

Intensité, Qualité, Excellence et classements des universités

Nicolas Carayol ^{◇,‡}, Ghislaine Filliatreau[◇], Agenor Lahatte[◇]

[◇] Observatoire des Sciences et Techniques (OST)

[‡] GREThA, Université Bordeaux IV - CNRS

Feb 2012

Abstract

Cet article présente, dans un abord volontairement non technique, un modèle générique permettant de comparer des établissements de recherche deux à deux. Ces relations bilatérales sont utilisées principalement pour construire des classements d'universités. Comparativement aux classements existant, notre approche permet de rationaliser l'information utilisée, et surtout, de mieux fonder les comparaisons réalisées. En effet, les classements proposés sont adossés à des jugements de valeur explicites, qui sont traduits en mesures précises, homogènes et contrôlées. Trois types de jugements de valeur de référence sont proposés: un premier se focalisant sur l'intensité de l'activité de recherche, un deuxième sur sa qualité et un troisième sur l'excellence. L'article discute les résultats obtenus pour les principales universités de recherche américaines.

Keywords: Relations de dominance, citations, impact, classements, classes de référence.

1 Introduction

Depuis quelques années, plusieurs classements mondiaux d'universités ont vu le jour, stimulés notamment par une recherche d'information croissante émanant d'étudiants de plus en plus mobiles et désireux de choisir au mieux leur école ou université. Soumis à une concurrence accrue, les responsables de ces institutions accordent désormais une forte attention à ces classements. Si la qualité de l'enseignement est importante pour ces étudiants, le prestige scientifique des établissements est également considéré comme un critère essentiel dans la plupart des classements. Dans un contexte de forte hiérarchisation des établissements, l'enjeu principal des classements d'universités est de parvenir à rendre compte à la fois de l'intensité de l'activité scientifique des établissements,¹ et de leur qualité dans une approche quantitative qui soit intelligible pour les acteurs et les utilisateurs. Pour tenter d'y parvenir, la plupart des classements consistent en un savant dosage de mesures diverses permettant de capturer séparément ces différents aspects. Par exemple, le classement de ARWU développé par l'Université Jiao Tong de Shanghai, agrège des mesures variées d'excellence (nombre de prix Nobel et médailles Fields, nombre d'articles dans *Science* et *Nature*, nombre de chercheurs fortement cités) et des mesures de volume d'activité (nombre d'articles publiés).

Notre méthodologie se distingue de ces approches en ceci qu'elle n'utilise que les informations relatives aux articles publiés et leur impact (pouvant être mesurés de différentes manières, comme indiqué plus bas) enregistrés dans le Web Of Science (Thomson Reuters), qui est maintenu et enrichi au sein de l'Observatoire des Sciences et Techniques. Notre approche s'avère ainsi plus économe en information, s'appuyant uniquement sur des sources contrôlées, plus homogènes, moins intermédiées, et plus robustes dans le temps. Il faut noter que nous utilisons cependant cette information de manière très fine, en considérant le nombre d'articles produits par chaque établissement, dans chaque discipline et pour tout niveau possible d'impact de ces articles.

¹Des problématiques similaires pourraient être développées pour des classements plus orientés vers la qualité de l'enseignement et/ou du placement des étudiants. Dans cet article nous nous limitons volontairement aux approches centrées sur la fonction recherche des universités.

Ce traitement très précis de l'information scientométrique nous permet notamment de rendre compte de la dimension de qualité et d'excellence de la production scientifique des établissements qui est difficile à appréhender dans une approche quantitative n'utilisant pas la revue par les pairs (l'intensité étant plus aisément appréhendée). L'impact des articles dans la communauté scientifique est calculé à partir des données relatives aux citations, soit directement reçues, soit au niveau du journal, lesquelles doivent ensuite être corrigées pour les nombreux biais disciplinaires. L'idée que l'impact scientifique peut être approximé par les citations peut être associée au travail pionnier du sociologue R. K. Merton (cf. Merton, 1973), lequel a été ensuite systématisé par des scientomètres tels Price (1965) et Garfield (1963).

Notre approche se distingue en outre des classements actuels car elle permet une grande transparence et une continuité entre les prémisses retenues et les résultats obtenus. En effet, chaque inférence d'ordre sur des couples d'établissements repose sur une unanimité de comparaison pour des classes bien définies de la fonction associée de valorisation de la qualité, ce qui revient ici à adosser les comparaisons d'universités à des jugements de valeur clairs relatifs à la prise en compte du volume, de la qualité et de l'excellence des publications. Nous modélisons ici trois types de jugements de valeur. Le premier, traduisant un intérêt pour l'intensité de l'activité de recherche, ne valorise pas particulièrement l'impact des articles scientifiques mais seulement le volume. Le second prend en considération l'impact des articles sans accorder de valorisation particulière aux articles les plus cités dans la discipline de référence. Le troisième donne un poids proportionnellement plus important aux articles les plus cités. Ainsi, en fonction des prémisses retenues, il est possible d'établir des relations de dominance² deux à deux au sein d'un ensemble d'universités dans un système de valeur explicite. Nous montrons ensuite comment utiliser ces relations pour construire des classements et des

²Nous utilisons le terme de dominance car, en termes généraux, cette théorie se présente comme une généralisation de la notion de dominance stochastique largement utilisée en économie dans la théorie de choix dans l'incertain (Rothschild et Stiglitz, 1970) et dans l'analyse de distributions de revenus ou de richesse (Atkinson, 1970). Elle étend la notion de dominance stochastique en cela qu'elle permet de comparer des profils d'institutions en valorisant aussi la quantité d'items (et non la seule qualité et sa distribution).

classes de référence.

Après avoir présenté succinctement notre méthode, nous présentons dans cet article quelques classements portant sur les principales universités de recherche des États-Unis obtenus grâce à cette méthodologie originale.

2 Méthodologie pour des classements fondés sur des relations de dominance

Cette section présente la méthodologie fondant les classements et classes de références, dans un abord non technique. Le lecteur pourra se tourner vers Carayol et Lahatte (2011) pour une présentation complète, et vers Carayol et al. (2011) pour une présentation en français. Nous présentons dans un premier temps la question de la mesure de la qualité des articles scientifiques. Nous introduisons ensuite les fondements des comparaisons bilatérales et leur jugement de valeur associés. Enfin, nous présentons une manière de traduire les comparaisons bilatérales en classements et classes de référence.

2.1 Approximer la qualité des articles

La mesure de la qualité sur laquelle notre méthode est fondée repose, notamment, sur une appréhension de la mesure de l'impact des articles, abordé via trois indices différents, tous basés sur des décomptes des citations reçues. Le premier est le nombre de citations reçues par l'article considéré sur une période donnée après publication. Un tel indicateur, dit de visibilité directe, est basé uniquement sur les références faites aux articles considérés. Il est très intuitif, mais s'avère être une mesure relativement bruitée - les chercheurs se plaignent souvent du fait que leurs "meilleurs" articles ne sont pas les plus cités et il arrive que certains articles soient cités bien au delà de leur contribution scientifique telle qu'estimée par les experts.

Une autre option consiste à utiliser l'impact moyen des articles publiés par le journal dans lequel l'article a paru. Cet indicateur indirect de l'impact d'un article traduit la capacité de ses auteurs à publier dans des journaux plus ou moins cités en moyenne, et donc plus ou moins prestigieux.

Ces deux premiers indices présentent l'inconvénient de favoriser, en termes d'impact, les spécialités ou les sous-disciplines les plus saillantes d'une discipline. L'impact moyen relatif, troisième mesure proposée ici, pallie un tel biais. C'est l'impact moyen du journal rapporté à l'impact moyen des journaux de sa (ou ses) spécialité(s) scientifique(s) de référence. Un tel indice semble être particulièrement utile quand on analyse la production scientifique d'institutions qui sont spécialisées dans des sous-domaines dont la visibilité moyenne n'est pas très élevée. Il permet en outre de prendre en compte la diversité des pratiques de citation dans les différentes spécialités d'une même discipline (par exemple, le fait qu'en mathématiques, la sous-catégorie mathématiques fondamentales est moins citée en moyenne que la sous-catégorie mathématiques appliquées).

Il pourrait sembler naturel d'utiliser l'impact d'un article, défini de l'une des manières présentées ci-avant, comme mesure de la qualité de ce dernier. Nous écartons toutefois cette hypothèse du fait de la variabilité de cet indice d'une discipline scientifique à l'autre (due à des différences de pratiques de citation ou de taux de couverture des bases bibliographiques). Aussi, nous proposons de mesurer la qualité des articles en se basant sur leur position dans la distribution au sein de leur discipline selon l'impact. La qualité s_a^k d'un article a dans une discipline donnée k , sera égale au plus grand s , tel que son impact est au moins aussi élevé que celui de $100 \times s\%$ des articles publiés dans cette discipline. Autrement dit, étant donné la distribution de la production scientifique dans la discipline, et pour une mesure d'impact donnée, la qualité d'un article est la probabilité qu'un article tiré au hasard dans la discipline soit d'impact inférieur.

Cette approche permet l'agrégation d'articles de disciplines scientifiques différentes pour tout niveau de qualité donné mais ne s'avère pas être indispensable dans le cas d'une comparaison intra-discipline.

2.2 Etablir des comparaisons bilatérales

Dans notre modèle, nous distinguons plusieurs types de comparaisons bilatérales entre institutions, chaque type de relation de dominance étant associé à une série d'hypothèses différentes sur la fonction implicite de valori-

sation de la qualité, $v(\cdot)$. Nous parlons ici de relations de dominance. Une relation de dominance de i sur j indique que la comparaison est unanime pour toute fonction $v(\cdot)$ dans une classes de fonctions bien définie. Chaque classe de fonction renvoie à un ensemble de valeurs relatif à une manière particulière de considérer la qualité des articles. Trois types de relations de dominance sont abordées: la dominance forte (notée \blacktriangleright), la dominance standard (notée \triangleright) et la dominance faible (notée \trianglerighteq).

La notion de dominance forte requiert la non négativité de la fonction $v(\cdot)$, en d'autres termes aucun article ne saurait contribuer négativement à la performance scientifique d'une institution. Cette propriété de non négativité s'avère être nécessaire pour chaque approche différente de la notion de valorisation, elle attribue un rôle très faible aux mesures d'impact. Elle peut ainsi être associé à une valorisation de l'*intensité* de publication.

Pour la dominance standard, la fonction $v(\cdot)$ doit être non négative et non décroissante, ce qui veut dire que dans un domaine donné on attribue une valeur la plus élevée aux articles dont l'impact est plus grand. Cette deuxième propriété semble cohérente avec les différentes approches possibles de la valeur de la production scientifique. Elle correspond à une valorisation de la *qualité*.

En plus des deux propriétés précédentes, l'hypothèse de convexité de $v(\cdot)$ est retenue dans la dominance faible. Cette dernière hypothèse implique que $v(\cdot)$ donne un poids proportionnel ou plus que proportionnel aux articles de plus forte qualité. La convexité peut être considérée comme une hypothèse largement partagée puisque les CEO des universités et leur conseil d'administration souhaitent la présence de leur institution dans les segments de forte visibilité académique. Elle correspond à une valorisation de l'*excellence*.

Dans Carayol et Lahatte (2011), un théorème établit les conditions nécessaires et suffisantes à chaque relation de dominance bilatérale, permettant d'établir ces relations de dominance en utilisant les uniques données de publication et d'impact, c'est-à-dire sans spécifier précisément la fonction de valorisation, mais seulement qu'elle doit appartenir à une classe de fonctions bien définie.

Pour la dominance forte, la condition nécessaire et suffisante est que l'institution dominante doit produire toujours plus que l'institution dominée pour tous les niveaux de qualité. Le volume de production joue donc un rôle important en dominance forte puisqu'une institution ne peut dominer fortement une autre si cette dernière a plus de publications pour au moins un niveau quelconque de qualité. Notons que la dominance forte traite tous les niveaux de qualité de la même manière (aucune prime n'est accordée aux articles les plus cités).

La relation de dominance standard attribue une valeur croissante implicite à la qualité. La condition nécessaire et suffisante pour établir une telle relation de dominance est que la performance cumulée de l'institution dominante soit supérieure à celle de l'institution dominée au dessus de tout niveau minimum de qualité possible. Elle véhicule donc l'idée qu'un article vaut toujours au moins autant qu'un autre article ayant un niveau d'impact plus faible. Ce régime de compensation permet d'établir la domination d'une institution qui a une performance relative plus élevée pour certains niveaux donnés de qualité.

Quant à la dominance faible, elle donne un rôle prépondérant aux articles de qualité élevée. Ici, la condition nécessaire et suffisante est que la performance cumulée de l'institution dominante, pondérée de la qualité, au dessus de tout niveau plancher de qualité, soit supérieure à celle de l'institution dominée. Cela implique qu'une institution ne peut pas être dominée par une autre si sa performance en quantité pondérée par sa qualité, pour un segment supérieur quelconque de qualité, est supérieure. La dominance faible dépend ainsi moins de la quantité, puisqu'elle accorde une importance maximale aux segments de plus forte visibilité.

Pour les comparaisons interdisciplinaires, nous agrégeons la valeur des productions des différentes disciplines. Cette opération ne pose pas de problème d'agrégation puisque la mesure de la qualité retenue corrige les différences disciplinaires (cf. supra).

Nous introduisons en outre la possibilité de n'établir ces relations de dominance que sur des segments de qualité haute, dont l'ampleur est ajustable, c'est-à-dire de limiter les comparaisons à des parts variables des meilleurs

articles dans une discipline donnée. Un paramètre (ϕ) permet de définir le niveau seuil de qualité au dessous duquel les publications ne seront pas prises en compte. Si $\phi = 1$ alors tous les articles sont comptabilisés. Plus le segment de part mondiale retenue est sélectif (ϕ petit), plus on se limite aux segments de plus forte visibilité en négligeant les articles de qualité plus faible (inférieure à $1 - \phi$).

2.3 Construire des classements et classes de référence à partir des relations de dominance

Les relations de dominances sont établies deux à deux. Nous montrons dans cette section que l'on peut utiliser les relations de dominance, telles que précédemment définies, pour classer des universités et pour construire des groupes de référence (ou groupes de pairs).

Une relation de dominance ne peut en général pas être établie pour n'importe quelle paire d'institutions. Lorsque c'est possible (on parle alors de relation complète), il est assez évident et naturel de déduire un classement à partir du complexe de comparaisons bilatérales. Dans le cas général, il s'agit néanmoins de classer des établissements qui ne peuvent pas toujours être comparés pour une relation de dominance donnée. Nous parlons alors de pseudo-classement, obtenus sur la base de scores attribués aux acteurs sur la base de leur position relative dans le complexe de dominances bilatérales. Le complexe des relations bilatérales de dominance peut être représenté sous la forme d'un réseau dirigé construit en compilant les relations de dominance deux à deux.³ Nous proposons ici un pseudo-classement construit sur deux critères simples. Le premier, le critère dominant, correspond au nombre d'autres institutions dominées. Le second critère est le nombre d'autres institutions qui dominent l'institution considérée (le score est pris ici par ordre décroissant).⁴

³Dans les représentations proposées à la section suivante, et par soucis de visibilité, nous avons enlevé les relations de dominance qui peuvent être simplement déduites par l'application de propriété de transitivité des relations de dominance;

⁴Dans Carayol et Lahatte (2011), nous proposons une méthode plus élaborée permettant de classer les établissements sur la base de leur position dans le réseau de dominances

Les pseudo-classements présentent l'avantage d'être toujours réalisables sur la base de n'importe quelle relation de dominance. Il faut cependant, dans la pratique se garder de construire des pseudo-classements lorsque la proportion de relations de dominance que l'on peut établir sur le nombre relations de dominance possibles est suffisamment élevée. Nous nommons cette mesure le taux de complétude.

Le fait de ne pas pouvoir toujours établir une relation de dominance entre deux institutions amène à considérer que ces dernières ont l'une vis-à-vis de l'autre le statut de "pair". Pour une institution donnée, nous pouvons donc constituer sur la base de cette idée l'ensemble de ses "pairs". Nous parlons de classe de référence de l'institution considérée associée à la relation de dominance utilisée: la classe de référence d'une institution comprend toutes les institutions qui ne peuvent se comparer à elle dans le type de dominance considéré (au sens qu'aucune relation de dominance ne peut être établie, ni dans un sens ni dans l'autre). Cela signifie qu'elles sont à la fois similaires mais dans le même temps différentes puisque cette impossibilité d'établir une relation de dominance indique que chacune des deux institutions a des points forts vis-à-vis de l'autre.

3 Application aux universités de recherche américaines

Nous illustrons notre méthodologie à partir des données de performances d'un échantillon d'universités américaines. Les données sont présentées dans un premier temps avant de discuter les résultats obtenus.

3.1 Les données mobilisées

L'échantillon expérimental est constitué de 112 universités américaines, apparaissant dans les meilleures positions du classement mondial des universités produit par l'université Jiao-Tong de Shanghai. Les publications de ces établissements ont été identifiées dans la version "off-line" de la base bilatérales.

Thomson-Reuter-Web of Science (WoS) maintenue et enrichie par l'OST. Pour cette étude pilote, nous avons choisi d'utiliser les données de publication de 2003, 2004 et 2005, avec une fenêtre de décompte de citations de 3 années. Sur la base de ces options, la production scientifique des 112 établissements considérés dans cette analyse totalise 329910 articles. Les scores de citations reçues par ces articles vont de 0 à 1292 et l'impact moyen des journaux dans lesquels ils sont publiés varie de 0 à 27.6.

L'affectation des articles publiés à des domaines scientifiques est basé sur le regroupement des journaux en 9 grandes disciplines. Les 8 premières (Biologie fondamentale, Recherche médicale, Biologie appliquée - écologie, Chimie, Physique, Sciences de l'univers, Sciences pour l'ingénieur, Mathématiques) correspondent à des champs scientifiques distincts alors que la dernière, appelée Multidisciplinaire, regroupe des journaux qui publient des articles pouvant être assignés à différentes disciplines. La non prise en compte des articles de cette catégorie introduirait un biais non négligeable puisqu'il éliminerait une part des articles importants de plusieurs disciplines. Trois de ces journaux (Proceedings of the National Academy of Sciences USA, Science et Nature) sont très prestigieux et leurs articles ont été affectés un à un à leur discipline principale de rattachement (sur une base lexicographique).

L'impact scientifique des articles des établissements est considéré au travers des trois indices d'impact exposés dans la Section 2.

3.2 Classements des universités US

Les Tableaux (1), (2) et (3) présentent les résultats obtenus pour les 15 premières universités au niveau interdisciplinaire (respectivement lorsque l'impact est mesuré par les citations directes, l'impact moyen et l'impact moyen relatif).⁵ Dans chaque tableau et pour chaque type de dominance, sont reportés le rang dans le pseudo-classement (σ_i) et la performance dans

⁵La Table (??) présente les taux de complétude par discipline (et pour toutes les disciplines), associés aux relations de dominance \blacktriangleright , \triangleright , et \trianglerighteq , pour les trois mesures d'impact, lorsque toutes les publications sont considérées ($\phi = 1$). Ces taux de complétude sont toujours en dessous de l'unité, ce qui implique que seuls des pseudo-classements sont possibles. Cependant, les taux de complétude des relations de dominance standard et dominance faible atteignent des niveaux acceptables (au dessus de 80%) pour permettre

le critère dominant (n_i) lorsque toutes les publications sont prises en compte ($\phi = 1$).⁶

| Relations de dominance basées sur les citations | | | | | | |
|---|-------|-------|-------|-------|-------|-------|
| | ▶ | | ▷ | | ⊇ | |
| | Rang | #Dom | Rang | #Dom | Rang | #Dom |
| | r_i | n_i | r_i | n_i | r_i | n_i |
| Harvard | 10 | 65 | 1 | 111 | 1 | 111 |
| Stanford | 11 | 59 | 2 | 107 | 2 | 107 |
| Seattle | 1 | 82 | 3 | 105 | 3 | 105 |
| UCLA | 7 | 72 | 3 | 105 | 3 | 105 |
| UM_Ann_Arbor | 7 | 72 | 5 | 104 | 6 | 104 |
| Berkeley | 17 | 41 | 6 | 104 | 5 | 105 |
| Johns_Hopkins | 19 | 41 | 7 | 103 | 7 | 104 |
| Pennsylvania | 28 | 29 | 8 | 101 | 8 | 103 |
| WI_Madison | 3 | 78 | 9 | 97 | 11 | 97 |
| MIT | 55 | 14 | 10 | 95 | 9 | 102 |
| Columbia | 13 | 48 | 11 | 94 | 10 | 98 |
| Cornell | 5 | 76 | 12 | 94 | 15 | 95 |
| Twin_Cities | 4 | 77 | 12 | 94 | 13 | 95 |
| DE_UCSD | 14 | 48 | 14 | 93 | 12 | 96 |
| UCSF | 48 | 17 | 15 | 91 | 13 | 95 |

Table 1: Top 15 du pseudo ranking des universités U.S. (toutes disciplines).

Bien que les différents classements exposés soient globalement proches, certaines institutions ont des rangs très différents selon le type de relation de dominance associée. Par exemple, l’université Harvard est classée première à la fois pour la dominance standard et la dominance faible (l’impact étant mesuré par les citations, cf Table (1) alors qu’elle n’est classée qu’en dixième position lorsque la dominance forte est utilisée, classement dans lequel elle est largement devancée par Seattle (Washington University) qui est alors classée première. De même, le MIT n’est classé qu’en cinquante cinquième position dans le classement fondé sur la relation de dominance forte, alors

de tels pseudo-classements.

⁶Par manque de place, tous les résultats ne peuvent être présentés dans l’article. Les résultats copncernant cette étude sont présentés dans Carayol et al (2011). Des résultats plus complets encore sont à la disposition de la communauté sur la page internet suivante: <http://carayol.u-bordeaux4.fr/ranking.html>.

| Relations de dominance basées sur l'impact moyen | | | | | | |
|--|-------|-------|-------|-------|-------|-------|
| | ▶ | | ▷ | | ⊇ | |
| | Rang | #Dom | Rang | #Dom | Rang | #Dom |
| | r_i | n_i | r_i | n_i | r_i | n_i |
| Harvard | 1 | 91 | 1 | 111 | 1 | 111 |
| Berkeley | 7 | 82 | 2 | 106 | 2 | 106 |
| Stanford | 11 | 79 | 2 | 106 | 2 | 106 |
| Seattle | 2 | 89 | 4 | 104 | 4 | 104 |
| UCLA | 5 | 87 | 5 | 103 | 5 | 103 |
| UM_Ann_Arbor | 2 | 89 | 6 | 102 | 6 | 102 |
| Johns_Hopkins | 6 | 83 | 7 | 101 | 7 | 102 |
| Pennsylvania | 13 | 77 | 8 | 100 | 8 | 102 |
| MIT | 20 | 62 | 9 | 97 | 9 | 101 |
| WI_Madison | 4 | 88 | 10 | 96 | 11 | 96 |
| Cornell | 9 | 80 | 11 | 94 | 12 | 95 |
| Columbia | 16 | 72 | 12 | 94 | 10 | 97 |
| DE_UCSD | 19 | 64 | 13 | 93 | 12 | 95 |
| Twin_Cities | 9 | 80 | 14 | 91 | 15 | 92 |
| UCSF | 31 | 47 | 15 | 91 | 14 | 95 |

Table 2: Top 15 du pseudo ranking des universités U.S. (toutes disciplines).

qu'il est classé dans les dix premiers pour les deux autres formes de dominance. Ces différences illustrent les effets de la prise en compte différentielle du volume, de la qualité ou de l'excellence dans les classements: comme on l'a vu plus haut, la dominance forte prend pleinement en compte l'effet pur du volume de publication, alors que la dominance standard et surtout la dominance faible font intervenir l'impact des publications, permettant à des institutions de taille éventuellement plus modeste de voir leur niveau d'excellence pleinement valorisé.

En outre, certaines institutions voient leur rang varier sensiblement en fonction de la mesure d'impact utilisée. Par exemple, l'Université Berkeley se classe sixième et cinquième (en dominance et en dominance faible) lorsque l'impact est mesuré par les citations, alors qu'elle se classe en troisième et seconde position (respectivement) lorsque l'impact moyen est utilisé. Ceci indique que les chercheurs de Berkeley sont particulièrement performants dans leur capacité à publier des articles dans les journaux les plus impor-

| Relations de dominance basées sur l'impact moyen relatif | | | | | | |
|--|-------|-------|-------|-------|-------|-------|
| | ▶ | | ▷ | | ⊇ | |
| | Rang | #Dom | Rang | #Dom | Rang | #Dom |
| | r_i | n_i | r_i | n_i | r_i | n_i |
| Harvard | 1 | 100 | 1 | 111 | 1 | 111 |
| Stanford | 10 | 83 | 2 | 107 | 2 | 107 |
| Seattle | 4 | 95 | 3 | 106 | 3 | 106 |
| UM_Ann_Arbor | 1 | 100 | 3 | 106 | 3 | 106 |
| Berkeley | 5 | 89 | 5 | 106 | 5 | 106 |
| UCLA | 3 | 97 | 6 | 105 | 6 | 105 |
| Johns_Hopkins | 6 | 89 | 7 | 101 | 9 | 101 |
| Pennsylvania | 13 | 79 | 8 | 100 | 7 | 102 |
| MIT | 19 | 72 | 9 | 97 | 8 | 101 |
| WI_Madison | 7 | 88 | 10 | 96 | 10 | 96 |
| Columbia | 15 | 79 | 11 | 95 | 10 | 96 |
| Cornell | 9 | 85 | 12 | 95 | 12 | 96 |
| DE_UCSD | 16 | 78 | 13 | 94 | 13 | 95 |
| Twin_Cities | 8 | 87 | 14 | 93 | 14 | 94 |
| UCSF | 30 | 52 | 15 | 92 | 15 | 93 |

Table 3: Top 15 du pseudo ranking des universités U.S. (toutes disciplines).

tants de leur discipline. Par ailleurs, si l'on utilise l'impact moyen relatif, Berkeley se retrouve à la quatrième position révélant ainsi que les chercheurs de Berkeley, ne sont pas seulement excellents dans leur capacité à publier dans les meilleurs journaux de leur spécialité, révélé par le classement selon l'impact moyen relatif, mais aussi qu'ils le font dans des spécialités de plus fort impact de leur discipline, ce qui leur donne un avantage pour les mesures d'impact direct.

3.3 Des réseaux de dominance et des classes de référence pour mieux apprécier le positionnement relatif des universités US

Les pseudo-classements ne rendent compte que d'une partie de l'information disponible, et il est souvent utile d'explorer également la structure des réseaux de dominance. La Figure 1 présente le réseau de dominance faible (impact mesuré par les citations) parmi les premières universités dans le pseudo-

classement associé. Nous observons qu'en dessous de l'Université Harvard, la structure de dominance fait apparaître une certaine complexité. En effet, aucune relation de dominance n'a pu être établie entre l'Université du Michigan (à Ann Harbor), l'Université Washington de Seattle, UCLA et l'Université Stanford. Cependant, la structure aval permet de voir que Stanford est mieux classée que les trois autres parce qu'elle domine Berkeley et le MIT, ce que les autres ne parviennent pas à faire.

Figures 1 et 2 ici

Lorsque l'on se limite aux articles appartenant aux dix pourcent le plus cités de chaque discipline ($\phi = 0.1$), le réseau de dominance entre ces mêmes institutions se modifie naturellement dans le sens d'une plus forte hiérarchisation.⁷ Désormais Berkeley domine Michigan et Seattle, et Stanford domine Michigan et le MIT domine UCLA et Johns Hopkins (cf. Figure 2). En outre, la hiérarchie s'est entièrement clarifiée puisqu'elle est maintenant linéaire depuis l'université de Pennsylvanie jusqu'à l'université du Wisconsin, en passant par l'université de Columbia et l'université de Californie à San Diego (UCSD).

Lorsque l'impact moyen est utilisé pour apprécier l'impact des articles (cf. Figure 3), on observe un déplacement de Berkeley dans la structure de dominance. En effet, Berkeley échappe maintenant à la domination de Stanford et domine désormais le MIT. Cette évolution est partiellement remise en cause lorsque l'impact relatif est utilisé (Figure 4). En outre, d'autres universités telles que Seattle et Michigan voient leurs positions réévaluées dans ce contexte. Ces deux universités dominent désormais UCLA qui apparaît donc relativement moins performante lorsqu'on se concentre sur sa capacité à placer des articles dans les meilleures revues de chaque spécialité.

Figures 3 et 4 ici

Pour un type de dominance donné, deux institutions qui n'ont entre elles aucune relation de dominance, appartiennent chacune à la classe de référence de l'autre. Ceci peut être déduit de la lecture des réseaux de dominance

⁷Lorsque ϕ diminue, plus de relations de dominance peuvent être établies alors qu'aucune ne peut être remise en cause. Le réseau de dominance tend donc naturellement à se hiérarchiser.

ajustés, par le fait qu'il n'y a pas de chemin entre ces deux institutions. La classe de référence d'une institution comprend toutes les institutions qui ne peuvent être comparées à elle, c'est à dire qu'elles sont à la fois proches et différenciées. Par exemple, Stanford et Seattle qui n'ont pas, entre elles, de relation de dominance faible (impact mesuré par les citations), appartiennent chacune à la classe de référence de l'autre. Lorsque l'on étudie la structure de la production scientifique de ces deux institutions, on s'aperçoit que, si ces deux universités ont des productions scientifiques proches leur permettant de se situer toutes les deux au sommet de la hiérarchie universitaire, leur structure de production au regard des citations directes sont effectivement différentes. Stanford a des performances relativement plus élevées parmi les publications les plus citées, alors que Seattle domine significativement dans des segments d'impact plus faible. Ainsi, lorsque l'on se limite aux publications faisant partie des dix pourcent les plus cités de chaque discipline, Seattle est dominée par Stanford: les deux institutions ne font plus partie de leurs classes de référence respectives.

4 Conclusion

Nous avons proposés différents outils de comparaison des productions scientifiques des universités (classements, classes, réseaux) qui reposent sur une théorie permettant d'établir différentes relations de dominance entre universités sur la base de leur production scientifique et de leur impact. Face à l'expansion récente des classements d'universités, cette approche est originale à plusieurs titres. Tout d'abord elle permet d'affiner les analyses de positionnement des institutions de recherche, en donnant aux utilisateurs potentiels une marge de liberté dans les critères utilisés. Ensuite, cette approche permet de relier des prémisses explicites sur la valorisation de l'impact scientifique à des comparaisons et des classements. Enfin, ce type d'approche permet à la fois de proposer une méthode de classements et une analyse des structures de réseaux entre institutions en établissant des relations de dominance deux à deux.

Nous avons illustré son application sur un large échantillon d'universités

de recherche américaines. Cette théorie pourrait être appliquée à de nombreux contextes dans lesquels un évaluateur, ou une communauté d'évaluateurs, souhaiterait(ent) comparer des individus ou des institutions, en fondant leurs jugements sur des principes simples, explicites et ajustables valorisant de différentes manières la quantité et la qualité de leur production scientifique.

Références

Atkinson, A. B., 1970, On the measurement of inequality, *Journal of Economic Theory* 2, 244-263.

Carayol, N., Lahatte A., 2011, Dominance relations and universities ranking, mimeo.

Carayol N. , Filliatreau G., Lahatte A., 2011, Impact, dominance et classements des universités, mimeo.

Cheng Y., Zitt M., 2009, Towards a global ranking by size-independent bibliometric measures: An experiment on selected US, French and Chinese Universities, mimeo OST.

Garfield, E. 1956, Citation Indexes for Science, *Science* 123, 61-62.

Merton R.K., 1973, *The sociology of science. Theoretical and empirical investigations*, University Press of Chicago, Chicago.

Price, de Solla, D.J., 1963, *Little Science, Big Science*, New York: Columbia University Press.

Quirk, J.P., Saposnick, R., 1962, Admissibility and measurable utility functions, *Review of Economic Studies* 29, 140-146.

Rothschild, M., Stiglitz, J. E., 1970, Increasing Risk: I. A definition, *Journal of Economic Theory* 2, 225-243.

Figures

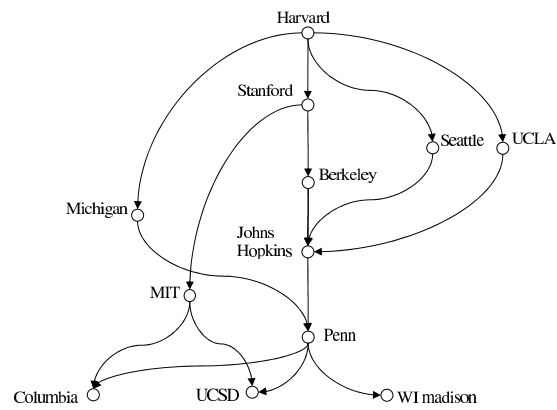


Figure 1: Réseau de dominance faible ajusté parmi les meilleures universités U.S. (toutes disciplines), quand l'impact est mesuré par les citations directes.

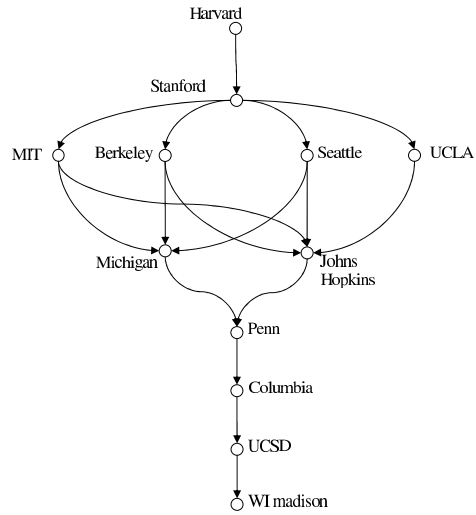


Figure 2: Réseau de dominance faible ajusté parmi les meilleures universités U.S. (toutes disciplines), quand l'impact est mesuré par l'impact moyen, en ne considérant que les articles qui ont un impact les situant dans les 10% supérieurs de la discipline.

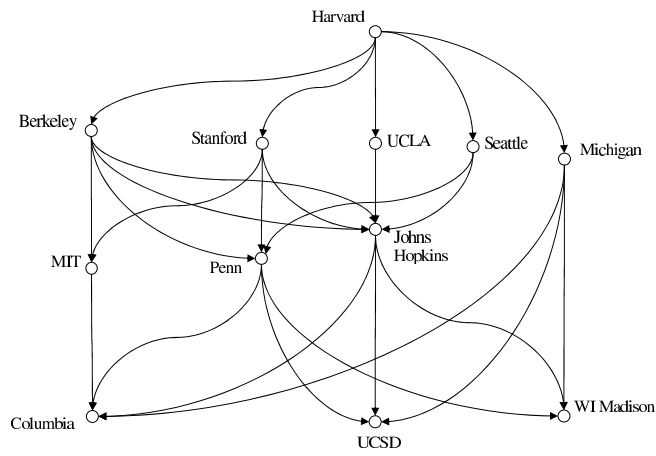


Figure 3: Réseau de dominance faible ajusté parmi les meilleures universités U.S. (toutes disciplines), quand l'impact est mesuré par l'impact moyen.

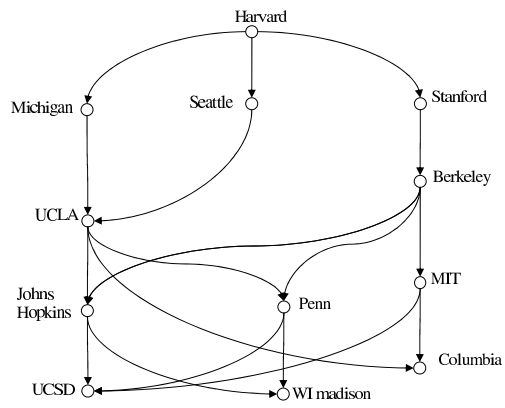


Figure 4: Réseau de dominance faible ajusté parmi les meilleures universités U.S. (toutes disciplines), quand l'impact est mesuré par l'impact moyen relatif.