

Utilité et limites des indicateurs bibliométriques : l'exemple des mathématiques

Jean-Marc Schlenker

Institut de Mathématiques
Université Toulouse III
<http://www.picard.ups-tlse.fr/~schlenker>

24 Mai 2009

Clé de voute du système académique.

- Les auteurs soumettent leurs articles,
- ils sont envoyé à des rapporteurs anonymes *compétents*,
- les responsables éditoriaux les acceptent ou pas suivant leur apport,
- *hiérarchie* des journaux.

Travail d'évaluation considérable (éditeurs, referees).

Permet de *sous-traiter* l'évaluation scientifique à des experts compétents, éditeurs et referee des revues spécialisées.

Utilisation possible des citations pour évaluer la visibilité des recherches ou (plutôt) des journaux.

Dans certaines disciplines (en France) le système des revues à comité de lecture n'existe pas vraiment.

Deux types principaux :

- basés sur les publications (pondération des journaux),
- basés sur les citations.

Les citations sont inutilisables dans certaines disciplines (maths) : facteur temps. Plus courant en biologie.

Les citations mesurent la *popularité*, les publis pondérées peuvent donner une indication plus qualitative.

La *pondération* des journaux compte : incite à publier beaucoup où à l'inverse à chercher des résultats de premier plan.

Journaux ou citations suivant discipline : bio vs maths.

Les maths ont des caractéristiques bibliométriques spécifiques.

- Revues fortement hiérarchisées.
- Forte hiérarchisation des résultats et large consensus dans la communauté mathématique.
- La publication des articles est lente ($\simeq 3$ ans).
- L'accumulation des citations est très lente ($\simeq 20$ ans) (\mapsto IF faibles des revues maths).
- Les dir. de thèses ne cosignent pas avec leurs doctorants.
- Les collaborations sont relativement rares, peu de co-auteurs.
- Relativement peu de publiés, souvent longues.

Grandes différences avec d'autres disciplines ; dans une moindre mesure à l'intérieur des mathématiques.

Quelques utilisations légitimes pour indications sur :

- institutions (part dans la production mondiale),
- labos/départements (si assez gros),
- études statistiques/économétriques “générales”.

En pratique, les évaluations à grande échelle (e.g. pays) se basent sur des critères bibliométriques “heuristiques”, même quand elles sont réalisées par des panels d’experts compétents.

Les indicateurs ne sont significatifs que pour des populations suffisantes (>> 100 chercheurs.an).

Même dans ces cas des précautions sont indispensables.

- Biais entre disciplines (eg maths vs biologie).
- Biais entre sous-disciplines (eg analyse vs théorie des nombres).
- Bon choix des indicateurs (eg MR vs WOS en maths).
- Pondération. Rôle suffisant des publications majeures : qualité vs quantité.
- Utiliser les journaux et non les citations (en maths) : “durée de vie” des articles très longue.

Le choix d'indicateurs inadaptés (eg WOS) fait exploser les biais et conduit à des indicateurs trompeurs.

Consensus dans la communauté mathématique : pas d'évaluations individuelles bibliométriques.

Elles sont très peu fiables (médailles Fields, vécu des commissions...).

- Les travaux les plus profonds ne se “voient” souvent pas (cf médaille Fields).
- Les citations n'ont une signification que sur la longue durée ($\simeq 20$ ans).
- Les biais entre sous-disciplines sont lourds et *doivent* être corrigés. Risque de déséquilibrer la recherche.

L'évaluation bibliométrique individuelle donne des *incitations perverses*.

- Travaux relativement superficiels, publiés et cités plus vite.
- Passer plus de temps à communiquer qu'à comprendre et chercher.
- Multiplier les articles.

Le problème disparaît (partiellement) pour les évaluations collectives.

Il est assez facile de “truquer” les données bibliométriques. Par exemple :

- Pour les journaux : encourager les citations, retarder les dates de publications.
- Pour les centres : inciter les chercheurs à maximiser leurs indices. Doubles affiliations douteuses.
- Pour les individus : “échanges” de citations, pressions diverses, optimisation des soumissions d’articles, articles jumeaux...

Indicateur de production (maths) *globales*.

Possibles : basés sur le MCQ de Math Reviews. Pondération MCQ² (↪ pondération forte des “meilleurs” articles).

$$\sum_{articles} \frac{MCQ^2 \times \#pages}{\#auteurs}$$

Stabilité/indicateurs : variation entre MCQ², MCQ ou IF. Biais systématiques.

Relative stabilité dans le temps, à partir d'une taille importante.

Biais possibles, à prendre en compte.

Fournissent quand même des indications utiles : comparatives entre pays/centres, dans le temps, etc

Indicateur par pays

Pays	Mean					
	Rg MCQ ²	Total MCQ ²	Rg MCQ	Total MCQ	Rg IF	Total IF
US	1	280236.94	1	252603.19	1	365079.34
F	2	99831.18	2	89355.25	2	110344.68
D	3	49639.95	3	51739.24	3	73256.26
UK	4	35953.25	4	36781.97	4	57705.52
J	5	30036.85	7	28524.63	7	35270.80
I	6	28985.20	5	30832.37	6	42690.82
PRC	7	24092.62	6	30011.98	5	47657.45
CA	8	23332.65	8	23980.07	8	34874.76
E	9	18611.18	9	20803.07	9	30443.30
IL	10	15515.79	10	14474.40	10	21548.40
AUS	11	8797.88	11	9578.82	11	14748.96
CH	12	8746.92	12	9022.30	13	12624.56
RS	13	8718.21	14	8564.48	14	12038.91
BR	14	8083.73	15	7917.07	15	11068.42
NL	15	8011.82	13	8777.04	12	14649.71
S	16	6739.02	16	7235.88	16	10382.41
DK	17	5615.24	21	4991.71	21	7194.77
PL	18	5582.50	17	6271.52	18	8551.10
B	19	4926.62	18	5476.83	17	9188.18
A	20	4446.38	20	5157.12	19	8136.17

Tab.: Indicateurs par pays, 2001–06

Données issues de MR (Dubois, Rochet, Schlenker).

Indicateur par institution

Univ.			Mean			
	Rg MCQ ²	Total MCQ ²	Rg MCQ	Total MCQ	Rg IF	Total IF
US-PRIN	1	13306.91	1	9338.57	1	12815.75
F-PARIS11	2	11775.70	2	8874.17	5	9693.73
US-MIT	3	10167.84	3	8523.69	2	11085.38
US-CA	4	9470.50	5	8040.04	3	11027.26
F-PARIS6	5	8912.29	4	8408.59	4	10483.94
US-CHI	6	8479.73	6	6075.45	9	8015.07
US-HRV	7	7378.19	11	5209.15	15	5995.08
US-NY	8	6987.13	10	5343.39	7	9014.28
US-MI	9	6697.51	7	5961.18	10	7825.62
US-STF	10	6615.42	8	5941.93	6	9595.09
US-UCLA	11	6195.40	13	5031.20	12	7063.67
CA-TRNT	12	6175.88	16	4565.75	24	5354.04
US-WI	13	6027.15	12	5062.31	11	7366.12
US-MN	14	5913.35	9	5596.68	8	8293.21
US-CAIT	15	5636.37	20	4129.30	28	5132.85
F-TOUL3	16	5590.39	14	4729.01	17	5682.38
US-PURD	17	5062.58	21	4075.22	32	4934.39
US-OHS	18	4967.37	18	4160.04	23	5420.67
US-IASP	19	4950.28	37	3321.80	51	3544.89
US-IL	20	4942.68	15	4714.72	16	5961.37

Tab.: Indicateurs par université, 2001–06

Données issues de MR (Dubois, Rochet, Schlenker).

Variation dans le temps, MCQ²

Univ.	year						Moyenne
	2001	2002	2003	2004	2005	2006	
CA-TRNT	739.2	1,087.9	1,518.4	684.7	1,312.3	833.4	1,029.3
F-PARIS11	1,952.7	2,502.0	1,069.2	2,115.9	2,799.8	1,336.1	1,962.6
F-PARIS6	1,318.2	1,704.2	1,019.9	1,595.9	1,409.7	1,864.3	1,485.4
F-TOUL3	733.9	753.6	528.0	1,288.5	1,424.7	861.7	931.7
US-CA	1,510.5	1,334.2	1,368.5	1,965.3	1,920.2	1,371.8	1,578.4
US-CAIT	344.2	324.2	1,175.0	1,194.1	940.7	1,658.3	939.4
US-CHI	1,501.2	1,028.3	1,447.8	1,690.0	1,274.8	1,537.7	1,413.3
US-HRV	1,325.9	1,517.0	1,073.6	645.1	926.7	1,889.8	1,229.7
US-IASP	733.8	840.0	966.9	828.1	965.0	616.5	825.0
US-IL	485.1	605.1	708.4	645.7	1,450.2	1,048.3	823.8
US-MI	1,057.7	1,432.0	1,258.9	892.5	911.5	1,144.8	1,116.3
US-MIT	1,201.0	1,259.1	1,875.8	2,040.2	2,065.2	1,726.5	1,694.6
US-MN	1,098.1	1,285.2	730.7	699.9	982.2	1,117.4	985.6
US-NY	912.9	1,082.1	1,607.1	1,178.1	1,222.8	984.2	1,164.5
US-OHS	933.3	776.2	489.9	845.6	883.8	1,038.6	827.9
US-PRIN	1,656.3	1,819.9	2,046.5	1,761.4	2,828.1	3,194.7	2,217.8
US-PURD	694.7	612.8	1,195.5	754.7	1,143.1	661.8	843.8
US-STF	1,128.1	767.7	669.6	766.6	1,304.5	1,979.0	1,102.6
US-UCLA	988.7	891.4	1,173.0	938.8	561.5	1,642.0	1,032.6
US-WI	1,072.5	561.1	917.9	859.3	1,211.7	1,404.7	1,004.5

Tab.: Production par université, 2001–06

Données issues de MR (Dubois, Rochet, Schlenker).

Variation dans le temps, MCQ

Univ.	year						Moyenne
	2001	2002	2003	2004	2005	2006	
CA-TRNT	520.9	718.0	1,099.3	571.8	901.1	754.6	761.0
F-PARIS11	1,348.7	1,856.1	976.9	1,647.2	1,888.3	1,157.1	1,479.0
F-PARIS6	1,372.7	1,487.7	1,055.5	1,448.3	1,399.7	1,644.7	1,401.4
F-TOUL3	647.6	713.1	599.0	1,054.7	1,025.3	689.3	788.2
US-CA	1,240.8	1,172.1	1,103.6	1,460.2	1,670.2	1,393.2	1,340.0
US-CAIT	304.4	339.1	817.5	773.0	780.8	1,114.5	688.2
US-CHI	1,114.0	833.3	958.8	1,111.0	1,007.3	1,051.1	1,012.6
US-HRV	825.2	1,034.1	690.6	574.8	758.2	1,326.2	868.2
US-IASP	441.0	525.4	629.7	523.0	725.6	477.3	553.6
US-IL	456.1	699.8	703.5	714.0	1,165.8	975.6	785.8
US-MI	939.0	1,144.0	1,103.0	868.4	848.9	1,057.8	993.5
US-MIT	1,076.6	966.0	1,489.4	1,723.1	1,695.5	1,573.0	1,420.6
US-MN	1,054.6	1,138.9	673.4	621.5	975.3	1,132.9	932.8
US-NY	673.9	879.2	1,172.7	857.2	973.4	786.9	890.6
US-OHS	833.5	626.7	494.0	673.6	851.6	680.6	693.3
US-PRIN	1,256.5	1,268.7	1,542.1	1,226.3	2,014.1	2,030.9	1,556.4
US-PURD	523.7	519.4	852.7	596.4	863.6	719.3	679.2
US-STF	1,044.0	712.5	628.3	763.5	1,202.8	1,590.7	990.3
US-UCLA	933.5	764.5	938.7	714.9	519.9	1,159.9	838.5
US-WI	884.2	497.5	861.0	796.5	903.3	1,120.0	843.7

Tab.: Production par université, 2001–06

Données issues de MR (Dubois, Rochet, Schlenker).

L'autonomie des universités ouvre de nouveaux champs d'application à la bibliométrie. Certaines sont dangereuses et/ou peuvent être utiles.

- Diagnostic des points forts d'une institution.
- “Calibration” des évaluations individuelles entre disciplines.
- Indicateurs de succès global d'une politique de recherche/discipline.

Nécessité absolue : *calibrer* les indicateurs bibliométriques. Ex : part dans la production mondiale pour des indices, ramené au nombre de chercheurs, etc.

- L'utilisation d'outils bibliométriques est irréversible, et ne fait souvent que formaliser des pratiques informelles et imprécises.
- Mais les domaines et les modalités d'applications doivent être soigneusement contrôlés.
- Les indicateurs doivent être adaptés aux spécificités disciplinaires.
- Seuls des scientifiques disciplinairement compétents peuvent les mettre au point et les utiliser sans danger.
- Les évaluations doivent *in fine* reposer sur des experts scientifiquement compétents, même s'ils utilisent des indicateurs quantitatifs.