Chapter 3

Research Methods and Techniques used in the Study

CHAPTER 3

RESEARCH METHODS AND TECHNIQUES

USED IN THE STUDY

Last decade has seen increasing demand and supply of research indicators to evaluate the research activities at institutional and national levels in all subject fields. Employers, policy makers, research managers have become very much interested in these indicators so as to get better idea about the return on investment.

All significant compilations of research indicators rely heavily on publication and citation data which is the raw material for bibliometric analysis. This does not mean that this is the only aspect of research output – patents, organization of meetings and lectures, guiding the Ph D students, social recognition, international awards, and editorial activities – related to the productivity of research also play an important role in overall assessment of institutes or individuals. Bibliometric analysis is thus only a partial indication but it points out one of the essential outcomes of research and is therefore worth using.

All over the world, bibliometric indicators are gaining increasing attention in national and international evaluations and research policy discussions. The basis for developing valid indicators is reliable data, good methods and advanced tools for analysis.

To arrive at an appropriate method for the present study, the researcher made a detailed study of the research methods/strategies commonly used.

Robson (2002) has divided the type of research studies by research purpose as well as by methods used. a) Research type by research purpose – Exploratory, Descriptive and Explanatory b) Research type by research method used – Historical, Comparative, Experimental, Case study, Survey and Archival.

It must be emphasized that no research method is inherently superior or inferior to any other. What is most important is not the label of the research method but whether it will enable one to answer a particular question and meet the objectives of the study. Each of the research methods can be used for exploratory, descriptive or explanatory research. It should be remembered that these methods are not mutually exclusive. For example, it is quite possible to use the survey method as a part of a case study.

The present study is a bibliometric study of one organization. Universally accepted definition does not exist for the term bibliometrics. Alan Pritchard (1969) defined bibliometrics as "the application of mathematics and statistical method to books and other media of communication". Bibliometrics is thus a measuring technique by which inter-connected aspects of written communication can be quantified. In the same year Robert A Fairthorne published a classic article "Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for bibliometric description and prediction" in which he used the term' bibliometric' and also acknowledged that Pritchard was the donor of this term.

According to Lancaster (1991) the tools used in bibliometric studies are : i) citation and reference analysis ii) document and content analysis iii) user studies and iv) circulation statistics. The present study uses two of these tools (content analysis and citation analysis) to fulfil its objectives, as mentioned in chapter 1.

Scope of the study

The present bibliometric study aims to measure the productivity of Physical Research Laboratory (PRL) scientists during a 10 year period (1997-2006) using the data of papers published in journals, conference proceedings and invited talks delivered. Thrust areas of research at PRL during this period have been found using content analysis of articles published in journals and allotting keywords to each of them. The study also tries to determine the usage of library collection by employing citation analysis to the bibliographies of the theses submitted by the Ph D students of PRL.

Physical Research Laboratory (PRL)

Known as the *cradle* of Space Sciences in India, the Physical Research Laboratory, Ahmedabad owes its existence to Dr Vikram A Sarabhai due to his deep interest in scientific research, his initiative and his outstanding powers of organization and management. It was founded following an agreement between the Ahmedabad Education Society and the Karmakshetra Educational Foundation in November 1947.

The initial focus was research on cosmic rays and the properties of the upper atmosphere. As a unit of the Department of Space, Government of India, PRL carries out fundamental research in select areas of Experimental and Theoretical Physics, Space and Atmospheric Sciences, Astronomy & Astrophysics and Planetary & Geosciences.

PRL has, from time to time also engaged itself in applied research problems relevant to the country's needs, particularly in the field of Space Science. PRL's vision is to undertake world-class research projects and make a mark for itself as a leading research institution on the global map. In addition to taking up key scientific projects for national development, PRL's mission is also to popularize science amongst the youth. There also exists a strong interaction with the neighboring educational institutions. At the various universities in Gujarat some of the advanced courses in physics have been taught many a times by members of the PRL.

In addition to taking up key scientific projects for national development, human resource development in several areas of above mentioned subject areas is also one of the priorities for PRL. There are about 140 scientists (60 are academic faculty and remaining are technical faculty and Post doctoral fellows) carrying out research in PRL. It has been offering the doctoral programme in various physics related fields since its inception. Accrediting universities with which it has signed the Memorandum of Understanding are Gujarat University, Nirma University, M. S. University of Baroda and Mohanlal Sukhadia University, Udaipur. Up till now 316 doctoral theses have been submitted by the PRL students. Every year about 15 students join for the Ph. D. program.

Period: 1997-2006

The period of study has been taken from 1997-2006. The landscape of scholarly communication witnessed a sea change during this period from print to electronic medium due to the Internet. Developed countries like USA, UK, Japan, and Germany were the first to adopt this change. The internet made it possible to disseminate the latest information to the scientists and students very quickly. Soon, publishers saw a huge opportunity and started offering the secondary databases and the journals in electronic medium. The electronic delivery of journals resulted in elimination of paper, storage and transportation costs and the ability to handle complex data, tables, moving pictures, sound, images and video clips. In addition, unlike sequential design of printed papers, web technology made it possible for the publishers to give interactive hyperlinks to related sources. The growth of the Internet witnessed emergence of several e-journals that were launched only for Internet without a printed counterpart. However, as the technology and popularity of Internet grew, several mainstream journals primarily available for print subscription also started appearing on the web. By 2001-02, the Indian publishers too had started providing the e-access to the print journals.

Keeping in mind this paradigm shift in scholarly communication, the scope of the present study is limited to the period 1997 to 2006. As this period witnessed the maximum flux, the period has been divided into two time slabs: pre-ejournals period (1997-2001) and ejournals period (2002-2006) for citation analysis of bibliographies of thesis submitted by PRL students.

Operational Definitions

Research word is composed of two words 're' and 'search' which means to search again. The concept of research took the shape of social and scientific investigation during the medieval times and developed into a full body of intellectual exercise only in the modern age. Research is a systematic investigation designed to develop or contribute to generalize

the observed phenomenon. Whenever traditional theory is found lacking in explaining the existing phenomena and a novel situation is faced, research originates. Research rejects either old theories or modifies them or suggests new theories. Thus research is a matter of raising a question and then trying to get an answer. Adding new knowledge to the existing corpus is the obvious function of any research. It inculcates scientific and inductive thinking and it promotes the development of rational thought process. It enables finding of solutions to problems and to resolving conflict in society. In this way it promotes progress of the society. Formal definition of the term 'research undertaken' used in the present study, is given below.

The Webster's International Dictionary (1986) defines research as "a careful, critical enquiry or examination in seeking facts or principles, diligent investigation in order to ascertain something."

International Encyclopaedia of Social Sciences (1968) defines research as "the manipulation of things, concepts or symbols for purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids construction of theory or in the practice of an art".

For the purpose of the present study, the term 'research undertaken' is considered as the research output of PRL comprising of the collection of research articles published in journals and conference proceedings, invited talks delivered and doctoral theses of PRL students. Other research outputs like technical notes, project reports have not been included for the study.

Data Collection for the study

Data for the study (papers published in journals and conference proceedings and invited talks delivered) was collected from the Annual Reports of PRL from 1997-98 to 2006-07. Data collection for the study was done during 2007-2010. For papers in journals, the

record consisted of names of the authors, name of the division, name of the journal, whether it is national or international, whether it has single author, double author or multi authors, whether the collaboration is international, national or domestic and the year of publication. For conference proceedings, data consisted of names of authors, name of the division, whether it has single author, double author or multi author, whether the conference was national or international, whether the collaboration was international, national or domestic and year of the conference. For Invited talks, the record consisted of name of the speaker, location of the talk – India or abroad and the year. Thus, such record was made for all the three components of the research output for all the years. This data was used to find the publication pattern of PRL scientists. To identify the active areas of research content analysis was done of the papers published in journals.

Content Analysis

Content analysis is a method for summarizing any form of content by counting various aspects of the content. This enables a more objective evaluation than comparing content based on the impressions. For example, an impressionistic summary of a TV program is not content analysis. Nor is a book review: it's an evaluation. The results of content analysis are numbers and percentages. Though it may seem crude and simplistic, the counting serves two purposes: to remove much of the subjectivity from summaries and to simplify the detection of trends. Thus content analysis requires extreme thoroughness. The content that is analysed can be in any form to begin with, but is often converted into written words before it is analysed. The original source can be printed publications, broadcast programs, other recordings, the internet, or live situations.

To identify the research trend in PRL during the study period, the researcher has carried out the document and content analysis of the research articles published in journals by providing the keywords to each article. The keywords were then used, to allot a PACS number (Physics and Astronomy Classification Scheme) to each article.

PACS is a hierarchical subject classification scheme designed to classify and categorize the literature of physics and astronomy. PACS provides an essential tool for classification and efficient retrieval of literature in physics and related fields. PACS contains 10 broad subject categories subdivided into narrower categories. PACS also includes detailed schedule for acoustics, geophysics, nanoscale science and technology supplement and an alphabetical topical index with corresponding PACS codes (AIP, 2006).

The PACS codes of all the articles are added up and then grouped. This is done for each year. Data for all the years is added up to arrive at top 25 subject headings put together corresponding to the PACS numbers and are indicated as the thrust areas of research carried out by PRL.

Citation Analysis

Citation analysis is one of the important tools of bibliometric analyses of the scholarly literature for a deeper understanding of scholarly activity and performance. Citation analysis studies the citations provided at the end of any scholarly communication and is generally regarded as a valuable tool for determining the impact of scholarly works. It examines the frequency and patterns of citations in journal articles and books. It uses citations in scholarly works to establish links to other works or other researchers.

In today's world of ever escalating cost of serials, citation analysis is also being used to determine which titles to purchase and which ones to discontinue.

Methods of citation analysis are unobtrusive and can be highly reliable, as it does not require the feedback from the users by way of questionnaire or interview, rather the data is derived from the actual use made as is reflected in the documents already submitted or published.

By examining the resources used (cited), present study aims to better understand and manage the library resources. Citation analysis is used to study the bibliographies of the

doctoral theses submitted by the Ph. D students of PRL during 1997-2006. This was done to find the usage of different types of documents in the library collection, whether there has been increase of use of electronic resources, whether there is increase in the use of non-subscribed journals, how far Ph. D. students cite the research done in PRL and to identify the gaps in journal subscriptions.

Data Analysis for the study

The research output of PRL scientists during the period of 1997-2006 was 2518 units out of which 1318 were papers published in journals, 436 papers in conference proceedings and 764 were the invited talks delivered. These have been used for identifying the publication pattern and research trends in PRL. With in PRL as there are various divisions, an attempt is also made to study division wise output and most prolific researchers in each division.

During the period of study, 68 theses were submitted by the Ph. D. students. These studies yielded a total of 10,864 citations for which citation analysis was done.

Excel software was used to enter the records of each year. Each record consisted of the name of the author/s of the article published, name of the journal, double/multi/single author (D/M/S), international/national journal (JI/JN), international/national/domestic collaboration (CI/CN/CP), division of the author and year of publication. Keywords were given to each article after reading the abstract and introduction of the paper. Then PACS number was allotted to each article.

Each category of this data was counted and sorted in descending order of number of times it appears in each year. Each year's data was then combined and computed in similar manner. As for example, PRL authors published 3, 4, and 2 articles in the Astrophysical Journal Letters in 2004, 2005 and 2006 respectively. So in three years 9 papers were published in this journal. Similarly, the choice of this journal to publish one's research output would be added up for all 10 years of the study. Each category of data – name of

the journal, author, research division, PACS code, collaboration, etc. was counted for each year. This data was then merged for all years to arrive at most preferred journals, most prolific researchers and the thrust areas of research (using the PACS number schedule). The authorship pattern (D/M/S), collaboration pattern (CI/CN/CP), pattern of papers in chapter of a book and international/national journals (CB/JI/JN), international/national conference proceedings (CPI/CPN) and invited talks delivered (TI/TN) were identified by carrying out similar computation.

Similarly, the records have been entered for each bibliography of the collected theses consisting of the name of the journal cited, year of publication and whether the author of the cited article is from PRL. The documents used were then tagged for print or electronic format using the license agreement with the publisher as the reference source. The data was also categorized according to the type of resource – books, journal articles, proceedings of a conference, doctoral theses, etc. (B, J, P, R, Th, St, Ep) cited by the students. The journals cited by the students were tagged as subscribed or non-subscribed (SJ/NSJ) going through the holdings database of the journals of the library for that year. This data is also first computed for one year and then it is merged for all the years to arrive at most cited journals in each subject area.

All this data was collected and analysed to identify which subject areas showed quickest adoption to electronic medium, which type of documents are most cited, which journals are most preferred by doctoral students and to identify the gaps in the library collection by looking through the non-subscribed titles in the cited list of journals.

To simplify the data handling, each data element - double authored paper, division name, international journal, - etc. was coded. These are listed below and have also been provided under the List of Abbreviations on page iv.

I Content analysis of research publications

D	–	Double authored paper
M	–	Multiple authored paper
S	–	Single authored paper
CB	-	Chapter of a book
JI	–	International Journal
JN	–	National Journal
CI	–	International Collaboration
CN	–	National Collaboration
CP	–	PRL Collaboration
CPI	–	International Conference Proceeding
CPN	–	National Conference Proceeding
TI	–	International Invited Talk
TN	–	National Invited Talk

II) Citation analysis of bibliographies of theses

Type of Documents

B	–	Book
Ep	–	E-print
J	-	Journal
M	-	Monograph
P	–	Proceedings
R	–	Report
St	-	Standard
Th	-	Thesis

Type of Journals

SJ	–	Subscribed journal
NSJ	–	Non-subscribed journal

Divisions

AAD	-	Astronomy & Astrophysics Division
GSDN	-	Geosciences Division
PLANEX	-	Planetary Exploration
PSDN	-	Planetary Science Division
SO-PH	-	Solar Physics Division
SOXS	-	Solar X-ray Spectrometer
SPA-SC	-	Space & Atmospheric Science Division
THE-PH	-	Theoretical Physics Division

Limitations of the study

The researcher has not included the number of projects completed and students guided by faculty members during the 10 year period. As it was not possible to get budget data for each year, researcher could not compute the return on investment (ROI) of the research done at PRL. As the researcher does not have access to either of the citation databases – Web of Science or Scopus, researcher could not get the citation data of PRL as a whole institute as well as that of the individual scientists. Not only the productivity, but it would have been possible to determine the impact of PRL as well as of each scientist. This aspect about research measurement remains to be done.

References

Fairthorne, R. A. (1969). Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for bibliometrics description and prediction. *Journal of Documentation, 25*(4), 319-343.

Webster's Third New International Dictionary. (1986). Massachusetts : Merriam-Webster Inc.

International Encyclopedia of the Social Sciences, (2007). *Michigan : Macmillan Reference USA.*

Pritchard (1969). Statistical Bibliography or bibliometrics ? *Journal of Documentation, 25,* 348-349

Robson, C. (2002). *Real World Research, 2nd ed.* Oxford : Blackwell Publishing.