

Rappels sur les indicateurs bibliométriques les plus couramment utilisés par l'OST

I. Les indicateurs de production scientifique

I.1. La part de publications

Pour un acteur géographique ou institutionnel donné et pour une année P de publication, l'indicateur « *part de publications* » est défini par le nombre de publications de l'acteur

(une institution, un pays...) rapporté au nombre de publications publiées la même année pour une référence donnée (le monde par exemple).

$$\text{Part de publications (\%)} = \frac{\text{Nombre de publications de l'année } P \text{ de l'acteur}}{\text{Nombre de publications de l'année } P \text{ de la référence}} \times 100$$

I.2. L'indice de spécialisation scientifique

Pour une année P de publication, l'indicateur « *indice de spécialisation scientifique* » est défini par la part de publications de l'acteur (une institution, un pays...) dans une discipline et

pour une référence donnée (le monde par exemple), rapportée à sa part de publications publiées la même année, dans l'ensemble des disciplines et pour la même référence.

$$\text{Indice de spécialisation} = \frac{\text{Part de publications (\%) de l'acteur dans la référence pour une discipline}}{\text{Part de publications (\%) de l'acteur dans la référence toutes disciplines confondues}}$$

Plus l'indice de spécialisation est supérieur (respectivement inférieur) à 1, plus l'acteur est « spécialisé » (respectivement

« sous-spécialisé ») par rapport à la référence, dans la discipline considérée.

II. Les indicateurs de visibilité

II.1. La part de citations observées

Le délai N pris en compte incluant l'année P de publication, la part de citations à N ans est le nombre de citations reçues en N années par les publications de l'acteur (une institution, un pays...) rapporté au nombre de citations reçues, pendant

la même durée N , par l'ensemble des publications publiées la même année dans une référence donnée (le monde ou une discipline par exemple).

$$\text{Part de citations observées à } N \text{ ans (\%)} = \frac{\text{Nombre de citations reçues par les publications de l'année } P \text{ de l'acteur pendant } N \text{ années}}{\text{Nombre de citations reçues par les publications de l'année } P \text{ de la référence pendant } N \text{ années}} \times 100$$

Lorsque $N = 2$, on parle de citations immédiates.

II.2. L'indice d'impact relatif (observé)

Le délai N pris en compte incluant l'année P de publication, l'indicateur « *indice d'impact relatif à N ans* » est défini par la part de citations reçues en N années par les publications de

l'acteur (une institution, un pays...) dans une référence donnée (le monde par exemple), rapportée à la part de ces mêmes publications dans la même référence.

$$\text{Indice d'impact relatif à } N \text{ ans} = \frac{\text{Part de citations observées (\% à } N \text{ ans de l'acteur dans une référence donnée)}}{\text{Part de publications (\%) de l'acteur dans la même référence}}$$

Par construction, l'indice d'impact relatif est égal à 1 pour la référence considérée (le monde par exemple). Lorsque l'indice est supérieur (respectivement inférieur) à 1,

les publications de l'acteur ont une meilleure (respectivement moins bonne) visibilité que la visibilité moyenne des publications de la référence.

II.3. L'indice d'impact espéré

Le délai N pris en compte incluant l'année P de publication, l'indicateur « *indice d'impact espéré à N ans* » est défini par la part de citations espérées à N années des publications de l'acteur (une institution, un pays...) dans une référence donnée (le monde par exemple), rapportée à la part de ces publications dans la même référence.

L'indice d'impact espéré est l'indice d'impact relatif qu'obtiendrait l'acteur si, dans chaque journal où il publie, si ses publications étaient citées comme la moyenne des publications du journal.

$$\text{Indice d'impact espéré à } N \text{ ans} = \frac{\text{Part de citations espérées à } N \text{ ans (\%) de l'acteur dans une référence donnée}}{\text{Part de publications (\%) de l'acteur dans la même référence}}$$

Par construction, l'indice d'impact espéré est égal à 1 pour la référence considérée (le monde par exemple). Lorsque l'indice est supérieur (respectivement inférieur) à 1,

les revues de publications de l'acteur ont une meilleure (respectivement moins bonne) visibilité espérée, en moyenne, que les revues de la référence.

II.4. Le ratio de citations relatif

Le délai N pris en compte incluant l'année P de publication, l'indicateur « *ratio de citations relatif à N ans* » (RCR) est défini par l'indice d'impact relatif de l'acteur (une institution, un pays...), rapporté à son indice d'impact espéré pour la même

référence (le monde par exemple).

Le ratio de citations relatif exprime le fait que les publications de l'acteur sont plus ou moins citées, en moyenne, que l'ensemble des publications des journaux dans lesquels il publie.

$$\text{Ratio de citations relatif à } N \text{ ans} = \frac{\text{Indice d'impact relatif à } N \text{ ans de l'acteur pour l'année } P}{\text{Indice d'impact espéré à } N \text{ ans de l'acteur pour l'année } P}$$

Lorsque le RCR est supérieur (respectivement inférieur) à 1, les publications de l'acteur ont une meilleure (respectivement

moins bonne) visibilité, en moyenne que les publications des journaux dans lesquels il a publié une année P donnée.

II.5. Les classes de visibilité scientifique

Le délai N pris en compte incluant l'année P de publication, les classes de visibilité scientifique, ou classes de citations, sont construites par percentiles décroissants de l'indice d'impact à N ans des publications dans la référence considérée (mondiale par exemple).

On peut établir les classes de différentes manières : par exemple,

- 1) par des classes englobantes : le TOP 1 des 1 % de publications les plus citées à N ans au niveau mondial une année P donnée, puis le TOP 5 des 5 % de publications les plus citées (englobant la classe précédente), puis le TOP 10 des 10 % de publications les plus citées (englobant les classes précédentes), etc.

2) par des classes juxtaposées: par exemple les six classes suivantes :

Classe 1 : les 5 % de publications les plus citées au niveau mondial

Classe 2 : les 5 % de publications suivantes

Classe 3 : les 10 % de publications suivantes

Classe 4 : les 20 % de publications suivantes

Classe 5 : les 20 % de publications suivantes

Classe 6 : les 40 % de publications les moins citées

Cette répartition est utilisée afin de respecter au mieux la distribution très dissymétrique des citations : beaucoup de publications reçoivent peu de citations et peu de publications sont très citées.

III. Les méthodes de décompte

III.1. Comment compter les publications produites en collaboration ?

Le plus souvent, une publication scientifique comporte plusieurs lignes d'adresses de laboratoires signataires, car elle a été produite par collaboration entre chercheurs de laboratoires appartenant à des institutions et des territoires différents. Se pose donc la question du mode de prise en compte de la publication pour chacun des laboratoires ayant participé à sa production.

Deux choix sont utilisés pour attribuer à un acteur (laboratoire, institution, territoire...) le décompte d'une publication dans laquelle on trouve son adresse : le compte de présence et le compte fractionnaire.

Le compte de présence est utilisé lorsque l'on s'intéresse à la **participation** d'un acteur à la production scientifique. On prend en compte pour 1 chacune des publications dans laquelle l'adresse de cet acteur apparaît, sans tenir compte du nombre total d'adresses de laboratoires signataires.

Le compte de présence est souvent privilégié pour son intuitivité. Il reflète la dimension cognitive de l'activité d'un acteur, et il ne décourage pas les collaborations. Toutefois, il n'est pas additif, notamment dans les changements d'échelle géographique.

Le compte fractionnaire est utilisé lorsque l'on s'intéresse à la **contribution** d'un acteur à la production scientifique, afin d'appréhender son poids scientifique. En ce cas, on prend en compte, pour chaque adresse de l'acteur, la fraction de compte que représente chaque adresse de la publication - étant donné le nombre total d'adresses qu'elle comporte.

En raison de ses propriétés de stabilité quand on change de niveau d'observation, on utilise souvent ce compte pour des études descriptives systématiques.

II.6. L'Indice d'activité par classe de visibilité

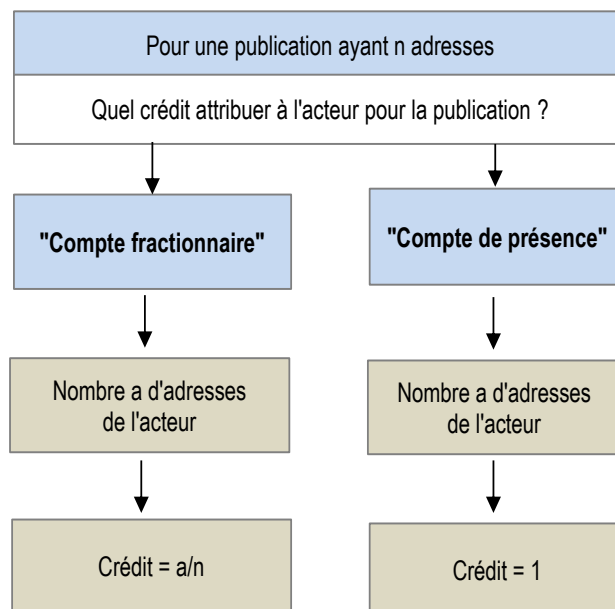
Pour une année P , l'indicateur « *indice d'activité* » est défini, pour une classe i de visibilité donnée, par la part des publications de l'acteur (une institution, un pays...) dans la classe de visibilité à N ans étudiée, rapportée à la part des publications de la référence (le monde par exemple) dans la même classe.

Indice d'activité dans une classe i =

$$\frac{\text{Part des publications (\%) de l'acteur dans la classe } i}{\text{Part des publications (\%) de la référence dans la classe } i}$$

Lorsque l'indice d'activité pour une classe de citations donnée, est supérieur (respectivement inférieur) à 1, l'acteur est plus (respectivement moins) présent qu'attendu d'après la valeur moyenne de la référence.

Figure 1 : le choix de comptage des publications ayant plusieurs adresses



III.2. Comment compter les publications ayant une multi-affiliation disciplinaire ?

Les publications sont étiquetées avec une ou plusieurs disciplines, en général celles de la revue scientifique support (mais dans certains cas par un étiquetage au niveau de la publication elle-même).

Deux choix sont utilisés pour décompter une publication affiliée à plusieurs disciplines : le compte disciplinaire de présence et le compte disciplinaire fractionnaire.

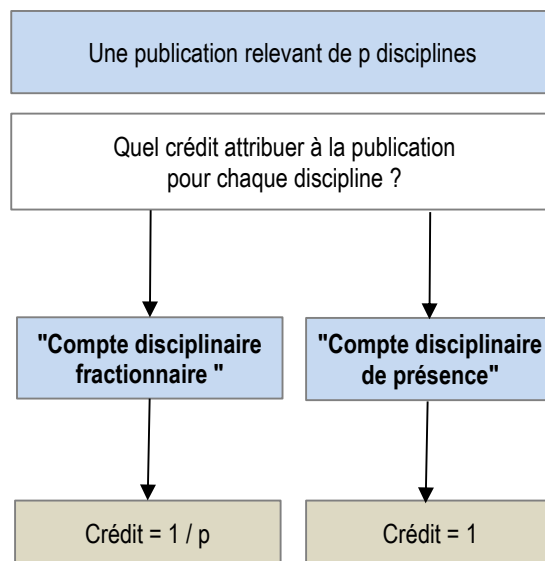
Quand on utilise le **compte disciplinaire de présence**, une publication étiquetée avec p disciplines est comptée pour 1 dans chacune de ces disciplines, dans une logique de participation.

Ce mode de comptage est utilisé pour son intuitivité. Il ne décourage pas les travaux dans les champs disciplinaires mixtes.

Quand on utilise le **compte disciplinaire fractionnaire**, une publication étiquetée avec p disciplines est comptée $1/p$ pour chaque discipline, dans une logique de contribution.

Ce mode de comptage est utilisé pour les travaux descriptifs systématiques car il évite de multiplier les doubles comptes disciplinaires.

Figure 2 : comptage des publications multi-affiliées par discipline



A lire également : les fiches méthodologiques et les fiches indicateurs sur www.obs-ost.fr

Ghislaine Filliatreau, Chris Roth
Ce document a bénéficié de l'assistance éditoriale de Marie-Laure Taillibert
Correspondance : ghislaine.filliatreau@obs-ost.fr