

Production coopérative d'indicateurs de politique scientifique  
des institutions de recherche françaises

**Indicateurs bibliométriques des institutions publiques  
de recherche françaises**

**(2000 – hors sciences humaines et sociales)**

Groupe « IPSIR » : publications scientifiques

Travail réalisé à l'OST

- ◆ Coordonné par : Michel Zitt (OST-Inra)  
Serge Bauin (UNIPS-CNRS)  
Ghislaine Filliatreau (OST)
  
- ◆ avec les représentants des institutions suivantes : CEA (F. Rousseau-Hans, D. de Prunelé, N. Reynier), Cemagref (N. Delherbe, O. Hologne), Cirad (L. Grasset), CNRS (M. Crance), Inra (E. Bassecoulard, C. Sireyjol, H. Pampouille), Inserm (N. Barberousse, N. Pinhas), IRD (M. Stjepanovic), Inria (J-C. Le Moal), Institut Pasteur (A. Gouyette, ), et la participation de la Conférence des Présidents d'Université (E. Esperet).
  
- ◆ la collaboration de S. Ramanana-Rahary (OST)
  
- ◆ et l'aide de N. Teixeira et E. Vachon (OST)

*Décembre 2004*

## INTRODUCTION

Dans le cadre de la "production coopérative d'indicateurs de politique scientifique", le groupe de travail sur les publications scientifiques ("IPSIR") a pour mission d'élaborer des indicateurs bibliométriques pour les organismes publics de recherche et les universités<sup>1</sup>.

Les indicateurs bibliométriques sont calculés à partir des articles publiés par les laboratoires de recherche dans les journaux scientifiques. Les articles scientifiques, qui contiennent des "connaissances certifiées", sont l'une des productions majeures des laboratoires de recherche académique, sachant qu'il y a, bien entendu, d'autres dimensions dans l'activité des laboratoires.

Les résultats présentés ici pour l'année 2000 font suite à une première étude exploratoire réalisée précédemment sur les publications de l'année 1997. Ils concernent une partie des institutions de recherche du secteur public et, sauf dans le cas des universités, ces institutions ont repéré elles-mêmes leurs articles dans le *Science Citation Index*, la base internationale de référence utilisée par l'OST.

## DONNÉES ET MÉTHODES

### *Source*

La base de données utilisée est le *Science Citation Index (SCI)* dans sa version la plus large (*Expanded*), qui fait référence pour la bibliométrie dans les domaines des sciences de la matière et de la vie. C'est une base très sélective, plutôt orientée vers l'activité scientifique académique et considérée comme représentative pour les disciplines bien internationalisées. Sa représentativité est moins bonne dans les disciplines appliquées, de terrain, à forte tradition nationale, et dans les disciplines (notamment l'informatique) où une part importante des informations passent par d'autres canaux que l'article.

Les SHS ne sont pas traitées dans ce rapport, notamment parce que les bases ISI de sciences humaines et sociales (SSCI ou AHCI) présentent dans certaines disciplines des biais de couverture considérables.

### *Types de documents*

Les indicateurs portent sur tous les types de document susceptibles d'être cités : articles, articles de synthèse, lettres, articles reprenant des comptes-rendus de congrès (CRC ou "*proceedings*") qui figurent par exemple dans des numéros spéciaux de revues régulières.

### *Classification disciplinaire*

La classification disciplinaire utilisée est la classification standard de l'OST, et résulte d'une agrégation des spécialités proposées par le SCI ("*subject categories*") et définies au niveau des journaux (voir Annexe). L'informatique n'y est pas classée dans les mathématiques mais dans les sciences pour l'ingénieur ; les sciences de l'univers, outre l'astronomie et les géosciences, comprennent les sciences et technologies de l'environnement et, au titre de la "terre fluide", l'océanographie, la biologie marine et l'hydrobiologie.

---

<sup>1</sup> Outre les institutions dont l'activité est étudiée dans ce rapport, le groupe de travail comprend l'Ifremer et le Cnes.

La catégorie ISI « multidisciplinaire » et la discipline OST correspondante sont hétérogènes, comprenant à la fois des revues très prestigieuses et de fort impact bibliométrique (Nature, Science, PNAS...) et des revues de vulgarisation. Dans cette catégorie, certains types d'indicateurs (« parts de marché » des citations) sont interprétables, mais d'autres (« part de marché » des publications, impact) sont plus délicats à interpréter. Nous avons choisi de ne pas faire apparaître la colonne « multidisciplinaire » dans les tableaux, mais les totaux en tiennent compte.

### ***Repérage des institutions***

Hors universités, les institutions participantes ont repéré les publications de leurs laboratoires dans le SCI. Le repérage des publications des universités a été coordonné par l'OST dans le cadre d'un contrat financé par le ministère en charge de la Recherche.

Dans cet exercice, les institutions se sont identifiées pour l'année 2000 au niveau de l'adresse des laboratoires signataires des articles, et non pas seulement au niveau de l'article.

#### Principe de repérage

Plusieurs cas de figure se présentent :

- 1 - le laboratoire - et donc son appartenance - est identifié sans ambiguïté.
- 2 - le laboratoire n'est pas identifié, mais le nom de l'institution figure dans l'adresse.
- 3 - l'adresse se réfère à une structure qui englobe plusieurs unités de recherche (par exemple un observatoire, un IFR, un site, ...)
- 4 - un auteur de l'organisme est signataire sans qu'on retrouve une adresse de laboratoire de l'organisme.

Le cas 4 n'est pas pris en compte, le cas 3 ne l'est que lorsque d'autres indices (le plus souvent le nom des auteurs) autorisent l'institution à conclure que l'article émane bien d'un de ses laboratoires.

Il faut noter qu'en conséquence un article d'un auteur étranger mentionnant une adresse institutionnelle française est décompté, mais un expatrié d'une institution française omettant de la mentionner comme affiliation n'est pas décompté. Dans le cas particulier des institutions à forte implantation à l'étranger, la présente version du rapport ne considère que la partie métropolitaine.

#### Détail par institution

CEA	Sont pris en compte les laboratoires du CEA civil et de défense, ainsi que ceux de l'IPSN, et les unités mixtes. Les adresses LURE et CYCERON ne sont pas prises en compte dans le périmètre strict choisi.
Cemagref	Les affiliations Cemagref, à quelques unités près, sont correctement identifiées dans le SCI.
CNRS	Sont prises en compte les unités évaluées et financées par le CNRS, qu'elles soient propres, mixtes ou associées.
Inra	Sont prises en compte les unités de recherche, propres ou associées, répertoriées dans l'annuaire des structures de l'Inra, y compris pour l'enseignement supérieur agricole. Ont été également incluses les publications où ne sont mentionnées que des

	superstructures ou filiales à participation Inra, si les publications en question figurent dans la base interne de l'Inra, Pubinra.
Inserm	Sont prises en compte les publications des "formations de recherche" (unités, équipes, centres d'investigation clinique...) de l'Inserm. Les documents déclarés par le laboratoire avec adresse de type hôpital ou service clinique, et répertoriés dans la Banque des Publications Inserm (BPI) ont été pris en compte. Les publications des chercheurs Inserm hors formations n'ont pas été comptabilisées.
IRD	Ne sont prises en compte, dans le présent exercice, que les adresses IRD correspondant aux unités propres ou mixtes localisées en France métropolitaine, ce qui représente environ les deux tiers des publications de l'IRD.
Universités	Ont été considérées toutes les structures évaluées et/ou financées par le ministère de l'Éducation nationale, de la Recherche et de la Technologie, ce qui inclus donc les unités mixtes. Cette mixité explique d'ailleurs les difficultés de repérage rencontrées pour des structures comme les instituts de recherche ou les observatoires, dans lesquels cohabitent des unités mixtes et des unités non mixtes. Ces difficultés n'ont pas toujours pu être résolues de manière incontestable et, dans un souci de repérage scrupuleux, le choix a été fait d'approcher les résultats par défaut lorsque l'adresse ne permettait pas de reconnaître l'appartenance universitaire du laboratoire. C'est le cas notamment pour un certain nombre d'adresses de laboratoires appartenant à des observatoires astronomiques ou des structures complexes comme le LURE.
Institut Pasteur	Pour l'Institut Pasteur comme pour l'IRD et le Cirad, l'exploitation est limitée dans cette version aux adresses françaises de l'institution.
Inria	Le principe de repérage dans des journaux scientifiques exclusivement sous-estime la production de l'Inria, spécialisé dans un domaine où d'autres modes de production sont très importants, notamment les communications dans des congrès.

Comme on le sait, dans le système français de production scientifique, *la mixité (c'est-à-dire la co-affiliation institutionnelle d'un même laboratoire à plusieurs institutions) est très importante*. La production des structures mixtes (unités de recherche dont l'affiliation est pluri-institutionnelle) est attribuée à chacune des institutions d'affiliation, ce qui veut dire - par exemple - que les données relatives aux universités recouvrent une grande part de celles du CNRS, puisque la majorité des unités CNRS sont mixtes avec l'université.

Pour fixer les idées sur la taille respective des productions des différentes institutions, le tableau 1 donne le nombre de publications par institution qui a été utilisé comme base pour les indicateurs de ce rapport.

**Tableau 1** : Nombre de publications où apparaît l'institution, poids dans la France, dans l'UE15 et dans le monde (en 2000, en %)

	Nombre	Part nationale	Part UE 15	Part mondiale
CEA	2 996	6,2	1,1	0,40
Cemagref	67	0,1	0,0	0,01
CNRS	24 167	49,9	8,5	3,23
Inra	2 451	5,1	0,9	0,33
Inserm	5 297	10,9	1,9	0,71
IRD	326	0,7	0,1	0,04
Universités	32 715	67,5	11,6	4,37
Cirad	162	0,3	0,1	0,02
Inria	363	0,7	0,1	0,05
Institut Pasteur	898	1,9	0,3	0,12
France		100,0	100,0	100,0
France (Nombre)		48 453	283 136	749 344

(Données ISI, traitement OST)

\* Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

### **Principes de comptage**

Selon les indicateurs, l'OST a utilisé deux types classiques de compte, le compte de présence ou le compte fractionnaire.

Le *compte de présence* est utilisé pour l'ensemble des indicateurs, à l'exception des ratios d'impact. Le *compte de présence dénombre les articles où l'institution apparaît au moins une fois dans la liste des adresses d'affiliation*, c'est une mesure de "participation" à la production scientifique. On ne fait pas la différence entre la présence d'un seul ou de plusieurs laboratoires de cette institution dans la liste des adresses.

En *compte de présence*, les nombres ou pourcentages ne peuvent être consolidés entre institutions, et la somme des parts françaises des différentes institutions est bien supérieure à 100 % en raison du grand nombre d'articles produits par des laboratoires mixtes, auxquels s'ajoutent bien entendu les articles produits en collaboration -et donc co-signés- par des laboratoires d'institutions différentes. Le compte de présence, contrairement au compte fractionnaire, avantage les institutions à forte mixité. Un autre avantage est sa simplicité d'interprétation.

L'absence d'additivité introduit une instabilité lorsque l'on change le niveau d'observation (du laboratoire à l'institution, de la région au pays, etc.) : par exemple l'indicateur UE15 est en général inférieur à la somme des indicateurs des pays composants.

Cela introduit aussi des biais par discipline, au profit de celles où les articles ont en moyenne un plus grand nombre d'institutions auteurs. Dans le présent rapport, lorsqu'une analyse est effectuée en compte de présence, la logique a été étendue à la définition de la discipline : les journaux sont assignés à plusieurs spécialités/disciplines scientifiques sont décomptés en entier dans chaque catégorie.

Le *compte fractionnaire* répond à une logique de "contribution" à l'activité scientifique, cette contribution étant calculée au prorata de la contribution relative d'un acteur dans la liste des adresses d'affiliation. Contrairement au compte de présence, le compte fractionnaire est consolidable à toutes les échelles. Bien adapté à la macro-analyse, il est couramment utilisé par l'OST, par exemple dans son rapport biennal. En micro-analyse d'institutions où celles-ci peuvent présenter de forts recouvrements (p.ex. CNRS-Université), le fractionnement pénalise le chiffre de production de chaque institution. Cet

inconvenient disparaît pratiquement pour les indicateurs d'impact mesurant la visibilité par les citations reçues.

### ***Contexte et limites des résultats***

Il faut interpréter avec prudence les données bibliométriques (par exemple, s'il s'agit d'établir un profil disciplinaire pour un institut de taille modeste), tout particulièrement pour certains indicateurs dont les distributions statistiques sont atypiques, comme les impacts, où les données relatives à un seul individu peuvent affecter la statistique générale.

L'interprétation de certains indicateurs comme les parts de marché doit naturellement tenir compte des tailles d'institutions. On pourra notamment se reporter au document "La recherche scientifique française : les enseignants-chercheurs et chercheurs des EPST", publié par l'OST en avril 2002<sup>2</sup> et portant sur les données de l'année 2000. En ne s'intéressant qu'aux enseignants-chercheurs hors SHS, on y voit par exemple que le personnel chercheurs des universités représente environ 30 140 personnes physiques ou 15 070 "équivalents temps plein recherche", contre environ 9 200 pour le CNRS, près de 1 960 pour l'Inserm, environ 1 500 pour l'Inra, 590 pour l'IRD, 320 pour l'Inria et moins de 100 pour le Cemagref.

Cependant, les données de personnels ne sont pas toujours simples à utiliser : ainsi, pour reprendre les cas du CNRS et des universités, l'annexe chiffrée du rapport d'activités 2000 du CNRS<sup>3</sup> montre que, hors SHS, 8 300 chercheurs du CNRS seulement sur les 9 400 que compte l'organisme travaillent dans des laboratoires repérables comme appartenant au CNRS alors que, dans ces mêmes laboratoires travaillent également de très nombreux enseignants-chercheurs, représentant 6 200 équivalents temps plein recherche. Par ailleurs, ces chiffres ne tiennent pas compte des nombreuses autres catégories de personnels présents dans les laboratoires et susceptibles de publier (ingénieurs de recherche, visiteurs étrangers, étudiants en thèse ou en stage etc.).

Un autre avertissement concerne les universités, qui constituent un ensemble d'établissements hétérogènes. À titre d'illustration nous avons inclus, dans le chapitre consacré aux comptages de citations, une comparaison entre deux universités qui montre bien que les chiffres globaux donnés pour les universités masquent une grande diversité.

### ***Indicateurs***

Nous présentons successivement les trois familles classiques d'indicateurs bibliométriques :

- la production scientifique : le décompte des publications et la spécialisation (section I),
- les collaborations scientifiques : les co-publications (section II),
- la visibilité : le comptage des citations et les différentes mesures d'impact (section III).

Chaque section inclut les précisions méthodologiques spécifiques.

---

<sup>2</sup> [http://www.obs-ost.fr/pub/DemoCher\\_300402.PDF](http://www.obs-ost.fr/pub/DemoCher_300402.PDF)

<sup>3</sup> <http://www.cnrs.fr/cw/fr/band/cnrs/rapports/rap00/chiffres.pdf>

## I. LA PRODUCTION SCIENTIFIQUE

Les tableaux I.1a et I.1b permettent une première approche des positionnements des institutions, en nombres et en pourcentages (non additifs) par rapport à la France

### Volume de production et profils disciplinaires

**Tableau I.1a** : Nombre de publications, par discipline (en 2000)

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths	Toutes
CEA	443	246	61	633	1 468	293	400	21	2 996
Cemagref	2	2	36	2	4	32	25		67
CNRS	4 533	1 943	1 501	6 291	6 652	3 004	3 039	1 570	24 167
Inra	1 104	461	1 251	233	14	146	84	23	2 451
Inserm	2 550	3 328	154	110	45	15	37	12	5 297
IRD	107	43	131	10		107	22	1	320
Universités	7 152	10 428	2 090	6 009	5 428	2 423	3 107	2 066	32 715
Cirad	59	14	115	18		13	11	1	162
Inria	8	3	5		18	5	284	100	363
Institut Pasteur	550	406	43	14	1	1	2	1	898
France	10 064	14 926	3 828	8 303	8 587	4 003	5 472	2 427	48 453

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

**Tableau I.1b** : Part nationale de publications, par discipline (en 2000, en %)

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths	Toutes
CEA	4,4	1,6	1,6	7,6	17,1	7,3	7,3	0,9	6,2
Cemagref	0,0	0,0	0,9	0,0	0,0	0,8	0,5		0,1
CNRS	45,0	13,0	39,2	75,8	77,5	75,0	55,5	64,7	49,9
Inra	11,0	3,1	32,7	2,8	0,2	3,6	1,5	0,9	5,1
Inserm	25,3	22,3	4,0	1,3	0,5	0,4	0,7	0,5	10,9
IRD	1,1	0,3	3,4	0,1		2,7	0,4	0,0	0,7
Universités	71,1	69,9	54,6	72,4	63,2	60,5	56,8	85,1	67,5
Cirad	0,6	0,1	3,0	0,2		0,3	0,2	0,0	0,3
Inria	0,1	0,0	0,1		0,2	0,1	5,2	4,1	0,7
Institut Pasteur	5,5	2,7	1,1	0,2	0,0	0,0	0,0	0,0	1,9
France	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Les universités signent ou co-signent les deux tiers des articles français, le CNRS la moitié. Les universités et le CNRS, tous deux à vocation multidisciplinaire et largement associés institutionnellement, sont naturellement présents partout, avec quelques nuances : le CNRS est par exemple très présent en sciences physiques et moins en recherche médicale, les universités sont très

actives en mathématiques, chimie, biologie fondamentale et recherche médicale, et moins actives en biologie appliquée.

Le CEA, dont la discipline dominante est la physique, a une activité importante en chimie, sciences de l'univers, sciences pour l'ingénieur et aussi biologie fondamentale.

Le Cemagref apparaît en biologie appliquée et sciences de l'univers, puis en sciences pour l'ingénieur.

Plusieurs instituts finalisés ont chacun deux disciplines majeures (tableau I.2) : biologie appliquée et sciences de l'univers pour l'IRD, biologie appliquée/écologie et biologie fondamentale pour l'Inra, biologie fondamentale et recherche médicale pour l'Inserm et l'Institut Pasteur, sciences pour l'ingénieur (comprenant l'informatique) et mathématiques pour l'Inria. Le Cirad est principalement présent en biologie appliquée/écologie, et à un moindre degré en biologie fondamentale.

**Tableau I.2** : Profils disciplinaires des institutions (en 2000, en %)

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths	Toutes
CEA	14,8	8,2	2,0	21,1	49,0	9,8	13,4	0,7	100,0
Cemagref	3,0	3,0	53,7	3,0	6,0	47,8	37,3		100,0
CNRS	18,8	8,0	6,2	26,0	27,5	12,4	12,6	6,5	100,0
Inra	45,0	18,8	51,0	9,5	0,6	6,0	3,4	0,9	100,0
Inserm	48,1	62,8	2,9	2,1	0,8	0,3	0,7	0,2	100,0
IRD	33,4	13,4	40,9	3,1		33,4	6,9	0,3	100,0
Universités	21,9	31,9	6,4	18,4	16,6	7,4	9,5	6,3	100,0
Cirad	36,4	8,6	71,0	11,1		8,0	6,8	0,6	100,0
Inria	2,2	0,8	1,4		5,0	1,4	78,2	27,5	100,0
Institut Pasteur	61,2	45,2	4,8	1,6	0,1	0,1	0,2	0,1	100,0
France	20,8	30,8	7,9	17,1	17,7	8,3	11,3	5,0	100,0

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

L'indice de spécialisation (tableau I.3) compare, pour chaque institution, son poids dans la France par discipline à son poids toutes disciplines confondues<sup>4</sup>. Il exprime l'engagement relatif de l'institution dans les divers secteurs de connaissance.

<sup>4</sup> C'est aussi le ratio de la part d'activité que l'institution consacre à une discipline à la part d'activité de la France dans cette discipline.

**Tableau I.3** : Indices de spécialisation par rapport à la France (en 2000)

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths	Toutes
CEA	0,71	0,27	0,26	1,23	2,76	1,18	1,18	0,14	1,00
Cemagref	0,14	0,10	6,80	0,17	0,34	5,78	3,30		1,00
CNRS	0,90	0,26	0,79	1,52	1,55	1,50	1,11	1,30	1,00
Inra	2,17	0,61	6,46	0,55	0,03	0,72	0,30	0,19	1,00
Inserm	2,32	2,04	0,37	0,12	0,05	0,03	0,06	0,05	1,00
IRD	1,61	0,44	5,18	0,18		4,05	0,61	0,06	1,00
Universités	1,05	1,03	0,81	1,07	0,94	0,90	0,84	1,26	1,00
Cirad	1,75	0,28	8,99	0,65		0,97	0,60	0,12	1,00
Inria	0,11	0,03	0,17		0,28	0,17	6,93	5,50	1,00
Institut Pasteur	2,95	1,47	0,61	0,09	0,01	0,01	0,02	0,02	1,00
France	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

(Données ISI, traitement OST)

*Grisés : chiffres peu significatifs, correspondant à moins de 50 publications*

*Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.*

On retrouve les points forts attendus pour les instituts spécialisés, qu'ils aient une activité bi-disciplinaire, comme la plupart des institutions spécialisées, ou un peu plus large, comme le CEA. Les universités et le CNRS, de profils logiquement assez proches et moins accentués, privilégient la physique et les sciences pour l'ingénieur pour le CNRS. Les deux institutions ont un ratio d'activité élevé en mathématiques, une spécialité française.

## Répartition régionale

La répartition régionale, en compte de présence, est présentée dans le tableau I.4.

**Tableau I.4** : Répartition régionale des publications, par institution (en 2000, en %)

Région	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	France
Ile-de-France	67,4	16,4	38,7	28,6	51,5	24,1	39,4	3,7	29,8	100,0	42,9
Champ.-Ard.			0,4	0,7	0,3		1,1				1,0
Picardie			0,6	0,8			1,3				1,3
Haute-Norm.			0,9	0,0	1,6		1,8				1,5
Centre	1,1	3,0	1,9	8,0	0,8	0,6	2,1		0,3		2,5
Basse-Norm.	4,5		1,6	0,2	0,5		1,7				1,5
Bourgogne	0,7		1,1	4,6	0,8		1,6				1,6
N.-P.-de-Calais			3,2	1,0	4,4		5,0				4,2
Lorraine			3,4	4,6	1,2	0,9	3,6		17,1		3,6
Alsace	0,1		7,2	0,5	8,0	0,6	5,6				5,4
Fr.-Comté			0,7	0,3	0,1		1,3				1,2
Pays de La Loire			1,9	6,0	2,3		3,1				3,3
Bretagne		4,5	3,4	8,9	1,5	10,0	3,5	9,9	15,7		4,0
Poitou-Charentes			1,8	2,4			1,6				1,6
Aquitaine	0,8	13,4	4,7	7,3	4,1	0,3	5,0				4,7
Midi-Pyrénées			6,5	9,9	3,8	2,5	5,0				6,0
Limousin			0,6	0,2			1,0				0,8
Rhône-Alpes	22,6	31,3	15,9	3,3	10,1	4,7	13,4		11,0		15,8
Auvergne		7,5	1,6	8,6	1,0		1,8				2,0
Lang.-Rouss.	1,5	17,9	6,1	9,2	4,3	48,8	5,8	87,0			5,8
PACA	5,5	7,5	9,0	5,5	8,4	10,6	6,8		29,8		8,8
Corse			0,1	0,0			0,2	0,6			0,1
France	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Rappel nombre	2 996	67	24 167	2 451	5 297	320	32 715	162	363	898	48 453

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 10 articles

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Ces données indiquent une certaine variété des situations par rapport à celle de la France, classiquement marquée par une forte concentration régionale des moyens en Ile-de-France. Ici, en compte de présence (logique de "participation") le poids de l'Ile-de-France est de 43 % pour l'ensemble traité (dernière colonne du tableau). Le chiffre serait moindre en compte fractionnaire, selon une logique de "contribution".

Par ordre de concentration décroissante sur ce critère du pourcentage en Ile-de-France, on trouve l'Institut Pasteur, à production exclusivement francilienne, puis le CEA francilien pour les deux tiers. L'Inserm est légèrement au-dessus de 50 %. Viennent ensuite le CNRS et les universités autour de 40 %. Ces institutions, ainsi que l'Inserm, ont cependant des implantations dans pratiquement toutes les régions.

Pour l'Inra, avec de multiples implantations régionales, et l'Inria, essentiellement présente sur quatre régions, l'Ile-de-France ne représente qu'environ 30 % de la production totale.

Enfin, il existe des institutions dont l'implantation principale n'est pas en Ile-de-France : Grenoble pour le Cemagref, Montpellier pour l'IRD et plus encore le Cirad.

## II – LES COOPERATIONS SCIENTIFIQUES

### 2.1 Collaboration entre institutions : co-publication et mixité institutionnelle

L'activité scientifique est par nature coopérative, que ce soit à l'intérieur des équipes de recherche ou entre équipes. On réserve le terme de "co-publications" aux publications résultant de collaborations entre laboratoires différents, collaborations ayant débouché sur la rédaction d'un article en commun. Ainsi la co-présence d'institutions sur un même article mais avec des adresses différentes identifie une "co-publication" tandis que la co-présence d'institutions sur un même article et sur une même adresse identifie une collaboration par mixité (co-présence structurelle ou co-affiliation)<sup>5</sup>. Les données sont présentées dans trois types de tableau, correspondant d'une part au total des co-présences et d'autre part à ses deux formes<sup>6</sup>. En "compte de présence" (donc en nombre d'articles) la somme des co-présences structurelles et des co-publications entre deux acteurs A et B peut être supérieure au nombre de leur co-présences totales.

La situation de la recherche a largement évolué depuis le rapport qui portait sur l'année 1997, en particulier du fait de la généralisation rapide de la mixité entre laboratoires. Cette mixité est devenue quasi systématique entre les universités et le CNRS, mais elle se généralise également entre les universités et d'autres EPST comme l'Inserm et, de plus en plus souvent, entre plusieurs de ces partenaires.

**Tableau II.1** : Co-présences totales entre les institutions participantes (nombres de publications en 2000)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	Tot .Partenar. FRA.	Rappel total Publi
CEA	337		1 756	47	131	7	1 253			17	2 155	2 996
Cemagref		1	22	9			33	3			51	67
CNRS	1 756	22	5 083	460	1 208	151	18 027	26	162	386	20 015	24 167
Inra	47	9	460	397	130	23	906	34	3	32	1 477	2 451
Inserm	131		1 208	130	652	7	4 687		2	225	5 002	5 297
IRD	7		151	23	7	18	152	28	1	2	240	320
Universités	1 253	33	18 027	906	4 687	152	9 479	40	95	461	26 427	32 715
Cirad		3	26	34		28	40	12			91	162
Inria			162	3	2	1	95		20		207	363
Institut Pasteur	17		386	32	225	2	461			149	661	898

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Attention : données non additives (compte de présence)

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Ces chiffres bruts peuvent être ramenés, pour chaque institution, à l'ensemble des partenariats qu'elle entretient avec les autres.

<sup>5</sup> Cette distinction a une limite : nous sommes conduits à considérer comme "co-publication" les cas où pour un seul acteur plusieurs affiliations institutionnelles sont mentionnées.

<sup>6</sup> Par exemple, une publication avec une adresse strictement CEA et une adresse de mixité CEA-CNRS va générer à la fois une co-publication CEA-CNRS et une co-présence institutionnelle CEA-CNRS

**Tableau II.2 :** Co-présences totales entre institutions, en % des partenariats 'nationaux' de chaque institution (en 2000)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	Tot. Partenar. FRA	Tot. Partenar. FRA	Rappel total Publi
CEA	15,6		81,5	2,2	6,1	0,3	58,1			0,8	100,0	2 155	2 996
Cemagref		2,0	43,1	17,6			64,7	5,9			100,0	51	67
CNRS	8,8	0,1	25,4	2,3	6,0	0,8	90,1	0,1	0,8	1,9	100,0	20 015	24 167
Inra	3,2	0,6	31,1	26,9	8,8	1,6	61,3	2,3	0,2	2,2	100,0	1 477	2 451
Inserm					13,								
	2,6		24,2	2,6	0	0,1	93,7		0,0	4,5	100,0	5 002	5 297
IRD	2,9		62,9	9,6	2,9	7,5	63,3	11,7	0,4	0,8	100,0	240	320
Universités					17,								
	4,7	0,1	68,2	3,4	7	0,6	35,9	0,2	0,4	1,7	100,0	26 427	32 715
Cirad						30,							
		3,3	28,6	37,4		8	44,0	13,2			100,0	91	162
Inria			78,3	1,4	1,0	0,5	45,9		9,7		100,0	207	363
Institut Pasteur	2,6		58,4	4,8	0	0,3	69,7			22,5	100,0	661	898

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Attention : données non additives (compte de présence)

La référence (total des partenariats français) est l'ensemble des « articles en partenariat français » avec au moins une adresse des institutions de la coopérative actuellement renseignés.

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Lecture : dans la ligne CEA, le chiffre de la colonne CNRS indique que pour 100 « articles en partenariat français » avec une adresse CEA, 81,5 ont une co-signature CNRS, alors que la colonne CEA indique que 15,6 ont une co-signature d'une autre adresse CEA

Ces co-présences se répartissent (avec recouvrements au niveau des articles) entre co-présences institutionnelles (mixité) et co-publications, présentées dans les tableaux suivants.

**Tableau II.3a :** Co-présences liées à la mixité entre institutions (nombres de publications en 2000)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	Tot. Partenar. FRA	Rappel total Publi
CEA			1 077	21	52		431				2 155	2 996
Cemagref				1			7				51	67
CNRS	1 077			230	500	101	17 022		116	308	20 015	24 167
Inra	21	1	230		51	11	563	12			1 477	2 451
Inserm	52		500	51		2	4273			120	5 002	5 297
IRD			101	11	2		81	18			240	320
Universités	431	7	17 022	563	4 273	81		3	33	294	26 427	32 715
Cirad				12		18	3				91	162
Inria			116				33				207	363
Institut Pasteur			308		120		294				661	898

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Attention : données non additives (compte de présence)

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

**Tableau II.3b** : Co-présences liées à la mixité entre institutions, en % des partenariats 'nationaux' de chaque institution (en 2000)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Univer-sités	Cirad	Inria	Institut Pasteur	Tot. Partenar. FRA	Tot Partenar. FRA	Rappel total Publi
CEA			50,0	1,0	2,4		20,0				100,0	2 155	2 996
Cemagref				2,0			13,7				100,0	51	67
CNRS	5,4			1,1	2,5	0,5	85,0		0,6	1,5	100,0	20 015	24 167
Inra	1,4	0,1	15,6		3,5	0,7	38,1	0,8			100,0	1 477	2 451
Inserm	1,0		10,0	1,0		0,0	85,4			2,4	100,0	5 002	5 297
IRD			42,1	4,6	0,8		33,8	7,5			100,0	240	320
Univer-sités	1,6	0,0	64,4	2,1	16,2	0,3	0,0	0,0	0,1	1,1	100,0	26 427	32 715
Cirad				13,2		19,8	3,3				100,0	91	162
Inria			56,0				15,9				100,0	207	363
Institut Pasteur			46,6		18,2		44,5				100,0	661	898

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Attention : données non additives (compte de présence)

La référence (total des partenariats français) est l'ensemble des « articles en partenariat français » avec au moins une adresse des institutions de la coopérative actuellement renseignés.

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Lecture : dans la ligne CEA, le chiffre de la colonne CNRS indique que pour 100 « articles en partenariat français » avec une adresse CEA, 50 % sont liés à la mixité institutionnelle CEA-CNRS.

**Tableau II.4a** : Co-présences liées aux co-publications entre les institutions participantes (nombres de publications en 2000)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	Tot. Partenar. FRA	Rappel total Publi
CEA	337		1 205	35	93	7	1 049			17	2 155	2 996
Cemagref		1	22	8			28	3			51	67
CNRS	1 205	22	5 083	319	875	86	7 088	26	71	195	20 015	24 167
Inra	35	8	319	397	93	15	559	26	3	32	1 477	2 451
Inserm	93		875	93	652	5	2 595		2	149	5 002	5 297
IRD	7		86	15	5	18	100	15	1	2	240	320
Universités	1 049	28	7 088	559	2 595	100	9 479	39	75	309	26 427	32 715
Cirad		3	26	26		15	39	12			91	162
Inria			71	3	2	1	75		20		207	363
Institut Pasteur	17		195	32	149	2	309			149	661	898

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

Attention : données non additives (compte de présence)

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

**Tableau II.4b** : Co-présences liées aux co-publications entre les institutions participantes (en % des partenariats 'nationaux' de chaque institution, en 2000)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	Total Coop-10	Total Coop-10
CEA	15,6		55,9	1,6	4,3	0,3	48,7			0,8	100,0	2 075
Cemagref		2,0	43,1	15,7			54,9	5,9			100,0	52
CNRS	6,0	0,1	25,4	1,6	4,4	0,4	35,4	0,1	0,4	1,0	100,0	19 568
Inra	2,4	0,5	21,6	26,9	6,3	1,0	37,8	1,8	0,2	2,2	100,0	1 293
Inserm	1,9		17,5	1,9	13,0	0,1	51,9		0,0	3,0	100,0	5 304
IRD	2,9		35,8	6,3	2,1	7,5	41,7	6,3	0,4	0,8	100,0	318
Universités	4,0	0,1	26,8	2,1	9,8	0,4	35,9	0,1	0,3	1,2	100,0	23 673
Cirad		3,3	28,6	28,6		16,5	42,9	13,2			100,0	94
Inria			34,3	1,4	1,0	0,5	36,2		9,7		100,0	204
Institut Pasteur	2,6		29,5	4,8	22,5	0,3	46,7			22,5	100,0	638

(Données ISI, traitement OST)

*Co-présences non additives (compte de présence)*

*La référence (total des partenariats français) ne couvre que les partenaires de la coopérative actuellement renseignés.*

*Grisés : chiffres peu significatifs, correspondant à moins de 50 articles*

*Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.*

*Lecture : dans la ligne CEA, le chiffre de la colonne CNRS indique que pour 100 articles en partenariat français avec adresse CEA, 55.9 % sont en co-opublication avec CNRS*

Selon les couples d'institutions, le rapport entre les deux types de co-publications (co-publication par collaboration et co-publication par mixité) est assez différent.

Le recouvrement massif du CNRS et des universités est majoritairement lié à la mixité entre ces deux institutions, mais on note également une intense activité de co-publication. Pour les autres partenariats, la place respective de la mixité et des collaborations est variable :

- Universités et Inserm : le recouvrement est principalement lié à la mixité
- CNRS et CEA : petite majorité de co-publications
- CEA et universités : majorité de co-publications
- CNRS et Inserm : majorité de co-publications
- Inserm et autres France : co-publications par construction
- Inra et universités : co-publications et recouvrement structurel à parts égales.

En observant les co-publications, on constate que l'Inserm et l'Inra ont pour partenaires de collaboration privilégiés les universités, puis des institutions hors coopérative et le CNRS. L'IRD, le Cemagref, l'Inria présentent les mêmes préférences, dans un ordre différent : les Universités, le CNRS puis les institutions hors coopérative. L'Institut Pasteur a comme premiers partenaires les Universités, le CNRS puis l'Inserm. Le CNRS est le premier partenaire du CEA, et a lui-même comme premiers partenaires les Universités, les institutions hors coopérative puis le CEA. Si on observe les partenariats des universités, aucun acteur en dehors du CNRS, avec lequel le recouvrement institutionnel est très large, ne représente plus de 17 % des relations. Pour le CNRS, aucun acteur hors universités ne représente plus de 11 % de ses partenariats. Dans les deux cas le second partenaire est constitué des institutions hors coopérative.

## 2.2 Collaborations européennes et internationales

Le taux de collaboration internationale est défini ici comme la proportion des articles de chaque institution qui ont au moins une adresse étrangère (y compris européenne) en co-signature.

**Tableau II.5 :** Taux de collaborations (%) européennes (UE15) et internationales (en 2000)

	Collaborations internationales	Collaborations européennes
CEA	51,4	30,9
Cemagref	25,4	17,9
CNRS	45,4	22,0
Inra	31,0	16,4
Inserm	34,9	18,4
IRD	61,6	20,6
Universités	36,6	17,5
Cirad	52,5	14,8
Inria	39,1	18,5
Institut Pasteur	50,4	25,3
France	39,2	19,9

*(Données ISI, traitement OST)*

*Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.*

Il est naturellement très élevé pour l'IRD (62 %), élevé (fourchette 50-52 %) pour le Cirad, le CEA et l'Institut Pasteur. On trouve ensuite le CNRS (45 %). De 39 % à 31 % on trouve dans l'ordre l'Inria, les universités, l'Inserm et l'Inra. Le Cemagref collabore à l'international sur un quart de ses articles.

Le taux de collaboration européenne est défini de manière analogue. En comparaison avec le taux de collaboration internationale totale, il s'avère particulièrement élevé pour le Cemagref et le CEA et, conformément aux attentes, faible pour l'IRD et le Cirad.

Qu'en est-il du profil international de collaboration ? Pour fixer les idées, rappelons que les principaux partenaires de la France, toutes disciplines confondues, sont dans l'ordre les Etats-Unis, l'Allemagne, le Royaume-Uni, l'Italie, l'Espagne, la Suisse, la Belgique, le Canada, la Russie et les Pays-Bas (voir tableau II.6).

Les universités et le CNRS ont des profils assez similaires et proches des profils moyens de la France qu'ils contribuent pour une large part à façonner. Leurs cinq premiers partenaires sont identiques à ceux du profil France. Les déviations positives au profil France sont enregistrées pour le CNRS au profit de la Russie, de la Chine, de quelques pays de l'Est européen - et pour les universités au profit du Maroc.

Les cinq premiers partenaires du CEA sont les Etats-Unis, l'Allemagne, le Royaume-Uni, l'Italie et la Suisse (dont le Cern). Par rapport au profil français, le CEA privilégie logiquement les grands pays physiciens que sont l'Allemagne, la Russie, l'Italie, la Suisse et le Japon, et parmi les pays de moindre taille scientifique la Grèce, la Pologne, le Portugal, la Norvège, le Danemark, et certains pays de l'Est européen.

Les premiers partenaires de l'Inra sont les Etats-Unis, le Royaume-Uni, l'Espagne, la Belgique, le Canada puis les Pays-Bas. Par rapport au profil moyen de la France, l'Inra montre une préférence pour le Danemark, l'Australie, la Belgique, la Suède, l'Australie - en partie un effet de spécialisation disciplinaire ; à l'opposé le déficit avec l'Allemagne et la Russie surtout, puis les Etats-Unis et le

Japon, sont marqués. Dans les partenariats plus modestes quantitativement, le lien avec les grands pays d'Amérique latine est remarquable.

L'Inserm a comme premiers partenaires les pays de pointe en recherche médicale et biologique, les Etats-Unis, le Royaume-Uni, l'Allemagne ainsi que l'Italie, Belgique, le Canada et la Suisse. Par rapport au profil moyen de la France, les pays nordiques et le Canada sont privilégiés, on note aussi l'intensité de la relation avec les Etats-Unis.

L'IRD, conformément à ses missions, a une configuration particulière de partenariats. Les premiers partenaires, en dehors des Etats-Unis, sont à parts égales le Royaume-Uni et le Brésil, puis la Belgique, l'Australie et l'Espagne. L'ensemble du profil est très différent du profil français, avec des partenaires forts en Afrique de l'Ouest, puis en Amérique latine (Brésil, Colombie, Mexique...). On note aussi une coopération proportionnellement importante avec l'Australie et l'Indonésie.

Le Cirad, comme l'IRD, diffère du profil moyen français. Les premiers partenaires sont les Etats-Unis, le Royaume-Uni, puis à égalité l'Espagne, la Belgique, les Pays-Bas et le Maroc. Sur des chiffres bruts très faibles, on note pour cette année le Sénégal et la Côte d'Ivoire, le Costa-Rica, le Cameroun ...

L'Inria présente aussi le profil particulier que l'on peut attendre de sa spécialisation disciplinaire. Ses premiers partenaires sont les Etats-Unis, le Royaume-Uni et les Pays-Bas, puis le Canada et l'Allemagne, l'Italie, l'Espagne et la Belgique. Par rapport au profil français, les déviations positives les plus marquées sont enregistrées pour le Portugal, les Pays-Bas, les Etats-Unis et Israël. On note également les liens avec les petits pays européens, la Norvège, ainsi que le Brésil et l'Argentine, sur des chiffres bruts très faibles.

L'Institut Pasteur a comme premiers partenaires les Etats-Unis, le Royaume-Uni, l'Allemagne, la Suisse et la Belgique. Les déviations positives par rapport au profil France sont notées pour Israël, le Brésil, la Chine, le Danemark, puis, à un niveau faible de co-publications brutes, certains pays d'Afrique et d'Amérique du Sud.

Pour le Cemagref, les chiffres absolus de coopération sont très faibles, compte tenu de la taille de l'organisme. Les premiers partenaires sont l'Italie, le Royaume-Uni, les Pays-Bas et l'Allemagne.

**Tableau II.6 :** Co-publications internationales (% des co-publications de chaque institution en 2000) par ordre des partenaires de la France

	CEA	Cemagref	CNRS	Inra	Inserm	IRD	Universités	Cirad	Inria	Institut Pasteur	France
USA	<b>29,9</b>	5,9	24,2	19,2	<b>32,4</b>	13,7	24,2	18,8	<b>35,9</b>	<b>28,0</b>	25,2
DEU	<b>24,4</b>	<b>17,6</b>	14,6	7,8	13,4	<b>3,6</b>	13,3	3,5	7,0	12,4	14,7
RU	<b>17,1</b>	<b>23,5</b>	11,5	<b>17,3</b>	<b>17,1</b>	9,1	12,7	5,9	<b>15,5</b>	<b>15,0</b>	13,9
ITA	<b>16,4</b>	<b>35,3</b>	9,4	7,8	<b>11,2</b>	5,1	9,8	2,4	4,2	6,0	10,3
ESP	5,8	5,9	<b>7,0</b>	<b>8,2</b>	5,5	6,1	6,5	4,7	4,2	3,8	6,7
CHE	<b>10,5</b>	5,9	6,1	4,7	<b>7,4</b>	<b>3,0</b>	6,4	1,2	2,8	<b>7,1</b>	6,6
BEL	<b>6,0</b>	5,9	4,6	<b>8,0</b>	<b>8,4</b>	<b>6,6</b>	<b>6,1</b>	4,7	4,2	<b>6,2</b>	5,9
CAN	4,0	<b>11,8</b>	4,6	<b>7,0</b>	<b>7,5</b>	<b>1,0</b>	5,4	2,4	<b>6,3</b>	4,0	5,5
CEI	<b>9,3</b>	<b>5,9</b>	<b>6,8</b>	<b>1,1</b>	2,5	<b>1,5</b>	4,9		3,5	<b>1,3</b>	5,3
NLD	<b>5,6</b>	<b>23,5</b>	3,5	<b>6,7</b>	<b>5,3</b>	<b>5,6</b>	4,3	4,7	<b>10,6</b>	<b>5,1</b>	5,0
JPN	7,1		<b>4,5</b>	3,0	<b>4,8</b>	<b>0,5</b>	4,0	2,4	1,4	<b>5,7</b>	4,3
SWE	<b>3,2</b>	<b>11,8</b>	2,3	<b>4,3</b>	<b>4,3</b>	<b>0,5</b>	3,1		2,8	<b>3,8</b>	3,2
POL	<b>6,6</b>		<b>3,7</b>	1,8	0,9		<b>3,2</b>			<b>0,7</b>	3,0
BRA	<b>4,4</b>		<b>2,9</b>	2,5	1,6	<b>9,1</b>	<b>2,7</b>	2,4	<b>2,8</b>	<b>4,4</b>	2,6
MAR	0,8		<b>2,7</b>	2,0	1,5	<b>0,5</b>	<b>2,9</b>	<b>4,7</b>	0,7	<b>0,4</b>	2,3
AUS	1,9		1,9	<b>3,3</b>	<b>2,6</b>	<b>6,1</b>	2,0	<b>2,4</b>	1,4	<b>1,3</b>	2,2
DNK	<b>4,0</b>		1,7	<b>3,7</b>	<b>2,5</b>	<b>0,5</b>	1,9		0,7	<b>2,6</b>	2,0
GRC	<b>5,0</b>	<b>17,6</b>	<b>2,0</b>	<b>2,1</b>	0,9	<b>1,5</b>	<b>1,8</b>		1,4	<b>0,7</b>	1,8
AUT	<b>2,5</b>		1,4	1,7	1,7	<b>0,5</b>	<b>1,8</b>	1,2	0,7	<b>0,9</b>	1,7
CHN	<b>3,1</b>		<b>2,1</b>	<b>2,0</b>	0,9	<b>0,5</b>	1,7	1,2	1,4	<b>2,2</b>	1,7
ISR	<b>1,9</b>		1,7	1,6	1,5	<b>1,5</b>	<b>1,8</b>		<b>2,1</b>	<b>3,5</b>	1,7
PRT	<b>3,1</b>		1,7	<b>1,8</b>	<b>1,7</b>	<b>1,0</b>	<b>1,7</b>		<b>4,2</b>	<b>1,8</b>	1,6
CSK	<b>2,2</b>		1,7	<b>1,1</b>	0,8	<b>1,0</b>	1,4		0,7	<b>0,7</b>	1,4
NOR	<b>2,7</b>	<b>5,9</b>	1,3	1,7	0,6		<b>1,4</b>		<b>1,4</b>	<b>0,2</b>	1,4
FIN	1,3	<b>11,8</b>	0,9	1,7	<b>2,5</b>	<b>1,0</b>	1,1		0,7	<b>0,9</b>	1,4
...											
Total*	1 540	17	1 0977	759	1 850	197	11 972	85	142	453	19 000

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 10 articles

Attention : données non additives (compte de présence)

Les chiffres en gras sont supérieurs à la moyenne française

Total\* : Nombre total de coopérations internationales

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

### III. LES INDICATEURS DE VISIBILITE : CITATIONS ET IMPACTS

Les articles scientifiques comportent une bibliographie, les "*articles cités*". Le nombre de références faites à un article ("*citations reçues*" par cet article) est, ainsi, une mesure de sa visibilité - on parle également de son "*impact*".

Les citations sont une mesure de l'utilité scientifique, de la visibilité et de l'influence internationale d'une publication. Elles dépendent néanmoins de nombreux facteurs, parmi lesquels le journal et la langue de publication, les habitudes de citations de la discipline, l'orientation de l'article : fondamentale, appliquée, méthodologique, etc. L'interprétation des citations comme mesure directe de "qualité" est donc incorrecte. Leur nature de phénomène social rend leur interprétation délicate (citations de complaisance, etc.) ; elles sont sujettes à d'importants effets d'amplification en faveur des articles les plus visibles.

Un paramètre essentiel est la variété des comportements de citation selon les domaines, avec pour résultat un nombre moyen de citations (émises et reçues) par article inégal, par exemple en biologie fondamentale et en mathématiques<sup>7</sup>. Il est donc utile de découper la science en domaines et important d'utiliser des indicateurs comparables d'un domaine à l'autre, comme l'impact relatif. Cependant, un domaine défini à un niveau donné (la discipline par exemple, qui est le niveau retenu pour les tables de ce rapport) recouvre une hétérogénéité aux niveaux inférieurs (la spécialité, le journal...). On est, en pratique, conduit à pratiquer diverses normalisations, et donc à considérer plusieurs indicateurs

**Avertissement** : dans le présent rapport, les chiffres de citation sont des citations à très court terme. Seules les citations de 2000 et 2001 ont pu être prises en compte : il faut considérer ces indicateurs avec prudence.

#### 3.1 Indicateurs de visibilité globale des institutions

##### *L'Impact Relatif par discipline*

En bibliométrie, "l'impact" désigne le nombre de citations par article. L'impact direct est le nombre moyen de citations par article pour un acteur donné (ici une des sept institutions), sur un mode de comptage et des fenêtres temporelles définies. L'impact "relatif" est un ratio de l'impact absolu de l'acteur pour un domaine à un impact moyen de référence pour le domaine (p.ex. l'impact de la France, du Monde...)<sup>8</sup>.

Un impact relatif supérieur à 1 indique une visibilité supérieure à la référence. Cette mesure est caractéristique d'un niveau d'observation donné, et permet la comparaison d'un domaine à un autre à ce niveau, par exemple entre disciplines, même s'il ne montre pas l'hétérogénéité des comportements à l'intérieur de la discipline (voir §3.2).

<sup>7</sup> Cela est dû aux pratiques de référencement du champ (listes de références citées), à des phénomènes de dynamique du champ ("âge" des citations et croissance), et à la position par rapport à d'autres champs dont il peut recevoir des flux de citations.

<sup>8</sup> C'est aussi, pour l'acteur, le ratio de sa part de marché des citations à sa part de marché des publications sur cette référence (le "marché" français, le "marché" mondial, etc.)

$$\text{impact acteur} = \text{cit}(\text{act}) / \text{pub}(\text{act})$$

$$\begin{aligned} \text{impact relatif acteur} &= ( \text{cit}(\text{act}) / \text{pub}(\text{act}) ) / ( \text{cit}(\text{ref}) / \text{pub}(\text{ref}) ) \\ &= ( \text{cit}(\text{act}) / \text{cit}(\text{ref}) ) / ( \text{pub}(\text{act}) / \text{pub}(\text{ref}) ) \end{aligned}$$

Pour un acteur, le fait de publier des articles peu visibles, par exemple en raison du choix du journal ou du thème de recherche, aura évidemment des conséquences très différentes sur l'indicateur de "part de marché des citations" et sur l'indicateur d'impact. Le fait d'ajouter un article de faible visibilité à un portefeuille de publications laisse pratiquement inchangée la "part de marché de citations", mais tend à abaisser l'indicateur d'impact.

**Tableau III.1 : Impact relatif par rapport au monde, par discipline (en 2000)**

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths
CEA	1,24	1,45	1,88	1,14	1,07	1,84	1,03	0,38
Cemagref	-	-	0,76	-	-	0,28	1,08	-
CNRS	1,17	1,48	1,48	1,18	1,11	1,25	1,19	1,12
Inra	0,79	0,80	1,13	0,77	0,48	0,50	1,38	0,41
Inserm	1,17	1,43	1,84	1,15	0,82	0,38	1,17	1,67
IRD	0,66	1,26	0,86	0,57	-	0,75	1,27	-
Universités	0,82	0,79	1,03	1,04	1,04	0,86	0,98	1,03
Cirad	0,42	0,22	0,76	0,61	-	0,40	2,38	-
Inria	0,15	0,20	0,12	-	0,41	0,00	1,04	1,20
Institut Pasteur	1,40	1,73	1,68	0,75	-	-	-	-
France	0,95	0,85	1,12	1,08	1,03	1,02	1,01	1,05
UE	0,94	0,93	1,18	1,17	1,15	1,03	1,10	1,07
Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

(Données ISI, traitement OST)

*Grisés : chiffres peu significatifs, correspondant à moins de 50 publications*

*- : chiffres correspondant à moins de 10 publications*

*Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.*

Dans ses domaines de spécialisation, la visibilité du CEA est particulièrement remarquable en sciences de l'univers, bonne en chimie et en physique, moyenne en sciences pour l'ingénieur. Pour les autres domaines, la visibilité du CEA est excellente dans les trois disciplines des sciences de la vie.

La visibilité de l'Inserm est excellente dans ses spécialités, biologie fondamentale et, plus encore, recherche médicale, qui sont pourtant des domaines où la concurrence américaine et anglaise est particulièrement vive. Elle est également très bonne en biologie appliquée, une discipline qui n'est pas au cœur de ses missions.

Le CNRS bénéficie d'une bonne visibilité dans tous les domaines, elle est même excellente sur des domaines où il est relativement peu spécialisé (recherche médicale, biologie appliquée).

La visibilité de l'Inra est élevée dans sa spécialité majeure, la biologie appliquée-écologie, et plus modeste en biologie fondamentale où l'investissement est plus récent. Dans les domaines hors spécialité, la visibilité est remarquable en sciences pour l'ingénieur.

L'Institut Pasteur enregistre des scores remarquables dans ses spécialités, en biologie fondamentale et surtout en recherche médicale.

L'IRD, dont la mission est plus centrée sur le développement, est un peu en retrait de la moyenne mondiale dans sa principale spécialité, la biologie appliquée, mais enregistre de très bons résultats en recherche médicale et sciences pour l'ingénieur (non significatif).

L'Inria fait un peu mieux que la moyenne mondiale en sciences pour l'ingénieur (où est classée l'informatique) et nettement mieux en mathématiques. En dehors de ces spécialités de l'institut, les données ne sont pas significatives.

L'interprétation des chiffres pour les universités pose un problème particulier. Leur visibilité est légèrement supérieure à la moyenne mondiale dans cinq disciplines et inférieure dans les autres. Mais

ces chiffres masquent une grande diversité car les universités sont un ensemble hétérogène, non seulement en production mais aussi en visibilité.

Pour illustrer ce point, le tableau suivant montre les écarts de visibilité entre deux universités A et B : alors que la différence de production n'est que de 1 à 2, la différence d'impact peut être considérable (sauf ici en physique mais sur des effectifs très faibles).

**Tableau III.2 :** Impact relatif par rapport au monde de deux universités à vocations différentes (en 2000)

		Biol.		Biol.			Sci. de	Sci. pour	
		Fondam.	Rech. Méd.	App.- Ecologie	Chimie	Physique	l'Univers	l'Ingénieur	Maths
UNIV_A	part mondiale*	2,0	3,5	2,8	1,9	0,5	0,8	1,8	3,3
	impact relatif	0,42	0,28	0,41	0,67	0,80	0,45	0,42	1,12
UNIV_B	part mondiale*	3,2	4,0	1,6	2,4	3,5	6,5	2,0	6,1
	impact relatif	1,11	0,80	1,08	0,83	0,79	1,11	2,07	1,92

(Données ISI, traitement OST)

(\*) part mondiale de publication, en 10<sup>000</sup>èmes

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

### 3.2 Une vision complémentaire : les indices d'activité des institutions dans les classes de visibilité

La distribution des citations, comme la plupart des distributions bibliométriques, est très asymétrique : une grande masse d'articles est peu ou pas cité alors qu'un petit nombre l'est très fortement.

On peut observer, pour chaque institution, quels pourcentages de ses articles font partie de la classe des articles « *les plus cités* » au niveau mondial, dans chaque discipline<sup>9</sup>. Ici, les articles sont répartis dans les classes suivantes :

- les 5 % d'articles les plus cités au niveau mondial
- puis les 5 % suivants
- les 10 % suivants
- les 20 % suivants
- les 20 % suivants
- les 40 % suivants.

Ce principe de classement a été appliqué de deux manières correspondant à deux niveaux de normalisation : discipline et spécialité. La méthode est détaillée en Annexe.

En pratique, la dernière (et parfois l'avant-dernière) classe(s) sont celle(s) des articles non-cités, très nombreux ici en raison de la fenêtre courte de citation.

Pour chaque classe de visibilité, "l'indice d'activité" est calculé. Cet indice 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.

<sup>9</sup> On peut, par exemple, observer l'indice d'activité de l'institution pour les articles de "bonne visibilité", définis par la classe des 35 % les plus cités. Supposons que l'institution A place 50 % des articles qu'elle signe ou co-signé (compte de présence) dans cet ensemble des 35 % les plus cités à l'échelle mondiale. L'indicateur proposé est " l'indice d'activité pour les articles de bonne visibilité", ici le ratio 50/35=1,43.

### Un indicateur d'"excellence" : la visibilité dans la classe des 5 % les plus cités

**Tableau III.3 :** Indice d'activité dans les 5 % les plus cités, par discipline - Normalisation par spécialité (en 2000)

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths
CEA	1,12	1,99	2,30	1,46	1,38	2,61	1,21	
Cemagref			0,65			0,70		
CNRS	1,12	1,47	1,54	1,24	1,09	1,18	1,33	1,25
Inra	0,74	1,06	1,42	1,36		0,93	0,96	1,20
Inserm	1,26	1,57	1,06	1,11	0,96		0,73	2,30
IRD	0,93	2,84	0,71			0,84	1,23	
Universités	0,88	1,06	1,13	1,05	1,09	0,98	1,24	1,17
Cirad			1,22			1,73		
Inria		6,78					1,42	1,10
Institut Pasteur	1,70	2,11	1,63					
MONDE	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

(Données ISI, traitement OST)

Grisé : discipline correspondant à moins de 50 publications

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Dans ce tableau d'indicateurs normalisés par spécialité, on note l'excellente performance du CEA, de l'Institut Pasteur, de l'Inserm. Pour les institutions de petite taille, les mesures doivent être interprétées avec prudence, la classe des 5 % ne mettant en jeu qu'un petit nombre de publications (une dizaine d'articles pour l'IRD). Une normalisation par discipline, plutôt que par spécialité, avantagerait plutôt les institutions de recherche fondamentale.

#### Le profil de citations par institution

Sur la base de la répartition en classes de citations, on peut également établir un profil de chaque institution. On porte en abscisse les rangs des classes par ordre croissant et en ordonnée l'indice d'activité. Les graphes de profil sont en Annexe.

On peut caractériser ces profils par leur tendance générale (descendante ou ascendante) et secondairement leur allure (régulière, en dôme ou en creux).

- profil descendant : l'acteur est d'autant plus présent que la classe est de forte visibilité. La performance est d'autant meilleure que la pente est plus forte. L'Institut Pasteur, le CEA, l'Inserm sont à classer dans cette catégorie, quelle que soit la normalisation. La pente de leur profil est nettement accentuée en normalisation par discipline, indiquant qu'ils publient dans des spécialités très visibles.

- profil plat : acteur présent dans toutes les tranches, suggérant un impact proche de la moyenne mondiale. Le CNRS, peu présent dans la ou les dernières tranches, est en position plus favorable que les Universités. En valeurs normalisées par spécialité, l'Inra a un profil plutôt plat, en creux, alors qu'en valeurs normalisées par discipline il est légèrement ascendant. Même tendance, avec un creux plus marqué, pour l'IRD.

- profil ascendant : performance d'autant plus faible que la pente est forte. Suggère un impact inférieur à la moyenne mondiale. Ce serait le cas du Cemagref, mais les chiffres sont peu significatifs.

### La catégorie des non cités

La distribution des citations varie selon la fenêtre considérée pour la mesure. Cette fenêtre est ici très courte. Les bornes de chaque classe varient, et c'est plus spécialement le cas pour les articles non cités, dont l'effectif diminue évidemment avec l'allongement de la fenêtre. Les différences entre disciplines sont considérables, en raison de pratiques de citation sensiblement différentes (nombre de citations émises et immédiateté).

**Tableau III.4 :** Indice d'activité dans la tranche des documents non cités (en 2000)

	Biol. Fondam.	Rech. Méd.	Biol. App.- Ecologie	Chimie	Physique	Sci. de l'Univers	Sci. pour l'Ingénieur	Maths
CEA	0,52	0,65	0,64	0,85	0,78	0,63	0,82	1,22
Cemagref	1,43	0,98	1,08	0,97	2,06	1,26	0,94	
CNRS	0,76	0,73	0,73	0,83	0,85	0,78	0,89	0,96
Inra	1,03	1,03	0,88	0,98	1,47	1,05	0,77	1,11
Inserm	0,70	0,67	0,68	0,74	0,96	1,21	0,78	0,67
IRD	0,85	0,87	0,90	0,58		0,90	0,75	1,35
Universités	0,94	1,00	0,89	0,88	0,87	0,89	0,94	0,99
Cirad	1,31	1,40	0,97	0,97		1,08	0,75	1,35
Inria	1,78	1,31	1,42		1,37	2,01	0,97	0,86
Institut Pasteur	0,56	0,61	0,58	0,41			0,69	1,35
MONDE	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

(Données ISI, traitement OST)

Grisé : discipline correspondant à moins de 50 publications

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

### Une décomposition de l'impact relatif : l'impact relatif espéré et le ratio de citation relative

Nous nous plaçons ici dans une logique de normalisation par journal scientifique, assimilé à une micro-spécialité avec ses caractéristiques bibliométriques propres. Dans cette optique, les citations reçues par un acteur pour un article sont considérées comme résultant de deux composantes : d'une part, le niveau de citation attendu (l'impact moyen du journal)<sup>10</sup>, d'autre part le ratio entre les citations reçues effectivement par l'article et le niveau de citation attendu. Ainsi, on va pouvoir caractériser, pour chaque institution, le RCR, qui est, par construction, le ratio de l'impact réel à l'impact espéré. Ainsi, le RCR indique, pour chaque acteur, la sur- ou sous- visibilité des articles de l'acteur par rapport à l'impact des journaux où ils sont publiés<sup>11</sup>. C'est une forme de normalisation de l'impact au niveau du journal.

<sup>10</sup> L'impact de chaque journal est ici calculé de manière différente de l'"*impact factor*" de l'ISI-JCR (types de documents pris en compte, calcul au niveau document par "année citée", gestion des fenêtres temporelles). L'"*impact espéré*" utilisé ici s'inscrit dans la même logique qu'un "*impact factor*" par acteur. Il existe des variantes du RCR avec des normalisations supplémentaires, en particulier par "type de document" (articles, articles de synthèse, lettres, etc.). Pour la présente étude, il nous a paru souhaitable de ne pas distinguer entre ces types en préservant l'avantage éventuel des "articles de synthèse", commandés par les éditeurs et traduisant une position de référence dans un domaine (malgré les biais éditoriaux), par rapport aux articles normaux.

<sup>11</sup> pour un acteur act

$$\text{impact brut IB}(\text{act}) = \text{cit}(\text{act}) / \text{pub}(\text{act})$$

$$\text{citations espérées acteur CE}(\text{act}) = \sum_{\text{journal}} [\text{pub}(\text{act}, \text{journal}) \times \text{impact}(\text{journal})]$$

pour l'ensemble des acteurs, CE( $\sum$  act) est identique à cit( $\sum$  act).

$$\text{impact espéré acteur IBE} = \sum_{\text{journal}} [\text{pub}(\text{act}, \text{journal}) \times \text{impact}(\text{journal})] / \text{pub}(\text{act})$$

Il faut noter que les articles de recherche fondamentale ont relativement plus de chances d'accéder à des revues de fort impact que les articles à caractère plus appliqué. En biologie, le type de modèle étudié (humain, animal, végétal) peut également conditionner dans une large mesure l'accès aux différents niveaux des revues. La capacité de l'Inserm, du CEA et du CNRS à accéder aux journaux de bonne visibilité, souvent à spectre large, est plus forte que celle d'organismes de recherche à vocation appliquée.

**Tableau III.5 :** Ratio de citation relative (en 2000)

RCR	Biol.		Biol.			Sci. de l'Univers	Sci. pour l'Ingénieur	Maths
	Fondam.	Rech. Méd.	App.- Ecologie	Chimie	Physique			
CEA	0,93	1,00	1,05	1,05	1,00	1,26	0,88	0,30
Cemagref	-	-	0,88	-	-	0,56	0,81	
CNRS	0,95	1,01	0,98	1,00	1,02	1,11	0,95	1,01
Inra	1,03	1,06	1,13	1,07	0,52	0,85	0,99	0,30
Inserm	0,97	0,95	0,99	0,92	1,18	0,53	0,64	0,92
IRD	1,15	1,89	0,88	0,70		1,04	0,95	-
Universités	0,90	0,96	0,96	0,96	1,02	1,07	0,90	1,00
Cirad	1,14	0,54	1,00	1,09		0,78	2,03	-
Inria	0,29	0,20	0,17		0,74	0,00	1,17	0,96
Institut Pasteur	1,03	0,88	0,95	0,63	-	-	-	-
France	0,94	0,99	1,01	0,98	1,01	1,09	0,92	0,99
UE	0,98	1,00	1,03	1,02	1,04	1,05	1,05	1,03
Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

(Données ISI, traitement OST)

Grisés : chiffres peu significatifs, correspondant à moins de 50 articles

- : chiffres correspondant à moins de 10 articles

Pour l'IRD, le Cirad et l'Institut Pasteur, ne sont prises en compte, dans le présent exercice, que les adresses correspondant aux implantations localisées en France métropolitaine. Dans le cas de l'IRD, cela représente environ les deux tiers des publications de l'organisme.

Rappelons que le ratio de citation relative dépend aussi du type de documents publiés (article, note, article de synthèse, etc.), avec une sur-visibilité des articles de synthèse. Le RCR est en général moins dispersé que l'impact espéré.

---


$$RCR = IB / IBE = CE / cit(act)$$

Le RCR se calcule indifféremment à partir des impacts bruts ou des impacts relatifs.

$$RCR = IR / IRE$$

L'impact global  $IR = IRE \times RCR$  traduit ainsi la combinaison de deux formes de concurrence :

- la concurrence pour accéder à la publication dans une revue de fort impact ("est-on présent dans les journaux de bonne visibilité ?") - caractérisée par l'"impact espéré".
- la concurrence pour la visibilité au sein de chaque micro-communauté de recherche à laquelle l'auteur s'adresse, micro-communauté supposée reflétée par un journal ("est-on bien ou moins bien visible dans les journaux où on est présent ?") - caractérisée par le RCR.

La "stratégie de publication" consiste notamment à se positionner dans ces deux formes de concurrence.

A titre d'exemple, observons quelques décompositions typiques de l'impact relatif réel, qui permettent de faire des hypothèses sur la manière dont est construite la visibilité des différentes institutions, selon qu'elle est plus liée à leur accès à des journaux de bonne visibilité ou à l'impact des articles au sein de ces journaux :

- le bon impact du CNRS est principalement attribuable à la visibilité des journaux dans lesquels il publie, le ratio de citation relative étant proche de l'unité, ou légèrement plus faible (biologie fondamentale, sciences pour l'ingénieur, multidisciplinaire). Une exception, les sciences de l'univers, où à la fois le RCR et l'accès aux journaux contribuent à la visibilité. Constat analogue pour l'Université, où le RCR tend plus nettement à diminuer l'impact global, sauf en physique, mathématiques, et sciences de l'univers.
- pour l'Institut Pasteur comme pour l'Inserm, c'est la capacité à publier dans des revues de très bon niveau qui détermine les excellentes performances, les RCR sont proches, et en général inférieurs, à la moyenne.
- l'Inra montre des RCR plutôt élevés, typiques d'un institut de recherche finalisée, où les performances à journal constant sont au-dessus de la moyenne, y compris en biologie fondamentale et recherche médicale, mais où l'accès aux revues les plus visibles est difficile. Dans la spécialité principale, la biologie appliquée, le RCR est au même niveau que l'impact global, l'"impact espéré", mesurant l'accès aux revues, est donc égal à l'unité. En revanche, dans la spécialité secondaire des sciences pour l'ingénieur, c'est l'accès à des revues visibles qui tire l'impact global vers le haut.
- la visibilité du CEA, en sciences de l'univers (où elle est extrêmement forte) et en chimie, passe à la fois par une capacité à investir des journaux de bonne visibilité et par une bonne performance relative dans ces journaux. En physique et en sciences de la vie, le RCR est moyen, en sciences pour l'ingénieur il est plus modeste, indiquant que c'est l'accès aux journaux de bon niveau qui garantit la visibilité.
- l'Inria construit sa visibilité avec le RCR en sciences pour l'ingénieur, et plutôt avec l'accès aux journaux de bonne visibilité en mathématiques.
- l'IRD montre un RCR excellent en recherche médicale et au-dessus de la moyenne en sciences de l'univers. Dans les deux cas, c'est l'accès aux revues qui tire l'impact global vers le bas (il reste très bon toutefois en recherche médicale). Pour autant que l'on puisse interpréter des effectifs faibles, le Cirad montre aussi des RCR neutres (biologie appliquée) ou de bon niveau (biologie fondamentale), mais sans accès aux revues les plus visibles. Enfin, les chiffres du Cemagref ne sont guère interprétables en raison des effectifs.

## CONCLUSION METHODOLOGIQUE

Le premier travail sur les données 1997, décidé entre les partenaires pour tester la faisabilité et l'intérêt des indicateurs bibliométriques déclinés par institution, avait été conclusif. Il montrait qu'une collaboration active entre les représentants des institutions partenaires était efficace, et devrait permettre une appropriation par chacune d'entre elles.

Les résultats présentés ici enrichissent la palette d'indicateurs - avec la limitation, pour les impacts, d'une fenêtre de temps très courte.

Compte tenu des différences dans la couverture et les méthodes de mesure, les résultats confirment les premiers travaux de 1997.

Même si les indicateurs de production ne peuvent être interprétés qu'en tenant compte de la dimension des institutions, le présent rapport ne comporte pas d'analyse de "productivité" en tant que telle, du fait de l'absence de données sur les ressources qui soient comparables entre elles et qui soient homogènes aux données de production étudiées ici - notamment en ce qui concerne les découpages disciplinaires.

Notons que les indicateurs relatifs de visibilité ici proposés sont, en première analyse, indépendants de la taille.

Enfin, une des conclusions importantes à tirer de ces résultats est que le niveau de référence des comparaisons (celui où s'effectue la normalisation) influence considérablement la mesure de performance. Ainsi, selon que l'on choisit de comparer des acteurs à l'intérieur de la discipline (où les spécialités les plus visibles seront favorisées), de la spécialité (où les journaux les plus visibles seront favorisés) ou du journal, les résultats sont très différents. En règle générale, plus le niveau de normalisation est global, plus les instituts de recherche fondamentaux sont favorisés. Au contraire, des ensembles de références étroits, comme la spécialité ou *a fortiori* le journal, favorisent les performances des instituts finalisés. Une normalisation à un niveau raisonnable est souhaitable (spécialité), sans toutefois négliger le fait que la visibilité de certaines spécialités est partiellement due à leur caractère central dans la distribution des flux de connaissance.

## ANNEXE

### Graphiques des indices d'activité des institutions dans les classes de citations

La distribution des citations, comme la plupart des distributions bibliométriques, est très asymétrique c'est-à-dire qu'une grande masse d'articles est peu ou pas citée alors qu'un petit nombre l'est très fortement. Nous avons ordonné tous les articles mondiaux par classe de visibilité :

- les 5 % d'articles les plus cités au niveau mondial
- puis les 5 % suivants
- les 10 % suivants
- les 20 % suivants
- les 20 % suivants
- les 40 % suivants.

Ce principe de classement a été appliqué de deux manières correspondant à deux niveaux de normalisation.

Dans le premier cas, le classement de base s'est opéré discipline par discipline. Au niveau global, toutes disciplines, la classe des 5 % les plus cités regroupe la classe des 5 % les plus cités en physique, la classe des 5 % les plus cités en biologie fondamentale, celle des 5 % les plus cités en recherche médicale, et ainsi de suite sur toutes les grandes disciplines. Nous sommes donc dans une logique de normalisation disciplinaire. Les courbes de distribution correspondantes sont en orange dans les graphes qui suivent.

Une seconde représentation repose sur un classement de base spécialité par spécialité (*subject category* ISI). Au niveau global, toutes disciplines, la classe des 5 % les plus cités regroupe la classe des 5 % les plus cités en « Anatomie morphologie », celle des 5 % les plus cités en « Biochimie, biologie moléculaire », puis en « Bio-ingénierie » et ainsi de suite sur toutes les spécialités. Cette seconde option renvoie à une logique de normalisation par spécialité. Les courbes de distribution correspondantes sont en vert dans les graphes qui suivent. Une autre normalisation, non présentée ici, est celle par journal, dans une logique analogue à celle du RCR.

La normalisation par spécialité neutralise l'effet de l'inégalité de visibilité entre spécialités au sein d'une même discipline. Il n'est pas rare que la normalisation par spécialité (ou à fortiori par journal) nivelle les hiérarchies et soit notamment plus favorable que la vision par discipline aux institutions de recherche finalisée.

En pratique, la dernière (et parfois l'avant-dernière) classe(s) est (sont) celle(s) des articles non cités, très nombreux en raison de la fenêtre courte de citation.

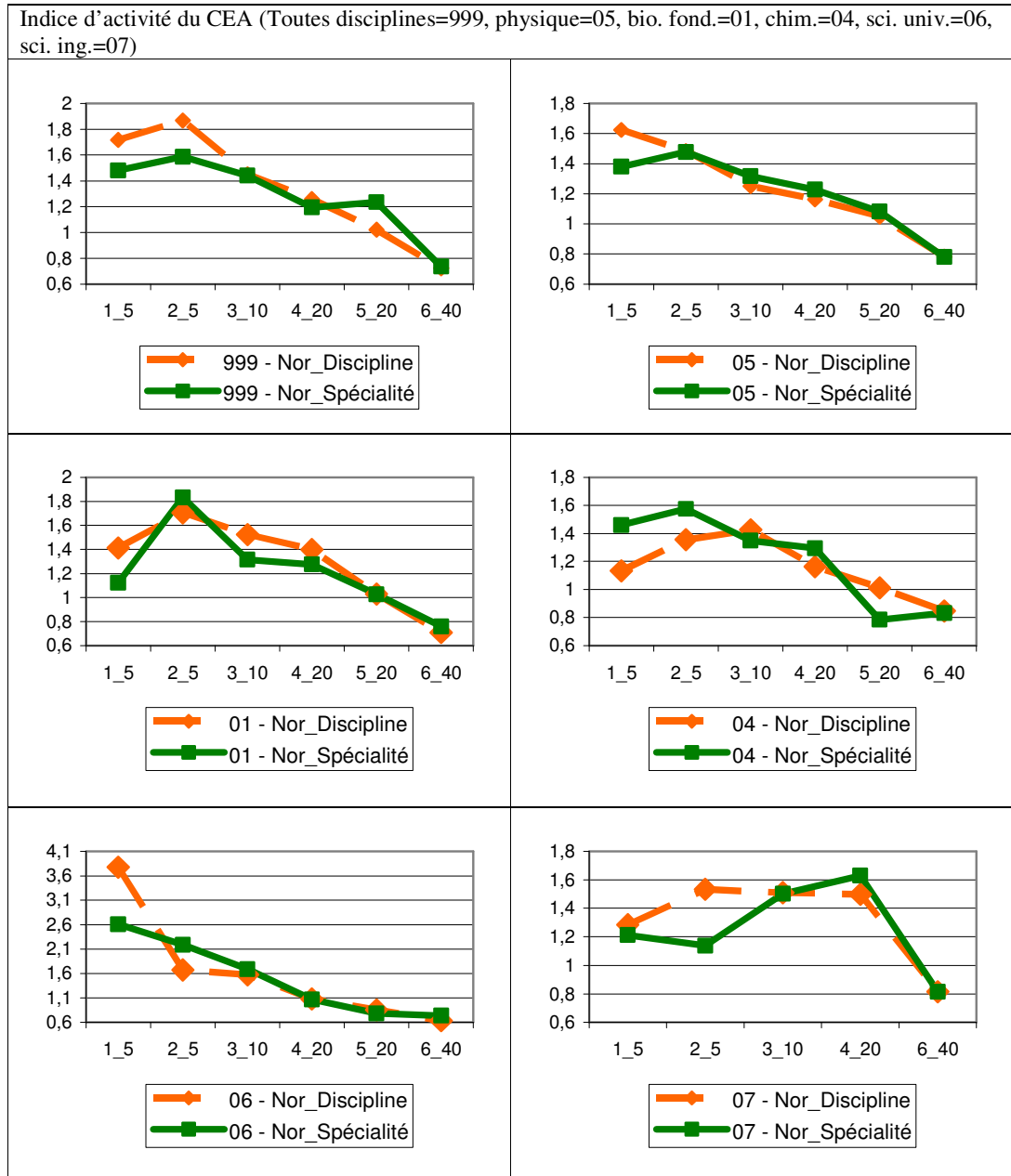
Pour chaque classe de visibilité, l'"indice d'activité" est calculé. Rappelons qu'il est égal au ratio entre le pourcentage de ses articles que l'acteur place dans cette classe, et le pourcentage des articles mondiaux dans cette classe. Par exemple, pour la classe d'"excellence" des 5 % les plus cités, ce dernier pourcentage est par construction proche de 5 %<sup>12</sup>.

---

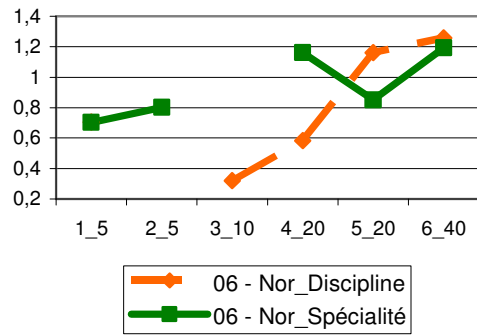
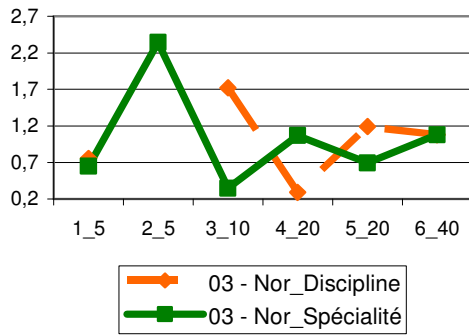
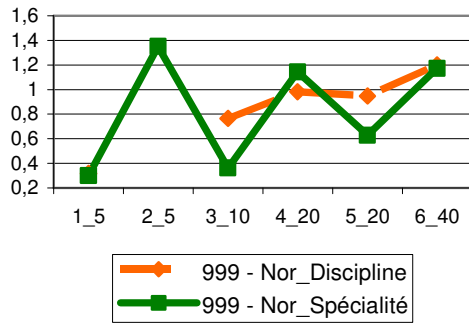
<sup>12</sup> L'effectif réel peut être différent de la valeur nominale, en raison des ex-aequo. L'indice d'activité, de manière identique, est le ratio de la part de l'institution dans les articles mondiaux de la classe, à la part de l'institution dans tous les articles mondiaux.

--- Orange : indicateur normalisé par discipline

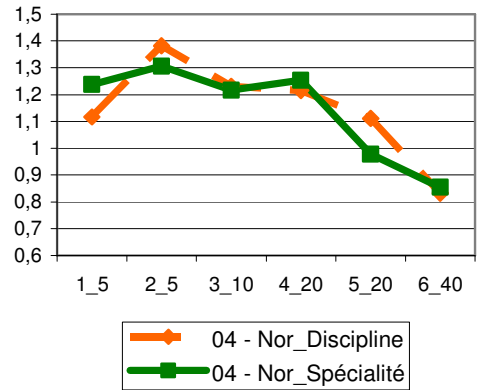
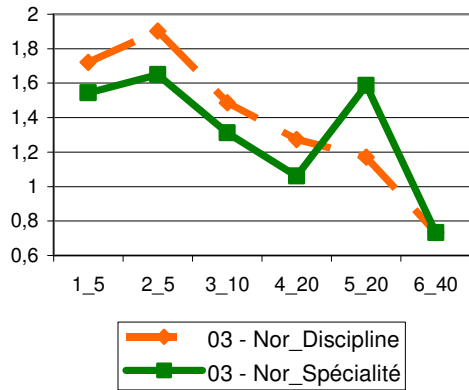
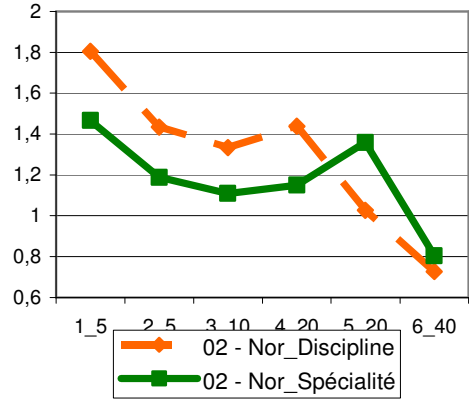
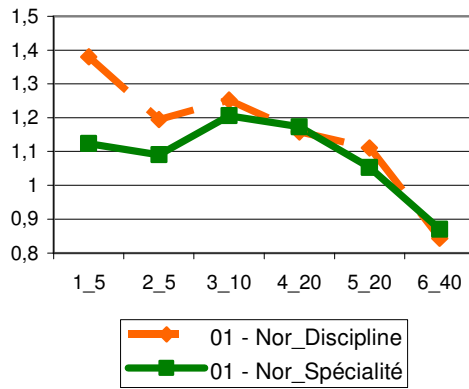
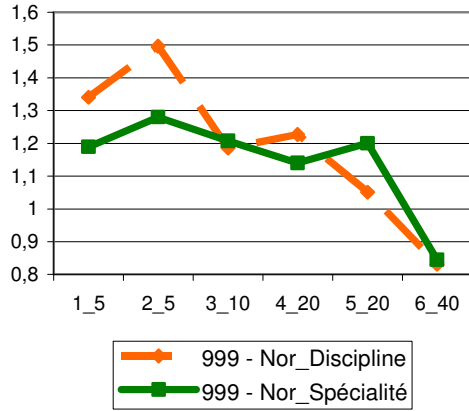
— Vert : indicateur normalisé par spécialité ISI



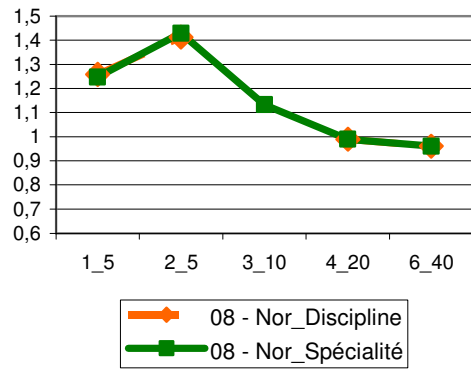
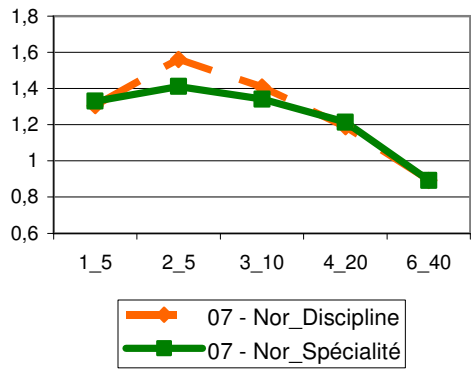
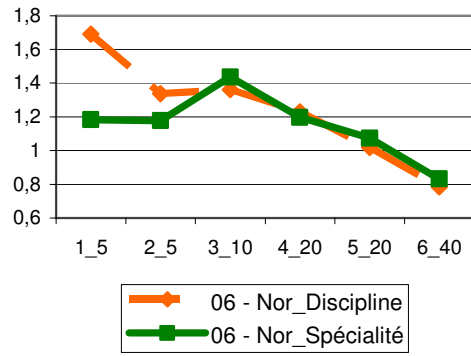
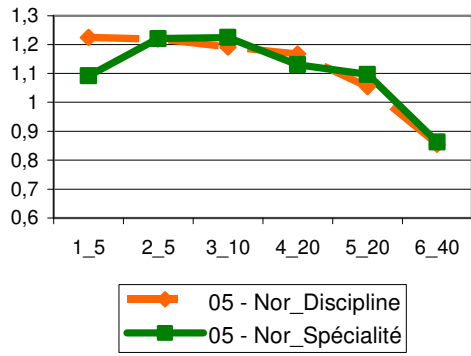
Indice d'activité du Cemagref (Toutes disciplines=999, bio. app.=03, sci. univ.=06)



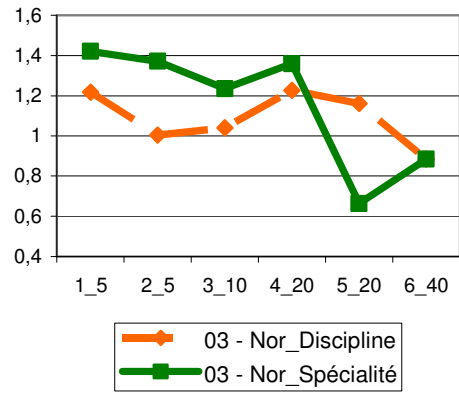
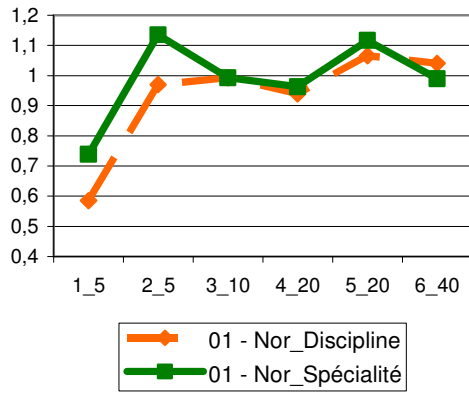
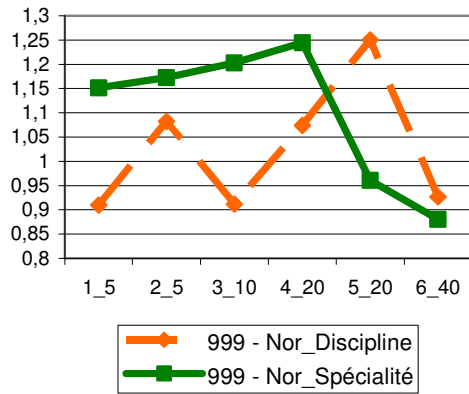
Indice d'activité du CNRS (Toutes disciplines=999, bio. fond.=01, rech. med.=02, bio. app.=03, chim.=04)



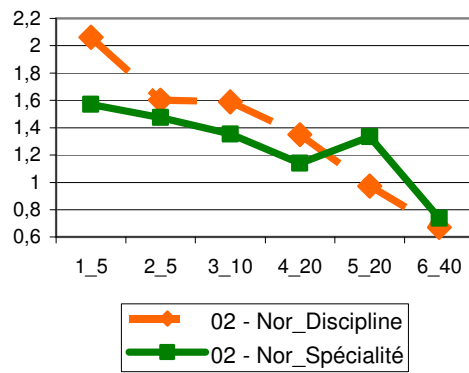
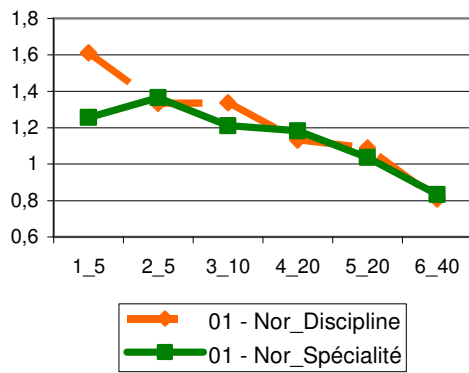
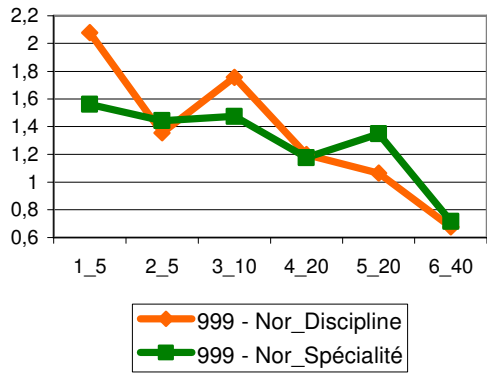
Indice d'activité du CNRS (physique=05, sci. univ.=06, sci. ing.=07, maths=08)



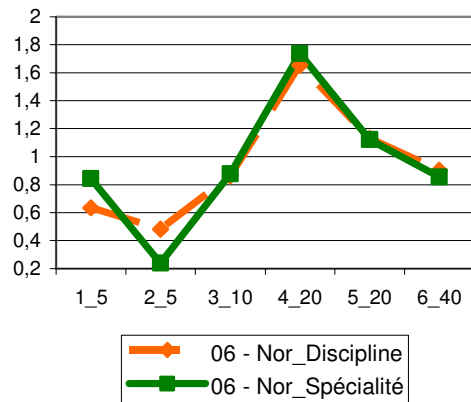
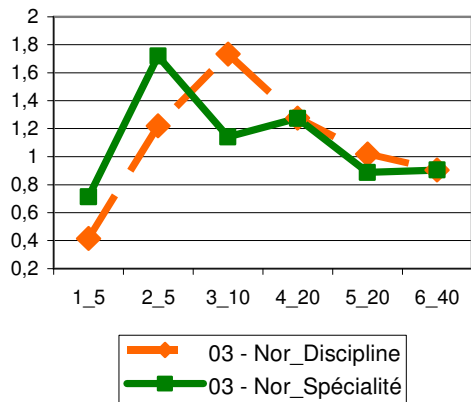
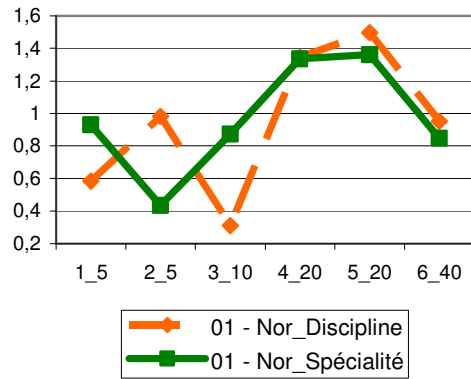
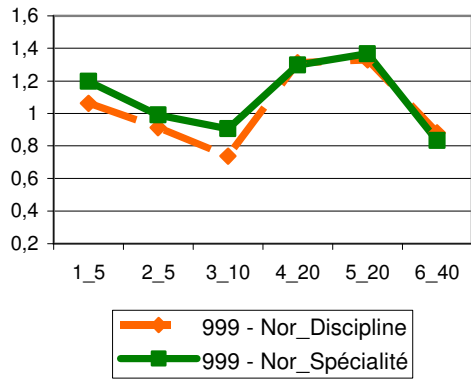
Indice d'activité de l'Inra (Toutes disciplines=999, bio. fond.=01, bio. app.=03)



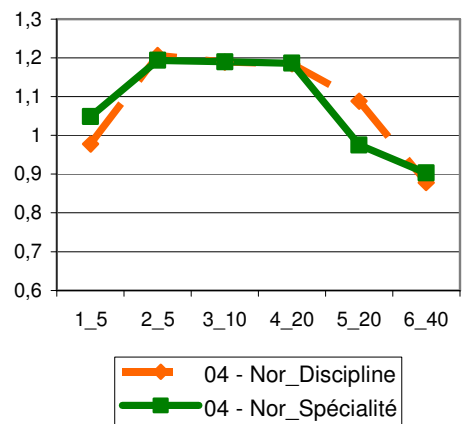
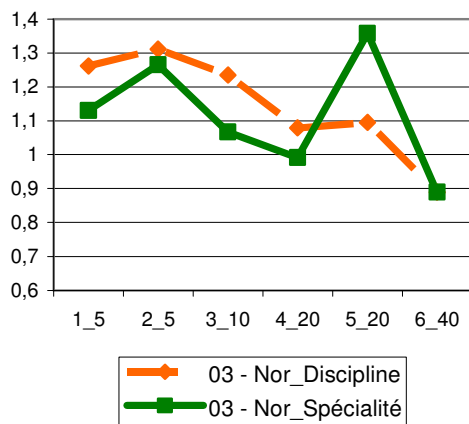
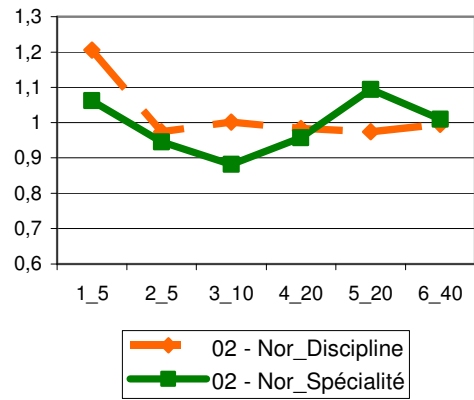
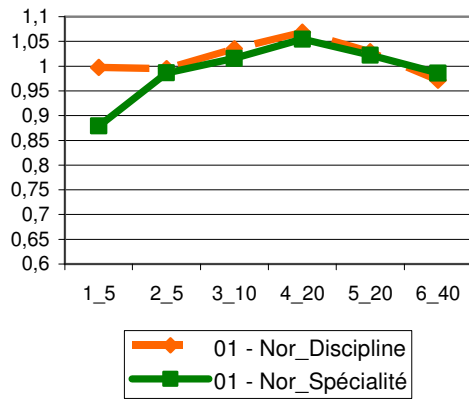
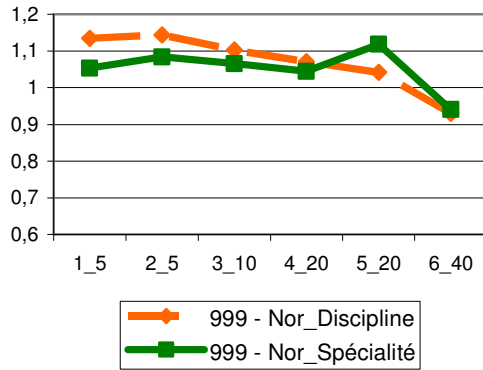
Indice d'activité de l'Inserm (Toutes disciplines=999, bio. fond.=01, rech. med.=02)



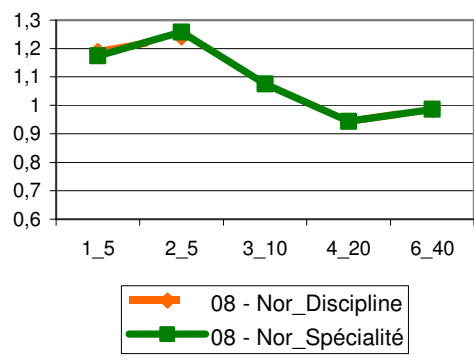
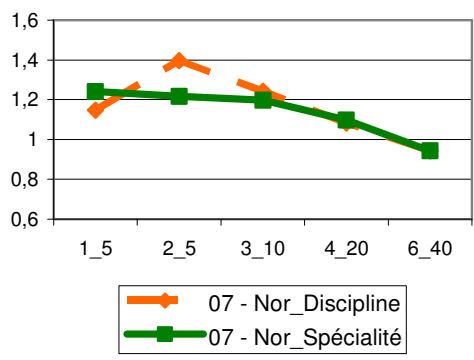
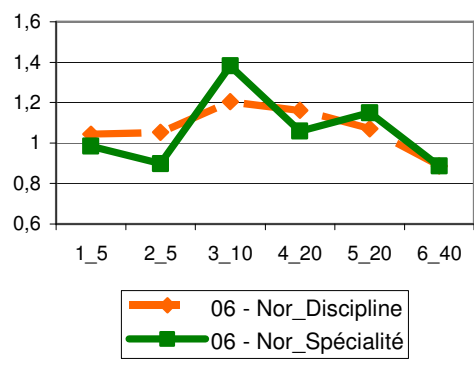
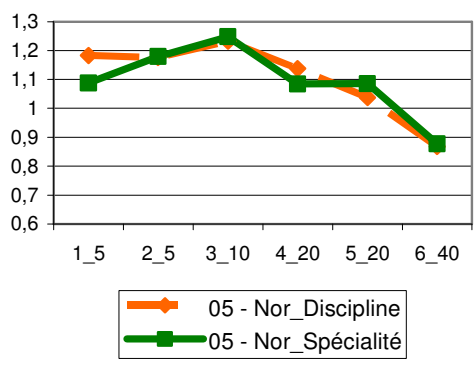
Indice d'activité de l'IRD (Toutes disciplines=999, bio. fond.=01, bio. app.=03, sci. univ.=06)



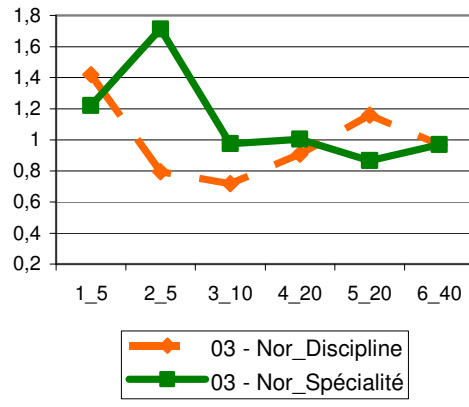
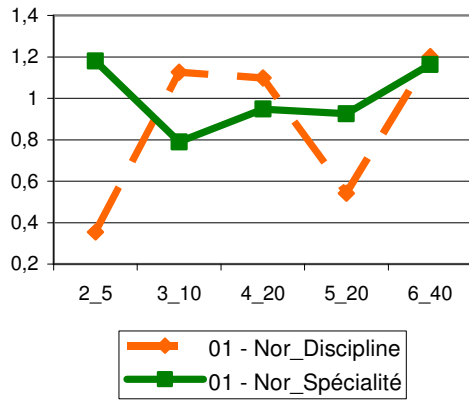
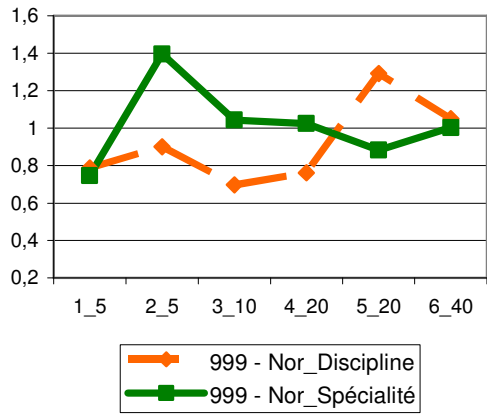
Indice d'activité des universités (Toutes disciplines=999, bio. fond.=01, rech. med.=02, bio. app.=03, chim.=04)



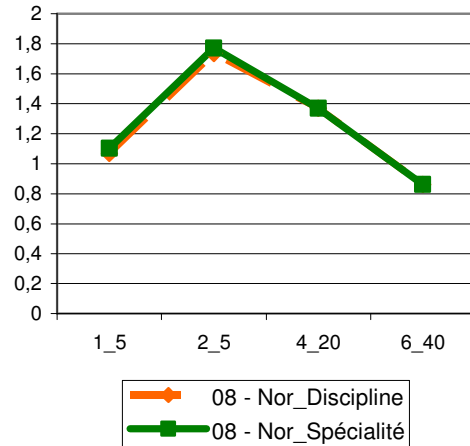
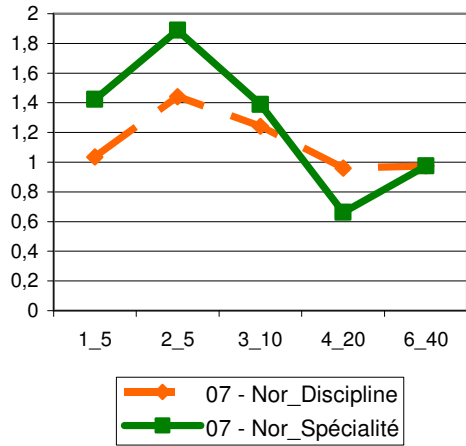
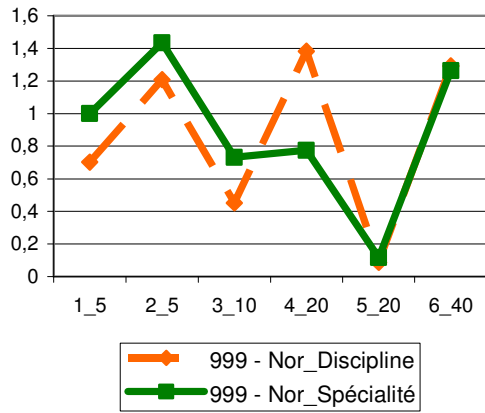
Indice d'activité des universités (physique=05, sci. univ.=06, sci. ing.=07, maths)



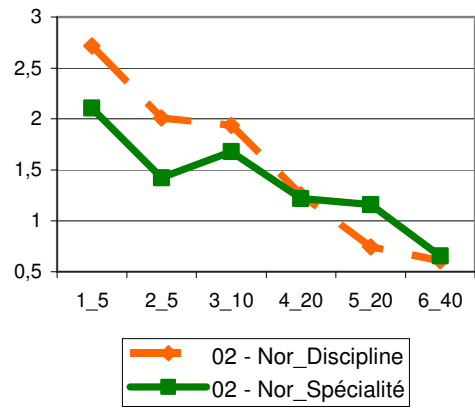
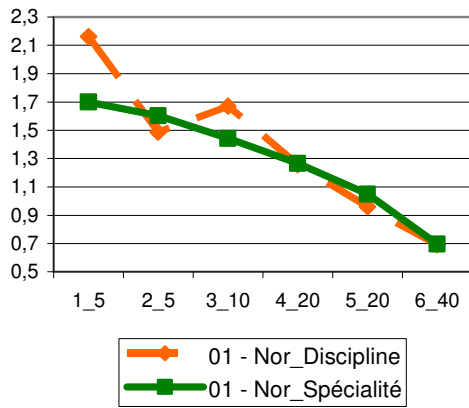
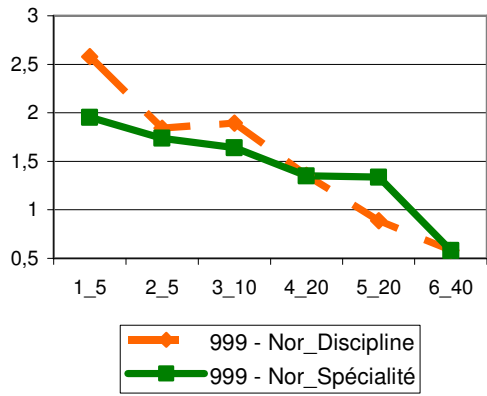
Indice d'activité du Cirad (Toutes disciplines=999, bio. fond.=01, bio. app.=03)



Indice d'activité de l'Inria (Toutes disciplines=999, sci. ing.=07, maths=08)



Indice d'activité de l'Institut Pasteur (Toutes disciplines bio. fond., rech. med.)



## ANNEXE

## Les disciplines OST comme agrégats des "spécialités scientifiques" du SCI

Tableau D.1a

<b>1. BIOLOGIE FONDAMENTALE</b>		<b>3. BIOLOGIE APPLIQUEE-ECOLOGIE</b>
Anatomie, morphologie	Neuro-Imagerie	Agriculture
Biochimie, biologie moléculaire	Microbiologie	Agriculture, multidisciplinaire
Bioingénierie	Microscopie	Agronomie gén. (agriculture, sc. sols)
Biologie cellulaire, histologie	Neurosciences	Biodiversité, biologie de la conservation
Biologie moléculaire et cellulaire	Nutrition, diététique	Biologie générale
Biomatériaux	Parasitologie	Biologie, autres
Biométhodes	Physiologie	Bois et textiles
Biophysique	Psychologie	Botanique, biologie végétale
Biotechnologie et microbiologie app.	Sciences comportementales	Ecologie
Embryologie	Systèmes reproducteurs	Entomologie
Génétique, hérédité	Techniques du laboratoire	Horticulture
Génie biomédical	Virologie	Mycologie
		Ornithologie
<b>2. RECHERCHE MEDICALE</b>		Sciences des productions animales
Allergologie	Médecine interne générale	Sciences et techniques agro-alim.
Andrologie	Méd. intégrative et de compl.	Sciences et techniques des pêches
Anesthésiologie	Médecine légale	Stations agricoles expérimentales
Cancérologie	Médecine tropicale	Sylviculture
Chimie, clinique et médecine	Médecine vétérinaire	Zoologie générale
Chirurgie	Neurologie clinique	
Dermatologie, vénérologie	Odontologie	<b>4. CHIMIE</b>
Endocrinologie	Ophthalmologie	Chimie analytique
Ethique médicale	Orthopédie	Chimie appliquée
Gastro-entérologie	Oto-rhino-laryngologie	Chimie générale
Gérontologie	Pathologie	Chimie minérale et nucléaire
Gynécologie, obstétrique	Pédiatrie	Chimie organique
Hématologie	Pharmacologie - pharmacie	Chimie physique
Immunologie	Pneumologie	Cristallographie
Maladies infectieuses	Psychiatrie	Electrochimie
Médecine cardiovasculaire	Radiologie, méd.nucléaire	Matériaux composites
Médecine cardiovasculaire 2	Réhabilitation	Matériaux/analyse
Médecine clinique, autres	Rhumatologie	Science des matériaux
Médecine d'urgence	Santé publique	Science des matériaux - bois, papier
Soins Intensifs	Santé publique 2	Science des matériaux - céramiques
Médecine de la dépendance	Toxicologie	Science des polymères
Médecine du sport	Transplantations	Traitements de surface
Médecine expérimentale	Urologie - néphrologie	

**Tableau D.1b**

<b>5. PHYSIQUE</b>		<b>7. SCIENCES POUR L'INGENIEUR</b>
Acoustique		Biocybernétique
Instrumentation		Composants
Optique		Rev. de synth. en inform. (CRC)
Physico-chimie		Contrôle
Physique appliquée		Contrôle 2
Physique des fluides et plasmas		Energie et carburants
Physique des particules		Génie maritime
Physique du solide		Génie aérospatial
Physique générale		Génie chimique
Physique mathématique		Génie chim. et thermodynamique
Physique nucléaire		Génie civil
Physique, autres		Génie de la construction
Spectroscopie		Génie électrique et électronique
		Génie industriel
<b>6. SCIENCES DE L'UNIVERS</b>		Génie mécanique
Astronomie et astrophysique		Génie métallurgique et minier
Biologie marine - hydrobiologie		Génie minier
Géographie		Génie pétrolier
Géologie		Informatique
Géosciences		Informatique/divers
Géotechnique		
Limnologie		<b>8. MATHEMATIQUES</b>
Météorologie		Mathématiques
Minéralogie		Mathématiques appliquées
Océanographie		Mathématiques générales
Paléontologie		Mathématiques théoriques
Ressources en eau		Mathématiques, autres
Sciences de l'environnement		Méthodes mathématiques (biologie et médecine)
Technologies de l'environnement		Méthodes mathématiques (sciences physiques)
Divers, Géophysique-Géochimie		Méthodes mathématiques (sciences sociales)
		Statistique et probabilités