Bibliometric Study of the Application of Computers in Synthetic Organic, Physical, Inorganic, and Analytical Chemistry Literature Abstracted by *Chemical Abstracts* in 1986

MING-YUEH TSAY

Library, Energy and Mining Research/Service Organization, ITRI, Building 64, 195, Section 4, Chung Hsing Road, Chutung, Hsinchu, Taiwan, ROC, and Department of Educational Media and Library Science, Tamkang University, Tamsui, Taipei, ROC

Received January 5, 1989

A bibliometric study of the application of computers in the analysis of chemistry, from 1984 to 1986, has been conducted. A total of 3817 references in *CA Selects*: *Computers in Chemistry*, 1986, were analyzed by a personal computer. Most of the articles appeared as journal papers, which contributed about 87% of the total literature. The Bradford–Zipf law was applied to investigate the journal literature. Thirty core journals were identified. Seven of the top 10 journals are physics oriented. *Analytical Chemistry* and *Analytica Chimica Acta* are devoted entirely to the subject. The vast majority, 6032 of 6972 authors, contributed only one article. The leading authors and their productivity were also studied. English is the predominant language of articles on this subject.

INTRODUCTION

The application of computer techniques to chemistry has become a popular topic during recent decades.[1] The literature describing this subject field has also grown rapidly.[2] A current-awareness bulletin on this subject has been published since 1978.[3] This paper explores some of the characteristics of the literature of this subject including its sources, journals, authors, and languages.

The data for this study were obtained from the bibliography mentioned above. *CA Selects*: *Computers in Chemistry* groups abstracts in three parts, titled "Biochemistry and Information Science", "Process Control", and "Analysis". Only the literature on the subject of computers in "Analysis", limited to abstracts published by *Chemical Abstracts* (*CA*) in 1986, will be investigated in this paper. The publication date of the original documents ranged from 1984 to 1986. "Analysis" contains abstracts of papers on the application of computers in synthetic organic, physical, inorganic, and analytical chemistry, including such studies as computer-assisted structure elucidation, computer-aided synthesis design, computer-aided evaluation, measurements, simulation, mathematical modeling, and programmed computations. *CA Selects* is a series of publications of the Chemical Abstracts Service, which covers the international literature of chemistry. Care has been exercised to examine the data collected to assure their identity. Subsequently, the data were sorted and manipulated by using a personal computer. By application of bibliometric techniques, especially Bradford's law, the results of this study were interpreted.

The objectives of this study are (1) to find the most significant primary journals on this subject, (2) to determine a nucleus of primary journals that contain a substantial portion of the journal literature on computers in the analysis of chemistry, (3) to find the productivity distribution of authors on this subject, and (4) to find the dispersion of kinds of languages and document types of the literature on computers in the analysis of chemistry.

LITERATURE SOURCES

Some of the detailed characteristics of the literature on the subject of computers in the analysis of chemistry will be discussed below. The references in the bibliography were classified according to their document types: journal articles, books, conferences, technical reports, patents, dissertations, and workshops. Table I shows the number of references from each source. As is common in many fields, the single most prevalent form of publication is the journal article, which contributes about 87% of the total literature. Conferences, the next most frequent type, account for about 5%. Technical reports, patents, books, dissertations, and workshops make up the remaining 5%.

Table I. Distribution by Document Type

document type	no. of articles	cum articles	%
journal	3333	3333	87.3
conference	200	3533	5.2
technical report	113	3646	3.0
patent	97	3743	2.5
book	45	3788	1.2
dissertation	25	3813	0.7
workshop	4	3817	0.1

BRADFORD PLOT

Bradford's law[4] has been widely used to study journal literature distribution. The journal literature of computers in the analysis of chemistry was fit to Bradford's law by plotting the cumulative number of papers for each journal versus the logarithm of its rank. This plot is presented in Figure 1. The graphical formulation of Bradford's law has a characteristic smooth S-shaped curve, with the final drop portion lying below the linear portion of the curve. The curve can be divided into three sections. Section one, the initial nonlinear portion of the curve, which is called the nucleus, contains the most productive sources and represents the higher density of the nuclear zone. Section two, a linear portion, is called the Zipf distribution. Section three, which is another nonlinear portion at the end of distribution, often called the "Groos droop", shows the departure from linearity for higher values of n.

As shown in Figure 1, the Bradford plot of the subject studied in this paper is very similar to the typical one. A saturation effect[5] can be observed at the beginning portion of the plot. There are 30 journals in the nucleus zone. The linear portion of the curve is within $30 \leq n \leq x$. The slope of this portion indicates that the total number of journal titles on this subject is 683 (according to $R(n) = K \log (n/s)$ $(c \leq n \leq k)$, where K is the slope of the linear portion of the Bradford–Zipf bibliograph and also the total number of journal titles that would be expected to publish papers on the subject and s is the value of n at the intersection of the straight portion of the bibliograph with the log (n) axis).[4] The anticipated number

0095-2338/89/1629-0156$01.50/0 © 1989 American Chemical Society

Table II. Distribution of Journal Articles

no. of articles	no. of journals	total no. of articles	%
1	306	306	9.2
2	112	224	6.7
3	65	195	5.8
4	41	164	4.9
5	21	105	3.2
6	21	126	3.8
7	15	105	3.2
8	8	64	1.9
9	10	90	2.7
≥10	84	1954	58.6
total	683	3333	100.0

Table III. Core Journals

rank	journal title	no. of articles	cum articles
1	*J. Chem. Phys.*	141	141
2	*Nucl. Instrum. Methods Phys. Res.*	71	212
3	*Phys. Lett.*	69	281
4	*Phys. Rev. A*	59	340
5	*Phys. Rev. B*	54	394
6	*Anal. Chem.*	50	444
7	*Nucl. Phys. B*	50	494
8	*Comput. Phys. Commun.*	45	539
9	*Anal. Chim. Acta*	44	583
10	*Mater. Res. Soc. Symp. Proc.*	44	627
11	*J. Appl. Phys.*	43	670
12	*Rev. Sci. Instrum.*	43	713
13	*Phys. Rev. Lett.*	42	755
14	*J. Phys. Chem.*	39	794
15	*J. Vac. Sci. Technol.*	33	827
16	*Chem. Phys. Lett.*	31	858
17	*Surf. Sci.*	31	889
18	*Appl. Opt.*	28	917
19	*J. Phys., Colloq.*	28	945
20	*Mol. Phys.*	28	973
21	*IEEE Trans. Electron Devices*	27	1000
22	*Phys. Rev. D*	25	1025
23	*Physica B+C (Amsterdam)*	25	1050
24	*Appl. Spectrosc.*	24	1074
25	*Fluid Phase Equilib.*	24	1098
26	*Comput. Chem.*	23	1121
27	*J. Magn. Reson.*	23	1144
28	*Fresenius' Z. Anal. Chem.*	22	1166
29	*J. Chromatogr.*	22	1188
30	*Phys. Status Solidi Appl. Res.*	22	1210

of journals can be estimated by the slope of the linear portion of the Bradford plot. Arbitrarily taking two points from the linear portion of Figure 1 gives

$$R(20) = 973$$

$$R(80) = 1914$$

$$K = \text{slope} = (1914 - 973)/[\log (80) - \log (20)] = 683$$

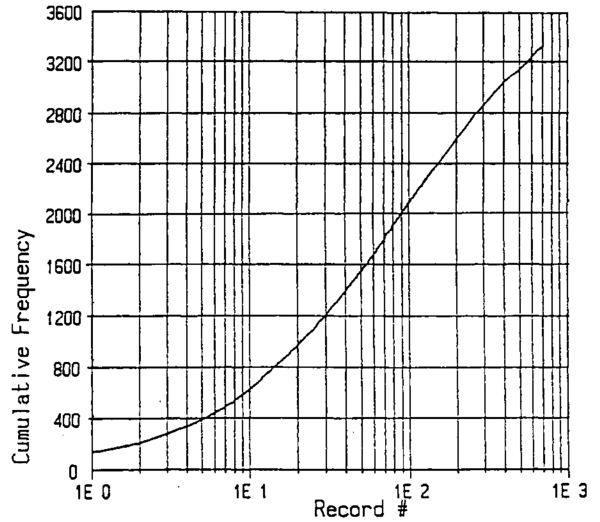

Figure 1. Bradford–Zipf distribution of journal literature.

The anticipated number of journals agrees perfectly with the actual number.

JOURNAL LITERATURE

On the basis of the Bradford plot the journal literature is further subdivided according to the number of papers published by each journal. As can be seen from Table II, a total of 683 journals are involved; about 300 [306 (45%)] of these contain only one paper. Eighty-four journals contain 10 or more papers. The top 30 core journals are listed in Table III along with the number of papers they published. The leading journal is *Journal of Chemical Physics*, publishing 141 papers (3.7% of the total, 4.2% of the journal papers). The top 30 journals in Table III contribute 1210 papers (32% of the total articles, 36% of the journal papers). They are clearly the core journals on the subject of computers in the analysis of chemistry. It is interesting to note that of these core journals in which papers on computers in the analysis of chemistry have appeared the majority of them are not devoted solely to chemistry. The subject scope of the top 10 journals is described in Table IV and has been drawn from Katz[6] and from the title page of each journal. Seven of the top 10 journals are physics oriented. This is of no surprise. Although computers in the analysis of chemistry is a chemistry subject, it is not uncommon that the subject has some connections with physics. *Journal of Chemical Physics*, which publishes the most papers on this subject, is an excellent example showing the relation between physics and chemistry. As the title indicates, *Analytical Chemistry* and *Analytica Chimica Acta* are devoted entirely to the subject of chemical analysis. Only *Computer Physics Communications* emphasizes the computational aspect. *Physics Letters, Physical Reviews A and B, Nuclear Physics*

Table IV. Subject Scope of the Top 10 Core Journals

rank	journal title	subject scope
1	*J. Chem. Phys.*	structure and spectra of atoms and molecules; statistical mechanics of gases, liquids, and solids; thermodynamical properties of chemical processes; and the interaction of radiation with matter
2	*Nucl. Instrum. Methods Phys. Res.*	instrument and techniques in nuclear physics and related areas; high-energy physics; nuclear spectroscopy; numerical and data analysis; and nuclear electronics
3	*Phys. Lett. B*	nuclear elementary particle and high-energy physics
4	*Phys. Rev. A*	general physics; atomic and molecular physics; fluids; thermodynamics; and plasmas
5	*Phys. Rev. B*	solid state; atoms in matter; Mössbauer effect; electrons and ion emission; range and energy loss; magnetic impurities; ferroelectrics; and spectra of solids
6	*Anal. Chem.*	theoretical and methodological analytical chemistry
7	*Nucl. Phys. B*	quantum field theory; study of fundamental particles
8	*Comput. Phys. Commun.*	theoretical physics, mathematics, computer science, and information engineering; computational physics and physical chemistry
9	*Anal. Chim. Acta*	aspects of modern chemical analysis both fundamental and applied
10	*Mater. Res. Soc. Symp. Proc.*	materials researches

Table V. Productivity of Authors of All Types of Documents

no. of articles	no. of authors	%
1	6032	86.52
2	701	10.05
3	154	2.21
4	52	0.75
5	18	0.26
6	11	0.16
7	2	0.03
8	1	0.01
9	1	0.01
total	6972	100.00

Table VI. Authors Contributing Five or More Articles

rank	author	no. of articles	rank	author	no. of articles
1	Wang, J.	9	18	Duesbery, M. S.	5
2	Hess, K.	8	19	Eckstein, W.	5
3	Jorgensen, W. L.	7	20	Enmanji, K.	5
4	Mazzone, A. M.	7	21	Kogut, J. B.	5
5	Baranyai, A.	6	22	Kowalski, B. R.	5
6	Binder, K.	6	23	Kozmutza, C.	5
7	Catlow, C. R.	6	24	Li, Y.	5
8	Evans, M. W.	6	25	Lugli, P.	5
9	Ferry, D. K.	6	26	Matsushita, M.	5
10	Jinno, K.	6	27	Meakin, P.	5
11	Kleinert, H.	6	28	Montvay, I.	5
12	Liu, Y.	6	29	Singh, J.	5
13	Ranft, J.	6	30	Sogo, K.	5
14	Ruff, I.	6	31	Teper, M.	5
15	Sasaki, H.	6	32	Thijssen, P. C.	5
16	Brennan, K.	5	33	Yokoyama, K.	5
17	Cormack, A. N.	5			

B, and *Nuclear Instruments and Methods in Physics Research* are totally physics oriented. *Materials Research Society Symposia Proceedings* is a conference series.

AUTHORS

A total of 6972 authors, including senior personal authors and coauthors, are cited in the bibliography. Table V presents the number of authors contributing one article, two articles, and so on. The articles can be in any of the six document types. The vast majority [6032 authors (87%)] contributed only one article. The largest number of articles by one author is nine. Thirty-three leading authors who contribute five or more articles are shown in Table VI. J. Wang is the most productive author; he contributed nine articles. K. Hess, who published eight papers, is next.

LANGUAGE

Table VII shows that English is the predominant language of articles on computers in the analysis of chemistry. English-language articles constitute 79% of the total (3030 of 3817). There are 787 non-English-language articles or 21% of the total, and they are written in 19 languages. It is clear that there are essentially only four languages for computers in the analysis of chemistry literature, namely, English, Russian, Japanese, and Chinese. Contributions in these languages result in about 94% of all material.

Table VII. Language Distribution

language	no. of articles	%
English	3030	79.4
Russian	296	7.8
Japanese	149	3.9
Chinese	128	3.4
German	91	2.4
French	22	0.6
Hungarian	20	0.5
Czech	14	0.4
European[a]	12	0.3
Spanish	11	0.3
Italian	8	0.2
Polish	8	0.2
Dutch	6	0.2
Romanian	6	0.2
Slovak	6	0.2
Korean	4	0.1
Bulgarian	3	0.1
Portuguese	1	0.03
Turkish	1	0.03
Ukrainian	1	0.03
total	3817	100.29

[a] These 12 articles were all in forms of patents published in *Eur. Pat. Appl.* The respectively corresponding language could not be identified.

CONCLUSION

It is often possible to gain insight into a field by studying some of the statistical characteristics of its literature,[7] and computers in chemistry is no exception. The following conclusions may be drawn from this study: (1) The journal literature of computers in the analysis of chemistry can match Bradford's law in both graphical and mathematical formulation. (2) Thirty core journals contributing significantly to the subject can be identified. Most of these journals cover both chemistry and physics. *Journal of Chemical Physics* is the leading journal and publishes the largest number of papers on the subject. (3) Most authors published only one paper on this subject in the period 1984–1986. The most productive author published about 10 papers in the same period. (4) Four essential languages are identified, namely, English, Russian, Japanese, and Chinese. English-language articles contributed about 80% of the total number of articles.

REFERENCES AND NOTES

(1) The American Chemical Society (ACS) has a Division of Computers in Chemistry that has been active since 1974 and regularly presents symposia at the ACS National Meetings.

(2) Symposium on the History of Computing in Chemistry. *Abstracts of Papers*, 192nd National Meeting of the American Chemical Society, Anaheim, CA; American Chemical Society: Washington, DC, 1986; COMP 1-6, HIST 1-10.

(3) *CA Selects: Computers in Chemistry*; American Chemical Society: Columbus, OH, 1986; 26 issues.

(4) Bradford, S. C. *Documentation*, 2nd ed.; Crosby Luckwood: London, 1953; p 154.

(5) Brookes, B. C. The Derivation and Application of the Bradford–Zipf Distribution. *J. Doc.* **1968**, *24*(4), 247–265.

(6) Katz, B.; Katz, L. S. *Magazines for Libraries*, 5th ed.; Bowker: New York, 1986.

(7) Hawkins, D. T. The Literature of Noble Gas Compounds. *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 190–199.