

NOTE MÉTHODOLOGIQUE B-5

Les publications scientifiques

1 – LA BIBLIOMÉTRIE SUR LES PUBLICATIONS SCIENTIFIQUES

Les publications dans les journaux scientifiques constituent pour les chercheurs un des principaux modes de diffusion de leurs travaux. Les notices qui décrivent ces publications sont enregistrées dans des bases de données bibliographiques. L'analyse "bibliométrique" mobilise les ressources de la statistique et de l'analyse des données pour traiter l'information contenue dans ces notices, qui comportent : des données sur les sources (journal, auteurs et leurs affiliations), des textes descriptifs (titre, mots-clés, résumé...), des indications de classement selon la nomenclature disciplinaire, et enfin des données sur les liens avec les autres publications (références bibliographiques citées).

L'exploitation de ces données afin de produire des indicateurs sur les publications scientifiques comprend typiquement l'analyse des volumes de publications, des volumes de citations qu'elles reçoivent et des coopérations scientifiques à travers les collaborations. Les biais des données, d'une part, les limites statistiques, d'autre part, rendent ces indicateurs très sensibles aux choix méthodologiques adoptés : leur interprétation demande donc une grande vigilance.

2 – LA SOURCE UTILISÉE

La base de données bibliographiques utilisée par l'OST est le *Web of Science*[®] (WoS) de Thomson Reuters, Philadelphie. C'est un outil de référence pour les producteurs d'indicateurs dans le monde ; les particularités et les biais de cette base revêtent donc une grande importance puisqu'ils affectent la plupart des statistiques internationales sur les publications scientifiques. On considère ici, principalement, la partie de la base dédiée aux sciences de la matière et de la vie : SCI-WoS. En effet, les sciences sociales et les humanités sont pour l'instant exclues des indicateurs bibliométriques standard proposés par l'OST, car les bases correspondantes du WoS présentent des biais, variables selon les spécialités, et qui peuvent être considérables.

Le Web of Science[®] s'est imposé en bibliométrie grâce à plusieurs caractéristiques. C'est une base qui assure le dépouillement quasi intégral d'environ 8 000 journaux scientifiques sélectionnés pour leurs règles de fonctionnement éditorial et leur niveau de visibilité internationale (nombre moyen de citations par article). Ce mode de sélection des revues ne garantit pas l'exacte représentativité de l'équilibre entre disciplines ou entre sous-disciplines. La base inclut également des comptes rendus de congrès et des revues électroniques.

Les critiques qui lui sont les plus couramment adressées concernent les biais en faveur de la science anglo-américaine, mais il existe aussi des biais contraires comme la sur-représentation de la littérature nationale de certains pays. L'écrasante majorité des publications répertoriées dans le Web of Science[®] est en langue anglaise, et cette proportion tend encore à s'accroître. C'est, dans quelques cas, le résultat d'un biais linguistique des bases de Thomson Reuters, mais c'est surtout le reflet de la domination d'un modèle international de la science.

Tableau 1

SCI-WoS : nombre et répartition des articles selon les disciplines (2001-2006)

| Discipline | Compte fractionnaire | | | | Compte de présence | | | |
|-----------------------------|----------------------|----------------|-----------------|--------------|--------------------|----------------|-----------------|--------------|
| | Nombre | | Répartition (%) | | Nombre | | Répartition (%) | |
| | 2001 | 2006 | 2001 | 2006 | 2001 | 2006 | 2001 | 2006 |
| Biologie fondamentale | 116 762 | 129 614 | 15,6 | 14,5 | 149 254 | 166 934 | 19,9 | 18,7 |
| Recherche médicale | 229 100 | 262 110 | 30,6 | 29,4 | 255 568 | 292 117 | 34,1 | 32,8 |
| Biologie appliquée-écologie | 52 099 | 59 613 | 7,0 | 6,7 | 67 063 | 76 029 | 9,0 | 8,5 |
| Chimie | 106 110 | 130 265 | 14,2 | 14,6 | 130 369 | 166 052 | 17,4 | 18,6 |
| Physique | 85 366 | 100 397 | 11,4 | 11,3 | 101 134 | 125 191 | 13,5 | 14,0 |
| Sciences de l'univers | 46 347 | 57 934 | 6,2 | 6,5 | 53 925 | 69 417 | 7,2 | 7,8 |
| Sciences pour l'ingénieur | 79 068 | 110 220 | 10,6 | 12,4 | 102 385 | 142 145 | 13,7 | 15,9 |
| Mathématiques | 22 563 | 27 914 | 3,0 | 3,1 | 26 304 | 33 749 | 3,5 | 3,8 |
| Multidisciplinaire | 11 429 | 13 197 | 1,5 | 1,5 | 12 590 | 18 373 | 1,7 | 2,1 |
| Toutes disciplines | 748 844 | 891 265 | 100,0 | 100,0 | 748 844 | 891 265 | 100,0 | 100,0 |

données Thomson Reuters, traitements OST

rapport OST-2008

- le total inclut les journaux de la spécialité "multidisciplinaire". Cette spécialité comprend des journaux de niveau scientifique très disparate

La représentativité de la base est peu contestée pour les domaines les plus internationalisés des sciences physiques ou de la biologie fondamentale. L'image peut être un peu moins fidèle pour des domaines présentant une forte spécificité nationale, une forte diffusion hors "articles scientifiques", un fort degré d'application, ou encore une taille trop faible. C'est le cas de certaines spécialités de la recherche médicale ou de la biologie appliquée.

D'autres problèmes viennent limiter la représentativité des bases bibliométriques en général, notamment dans les disciplines à cycle court, où la communication se fait autant par voie de conférences que d'articles (notamment en informatique).

3 – DE LA SOURCE AUX INDICATEURS

3.1. Les journaux et les documents sélectionnés

La couverture en journaux de la base évolue en fonction de l'évolution de la visibilité internationale des journaux scientifiques. Ici, on a suivi le principe de "l'ensemble dynamique de journaux", plus représentatif au fil du temps que la solution alternative "ensemble constant de journaux". Pour cette raison, les indicateurs sont en général lissés sur trois ans (dans les tableaux, la dernière année d'information disponible est utilisée pour dater l'indicateur : 2006 pour 2004-2006). L'année correspond à la date de publication des articles, et non à la date d'entrée en base. Cinq types de documents ont été retenus : articles, articles de synthèse ("reviews"), lettres, notes, articles issus de comptes rendus de congrès. Ces derniers apparaissent souvent sous forme de sélections dans des numéros spéciaux. La proportion de congrès couverts de cette manière reste faible.

Les indicateurs standard sont fondés sur le périmètre complet ("Science Citation Index Expanded") de la base. Cependant, des travaux bibliométriques ont établi que la population de journaux couverte était très hétérogène, avec une majorité de journaux internationaux mais aussi une minorité de journaux dont les auteurs ou l'audience se limitaient à un ou quelques pays. Dans cette minorité, de nombreux journaux ont un impact faible et assez peu fiable. Les effets de cette sous-population sur les indicateurs bibliométriques sont loin d'être négligeables. Un exemple classique est celui de la Russie, dont la littérature est largement représentée dans la base par des journaux nationaux avec des auteurs nationaux. Enlever cette sous-population diminue l'indicateur de part de publications, mais augmente fortement l'indicateur d'impact des publications de ce pays. Aussi on trouvera dans cet ouvrage des tableaux d'indicateurs calculés en périmètre ajusté. Le périmètre ajusté est obtenu en enlevant de la base les journaux présentant à la fois un faible indice d'internationalisation et un faible indice d'impact. Le degré d'internationalisation est défini par le profil des pays-auteurs dans le journal, par comparaison avec le profil moyen dans la discipline ou la spécialité.

Ceci limite les mesures à la littérature scientifique la plus internationalisée, avec l'effet indirect de modifier l'équilibre entre disciplines, qui sont inégalement internationalisées : par exemple, les mathématiques et la biologie fondamentale sont très internationalisées, la chimie et la biologie appliquée-écologie le sont moins. L'effet sur les indicateurs nationaux est important pour plusieurs pays émergents dotés d'une activité éditoriale propre, mais peu visible, notamment en comparaison avec des pays très performants comme la Suisse ou les États-Unis.

Les seuils choisis pour construire le périmètre ajusté sont assez stricts. Les journaux les plus "nationaux" sont donc écartés, sauf s'ils figurent dans les 50 % (environ) les plus visibles. Ceci conduit, pour l'année 2005 par exemple, à soustraire de la base 15,7 % des journaux. En termes d'articles, 11,0 % sont écartés, représentant l'intersection entre les supports les moins internationalisés (13,7 % des articles) et les moins visibles (33,5 % des articles).

Tableau 2

SCI-WoS : nombre et répartition des articles selon la discipline et le périmètre (2001-2006)

| Discipline | Périmètre ajusté | | | | Périmètre standard | | Périmètre ajusté/périmètre standard (%) | |
|-----------------------------|------------------|----------------|-----------------|--------------|--------------------|----------------|---|-------------|
| | Nombre | | Répartition (%) | | Nombre | | | |
| | 2001 | 2006 | 2001 | 2006 | 2001 | 2006 | 2001 | 2006 |
| Biologie fondamentale | 110 040 | 124 227 | 16,9 | 15,5 | 116 762 | 129 614 | 94,2 | 95,8 |
| Recherche médicale | 193 678 | 231 805 | 29,7 | 29,0 | 229 100 | 262 110 | 84,5 | 88,4 |
| Biologie appliquée-écologie | 42 878 | 51 138 | 6,6 | 6,4 | 52 099 | 59 613 | 82,3 | 85,8 |
| Chimie | 90 387 | 113 499 | 13,9 | 14,2 | 106 110 | 130 265 | 85,2 | 87,1 |
| Physique | 75 121 | 89 078 | 11,5 | 11,1 | 85 366 | 100 397 | 88,0 | 88,7 |
| Sciences de l'univers | 42 284 | 53 404 | 6,5 | 6,7 | 46 347 | 57 934 | 91,2 | 92,2 |
| Sciences pour l'ingénieur | 67 953 | 100 068 | 10,4 | 12,5 | 79 068 | 110 220 | 85,9 | 90,8 |
| Mathématiques | 21 213 | 26 227 | 3,3 | 3,3 | 22 563 | 27 914 | 94,0 | 94,0 |
| Multidisciplinaire | 7 891 | 10 911 | 1,2 | 1,4 | 11 429 | 13 197 | 69,0 | 82,7 |
| Toutes disciplines | 651 445 | 800 356 | 100,0 | 100,0 | 748 844 | 891 265 | 87,0 | 89,8 |

données Thomson Reuters, traitements OST

rapport OST-2008

3.2. Le regroupement des journaux en disciplines et sous-disciplines

Les huit disciplines académiques (voir la fiche de nomenclature A4), définies à partir des “spécialités disciplinaires” (près de 170 spécialités sont assignées aux journaux par *Thomson Reuters* (“subject categories”)). Dans la base, les journaux peuvent être attribués à plusieurs spécialités (jusqu’à 6). En compte fractionnaire, les articles des journaux multi-attribués sont fractionnés entre spécialités. En compte de présence, ils sont intégralement comptés dans chaque spécialité d’appartenance. La rubrique “multidisciplinaire”, très hétérogène, n’est pas isolée dans les tableaux mais contribue aux totaux “Toutes disciplines”.

La classification en sous-disciplines a été faite à partir d’une carte des proximités entre spécialités. Des experts scientifiques ont été associés à la discussion des résultats.

Tableau 3

SCI-WoS : nombre et répartition des articles par sous-discipline au niveau mondial (2001-2006)

| Sous-discipline | Compte fractionnaire | | | | Compte de présence | | | |
|--|----------------------|----------------|-----------------|--------------|--------------------|------------------|-----------------|--------------|
| | Nombre | | Répartition (%) | | Nombre | | Répartition (%) | |
| | 2001 | 2006 | 2001 | 2006 | 2001 | 2006 | 2001 | 2006 |
| Biochimie | 48 694 | 50 085 | 6,5 | 5,6 | 65 314 | 67 505 | 8,7 | 7,6 |
| Bioingénierie | 16 959 | 20 024 | 2,3 | 2,2 | 24 856 | 30 555 | 3,3 | 3,4 |
| Biotechnologie, génétique | 14 763 | 17 277 | 2,0 | 1,9 | 24 136 | 29 029 | 3,2 | 3,3 |
| Microbiologie et virologie, immunologie | 27 032 | 30 323 | 3,6 | 3,4 | 35 498 | 38 969 | 4,7 | 4,4 |
| Neurosciences, sciences comportementales | 39 688 | 45 075 | 5,3 | 5,1 | 52 472 | 59 627 | 7,0 | 6,7 |
| Endocrinologie | 12 274 | 14 175 | 1,6 | 1,6 | 15 282 | 17 691 | 2,0 | 2,0 |
| Reproduction, biologie du développement | 12 085 | 13 735 | 1,6 | 1,5 | 15 808 | 18 933 | 2,1 | 2,1 |
| Cancérologie | 13 827 | 17 675 | 1,8 | 2,0 | 18 276 | 23 264 | 2,4 | 2,6 |
| Cardiologie-pneumologie | 27 417 | 30 555 | 3,7 | 3,4 | 33 635 | 37 788 | 4,5 | 4,2 |
| Chirurgie, gastro-entérologie, urologie | 29 797 | 32 933 | 4,0 | 3,7 | 39 673 | 43 852 | 5,3 | 4,9 |
| Divers médical | 63 366 | 66 747 | 8,5 | 7,5 | 73 557 | 77 495 | 9,8 | 8,7 |
| Pharmacie, toxicologie | 19 569 | 24 588 | 2,6 | 2,8 | 29 693 | 35 672 | 4,0 | 4,0 |
| Santé publique et divers | 24 992 | 32 773 | 3,3 | 3,7 | 34 282 | 44 602 | 4,6 | 5,0 |
| Agriculture, biologie végétale | 18 741 | 20 940 | 2,5 | 2,3 | 22 510 | 24 990 | 3,0 | 2,8 |
| Agroalimentaire | 13 684 | 16 695 | 1,8 | 1,9 | 18 054 | 22 562 | 2,4 | 2,5 |
| Écologie, biologie marine | 21 144 | 25 538 | 2,8 | 2,9 | 26 433 | 32 201 | 3,5 | 3,6 |
| Chimie générale | 16 875 | 21 512 | 2,3 | 2,4 | 20 198 | 26 433 | 2,7 | 3,0 |
| Chimie analytique | 12 204 | 14 610 | 1,6 | 1,6 | 17 525 | 22 030 | 2,3 | 2,5 |
| Chimie organique, minérale, nucléaire | 20 068 | 22 842 | 2,7 | 2,6 | 23 446 | 27 275 | 3,1 | 3,1 |
| Matériaux, polymères | 35 428 | 46 470 | 4,7 | 5,2 | 45 449 | 62 122 | 6,1 | 7,0 |
| Physique du solide | 56 218 | 69 187 | 7,5 | 7,8 | 71 706 | 91 984 | 9,6 | 10,3 |
| Physique particules et nucléaire | 33 705 | 38 787 | 4,5 | 4,4 | 42 137 | 52 337 | 5,6 | 5,9 |
| Physique générale | 22 714 | 26 926 | 3,0 | 3,0 | 25 083 | 30 448 | 3,3 | 3,4 |
| Environnement | 11 610 | 16 142 | 1,6 | 1,8 | 17 354 | 24 596 | 2,3 | 2,8 |
| Astronomie, astrophysique | 9 340 | 11 193 | 1,2 | 1,3 | 11 761 | 13 129 | 1,6 | 1,5 |
| Géosciences | 22 654 | 27 209 | 3,0 | 3,1 | 27 910 | 33 454 | 3,7 | 3,8 |
| Énergie, génie chimique et industriel | 24 489 | 29 556 | 3,3 | 3,3 | 34 144 | 43 054 | 4,6 | 4,8 |
| Génie civil, minier | 6 314 | 8 060 | 0,8 | 0,9 | 10 150 | 13 103 | 1,4 | 1,5 |
| STIC : génie électrique et électronique | 14 610 | 16 733 | 2,0 | 1,9 | 24 265 | 29 376 | 3,2 | 3,3 |
| STIC : informatique, télécommunications | 13 608 | 28 913 | 1,8 | 3,2 | 18 546 | 34 980 | 2,5 | 3,9 |
| STIC : intelligence artificielle | 10 381 | 15 838 | 1,4 | 1,8 | 14 457 | 20 846 | 1,9 | 2,3 |
| Mathématiques | 22 563 | 27 914 | 3,0 | 3,1 | 26 304 | 33 749 | 3,5 | 3,8 |
| Multidisciplinaire et divers | 12 032 | 10 238 | 1,6 | 1,1 | 14 059 | 10 693 | 1,9 | 1,2 |
| Toutes sous-disciplines | 748 844 | 891 265 | 100,0 | 100,0 | 748 844 | 1 174 344 | 100,0 | 100,0 |

données Thomson Reuters, traitements OST

rapport OST-2008

3.3. Le principe de comptage

Les statistiques par type d'acteur (pays, région) reposent non sur la nationalité des auteurs mais sur l'adresse des laboratoires et des institutions indiquée par chaque auteur. En d'autres termes, l'article d'un chercheur français en détachement aux États-Unis comptera uniquement pour le laboratoire américain d'accueil s'il omet de préciser l'adresse de son institution d'origine.

Les articles scientifiques étant souvent cosignés par plusieurs auteurs et plusieurs institutions, plusieurs options de comptage existent. Dans une logique de "contribution" à la science mondiale, chaque article est fractionné au prorata du nombre d'adresses différentes indiquées par ses auteurs, de manière à ce que la somme des adresses soit de 100 %. Ce principe est également appliqué aux articles d'un journal scientifique appartenant à plusieurs spécialités. Ce type de compte, dit "fractionnaire", où chaque article a un poids unitaire, est additif à toutes les échelles et bien adapté à la macroanalyse.

L'autre logique est celle de la "participation" à la science mondiale, qui repose sur un décompte en compte "entier-distinct" ou "de présence" : tout acteur est crédité d'une participation unitaire à une adresse dès lors que sa participation est attestée par une adresse. L'indicateur reflétant la participation est supérieur à celui de la contribution, par exemple la France peut être présente dans 8 % des publications mondiales d'un domaine mais n'y contribuer que pour 5 % quand on considère le fractionnement par article. Le compte de présence produit donc des participations d'acteurs dont la somme est supérieure à 100 % et les valeurs varient à chaque changement d'échelle. Malgré cet inconvénient, le compte de présence est préférable pour la microanalyse. Il est aussi plus facilement interprétable pour les copublications.

4 – LES INDICATEURS PRÉSENTÉS

Il convient de noter que des précautions s'imposent pour interpréter les indicateurs calculés pour de petites entités (petits pays, petites disciplines) qui sont sensibles à l'activité d'un nombre limité d'acteurs et susceptibles d'importantes fluctuations.

4.1. Les indicateurs de production scientifique

La part mondiale, calculée à partir du dénombrement des publications, est le plus simple des indicateurs comparatifs de production. Le maintien de leur part mondiale par les systèmes de recherche des principaux pays scientifiques exige un effort soutenu, en raison de l'émergence de nouveaux acteurs, comme la Chine et l'Asie du Sud-Est.

4.2. Les indicateurs de spécialisation

L'indice de spécialisation rapporte le poids d'un pays dans une discipline au poids du pays toutes disciplines confondues, ou encore le poids de la discipline dans ce pays au poids de la discipline dans le monde. Cet indice est l'un de ceux qui permet d'établir et de comparer le profil disciplinaire de différents pays. La dispersion de ces indices donne, pour chaque pays, une idée de sa stratégie de plus ou moins grande spécialisation globale : investissement réparti ou au contraire stratégie de "créneau", un choix souvent contraint pour les pays de petite taille. La prédominance d'un pays (par exemple les États-Unis) tend à peser sur l'amplitude de l'indice de spécialisation correspondant. Ici, la valeur neutre est égale à 1 ; une valeur supérieure à 1 signifie une spécialisation ; inférieure à 1, il s'agit d'une sous-spécialisation.

4.3. Les indicateurs de visibilité : citation et impact

Les publications scientifiques comportent une liste de références bibliographiques (ou "citations émises") aux travaux antérieurs. On peut ainsi, pour une publication donnée, évaluer après un certain délai (la "fenêtre de citation"), le nombre des citations reçues. Le "facteur d'impact" des journaux, (*Impact Factor* dans le *Journal of Citation Report-JCR*), bien connu des scientifiques qui en tiennent souvent compte pour orienter leur stratégie de publication, a popularisé la notion "d'impact", définie par le nombre moyen de citations par article, calculé selon diverses conventions. Les comportements de citation, dont dépend l'interprétation des impacts, ont donné lieu à de nombreux travaux en sociologie des sciences et en bibliométrie. Le choix des périodes de référence affecte également la valeur des indicateurs. Les impacts varient selon la langue, le type de recherche, le champ scientifique. La variété des caractéristiques et habitudes de citation des communautés scientifiques interdit les comparaisons directes entre disciplines. Enfin, les citations sont marquées par des effets d'auto-renforcement très sensibles ("le succès entraîne le succès").

Un indicateur d'impact nécessite une désagrégation et/ou une normalisation par discipline ou domaine. Le principal indicateur d'impact utilisé par l'OST est l'indicateur d'impact relatif qui est le ratio, par exemple pour un pays dans un domaine donné, de la part mondiale en citations recueillies et de la part mondiale correspondante en publications ; c'est aussi le ratio du nombre moyen de citations par publication, recueilli par ce pays dans ce domaine, par rapport à la moyenne mondiale. C'est une mesure commode mais qui n'a pas les propriétés d'une moyenne pondérée. Elle permet de comparer des disciplines entre elles, mais reste sensible à la composition de ces disciplines en spécialités au comportement de citation différent.

Le "ratio de citation relatif" (RCR) apporte une information complémentaire. Le RCR est défini comme le ratio de l'impact réel des publications d'un acteur sur l'"impact espéré" qu'auraient ces publications si elles recevaient l'impact moyen du journal dans lequel elles sont publiées. L'impact moyen du journal, bien que calculé avec des options différentes du "facteur d'impact" de *Thomson Reuters*, reflète la même réalité que ce dernier. L'"impact espéré" exprime la capacité de l'acteur à publier dans des journaux d'un certain niveau de visibilité, quel que soit le niveau de citation relatif qu'il atteint dans chacun de ces journaux. C'est une mesure de compétitivité pour l'accès à des journaux prestigieux. Le RCR

exprime la position de l'acteur au sein de chaque journal, indépendamment de sa capacité à être publié dans tel ou tel support. Par exemple, l'impact espéré discrimine en général les acteurs selon leur orientation de recherche vers le fondamental (accès facilité aux revues prestigieuses) ou l'appliqué. Le RCR exprime la visibilité de chaque acteur dans le domaine où il exerce, indépendamment de la visibilité propre du domaine, par exemple fondamental ou appliqué. Les deux mesures, impact espéré et RCR (dont par définition le produit est l'impact global), sont nécessaires pour caractériser la position et la stratégie de publication d'un acteur.

Par ailleurs, la part mondiale de citations (numérateur des indices d'impact relatif) constitue par elle-même un indicateur intéressant, moins sensible que les impacts à la délimitation des frontières du Web of Science® (c'est-à-dire la frange de journaux de faible impact). Cette robustesse peut être mise à profit dans les comparaisons chronologiques.

Tableau 4

SCI-WoS : nombre moyen de citations par publication (indice d'impact direct) selon différentes fenêtres de citations (2002)

| Discipline | Monde : indice d'impact direct (2002) | | | |
|-----------------------------|---------------------------------------|-------------|-------------|-------------|
| | Fenêtre de citations à : | | | |
| | 2 ans (immédiates) | 3 ans | 4 ans | 5 ans |
| Biologie fondamentale | 3,24 | 7,13 | 11,11 | 14,86 |
| Recherche médicale | 1,85 | 4,32 | 6,94 | 9,43 |
| Biologie appliquée-écologie | 1,02 | 2,47 | 4,17 | 5,91 |
| Chimie | 1,54 | 3,44 | 5,44 | 7,41 |
| Physique | 1,80 | 3,66 | 5,45 | 7,11 |
| Sciences de l'univers | 1,56 | 3,48 | 5,61 | 7,72 |
| Sciences pour l'ingénieur | 0,48 | 1,23 | 2,14 | 3,09 |
| Mathématiques | 0,43 | 1,04 | 1,76 | 2,51 |
| Multidisciplinaire | 7,27 | 15,37 | 23,71 | 31,74 |
| Toutes disciplines | 1,83 | 4,10 | 6,49 | 8,79 |

données Thomson Reuters, traitements OST

rapport OST-2008

4.4. L'indice d'activité par classe de citation

Les indicateurs de citation tels que l'impact relatif ou le RCR donnent une idée globale de la visibilité d'un acteur. On peut aussi s'interroger sur la manière dont un niveau de visibilité donné est atteint, soit par une présence particulière dans les classes de forte visibilité ("recherche d'excellence"), soit par l'investissement des classes compétitives sans être les plus visibles. On peut décrire cette activité par un profil dans des "classes de citation" de niveau fixé par convention aux x % d'articles les plus visibles, selon des classes de visibilité décroissante.

Les publications sont réparties en six classes selon le nombre de citations reçues :

- Classe 1 (notée 1_5) : les 5 % d'articles les plus cités au niveau mondial
- Classe 2 (notée 2_5) : les 5 % d'articles suivants
- Classe 3 (notée 3_10) : les 10 % d'articles suivants
- Classe 4 (notée 4_20) : les 20 % d'articles suivants
- Classe 5 (notée 5_20) : les 20 % d'articles suivants
- Classe 6 (notée 6_40) : les 40 % d'articles les moins cités au niveau mondial.

Cette répartition respecte approximativement la distribution très dissymétrique des citations ; de nombreux articles ne reçoivent que peu de citations et peu d'articles sont très cités. La classe notée 6_40 correspond aux articles peu ou pas cités dans la fenêtre de citation considérée.

L'indice d'activité par classe de citation de l'acteur dans la discipline considérée est calculé à partir de la répartition établie pour la référence mondiale. Pour chacune des 6 classes de visibilité, l'indice d'activité de l'acteur est égal au ratio entre le pourcentage des articles que l'acteur place dans cette classe, et le pourcentage des articles mondiaux dans cette classe, pourcentage qui peut différer de la valeur nominale en raison des ex aequo.

4.5. Les indicateurs relationnels

En règle générale, la communauté scientifique est fortement interconnectée, une partie de ces relations gardant une trace exploitable dans les publications, notamment sous forme de cosignatures d'articles (copublications).

La quantification des copublications suppose une série de choix méthodologiques. Le premier concerne le type de comptage, fractionnaire ou présence, mentionné plus haut. Les copublications sont présentées dans ce rapport en compte de présence, plus intuitif par rapport à la notion de collaboration : on fait l'hypothèse que le fait de cosigner suppose l'établissement d'un "lien" entre cosignataires, indépendamment de la présence d'autres cosignataires parfois très nombreux.

Les chiffres de copublications étant particulièrement faibles et fluctuants pour certains pays/disciplines, on prendra garde à l'interprétation des valeurs et des évolutions de cet indicateur très sensible.