Technological Forecasting & Social Change 78 (2011) 346–357

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Innovation forecasting: A case study of the management of engineering and technology literature

Scott W. Cunningham*, Jan Kwakkel

Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Article history:
Received 16 December 2009
Received in revised form 1 November 2010
Accepted 5 November 2010

Keywords:
Innovation forecasting
Management of technology
Engineering management
Time series analysis
Spreadsheets
Probability models

ABSTRACT

The following paper contributes to the methodology of innovation forecasting. The paper analyzes the literature of engineering and technology management. A brief history and justification for interest in engineering and technology management is presented. The field has a sixty year history of interdisciplinary, and is therefore a ripe source for closer investigation into time trends of knowledge. The paper reviews the literature of innovation forecasting, examining a range of theoretical and methodological literatures interested in the evolution of knowledge. A new application of a model, suitable for sparse and count-like publication data, is presented. A mathematical presentation of the model is offered. A discussion is offered on how the model may be implemented in an approachable way within spreadsheet software. A time history of engineering management literature is extracted from a database and analyzed using the model. A projection of keyword growth is offered, and key features of the emerging knowledge base within engineering management are discussed. Recommendations for future research, as well as for those monitoring the status of the discipline of engineering management, are made.

1. Introduction

In this paper we examine the evolution of knowledge in the field of engineering and technology management. The goal of this survey is to assist journal editors, conference schedulers, and others involved in the building of the engineering management community to better assess the current state of knowledge in the field. The presented technique aims at providing a way to rapidly scan or monitor content areas of interest, highlighting if and in which direction further investigations are warranted. The aim of this technique is not to retrieve or model content or to produce maps of science, but rather to trace the dynamics of areas of interest over time. In addition to this practical concern, there are two intellectual reasons to be curious about the evolution of knowledge in the field of engineering and technology management. The first stems from curiosity about engineering management itself; a new and multidisciplinary field embodying a new perspective on the design, development and implementation of new technologies. The second reason stems from a more general investigation into the dynamic and evolutionary character of knowledge.

1.1. Motivation for the research

A brief sketch of engineering management follows; our goal is to chart the trajectory of the field and make it further clear why it is interesting to track the evolution of knowledge in this field. The years prior to the 1950s are known by some as the “Golden Age of Engineering.” The scale, scope and success of “heavy engineering” projects provided the discipline with new levels of prestige.

* Corresponding author.
E-mail address: S.Cunningham@tudelft.nl (S.W. Cunningham).

0040-1625/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.techfore.2010.11.001

Engineering systems analysis principles diffused to the government and military, which were interested in enhancing their investments in capital outlays. Early analyses attempted to quantify engineering and technological performance with a wider range of non-economic criteria. Also around this time there were early institutionalized attempts to foster collaboration between the engineering and management disciplines; many of these efforts failed. Early multidisciplinary faculties, such as Carnegie Mellon's "engineering and public policy department" were founded in the early 70s. A general enrichment of engineering education was also pursued, for instance through the Accreditation Board for Engineering and Technology (ABET), around this time. However, a backlash against systems thinking was also growing. Critiques both methodological and normative were launched. Methodologically there were concerns about the lack of holism, as well as a concern about the capability of systems analysis approaches in providing definitive advice for large scale societal problems. Normatively there were concerns that this perspective operated against a peaceful and pluralistic society. By the 1980s engineering management was established as a mainstream field of scientific publication. Although the field had been present in publication databases since the early 50s [3–6] the late 80s saw an explosion in both numbers of publications as well as the variety of sources reporting on engineering management.

Between the decade of the 80s and 90s, engineering management saw a more than five-fold increase in papers and sources, a growth which had not been seen before or since. The 1990s and 2000s saw increasing specialization and professionalization of the field with the founding of dedicated degree programs in management of engineering and technology

Table 1 illustrates the foregoing overview of the development of engineering management as reflected by publications and journals. The table shows both the number of papers per decade, as well as the numbers of associated journals per decade. Also listed is the number of publications associated with the single leading journal of that decade. Table 2 summarizes the timeline. As a field, the management of engineering and technology is interesting because of its extensive history as well as its interdisciplinary character. The fact that there were early, yet unsuccessful, attempts at institutional collaboration makes a closer examination of the evolution and integration of engineering and management knowledge even more interesting.

Dynamic theories of knowledge generation are of intrinsic interest to a number of different fields of research. The theory of dynamic capabilities is used by practitioners in strategic and technology management; these researchers are interested in how firms acquire and broker knowledge for economic advantage [10,11]. This theory within the strategic management literature draws upon evolutionary economic approaches [12]. The community of practice literature examines the technological and social strategies adopted by networks of experts attempting to remain productive while dealing with rapid changes in knowledge [13]. The field of scientometrics attempts to create both static and dynamic measures of knowledge production. Previous scientometric research into dynamic measures of knowledge production has examined a range of alternative measures including word usage [14], citation usage [15,16], and patterns of collaboration [17]. Scientometricians often use a sociology of science perspective when modeling the growth of knowledge [18]. This perspective, in turn, is strongly influenced by semiotics and linguistics [18,19].

2. The data

In this section we discuss the choice of data, and the query strategy. We review the time history of engineering management publication. We discuss measures of content for this collection of articles, and the indexing methodology which is used to produce tables for detailed analysis.

Routine scanning and monitoring activities for science and technology require an explicit focus on a smaller subset of science and technology as a whole. In this case, the explicit focus is on the management of engineering and technology. The level focus varies, but often five to ten thousand articles are a good subset. This specific focus is very different from national studies of science and technology, where a substantial fraction of the science base must be investigated for performance or benchmarking purposes. This specific focus is also very different from information retrieval applications, where by definition, the entire database is designed to be efficiently indexed and retrieved.

The data is collected from the ISI Web of Science database, searching all topic fields using the terms:

"management of technology" or "technology management" or "management of engineering" or "engineering management".

This query results in finding any one of the listed set of phrases present anywhere in the document. We do not limit publication type to journals, and therefore include conference proceedings, editorials and reviews. The ISI database was selected because of its

Table 1
ISI publications in engineering and technology management.

Decade	Papers	Single leading source	Journals
Before 1950	0	0	0
1950s	14	8	3
1960s	18	4	10
1970s	45	5	30
1980s	79	11	39
1990s	530	55	238
2000s	1182	37	>500

Table 2
Timeline of development in engineering management.

Decades	Event
Before 1950	Golden age of engineering [1]
1950s	Birth of systems analysis [2]; early engineering management publication [3–6]
1960s	Development of systems management [2,7]; Early institutionalized attempts at multidisciplinary collaboration [8];
1970s	Founding of early multidisciplinary faculties [8]; critiques of systems approaches [2,9]
1980s	Engineering management as a mainstream field for scientific publication; inclusion of engineering management indexing schemes; enrichment of engineering education
1990s	Offering of specialized engineering management degrees; launching of engineering management conferences; diffusion of ideas to industry
2000s	Increasing professionalization and differentiation

long indexing history, its familiarity to researchers in innovation policy research, and its broad sampling of relevant publication outlets.

These words are selected based on a naïve semantic model of the field. The combination of "engineering management" and "management of technology" into a single field with common interests is upheld by many practitioners in the field, for instance by the Portland International Conference of Management of Engineering and Technology (PICMET). Nonetheless, as will be shown, there are subtle differences in the usage of these terms, which must nuance our interpretation of the findings. On the other hand, the merit of the technique to be discussed is a quick and iterative scanning of databases of interest to highlight topics of interest. Given the approach it is readily possible to grow, expand or otherwise evolve the query as new information comes to light.

This query results in 2580 documents. These documents span a seventy year (see Table 3). The earliest document in the collection is a German language publication on technology and water management [20]. Other early contributions were from a special issue of the Proceedings of the Institute of Radio Engineers [3,4,6]. This foresighted special issue examined the challenges of management in the emerging electronics industry. (The Institute of Radio Engineers is an early precursor to the modern IEEE organization.)

Deeper insights into the evolution of knowledge in this field are afforded by looking at content terms in the documents. Several content measures are possible when using the ISI database. Words and phrases from titles and abstracts might be used. Subject categories, which are inferred based upon the journal of publication, may be used. Author provided keywords might be used. ISI also provides their own keywords which are harvested from the document according to an ISI-specific ontology.

For this research we use the author provided keywords, which have the advantage of providing a succinct, easily interpretable measure of content for the document. The measure, although limited, might be supplemented in future by additional sources of information. Author supplied keywords may not be the best choice for indexing and retrieval purposes. A richer source of semantic information might instead be obtained by fully indexing the article abstracts. However this would undoubtedly reveal that documents contain a semantic web of concepts, which over time and document collections is growing and changing dynamically. For the purposes of this paper we focus more closely on the dynamics of individual terms and phrases as revealed by author provided keywords. The primary threat to using author provided keywords is the fact that the authors may have an idiosyncratic concept of content which does not reflect the field as a whole.

Another limitation of using author provided keywords is the fact that this searchable field was only included in the ISI database for records from 1990 onward. Furthermore, not every author provides keywords to his or her documents. We do however have a total of 9273 keywords with which to appraise the content of these articles: roughly four keywords per article.

Use of keywords varies subtly according to whether the article is self-identified as "engineering management" or "technology management." A chi-squared test performed on the expected and observed use of keywords in the collection results in a chi-squared statistic of 11,851, with 9070 degrees of freedom. This result would be expected by chance in less than 5% of cases.

The observed occurrences of the top ten keywords across the three corpuses of engineering management, technology management and management of technology are given in Table 4. The expected occurrence of these keywords, assuming that the keywords are independently distributed of the corpus, is shown in Table 5. The principal differences in content lie with engineering management, although to a lesser extent there are differences in content between the "technology management" and "management of technology" corpuses.

Table 3
History of engineering management publication.

Decade	Publications
1930s	1
1940s	3
1950s	16
1960s	43
1970s	75
1980s	143
1990s	850
2000s	1460 (incomplete)

Table 4
Observed occurrences of keywords.

Phrase or query term	Total occurrence	Corpus (observed counts)		
		Engineering management	Technology management	Management of technology
Engineering management	64	59	2	3
Technology management	292	4	268	20
Innovation	143	6	105	32
Project management	40	15	15	10
Performance	91	7	60	24
Engineering education	12	10	2	0
Technology	70	4	46	20
Software engineering management	8	8	0	0
Strategy	55	3	41	11
Information technology	43	1	31	11

By inspecting the observed versus expected matrices it is possible to assign keywords to the corpus which emphasizes the keyword more than is expected given statistical independence (Table 6). The engineering management corpus emphasizes engineering, which is unsurprising. However it also emphasizes "project management." The management of technology corpus emphasizes technology, again unsurprising, but it also emphasizes "innovation," "performance," and "strategy." Evidence of such subtle differences alerts the analyst to the necessity of treating the source query with care.

In the following paragraphs we survey the leading key terms and phrases in order to get a general perspective of the field prior to creating a dynamic model of keyword usage. We split the frequency of the keywords directly related to the query from the other keywords for it is tautological to claim that the superset will grow faster than any of its subsets. For this same reason, we also present the results of the method for the keywords directly related to the query separate from the remainder of the keywords. This is shown in Table 7. Note that there were no documents in the corpus which used multiple query-related keywords.

Table 8 shows the frequency of the keywords in the dataset. We choose the original query terms to tabulate over time, since as demonstrated, differential growth in the query components can affect over-all content. In addition, we include the top thirty-three words or phrases in the data set, resulting in thirty-five keywords for tabulation. This represents roughly 1 in 4 of the total keywords.

In general, a given body of literature is constantly innovating new keywords, resulting in any fixed list growing rapidly out of date. Thus, a catch-all keyword, representing all others not listed, is also included in the list. This is an important indicator of the presence of content change. These keywords are presented by rank and frequency in Table 9.

The documents are tabulated across all years of the data using these thirty-five keywords and the catch-all phrase The resulting table shows the dynamic evolution of keywords over time. The perl scripting language was used to produce the tables based on the full downloaded ISI records.

3. Method

In this section we discuss features of the data, and note previous research involving the modeling publication counts. Previous research includes trend extrapolation, population modeling, and linear dynamical models. Finally, the model itself is described.

3.1. Sparse and count-like characteristics of the data

The data is sparse and count-like in character. The average rate of keyword usage in any given year is 1.80. The probability of any keyword–year combination being absent from the data is 55%. Particularly in the earlier years, the use of keywords was infrequent. These stylized facts suggest the use of Poisson models for modeling the data.

Table 5
Expected occurrences of keywords.

Phrase or query term	Engineering management	Technology management	Management of technology	Partial chi-squared value
Engineering management	10.72533	37	15.93616	2332.443
Technology management	48.93433	170	72.70872	2010.613
Innovation	23.96441	83	35.60735	321.0601
Project management	6.703332	23	9.960099	69.04959
Performance	15.25008	53	22.65923	67.57893
Engineering education	2.011	7	2.98803	63.88617
Technology	11.73083	41	17.43017	59.41612
Software engineering management	1.340666	4.7	1.99202	44.3812
Strategy	9.217082	32	13.69514	38.43117
Information technology	7.206082	25	10.70711	38.34298

Table 6
Relative emphasis by corpus.

Phrase or query term	Emphasis
Engineering management	Engineering management
Project management	Engineering management
Engineering education	Engineering management
Software engineering management	Engineering management
Technology management	Technology management
Innovation	Technology management
Performance	Technology management
Technology	Technology management
Strategy	Technology management
Information technology	Technology management

Normal approximations to the data, even using continuity corrections, fail when rates are less than five [21]. There are concerns with error-modeling, parsimony and predictive validity when using Gaussian models when these models are distributionally inappropriate. Gaussian models over-weight high count and high variance years. As a result critical information at the start of new publication trends is effectively discarded. Gaussian distributions, when used inappropriately, result in excess model parameters, and therefore an inability to generalize models when new data is present. Gaussian noise, when added to model structure, may result in negative predictions. Predictions of "negative" publication count are suspect both conceptually as well as validity-wise.

Unfortunately there are few existing Poisson models to be used as exemplars in the scientometrics, bibliometrics or informetrics literatures. An exception is [22] where the authors model the distribution of papers by single authors as a Poisson distribution, and the distribution of papers across authors as a gamma distribution. The authors present evidence in light of the analysis of actual production data [23].

3.2. Innovation forecasting

An early effort to model and forecast publication and patent counts was provided by [24]. These authors model the early stages of the innovation cycle using a broad spectrum of indicators and methods. In particular the authors draw upon established models of trend extrapolation, such as the Fisher–Pry model. Given the important role of these models in innovation forecasting, a closer survey of these models is given next.

3.2.1. Trend Models

Trend models examine the adoption or diffusion of new technologies into the marketplace. These models are accompanied by techniques for non-linear regression, allowing the structural models to be fit to real data. Estimates of future growth or diffusion based on data are then possible. Thus, the principal character of these models is their use in modeling directly observed indicators of growth and substitution.

Four major kinds of market models include the Fisher–Pry, Pearl, Gompertz and Bass models [25,26]. These models differ in their postulated underlying non-linear processes of diffusion and saturation. The underlying process is often justified using dynamic models of population growth.

For instance, the Bass model is intended to describe the first adoption of a new technology. The model assumes that there are two key adoption processes. Some fraction of consumers will unconditionally adopt a new technology at a given period of time. This fraction is known as the "coefficient of internal innovation." Another fraction of consumers will adopt a new technology only if their peers have adopted the technology. This fraction is known as the "coefficient of external innovation." Together the internal and external coefficients determine the ultimate speed and extent of technology adoption. Since the model describes the first adoption only, once all prospective customers have adopted the technology predicted new adoptions then cease. The Bass model is closely related to the Pearl and Fisher–Pry models of technology adoption. The difference is that the Bass model predicts the rate of new adoption, while the Pearl and Fisher–Pry models forecast the cumulative new adoptions.

3.2.2. Population dynamics

Population dynamic models are more fundamental in character than either innovation forecasting or market extrapolation models. It is helpful however to examine the assumptions behind these models to better understand their use and applicability.

Table 7
Position of query related keywords among the leading keywords.

Rank	Keyword	Frequency
1	Technology management	292
7	Engineering management	64
11	Management of technology	46

Table 8
Leading keywords in the data.

Rank	Word or Phrase	Frequency in the Data
1	Technology management	292
2	Innovation	143
3	Management	92
4	Performance	91
5	Model	78
6	Technology	70
7	Engineering management	64
8	Strategy	55
9	Knowledge management	49
10	Systems	48
11	Management of technology	46
12	Information technology	43
15 (tied)	Firms	42
15 (tied)	Knowledge	42
15 (tied)	Product development	42
16	Perspective	41
17	Project management	40
18	Firm	39
19	Industry	33
20	Framework	30
22 (tied)	Competitive advantage	29
22 (tied)	Design	29
25 (tied)	Business	27
25 (tied)	Technology strategy	27
25 (tied)	Technology transfer	27
27 (tied)	Impact	26
27 (tied)	Networks	26
28	Strategies	25
29	Information	23
35 (tied)	Capabilities	21
35 (tied)	Evolution	21
35 (tied)	Implementation	21
35 (tied)	Quality	21
35 (tied)	Strategic management	21
	All others	7549

The structural character of trend models implies a deterministic or stochastic dynamic at work. These models are written in terms of differential equations, where the pattern of growth is modeled as an interaction between a positive feedback loop of growth, and a negative feedback loop of saturation. Stochastic variants of these models are possible, where the positive feedback loop is described as “growth,” and the negative feedback loop is described as “saturation” [27].

These models generate the structural forms associated with trend extrapolation models, including exponential and hyperbolic growth. In addition, in their stochastic forms, they duplicate probabilistic processes such as the normal, gamma and beta distribution. The basic character of these models involves explanation of future growth based on current, observed phenomenon. This explanatory form allows a fairly deep explanation of how and why growth occurs in terms of population-level forces and incentives [28]. Thus, similar surface level dynamics may arise despite fundamental differences in the underlying process itself.

3.2.3. Linear dynamical systems

The fundamental character of linear dynamical systems is an explanation of dynamic behavior based upon partially unobserved characteristics of the system. The system may contain the presence of noise which may alter either the dynamic, or the observation and measurement of the system. The Kalman filter is an efficient mechanism for estimating the parameters of a linear dynamic system given data.

A data-oriented review of linear dynamic systems is provided by Roweis and Gharamani [29]. This review makes clear the relationships between linear dynamic systems and other statistical models. Most interesting are the connections between linear dynamic systems and factor analysis. Previous bibliometric work has attempted to model the evolution of semantic spaces of

Table 9
Model quality of fit.

No	Description of model	Parameters	Log likelihood	AIC
1	Constant rates	35	−3379	6964
2	Exponential	70	−1178	2496
3	Second-order exponential	105	−1130	2330

documents and concepts [14,19]. This work is challenged however by creating a common basis for comparing the evolution of the system over time. Linear dynamical systems are the logical extension of science maps to consider dynamical phenomenon.

A basic description of the linear dynamical system is as follows [29]. The linear dynamical system is a discrete time dynamical system with Gaussian noise. The system consists of an underlying state matrix x, which might be of any dimension. We withhold indices on this state matrix without loss of generality. The matrix which describes the transition of this state over time is the matrix A. The subscripts $(0, 1, \ldots t, t+1)$ describe the time evolution of this system. A source of Gaussian noise (wt) is added to the system. We may specify this as a stationary Gaussian distribution with mean zero and covariance matrix Q, without any loss of generality.

Eq. (1). The linear dynamical system

$$x_{t+1} = Ax_t + w_t = Ax_t + w_0,\ w_0 \sim \mathcal{N}(0, Q)$$
$$y_t = Cx_t + v_t = Cx_t + v_0,\ v_0 \sim \mathcal{N}(0, R)$$

This hidden state vector is transformed into the observable state of the system through matrix C. Here too there is a source of Gaussian noise (vt). This is a stationary Gaussian distribution with a mean zero and covariance matrix R. This equation is of particular and interest in forecasting when the underlying state (x) is of much smaller dimensionality than the observed vector (y). Thus complex dynamics are explained and anticipated with a relatively simple description of the system.

3.3. The model

Having reviewed innovation forecasting, trend extrapolation, population dynamics, and linear dynamical systems we can now present the proposed model. This model combines ideas from each of these in pursuit of a model of publication. The innovation forecasting perspective is valued because of its conceptualization of publication within a larger context of innovation. The trend extrapolation perspective is useful because of its capability for non-linear modeling of observed phenomenon. The dynamical systems perspective is useful because the potential for explaining the underlying dynamics in terms of population-level forces. The introduction of stochastic elements is also useful here. The linear dynamical systems perspective is useful because it provides the capability for modeling growth based on unobserved, underlying factors. As noted, previous scientometrics approaches have attempted to model the dynamics of knowledge over time.

The model is described in Eq. (Eq. (2). Suppose that there are n keywords and t years in the data. Each keyword has the same explanation dependent on parameters, and the keywords pattern over time is independent of one another conditioned on the model. Thus, we may drop the index n from the equation without loss of generality.

Eq. (2). The generative model of the data

$$\lambda_t = exp\left(a + bt + ct^2\right)$$
$$y_t = Pois(\lambda_t)$$

The mean rate of publication is an exponential function of time. There is a constant term (a), an exponential growth term (b), and further we expect a higher-order term reflecting for instance the increasing difficulty of publishing in a mature and therefore saturated field (c). The resulting dynamic equation of expected publication over time is a non-homogeneous differential equation.

This mean rate is then entered as a parameter in a Poisson random variable. Thus, any given realization of the process is discrete, and subject to natural variation year over year. The exponential link in these equations is associated with Poisson models, and guarantees that the domain of the function remains positive [30]. The Poisson is always associated with non-negative rates.

The model very closely approximates the Bass model. A small, fifth-order approximation is needed to make the two formulations equivalent. In practice, the differences are very small, much smaller than the trend error itself. Eq. (Eq. 3) shows the model; it entails regressing the current publication levels on the cumulative publication up to all previous years.

Eq. 3. A Bass-equivalent model

$$y_t = a + b\sum_{i=0}^{t-1}(y_t) + c\sum_{i=0}^{t-1}\left(y_i^2\right)$$

Given the close approximation, it is therefore worth comparing the dynamics of the Bass model with a prospective model of innovation in science. Like the Bass model, there may be a pool of prospective scientists able and willing to adopt new scientific ideas. These ideas are partially and incompletely measured by keyword usage in scientific articles. A fraction of these scientists may spontaneously experiment with new ideas. A separate fraction may emulate other scientists adopting new ideas only once they are tried and proven by others. Unlike the Bass model, the underlying scientific population may be rapidly growing. Further, the quantity of publication per year per scientist may also be increasing, if only because scientific research is becoming better instrumented. Thus, the Bass model may be translated to create a rough, if ready approximation of the dynamics of scientific publication.

3.4. Implementation details

One way of estimating the linear dynamic system is to approach the problem as a matter of non-linear optimization. The problem may then be solved using Frontline Solver, in the Excel package [31]. This approach is suitable for smaller problems, and has the advantage of rapid prototyping, ease of access and accountability [32].

An overview of the spreadsheet layout is given in Fig.1. The model consists of the following elements:

— Objective function
— Decision variables (*C*, parameter values)
— Derived variables (*X*, logged rates)
— Parameters (*A*, dynamic structure)
— Expectations (λ, Poisson means)
— Actual data
— Penalty function (log likelihood)

4. Results

The first task is to evaluate of possible specifications consistent with the model described in the previous section. Two such models are given (Table 9). The parameters of each of the models are given, as is the log likelihood of the best fitting model. Increasing the dimensionality of the *C* matrix is more expensive in parameters than adding additional parameters to the state matrix *A*. This is because the *C* matrix is sized proportional to the keywords (30 in this example), while the *A* matrix only adds a new parameter or two.

The first, the null hypothesis, is a simple static model without changes in rates over time. The second model presumes that the growth rates are exponential over time. The second-order exponential model tested involves a deceleration of growth over time.

As can be seen from Table 9, the proposed model provides the highest likelihood, and the most informative model as indicated by Aikake's Information Criterion (AIC). A component of saturation appears to be an important explanation of the dynamic in this data. The null hypothesis, a static usage of keywords, is handily rejected. Most of the lack of fit is contributed by the composite term "all others" which contributes −195 units of log likelihood out of the total −1130. A closer look at the time trajectory of this poorly fitting catch-all term is shown in Fig. 2. The quality of the fit appears adequate.

The dynamic model of Eq. (2) is suggested solely as a descriptive model of the data. As can be seen from the model results, the resultant quality of fit is high despite relatively few required parameters of the model. We suggest that the model is suitable for

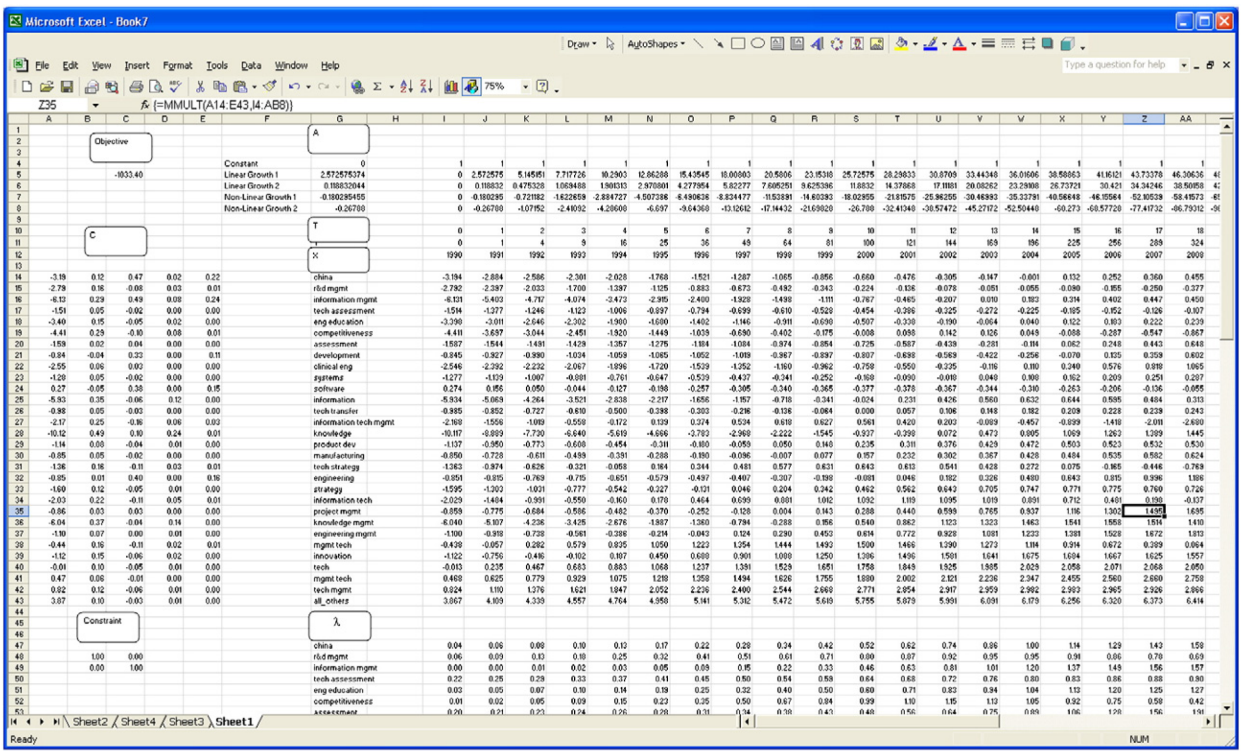

Fig. 1. Overview of the spreadsheet layout results.

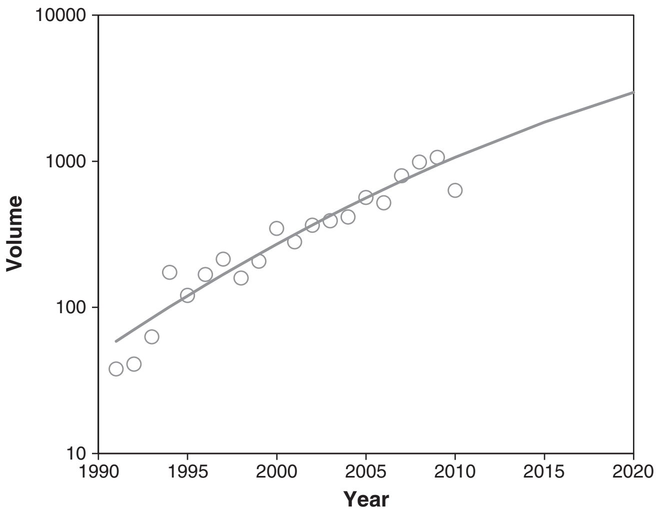

Fig. 2. Trajectory of the "all others" keyword.

generalization over the short term; as here with twenty years of data, a four or five year outlook seems reasonable. Despite these notes about model adequacy, a general theory of publication dynamics is still needed. Confronting theory and data would undoubtedly enrich both endeavors.

Table 10
Projection of query related keyword usage.

Keywords	2010 (incomplete, actual)	2010 (projected)	2015 (projected)
Engineering management	5	11.5	27.8
Management of technology	1	2.2	0.2
Technology management	19	20.5	13.8

Table 11
Projection in keyword usage.

	2010 (actual)	2010 (projected)	2015
Business	5	3.5	4.0
Capabilities	2	2.2	0.5
Competitive advantage	2	2.0	0.7
Design	2	3.0	2.2
Evolution	0	1.0	0.1
Firm or firms	8	9.6	11.7
Framework	5	6.4	14.4
Impact	4	3.6	5.1
Implementation	0	2.2	2.1
Industry	7	9.1	76.2
Information	2	1.7	0.7
Information technology	0	2.8	1.1
Innovation	4	13.2	10.0
Knowledge	5	7.4	10.3
Knowledge management	2	5.5	1.9
Management	6	9.8	9.5
Model	11	9.6	10.9
Networks	2	3.0	2.5
Performance	5	13.6	13.1
Perspective	4	7.0	24.7
Product development	0	2.7	0.8
Project management	3	6.4	13.6
Quality	2	3.1	4.8
Strategic management	3	2.6	5.5
Strategy or strategies	5	4.5	1.2
Systems	3	6.6	7.8
Technology	5	5.3	2.1
Technology strategy	1	1.0	0.4
Technology transfer	5	3.2	9.2

Table 12
Query related keyword velocity and momentum.

Keyword	Velocity	Momentum
Engineering management	242%	27.79
Management of technology	9%	0.2
Technology management	67%	13.82

We now project the model forward using the best fitting model. This enables us to draw more acute conclusions concerning the current and future state of the field of engineering and technology management. Tables 10 and 11 show a projection to 2015, alongside actual and projected 2010 data. As can be seen, there are some consequential shifts in the perceived content of these papers. The most significant changes as represented by this projection include the rate of new keyword usage, and shifts in the relative usage of “management of technology” and “technology management.” By 2015, an increase in 74% in the all others category is seen, reflecting the continual change of the field. While usage of the words “technology management” and “management of technology” has significantly declined, drops are replaced with corresponding increases in the use of the cognate “engineering management.” The relative merits of one of these terms over the other remain unclear, but as was noted earlier there are distinct interests and precommitments associated with each.

Tables 12 and 13 show two derived measures of keyword growth: velocity and momentum. Velocity is the comparative growth (or decline) of keyword usage from year 2010 (as estimated) to 2015 in the model. Momentum is the growth rate multiplied by the 2010 estimated rate of term usage. Both velocity and momentum are needed to assess the significance of keyword usage. Velocity pinpoints words showing significant growth of keywords, while momentum shows those fast growing keywords that have reached a significant level of adoption.

A simple 2×2 framework affords a monitoring scheme for further decomposing categories of growth by keywords (Table 14). Low growth and low momentum represents specialized concepts used by relatively few authors. High growth but low momentum represents a source of emerging concepts. High momentum but low velocity represents a stable of core concepts in the field. Both high growth and high momentum keywords represent the most significant source of growth in the field.

The framework given in Table 14 is applied by finding a breakpoint in velocity and momentum, and placing keywords in the corresponding four categories according to their growth characteristics. The resultant portfolio allows us to categorize the keywords appropriately. Note the strong correlation between velocity and momentum. Despite this, simple analytic procedures can be used to find breakpoints for velocity and momentum.

Table 13
Keyword velocity and momentum.

	Velocity	Momentum
Business	114.6%	4.01
Capabilities	24.2%	0.54
Competitive advantage	34.5%	0.68
Design	72.6%	2.19
Evolution	9.4%	0.10
Firm or firms	122.6%	11.75
Framework	227.2%	14.43
Impact	140.5%	5.05
Implementation	98.1%	2.14
Industry	836.7%	76.18
Information	39.7%	0.69
Information technology	38.6%	1.07
Innovation	76.2%	10.04
Knowledge	138.7%	10.33
Knowledge management	33.9%	1.85
Management	96.9%	9.50
Model	113.5%	10.86
Networks	81.5%	2.46
Performance	96.9%	13.14
Perspective	350.9%	24.66
Product development	31.6%	0.84
Project management	210.7%	13.57
Quality	153.5%	4.79
Strategic management	213.0%	5.45
Strategy or strategies	26.6%	120.3%
Systems	117.9%	783.8%
Technology	40.0%	210.1%
Technology strategy	38.5%	37.2%
Technology transfer	287.9%	924.5%

Table 14
2 × 2 framework for monitoring keyword usage.

		Momentum	
		Low	High
Velocity	High	Emerging concepts	Growth concepts
	Low	Specialized concepts	Core concepts

Table 15 shows the resulting portfolio. This portfolio is likely to change over time. The underlying dynamics of the model suggest a variable rate of velocity over time. Furthermore, the dynamics implied by the model suggest a variable relationship between velocity and momentum as well.

5. Closing remarks

The results suggest that technology and engineering management is growing more applied. Attention to business, firms and industries is increasingly more prevalent. Project management is on the rise as a topic for research. Likewise, implementation and quality are increasing in prominence in the literature.

The task of innovation forecasting presents multiple challenges. One challenge entails treatment of the unique characteristics of publication count data. Another entails the need for constant scanning of new and emerging terms from the domain of low frequency keywords. This paper presented a mathematical model for use in innovation forecasting. The presented model aims at providing a way to rapidly scan or monitor content areas of interest. The model allows effective treatment of Poisson distributed data. Furthermore, when coupled with indexing software and spreadsheet modeling affords a relatively rapid technique for keeping atop new and emerging concepts within a literature base. We suggest that the model and implementation provide useful desktop monitoring procedures for continuous monitoring of emerging fields of new technology.

The technique is illustrated by focusing on a particular subset of science and technology as a whole — in this case the fields of engineering and technology management. Term growth in that field is naturally limited by the query itself — no term within the set engineering and technology management can grow faster than the collection as a whole. In this paper then, we examine the internal growth of a key term relative to the field as a whole. Nonetheless, it may be helpful for some applications to examine the external growth of the term. A growth perspective from within the query, as well as from science and technology as a whole, can be helpful.

Further exploration of keywords beyond the top thirty is needed. Of the thirteen terms identified as high growth concepts, nine were from the second decade of ranked keywords. Undoubtedly even more significant keywords from the third decade and below are likely to emerge.

It is worth comparing the dynamics of the Bass model [26] with this prospective model of innovation dynamics in science. In the Bass model, there may be a pool of prospective scientists able and willing to adopt new scientific ideas. These ideas are partially and incompletely measured by keyword usage in scientific articles. A fraction of these scientists may spontaneously experiment with new ideas. A separate fraction may emulate other scientists adopting new ideas only once they are tried and proven by

Table 15
Portfolio of leading keywords.

Emerging concepts	Growth concepts
Business	Engineering management
Framework	Firm or firms
Impact	Knowledge
Implementation	Management
Industry	Model
Perspective	Performance
Quality	Project management
Strategic management	systems
Technology transfer	
Specialized concepts	**Core concepts**
Capabilities	Innovation
Competitive advantage	Knowledge management
Design	Management of technology
Evolution	Strategy or strategies
Information	Technology
Information technology	Technology management
Networks	
Product development	
Technology strategy	

others. However, unlike the Bass model, the underlying scientific population may itself be rapidly growing. (Bass himself notes that the market potential may be slowly changing over time.) Further, the quantity of publication per year per scientist may also be increasing, if only because scientific research is becoming better instrumented. Thus, the Bass model may be translated to create a rough, if ready approximation of the dynamics of scientific publication.

Furthermore, a more complete investigation of the construct validity of keywords is needed. Previous research has suggested that words and phrases within titles and abstracts provide a model of content which may be generalized across researchers [33,34]. While such a framework would not obviate the need for keywords, additional confidence in their use and significance could be achieved. Additional work on joint models of both semantics as well as temporal dynamics, for the purposes of innovation monitoring and scanning, is also needed.

References

[1] S.C. Florman, The Civilized Engineer, St. Martin's Griffin, 1988.
[2] F. Capra, The Web of Life: a New Scientific Understanding of Living Systems, Anchor Books, New York, 1996.
[3] W.A. MacDonald, Production aspects of engineering management in the electronics industry, Proceedings of the Institute of Radio Engineers 41 (3) (1953) 425-425.
[4] M.J. Kelly, Research and development problems of engineering management in the electronics industry, Proceedings of the Institute of Radio Engineers 41 (3) (1953) 425-425.
[5] H. Pratt, General problems of engineering management facing the electronics industry, Proceedings of the Institute of Radio Engineers 41 (3) (1953) 425-425.
[6] D.L. Putt, What the military services expect from engineering management in the electronics industry, Proceedings of the Institute of Radio Engineers 41 (3) (1953) 425-425.
[7] C. Freeman, Measure of Output of Research and Experimental Development, UNESCO, 1969.
[8] M.G. Morgan, Carnegie Mellon's department of engineering and public policy, International Journal of Technology, Policy and Management 1 (2) (2001) 138–150.
[9] H.W.J. Rittel, M.W. Webber, Dilemmas in a general theory of planning, Policy Sciences 4 (1973) 155–169.
[10] D.J. Teece, G.P. Pisano, A. Shuen, Dynamic capabilities and strategic management, Strategic Management Journal 18 (7) (1997) 509–533.
[11] F. Malerba, Sectoral systems of innovation and production, Research Policy 31 (2002) 247–264.
[12] M. Zollo, S.G. Winter, Deliverate learning and the evolution of dynamic capabilities, Organization Science 13 (3) (2002) 339–351.
[13] D. Fuhr, F. Fuchs-Kittowski, Against hierarchy and chaos: knowledge coproduction in nets of experts, Journal of Universal Computer SCience 10 (3) (2004) 176–185.
[14] E.C.M. Noyons, A.F.J. van Raan, Monitoring scientific developments from a dynamic perspective: self-organized structure to map neural network research, Journal of the American Society for Information Science 49 (1) (1988) 68-61.
[15] E. Garfield, Historiographic mapping of knowledge domains literature, Journal of Information Science 30 (2) (2004) 119–145.
[16] K.W. Boyack, K. Borner, R. Klavans, Mapping the evolution and structure of chemistry research, Proceedings of ISSI 2007: 11th International Conference of the International Society for Scientometrics and Informetrics 1 (2007) 112–123.
[17] E.O. Reid, Evolution of a body of knowledge: an analysis of terrorism research, Information Processing and Management 33 (1) (1997) 91–106.
[18] M. Zitt, A simple method for dynamic scientometrics using lexical analysis, Scientometrics 22 (1) (1991) 229–252.
[19] A. Rip, J.P. Courtial, Co-word maps of biotechnology: an example of cognitive scientometrics, Scientometrics 6 (1984) 265–269.
[20] O. Schoenefeldt, F. Alten, The significance of technology and pedology for German water resource management, Angewandte Chemie 48 (1935) 101–110.
[21] M.L. Berenson, et al., The normal approximation to the binominal and Poisson distributions, Business Statistics, Prentice-Hall, Inc, 1998.
[22] J.C. Huber, Cumulative advantage and success-breeds-success: the value of time pattern analysis, Journal of the American Society for Information Science 49 (5) (1998) 471–476.
[23] S.J. Bensman, Urquhart and probability: the transition from librarianship to library and information science, Journal of the American Society for Information Science and Technology 56 (2) (2005) 189–214.
[24] R.J. Watts, A.L. Porter, Innovation forecasting, Technological Forecasting and Social Change 56 (1) (1998) 25–47.
[25] A.L. Porter, et al., Forecasting and Management of Technology, John Wiley and Sons, New York, 1991.
[26] F. Bass, A new product growth model for consumer durables, Management Science 15 (5) (1969) 215–227.
[27] L. Cobb, Stochastic differential equations for the social sciences, in: L. Cobb, R.M. Thrall (Eds.), Mathematical Frontiers of the Social ad Policy Sciences, Westview Press, 1981.
[28] J. Maynard Smith, Evolution and the Theory of Games, Cambridge University Press, Cambridge, 1982.
[29] S. Roweis, Z. Ghahramani, A unifying review of linear Gaussian models, Neural Coputation 11 (2) (1999) 305–345.
[30] P. McCullagh, J. Nelder, Generaized Linear Models, Chapman and Hall, London, 1989.
[31] D. Fylstra, et al., Design and use of the Microsoft Excel Solver, INFORMS Interfaces 6 (1998) 29–55.
[32] C. Ragsdale, Spreadsheet Modeling and Decision Analysis, South-Western College Pub, 2003.
[33] S.W. Cunningham, The content evaluation of British scientific research, Science Policy Research Unit, University of Sussex, 1996.
[34] S. Deerwester, et al., Indexing by latent semantic indexing, Journal of the American Society for Information Science 41 (6) (1990) 391–407.

Scott Cunningham received a Ph.D. in Science, Technology and Innovation Policy from the Science Policy Research Unit. He has worked as a knowledge discovery analyst for software and database companies, and has helped in advising the British and Malaysian governments on their national profiles of scientific activity. He currently works at the Delft University of Technology, as an assistant professor.

Jan Kwakkel is a Ph.D. student at Delft University of Technology, The Netherlands. His research interests are handling uncertainty in model-based public policy analysis and scientometrics. He will defend a Ph.D. on the topic of the treatment of uncertainty in airport strategic in December 2010.