Dynamic Patterns in Processes of Science Systems: Science Academies & their Journals-An illustrative example

Ashok Jain[1] and Mrinal Kanti Das[2]

[1]*EMPI Business School, New Delhi, INDIA.*

[2] *Institute of Informatics & Communication, University of Delhi South Campus, Benito-Juarez Road, New Delhi-110021, INDIA.*

ABSTRACT

The paper presents a quantitative method of comparing the hitherto less explored processes embedded in science activities performed by different organisations. The method requires only time series data of the output of activity. As an illustration, the method is applied to compare the processes embedded in activity of academies in the US, Japan, China (Taiwan) and India; the chosen activity is meant to disseminate results of scientific results through English language journals. The activity is viewed as an activity performed by a complex system involving interactions at different times between paper contributors, finance providers, peers, readers, organisations and structures responsible for its publication and distribution. The paper estimates Permutation Entropy as a complexity index to characterise processes embedded in the activity. The method circumvents the need of measurable data on individual actors and agencies involved in the activity. Results reveal similarities in processes adopted by academies in US and Japan and hence they constitute a cluster; academies of Taiwan and India lay out side this cluster. Referring to literature on organisational learning, it is pointed out that complexity index also reflects the ability of the system to learn from experience, an observation that has policy implications. Finally it is noted that as the method requires only time series data of the activity out put, it can be applied for inter country comparisons of processes and learning abilities embedded in other science and technology sub systems of National Innovation System and can in principle be used as an additional tool for cross-country comparisons of National Innovation Systems.

Key words: Science Academies, National Innovation System, time series, complexity index, Permutation entropy.

INTRODUCTION

Reviewing developments in Scientometrics, Mingers and Leydesdorff [1] have identified two emerging dimensions in Scientometrics studies as a) study of processes of science activity and b) publication of science journals. The present paper is an attempt to study some aspects of both these dimensions.

We interpret the term science activity to refer to all such activities that contribute in one way or the other to development of science. and technology included in National Innovation System[2-4]. The activity of publication of a science journal is also viewed as a science activity; it contributes to development of science through dissemination of peer evaluated research results amongst the scientific community.

Publication of journal is viewed as an activity of a Science Academy /Society/ Association/Institution performed by a group of scientists such as an editorial board to help development of science in some desired area or areas of research. The main components of their activity are: receiving research papers received from scientists working in various research organisations such as universities and research institutions, getting them peer-evaluated and, passing on the selected papers to a printer/distributor for conversion into a journal its and marketing.

***Address for correspondence:**

Author Name: Mrinal Kanti Das

Author Department and address: Institute of Informatics & Communication, University of Delhi South Campus, New Delhi-110021.

E-mail: dasmkd11@gmail.com

Phone no: 09643207600

Access this article online

Official Publication of Phcog.Net

Website: www.jscires.org

DOI: 10.5530/jscires.5.3.2

From an organisational perspective, one may consider the journal publication activity to involves interactions between a network of several organisations; a nodal coordinating organisation (Science Academy /Society/ Association/Institution), a dispersed set of organisations where research is performed (such as Universities, Research Institutions) and an organisation for printing and marketing of the journal.

Succinctly stated, the journal publication science activity is:

Conversion by organisation such as Science Academy of research information generated by scientists in various organisations into a value added 'information product' packaged and marketed as a Journal by a professional organisation[1].One expects a Science Academy to change its processes of journal publication activity to respond to changing requirements of its 'customers' namely the scientific community.

Different Academies may adopt different processes with differing abilities to respond to change, in other words the dynamics of processes embedded in journal publication activity may differ between academies.

This paper is concerned specifically with the dynamics of interactive processes embedded in the journal publication activity of Science Academies.

Objective of this paper is to propose a mathematical technique to estimate an index that characterises the dynamics of processes adopted by different Science Academies in bringing out a journal.

It is emphasised that comparative study of all aspects of Science Academies or of Journals brought out by them is not the objective of this paper. Journals of Science Academies taken up in this paper are only to illustrate application of proposed mathematical technique to estimate an index of dynamics of Journal publication activity performed by Science Academy using time series data of activity's out put (number of papers contained in the journal).

1 Generic similarity between this science activity and economic activity of industrial production may be noted. Both require financial and other resources and convert an input into value added product for which there is a demand. In the former case the activity is performed by say a Science Academy and in latter by an enterprise such as a firm; former deals with soft information product, the latter with hard product or service. An industrial firm has to introduce change in its working processes to respond to changing demands/ requirements of its customers and economic environment.

We start by examining similarities between dynamics of processes involved in publication of journal and dynamics of processes encountered in climatology, biological and astrophysical studies. This has enabled us to select mathematical technique appropriate for studying the dynamics of journal publication activity. The selected technique of estimating Permutation Entropy (P_E) from activity's time series out put data as an index of dynamics as is explained. Data of journals of four scientific academies and their trends are briefly presented subsequently. Results of P_E's estimated from data of these journals are presented along with observations on working of Science Academies. While concluding findings of this paper, some observations on studies on dynamics of science processes that could be taken up in future are indicated.

Dynamics of Processes and Analysis Methodology

Clearly the process involves interactions to and fro flow of information (feed back loops) between Science Academy that brings out a journal and researchers working in various national and international organisations such as universities and research institutions. The compositions of researchers that interact with Science Academy is not fixed, keeps changing with changes in research priorities and thus feed back to the Science Academy in terms of when paper would be received from where and in what area also keeps changing.

The dynamics of a system in which interactions between constituents of a system change over time and have feed back loops between actions of various constituents are characteristic of a complex non linear system. It is in this sense that process of bringing out a journal by Academies has to be mathematically treated by tools and techniques used in the study of complex systems.

Such mathematical techniques have emerged in similar situations encountered for characterisation of phenomena for example, in non-linear dynamical systems, in climatology, in biological and astrophysical systems. One such technique is to measure the complexity that gets embedded in the system during its functioning in terms of 'Permutation Entropy' (hereafter referred to as P_E). Complexity measurement P_E thus serves as an index to characterise the functioning of the system or the dynamics of involved processes.

Indices in Scientometrics provide comparative information on science and technology performance of individuals, organizations, regions, countries and groups disaggregated into knowledge categories. Like wise, P_E is an index that is meant to facilitate comparison of dynamics of processes in publication of journals adopted by different Science Academies.[2]

Mathematical Formulation

Several methods are available for quantitative studies of complex systems. These are usually based on a system construct or model of architecture of individual constituents and rules governing interaction between the constituents stipulated by a theory (as causal laws in physical sciences or dependent and independent variables in social sciences).

Time series data of a single observable (output of similar tasks performed by different systems) can then be used to verify the validity and applicability of model in different conditions. The possibility of uncertainty or noise entering the operation of the system due to variations in operational or measurement environment is recognized. To quantitatively deal with such uncertainty in a time varying or operational system, it is customary to estimate the probability of possible states the system can attain estimated as Entropy.[6]

Assuming that the observable takes only discrete values, and the various events or states of the system that has M degrees of freedom, the probability p of $X(t)$ can be mathematically expressed as:

$$p = \{p_i : i = 1, 2, \ldots.., M\}, \quad (1)$$

The uncertainty is expressed as Shannon's entropy given by the following equation:

$$S[P] = -\sum_{i=1}^{M} p_i ln\left(p_i\right). \quad (2)$$

If $S[P] = S_{min} = 0$, then the system is deterministic, as it is possible to predict outcome of an event with a probability $p_i = 1$. Further, in the case of a perfect dice the probability, p_i, of observing either 1,2,3,4,5 or 6 is 1/6. Here, the sample space size $M = 6$, and since the outcome of an event of throwing a perfect dice is independent of any other event, the probability is uniformly distributed over all possible outcome i.e., $p_i = 1/6, \ \forall i = 1,2,3,\ldots,6$. Therefore , in general, if the probability function is uniformly distributed over various degree of freedom i.e., $\left(p_i = \left\{\frac{1}{M} \forall i = 1,\ldots,M\right\}\right)$ then the maximum uncertainty of the system is $S[p] = S_{max} = \ln M$. In such a case minimal information is available for the system.

In general, many large systems in nature have long-range connectivity amongst constituents; characterization of the functioning of such systems on the basis of understanding of individual constituents is not possible.[7-9] To characterize such systems algorithms have however been developed during the last two decades that decipher complexity parameters such as, fractality, Lyapunov exponents, etc. from time series data of a single variable; these algorithms, compared to earlier methods, are more robust against noise in the data.[10-12]

In recent years symbolic dynamics has emerged as an alternate approach to deal with such non-liner complex systems encountered in many physical, biological and social systems.[13] Symbolic dynamics links raw time series measurements to a series of 'discredited' or 'discarded' symbols or patterns that the output generating dynamics embeds over time in the system.

We adopt Symbolic dynamics approach to characterize complex systems that publish Journals. In addition to it being independent of knowledge of constituents, models of interactions between them and prior assumptions regarding probability distribution function of output resulting from interactions, it promises to capture hitherto neglected possible feature of 'self learning' capability expected to get embedded in a system or organization involving human intelligence; 'learning' from out put(s) of an event or events,[3] may modify the probability of occurrence of next and subsequent events in such a way that over time, the number of probable states or patterns get discarded moving the dynamical behaviour of the system to a higher 'deterministic' level.

Bandt and Pompe[14] have provided a method of deciphering symbols or patterns in terms of a 'complexity index' by estimating 'permutation entropy' (P_E) from time series. Unlike estimation of Shannon entropy the method does not require assumptions, models and theories and uses only information on temporal relationship between neighbouring values in a time series data.[15] Further, the

2 Somewhat related project recently launched for ranking of different Book Publishers using Book Citation index 2009-2013 may also be noted[5]

3 It should be noted that 'learning' is not always in coded form but more often in tacit form embedded in people involved in operations of the system.

parametric behaviour of is quite analogous to Lyapunov exponents and is therefore structurally robust to noise in a real world time series.

Representing increase in neighbouring values in the data as 0 and 1, the dynamics of the system may lead to two states X_1 and X_2:

$$X_1 = \{1,0,1,0,1,1\} \text{ and } X_2 = \{1,1,1,0,1,0\}$$

Clearly several other combinations of 1,0 are possible each leading to a set of possible states. (Classical estimation of Shannon entropy does not decipher the probable occurrence of such states and treats $S[p(X_1)] = S[p(X_2)]$.

Consider the general time series

$$X = \{x_t : t = 1, 2\ldots, N\} \quad (4)$$

At each time $t = q$, a vector composed of the n-th subsequent values is constructed:

$$q \vdash \left(x_q, x_{q+1}, \ldots, x_{q+(n-2)}, x_{q+(n-1)}\right) \quad (5)$$

where n termed as dimension corresponds to quantum of information contained in each vector. An ordinal pattern defined as the permutation

$$\pi = (r_0 r_1 \ldots r_{n-1})$$

of (01......[n−1]) is associated to this vector, such that

$$x_{q+r_0} \le x_{q+r_1} \le \ldots \le x_{q+r_{n-2}} \le x_{q+r_{n-1}} \quad (6)$$

In other words sorting values of each vector in an ascending order, and creating a permutation pattern π with the offset of the permuted values. For instance, if $x = (2, 6, 9, 10, 7, 10, 3)$, then it is readily observed that for n = 3, 2 patterns (0, 1, 2) correspond to $x_t < x_{t+1} < x_{t+2}$; 2 patterns (2, 0, 1) correspond to $x_{t+2} < x_t < x_{t+1}$ and 1 pattern (1, 0, 2) corresponds to $x_{t+1} < x_t < x_{t+2}$ results in permutation entropy $P_E = 1.5219$. Therefore for the time series $\{x_t\}_{t=1\ldots N}$, all permutation n! of π are considered to define the relative frequency.

$$p(\pi) = \frac{\#\{t \mid 0 \le t \le N-n, (x_{t+1}, \ldots, x_{t+n}) \text{ has type } \pi\}}{N-n+1} \quad (7)$$

The permutation entropy P_E is defined as:

$$P_E = -\sum p(\pi)\, ln p(\pi) \quad (8)$$

where for a given order n, the sum is taken over all $n!$ permutations π.

One can then infer that during the same period of operation, systems with lower P_E values would have attained higher determinisms or certainty in giving an output compared to those with higher P_E values.

To apply this approach to the systems that publish Journals, we estimate P_E from time series data of number papers/year contained in journals from different countries. The calculated complexity indices would then allow a comparative study of nature of complexity or level of 'determinism' inherent in the dynamics of journal publishing systems of different countries.

Publication Data sources

In our analysis, readily available time series data of the following English language science journals of four countries has been used:

US: American Journal of Physics (AmJP), established in 1933 by American Association of Physics Teachers under the title the American Physics Teacher (Volumes 1 through 7), the name was changed to the American Journal of Physics in 1940.

Japan: Progress of Theoretical Physics (PTP) founded by Hideki YUKAWA in 1946. PTP terminated in December 2012 and merged into Progress in Theoretical and Experimental Physics (PTEP) in January 2013, as a Journal of Physical Society of Japan.

China: Chinese Journal of Physics (CJP) a peer reviewed journal published by The Physical Society of Republic of China in Taiwan since 1963.

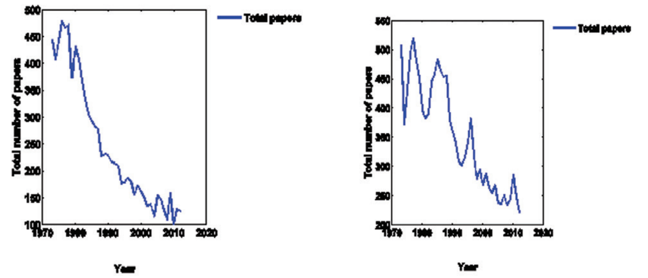
India: Pramana a Journal of Physics published by the Indian Academy of Sciences in collaboration with Indian National Science Academy and Indian Physics Association.

Data consists of total number of papers/year and papers by (i) single author, (ii) double authors, (iii) triple authors, (iv) more than triple authors, contained in journals brought out by the above academies during 1973-2012.

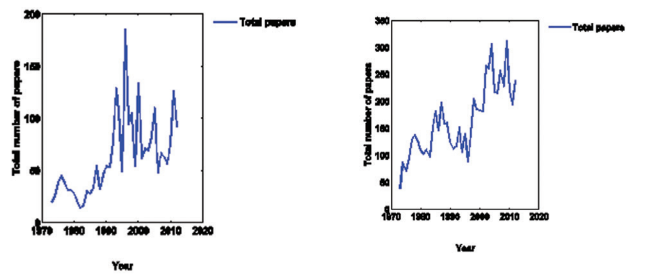
(a)Time variations in total number of papers are shown in Figure 1a-d.

The differences in up and down pattern of variation in different journals are evident.

The pattern of up and down variation in number of research papers by single author (S), double author (D),



(a) Progress of Theoretical Physics (PTP) (b) American Journal of Physics(AmJP)



(c) Chinese Journal of Physics (CJP) (d) Pramana (P)

Figure 1: Time series of total number of research publication.

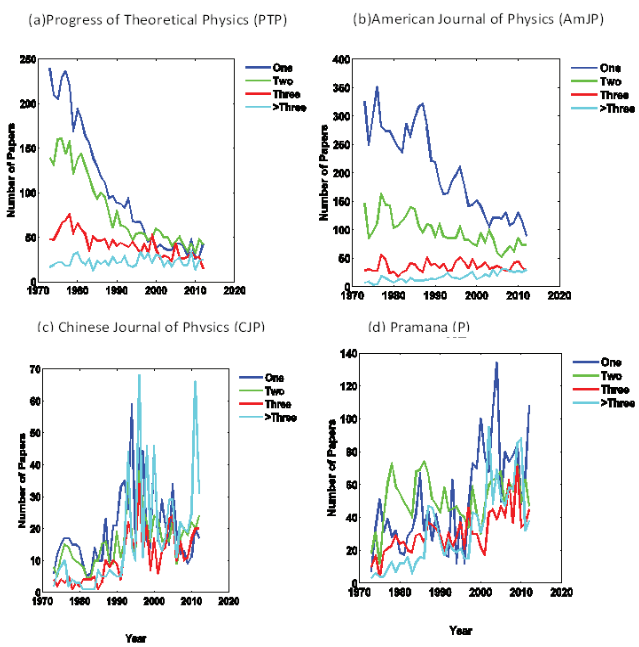


Figure 2: Time series of S, D, T, and M.

triple authors (T) and more than triple authors (M), contained in the journals are shown in Fig.2a-d. Again differences in the patterns of up and down variations are obvious.

(b) Analysis of trend based on pi-chart:
π-Charts of cumulative time series data for time intervals 1973-1982, 1983-1992, 1993-2002 and 2003-2012 are presented in Figure 3-6 to show general trends in S, D,

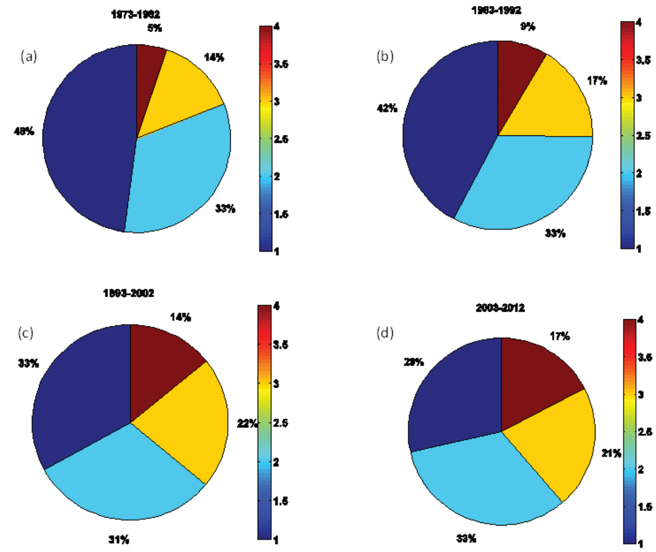


Figure 3: π−chart for the journal 'Progress of Theoretical.

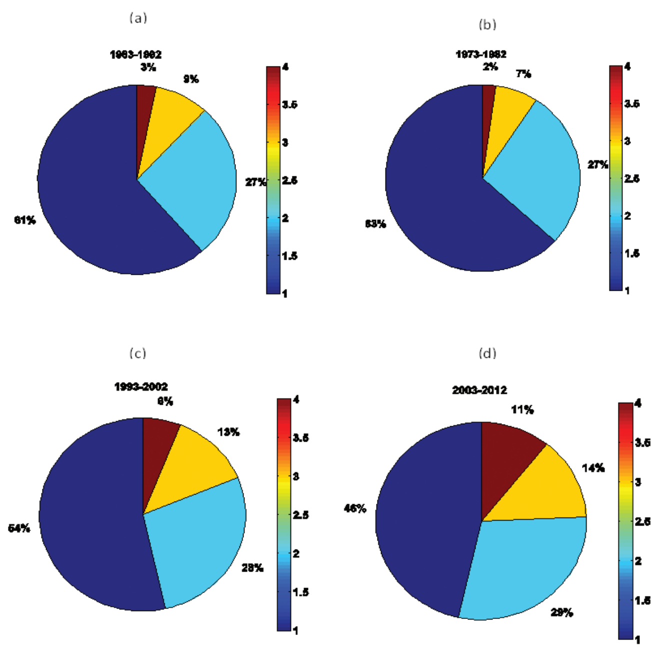


Figure 4: π−chart for the journal 'American Journal of Physics.

T, and M published papers in the above four journals, we construct π-chart.

Results are shown in Figure 5-6.

Data of these journals also confirm the well-known observation on trends in single and multi-author papers. In case of PTP in each decade, the percentage of research articles published by single author decreases, the percentage of research article by double authors remains almost constant at a level of ~33%, while there is a noticeable

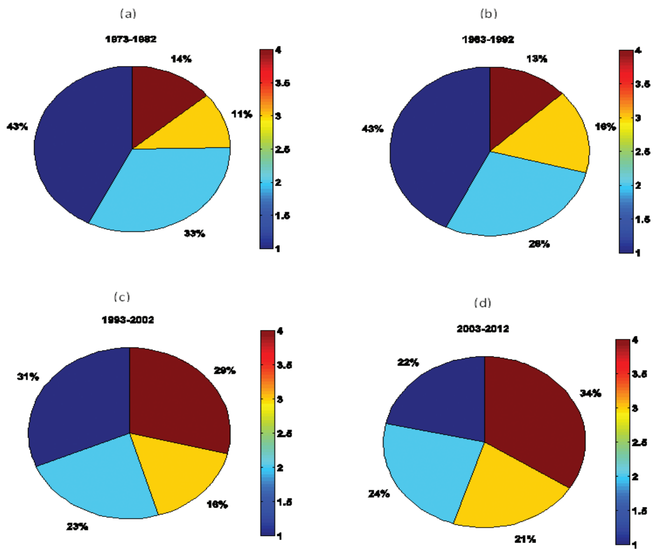


Figure 5: π–chart for the journal 'Chinese Journal of Physics (CJP)'.

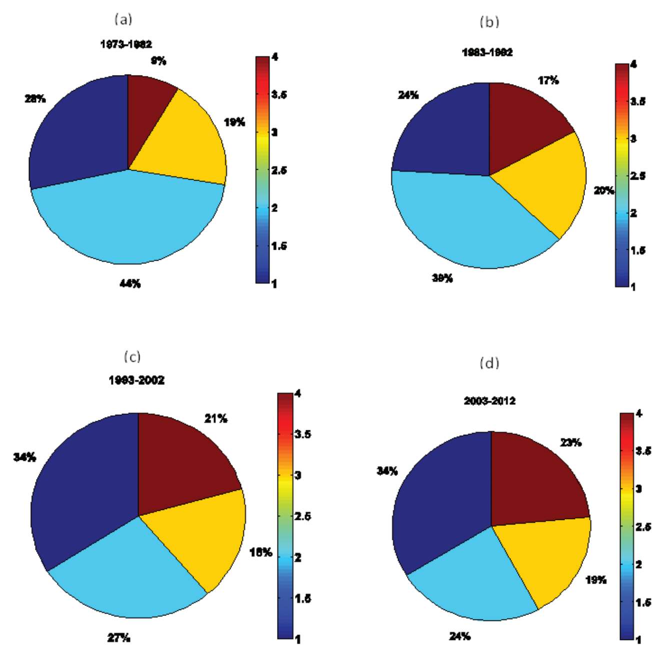


Figure 6: π–chart for the journal 'Pramana (P)'.

increase in percentage of research article by three and more than three authors (Figure 3).

In case of AmJP (Figure 4), a drop in percentage of papers contributed by single author from 63% during 1973-1982 to 46% in the period 2003-2012 is observed but , increase is observed in papers by two, three and more than three authors, a trend similar to that in PTP

In CJP (Figure 5) and P (Figure 6), it is observed that the percentage of research papers due to single author decreases in each decade, whereas the percentage of research papers contributed by two authors decreases marginally. In CJP a growth in the percentage of papers published by three or more authors is observed, but in P the contribution of research papers by three authors remains at a constant level in each decade. Table 1 summarizes the general trend of percentage of research papers published by single author, double authors , three authors and more than three authors in PTP, AmJP, CJP and P in each decade of the period 1973-2012.

Table 1: General Trends in single, double, triple and more co-authored Papers each decade of 1973-2012

Authors	Journals			
	PTP	AmJP	CJP	P
1	Decreasing	Decreasing	Decreasing	Constant
2	Constant	Constant	Constant	Decreasing
3	Increasing	Increasing	Constant	Constant
>3	increasing	Increasing	Increasing	Increasing

Indexing dynamics of processes of publication by Permutation entropy

For deciphering dynamical behaviour of the journal publishing activities of academies, the selected mathematical technique requires capturing only increase and decrease in the number of papers i.e., patterns (Figure 7). We reiterate that the purpose is to compare the dynamics or internal processes of journal publication systems of different academies and not the contents of outputs.

To estimate permutation entropy (P_E), we consider the order $n = 3$, that is three successive data values, then the possible types of motifs or **ordinal patterns** are as shown in Figure 7.

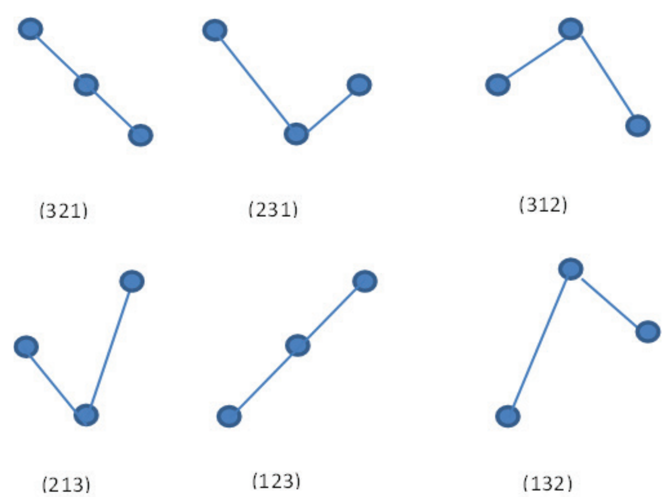


Figure 7: Different ordinal pattern for $n = 3$.

Table 2: Complexity measure of Research Publications					
Journal	**Single Author (S)**	**Double Author (D)**	**Triple Author (T)**	**More than Triple Author (M)**	**Total No. of Papers**
Progress of Theoretical Physics (PTP)	1.5826	1.7326	1.7597	1.7314	1.5143
American Journal of Physics(AmJP)	1.6758	1.6945	1.7026	1.7125	1.6505
Chinese Journal of Physics(CJP)	1.7440	1.7105	1.6954	1.6861	1.7559
Pramana (P)	1.7548	1.6964	1.7515	1.7631	1.7251

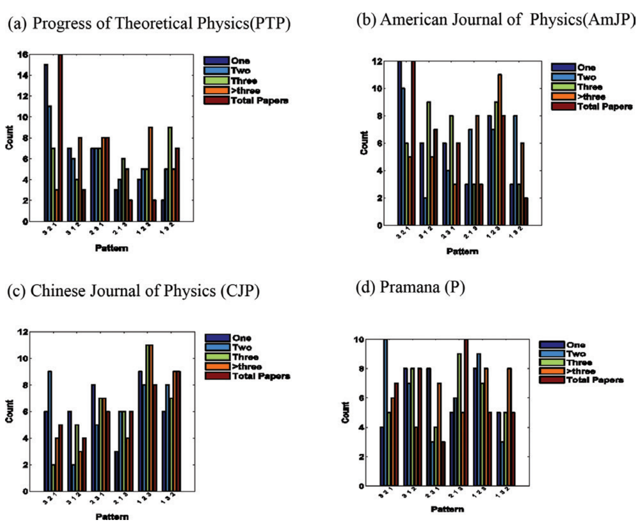


Figure 8: Motifs or ordinal pattern count for S, D, T and M time series in different journals.

The following bar charts are obtained by counting the number of various ordinal patterns in time series of publication data of a journal:

It is quite evident that the count of motifs for S, D, T, and M time series along with that of total number of publication in different journals are different thereby discriminating the inherent processes involved in the dynamics of publications in these journals. Therefore knowing the frequency of different motifs in a time series, using eq. (8), the complexity, P_E, has been computed for various time series of Figure 1a-d and the time series of S, D, T and M of Figure 2a-d. The result of computation of P_E for $n = 3$, are shown in Table 2.

From data on total number of papers/year it is observed that the value is lowest for PTP and highest for CJP. Further, value of for P_E is comparable to that of CJP.

There is an increasing trend in the complexity measure for the time series S, D and T in the journal PTP and AmJP while a marginal/ nominal decrease in the value of is observed in the journals CJP and P.

In case of more than triple author, the value of is lowest in CJP, while the highest value occurs in P. A comparison of values for the time series S in the above four journals suggest lowest value of complexity in PTP.

From quantitative estimate of complexity indices presented in Table 2, it is observed that the processes involved in the publication of research papers in PTP and AmJP are similar and have lesser embedded complexity than CJP and P; in other words the former Academies are better organised than the latter.

CONCLUSION

Adopting an organisational perspective, the activity of an organisation such as Science Academy of bringing out a journal has been shown as that of a dynamically evolving nonlinear complex system, out put of the activity of which is captured as time series data.

In this framework, the computation of from a time series data quantitatively indexes the complexity inherent in the process. The ability of the process to discard over time, a number of 'patterns' of the type shown in Figure 8, decreases P_E.

We note in the last column of Table 2 that both PTP and AmJP systems have over time lowered their values of P_E of compared to CJP and P.

Following the literature on organizational learning[16] ability to discard patterns over time for attaining lower values of P_E may be interpreted as showing the capability of a system to learn from experience.

The results show similarities in dynamics of processes inherent in the functioning of academies bringing PTP and AmJP. Both with lower values of P_E have stronger embedded self-learning mechanisms (it being strongest for PTP), compared to CJP and P with higher values of P_E.

Higher learning capabilities have enabled these academies in Japan and the US to respond to changing requirements of researchers and economic environment. These changes relate to i) researchers now increasingly seek information disaggregated in specialised areas of research and ii) economic dimension of journal publication activity is being increasingly recognised. Academies in Japan and US have accordingly shifted their publication activities away from bringing out papers that cover a range of areas (note decreasing number of papers in Figure 1a and Figure 1b) to bringing out journals in specialized areas of research increasingly and tying up with commercial publishers for the task of printing and marketing of their journals. Better response to emerging trends in single and co-authored papers is also seen in case of PTP and AmJp, a general declining percentage of single author papers, constant level of two authored papers and increasing trend in three and more co-authored papers.

On the other hand one notes similarities between Academies of Taiwan and India in bringing out CJP and P respectively. Both show comparable values less than those of Academies in Japan and US that brought PTP and AmJp. Academies responsible for CJP and P appear to have thus far felt no need to change. Trends in single and co-authored papers also show similarities in Taiwanese and Indian academies.

It may be pointed out that the method of analysis used in the present paper can be applied to time series output data of activities of other systems or organisations such as Universities and Research institutions to characterise the embedded processes in their functioning and thereby drawing inferences on their capability to learn from experience. This would require universities and research organisations to maintain robust time series data of the output of their research activities- A task often not paid due attention in science policies and practices.

The conclusions drawn in this paper indicate that the policy makers need to look more carefully at the learning capabilities of science organisations and institutions so that they are better prepared to respond to changing trends and demands of research. This becomes all the more important as it is recognized that 'the science system' is influenced by, amongst other factors, its historically inherited advantages and disadvantages and capacity of involved actors or agents to 'learn' from working in changing economic and social environment'.[17]

REFERENCES

1. Mingers J, Leydesdorff. A Review of Theory and Practice in Scientometrics. Euro J Op Res. 2015 (accepted for publication) http://arxiv.org/pdf/1501.05462.pdf
2. Freeman C. The Economics of Industrial Innovation.1974: Second Ed., Cambridge (Mass:)MIT Press.
3. Godin B. National Innovation System: The System Approach in Historical Perspective. Science, Technology and Human val.: 2009:34:476.2009; B. Godin (2015), Innovation Contested–The Idea of Innovation Over the Centuries. London: Routledge, 2015.
4. Nelson RR. National Innovation Systems: A comparative Analysis, New York: Oxford Univ. Press. 1993.
5. Daniel T, Nicolás R, Evaristo J, Enrique F. The BiPublishers ranking: Main results and methodological problems when constructing rankings of academic publishers. 2015 http://arxiv.org/pdf/1505.01074.pdf.
6. Shannon C. A Mathematical Theory of Communication. Bell Syst Tech J. 1948:27:379.
7. Shuangcheng L, Qiaofu Z, Shaohong W, Erfu D. Measurement of Climate Complexity using Sample Entropy. Int J Climatol. 2006:26(15):2131.
8. Mitchell M, Complexity A Guided Tour. 2009, Oxford University Press.
9. Morabito FC, Labate D, Foresta FL, Bramanti A, Morabito G, Palamara I. Multivariate Multi-Scale Permutation Entropy for Complexity Analysis of Alzhemer's Disease EEG. Entropy. 2012:14(7):1186.
10. Grassberger P, Procaccia I. Phys Rev Lett. 1983:50:346.
11. Eckmann JP, Ruelle D. Ergodic theory of chaos and strange attractors. Rev Mod Phys. 1985:57(3):617.
12. Kantz H, Schreiber T. Nonlinear Time Series Analysis.1997, Cambridge University Press, Cambridge.
13. Daw CS, Kennel MB, Finney CEA, Connoly FT. A review of symbolic analysis of experimental data. Phys Rev E. 1998:57:2811.
14. Bandt C, Pompe B. Permutation entropy-a natural complexity measure for time series. Phys Rev Lett. 2002:88(17):1741.
15. Zanin M, Zunino L, Rosso OA, Papo D. Permutation Entropy and Its Main Biomedical and Econophysics Application: A Review. Entropy. 2012:14(8):1533.
16. Argyris C, Schön D. *Organizational learning: A theory of action perspective,* 1978, Reading, Mass: Addison Wesley; *Organizational learning II: Theory, method and practice,* 1996, Reading, Mass: Addison Wesley.
17. Slavo Radosevic, Esin Yoruk, Scientometrics. 2014;101:1344. org/pdf/1501.05462.pdf

How to cite this article: Jain A and Das MK. Dynamic Patterns in Processes of Science Systems: Science Academies & their Journals-An illustrative example J Scientometric Res. 2016;5(3):189-196. Full text available at http://www.jscires.org/v5/i1