Bradford's law of scattering for climate-friendly technologies and metainformational effect of time

Irina L. Zerchaninova

The well-known Bradford's distribution of scientific publications on a subject throughout three zones of scattering is associated with the relationship between past, present and future in the development of the subject. The conducted entropy-based analysis of this relationship on climate-friendly technologies showed an effect of potential interest to other subjects and research fields.

Generally, the empirical Bradford's Law of Scattering states that the publications in periodicals on a particular subject can be divided into three zones of scattering, where the 1st (nuclear or the core), 2nd (central) and 3rd (external) zones cover periodicals respectively of the most, larger and lesser productivity to the subject, each zone contains roughly the same sum of publications, the ratio between numbers of the periodicals in these zones is close to 1 : q : q^2 (ideal), q is called Bradford multiplier, q → 5[1]. There have been created a variety of mathematical models and concepts for this law[2-8]. Among them, the Brookes' model (which will be referred next) illustrates the exponential character of the 1st zone and linear one of the next if to plot the ranks of periodicals in descending order of productivity logarithmically along the horizontal axis and related running sums of publications along the vertical axis. However, because the 3rd zone is a zone of reduction and the sum of the publications in the leading periodical (the 1st point on the plot) is considerably above the following, the whole graph bears a resemblance to an extended logistic function or a spiral unit. In practice, these zones' unbalance type is used to briefly outline the type of subject development. The heavy core gives evidence of a conservative area's, the dominant external zone counts in favour of dynamically progressing character, the most stable (harmonious) development type is confirmed by the ideal ratio. The ratio close to ideal indicates the evolutional (after birth and establishing, and before exhaust/ transformation) phase of development. The Bradford's law is true if the selected publications compose a unified area, and backwards. This law (similar ratio) is extended to sets other than ones of scientific publications in periodicals.

Actually, the Bradford's ratio close to ideal is rather hard to attain because it requires not only the evolutional phase for subject development but also general samples. Identifying the 3rd zone for a new area is sometimes a separate scientific task in view of the high degree of scattering. Researchers often divide the sets into three approximately equal parts and calculate the deviations. To date, the law has not justified theoretically, though attempts were made to connect it with the area's information entropy, and some studies revealed that the ideal ratio would be correlated with its maximum. The Brookes' model in fact discloses the relationship between cumulative amount of descriptions of elements and maximum allowable information entropy of this amount. Because the 1st, 2nd and 3rd zones reflect respectively the basic (formed), most intensive (actively forming) and new (forming and possibly forming) directions in the development of a subject to a great extent, these zones can be associated with the past, present and future of this development, but with caution to interpretations in the borderlines. The information entropy of a previous zone is fatally less than one of the followed, however, the "past", "present" and "future" are formed in interacting with one another, the "past" is also uncertain, and the knowledge is objective and comprehensive. It bears a general resemblance to time's arrow in the context of many discussions. It is interesting and has a potential for further studies, specifically, if to find subjects at the evolutional development phase, for which collecting general samples is really possible. One such a possibility was provided appreciably owing to the author's participation in the IEA/OECD Greenhouse Gas Technology Information Exchange (GREENTIE) and the EC-FP6 renewable energy (green) certification project RusTREC.

Implementing RusTREC included a call to briefly outline the development type of technological areas at the intersection of different climate-friendly technologies, namely, renewable energy (RE) & energy efficiency (EE), RE & targeted greenhouse gas (GHG) mitigation. The goal originated from doubts about the necessity of mechanisms to stimulate RE in Russia beyond EE and Kyoto's lines. Resorting to the Bradford-Brookes' model showed that the area of EE technologies had formed and was evolving as conservative (buildings envelope materials and the like in the core), while its subarea toward GHG mitigation (EE projects specially demonstrated climatic effect) could be defined as nearing harmonious, where RE technologies as primary did not play a sizable role. The RE technologies toward GHG mitigation demonstrated within RE projects had not form a unified area yet, so did the existing GHG mitigating technologies. At that time, the set of RE technologies demonstrated within RE projects proved both integral and dynamically progressing with the heading wind energy, while demonstrated RE technologies of higher efficiency (efficient) were still approaching to become a unified area, in which bioenergy ones presented the most and already established a conservative or slow progressing (sub)area with forest waste burning as leader[9]. These estimates were given for the year of 2006, based on the open data available through GREENTIE (IEA/OECD countries'), and used as a pro argument to RE specific stimulating mechanisms[10]. Indeed, the RE support in the taken countries has illustrated the effectiveness of these mechanisms under the fact that RE technologies objectively presented a minor part within EE and Kyoto's projects not formed conceptually.

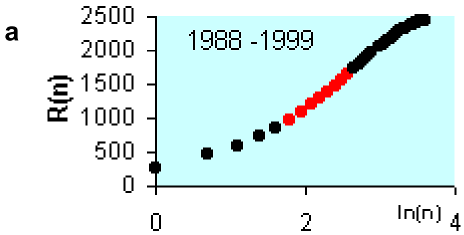
Among sets for that research three ones turned out close to ideal in view of satisfiability of the Bradford's ratio. These sets (a,b,c) and another one for comparing (d) are plotted in Figure 1, related rank-ordered data can be seen in tables 1-4, where d is amount of descriptions of technologies within demonstration projects (abstracts), f is number of technological directions containing indicated amount (or their frequency to indicated amount). The Bradford-Brookes' model was to be adapted to the used data[11-14]: technology descriptions in project abstracts against scientific publications in periodicals, technological directions (rubrics of technological classifications) against periodicals, about 20 ÷ 150 directions in new technological areas against about 150 ÷ 1000 periodicals. The findings were[15]: conjunction point between exponential and linear curves (it was denied by Brookes), k (slope of linear curve) → D_c (cumulative amount of descriptions of demonstrated technologies in the core), q → e (2.718282), the same sign (plus or minus) for empirical deviations Δk and Δq. These findings were obtained for the three points related to the close-to-ideal cases (Fig. 1) ($R^2 = 0.9999 \div 1$), confirmed by the less ideal cases, and deduced using elementary calculations to define $k \approx D_c/1$ (number of directions in a small core $1 < \ln(N_c) < 2$), as well as the Kozachkov's equation $q = e^{Dc/k}$ (modified designations) under his assumption of linear character of the Bradford's distribution at the area's development stage when single leader is replaced by collective, together with other equations of his, according to which e is lower value of q if to divide a set into three zones of scattering.

Additionally, the dependence between information entropy of the core $H_c = - \sum d(n)/D_c \ln(d(n)/D_c)$, Bradford multiplier q and maximum allowable information entropy of technological area ln(N) was explored, basing on the three points for the same close-to-ideal cases; the found formula was: $H_c q = \ln(N)^2 - 4$, where N is total number of technological directions of the area. In comparison with the standard formula for geometric progression $N = N_c (q^3 - 1)/(q - 1)$ (see tables 1-4), it gives q valuations as q(a) = 2.1 to 2.1, q(b) = 2.4 to 2.4, q(c) = 3.14 to 3.25, q(d) = 4.31 to 4.22, which can be referred to good approximation. This formula was arrived at on the areas, standing at the evolutional development phase, and under the samples close to general. The performed search for an ideal case allowed to assume that ln(N) → 4, $N_c = 5$, linear curve intersects the horizontal axis at the point 0.61, $k = 3.39D_c/3.39$, $k_c \approx 0.39D_c$ if to approximate R(n) for the core (the 1st zone of scattering) by a separate linear curve. Hence, ideally, $H_c q = \ln(N)$, and the formula for practical calculations (considering deviations) could be obtained through $½(4) = (H_c q)/2 = \ln(N) - 2$. The formula for the ideal case is also possible to be deduced from the cybernetic Law of Requisite Variety, as $\ln(H_c) = \ln\ln(N) - \ln(q)$, where lnln(N) is disturbance variety, ln(q) is regulation variety, $\ln(H_c)$ is minimum regulation output variety at the metaentropic level of description of a subject (technological area, scientific area, or any other subject), which could be looked upon as a subject of a regulation as well.

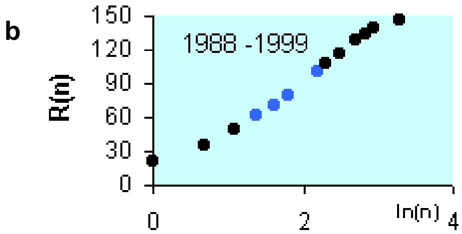
Involvement of metaentropic level enabled to reveal values shown in Figure 2. By defining information through zones of scattering as an information entropy increment, the e logarithm of increment of information entropy lnΔH can be related to metainformation about the subject, which is obtainable from a zone at the entropic level of subject's consideration, and the increment of e logarithm of information entropy ΔlnH can be related to metainformation about the same subject, which is obtainable from a zone at the metaentropic level of subject's consideration, where $\ln\Delta H = \Delta\ln H = \ln H_c - 0$ for the 1st zone. The illustrated values are caused by common attributes of logarithms, the entropy is implicit information, and not only technological classifications may consist of 20 ÷ 150 directions at least in the top classification rubrics. Therefore, the following statement is credible: if descriptions of elements pertaining to a one subject can be integrated into N number of directions so that directions ranked in descending order of productivity to the subject arrange three zones by approximately equal amounts of such descriptions, and zone's direction numbers form geometric progression with multiplier equal to e logarithm base, it is impossible at the entropic level of subject's consideration to obtain tangibly nonzero metainformation about the subject from the 3rd zone, but it is possible at the metaentropic level to obtain this metainformation as close to metainformation from the 1st zone, which is coincident for both levels and tends to ln(lnN/e).

Further examination of the cases presented in Figure 1 (tables 1-4) in view of this statement has indicated values, some of which are given in Table 5 (a, b, c, d are identical). It is seen that roughly the deviations from the ideal proportion at the entropic level could be compensated at the metaentropic level. As example (set a), -0.07 + 0.47 = 0.40, -0.26 +0.26 = 0. With regard to the interpretations above, another statement can be done: if a subject is being at the evolutional development phase, it is hardly possible to obtain metainformation about a future of its development, stationing at an informational level, and it is possible to obtain at related metainformational level as metainformation about the future as metainformation about the past, while the latter is equal at both levels. Each statement reflects the same and may be called as a "metainformational effect of time". Applications of this effect would be useful for technological monitoring, particularly, on GHG mitigation, energy efficiency, renewable energy. Besides, there are parallels between information entropy and thermodynamic entropy, as well as between their adaptations applied in the evolution theory, regulation processes and other fields.

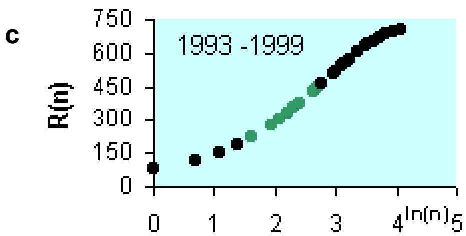
Essential remarks. "5" in Figure 2 is not surely N_c, it might also be q = 5 as median of possible q values accordingly the referred Kozachkov's equations (equal to the Bradford multiplier for periodicals). On the assumption that accessible H_{max} for the considered distributions is slightly less than ln(N), some ideal proportions will be still closer to the golden ratio's. The quantity of new elements and growth of entropy seem to be limited "at the top". We are dependent of our abilities to perceive subjects, or to distinguish between subjects, actually, operating not with elements, pertaining to them, but with descriptions of these elements. Information is a smoothed-out reality, metainformation is a smoothed-out information, and, probably, so on. What could be seen if to observe us in a development process as "smoothed-out", or under "macroscope"? And, whether it is conceivable that energy as though dissipated at one level is virtually recovered at other level.



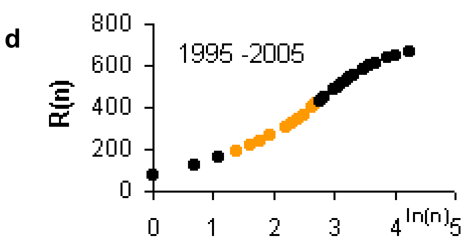
R(n) = 266.8 $n^{0.73}$ for 1 ≤ n ≤ 5
R(n) = 850.4 ln(n/1.86) for 5 ≤ n ≤ 37 $R^2 > 0.99$



R(n) = 20.96 $n^{0.79}$ for 1 ≤ n ≤ 3
R(n) = 47.9 ln(n/1.1) for 3 ≤ n ≤ 27 $R^2 > 0.99$



R(n) = 76.3 $n^{0.64}$ for 1 ≤ n ≤ 4
R(n) = 212.2 ln(n/1.78) for 4 ≤ n ≤ 59 $R^2 > 0.99$



R(n) = 79.6 $n^{0.65}$ for 1 ≤ n ≤ 3
R(n) = 183.4 ln(n/1.5) for 3 ≤ n ≤ 69 $R^2 > 0.98$

Figure 1 Distribution of cumulative amount of descriptions of demonstrated technologies over number of technological directions in descending order of productivity for the technological area: energy-efficient technologies (**a**), energy-efficient technologies toward GHG mitigation (**b**), renewable energy technologies (**c**), efficient renewable energy technologies (**d**). The coloured points on the plots show zones of intensive current development, the non-coloured points before show zones of basic directions, and the non-coloured points after show zones of new directions. In relation to distributions (a), (b), (c), the ratio between zones is close to the Bradford's.

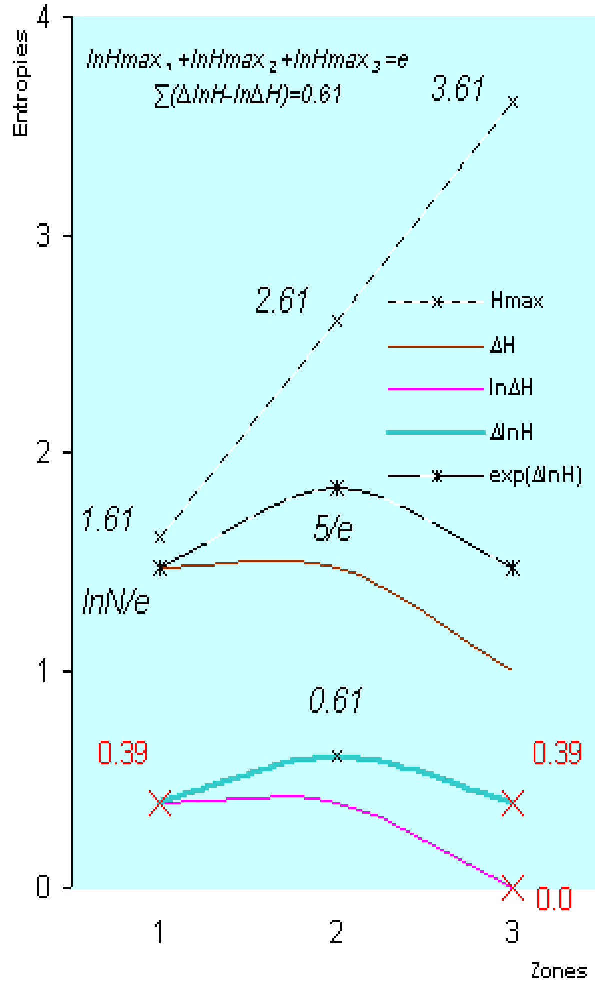


Figure 2 Rounded information entropies toward distribution of cumulative amount of descriptions of elements over number of related directions in descending order of productivity for a subject, if the ratio between zones 1, 2 and 3 is equal to the Bradford's (i), Bradford multiplier q=e, where e is natural logarithm base (ii), and the elements can be intergrated into ($5+5e+5e^2$) directions (iii).

Zones 1, 2 and 3 can be interpreted respectively as "past", "present" and "future" of the development of a subject with the currently available knowledge.

Table 1 Data for Figure 1a

Zones, q≈2.1	f	d	fd	n	R(n)
1	1	263	263	1	263
	1	193	193	2	456
	1	141	141	3	597
	1	139	139	4	736
	1	127	127	5	863
2	1	116	116	6	979
	1	110	110	7	1089
	1	109	109	8	1198
	1	97	97	9	1295
	1	93	93	10	1388
	1	92	92	11	1480
	1	89	89	12	1569
	1	88	88	13	1657
3	1	76	76	14	1733
	1	67	67	15	1800
	1	58	58	16	1858
	1	57	57	17	1915
	1	49	49	18	1964
	2	45	90	20	2054
	1	43	43	21	2097
	1	42	42	22	2139
	1	39	39	23	2178
	1	38	38	24	2216
	1	37	37	25	2253
	1	34	34	26	2287
	1	24	24	27	2311
	1	22	22	28	2333
	2	20	40	30	2373
	1	19	19	31	2392
	1	16	16	32	2408
	1	14	14	33	2422
	1	6	6	34	2428
	1	5	5	35	2433
	1	3	3	36	2436
	1	2	2	37	2438

Table 2 Data for Figure 1b

Zones, q≈2.4	f	d	fd	n	R(n)
1	1	21	21	1	21
	1	15	15	2	36
	1	14	14	3	50
2	1	11	11	4	61
	1	10	10	5	71
	1	9	9	6	80
	3	7	21	9	101
3	1	6	6	10	107
	2	5	10	12	117
	3	4	12	15	129
	2	3	6	17	135
	2	2	4	19	139
	8	1	8	27	147

Table 3 Data for Figure 1c

Zones, q≈3.25	f	d	fd	n	R(n)
1	1	77	77	1	77
	1	40	40	2	117
	1	37	37	3	154
	1	34	34	4	188
2	1	31	31	5	219
	2	30	60	7	279
	1	28	28	8	307
	1	27	27	9	334
	1	21	21	10	355
	1	20	20	11	375
	3	19	57	14	432
	1	18	18	15	450
3	1	15	15	16	465
	3	14	42	19	507
	1	13	13	20	520
	2	12	24	22	544
	1	11	11	23	555
	1	10	10	24	565
	1	9	9	25	574
	4	8	32	29	606
	4	7	28	33	634
	1	6	6	34	640
	3	5	15	37	655
	4	4	16	41	671
	5	3	15	46	686
	6	2	12	52	698
	7	1	7	59	705

Table 4 Data for Figure 1d

Zones, q≈4.22	f	d	fd	n	R(n)
1	1	79	79	1	79
	1	49	49	2	128
	1	33	33	3	161
2	1	29	29	4	190
	1	27	27	5	217
	1	25	25	6	242
	1	24	24	7	266
	2	21	42	9	308
	1	20	20	10	328
	1	19	19	11	347
	1	18	18	12	365
	2	17	34	14	399
	1	16	16	15	415
3	1	15	15	16	430
	1	14	14	17	444
	3	13	39	20	483
	1	11	11	21	494
	2	10	20	23	514
	1	9	9	24	523
	2	8	16	26	539
	2	7	14	28	553
	5	6	30	33	583
	3	5	15	36	598
	4	4	16	40	614
	8	3	24	48	638
	7	2	14	55	652
	14	1	14	69	666

Zones	a		b		c		d	
	lnΔH	ΔlnH	lnΔH	ΔlnH	lnΔH	ΔlnH	lnΔH	ΔlnH
1	0.45	0.45	0.08	0.08	0.28	0.28	0.03	0.03
2	-0.07	0.47	0.04	0.68	0.25	0.68	0.43	0.91
3	-0.26	0.26	-0.23	0.32	-0.02	0.32	0.07	0.35

Table 5 Rounded values of lnΔH and ΔlnH for the technological areas shown in Fig. 1 (data of the tables 1-4)

My thanks to the IEA/OECD EETIC staff, who always answered my questions.

1. Garfield, E. Bradford's law and related statistical patterns. *Essays of an information scientist* **19**, 5-12 (1980).
2. Rao, I.K.R. Classical laws of informetrics: an overview. *Paper of DRTC Workshop on Informetrics and Scientometrics* (16-19 March 1998, Bangalore). 37 p.
3. Ailamazian A.K., Stasj, E.V. Informatika i teoriia razvitiia [Informatics and development theory]. Moscow: *Nauka*, 1989. 174 p.
4. Kozachkov, L.S. Sistemy potokov nauchnoi informatsii [Systems of scientific information flows]. Kiev: *Naukova Dumka*, 1973. 199 p.
5. Kozachkov, L.S. Informatsionnye sistemy s ierarhicheskoi "rangovoi" strukturoi [Information systems with hierarchic "rank" structure]. *Scientific-Technical Information, Series 2 (Information processes and systems) of VINITI [RAN]*, *Moscow*, **8**, 15-24 (1978).
6. Bensman, S.J. Bradford's law and fuzzy sets: statistical implications for library analysis. *IFLA Journal* **27(4)**, 238-246 (2001).
7. Jablonsky, A.I. Modeli i metody issledovanija nauki [Models and methods in research of science]. Moscow: *URSS*, 2001. 400 p.
8. Hjørland, B., Nikolaisen, J. Bradford's law of scattering: ambiguities in the concept of "subject". *CoLIS 2005*, *LNCS 3507*, Springer-Verlag, Berlin Heidelberg, 96-106 (2005).
9. Zerchaninova, I., Kurochka, I., Novikov, A. Some possibilities to stimulate development of bioenergy. *Presentation at 8th International seminar on bio-fuels in clean power production & transport (29-30 November 2005, Moscow).* CD-ROM of *International Science and Technology Center (ISTC)*, also on http://www.etaflorence.it/rustrec (Publications).
10. Renewable energy certificates as instrument to monitor and stimulate renewable energy development in Russia (RusTREC). Inception report. *FP6-2002-INCO-Russia_NIS-SSA-4-SSA, contract 012012.* 2006. 78 p.
11. GREENTIE Directory, on CD-ROM, *IEA/OECD Energy and Environmental Technologies Information Centres (EETIC)*. [Harwell]:Sittard, 1999.
12. CADDET Register, on CD-ROM, *IEA/OECD Energy and Environmental Technologies Information Centres (EETIC)*. [Harwell]:Sittard, 1999.
13. GREENTIE Directory, on http://www.greentie.org/directory/index.php (last update of 2005).
14. CADDET InfoStore, on http://www.caddet.org/infostore/index.php (last update of 2005).
15. Zerchaninova, I.L. Vozobnovliaemaia energetika: globaljnye tendentsii i stimuly razvitiia [Renewable power engineering: global tendencies and development stimulus]. *International Scientific Conference "Global problems of modern power engineering security"* (4-6 Apr. 2006, Moscow). *International Independent University of Environmental and Political Sciences (IIUEPS)*, 444-459 (2006), also on http://www.etaflorence.it/rustrec (Publications).

Irina L. Zerchaninova

Bradford's law of scattering for climate-friendly technologies and metainformational effect of time (2008)

The most frequently asked questions. Part Three[i]

Q Nevertheless, how come that the world exists in the two formats on equal terms, but energy dissipated within "our" format is recovering through electromagnetic radiations within "outside" format?[ii]
A First, our world might "exist" in more number of formats … We may receive some "outside" radiation too.

Besides, if we subtract the entropies of information level from the entropies of metainformation level (two curves at the bottom of Figure 2 plus Table 5 in the discussed paper), we will see that "reserves" of metainformation level are "accumulated" if though they should provide knowing our future more than knowing our past – as if there were some reverse process (moving from right to left). To date, no one of the readers paid attention to the fact that in the given case metaentropies, which were entropies at the compressed level, had their values more than the entropies. Maybe, electromagnetic radiations are to provide such a reserve for reverse process – among various tasks. Maybe, not … However, in any case, the metainformation (metaentropic) level has the real "reserve". And the above mentioned subtraction tends to knowing future more than knowing present, while knowing past is getting equal close to zero.

Q What is semantic information in view of the supposed parallels?
A Semantic information does not leave its carrier[iii], being transferred. It is multiplied by twice. Perhaps, this information can be multiplied to infinity if there are an infinite number of carriers, which are "ready" to record it on. Therefore, those microscopic objects which seem to be smeared-out over space are "candidates" here.

Q But semantic information has no mass?
A If to follow at least the most of concepts of information, any kind of information must have its carrier where carrier is some materialised, while information itself has no mass. However, information may have volume. Moreover, information can vanish. For example, it can be deleted from its carrier. Cases are also possible when information was not managed to be recorded, or all its carriers were destroyed.

Q Do you believe that semantic information is the base?
A Information (idea) is in the beginning of any life cycle. All other is forming around it. And the information role of some microscopic objects would be a credible speculation.

Q What is your attitude to everettics?
A If you replace, for example, the frequencies with the relative frequencies and calculate entropies, you will receive an option of multiverse.

However, you could note that not the only this or that probability is being realised in here … Nevertheless, anyone will hardly be able to deal with more than one potentiality full functional. In such a sense, any probability might be realised. It seems that such an interpretation may be referred to any type of subject.

---

[i] It's not the final part.
[ii] Faq-to-faq.
[iii] Not to confuse with information source. The difference between information carrier and information source looks like the difference between energy carrier and energy source.