Jointly published by Elsevier Science Ltd, Oxford and Akadémiai Kiadó, Budapest

Scientometrics, Vol. 40, No. 3 (1997) 507–528

GROWTH OF RESEARCH LITERATURE IN SCIENTIFIC SPECIALITIES. A MODELLING PERSPECTIVE

B. M. GUPTA,* PRAVEEN SHARMA,* C. R. KARISIDDAPPA**

**Scientometrics and Informetrics Group, National Institute of Science, Technology and Development Studies,*
Dr K. S. Krisnnan Marg, New Delhi 110012 (India)
***Department of Library & Information Science, Karnataka University, Dharwad 580003 (India)*

(Received August 30, 1997)

The paper discusses the application of three well known diffusion models and their modified versions to the growth of publication data in four selected fields of S&T. It is observed that all the three models in their modified versions generally improve their performance in terms of parameter values, fit statistics, and graphical fit to the data. The most appropriate model is generally seen to be the modified exponential-logistic model.

1. Growth of knowledge

The understanding of the process of growth of knowledge in research specialities and its modelling has challenged bibliometricians and sociologists for long. Over the years, some literature has appeared in this area. *Gilbert*[1] has reviewed the existing literature on the indicators of growth of knowledge in scientific specialities and lists many ways of measuring it, noting their strengths and limitations and commenting on their use.

There are two approaches that have normally been considered in understanding knowledge growth: Qualitative and quantitative. Qualitative approach suggests structural or descriptive models of knowledge growth, while descriptive models use social phenomenon to explain diffusion and creation of knowledge. Quantitative approach is a more recent phenomenon, and have relied on summarisation statistics to describe observed behaviour, while others apply growth and technology diffusion models and bibliometric/scientometric techniques.

The growth of scientific knowledge generally takes the form of logistic curve. The successive phases of knowledge growth represented by logistic growth curve are: (a) a

0138–9130/97/US $ 15.00

preliminary period of growth in which absolute increments are small, although the rate of increase is large but steadily decreases; (b) a period of exponential growth when the number of publications in the field doubles at regular intervals as a result of a constant rate of growth that produces increasing amount of absolute growth; (c) a period when the rate of growth declines but the annual increment remains absolutely constant; and (d) a final period when both the rate of increase and the absolute increase decline and eventually approaches zero.[2,3]

The fact that knowledge in scientific specialities exhibits logistic growth indicates that scientific growth is a social process as well as cognitive one. The growth of scientific knowledge is a kind of diffusion process in which ideas are transmitted from person to person, similar to the diffusion of innovations which is also shown to follow logistic growth, curve, according to *Rogers*.[4] In such studies, the exponential increase in the number of adopters has been perceived as a social influence process. When members of a social system are in communication with one another, a kind of "contagion" effect occurs in which individuals in a social system who have adopted an innovation influence those who have not yet adopted it. One can also say that the growth of scientific publications can also be interpreted as a contagions process in which early adopters (authors) of ideas influence later adopters (authors), which, in turn, creates an exponential growth of research papers.

Contagion refers to a specific process. For example, when we refer to a disease as contagious we mean that the disease spreads through personal contact. If a fad is contagious, then we imply that individuals will take up a fad if they see someone else doing it. In the diffusion of innovation or scientific knowledge, contagion refers to individuals monitoring others and imitate their behaviour to adopt or not to adopt. Contagion can be called a social process of how individuals form opinions and eventually adopt or not to adopt an innovation or idea. Contagion can be called a lens through which individuals monitor the behaviour of others, and it leads to influence in adoption behaviour. Contagion can occur via cohesion (direct ties), structural equivalence, social proximity, popularity (centrality), or it can be system wide (using system adoption level as a measure). These four contagion processes define those other individuals in the social system who influence an individual's behaviour.[5]

2. Diffusion models

Recent years have seen increasing interest in the diffusion process, the process by which an innovation spreads and grows. The interest is manifested by the number of publications and research disciplines concerned with this phenomena. By 1968, nearly

1100 publications were available on the diffusion of new ideas, practices, technologies, and products; the number up to 1971 was more than 1500, an increase of more than 36%.[6] There are, at present, more than 18 research disciplines in social sciences sharing this concept but looking at its different aspects and applications.[7]

The rich and multidisciplinary literature on the diffusion of innovations, ideas, practices and technologies, reflects in general, two broad and distinct approaches. The first approach focuses on spatial aspects of the diffusion process and an examination of socio-economic factors that influences it. Geographers, sociologists, and development planners seem to be interested in this approach. In the second approach the main focus is on the study of the time pattern of the spread of innovation at a macro-level. This approach has been adopted primarily by technology planners, market analysists, and industrial researchers.[8]

Although a wide variety of innovations and diffusion processes have been investigated, one research finding keeps on recurring: if the cumulative adoption time path or temporal pattern of a diffusion process is plotted, the resulting distribution can generally be described as taking the form of an *S*-shaped (sigmoid) curve. Much of the early research on diffusion processes focused on describing observed diffused patterns in terms of prespecified trend or distribution functions. For example, cumulative normal, Gompertz, and logistic distribution functions have all been used to model diffusion processes because each gives rise to a *S*-shaped curve.[24] However, because any unimodel distribution function will generate an *S*-shaped curve, it is not often possible to empirically determine which of the several competing trends of distribution functions best describe a given diffusion curve.[25]

The existing *S*-shaped technology diffusion models can be broadly classified into two classes:(a) models that consider diffusion in terms of adopters and non-adopters of technology; (b) models that consider diffusion as a process of substitution, an existing technology being replaced by a new technology. Our further discussion in this paper will be based on the first kind of model.

Different hypotheses and interpretations, as available in the literature, are presented below which explain the *S*-shaped notion of a diffusion curve:

(a) *Economic Approach*: *Mansfield*[9] considered diffusion of technology in terms of economic advantage, investments and uncertainties associated with the introduction of new technologies. Similarly, *Griliches*,[10] *Robinson* and *Lakhani*[11] and *Brown*[12] proposed a supply and demand rationale in a diffusion explanation.

(b) *Learning Approaches*: *Cassetti* and *Semple*[13] and *Sahal*[14] employed a learning perspective in explaining diffusion patterns. This approach incorporates concepts of learning process at different stages of technology development (invention, innovation and motivation) as well as cognitive processes involved in adoption of new products or processes (word of mouth, imitation,etc.)

(c) *Self-Organising Systems Approach*: In this approach, innovative systems are taken as inherently 'untidy' systems and the processes of invention, innovation, and economic development[15] are expressed in terms of evolution of a self-organizing system.[16]

The initial work of fitting data into *S*-shaped model was carried out by *Mansfield*[9] in historical technological substitution data on rail roads, coal, steel, and breweries. *Blackman*[17,18] studied the innovation dynamics in aircraft jet engine market and in electric utility and automotive sectors. Other applications include the study of industrial technology by *Nevers*,[19] medical innovations by *Easingwood et al.*,[20] energy efficient innovations by *Teotia* and *Raju*,[21] telecommunication innovations by *Bewley* and *Fiebig*,[22] and agricultural innovations by *McGowan*.[23]

3. Developing the basic diffusion model

To apply and interpret results of any diffusion model, one must first understand its conceptual as well as mathematical foundation. Such knowledge can be obtained by first establishing the basic or fundamental diffusion model and then examining its major components and assumptions underlying the model formulation. *Rogers* and *Shoemaker*[6] have defined diffusion as the process by which innovations spread among the members of a social system. The diffusion process is to be distinguished from the adoption process which refers to the sequences of stages through which the adoption unit progresses from first awareness of the innovation to final acceptance. Implicit in the definition of diffusion process is also the assumption that diffusion of any innovation is not a result of the physical diffusion of innovation from a region of high density to a region of low density of adoption, but instead is the movement of the innovation from adopter to non-adopter. To begin with the development of the model, the following notion is introduced:[26]

Let $n(t)$ be proportion of adopters at time t; $N(t)$ be cumulative number of adopters at time t;Then at any time T

$$N(T) = n(t)\ dt, \quad (1)$$

or at any time t,

$$n(t) = dN(t)/dt, \quad (2)$$

and $n(t)$ has a maximum when,

$$dn(t^*)/dt = 0 \text{ at } t=t^* \quad (3)$$

provided we assume (a) $n(t)$ and $N(t)$ are continuous functions and their derivatives exist at all points; and (b) $n(t)$ is a unimodel function (for a unique maximum).

In order to derive an explicit formulation for $n(t)$ and $N(t)$, it is necessary to provide a structure for the rate in Eq.(2).

Suppose the rate of diffusion at any time t is directly proportion to the proportional of potential adopters available at that time; in other words, as the cumulative proportion of adopters approaches its ceiling say M, the rate of diffusion decreases proportionally. Mathematically, such a situation may be written as:

$$n(t) = dN(t)/dt \text{ is proportional to } (M-n(t)) \;..... \qquad (4)$$

The statement in Eq. (4) implies that there exists some proportionality measure relating to the diffusion rate and the potential proportion of adopters.

If $g(t)$ is used to represent this "constant" of proportionality, the resultant Eq.(4) may be rewritten as:

$$n(t) = dN(t)/dt = g(t)[M-n(t)] \;..... \qquad (5)$$

Equation (5) is the rate equation for the diffusion process. The rate of diffusion or proportion of adoptions at time t is controlled by the proportionality "constant" $g(t)$, the value of which depends upon the specific innovation, social system in which it is diffused and the channels and change agents used to diffuse it.

$g(t) = f($ innovation, social systems, channels, and change agents)

Hence, $g(t)$ is "constant" only when the above characteristics of diffusion process are delineated.

Broadly speaking the two channels through which potential adopters are converted into adopters category are: (a) internal influence channel, in which it is the result of internal interaction between those who have adopted the product and those who are yet to adopt; (b) external influence channel, in which the communication media and other external factors play a significant part; (c) mixed influence channel, in which a mix of the internal influence and external influence factors play a significant role.

In the absence of empirical and mathematical functions for $g(t)$, two basic approaches generally used to represent $g(t)$ are: (a) $g(t)$ as a function of time; and (b) $g(t)$ as a function of the number of previous adopters. In order to demonstrate the relationship between the existing models, the latter approach will be the one followed here. Then, specifically, $g(t)$ can be expressed as a function of $n(t)$ such that,

$$g(t) = p + q\, n(t) + r\, n(t)^2 + \;..... \qquad (6)$$

However, for reasons such as convenience, a desire to retain analytical parasimoney, and facilitation of interpretation and parameter estimation, $g(t)$ has been typically formulated as either,

$$g(t) = p, \tag{7}$$
$$g(t) = q\,N(t), \tag{8}$$
$$g(t) = [p + q\,N(t)] \tag{9}$$

where p and q are treated as model coefficients or parameters.

Using these values of $g(t)$, the fundamental diffusion model can be expressed and referred as:

$dN(t)/dt = p\,(M\text{-}n(t))$ referred as external influence model

$dN(t)/dt = q\,N(t)\,[M\text{-}n(t)]$ referred as internal influence model

$dN(t)/dt = (p + q\,N(t))\,[M\text{-}n(t)]$ referred as mixed influence model

The three mathematical functions that represent these processes are presented in Table 1.

Table 1
Mathematical models of the diffusion of innovations

	Internal	External	Mixed
Cumulative function	$\dfrac{M}{1+\dfrac{N\text{ - }N_0}{N_0}\mathrm{e}^{-qn(t-t_0)}}$	$M(1-\mathrm{e}^{-pt})$	$\dfrac{N\dfrac{p(N-N_0)}{p+qN_0}\mathrm{e}^{[-(p+qN)(t-t_0)]}}{1+\dfrac{q(N\text{ - }N_0)}{p+qN_0}\mathrm{e}^{[-(p+qN)(t-t_0)]}}$
Derivative	$qN(t)[N\text{-}N(t)]$	$p[N\text{-}N(t)]$	$(p+qN(t)[N\text{-}N(t)]$
Diffusion of	Adoption	Awareness	Adoption and awareness
Type of communication	Interpersonal	Mass media	Interpersonal and mass media

Note: $N_0 = N(t = t_0)$, is the number of initial adopters (adopters at time t_0); M is the population size; p and q are model parameters.

4. Discussion on models

Bass model

Bass[27] described diffusion of technology in terms of parameters p and q, which he called coefficients of innovation and imitation, respectively. Bass model is mathematically expressed as:

$$dN(t)/dt = [(p + q(n(t)/M))]\,(M\text{-}n(t)) \;.... \tag{10}$$

Bass model implies that external influence relates to innovative factors and internal influence to imitative factors. The term, coefficient of 'innovation' or p used by *Bass* therefore implies influence originating from the innovation itself, i.e. external to the use of adopter. The internal coefficient, q on the other hand, represents horizontal communication. Such influences are based on personal contacts, unstructured or informal channels of communication. Here $[p + q\ (n(t)/M)]$ may also be referred as conversion factor, a fraction by which any time t potential adopters are converted to the adopter's category.

Mansfield model

The model that captures the dynamics of the internal influence is the well known Mansfield model which can be mathematically expressed as:

$$dN/dt = q\ [N(t)/M)\ (M-N(t)]\ .... \qquad (11)$$

Where M is considered as the total number of possible adopters, $N(t)$ is the number of adopters at time t, and q is the coefficient of internal influence. Here $q[N(t)/M]$ may be referred as conversion factor, a fraction by which any time t potential adopters are converted to adopters category.

Exponential logistic model

A useful model expressed in the form of exponential-logistic differential equation as suggested by *Sharma et al.*[28] is presented here. The model is mathematically expressed as:

$$N(t+1) = N(t)\ e^{b-cN(t)} \qquad (12)$$

where $N(t)$ represents the total number of adopters at a particular time step and b and c are real parameters which are constant. The number of adopters at a time step $(t+1)$ are completely determined by its number at time step t. The coefficients b and c would be both positive if the system has a saturation. The ratio b/c will provide an estimate of the total market M. The coefficient b is always positive for growth while the parameter c could be negative or very small if the saturation is not indicated by the data.

5. Modification in the existing models

The three models described above and their variants have been used in the literature for analysing the diffusion several technologies, products and ideas. *Sharma* and

Bhargava[29] re-examined the role of influencing factors in the existing models by posing the following question. The question relates to the homogeneity of the influence coefficient *[g(t)]* over the group of adopters, who adopt the product at various temporal stages.

To explain this we recall that any time *t*, the number of adopters *N(t)* is given by

$$N(t) = n(t)+n(t-1)+n(t-2)+\dots\ n(2)+n(1) \quad (13)$$

In the Mansfield model, the conversion factor is taken to be *qN(t)/N*, i.e. all categories of adopters, *n(1), n(2), ... n(t)* are assumed to be equally effective. Thus, influence represented by various terms as *n(1)[M-n(t)], n(2)[M-N(t)]..........,n(t)[M-n(t)]* has the identical weightage. This is reflected by the constancy of internal influence coefficient and implies that there is complete and uniform mixing of members of the social system - prior and potential adopters. In their paper, *Sharma* and *Bhargava*[29] ask "why should one give equal weightage to the opinion of those who had adopted a product in the recent past and to those who had adopted it earlier? They, therefore, propose changes in the existing models presented in earlier section, and assume that the adopter categories *n(t)*, *n(t-1) ... n(1)* influence the potential adopters with varying degrees of effectiveness or weight. They also assume that the influence of existing adopters on the likely potential adopters decreases with time.

To incorporate the varying effectiveness of various year-wise categories of adopters: *n(t), n(t-1), ...n(1)*, *Sharma* and *Bhargava*[29–30] have suggested an improved methodology in which they have replaced *N(t)* by *NS*, where *NS* is given as:

$$NS = \sum_{i=0}^{i=t-1} n(t-i)w^i \quad (14)$$

$$NS = [n(t) + wn(t-1) + w^2 n(t-2) + \dots + w^{n-1}\ n(1)]$$

Here, *w* is a weight factor, which is to be calculated by trial and error method. *Sharma* and *Bhargava* tried different values of *w*, but finally they found that when *w*= 1/2, the results were found to be more encouraging. If *w*=1/2, we can then say that effectiveness of influencing factors decreased by 50 per cent each year.

Using this innovative method, we have found that the models become quiet effective in capturing not only the growth of research publications but also their yearly fluctuations. The overall performances of the three models have also improved in terms of parameter values, fit statistics, and graphical fit to the growth data. In the light of the above suggested modification, the various model equations of Bass, Mansfield, and exponential - logistic model can be rewritten as follows:

Bass model equation is rewritten in its modified form as:

$$dN(t)/dt = [p + q(NS/M) (M-N(t)] \tag{15}$$

Mansfield model equation is rewritten in its modified form as:

$$dN(t)/dt = q(NS/M) (M-N(t)) \tag{16}$$

Exponential-logistic model is rewritten in its modified form as:

$$dN(t)/dt = NS\, e^{b-cN} - NS \tag{17}$$

6. Growth of research output in scientific specialities

We have seen in section 1 that the growth of scientific knowledge in a scientific speciality is similar to diffusion of innovations and follows the logistic growth curve. Faced with the problems of defining and measuring knowledge, most studies measure the growth of literature (or number of publications). In quantitative studies, growth in the number of publications is therefore taken as a measure or operational definition of growth of knowledge. As a result we can apply some of the diffusion models to the growth of scientific papers in a scientific speciality.

A few researchers in the past have studied growth of research papers in a scientific speciality, specially from the point of view of modelling their growth data and fitting the appropriate models. *Sterman*[31] studied and modelled growth of a paradigm or metaphor using systems dynamics approach. He indicated that the number of puzzles solved within a "paradigm" or "metaphor" when plotted against time gives a logistic *S* curve. This *S* curve arises essentially due to interplay between the complexities of puzzles that a metaphor can solve and the confidence level amongst the practitioners. Using, Sterman ideas as a base, *Jain* and *Garg*[32] have studied, for the first time, growth of research papers in laser research from the modelling perspective. They have proposed a model, which is found to be similar to Bass model, to interpret the past and future growth trends of literature output on the basis of characteristic features of scientific specialities. On the similar analogy, *Garg et al.* studied the growth of publications on world solar power research using Bass model and *Gupta et al.*[34] on world theoretical population genetics research, using a modified exponential logistic model.

7. Methodology and database

The simple and best method to collect growth data on a scientific speciality is either through computerised databases on various subject fields (available in CD-ROM media and online) or printed abstracting services and bibliographies. There are, however, a few databases, abstracting services, and printed bibliographies which have a long history of coverage of the developments in a subject field. For modelling the growth data, it is necessary to have long time series data on the growth of publications in a subject field, depicting their different stages of the development. Keeping this viewpoint in mind, we have selected the following four subject fields for analysis:

(1) Physics research output as reflected in *Physics Abstracts* from 1907–1994 (referred as PHY).
(2) Chemical sciences research output as reflected in *Chemical Abstracts* for the period 1901–1994 (referred as CHEM).
(3) Electrical and electronic engineering research output as reflected in *Electrical and Electronic Abstracts* for the period 1907–1994 (referred as ENG).
(4) Theoretical population genetics research output as reflected in the "*Bibliography of Theoretical Population Genetics Research*", compiled by *Felsenteen*[35] for the period 1907–1980 (referred as GEN).

In order to understand the literature growth characteristics in individual subject fields, it would be useful to know the initial number of publications (N_0), and data time span and cumulative number of publications in the period covered.

Sub.	N_0	N	Data span
PHY	2132	**174237**	88 years
CHEM	11847	15517854	88 years
ENG	1474	187725	88 years
GEN	1	7663	74 years

Data for each subject field were cumulated before modelling exercise was undertaken. Although, many diffusion models, as applied across magnitude of problems and disciplines are available in the literature, we have selected the following models for analysing data in the four choosen subject fields.

(1) Bass Model (Model 1);
(2) Modified Bass Model (Model 2);
(3) Mansfield Model (Model 3);
(4) Modified Mansfield Model (Model 4);
(5) Exponential - Logistic Model (Model 5);
(6) Modified Exponential - Logistic Model (Model 6).

7. Results and interpretations

Bass model

This is a mixed influence model and has been tried for modelling the growth of research output in a speciality. This model can be mathematically expressed as:

$$dN(t)/dt = [p + q(N(t)/M] [M-N(t)]$$

This model was applied to the growth of literature in the four subject fields mentioned above. The results and parameter values obtained from this model are given below:

Subject	Value of parameters		
	p	*q*	*M*(Thousands)
PHY	.31d-3(0)	.094(.003)	6602.99(257.3)
CHEM	-.181d-3(0)	.065(.002)	35358.63(2327.1)
ENG	-.289d-3(0)	.008(.003)	4008.19(380.1)
GEN	-.338d-3(0)	.093(.004)	33.57(6.6)

Note: *d* stands for power of 10 as base.

The values of R^2 and F which indicate the fit of the model are given below.

Subject	R^2	F
PHY	0.984	2768.26
CHEM	0.977	2355.70
ENG	0.984	2834.08
GEN	0.979	1713.05

The Bass model expects the value of parameters p and q to be positive and less than 1. We have observed in the analysis that the value of parameters p and q is less than 1, but the value of parameter p is found to be negative in three subjects, except physics. It is also observed that the value of M which is an indicator of the highest number of papers a subject can have (saturation limit), is small. The value of R^2 in the four subject fields ranges from 0.977 to 0.984. Figure 1 presents the data and model fit for Bass model.

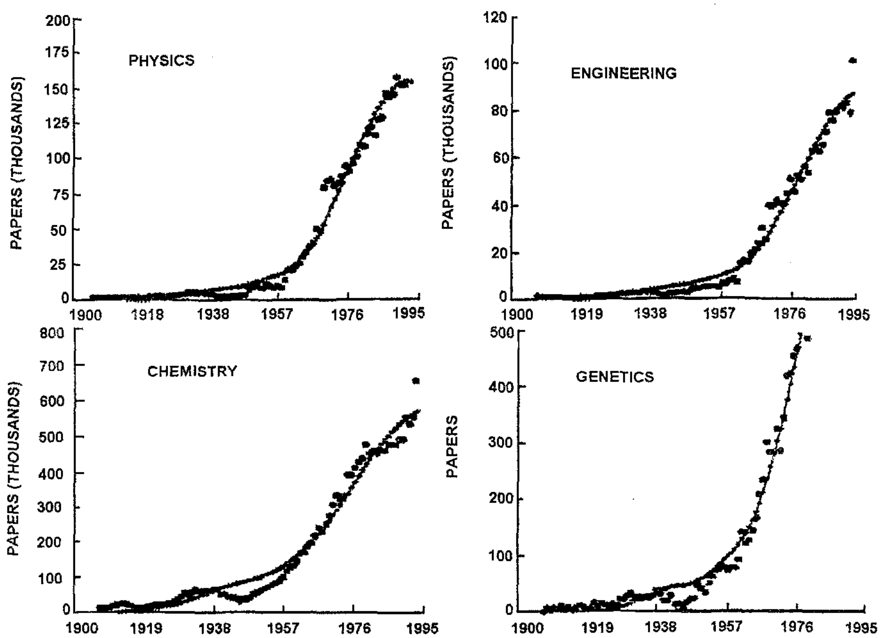

Fig. 1. Data and model fit for Bass model

Modified bass model

The equation for modified Bass model is written as:

$$dN/dt = [p + q(NS/M) [M-N]$$

When modified Bass model was applied to the growth of literature in the four subject fields, the result and parameter values obtained are given below:

Subject	Value of parameters		
	p	q	M(Thousands)
PHY	.202d-5(0)	.538(.007)	63096.42(18967.12)
CHEM	.910d-6(0)	.524(.007)	701044.87(513780)
ENG	.143d-5(0)	.533(.008)	55603.82(31265.09)
GEN	-.289d-4(0)	.578(.011)	68.790(12.764)

The value of R^2 and F which indicate the fit of the model are given below in the table.

Subject	R^2	F
PHY	0.998	19083.17
CHEM	0.998	28616.76
ENG	0.997	16283.63
GEN	0.997	12986.47

This parameter values obtained from this model indicate that p and q are less than 1. The value of p is found to be positive in first three subjects, giving negative value in theoretical population genetics. The value of R^2 and F has improved and now ranges between 0.997 and 0.998. The value of M obtained from this model is observed to be larger than the value obtained in the previous model. Figure 2 gives the model fit for the data.

Mansfield model

The Mansfield model is a widely used model and is basically an internal influence diffusion model. The model is mathematically expressed as:

$$dN/dt = q\ (N/M)\ [M-N]$$

This model, when applied to four subject fields under consideration, gave the following results and parameter values.

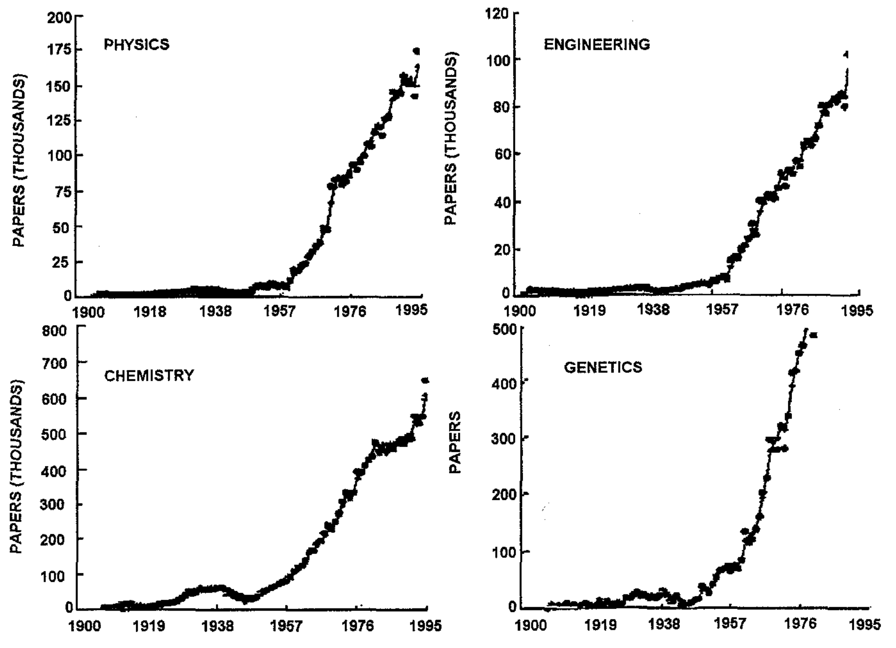

Fig. 2. Data and model fit for modified Bass model

Subject	Value of parameters	
	q	M(Thousands)
PHY	0.091(0.002)	6831(269)
CHEM	0.063(0.002)	37068(2297)
ENG	0.084(0.002)	4214(217)
GEN	0.084(0.003)	52(17)

The values of R^2 and F which indicate the goodness of fit of the model are given below:

Subject	R^2	F
PHY	0.983	3997.772
CHEM	0.976	3505.204
ENG	0.983	4072.142
GEN	0.977	2299.989

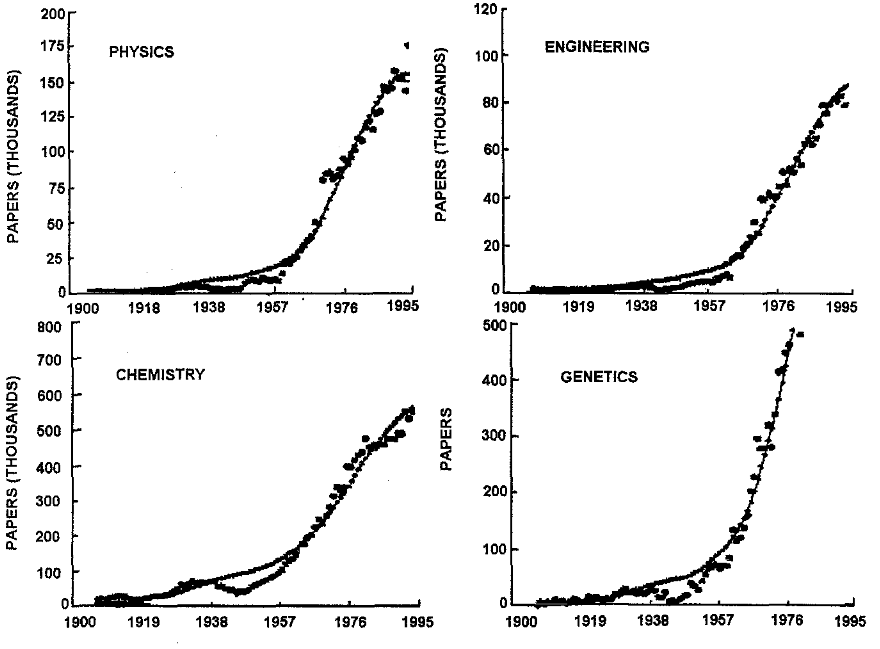

Fig. 3. Data and model fit for Mansfield model

The model expects the value of parameter q to be positive and M indicates the saturation limit on the total number of papers expected in individual subject field. The value of parameter q is found to be positive in all the four subject fields. The value of R^2 obtained is found to range between 0.976 to 0.983 in the four subject fields. Figure 3 presents graphical fit of the data and estimated values for the model. From the graphical fit, it is observed that the model is not able to capture the fluctuation in the growth data.

Modified Mansfield model

This model is mathematically expressed as:

$$dN/dt = q\ (NS/M)\ [M-N]$$

The model when applied to the growth of literature in the four subject fields gave the following results and parameter values.

Subject	Value of parameters	
	q	M(Thousands)
PHY	0.539(0.006)	60834(16324)
CHEM	0.526(0.005)	597072(313392)
ENG	0.534(0.007)	514559(24416)
GEN	0.568(0.008)	80.14(15)

The values of R^2 and F which indicate the goodness-of-fit of the model are given separately:

Subject	R^2	F
PHY	0.998	28915.61
CHEM	0.998	43327.70
ENG	0.997	24672.44
GEN	0.997	19089.77

The value of q is again found to be positive in all the four subject fields. The value of M obtained is large in all the subject fields. The value of R^2 ranges between 0.997 and 0.998, fairly close to 1. Figure 4 presents the fit of data and estimated values of this model. It indicates that the model is able to capture the fluctuations in the data to a very large extent.

Exponential-logistic model

This model is mathematically expressed as:

$$dN/dt = N\ e^{b-cN} - N$$

When this model was applied to the growth of literature in the four subject fields under consideration, the results and parameter values obtained are given below:

Subject	Value of parameters	
	b	*c*
PHY	0.087(0.002)	0.013(0.001)
CHEM	0.061(0.002)	0.002(0)
ENG	0.081(0.002)	0.019(0.001)
GEN	0.081(0.003)	0.001(0.001)

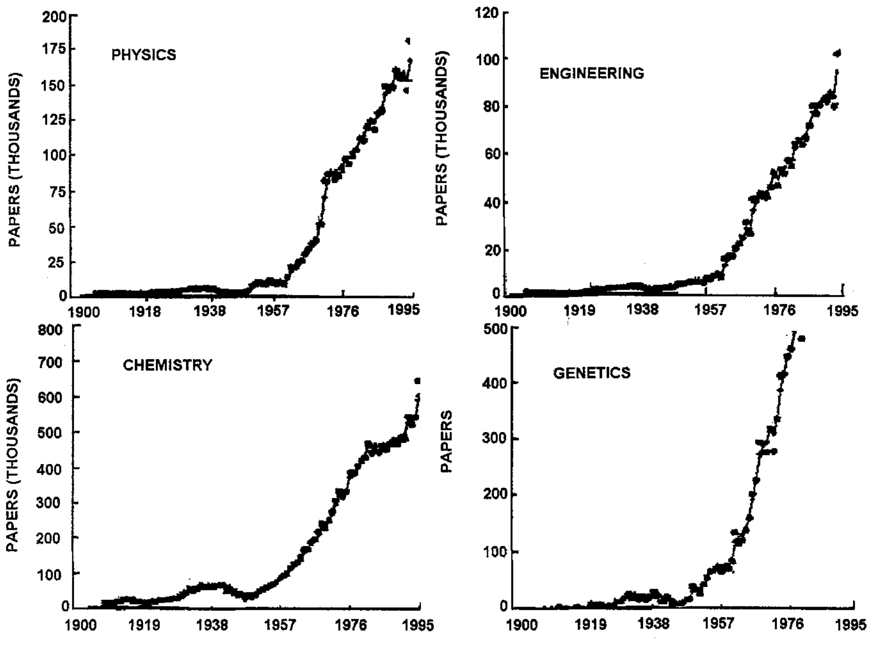

Fig. 4. Data and model fit for modified Mansfield model

The values of R^2 and F obtained in this model are as follows:

Subject	R^2	F
PHY	0.983	4015.557
CHEM	0.976	3500.153
ENG	0.983	.4081.385
GEN	0.977	2298.415

The model normally expects values of parameters b and c to be positive. The value of b is always positive for growth while that of c could be negative (or very small), if the saturation in the growth of the field is not indicated in the data. In the application of this model to the four subject fields the value of parameters b and c obtained are positive in all the four subject fields. The value of R^2 obtained for the four subject fields ranges from 0.976 to 0.983. Figure 5 presents the plot of data and estimate values for this model. It indicates that the model is not able to capture the fluctuations in the growth data.

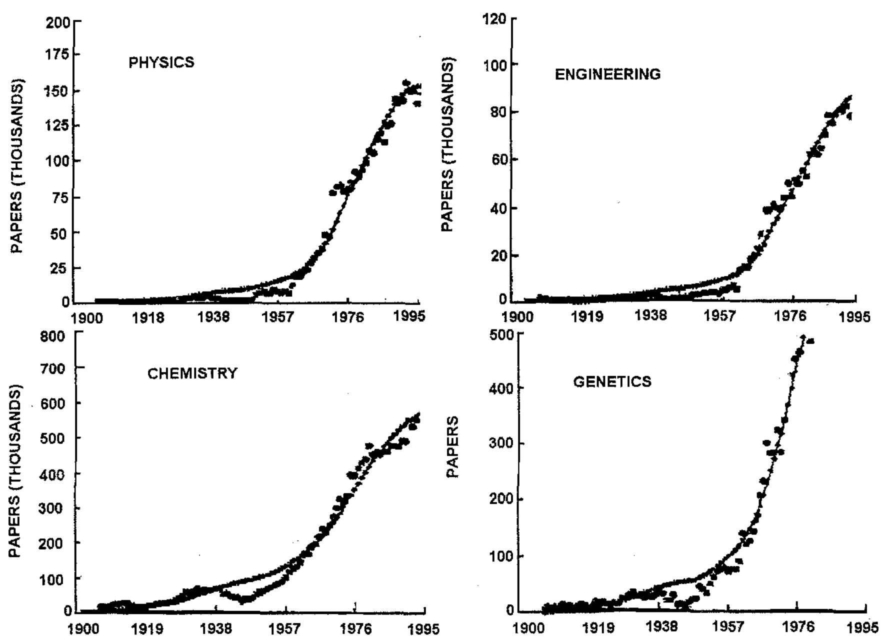

Fig. 5. Data and model fit for exponential logistic model

Modified exponential logistic model

The modified exponential logistic model is mathematically expressed as :

$$dN/dt = NS\, e^{b-cN} - NS\,,$$

where NS is, as defined in Eq.(14). When this model was applied to the growth of publications in the four choosen subject fields, the results and parameters values obtained are given below. The value of R^2 and F obtained from this model are also given separately.

Subject	Value of parameters	
	b	c
PHY	0.431(0.004)	0.006(0.002)
CHEM	0.423(0.003)	0.001(0)
ENG	0.428(0.004)	0.007(0.003)
GEN	0.450(0.005)	0.005(0.001)

The values of parameters b and c are again found to be positive in all the four subject fields. The value of R^2 has improved in all the four subject fields and now ranges from 0.997 to 0.998. Figure 6 presents the data and fit of the model. The graph indicates that the model is able to capture fairly well the fluctuations in the data.

Subject	R^2	F
PHY	0.998	28933.475
CHEM	0.998	43341.044
ENG	0.997	24672.459
GEN	0.997	19067.959

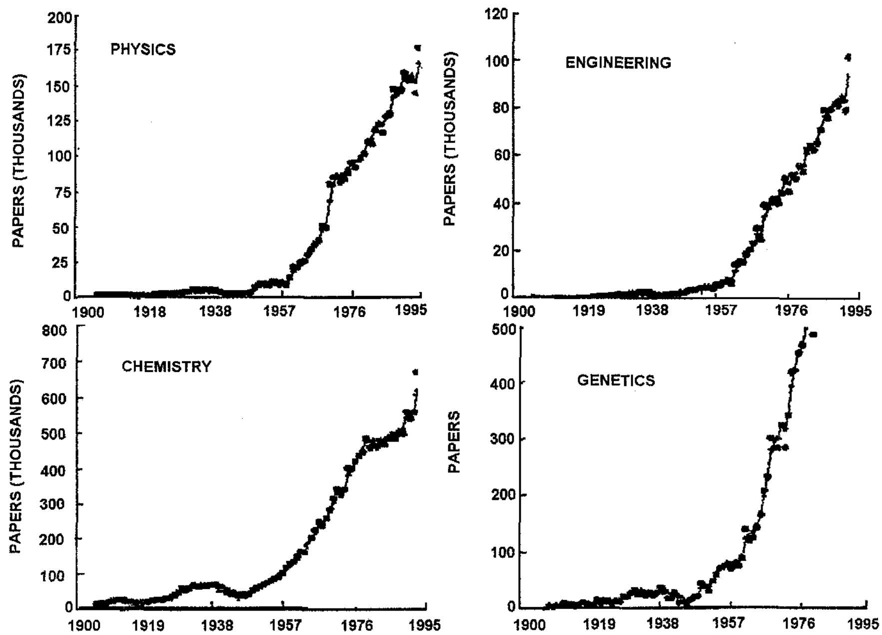

Fig. 6. Data and model fit for modified exponential logistic model

8. Conclusions

In this paper, we have focused on the applicability of six well established diffusion models to the growth of publication data in four choosen subject fields of science and technology. The results of the application of these models showed the same trend in all the four subject fields. The models have been discussed in the light of following factors: parameter values, fit statistics, and graphical fit to the data. In the original versions of the three models used namely Bass, Mansfield and Exponential-Logistic the values of parameters obtained are generally not found to be satisfactory to the model expectations. The results of the fit statistics mainly the value of R^2 and F were also not very good. From the graphical fit, it is observed that the models were not able to capture the fluctuations in the growth data. The drawbacks observed in these models were overcome by using the modified versions of these models. The application of the modified versions of these models has generally improved the parameter values, fit statistics and the graphical fit to the data. However, the range of improvements in the model differed from one model to another. The best improvement, however, seems to be observed in the modified Exponential - Logistic Model which gave the parameter values as expected from the data. The value of R^2 have also considerably improved and are quite close to 1. The fluctuations in the growth data were also fully captured by this model.

*

The authors express their sincere thanks to Dr Ashok *Jain*, Director, NISTADS, for encourement and support in writing this paper.

References

1. GILBERT, G. N. (1978), Measuring the growth of science: A review of indices, *Scientometrics,* 1(1):9–34.
2. SOLLA PRICE, D. J. DE. (1961), *Science Since Babylon,* New Haven, Connecticut; Yale University Press.
3. SOLLA PRICE, D. J. DE. (1963), *Little Science, Big Science,* New York; Columbia University Press.
4. ROGERS, E. M. (1962), *Diffusion of Innovations,* New York; Free Press of Glencoe.
5. VALENTE, T. W. (1995) *Network Models of the Diffusion of Innovations.* NJ; Hampton Press, Inc.
6. Rogers, E. M., SHOEMAKER, F. F. (1971), *Communication of Innovations: A Cross Cultural Approach,* New York; The Free Press.
7. ALBER, R., ADAMS, J. S., GOULD, P. (1971), *Spatial Organisation,* Englewood Cliffs, N J: Prentice Hall.
8. KUMAR, V., KUMAR, U. (1992), Innovation diffusion: some new technology substitution models, *Journal of Mathematical Sociology,* 17(2–3): 175–194.
9. MANSFIELD, E. (1961), Technical change and the rate of immitation, *Econometrica,* 29:741–766.
10. GRILICHES, Z. (1957), Hybrid corn: An exploration in the economics of technological change, *Econometrica,* 25: 501–522.
11. ROBINSON, B., LAKHANI, C. (1975), Dynamic Price models for new product planning, *Management Science,* 21:1113–1122.
12. BROWN, L. A. (1981), *Innovation Diffusion: A New Perspective,* New York; Methuen.
13. CASETTI, E., SAMPLE, R. K. (1969), Concerning and testing of spatial diffusion hypothesis, *Geographical Analysis,* 1:254–259.
14. SAHAL, D. (1981), *Patterns of Technological Innovations,* Rading, MA; Addison-Wesley.
15. SAHAL, D. (1983), Invention, innovation and economic evaluation, *Technological Forecasting and Social Change,* 23:213–235.
16. KARMESHU; BHARGAVA, S. C., JAIN, V. P. (1985), *Urban Economics,* 15:137–141.
17. BLACKMAN, A. W. JR. (1971), The rate of innovation in the commercial aircraft jet engine market, *Technological Forecasting and Social Change,* 2:3–4.
18. BLACKMAN, A. W. JR. (1972), A mathematical model for trend forecast, *Technological Forecasting and Social Change,* 3: 441–452.
19. NEVERS, J. V. (1972), Extensions of a new product growth model, *Sloan Management Review,* 13: 78–79.

20. EASINGWOOD, C. J., MAHAJAN, V., MULLER, E. A. (1981), Non-symmetric responding logistic model for technological substitution, *Technological Forecasting and Social Change,* 20:199–213.
21. TEOTIA, A. P. S., RAJU, P. S. (1986), Forecasting the market penetration of new technologies using a combination of economic cost and diffusion models, *Journal of Product Innovation Management,* 3: 225–237.
22. BEWLEY, R., FIEBIG, D. G. (1986), A flexible logistic growth model with applications in telecommunications, *International Journal of Forecasting,* 3: 225–237.
23. MCGOWAN, I. (1986), The use of growth curves in forecasting market developments, *Journal of Forecasting,* 5(1): 69–71.
24. HAMBLIN, R. L; JACOBSIN, R. B., MILLER, J. L. L. (1973), *A Mathematical Theory of Social Change,* New York; John Wiley.
25. MAHAJAN, V., PETERSON, R. A. (1985), *Models for Innovation Diffusion,* Beverly Hill; Sage Publications, p.10.
26. MAHAJAN, V., SCHOLEMAN, M. E. F. (1977), Generalized model for the time pattern of the diffusion process, *IEEE Transactions on Engineering Management,* EM 24(1):12–18.
27. BASS, F. M. (1969), A new product growth model for consumer durables, *Management Science,* 15: 215–227.
28. SHARMA, L., BASU, A., BHARGAVA, S. C. (1993), A new model of innovation diffusion, *Journal of Scientific and Industrial Research,* 52 (3): 151–158.
29. SHARMA, P., BHARGAVA, S. C. (1994), A non-homogenous non-uniform influence model of innovation diffusion, *Technological Forecasting and Social Change,* 46:279–288.
30. SHARMA, P., BHARGAVA, S. C. (1996), Study on diffusion pattern and forecast of Indian electronic consumer durables, *Journal of Scientific and Industrial Research,* 55(7):483–489.
31. STERMAN, J. D., (1985), The growth of knowledge: Testing a theory of scientific revolution with a formal model, *Technological Forecasting and Social Change,* 28: 93.
32. JAIN, A., GARG, K. C. (1992), Laser research in India: Scientometric study and model projections, *Scientometrics,* 23: 395–415.
33. GARG, K. C., SHARMA, P., SHARMA, L. (1993), Bradford's law in relation to the evolution of a field: A case study of solar power research, *Scientometrics,* 27:145–156.
34. GUPTA, B. M., SHARMA, L., KARISIDDAPPA, C. R. (1995), Modelling the growth of papers in a scientific speciality, *Scientometrics,* 33:187–201.
35. FALSENTEEN, J. (1981), *Bibliography of Theoretical Population Genetics,* Pennsylvania; Dowden, Hutshinson & Ross, Inc.