

優れた論文とは何か*

—文献引用から見た評価あれこれ—

廣田 勇**

1. はじめに

10年余りに、現在は廃刊となっている科学雑誌「自然」(中央公論社)に『論文の寿命』と題する小文を寄せたことがある(廣田, 1984)。その要旨は、「長年にわたり読み継がれ、数多く引用される論文こそ優れた論文である」との見地から、気象学における論文引用の実態調査を行ない、「論文の生き長らえ方」を論じたものであった。

近年、大学制度改革が急ピッチで進み、大学院重点化や研究所の改組などが行なわれつつあることに伴ない、各組織・機関における「外部評価」や「自己評価」が盛んになってきた。そこで言う「評価」には、組織の構成や運営のあり方も含まれてはいるが、大学の場合、評価さるべき対象は何と言ってもまず第一に研究成果(およびその教育への反映)であろう。

このような状況をふまえ、この小文では、研究成果の客観的評価を行なうひとつの切り口として、「論文の引用」を再び取り上げ、12年前の小論を敷衍する形で、最新の調査も加え、いささかの議論の展開を試みることにする。読者各位のそれぞれの立場から、気象学における研究評価の参考にしていただければ幸いである。

2. 評価の難しさ

ひとつ口に研究成果の評価と言っても、それには実に難しい要素が多々含まれている。まず第一に、個々の論文の価値判断と研究者の業績評価とではその意味あいがかかなり異っている。前者の例としては、ジャーナル投稿論文の査読とか大学院生の学位論文審査などが

あげられよう。これは純粹にその論文の内容だけについての判断が要求される事柄であり、それまでの著者の履歴は殆ど関係がない。これに対し、後者の場合、たとえば科学研究費の申請とか、昇格人事とか、或いは学会賞の選定などの審査基準は、それまでに当人が発表してきたひとまとまりの仕事が、ある時間の経過の中で、その研究分野においてどう貢献してきたかの「積分值」に重点が置かれることとなる。

さてそれではその「積分」をどう実行するのか。発表論文の数を見るだけでは不十分であろう。純粹に個々の論文の内容の「絶対評価」を合算できればよいのだが、それとても常に主観がつかまとう。「あの人は良くやっている」だけでは説得力に欠ける。ことほど左様に評価というのは難しい。

そこで、考えられるひとつの客観的判断方法が「引用による評価」なのである。

3. 定量化の試み

そもそも、研究論文を印刷公表することの最大の意義は、著者の主張を表現することと同時に、その論文に書かれた成果が後に続く研究に様々なかたちで影響を与えることであり、従って、その影響力を具体的に見るひとつの目安が文献引用だということになる。

このことを示すため、ひとつの私案として次のような「影響力公式」を考えてみよう。

$$N(y) = Q \cdot J \cdot M \cdot P$$

$N(y)$: ある論文が y 年後に引用される回数

Q : その論文の絶対的価値(優秀度)

J : 発表の場(ジャーナル)の周知度

M : その研究分野の母集団(研究者数)の大きさ

P : その論文が着目される単純(サイコロ的)確率

ここで Q が求めたい未知数である。 J については、

* On the evaluation of scientific research works.

** Isamu Hirota, 京都大学理学研究科地球物理学教室.

© 1996 日本気象学会

難しい問題は多々あるが、ここではとりあえず、査読制度のある国際誌なら一応ほぼ一定と考えておこう。ただし、シンポジウムのプロシーディングスとか研究費の(和文)報告書だけでは力不足である。Mの母集団の定義は数学、化学、地球科学といった大分野別、気象学、海洋学、地震学のような学会単位、更には、同じ気象の中での力学、雲物理、大気放射といった区分、のどれで見てもよい。Mが大きければそのテーマの論文数も多いから、Nも増えると思われる。ところが一方、同業者が多ければそのぶんだけ競争に勝つことが難しいから、着目される確率PはMに逆比例する。つまり $M \cdot P = \text{一定}$ ということになり、結局上記公式は母集団サイズには関係しないことがわかる。従って、(右辺と左辺をいれかえて)、

$$Q = kN \quad (k = \text{一定})$$

と書ける。言い換えれば、定量化の難しいQ(論文の価値)が、計測可能な引用回数Nによって評価できるわけである(蛇足を加えるなら、「影響力公式」の右辺には著者の年齢性別所属肩書などの因子の入り込む余地が無いことにも注意)。

このような見地に立って、以下の各節では、論文引用の意味内容をさらに詳しく分析することから始め、引用の実態の統計調査に基づく考察、更には研究成果評価法の具体例についても筆を進めてみたいと思う。

4. 引用の性格

言うまでもなく、個々の論文における文献引用とは飾り物ではない。その研究内容に直結した必然性をもって議論の引き合いに出されたものなのである。実際、ジャーナルのレフェリーシートには「文献の引用は適切か」のチェック項目がある。

しかしながら、同じ引き合いに出されると言っても、その意味内容は多種多様である。前掲の旧著『論文の寿命』で試みた「引用の分類」をここで簡単に整理してみると；

(A) 時流型：最近流行のテーマ或いはプロジェクト等に関連したもの。同工異曲の論文が数多くあり、印刷後すぐに引用される度数は多いが、競争相手が多いため淘汰も激しく、長続きするものは少ない。

(B) 集成型：数多いA型論文の問題点は何年後に少数の優れた論文によって一応の決着がつけられ、以後引用されるのはこの集成型に限られる。このタイプには、総合報告や教科書も含まれる。この種の文献の読まれ方は10年以上に及ぶと見てよい。

(C) パイオニア型：ある問題(或いは現象)を初めて手がけた論文。たとえその内容が素朴で不完全であっても後のちまで引用される頻度は大きい。事実の発見・記述、技術手法の開発、概念の提出など、術語に原著者の名が冠せられることも稀ではない。

(D) 温故知新型：発表当時は時流に乗らずとも、後年になってその先見性が発掘・再評価され、新たな脚光を浴びるもの。数十年の時間差のある例も決して珍しくない。

(E) 古典的名著型：上記(C)のパイオニア型を更に古く遡ったもの。気象学の場合、19世紀の古典物理学論文(たとえば Rayleigh, Kelvin, Helmholtz など)の引用が時折見られる。

このような分類からわかるように、引用による評価においては、単に被引用度数のみではなく、歳月の重みをかけた価値判断基準が大切である。換言すれば、それぞれの論文には寿命があり、長寿を保っている論文がすなわち優れた仕事である、と考えてよからう。

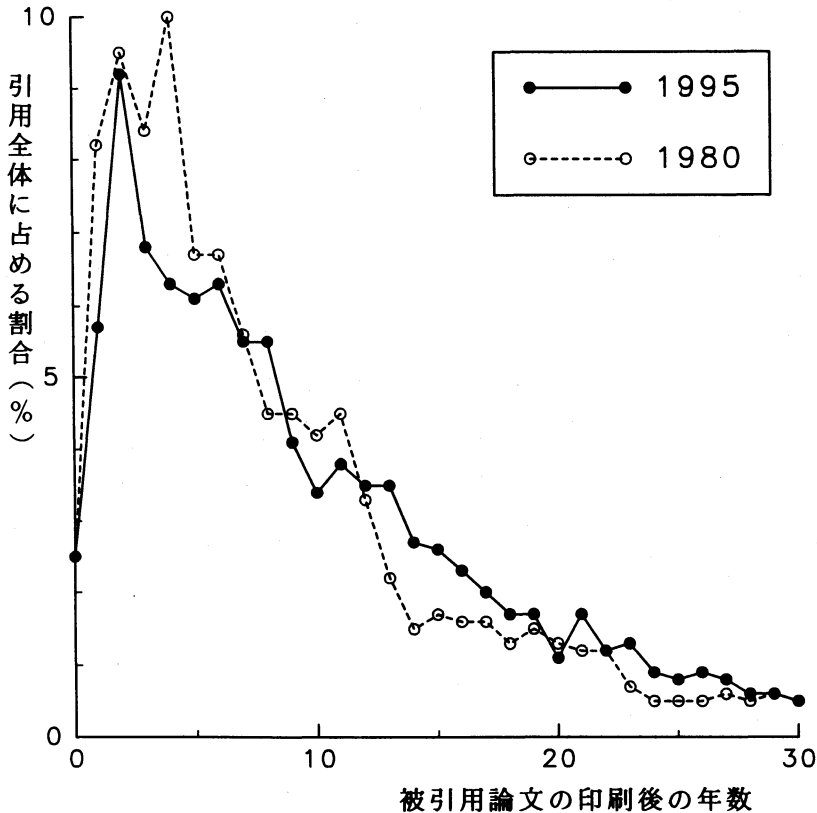
5. 文献引用の実態

気象学分野における引用の実態調査として、『論文の寿命』では、代表的国際論文誌のひとつであるアメリカ気象学会誌 Journal of the Atmospheric Sciences (略称 JAS) を選び、その1980年について、掲載論文225編の参考文献総数4,686編の年号分布を調べ統計を行なった。今回は、それに加えて、同誌の最近年(1995年)について同様の統計を行ない、両者間の異同の検討をも試みた(因に1995年の場合、掲載論文数303、引用論文総数9,417である)。

第1図にその統計結果を示す。横軸には引用された論文の発表後の年数(つまり論文の年齢)、縦軸にはその年齢ごとの引用数を全体に対する%で表示してある。図中の白丸が1980年、黒丸が1995年の数値である。

この図を見て、まず第一に言えることは、2本の線が大まかな点で良く似ていること、すなわち、発表後数年の論文の引用が圧倒的に多く、それ以後は急激に減少していることである。ここには図示はしないが、この縦軸を対数目盛にして書き直せば、両者とも(多少のバラツキはあるが)ほぼ直線に乗っているので、引用頻度は、「論文の年齢に応じた指数関数的減衰」であることがわかる(その減衰率は、大まかに言って、約20年で1/10となる見当である)。

年齢の若い論文の引用の殆んどは、前節の分類で言えば(A)の「時流型」に当たると考えてよい。これに



第1図 JAS論文における引用文献の年別比率。

対し、年令が10年以上の引用率の積算は両年の平均で34%であるから、上記の「集成型」や「パイオニア型」などが引用全体の丁度1/3を占めていると言える。

ところが、この図をもう少し細かく見れば、1980年と1995年の間にいくつかの相異点が見出される。

まず、時流型引用と判断される過去数年の論文の引用ピークが、1980年の場合2～4年の幅にわたっているのに対し、1995年の例では2年目にのみピークが目立ち、3年目以上から急速に落ちこんでいる。このことは、論文1編当りの引用文献数の単純平均値が、1980年では $4686/225 \approx 21$ であるのに対し、1995年では $9417/303 \approx 31$ 、とかなり大きな差を示していることと無関係ではないように思われる。つまり、最近の傾向のひとつとして、「出版直後の論文を数多く引用するような」、更には「目先のテーマに追従しているような」研究論文が増えていることを意味するのではないかと危惧される。これは、引用云々以前の問題として、最近の論文には書く側に「時流型」が多い、ということであり、誤解を恐れず敢えて言うなら、そこに最近の「ブ

ロジェクトばやり」の風潮の反映すら感じられるのである。

一方、その反面、絶対数は必ずしも多くはないが、論文年令が15年前後の引用率はむしろ1995年の場合のほうが大きいように見える。この事実から、時流型とは別に、ある程度古いテーマをじっくり考え直そうとする「温故知新型」の研究も健在だ、と考えるのは好意的解釈に過ぎるであろうか。

いずれにせよ、この1枚の図からだけでも様々な特徴が読み取れることを御理解いただけるものと思う。

6. 引用による評価

ここまでの話は、主として文献引用の実態を全体的にながめたものであった。

それでは次に、いよいよ引用を通して個々の研究者(或いは研究論文そのもの)を評価する、という問題に踏み込んでみよう。この議論は、「他者評価」と「自己評価」の両方に共通であることを予め注意しておきたい。更にまた、「評価」という言葉(考え方)にどうし

でも馴染めないと言うのであれば、それを「学問の自律性」と読みかえても良い。私の経験から見て「評価を軽視する人は学問的レベルも低い」と断言して差支えないからである。

まず引用の「仕方」から見ることにする。

毎月手許に届く新着雑誌を開くとき、個々の論文を精読する前に、私はまず、論文題目でふり目をかけ、次いでアブストラクトと主要な数式・図表に目を通し、さらに文献リストを一瞥することになっている。この感覚は、実は昔の商人が新顔の客の「足もとを見る」と言ったのとかかり良く似ている。つまり文献引用リスト（特にその年号分布）はその論文の性格を端的に示しているからである。もし完全に独創的な研究というものがあつたとしたら、引用文献など不要なはずである。現実には、そのような論文など無いが、その論文自身が時流型か否かは文献リストを見れば一目瞭然である。古い年号が多く並んでいたら、それなりに意義のある論文ではないかとの興味がわく。

この見方を私自身の自己評価にあてはめてみるため、これまでに自分が書いたジャーナル論文約60編における引用総数約900編について第1図と同様な調査を試みた。

その結果は、論文年令5～8年のものが2%位づつ多いこと以外は、1980年の分布に非常に良く似ていた。これは、私の引用の仕方、ひいては仕事の性格が、良く言って標準的、悪く言えば平凡、ということになろう。（約900編の年令調べは暗算で行ないつつ2時間弱の作業時間で済んだ。ひまなとき各自で試みられることをおすすめする）。

次に、個人別の被引用に関しては、Science Citation Index（略称SCI）という年刊の出版物を利用するのが便利である。このSCIは、前年に発刊された各種科学雑誌の文献引用リストをすべて整理し、引用された著者名のアルファベット順に、何年の論文が誰によってどの雑誌に現われたかを網羅したものであり、大きな図書館には大抵備えつけてある。（ただし、整理法は論文の第一著者名のみによっていることに注意。従って、以下の数値も第一著者の論文に限定されている。）このSCIを毎年調べれば、様々な客観的診断ができるわけである。

まず、被引用数の積算値の目安として、有馬・金田（1984）によれば、東京大学理学部物理学科教官の定年（60才）までの被引用総数はひとり当たり約490回であるという。ひとりの活動期間を約35年と見れば、1

年当りおよそ15回ということになる。もちろん、物理学と気象学とでは、研究の形態が多少異なるから、単純な比較は出来ないが、一例として、（1980年版のSCIで見たかぎりにおいて）、チャーニー、ロレンツ両巨頭が年間約150回、リンツェン、ホルトンクラスで約100回、という数字は目安のひとつとなろう。日本人については礼儀上触れないが、恥をしのんで自分自身について言えば、ここ数年の年間平均値が約40回、過去30年間の総計は約700回に過ぎない（私の論文のうち、約半数は気象集誌に発表したものであるが、気象集誌の国際的周知度が低いからという言い訳をするつもりはない）。

有馬・金田の報告でもうひとつの興味ある指摘は、研究者の活動力を年令別に見た統計である。同じく東大物理教官の平均像として、1年間の被引用総数の年令別変化のグラフが示されており、そのピークは30代後半にあつて以後減少している。これは、勢いのある30代前半の仕事が数年後に多くの評価を受けたことと、逆に40代50代になると活動力の低下に加えて、昔の仕事が人々から忘れ去られてゆくことの両面を、恐ろしいまでに立証した統計結果であると言えよう。

良く言われるように、気象学（ひろくは地球物理学）は、物理学や数学とはちがって、経験がものをいう性格を含んでいるため、もし同種の統計を気象学について行なってみたとしたら、活動力のピークはもっと高年令側にシフトしているであろうと推測される。

この節の最後にもうひとつ、引用の価値は回数のみでは測れないという人のために、再び寿命の話に戻そう。論文を書きはじめてから十数年以上のキャリアのある人ならば、各自の論文の引用のされ方について第1図の上に重ねたプロットが出来るはずである（ただし、ここ数年以内に書いた論文が今から十年後にどう読まれているかは俄かに予測ができないので、年令の若い論文は除かねばならない）。

このような統計を自分の1985年以前の論文数十篇の被引用につき、印刷後10年間に限って調べてみたところ、ピークはやはり4年目にあるが（13%）、5～10年目は殆ど変わらず毎年9～10%を占めていることがわかった。第1図の全体平均にくらべ、自分の論文は（内容は地味であっても）寿命がある程度長いことを確認できたので成果の不足をお許し願おう。

これに対し、複数の論文についての平均値ではなく、特定の個々の論文の永続性に着目する人もおられよう。一例として、1949年のイーディの傾圧不安定理論

についてみれば、印刷後30年の1979年に24回の引用、半世紀近く経った現在でもなお年間約20回の引用があり、まさに文句なしの長寿論文と言える。世間一般では「一発屋」といえばマイナスイメージであるが、イーディのような例は立派すぎるほどの好例であり、その評価はデルタ関数的というよりむしろ長く美しい光芒を残す巨大彗星にも擬せられよう。業績は論文数だけでは測れない、と主張する人は是非自らこの例に倣って貰いたいものである。

7. おわりに

この小論は、客観性を重視するあまり、数字にこだわりすぎたかも知れない。しかし我々は決して「論文書きゲーム」をやっているわけではない。大切にしているのは気象学そのものである。その際、研究(論文)の価値判断とは、本来、「絶対的」なものであることを私自身重々承知している。研究の発想、問題提起、手法、結果、解釈、洞察、それぞれの部分で、流石、と唸らせ敬服させられるような優れた論文というのは確かに存在する。論文とはまさに作品であり、その効果は優れた美術や音楽や文学に接したとき感動を受けることと全く同じと言ってよい。それはある意味で主観的ではあろうが、総合的絶対評価が普遍性を持つに至るとき、そこに自ずと客観性が生まれてくるものであり、時間の経過に伴う歴史的淘汰作用がその実証となる。そう考えれば、本論で述べた定量化の試みも、最終的には総合判断のひとつの手段と見なされよう。

最後に、これから将来にわたって評価の試練を受け

るであろう若い研究者に対してひとこと述べておきたい。まず第一に、研究成果は論文としてどしどし発表しその内容評価を世に問うべきである。若書きで未熟な部分があっても恥じることはない。よしんば玉石混淆であってもかまわない。優れた作品は後のちまで読まれ続け、そうでないものは忘れ去られるだけである。古い駄作が後になって非難されることは決してない。もちろん、だからと言って、当世流行のプロジェクトの末端につながって泡のように消える論文ばかりでは困る。難しいではあろうが、先を見据えた永続性のあるテーマに取り組んでももらいたいと思う。本当に良い仕事は、必ず他の研究者にインパクトを与え、その効果はまた次の自分の研究に何らかの形で反映されてくるにちがいない。本論で述べた引用の実態とは、まさに、その結果を映し出す鏡なのである。

謝辞

本文の草稿に対し、同僚の淡路敏之氏から貴重なコメントをいただきました。原稿のワープロ清書は江戸領子さんにお願ひしました。あわせてお礼申し上げます。

参考文献

- 有馬朗人, 金田康正, 1984: 研究者の活動力を測る—発表論文数と論文被引用数による実績評価の試み—, 科学, 6, 360-365.
廣田 勇, 1984: 論文の寿命—文献の引用に見る不易流行—, 自然, 5, 77-81.