CHAPTER 1

Introduction

1. The Problem

Scientific and technical study is in a state of continuous flux. Sectors once considered to be strategic become obsolete, while new and promising fields appear. The importance of the analysis of the dynamics of science and technology is scarcely in doubt and has accordingly attracted much interest. A quantitative concern with the scientific change is to be found in range of disciplinary approaches[1]. It has been of a central concern of philosophers of science (*see* Popper[2]; Lakatos[3]; Toulmin[4]; and Hesse[5]). Sociologists of science have displayed a similar concern in their numerous studies of scientific controversies (e.g. Collins[6] and Mackenzie[7]), specially growth studies (e.g. Mullins[8]) and day-to-day work in scientific laboratories (e.g. Latour and Woolgar[9]). Such an interest has also been depicted in the work of many historians of science (e.g. Kuhn[10]; Shapin[11]; and Thackray[12]), and writers on science policy and political dimensions of science (Weingart[13] and members of the Starnberg School; Rip[14]; Bohme, van den Daele, Hohlfeld, Krohn and Schafer[15]).

In philosophy of science, scholars focused their attention on problems in the origin, substantiation and development of scientific knowledge, its functions, structure and dynamics. In each historical period, various problems and objects were at the center of science investigation and new models of development were discussed. By the second half of 20th century, methodological conceptions and models were more actively discussed in the philosophy of science, where even greater importance was attached to such concepts, as scientific speciality, science paradigm, research program, and social group. Parallel to this, some works appeared on the horizon focusing on the problems in the formulation and development of new scientific directions and producing maps of scientific disciplines, as well as studying scientific knowledge and scientific activity in their sociological aspects, leaning upon various forms of scientific publications or studying group of scientists and types of communication in science. At the same time, one also started witnessing convergence and interaction between the philosophical and sociological approaches in science. However, from 1960's onwards the philosophical-methodological approach to this range of problem has been increasingly supplemented by historical, sociological and bibliometric approaches. The trend towards their rapprochement and integration became even more distinct[16].

Under, bibliometric approach, a series of new methods and techniques were developed and used by scholars such as D. de Solla Price, Eugene Garfield, Henry Small, Belvith Griffith, R. Merton, Diana Crane, N. Mullins, S. Cole, J. Cole, H. Zuckerman, W. Mulkay and many others, who threw light and generated information on the analysis of science and social and cognitive forms of knowledge representation at the level of scientific speciality, in both natural and social science.

The bibliometric approach to science and technology is primarily based in quantitative characteristics, attributes or objects of documentary flows. It is primarily based on the analysis of the bibliographic data on publications, which comprises attributes such as title (including keywords and phrases), authors (including their address, co-authors and reputation), book, journal (title, subject, and origin of country), etc. The objects of study in bibliometric studies are often grouped under various attributes. The methodology of science views scientific knowledge and its development through the presentation of objects described in a scientific publication.

Scientific publication is by no means an invariable piece of information that can be published, stored, retrieved and delivered on request. A scientific publication is a kind of written material containing information with respect of scientific activities – either in its physical form, or its electronic equivalent in a computerized database[17]. It is assumed that it is a changing particle of the social web of science, produced under social conditions and used under others. A scientific paper or text not only reveals the world-building strategy of its authors, but also the nature and force of the building blocks derived from the domain of science from which it draws and to which it contributes. Thus, it provides access to the dynamics of science to the shared worlds that constitutes a means of mutual (or evolving) control[18].

Each publication is uniquely represented by the bibliographical information it contains. In addition, each distinct type of publication has its own set of common characteristics and attributes. Two basic units of bibliographic information are generally used: (a) items, which refer to information in or about publications. The principle items of analysis are author names, addresses, citations, and keywords, as well selected words from the title, abstract, or the full text. Abstracting and indexing services often add further bibliographical information such as type of publication, language, one or more controlled terms (keywords given by data producers), subject classification terms or codes and (b) entities, which refer to (aggregates of) publications. These entities represent a set of publications, starting from the micro-level (a single author) via meso-levels (varying in size from university departments, scientific journals to research specialities) to macro-level entities (scientific sub-fields or countries). The information required for bibliometric analysis is derived from not only primary scientific journals, but also from indexing and abstracting journals and databases.

There are two common approaches to the quantification of information flow[19]: (a) The first built directly on the work of Price and his predecessors consists in following the dynamics of separate items or objects: publication as a whole or its attributes such as authors' names, addresses, citations and keywords, as well selected words from title, abstract, and in text, journal title, its country and subject, etc.; and

(b) The second approach consists in identification of links between objects, their correlation and classification.

In the first approach, depending upon the choice of the object for a bibliometric study, different pictures of dynamics can be obtained as regards to the state of science and technology. For example, one can model the growth of a scientific field and the direction it is going to follow, by measuring its annual publication data. Similarly, if the bibliometric study has its object of study, e.g. keywords, then alternations in keyword frequencies can reflect both some substantial regrouping of research directions in science, and (much more rarely) linguistic changes. Similarly, the study of the distribution of productivity of authors and articles in journals in different time periods can be considered in terms of indicators, which can throw light on the level of development in the scientific specialities. The use of citation data and their age distribution over different time periods can also reveal the level of obsolescence and the structural changes taking place in the field from time to time.

Under the second approach, the analytical bibliometric procedures are not directed at obtaining characteristics, but to gain statistics relating to quantitative indicators of information files and also to identify with their help qualitative structures of science and its individual fields, by mapping them. The concept structure implies information on relations among constituting elements. Bibliometrics has a variety of informations at its disposal to study the relational aspects of study. The most important are: the authors of co-authored publications, their addresses, keywords from the text, and the reference lists. Each aspect reflects a different aspect of science. For example, the address can be used for assessment of international cooperation, while the citation relation may indicate intellectual links between sets of publications.

Bibliometric maps can be built for exploring the underlying structure of similarities and interrelationships between items and / or entities. So far, four types of bibliometric maps have been frequently used: (a) journal-to-journal citation maps[20], indicating the magnitude of subject relations between journals and subsequently reflect the macro-level structure of scientific activities; (b) co-citation maps[21], depicting the extent to which two articles are cited together in other articles. Clustering procedures make it possible to obtain co-citation clusters which reflect the structure and content of research-front specialities; (c) co-word maps, where the count of word co-occurrences are used to construct a co-word structure representing research themes in a discipline and their inter-relations thus capturing the cognitive structure of speciality[22]; (d) co-classification maps are based on the analysis of the co-occurrences of subject classification terms. The classification terms mostly represent cognitive elements, which relate to scientific topics, specialities, or fields[23].

1.1 Need for the study

Most of the earlier studies in the field of bibliometrics have focussed on the first approach, where individual attributes or objects of publications are utilized to study quantitatively the research activities in the scientific specialities. These studies, however, suffer from certain shortcomings and consequently convey limited meaning. These studies have been confined to fixed time intervals, which varied from one study to another. These studies, therefore, provided only a sunspot picture of the research activities of the scientific specialities in a particular period of time. More recently, studies started appearing in the field focussing on study of the dynamics of change in scientific specialities, by measuring their research activities in successive time intervals. These studies were largely confined themselves to only a small span of the life cycle of the scientific specialities. These studies were mostly unidirectional and did not present different perspectives simultaneously. As a result, the need was felt to conduct the present study.

1.2 Statement of the problem

The focus of the present study is on the application of bibliometrics methods to study quantitatively the growth and dynamics of scientific activities in theoretical population genetics speciality from almost since its inception, i.e., 1881 to 1980. The dynamics of change have been studied, by quantitatively analyzing the various objects or features of publications, in successive time period blocks. Today being the age of genetics, the speciality of "Theoretical Population Genetics" was chosen for this study. Another major reason for selecting the speciality was the availability of a source of information, which could provide us a comprehensive output of literature for about 100 years in this scientific speciality. The availability of data for such a long period on a speciality helped us in studying the dynamics of change in activities of this scientific speciality effectively, which is normally not possible due to paucity of data for such a long period.

2. Objectives of the Study

The main objectives of the study are:

(i) To evaluate the growth of publications and authors in the field of theoretical population genetics from 1907 to 1980, and to evolve a suitable model for this growth;

(ii) To determine the changes in the pattern of collaboration between authors and also between countries with time, as reflected in the publications of theoretical population genetics speciality;

(iii) To determine the obsolescence of literature through the age of citations appearing in the publications of this speciality ; and

(iv) To study the productivity distribution with time of authors in this field, using different approaches.

3. The Hypotheses

The following hypotheses are proposed to be tested:

A. Whether the growth of literature and authors follows an exponential or logistic growth model, as is generally believed?

B. Whether the extent of collaboration increases with the growth of the publications?

C. Whether the logistic model or any other model is applicable to the decline in percentage share of single-authored publications with time?

D. Whether any one or more than one statistical probability distribution is applicable to the number of authors per publications with time?

E. Whether the authors prefer to collaborate with authors at the same productivity level?

F. Whether the impact of funding and collaboration tends to increase the productivity of individual authors?

G. Whether the nature and types of collaboration as reflected in publications changes with time?

H. Whether any statistical probability distribution is applicable in the age distribution of citations with time?

I. Whether there is any relation between growth rates, obsolescence rate, and half-life?

J. Whether Lotka's law and any statistical probability distribution describe the scientific productivity distribution of authors?

K. Whether there is any relation between scientific productivity distribution and development of the TPG speciality?

L. Whether there is relation between extent of duration of participation and the scientific productivity distribution of authors?

M. Whether there is any statistical regularity in the frequency distribution of the speed of contribution by authors?

4. Database and Methodology

4.1 Database used

The literature on the speciality of theoretical population genetics is scattered across different types of publications and is in different languages. Although there are a few abstracting and indexing services and databases which cover literature on this speciality, we have only used the database available in the form of a printed bibliography entitled " Bibliography of Theoretical Population Genetics", compiled by Felsenstein and published in 1981. It is the most important bibliographical source of information on this speciality, covering literature from 1881 to 1980, almost since the inception of the speciality. The comprehensiveness of this source has been checked with the major review articles published in the field and also from publications depicting the history of theoretical population genetics speciality.

4.2. Methodology

Numerous methods are available for analysis of activities in scientific specialties. They can be grouped under qualitative and quantitative methods. The method used in the present study is the quantitative one, and is termed as the "bibliometric approach". The application of several mathematical and statistical techniques, methods and models has been studied, and is described here.

In the second chapter, selected linear and non-linear growth models have been introduced and studied for their application to the cumulative growth of publications and new authors in theoretical population genetics speciality. The growth models are selected on the basis of fit statistics and visual inspection of graphical presentation of the observed and estimated data values obtained from the application of different growth models. A methodology for identification and classification of growth models, based on two growth rate functions, has been described and utilized for identification of best model(s) describing the cumulative growth of publications and authors in this field.

In the third chapter on collaborations in the area of the theoretical population genetics, the growth in the proportion of collaborative publications in total publications in theoretical population genetics speciality has been considered in block of five-year publication data during 1896-1980. The growth in the proportion of various types of collaborative publications has been explored. In order to get a picture of relative growth of single-authored and multi-authored publications; collaborative publications have been analyzed in the form of relative frequency of publications by number of authors for different period blocks during 1931-35 to 1976-80. The decline in percentage share of single-authored publications with time in theoretical population genetics speciality has been studied and modelled to find whether any particular growth model fits in the publication data.

Some mathematical measures, such as degree of collaboration, collaboration index, and collaboration coefficient have been computed for different period blocks during 1916-20 to 1976-80 to study the increase in the extent of collaboration and degree of professionalisation in theoretical population genetics speciality with time. The applicability of selected statistical probability distributions has been studied, for their goodness-of-fit, in publication data on the number of authors per publication in different period blocks during 1941-1980. The statistical probability distributions, showing best fit on the publication data have been selected on the basis of basic statistics and parameter values obtained from the application of different models and through the use of statistical test. It has also been explored whether the average number of collaborators per author can be used as a better substitute in Lotka's law, because it takes into consideration the productivity-collaboration linkages. The relation between collaborating authors and collaborators has been explored, using a new index called the "homophily index", depicting the levels of productivity (measured by rank) at which the authors collaborate among themselves, using analogy from a social psychological theory. Three characteristic properties of the collaborative structure in theoretical population genetics speciality have been studied through "Birds of the feather flock together" phenomenon, edge effect, and inverse relationship features depicted through the three-dimensional representation of homophily indices.

For studying the nature and types of collaboration and the impact of funding and collaboration on the productivity of authors, the selected data from 11 core journals in the discipline of theoretical population genetics have been used in the period blocks of five years from 1956-60 to 1976-80. This sample consisted of 2502 publications, of which 1474 publications were single-authored (468 publications of which were funded) and 1028 publications were multi-authored (667 publications of which were funded). To find the productivity of individual authors, the normal count method has been used for counting their productivity. To evaluate the impact of funding on the productivity of individual authors, the data in each period block were subdivided into two subgroups - funded and non-funded - and their productivity distributions have been considered, besides looking into the average productivity per author. For studying the effect of collaboration on productivity of authors, the data in each period block were divided into two subgroups - collaborated and non-collaborated - and their productivity distributions have been studied, besides looking into the average productivity per author.

To evaluate the nature and types of collaboration, the affiliations of authors, as reflected in publications has been culled from primary journals and analyzed. The publications have been de-segregated under three types: internal collaboration, domestic collaboration, and international collaboration. For simplification, the total publications have been divided into two groups: A and B. In Group A, publications from United States have been included, as it constitutes the largest block of publications. In Group B, publications from all other 42 countries have been included, which together contributed 47.04% to the total publications.

In the fourth chapter on growth and obsolescence of literature in theoretical population genetics, a synchronous analysis of the selected publication data has been done. For studying the age of citations in source journal papers, data have been culled from the core journals pertaining to the years 1929, 1939, 1949, 1959, 1969, and 1979. These data sets have been selected because they represent different phases of development of theoretical population genetics speciality. Different mathematical indices, such as half-life and immediacy index have been explored, for studying the extent of obsolescence in the TPG speciality with time. The densities of the age of references cited in source papers of TPG pertaining to years 1929, 1939, 1949, 1959, 1969, and 1979 have been explored for finding the statistical probability distribution best fitting in these data sets. To study the relation between growth rate and obsolescence rate, two other types of data have been collected: (a) data related to the growth in the number of publications from 1931 to 1980; and (b) data on age distribution of citations appearing in publications for the source years 1931, 1937, 1943, 1949, 1955. 1961, 1967, 1973, and 1979, representing different phases of development of TPG. The growth rate and obsolescence rate, using both regression and exponential methods, and half-life has been computed from age distribution of citations and publications for different source years. A methodology and relationship between growth rate, obsolescence rates, and half-life has been suggested and used.

In the final chapter on scientific productivity of authors, an overall productivity distribution of all authors participating in theoretical population genetics speciality during 1881 - 1980 are studied, and the applicability of Lotka's law and other statistical distributions has been explored for their fit. Distribution of scientific productivity of all authors, given the equality of chances of their participation has been considered, to find the impact of longer activity of authors on their productivity distribution. Then, the distribution of the productivity of different group of authors, active for the same length of time, using different criteria has been evaluated, and the limitation in such analyses have been pointed out.

The distribution of productivity of authors in three different phases of the development of theoretical population genetics pertaining to period blocks 1921-50, 1951-65, and 1966-80 has been considered in terms of scattering and concentration indices and relation to its developments cycle. Next the frequency distribution of the scientific productivity of authors having equal duration of participation has been studied. The distribution of scientific productivity of authors having one year, three years, five years, seven years, 10-14 years, 20-24 years, and 30-34 years of participation has been analyzed. The data on the duration of participation and the number of authors involved have been tested in understanding the statistical model best fitting these. Lastly, the speed of publications by authors has been studied and modelled for two groups of authors who have published: (i) 2 to 7 papers, and (ii) 8 or more papers, separately.

4.3 The scope of theoretical population genetics studied

This study is proposed on theoretical population genetics speciality, defined to include theoretical work in quantitative genetics and in statistical human genetics. For analysis, literature covered should have theoretical contents. The theory must deal with genes segregating in populations not solely with statistical analysis or ecological interactions. The field is multidisciplinary in character having roots in many disciplines, such as mathematics and statistics, agriculture, botany, zoology, and medicine.

The science of heredity, namely genetics, had its beginning in 1866, when Gregor Mandel published the results of his experiments on the pea plant. The mathematical rules, which were usually formulated, went unnoticed for three decades and were rediscovered in 1900. These principles of inheritance which Mandel recognized became the foundations of genetics and later a major factor in the development of biology.

Application of genetical principles to entire population of organisms constitute the subject of population genetics, where population refers to a group of organisms of the same species living within a prescribed geographical area. Population genetics considers as they occur in population and as they changed in evolution. Three major areas of population genetics are (i) the study of genetic polymorphism's, (ii) study of evolution, and (iii) the development of the theory of population dynamics.

The population genetics aims at studying the frequency of genes and their equilibrium, the estimation of equilibrium frequencies in natural population, changes in the gene frequencies due to mutation, selection, migration, etc., genetic structure of population, speciation, and evolution, and even extrapolates the topic such such as cultural and biological evolution and the effect of deleterious genes and eugenics.

While moving from chromosomal or molecular genetics to population genetics, the major shifts in the orientation of the subject are necessary. Firstly, the emphasis changes from the individual - individual base pair, gene, locus, or strain - to large collection of individuals. Secondly, instead of considering stable genetics markers, pure lines, uniform enzymes, assays, etc., the focus is on the variation and the amount of variability within a population. Thirdly, time scale to be considered is also different, as some of the phenomena in population genetics may take thousand years to occur.

The most salient feature of natural population is its diversity, which is obvious when we consider any specific plant, animal or even human being. The sources of variability and many of the mechanisms of evolution remained unclear until post - Mandelian population concepts developed by Haldane, Fischer, and Wright.

The forces of natural selection, mutation, and migration acting in a given population of given size within a given breeding structure and reproducing by known rules of genetic behavior are responsible for the genetic constitution of a population. The study of quantitative relationship between the forces and the resultant distribution of a genotype is known as theoretical population genetics. This subject mainly involves determining the rate of change of genetic structure of population or the equilibrium conditions in the form of mathematical structures in which the values of population size, mutation rates, selection intensities, migration rates, recombination rates or other relevant parameters are substituted. Therefore, theoretical population is the study of the possible rather than the actual.

The subject population genetics progressed very rapidly during the first thirty years after the rediscovery of Mandalism. It concentrated mainly on the population dynamics of single-locus systems under the constant force of natural selection, mutation, and breeding structure of population. In the subsequent years, emphasis shifted to a more complex genetic systems (many loci and many alleles) and the environment varying in space and time.

The discovery of protein polymorphism in 1966 by Lewontin and Hubby led to the experimental evidence of the existence of a large amount of genetic variation in natural populations in a relatively very short span of time in a vast array of texa. Since the 1970's, the discovery of molecular biology and the refinement of techniques of studying genetic variation led to the availability of large data. This, in turn, has enabled the scientists to test various hypotheses about forces operating on these variations. This has greatly helped in determining the rate of natural selection, mutation drifts, etc., and their effects on the natural population.

The importance of theoretical population genetics lies in the fact that subject has the potential to predict the genetic consequences of a particular course of action in plants and animal breeding, and in even in human society.

A major issue, which has kept theoretical population geneticians, busy in the debate between "balance" and "neutral" schools for a long time now. According to the neutral theory which is essentially Non-Darwinian, it is proposed that most of the differences at the molecular level (amino acids in proteins) in different species have occurred without natural selection and are, therefore selectively neutral. On the other hand, balance school regards the standing of genetic diversity within a population as being a consequence of balancing forms of natural selection.

Today the population genetics stands on a threshold of a new period of scientific advance. The traditional theoretical and empirical methods of genetics and population genetics have been greatly supplemented and strengthened by the powerful tools of molecular biology to provide a closer view of genetic diversity and the evolutionary mechanisms that shape genetic diversity. Population genetics is in a position today to provide the bases for understanding environmental adaptations and theoretical framework for deeper understanding of improvement programs of plants and animals, as well as genetic consequences and then possible impact on the human society of any event affecting human genetics.

4.4 Characteristics of literature used as reflected in the source

It will also be not out of context to analyze here the characteristics of literature covered in the bibliography, with a view to gather ideas on the nature, format, language, and output of various kinds of literature in the field.

Growth of Publications and Authors

The annual growth of publications and authors during 1907 to 1980 in the field of theoretical population genetics is shown in Figure 1.1. A break-up of the growth of publications and new authors in blocks of ten years is presented in Table 1.1.

Table 1.1
Growth of publications and new authors in blocks of ten years during 1881-1980

Period blocks	Number of publications	Number of new authors
1881-1890	3	2
1891-1900	8	5
1901-1910	36	12
1911-1920	77	32
1921-1930	164	68
1931-1940	282	78
1941-1950	222	82
1951-1960	647	292
1961-1970	2198	940
1971-1980	4143	1698

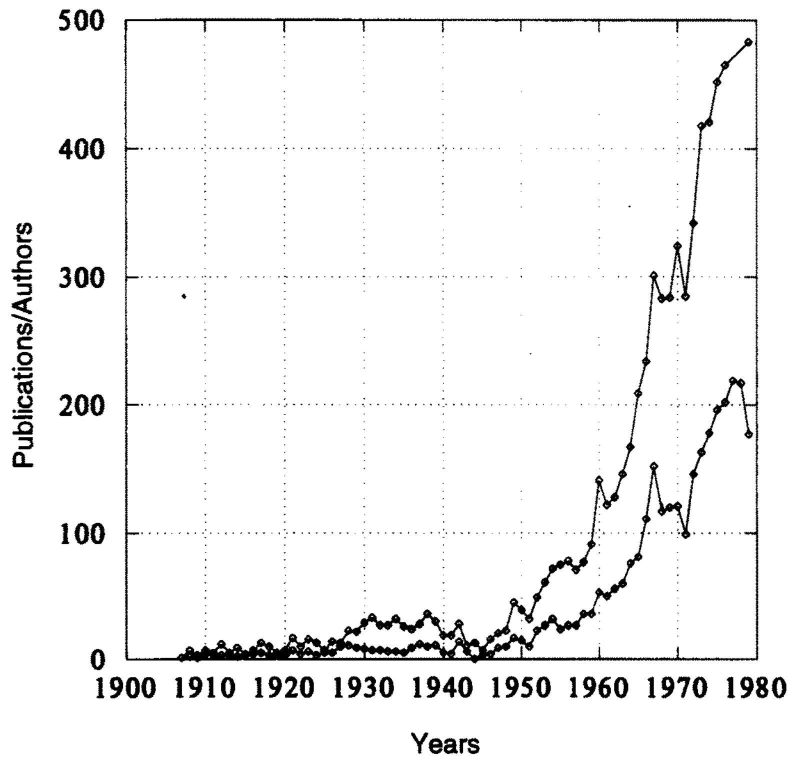

Fig. 1.1. Annual growth of publications and new authors during 1907-1980

It is observed from this table that the publications are increasing at a fast rate and appears to be doubling from 1901-10 to 1931-40. Then the increase in the growth of the publication slowed down, because of World War, and then again started increasing fast. As a result, it showed a tripling growth from 1941-50 to 1961-70, and then doubling again up to 1971-80.

In the case of new authors, substantial increase is only noticed from 1901-10 to 1921-30, and then from 1941-50 to 1961-70. We can clearly delineate different phases in the development of theoretical population genetics. The period from 1901 to 1920 represents the stage where there are too many fluctuations and the number of publications and new authors are also very small.

The first phase, spanning its publication activities from 1921 to 1950, indicates the initial development of the area, with quiet week activity. Here there are many transient authors and few authors with more than short activity. Although the number of publications and new authors entering in the field at this stage are small, but the overall growth rate is comparatively very high and also fluctuating within the phase.

The second phase, spanning from 1951 to 1965, is known as the pioneering stage, illustrating the exploding learning processes of few pioneering contributors. As a result, one notices a sudden spurt in the number of publications and new authors here, indicating the rapid growth of theoretical work in population genetics. The rate of growth of publications and authors is around 10 % to 15% during this phase.

The third phase, spanning from 1966-80, symbolizes the stage, where the area has attracted many new authors, but at the same time explosion in the learning curves of the pioneers has probably reduced. However, there is an extensive activity of the followers of the pioneers. This is reflected in improvement, completion, relativization, diffusion, and applications of new knowledge. As a result, we find that although there is an exponential growth of publications and authors again, their growth rate has comparatively fallen down to around 5%.

The last phase 4, in which normally the perceived interest in the field declines and falls sharply and consequently opportunities decline and authors start migrating to other fields is still missing in our data on theoretical population genetics publications.

Considering the format of publications from 1901 to 1980, we can classify the total publications under following types: journal papers (83.54%), papers in books (6.50%), conference/seminar papers (4.61%), books (2.48%), dissertations (1.95%), etc. A chronological break-up of publications, as shown in Table 1.2 indicates that share of publications reflected through journals continue to remain dominant during the entire period, although between two fluctuating extremes, its contribution was 97.22% (during 1901-10) and 78.67% (during 1951-60). The percentage contribution of seminar/conference papers and papers in books have consistently increased, while that of books, dissertations, and others first increased and then dropped during the last block year 1971-80.

Table 1.2
Growth of publications by types of format in blocks of ten years during 1901-1980

Period blocks	Number of publications of different format						
	JP	CSP	BP	Books	Dissertation s	RP	Others
1901-10	35	1	-	-	-	-	-
1911-20	65	-	3	5	-	3	1
1921-30	151	-	4	8	-	-	1
1931-40	259	6	2	13	2	-	-
1941-50	190	8	10	11	3	-	-
1951-60	509	58	46	14	13	3	4
1961-70	1808	114	112	58	73	22	11
1971-80	3473	171	328	84	61	16	10
Total	6490	358	505	193	152	44	27

JA = Journal papers; CSP = Conference/seminar papers; BP = Papers in books; RP = Papers in reports.

The languages in which theoretical population genetics literature are published includes English, German, French, Russian, Japanese, Spanish, Polish, Czech, Chinese, Dutch, Romanian, Hungarian, Serbo-Croatian, Slovak, and Turkish. A language-wise break-up of publications of theoretical population genetics is shown in Table 1.3.

Table 1.3
Growth of publications by types of language in blocks of ten years during 1901-1980

Period blocks	Total publications	Number of publications of different languages					
		English	German	French	Russian	Japanese	Others
1901-10	37	30	6				
1911-20	77	60	15	1			1
1921-30	164	89	65	5	3		2
1931-40	282	205	47	18	11	1	
1941-50	222	205	4	11	1	1	
1951-60	647	590	21	14	5	14	3
1961-70	2198	1889	84	81	66	41	37
1971-80	4143	3594	136	107	226	41	39
Total	7769	66662	378	237	312	98	82

Table 1.4
List of authors having more than 50 publications during 1881-1980

S.No	Range of publication output	Name	Period of contribution	Number of publications
1	Above 100	Kimura, M	1930-80	199
		Nei, M	1953-80	120
		Wright, S	1917-80	123
2	91-100	Maruyuma, Y	1963-80	98
		Morton, N.E	1955-80	98
		Karlin, S.	1962-80	84
3	81-90	Karlin, S.	1962-80	84
4	71-80	Crow, J.F.	1945-79	72
		Ohta, T.	1966-80	75
		Robertson, A	1949-80	79
5	61-70	Cavallis-Sforza, L.L.	1959-79	70
		Feldman, M.W	1966-80	62
		Hill, H.W	1965-80	62
6	51-60	Chakraborty, R	1970-80	56
		Charlesworth, B	1970-80	53
		Cockerham, C.C.	1952-80	59
		Ewens, W.J.	1961-79	58
		Lewontin, R.C.	1953-79	51
		Li, C.C	1948-78	53
		Mather, K.	1936-79	52
		Smith, C.A.B	1947-80	54

Productivity of Authors

There are in all 3209 authors who have contributed 7780 publications in theoretical population genetics from 1881 to 1980. A strong concentration/scattering effect is observed in the distribution of scientific productivity of authors. It has been observed that 93.27% of the total authors, who have contributed from 1 to 10 publications, have contributed 52.66% to total publications. On the other hand, only 1.85% of the total authors has contributed 24.53% to total publications. A list of important authors having more than 50 publications during 1881-1980, along with their output and period of contribution is given in Table 1.4.

Table 1.5
List of journals which published 80 and above papers during 1881-1980

S.No.	Name of the journal	Period covered	Number of papers
1	Genetics	1916-80	649
2	Biometrics	1948-80	355
3	Heredity	1947-80	350
4	American Journal of Human Genetics	1949-80	332
5	American Naturalist	1910-80	327
6	Theoretical Population Biology	1970-80	260
7	Annals of Human Genetics	1953-80	226
8	Evolution	1948-80	146
9	Nature	1930-80	145
10	Journal of Theoretical Biology	1961-80	141
11	Annual Report of the National Institute of Genetics Japan	1951-79	133
12	Theoretical and Applied Genetics	1968-80	132
13	Genetical Research	1961-80	131
14	Proceedings of the National Academy of Sciences USA	1920-80	131
15	Advances in Applied Probability	1970-80	107
16	Journal of Genetics	1922-77	92
17	Genetika	1952-80	85
18	Journal of Heredity	1917-80	83
19	Annals of Eugenics	1930-54	80

Productivity of Journals

If we look at the scattering of publications in various journals, we find again a strong concentration/dispersal effect in the data. In all 6745 publications are originated in 659 journals. On one hand one observes that 88.12% of the total journals (having contributed from 1 to 10 articles) have in all contributed only 17.77% to the total articles. On the other hand, 2.13% of the total journals (having contributed more than 100 papers) have contributed 48.59% to the total articles. A list of the journals, which published 80 and above papers during 1881-1980 is given in Table 1.5.

5. The Chapterization

The complete study has been divided into the following six chapters:

Chapter 1: Introduction

It provides an overview of bibliometrics research, and discusses the basis of the present study, its objectives, scope, hypothesis and explains the methodology and the source material used for the collection of data.

Chapter 2: Growth of literature and authors in the area of theoretical population genetics

This chapter describes different approaches for studying the growth of scientific knowledge, as reflected in publications and new authors. It explores the application of selected growth models to the cumulated growth of publications and new authors in theoretical population genetics speciality from 1907 to 1980. It also describes the criteria on which growth models are selected for their possible applications in theoretical population genetics literature.

Chapter 3: Collaborations in the area of theoretical population genetics

This chapter focuses on the growth and proportion of multi-authors publications in theoretical population genetics speciality, with a view to find out the extent of collaboration in this field. It studies the applicability of suitable mathematical model in the decline of single-author publications and selected probability distributions to the number of authors per publication with time. At what levels of productivity, authors collaborate among themselves are also explored. It also studies the impact of funding and collaboration on the productivity of authors with time. Lastly, the nature and types of collaboration, as reflected in publications, is also studied with time.

Chapter 4: Growth and Obsolescence of literature in the field of theoretical population genetics

This chapter deals with the obsolescence of literature with time, as reflected in age of citations appearing in publications in theoretical population genetics speciality. The probability distribution functions that best fit in the age of citations are explored. The relation between the growth rate and obsolescence rate of literature is also studied.

Chapter 5: Scientific productivity of authors in area of theoretical population genetics

This chapter describes the distribution of scientific productivity of authors in theoretical population genetics, as reflected in its publications from 1881 to 1980, using different approaches. It specifically aims to study: (1) the productivity distribution of authors, using cross-section type of approach; (2) the productivity distribution of various cohorts of authors having same duration of participation; (3) the relation between extent of the duration of participation and the productivity distribution of authors, considered in cohorts; and (4) the frequency distribution of the speed of publications by authors.

Chapter 6: Findings and conclusions

The major findings from different chapters are described. The conclusions arrived are presented.

References

1. Callon, Michel, Law, John and Rip, Arie. Quantitative scientometrics. In. Callon, Michel, Law, John and Rip, Arie, eds. *Mapping the Dynamics of Science and Technology*. London; The MacMillan Press, 1986, p.103.

2. Popper, Karl. *The Logic of Scientific Discovery*. London; Hutchinson, 1959.

3. Lakatos, Imre. Falsification and methodology of scientific research program. In. Lakatos, Imre and Musgrave, A., eds. *Criticism and Growth of Knowledge*. Cambridge University Press, 1970, 91-195.

4. Toulmin, S.E. *Human Understanding. Part 1*. Oxford University Press, 1972.

5. Hesse, Mary B. *The Structure of Scientific Inference*. London; MacMillan, 1974.

6. Collins, H.M. The seven sexes: A study in the sociology of a phenomenon or the replication of experiments in physics. *Sociology* 9(1975): 205-224.

7. MacKenzie, Donald. Statistical theory and social interests: A case study. *Social Studies of Science* 8(1978): 35-83.

8. Mullins, N.C. A model for the development of a scientific speciality: The phage group and the origins of molecular biology. *Minerva* 10(1972): 51-82.

9. Latour, Bruno and Woolgar, Steven. *Laboratory Life: The Social Construction of Scientific Facts*. London; Sage, 1979.

10. Kuhn, Thomas S. *The Structure of Scientific Revolution*. Chicago; Chicago University Press, 1970.

11. Shapin, Steven. The politics of observation: Cerebral anatomy and social interests in the Edinburg phrenology disputes. In. Wallis, Roy G., ed. *On the Margins of Science: The Social Construction of Rejected Knowledge.* University of Keele. 1979, 139-178 (Sociological Review Monograph)

12. Thackray, Arnold. *John Dalton: Critical Assessments of His Life and Science.* Cambridge, Mass.; Harvard University Press, 1972.

13. Weingart, Peter. The scientific power elite: A chimera. The deinstitutionalization and politicization of science. In. Elias, Norbert, Martins, Hermino, and Whitley, Richard M., eds. *Scientific Establishments and Hierarchies.* Dordrecht and Boston; Reidel, 1982, 71-87 (Sociology of Science Yearbook, Volume 6)

14. Rip, Arie. A cognitive approach to science policy. *Research Policy* 10(1981): 294-311.

15. Bohme, Gernot, Daele, Wolfgang van den , Hohlfeld, Rainer and Schafer, Wolf. *Finalization in Science: The Social Orientation of Scientific Progress.* Dordrecht, Boston and Lancaster; Reidel, 1983.

16. Marshakova-Shaikevich, I.V. Bibliometrics as a research technique in epistemology and philosophy of science. *International Forum on Information and Documentation* 18(October 1993): 3-9

17. Introduction to scientometric studies of science. In. Tijssen, Robert J.W. *Cartography of Science.* Leiden University; DSWO Press, 1992.

18. Callon, Michel, Law, John and Rip. Arie. How to study the force of science. In. Callon, Michel, Law, John and Rip. Arie, eds. *Mapping the Dynamics of Science and Technology.* London; The MacMillan Press, 1986.

19. Rappa, Michael A. and Debackere, Koenraad. Monitoring progress in R&D communities. In. Weingart, Peter et al, eds. *Representations of Science and Technology.* Leiden University; DSWO Press, 1992, 253-265.

20. Price, D. de Solla. Network of scientific papers. *Science* 149(1965): 510-515.

21. Small, H. Co-citation in the scientific literature: A new measure of the relationship between publications. *Journal of the American Society for Information Science* 24(1973): 265-269.

22. Callon, M., Courtial, J.-P., Turner, W.A. and Bauin, S. From translations to problematic networks: An introduction to co-word analysis. *Social Science Information* 22(1983): 191-235.

23. Van Raan, A.F.J. and Peters, H.P.F. Dynamics of a scientific field analyzed by co-subfield structures. *Scientometrics* 15(1989): 607-620.

24. Felsenstein, Joseph. *Bibliography of Theoretical Population Genetics.* Pennsylvania; Dowden, Hutchinson & Ross, Inc., 1981, 866p

25. Lewontin, R.C. Population genetics. *Annual Review of Genetics* 1(1967): 37-70; 6(1973): 1-16; 7(1974): 1-17; 19(1985): 81-102.

26. Crow, James F. Population genetics history: A personal view. *Annual Review of Genetics* 21(1987): 1-22.