

évaluent la longitude de leur position en comparant leur heure locale à celle d'un méridien de référence [Greenwich, ou Paris pour les Français], qu'un garde-temps embarqué – une montre – est censé indiquer : à raison de 15 degrés pour une heure, leur différence correspond à l'angle entre le méridien de leur position et celui de référence. Reçus à bord quasi instantanément, les signaux horaires permettront de corriger les indications du garde-temps au cours de la navigation, ce qui était impossible auparavant. Le 22 juin 1910, une nouvelle installation, avec des locaux souterrains et une antenne à six fils de 450 mètres chacun, est inaugurée par le Bureau des longitudes et son président Henri Poincaré. La première émission officielle de signaux horaires a eu lieu le 23 mai précédent.

Dès ce moment, Paris peut prétendre à un rôle international de premier plan dans le domaine de l'heure. Encore faut-il que la France consente, sans plus tarder, à entrer dans le système horaire mondial. Même si cet argument, de l'ordre du pari, ne peut être produit à l'appui de la réforme, il est clair que les performances de la station radiotélégraphique de la tour Eiffel sont l'élément déclencheur d'une décision qui, de toute façon, ne peut être encore longtemps différée.

Le 20 juillet 1910, le ministre Gaston Doumergue adresse au président de la commission sénatoriale concernée une lettre qui met fin à un blocage de 12 ans : « Tous les départements ministériels se sont montrés favorables à la mesure projetée, et le gouvernement, qui vient d'en délibérer, a reconnu, sans aucune réserve, l'intérêt qu'il y avait, au point de vue des relations internationales, à adopter un système aujourd'hui admis par tous les pays de l'Europe en dehors de la France, du Portugal et de la Grèce. » Le ministre demande à la commission de déposer son rapport « dans le plus bref délai possible ».

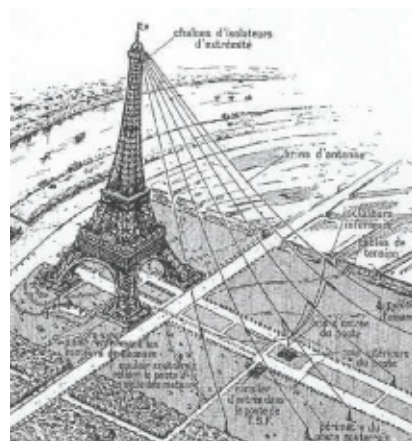
Le géodésien Charles Lallemand, nommé commissaire du gouvernement, défend le projet Boudenoit au Sénat le 26 janvier, puis le 10 février 1911. Entre-temps, une nouvelle rédaction a été proposée, qui vise à limiter l'emploi de l'heure de Greenwich aux chemins de fer, postes et télégraphes, celle

de Paris étant maintenue partout ailleurs. Lallemand n'a aucun mal à faire repousser ce texte qui réintroduit une dualité d'heures et double le décalage horaire dans les gares. Le projet Boudenoit est adopté par le Sénat dans sa rédaction initiale, tandis que la loi est définitivement votée à la Chambre le 9 mars suivant et publiée le lendemain au *Journal officiel*. Son application est immédiate : dans la nuit du 10 au 11 mars 1911, le temps est suspendu à minuit pendant 9 minutes et 21 secondes, tandis qu'enfin disparaît l'heure du chemin de fer.

65 millièmes de seconde d'écart

Le libellé de la loi ne trompe personne, mais il a dû faire sourire outre-Manche. Déclarer explicitement que l'heure légale française est celle de Paris, tout en annonçant implicitement qu'il s'agit de celle de Greenwich, relève de l'exploit ! Certains ironisent à ce sujet, d'autres parlent d'abdication, mais l'opération est un succès total : une *Conférence internationale de l'heure*, réunie à l'Observatoire de Paris en octobre 1912, prévoit la création d'une *Commission internationale de l'heure* dont l'organe exécutif, le *Bureau international de l'heure*, siégera à Paris ; il y sera officiellement établi en 1919. Paris est aussi désignée comme futur *Centre horaire international*, et la station centrale de signaux horaires sera celle de la tour Eiffel.

Le 9 mars 1911, la France a donc rejoint le concert des nations... à ceci près que les 9 minutes 21 secondes ne correspondent pas exactement à l'écart en longitude entre les observatoires de Paris et de Greenwich. En réalité, il s'en faut de 65 millièmes de seconde que la France soit strictement rattachée au système horaire mondial. Cette anomalie ne sera corrigée que par le décret du 9 août 1978 qui abroge la loi de 1911 et légalise en France un nouveau temps international, le temps universel coordonné, fondé sur le temps atomique et connecté à la rotation de la Terre. Ironie du sort, ce temps est élaboré au Bureau international des poids et mesures, près de Paris... et le méridien historique de Greenwich n'intervient plus dans cette affaire.



4. L'INSTALLATION RADIO de la tour Eiffel en 1909. La transmission d'informations – notamment l'heure – pouvait atteindre 3 000 kilomètres, et plus de 4 500 kilomètres en 1910.

L'AUTEUR



Jacques GAPAILLARD est professeur émérite à l'Université de Nantes où il a enseigné les mathématiques et l'histoire de l'astronomie [Centre François Viète].

✓ BIBLIOGRAPHIE

J. Gapaillard, *Une histoire de l'heure en France*, Vuibert, à paraître.

P. Galison, *L'empire du temps. Les horloges d'Einstein et les cartes de Poincaré*, Gallimard, 2007.

✓ À ÉCOUTER

Jacques Gapaillard sera l'invité de l'émission *La marche des sciences* du jeudi 3 mars 2011, sur France Culture de 14h à 15h. <http://www.franceculture.com>

LOGIQUE & CALCUL

Mesurer les chercheurs

La folie évaluatrice dans le monde de la recherche scientifique a provoqué une multiplication des méthodes. L'indicateur de Hirsch est devenu le moyen le plus expéditif de noter un chercheur.

Jean-Paul DELAHAYE

Il est légitime et utile d'évaluer la qualité d'un chercheur. Mais comment le faire quand si peu de gens ont la compétence nécessaire ? Si l'on s'adresse à des spécialistes, on n'échappera pas à la subjectivité, aux querelles d'école, voire à la jalousie. L'évaluation humaine par les pairs restera essentielle, mais selon l'opinion qu'« un chiffre, même médiocre, vaut mieux que pas de chiffres du tout », on a voulu compléter et contrôler le jugement humain avec du concret et du mesurable.

Il est naturel de s'appuyer sur les écrits d'un chercheur pour le noter. Il existe aujourd'hui une multitude de bases bibliographiques qui recensent les articles, les congrès et les livres et permettent, en tenant compte d'une grande quantité de publications, de calculer le nombre *N* de travaux d'un chercheur.

Mais les homonymies, les erreurs de frappe, les obstacles typographiques créés par les lettres accentuées, les traits d'union et les abréviations des prénoms, la non-exhaustivité des bases bibliographiques, tout cela engendre des erreurs qui parfois fau-

sent gravement les résultats [voir plus loin l'exemple de Grigori Perelman]. Par ailleurs, certains journaux sont bien plus exigeants que d'autres, et il serait injuste de comptabiliser de la même façon un article dans une revue qui publie tout texte soumis et un autre dans une revue qui n'en retient qu'un sur 100.

Une solution consiste à classer les revues et publications en catégories et à pondérer les listes des écrits d'un chercheur en fonction de ce classement. Cette solution largement utilisée est mauvaise, car, si le classement des revues est fait par des comités réunis pour cela, on risque de constater la surévaluation des revues auxquelles participent les membres du comité et la pénalisation des revues associées aux écoles et sensibilités concurrentes mal représentées.

La solution du calcul de « facteurs d'impact » des revues ne semble guère meilleure, les méthodes proposées pour les définir étant médiocres, voire absurdes. Le facteur d'impact détermine la valeur d'une revue d'après le nombre de ses articles qui sont cités dans les deux années suivant

leur parution. Ce calcul fixe la valeur de la revue et donc le coefficient associé aux articles qu'elle publie.

Cette méthode est absurde : en mathématiques par exemple, plus de 90 pour cent des citations portent sur des articles publiés plus de deux ans auparavant. Une autre absurdité du facteur d'impact est qu'il est calculé en ne prenant en compte que certaines revues, déterminées subjectivement et en favorisant bien sûr les revues de langue anglaise : le facteur d'impact contient donc la composante subjective qu'il prétendait éviter.

Le contestable « facteur d'impact »

Autre problème dont les spécialistes de bibliométrie ont démontré la gravité : un article dans une « bonne revue » peut être moins bon, par exemple par le décompte des citations qu'on en fait, qu'un article paru dans une revue jugée moyenne ou médiocre. Il n'est donc pas satisfaisant de se fonder sur le classement des comités ou les fac-



teurs d'impact des revues pour juger la qualité des publications d'un chercheur. En 2007, la *European Association of Science Editors* (EASE) a d'ailleurs émis l'avis suivant : « Le facteur d'impact n'est pas toujours un instrument fiable. En conséquence, on ne doit l'utiliser – avec précaution – que pour comparer l'influence des revues, mais pas pour évaluer un article et encore moins un chercheur ou un programme de recherche. »

Une idée plus simple et sans doute bien plus saine est de prendre en compte le nombre total de citations (NTC) faites aux travaux du chercheur à évaluer. Si un chercheur est cité, c'est que ses travaux sont reconnus et appréciés. Plus un chercheur est cité, plus son NTC augmente, plus il devient clair qu'il est efficace, a de l'influence et produit des idées et résultats pertinents. Au moins en première approximation, il semble raisonnable de jauger un chercheur à partir de son NTC.

Calculer le NTC d'un chercheur exige de disposer de bases de données contenant les listes d'articles publiés mais aussi, pour chaque publication, la liste des publications qu'elle mentionne. De l'ensemble

de ces listes on déduit, grâce à des programmes informatiques, le nombre de citations renvoyant à une publication donnée. Les trois principales bases de données sont la propriété d'acteurs privés : *Thompson Reuters* dispose de la base *SCI (Science Citation Index)*, via le *Web of Science* ; *Elsevier* a lancé en 2004 la base *Scopus* ; *Google* propose *Google Scholar* dont l'accès, contrairement aux deux autres, est gratuit.

Compter les citations

Les deux premières incluent surtout des journaux de recherche rigoureusement sélectionnés et mesurent donc l'activité de recherche dans un sens étroit. La base *Google Scholar*, plus large, donnera une idée de l'activité du chercheur en prenant mieux en compte les actes de conférences et congrès (très importants en informatique), ainsi qu'une part du travail de publication didactique. Selon les domaines et selon ce qu'on souhaite mesurer, on choisira l'une ou l'autre, mais toutes les trois sont assez bonnes et donnent des évaluations rarement contradictoires.

Considérons quelques exemples calculés en utilisant la base *Google Scholar* et la page Internet qui en synthétise les résultats (<http://quadsearch.csd.auth.gr/index.php?lan=1&s=2>).

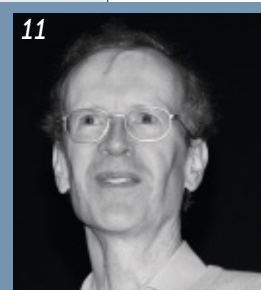
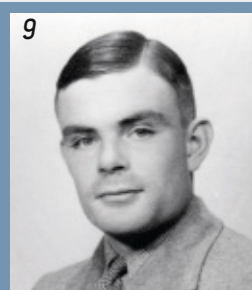
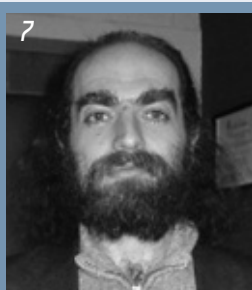
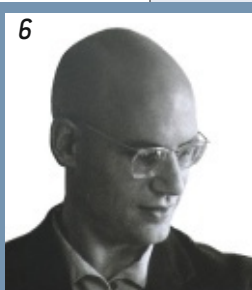
Nous trouvons que les travaux de Paul Erdős ont été cités 15 040 fois, chiffre considérable, mais peu étonnant puisqu'il est sans doute le plus prolifique de tous les mathématiciens du XX^e siècle avec plus de 1 500 articles à son actif (dont 975 sont identifiés par le système). Jean Dieudonné, un des piliers du groupe Bourbaki, a lui été cité 12 580 fois. L'influent Nicolas Bourbaki obtient 16 127. Alexandre Grothendieck, mathématicien français vénéré, a un NTC de 9 110 et Cédric Villani, l'un des médaillés Fields 2010, un NTC de 3 426. Notons que seuls les mathématiciens de bonne renommée atteignent un NTC de 1 000.

Quelques tests et l'examen de ces nombres montrent que le nombre total des citations NTC est une note imparfaite attribuée à un chercheur. D'abord, les problèmes typographiques sont difficiles à éliminer. Une requête pour « N Bourbaki » donne 25 642,

1. Le classement de quelques mathématiciens

Pour chacun de ces 11 mathématiciens renommés, on a noté le nombre *N* des publications repérées, le nombre total de citations (NTC) qu'elles ont reçu, l'indice *h*, l'indice *g* et le taux moyen de citations (TMC). Ces données brutes ont été calculées le 29 décembre 2010 à l'aide du site *Uad Search* (<http://quadsearch.csd.auth.gr/index.php?lan=1&s=2>) qui utilise la base bibliographique *Google Scholar*. Paul Erdős est le chercheur de notre liste qui a le plus publié, mais c'est lui qui a le taux moyen de citations par article le plus faible. Les deux plus célèbres mathématiciens de notre sélection sont sans doute Andrew Wiles et Grigori Perelman à cause de leur démonstration, pour l'un du grand théorème de Fermat, pour l'autre de la conjecture de Poincaré. Or ce sont les moins cités, ceux qui ont le moins de publications identifiées et les plus petits indices *h* et *g* !

	<i>N</i>	NTC	Indice <i>h</i>	Indice <i>g</i>	TMC
1. Nicolas Bourbaki	355	16127	47	125	45,4
2. Alain Connes	329	19892	61	138	59,4
3. Jean Dieudonné	310	12580	39	111	40,6
4. Paul Erdős	975	15040	60	110	15,4
5. Kurt Gödel	215	7471	28	85	34,6
6. A. Grothendieck	136	9110	32	95	66,9
7. Grisha Perelman	20	1643	11	20	82,1
8. Terence Tao	445	13387	49	107	30,1
9. Alan Turing	100	14359	22	100	143,6
10. Cédric Villani	102	3426	28	58	33,6
11. Andrew Wiles	88	3277	16	57	37,2



ce qui est presque le double de ce qu'on obtient avec la requête « Nicolas Bourbaki ». L'écart, qui n'est pas aussi grand pour les autres chercheurs, est sans doute dû au fait que les ouvrages de Bourbaki ne font pas apparaître complètement le prénom Nicolas.

Autre difficulté : une partie des citations comptabilisées pour Bourbaki devrait être reportée sur Jean Dieudonné, puisqu'il est l'un des auteurs anonymes des ouvrages de Bourbaki. Bien sûr, personne ne sait comment opérer ce report ! Le même problème se pose, cette fois à une bien plus grande échelle, pour l'encyclopédie Wiki-

pédia dont l'impact est devenu considérable... et impossible à prendre en compte.

Cédric Villani, le plus jeune des chercheurs mentionnés ici, est défavorisé puisque sa carrière est loin d'être terminée.

Par ailleurs, d'une discipline à l'autre, les usages en matière de publication et de citation sont très différents. Il est donc absurde de comparer des chercheurs de disciplines différentes en considérant leur NTC (voir la figure 3). En mathématiques, par exemple, on publie moins et les articles sont plus souvent publiés avec un seul nom d'auteur, alors qu'en physique des articles ayant plusieurs

dizaines de signataires ne sont pas rares. Ce problème des articles signés à plusieurs se retrouvera plus loin pour l'indice h .

En examinant, publication par publication, le nombre des citations faites à chacune, on remarque que, même chez les plus renommés des scientifiques, nombreux sont les articles à ne jamais être cités. Certes, il est faux de croire que les articles jamais cités sont inutiles, mais je doute que beaucoup de chercheurs sachent qu'une part importante de leurs travaux n'est jamais mentionnée, et cela même s'ils ont été publiés dans des revues prestigieuses.

2. Articles, citations, indice h

Considérons un chercheur ayant N publications $P_1, P_2, \dots, P_N$ à son actif. Notons C_i le nombre de fois que la publication P_i a été citée dans une autre publication. On supposera que l'on a classé les publications en plaçant les plus citées en tête : $C_1 \geq C_2 \geq C_3 \geq \dots \geq C_N$. Le nombre total NTC de citations de ce chercheur est : $C_1 + C_2 + \dots + C_N$. L'indice h de ce chercheur est par

définition le plus grand i tel que $i \leq C_i$. Géométriquement, l'indice h est donc l'abscisse du point le plus à droite du graphe de la fonction $i \rightarrow C_i$, placé au-dessus de la droite $y = x$, c'est-à-dire en gros l'abscisse de l'intersection de la droite avec la courbe (voir le graphique de gauche ci-dessous).

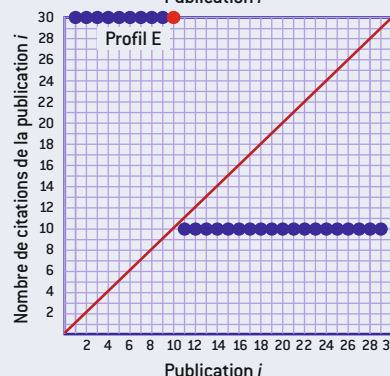
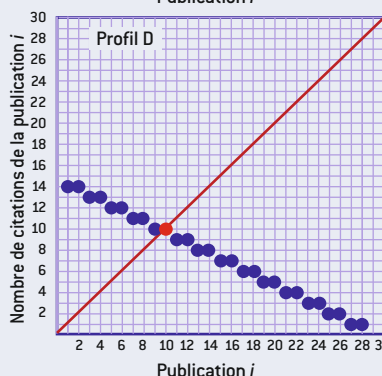
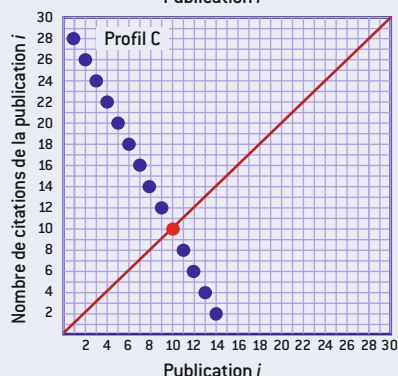
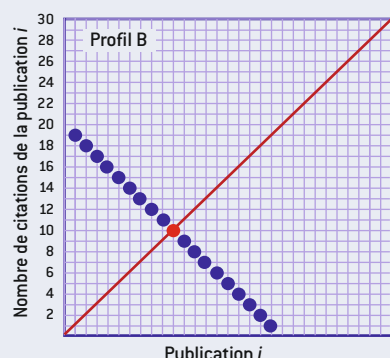
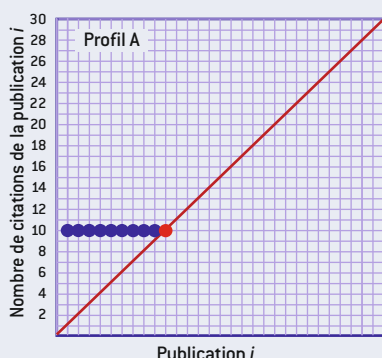
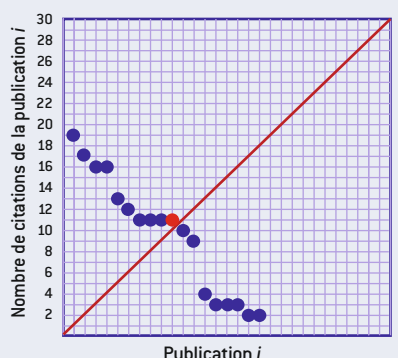
Avoir un indice h de 10 peut résulter de profils très différents (c'est

un reproche fait à l'indice h). Examinons cinq profils.

- Profil A. C'est le « chercheur optimal », il obtient 10 avec 10 articles cités chacun 10 fois. Nul ne peut avoir un indice h de 10 en étant moins cité.
- Profil B. C'est le « chercheur linéaire équilibré » : en moyenne, chacune de ses 19 publications est citée 10 fois et il a autant de citations inutiles (45) avant le point d'intersection de sa

courbe avec la droite $y = x$ qu'après. Son NTC est $1 + 2 + 3 + \dots + 19 = 190$. Comme on le voit ci-contre (modèle A), il peut espérer gagner un point d'indice h tous les deux ans.

- Profil C. Encore un « chercheur linéaire ». Cette fois, 90 des citations de ses neuf premières publications sont inutiles (à son indice h de 10) et 30 autres après le point d'intersection. Conseil : il doit éviter les



Le nombre total de citations NTC d'un chercheur est informatif, mais son domaine de variation est assez large et il a plusieurs défauts qui le rendent inapte à opérer des comparaisons fines. Un chercheur ayant publié un seul travail dans un domaine central ou à la mode et qui, du coup, est très visible et largement cité sera considéré comme l'égal d'un autre qui aura publié une série de travaux difficiles et plus spécialisés, donc moins cités, alors que la contribution scientifique réelle du second peut être très supérieure. Un chercheur publiant des articles de synthèse sera favorisé, même s'il est peu

créatif. Les auteurs de livres seront en général défavorisés car, d'une part, les livres sont moins bien pris en compte que les articles dans les bases de données et, d'autre part, le travail pour écrire un livre est en général bien plus grand que pour un article.

L'indice h de J. Hirsch

C'est sans doute pour cela qu'en 2005, Jorge Hirsch, physicien de l'Université de Californie à San Diego, a introduit une nouvelle idée qui, depuis, est devenue le critère préféré pour évaluer les chercheurs. L'idée de l'indice

de Hirsch a donné lieu à tant d'études et de travaux de bibliométrie qu'aujourd'hui la publication de J. Hirsch de 2005 est citée 1 252 fois alors que sa publication la plus citée en physique ne l'est que 410 fois... ce qui est pourtant très bien !

L'indice h (ou h -index en anglais) de Hirsch prend en considération uniquement les meilleures publications d'un chercheur. L'indice h d'un chercheur est le plus grand entier k tel que les k articles les plus cités de ce chercheur ont chacun été cités au moins k fois. Pour avoir un indice h de 1, il faut avoir 1 publication citée 1 fois. Pour avoir un indice h de 2,

publications trop citées (par exemple en les fractionnant).

- Profil D. À nouveau un « chercheur linéaire », mais ses publications après l'intersection sont les plus inutiles et gâchent les citations (90 pertes après le point d'intersection, et 20 avant). Conseil : il doit se faire citer plus.
- Profil E. Ce chercheur est le plus « maladroit » de notre lot : 200 citations sont inutiles avant l'intersection,

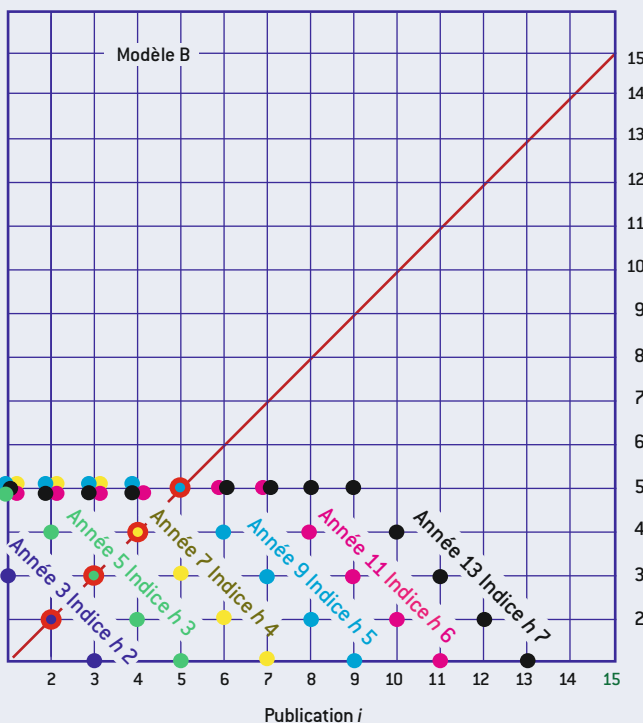
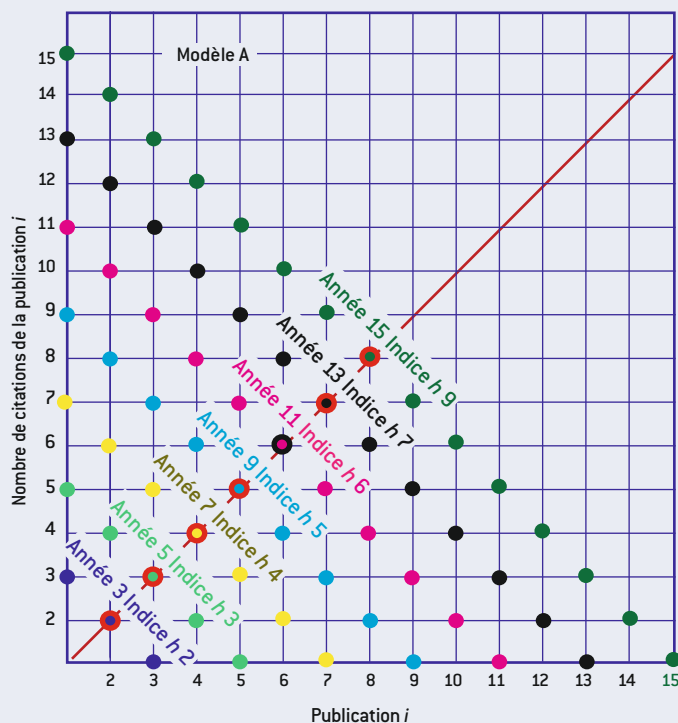
et 200 après. On imagine sans mal des cas pires encore.

Examinons maintenant l'évolution de l'indice h de chercheurs réguliers. Imaginons un chercheur (modèle A) qui publie un article chaque année et supposons que dès l'année où l'article a été publié, il est cité une fois par an exactement (y compris l'année de sa publication). Au bout d'un an, l'indice h

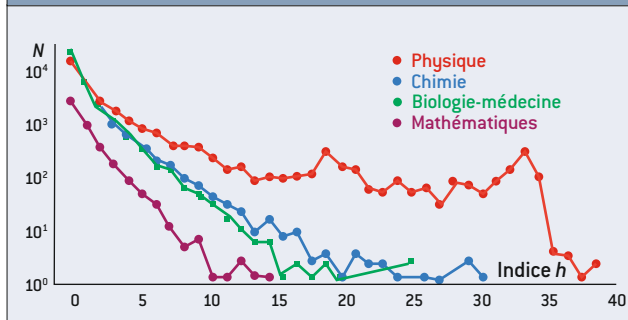
de notre chercheur imaginaire est 1 (il a un article cité une fois). Il lui faut attendre la fin de sa troisième année pour avoir un indice h de 2 et la fin de l'année $2k - 1$ pour avoir un indice h égal à k . L'indice h d'un tel chercheur croît linéairement.

L'hypothèse de citations régulières et indéfiniment prolongées pour chaque article publié est assez optimiste. En effet, les phénomènes de

mode et l'oubli ont pour conséquence qu'après avoir atteint un maximum, un article verra en général l'intérêt qu'il suscite diminuer. Si l'on suppose qu'au bout de cinq ans, on cesse de mentionner les articles de notre chercheur imaginaire (modèle B), cela aura un grave impact sur son indice h qui, après avoir atteint 5, cessera de croître, car aucun de ses articles ne sera jamais cité plus de cinq fois.



3. D'une discipline à l'autre



D'une discipline scientifique à l'autre, les usages en matière de publications et de citations sont très différents. Ainsi, les mathématiciens publient moins d'articles que les physiciens ou les biologistes et publient souvent seuls, ce qui n'est pas le cas en physique et encore moins en chimie. Le nombre de citations que reçoit en moyenne

un article varie aussi sensiblement d'une discipline à l'autre. L'indice h des prix Nobel de biologie vaut ainsi presque le double de celui constaté en physique. Le graphique ci-contre du nombre N de chercheurs brésiliens ayant un indice h donné est tiré d'un article de Pablo Batista *et al.* publié dans la revue *Scientometrics* en 2006.

il faut avoir au moins 2 publications citées au moins 2 fois chacune, etc.

L'indice h est donc majoré par le nombre de publications du chercheur : si vous n'avez publié que 5 articles, votre indice h ne pourra pas dépasser 5. Par ailleurs, si vous publiez beaucoup, mais que vous n'êtes jamais cité, votre indice h sera nul. Autre conséquence (voulu par J. Hirsch) de la définition : si vous avez un très bon NTC, mais que cela n'est dû qu'à un seul article, votre indice h vaudra 1.

Plus on avance, plus il est difficile de faire croître son indice h . Le nombre minimal de citations qu'il faut pour atteindre un indice h de k est k^2 . En général, on constate que le NTC d'un chercheur est 3 à 5 fois le carré de son indice h . Ce facteur est bien supérieur à 5 dans le cas où les meilleurs travaux sont beaucoup cités, ce qui est souvent le cas chez les chercheurs exceptionnels. Pour doubler son indice h , il faut être cité 4 fois plus, et pour le multiplier par 10, il faut que la notoriété des travaux soit multipliée par 100. Cet indicateur composite prend en compte l'intensité du travail d'un chercheur et sa qualité.

Voici un petit tableau pour les cinq mathématiciens mentionnés plus haut :

	N	NTC	Indice h
P. Erdős	975	15 040	60
N. Bourbaki	355	16 127	47
J. Dieudonné	310	12 580	39
A. Grothendieck	136	9 110	32
C. Villani	102	3 426	28

En dépit d'opposants déterminés, l'indice h constitue, selon les spécialistes de bibliométrie, une mesure efficace et assez robuste, qui donne une idée grossière et instantanée de la productivité et de la qualité du travail d'un chercheur. Cependant :

– Il ne faut pas oublier qu'il dépend de la base bibliographique permettant son calcul.

– Il peut y avoir des erreurs dues en particulier aux homonymies et aux difficultés typographiques. Il faut contrôler au moins les k premières publications pour un indice h valant k ; cet examen est plus facile que pour le NTC, car l'indice h ne dépend que d'un petit nombre de publications.

– Il est absurde de comparer des chercheurs de disciplines différentes, car les usages concernant les publications et les citations varient d'un domaine à l'autre. Les facteurs correctifs qu'on a tenté de calculer restent peu sûrs.

– De nombreux biais favorisent ou défavorisent un chercheur. En particulier, on doit prendre en compte l'âge d'un chercheur et le fait qu'il publie seul ou pas.

– Un seul nombre ne peut pas et ne pourra jamais évaluer à la fois la créativité, la persévérance, le rayonnement, l'influence, l'aptitude à la synthèse d'un individu. En conséquence, tout ramener à une seule note ne donnera jamais d'évaluation fine de la personnalité et des accomplissements précis d'un chercheur. C'est essentiel à retenir pour un comité qui doit recruter, attribuer une promotion ou décider d'un financement.

Dans son article de 2005, J. Hirsch a proposé une série d'arguments en faveur de son invention. Selon lui, parce qu'il limite l'importance des *hits* (publications citées de très nombreuses fois), l'indice h donne un classement plus juste des chercheurs que le NTC qui par ailleurs possède un pouvoir discriminant illusoire. De même, il n'a pas le défaut du nombre N de publications (car un chercheur qui publie beau-

coup sans jamais être cité n'est probablement pas un bon chercheur).

L'indice h est un meilleur indicateur que le nombre d'articles cités au moins K fois, qui décompterait aussi les bons articles d'un chercheur, mais dépendrait d'un facteur arbitraire K . Ce n'est pas le cas de l'indice h .

L'indice h d'un chercheur ne peut qu'augmenter. Avec des hypothèses simples sur la productivité d'un chercheur et l'écho de son travail, on aboutit à un modèle théorique de chercheur ayant un indice h qui croît linéairement tout au long de sa carrière, ou alors qui croît jusqu'à un maximum que le chercheur ne dépasse plus (voir la figure 2).

Du pour et du contre

D'après J. Hirsch, l'indice h est peu sensible (du moins quand il n'est pas trop petit) au phénomène de l'autocitation et aux petites erreurs que contient nécessairement une base de données bibliographique.

En examinant par exemple une série de lauréats du prix Nobel de physique, J. Hirsch a constaté qu'ils possèdent tous un indice h supérieur à 20. D'autres tests ont confirmé dans d'autres disciplines que les personnalités les plus en vue ont tous de bons indices h (pour leur discipline).

J. Hirsch a aussi mené une étude statistique pour évaluer la capacité prédictive de l'indice h . Prenant en compte l'indice h au bout de 12 années de recherche d'une série de scientifiques, il a calculé la capacité de son indicateur à prévoir la qualité du travail dans les 12 années suivantes. Son étude semble concluante. Assez étrangement même, la corrélation entre les indices h au bout des

12 années et le NTC au bout de 24 ans est meilleure que la corrélation entre le NTC au bout de 12 ans et le NTC au bout de 24 ans : pour savoir si un chercheur sera bon (au sens du NTC) dans la suite de sa carrière, il vaut mieux se fier à son indice h actuel qu'à son NTC actuel ! Parmi les défauts de l'indice h , il y a son imprécision : c'est un entier, qui interdit de faire des comparaisons fines. Mais il est probable que les nombreuses méthodes qui tentent de donner des notes plus fines (non entières) créent une précision illusoire.

On lui reproche aussi d'être victime de l'effet Matthieu, ainsi nommé à cause de la phrase de l'Évangile selon saint Matthieu : « À celui qui a, il sera beaucoup donné et il vivra dans l'abondance, mais à celui qui n'a rien, il sera tout pris, même ce qu'il possédait. » Autrement dit, on citera une publication d'autant plus facilement qu'elle l'a déjà été beaucoup.

Le plus grave défaut de l'indice h est qu'il incite à découper en plusieurs petits articles ce qui aurait dû être un seul article et à renoncer à écrire des livres qui rapportent trop peu. Son utilisation encourage les pratiques de copinage : on citera largement ses amis et on les fera cosigner, pour peu qu'ils fassent de même. Enfin, son utilisation conduit à choisir des sujets faciles et à la mode qui assurent qu'on sera cité, voire suggère la publication d'articles provocants défendant des thèses auxquelles on ne croit pas soi-même, mais dont on est certain qu'elles feront réagir et donc qu'elles seront citées.

On peut ajouter un autre conseil encore : publiez un article avec une erreur et faites un correctif, vous aurez ainsi deux publications d'un coup : si l'article intervient dans votre indice h , il comptera double puisque ceux qui le citeront mentionneront aussi le correctif. N'oubliez pas non plus que lorsque vous vous citez, il faut choisir ceux de vos articles proches du seuil de votre indice h : ceux trop cités n'ont pas besoin de l'être plus, et ceux trop peu cités risquent fort de ne l'être jamais assez.

L'indice g est un peu compliqué, mais intéressant. C'est le plus grand entier g tel que les g articles les plus cités sont cités au moins g^2 fois au total. Cet indicateur, toujours supérieur ou égal à l'indice h , prend mieux

en compte les *hits* d'un chercheur. Notons un petit paradoxe autoréférent : J. Hirsch aurait mieux fait de définir cet indice g qui le favorise bien plus que l'indice h , puisque son article sur l'indice h est un *hit* dans sa liste de publications...

Oubliés et tricheurs

Pour insister sur la méfiance qu'on doit éprouver à l'égard de l'indice h , mentionnons deux cas extrêmes.

Grigori Perelman qui a démontré la conjecture de Poincaré est l'un des dix plus grands mathématiciens vivants ; pourtant son indice h , calculé avec la base *Google Scholar*, vaut 0. La raison est que celui qu'on nomme Grigori Perelman en Occident, et c'est sous ce seul nom qu'il est mentionné dans *Wikipédia*, est enregistré avec le prénom Grisha ! Notons que même avec le prénom corrigé, l'indice h de Perelman reste faible.

L'exemple de Ike Antkare est inverse. L'indice h de Ike Antkare vaut 95, ce qui en fait l'un des meilleurs chercheurs de tous les temps. Mais cet exploit n'est pas de nature scientifique : Ike Antkare n'existe pas ! Son indice h résulte d'une manipulation délibérée menée par Cyril Labbé, de l'Université de Grenoble. C. Labbé a créé 101 faux articles ayant Ike Antkare comme auteur. Cent de ces articles citaient tous les autres, et le 101^e citait des articles réels déjà référencés par *Google Scholar* pour faire croire que Ike Antkare était bien connecté au monde réel universitaire. C. Labbé a ensuite mis en ligne des pages Internet où ont été déposés ces 101 articles. *Google Scholar*, piégé, a pris en compte ces pages et ces articles... du moins presque tous. Ainsi, pour *Google Scholar*, Ike Antkare a publié 95 articles, chacun ayant été cité la même année au moins 95 fois, et il faut donc lui attribuer un indice h égal à 95.

Cet exploit illustre qu'il est facile de tricher et démontre à ceux qui utilisent l'indice h pour évaluer les chercheurs qu'ils doivent se méfier. Faire de la bonne recherche n'a jamais signifié publier une multitude de petits articles en y éparpillant ses idées pour accumuler des points, comme un automobiliste accumule les points d'une station-service pour gagner des verres à whisky. ■

L'AUTEUR



Jean-Paul DELAHAYE est professeur à l'Université de Lille et chercheur au Laboratoire d'informatique fondamentale de Lille (LIFL).

✓ BIBLIOGRAPHIE

Rapport de l'Académie des sciences, *Du bon usage de la bibliométrie pour l'évaluation individuelle des chercheurs*, 17 janvier 2011 [disponible sur <http://www.academie-sciences.fr>]

N. De Bellis, *Bibliometrics and Citation Analysis : From the Science Citation Index to Cybermetrics*, The Scarecrow Press, 2009.

C. Labbé, *Ike Antkare, one of the great stars in the scientific firmament*, ISSI Newsletter, vol. 6(2), 2009 : <http://membres-lig.imag.fr/labbe/Publi/IkeAntkareSub.pdf>

G. Woeginger, *An axiomatic characterization of the Hirsch h -index*, *Mathematical Social Sciences*, vol. 56, pp. 224-232, 2008.

J. Hirsch, *Does the h -index have predictive power ?* *PNAS*, vol. 104/49, pp 19193-19198, 2007. [arxiv.org/abs/0708.0646].

A.-M. Kermarrec *et al.*, *Que mesurent les indicateurs bibliométriques ?*, Rapport INRIA, 2007 : www.inria.fr/content/download/6425/58287/version/1/.../ce_indicateurs.pdf

L. Egghe, *Theory and practice of the g -index*, *Scientometrics*, vol. 69(1), pp. 131-152, 2006.

J. Hirsch, *An index to quantify an individual's scientific research output*, *PNAS*, vol. 102/46, pp. 16 569-16 572, 2005. [arxiv.org/abs/physics/0508025].

ART & SCIENCE

L'homme descend du crocodile...

... ou comment une statue océanienne conduit à revisiter le darwinisme et les hypothèses de ceux qui s'y opposent.

Philippe CHARLIER

A la fin des années 1990, le marchand d'art David Godreuil parcourt la vallée du fleuve Sépik et sa région, au Nord de la Papouasie-Nouvelle-Guinée. Dans le village de Yenchan, il collecte une grande statue ancienne, encore couverte de quelques restes de peintures végétales et minérales (voir page ci-contre). L'individu figuré, d'une soixantaine de centimètres de hauteur, est un être mythique, un ancêtre du clan, dont la filiation avec le crocodile est visible sous la forme d'un long prolongement caudal en queue de saurien garni de renflements suggestifs. De fait, le crocodile est au cœur de nombreuses croyances de ces populations, une légende affirmant même que le paradis et les enfers correspondent chacun à une partie de la mâchoire entrouverte d'un de ces reptiles.

Cependant, s'agit-il réellement d'un sujet hybride, destiné à rappeler à l'homme son ascendance animale lacustre, ou plutôt de la figuration d'une anomalie anatomique ? Il est en effet possible de voir dans cette statue la représentation d'une véritable queue chez un être humain.

Au cours de leur développement, les embryons humains sont un temps dotés, de façon physiologique et donc normale, d'une queue pouvant mesurer jusqu'à un sixième de leur taille totale. Toutefois, cet appendice s'atrophie au fil de la croissance, jusqu'à disparaître totalement au stade fœtal. Seul persiste chez l'adulte un bourgeon minuscule constitué de une à trois vertèbres coccygiennes (le coccyx), l'ensemble mesurant moins de trois centimètres de longueur.

Mais cette disparition souffre quelques exceptions. En effet, chez certains individus,

pour des raisons hormonales ou génétiques, cette queue embryonnaire peut perdurer (on parle de queue vestigiale) sous des formes et des aspects notablement différents d'un sujet à l'autre. Ce peut être une queue dure constituée de vertèbres accompagnées de tissus musculaires autorisant une mobilité, ou bien une queue molle uniquement composée de tissus vasculo-nerveux, avec parfois une matrice cartilagineuse. Dans certains cas, une forte pilosité recouvre ces queues ves-

Une pilosité importante confère parfois aux queues vestigiales de certains individus un aspect animal.

tigiales, leur conférant un aspect encore plus animal. Les garçons sont deux fois plus souvent porteurs d'une telle malformation que les filles.

Une prise en charge chirurgicale est la règle, l'opération étant le plus souvent effectuée peu de temps après la naissance. Cependant, la situation est plus complexe lorsque d'autres malformations sont associées : une fistule anale (une communication anormale entre deux zones d'ordinaire bien séparées), une anomalie du système génito-urinaire, une lésion vertébrale... De telles observations sont nombreuses dans la littérature médicale, avec des cas décrits en ex-Indochine, en Inde, au Japon, en France, aux États-Unis, etc.

Dans ce dernier pays, le courant créationniste a récupéré ces malformations pour étayer diverses théories fantaisistes. Le point de départ fut un article paru dans le *New England Journal of Medicine* en 1982, où Fred Ledley, alors à l'Institut médical Howard Hughes, aux États-Unis, postulait, à partir du cas d'un de ses anciens patients, que l'observation de queues vestigiales chez des nouveau-nés mettait en évidence « la relation entre humains modernes et leurs ancêtres primitifs ». Les créationnistes ont réagi en affirmant que de telles malformations ne peuvent être que le fruit du hasard (la preuve étant leur diversité) et non le résultat de l'expression accidentelle d'un gène archaïque.

Rappelons que le créationnisme est une doctrine religieuse (surtout chrétienne, mais aussi partagée par quelques courants musulmans et juifs) s'opposant à l'évolutionnisme darwinien. Selon les créationnistes, la vie et la Terre ont été créées par Dieu comme le racontent les livres saints. Toutes les traces plaident pour une évolution des espèces (les fossiles, par exemple) ne seraient que des leurres disposés par Dieu pour fausser le jugement des hommes et les éloigner de la Vérité. Pourtant, l'homme a bien un ancêtre commun avec le crocodile !

Philippe CHARLIER travaille au Service de médecine légale et d'anatomie/cytologie pathologiques, à l'Hôpital universitaire Raymond Poincaré (AP-HP, UVSQ), à Garches.

Ph. Charlier, *Les monstres humains. Les malformations humaines dans l'Antiquité*. Paris, Fayard, 2008.

F. D. Ledley, *Evolution and the human tail : a case report*, *New England Journal of Medicine*, vol. 306(20), pp. 1212-1215, 1982.