantitative evaluation of foodstuffs [26,39]. Food safety and quality are also a priority with authenticity and adulteration of foodstuffs as main topics of interest [40]. Finally, at-line, in-line or online applications are also expected to be progressively implemented in the food industry as part of a process monitoring and control strategy [28,29,41]. The increasing complexity of both technology and applications requires chemometrics tools to process and analyse the generated chemical data. On the other hand, a paper describing the fundamentals and use of PCA was also represented as one of the most cited papers (second most cited) and indicated the wide use of this unsupervised technique as an exploratory tool for complex chemical datasets [24].

The above mentioned and the exploration of the most recently published articles in 2021 indicates that PCA, PLS in its regression (PLSR) and discrimination versions (PLS-DA) are still the preferred chemometrics tools employed. The orthogonal variation of PLS (OPLS) and more advanced algorithms such as SVM (for both regression and classification), random forests (RF) or ANN were also widely employed. Other commonly observed techniques included soft independent modelling of class analogy (SIMCA) or parallel factor analysis (PARAFAC) for three-dimensional data. Finally, the wide presence of variable importance in projection (VIP) as a variable selection tool is worth mentioning. Applications for geographical or product discrimination (authentication) and adulteration seem to be still the focus in the most recent body of literature [42–46]. Together with this, interesting applications to monitor food production processes such as frying oil processes or roasting were observed [43,47]. Discrimination between different stages of the production process [48], thermal degradation [49] or food profiling at elemental and nutritional levels [45,50] was also observed. A substantial number of studies involving spectroscopy work were identified. The high presence of imaging applications using both hyperspectral and digital data [51,52], the use of less common Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) [50] or surface-enhanced Raman spectroscopy (SERS) [53], and finally, applications making use of NMR data [54,55] indicate the high scientific interest of these applications which are necessarily accompanied of chemometrics work.

The keyword analysis in bibliometric studies provides an enhanced visualization of the most frequent keywords, and an overview of the co-occurrence between keywords. The analysis can be used to identify the research activity, and the progression of the research performed over time. In this case, a progression was observed with, for example, the chemometrics tools used, the spectroscopy techniques applied or even the analytical techniques employed to generate the chemical data. Regarding the chemometrics applications employed, it appears that PCA and PLS were widely used in the initial period and are still the option of choice for most chemometrics applications, in accordance with the most recent publications. Discriminant analysis appeared in the second period to became also a widely used chemometric tool with multivariate data analysis (MVDA). The appearance of this latter term shows the multivariate nature of most chemometrics applications. Despite the absence of a reference to this later term in the original definition of chemometrics, it is widely accepted that most of chemometrics applications provide a multivariate approach to the data [1,5,7]. Interestingly, alternative techniques such as those reported in the above paragraph for regression and discrimination were represented neither in the keyword's networks nor in the most frequently used keywords (with the only exception of ANN in the 2001–2010 period). This, although unexpected, indicates the power and wide acceptance of PCA, PLS and DA as chemometrics tools. As previously discussed, a wide number of alternative chemometrics tools were observed in the most recent publications indicating that, despite their absence in the keyword analysis, they are preferred and widely used by the chemometrics community. The combined use of chemometrics and machine learning tools to handle larger and more complex data sets seems to be common practice and is expected to become more relevant.

Applications for varying foodstuffs were observed, which shows the wide spectrum of suitable applications for chemometrics work and the importance of chemometrics as a discipline within food science and technology, where the chemical evaluation of food components plays a central role. Olive oil appeared as the foodstuff with the highest presence, with applications for fatty acids, phenolic compounds or research focusing on wine and honey. The main applications dealt with assessing of geographical identification or authentication and adulteration of foodstuffs with non-permitted adulterants to ensure food quality and safety. The appearance of more recent spectroscopy applications was also observed, such inductively coupled plasma – optical emission spectrometry (ICP-OES) for elemental content analysis. The higher presence in the most recent period of mass spectrometry to complement liquid and gas chromatography analysis or even the use of metabolomics approaches is also worth mentioning.

The evaluation of the most prolific authors showed the presence of primarily food scientists and chemists making use of chemometrics approaches. However, pure chemometrics work was also observed in some cases within the group of most prolific authors. This is again points to the observed trend of the widespread use of chemometrics tools by researchers not devoted to pure chemometrics work [1,11]. Moreover, the most relevant topics identified were represented among the research works signed by these authors. Most of the authors published papers using spectroscopy measurements, and to a lesser extent, only chemical data obtained from an ample variety of foodstuffs.

Interesting patterns were observed for the most prolific countries and for their collaboration. Spain appeared as one of the most prolific countries during the evaluation period. This could be due to the relevance of the agri-food industry for the country's economy [56]. Interestingly, China and Brazil became the two most prolific countries publishing chemometrics work for food applications in the last and more recent period evaluated. These emerging countries, belonging to the BRICS group and coined as the leading economies in the future, have also sustained their economies in a growing and prolific agri-food industry [57]. In addition, the movement of chemometrics work from the original north European countries, where the field was initially developed, to southern European and other countries was not clearly observed. Countries where the field was initially developed appeared within the most prolific but not in the first positions. Interestingly, if the number of publications was standardised per million inhabitants, countries such as Sweden, Denmark, or even the Netherlands and Belgium appeared in higher positions (with a higher number of papers per million inhabitants), which could be an indication of the vital role played by these countries when developing chemometrics as a field. In addition, the appearance of Ireland as one of the most prolific countries in both number of publications and collaboration activity also deserves attention. Regarding collaboration, USA and China showed a solid and prolific collaboration activity, with China appearing as the overall leading country in terms of publications in collaboration. Intense collaboration was also observed among neighbouring countries such as Spain and Italy or Portugal, France and Belgium or Malaysia and Indonesia. This can be assigned to the production of similar foodstuffs and, therefore, the need for similar chemometrics applications.

5. Conclusions

A bibliometric evaluation of the use of chemometrics in the food science and technology research area was presented in this study. The results showed chemometrics as a prolific and growing field with multiple applications identified for a broad spectrum of foodstuffs. Increased funding activity in papers making use of chemometrics approaches was also observed. The presence of a rising number of non-pure chemometricians using chemometrics tools could also be depicted. PCA, PLS and DA are still the chemometrics techniques of choice for most applications. More recent chemometrics tools were also observed in the most recent literature and are expected to gain traction in the coming years. Spectroscopy work appears as the field making more intense use of chemometrics. However, chemometric applications were also observed for chemical data obtained from non-spectroscopy instrumentation. Finally, a very intense publication intensity was observed for China and Brazil, placing them as the current leading countries in the field.

Limitations of the study

Only the articles that included the term chemometric or chemometrics were analyzed in this study. This excluded those articles that did not include the abovementioned terms in the title, abstract or keywords, despite doing chemometrics work. However, bibliometric studies aim to provide a representative sample that can be solidly used to extract conclusions from the records obtained. The records obtained were those where the authors decided that chemometrics work was relevant enough to include the terms in the selected fields. Moreover, including some of the identified methods within chemometrics tools is also debatable. Some of these methods were developed or proposed by statisticians or machine learning scientists and could be assigned to these fields, whereas other methods are more clearly assigned to chemometrics. It is not the intention of the authors to classify these methods in either field and instead indicate that the wide use of these methods in work devoted to chemometrics makes them relevant for the field.

Author statement

Jose Luis Aleixandre-Tudo: Conceptualization, Funding acquisition, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing; Lourdes Castello-Cogollos: Data curation, Formal analysis, methodology, Visualization, Writing – review and editing; Jose Luis Aleixandre: Conceptualization, Supervision, Writing - review & editing; Rafael Aleixandre-Benavent: Conceptualization, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors want to gratefully acknowledge the Spanish Ministry of Science, Innovation and Universities for the financial support received under the Beatriz Galindo program (BG20/00021).

References

[1] R.G. Brereton, The evolution of chemometrics, Anal. Methods 5 (2013) 3785–3789, https://doi.org/10.1039/c3ay90051g.
[2] C.Q. Li, N. Xiao, Y. Wen, S.H. He, Y. Da Xu, Y.W. Lin, H.D. Li, Q.S. Xu, Collaboration patterns and network in chemometrics, Chemometr. Intell. Lab. Syst. 191 (2019) 21–29, https://doi.org/10.1016/j.chemolab.2019.05.011.
[3] S.D. Brown, B. Kowalski, The Chemometrics Revolution Re-examined, 2016, pp. 1–23, https://doi.org/10.1002/cem.2856.
[4] L.K.D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.
[5] S. Wold, Chemometrics; what do we mean with it, and what do we want from it? Chemometr. Intell. Lab. Syst. 30 (1995) 109–115, https://doi.org/10.1016/0169-7439(95)00042-9.
[6] O.Y. Rodionova, A.L. Pomerantsev, Chemometrics: achievements and prospects, Russ. Chem. Rev. 75 (2006) 271–287, https://doi.org/10.1070/rc2006v075n04abeh003599.
[7] R. Bro, Multivariate calibration: what is in chemometrics for the analytical chemist? Anal. Chim. Acta 500 (2003) 185–194, https://doi.org/10.1016/S0003-2670(03)00681-0.
[8] P.C. Jurs, B.R. Kowalski, T.L. Isenhour, C.N. Reilley, Computerized learning machines applied to chemical problems molecular structure parameters from low resolution mass spectrometry, Anal. Chem. 42 (1970) 1387–1394, https://doi.org/10.1021/ac60294a015.
[9] P.H. Weiner, E.R. Malinowski, An investigation of the Van der Waals effect in nuclear magnetic resonance spectroscopy by factor analysis and the prediction of diamagnetic and paramagnetic susceptibilities, J. Phys. Chem. 75 (1971) 3160–3163, https://doi.org/10.1021/j100689a022.
[10] D.L. Massart, The use of information theory for evaluating the quality of thin-layer chromatographic separations, J. Chromatogr., A 79 (1973) 157–163.
[11] P. Geladi, P.K. Hopke, Editorial: is there a future for chemometrics? Are we still needed? J. Chemom. 22 (2008) 289–290, https://doi.org/10.1002/cem.1141.
[12] P.G. S Wold, K. Esbensen, Principal component analysis, Chemometr. Intell. Lab. Syst. 2 (1987) 37–52.
[13] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[14] S. Wold, Pattern recognition by means of disjoint principal components models, Pattern Recogn. 8 (1976) 127–139.
[15] G.R. Flåten, B. Grung, O.M. Kvalheim, A method for validation of reference sets in SIMCA modelling, Chemometr. Intell. Lab. Syst. 72 (2004) 101–109, https://doi.org/10.1016/j.chemolab.2004.03.003.
[16] R. Bro, PARAFAC. Tutorial and applications, Chemometr. Intell. Lab. Syst. 38 (1997) 149–171, https://doi.org/10.1016/S0169-7439(97)00032-4.
[17] A.D. Olinsky, J.T. Quinn, P.M. Mangiameli, S.K. Chen, A genetic algorithm approach to nonlinear least squares estimation, Int. J. Math. Educ. Sci. Technol. 35 (2004) 207–217, https://doi.org/10.1080/00207390310001638331.
[18] C.K. Chui, Wavelets: a Mathematical Tool for Signal Analysis, Society for Industrial and Applied Mathematics., 1992.
[19] M.C. Ferreira, R. Kiralj, The past, present, and future of chemometrics worldwide: some etymological, linguistic, and bibliometric investigations, J. Chemom. (2007) 247–272, https://doi.org/10.1002/cem.
[20] V. Batagelj, A. Mrvar,Pajek: Analysis and Visualization of Large Networks, Springer, Berlin, Heidelberg.
[21] T. Kamada, S. Kawai, An algorithm for drawing general undirected graphs, Inf. Process. Lett. 31 (1) (1989) 7–15.
[22] S. Cheong, Y. Si, Accelerating the kamada-kawai algorithm for boundary detection in a mobile ad hoc network, ACM Trans. Sens. Netw. 13 (1) (2016) 1–29, 13.
[23] B.M. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K.I. Theron, J. Lammertyn, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review, Postharvest Biol. Technol. 46 (2007) 99–118, https://doi.org/10.1016/j.postharvbio.2007.06.024.
[24] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (9) (2014) 2812–2831, https://doi.org/10.1039/c3ay41907j.
[25] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, G. Downey, J.M. Frias, Hyperspectral imaging - an emerging process analytical tool for food quality and safety control, Trends Food Sci. Technol. 18 (2007) 590–598, https://doi.org/10.1016/j.tifs.2007.06.001.
[26] J.L. Aleixandre-Tudó, L. Castelló-Cogollos, J.L. Aleixandre, R. Aleixandre-Benavent, Bibliometric insights into the spectroscopy research field : a food science and technology case study, Appl. Spectrosc. Rev. (2019) 1–34, https://doi.org/10.1080/05704928.2019.1694936, 0.
[27] M. Fire, C. Guestrin, Over-optimization of academic publishing metrics: observing Goodhart's Law in action, GigaScience 8 (2019) 1–20, https://doi.org/10.1093/gigascience/giz053.
[28] J.U. Porep, D.R. Kammerer, R. Carle, On-line application of near infrared (NIR) spectroscopy in food production, Trends Food Sci. Technol. 46 (2015) 211–230, https://doi.org/10.1016/j.tifs.2015.10.002.
[29] N.D. Lourenço, J.A. Lopes, C.F. Almeida, M.C. Sarraguça, H.M. Pinheiro, Bioreactor monitoring with spectroscopy and chemometrics: a review, Anal. Bioanal. Chem. 404 (2012) 1211–1237, https://doi.org/10.1007/s00216-012-6073-9.
[30] P.F. Ndlovu, L.S. Magwaza, S.Z. Tesfay, R.R. Mphahlele, Vis-NIR spectroscopic and chemometric models for detecting contamination of premium green banana flour with wheat by quantifying resistant starch content, J. Food Compos. Anal. 102 (2021) 104035, https://doi.org/10.1016/j.jfca.2021.104035.
[31] A.O. Barimah, Z. Guo, A.A. Agyekum, C. Guo, P. Chen, H.R. El-Seedi, X. Zou, Q. Chen, Sensitive label-free Cu2O/Ag fused chemometrics SERS sensor for rapid detection of total arsenic in tea, Food Control 130 (2021) 108341, https://doi.org/10.1016/j.foodcont.2021.108341.
[32] J.L. Aleixandre-Tudó, L. Castelló-Cogollos, J.L. Aleixandre, R. Aleixandre-Benavent, Renewable energies: worldwide trends in research, funding and international collaboration, Renew. Energy 139 (2019) 268–278, https://doi.org/10.1016/j.renene.2019.02.079.
[33] J.L. Aleixandre-Tudó, L. Castelló-Cogollos, J.L. Aleixandre, R. Aleixandre-Benavent, Tendencies and challenges in worldwide scientific research on probiotics, probiotics antimicrob, Proteins 12 (2020) 785–797, https://doi.org/10.1007/s12602-019-09591-0.
[34] J.L. Aleixandre-Tudó, M. Bolaños-Pizarro, J.L. Aleixandre, R. Aleixandre-Benavent, Worldwide scientific research on nanotechnology: a bibliometric analysis of tendencies, funding, and challenges, J. Agric. Food Chem. 68 (2020) 9158–9170, https://doi.org/10.1021/acs.jafc.0c02141.
[35] J.L. Aleixandre-Tudó, L. Castelló-Cogollos, J.L. Aleixandre, R. Aleixandre-Benavent, Emerging topics in scientific research on global water-use efficiency, J. Agric. Sci. 157 (2019) 480–492, https://doi.org/10.1017/S0021859619000789.

[36] J.M. Amigo, I. Martí, A. Gowen, Hyperspectral imaging and chemometrics. A perfect combination for the analysis of food structure, composition and quality, Data Handling Sci. Technol. 28 (2013) 343–370, https://doi.org/10.1016/B978-0-444-59528-7.00009-0.

[37] Y. Liu, H. Pu, D.W. Sun, Hyperspectral imaging technique for evaluating food quality and safety during various processes: a review of recent applications, Trends Food Sci. Technol. 69 (2017) 25–35, https://doi.org/10.1016/j.tifs.2017.08.013.

[38] Y.Z. Feng, D.W. Sun, Application of hyperspectral imaging in food safety inspection and control: a review, Crit. Rev. Food Sci. Nutr. 52 (2012) 1039–1058, https://doi.org/10.1080/10408398.2011.651542.

[39] H. Cen, Y. He, Theory and application of near infrared reflectance spectroscopy in determination of food quality, Trends Food Sci. Technol. 18 (2007) 72–83, https://doi.org/10.1016/j.tifs.2006.09.003.

[40] J.H. Qu, D. Liu, J.H. Cheng, D.W. Sun, J. Ma, H. Pu, X.A. Zeng, Applications of near-infrared spectroscopy in food safety evaluation and control: a review of recent research advances, Crit. Rev. Food Sci. Nutr. 55 (2015) 1939–1954, https://doi.org/10.1080/10408398.2013.871693.

[41] B. Boldrini, W. Kessler, K. Rebner, R.W. Kessler, Hyperspectral imaging: a review of best practice, performance and pitfalls for in-line and on-line applications, J. Near Infrared Spectrosc. 20 (2012) 483–508, https://doi.org/10.1255/jnirs.1003.

[42] A. Amirvaresi, N. Nikounezhad, M. Amirahmadi, B. Daraei, H. Parastar, Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection, Food Chem. 344 (2021), https://doi.org/10.1016/j.foodchem.2020.128647.

[43] Y. Chang, L.Y. Chan, F. Kong, G. Zhang, H. Peng, An innovative approach for real-time authentication of cocoa butter using a combination of rapid evaporative ionization mass spectrometry and chemometrics, Food Control 133 (2022), https://doi.org/10.1016/j.foodcont.2021.108617.

[44] T. Shi, G. Wu, Q. Jin, X. Wang, Camellia oil adulteration detection using fatty acid ratios and tocopherol compositions with chemometrics, Food Control 133 (2022), https://doi.org/10.1016/j.foodcont.2021.108565.

[45] X.H. Zhang, X.D. Qing, S.T. Mu, D. Wang, J.J. Zheng, S.J. Zhou, C. Kang, Z. Liu, Authentication of honey of different nectar sources and antioxidant property evaluation by phenolic composition analysis with chemometrics, Food Control 124 (2021), https://doi.org/10.1016/j.foodcont.2021.107900.

[46] A. Rohman, A.R. Putri, Irnawati, A. Windarsih, K. Nisa, L.A. Lestari, The employment of analytical techniques and chemometrics for authentication of fish oils: a review, Food Control 124 (2021), https://doi.org/10.1016/j.foodcont.2021.107864.

[47] Z. Li, Y. Sun, F. Dong, Q. Yu, L. Fan, J. Zhao, M. Wang, Discrimination of heating and frying vegetable oils based on UPLC/Q-TOF MSMS and chemometrics, Food Res. Int. 140 (2021), https://doi.org/10.1016/j.foodres.2020.109874.

[48] F. He, J. Duan, J. Zhao, H. Li, J. Sun, M. Huang, Different distillation stages Baijiu classification by spectrometry and gas chromatography-olfactometry-mass spectrometry combined with chemometric strategies, Food Chem. 365 (2021) 130430, https://doi.org/10.1016/j.foodchem.2021.130430.

[49] A. Michalska-Ciechanowska, J. Brzezowska, A. Wojdyło, A. Gajewicz-Skretna, E. Ciska, J. Majerska, Chemometric contribution for deeper understanding of thermally-induced changes of polyphenolics and the formation of hydroxymethyl-L-furfural in chokeberry powders, Food Chem. 342 (2021), https://doi.org/10.1016/j.foodchem.2020.128335.

[50] O.Y. Song, M.A. Islam, J.H. Son, J.Y. Jeong, H.E. Kim, L.S. Yeon, N. Khan, N. Jamila, K.S. Kim, Elemental composition of pork meat from conventional and animal welfare farms by inductively coupled plasma-optical emission spectrometry (ICP-OES) and ICP-mass spectrometry (ICP-MS) and their authentication via multivariate chemometric analysis, Meat Sci. 172 (2021), https://doi.org/10.1016/j.meatsci.2020.108344.

[51] E.M. Achata, C. Esquerre, K.S. Ojha, B.K. Tiwari, C.P. O'Donnell, Development of NIR-HSI and chemometrics process analytical technology for drying of beef jerky, Innovat. Food Sci. Emerg. Technol. 69 (2021), https://doi.org/10.1016/j.ifset.2021.102611.

[52] R.S. de A. Azevedo, J.H. Santos Neto, L.O. dos Santos, A. de A. Gomes, D. de A. Santana, A.M.P. dos Santos, Homogeneity and stability assessment of a candidate to pumpkin seed flour reference material by means of computer vision based chemometrics assisted approach, Food Chem. 368 (2022), https://doi.org/10.1016/j.foodchem.2021.130842.

[53] A. Zhu, Y. Xu, S. Ali, Q. Ouyang, Q. Chen, Au@Ag nanoflowers based SERS coupled chemometric algorithms for determination of organochlorine pesticides in milk, LWT (Lebensm.-Wiss. & Technol.) 150 (2021) 111978, https://doi.org/10.1016/j.lwt.2021.111978.

[54] M. Vasas, F. Tang, E. Hatzakis, Application of nmr and chemometrics for the profiling and classification of ale and lager american craft beer, Foods 10 (2021) 1–10, https://doi.org/10.3390/foods10040807.

[55] A. Mascellani, G. Hoca, M. Babisz, P. Krska, P. Kloucek, J. Havlik, 1H NMR chemometric models for classification of Czech wine type and variety, Food Chem. 339 (2021) 1–9, https://doi.org/10.1016/j.foodchem.2020.127852.

[56] P.M.B. Albino, F.J.A. Ulibarrena, K.S. Elorz, The impact of the sector of activity and of the region of operation on the competitiveness of the Spanish Agri-Food Industry: a Shift-Share analysis/El impacto del sector de actividad y de la región de actuación sobre la competitividad de la industria a, Brazilian J. Bus. 3 (2021) 1979–2001, https://doi.org/10.34140/bjbv3n2-045.

[57] M.Q. de Monteiro Jales, M.S. Jank, S. Yao, C.A. Carter, Agriculture in Brazil and China: Challenges and Opportunities, 2006. Occasional Paper.