

REPÉRAGE DES PUBLICATIONS PAR LES ÉTABLISSEMENTS

Laure Martineau, Suzy Ramanana-Rahary, Françoise Rojouan

REPÉRAGE AUTOMATIQUE/ REPÉRAGE ÉTABLISSEMENT

- Projet U – Multirank
- 6 établissements ont participé au test
 - Trois universités de recherche nationales, 1 université de rayonnement régional, une école d'ingénieur, un des centres de recherche d'un organisme national
- sur 6 années de publications : période 2004 à 2009
- Analyse comparative quantitative (2004-2008)
- Analyse comparative qualitative (2009)

REPÉRAGE AUTOMATIQUE/ REPÉRAGE ÉTABLISSEMENT

- Repérage automatique (TD)
 - Sur les adresses des auteurs dans la base de données
 - Liste d'occurrences simples combinant le(s) nom(s) officiel(s) des établissements (univ + nom – univ + ville.....)
- Repérage par les établissements (BU)
 - personnes connaissant bien leur établissement et habituées depuis plusieurs années à l'exercice
 - Validation du repérage automatique

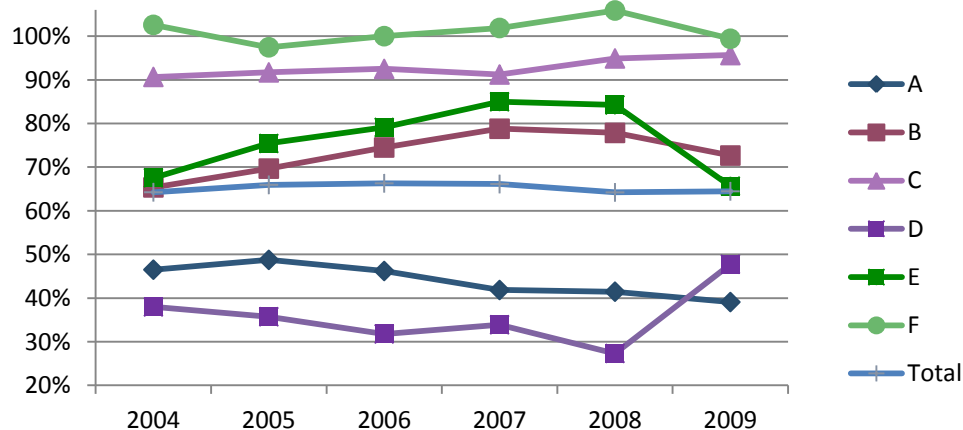
REPÉRAGE AUTOMATIQUE/ REPÉRAGE ÉTABLISSEMENT

Volume des publications attribuées aux établissements

Etablissements	Publications 2004-2009		
	repérage auto (TD)	repérage étab. (BU)	TD/BU
A	3622	8309	44%
B	6159	8384	73%
C	5439	5858	93%
D	1178	3319	35%
E	984	1298	76%
F	920	909	101%
Total	18302	28077	65%

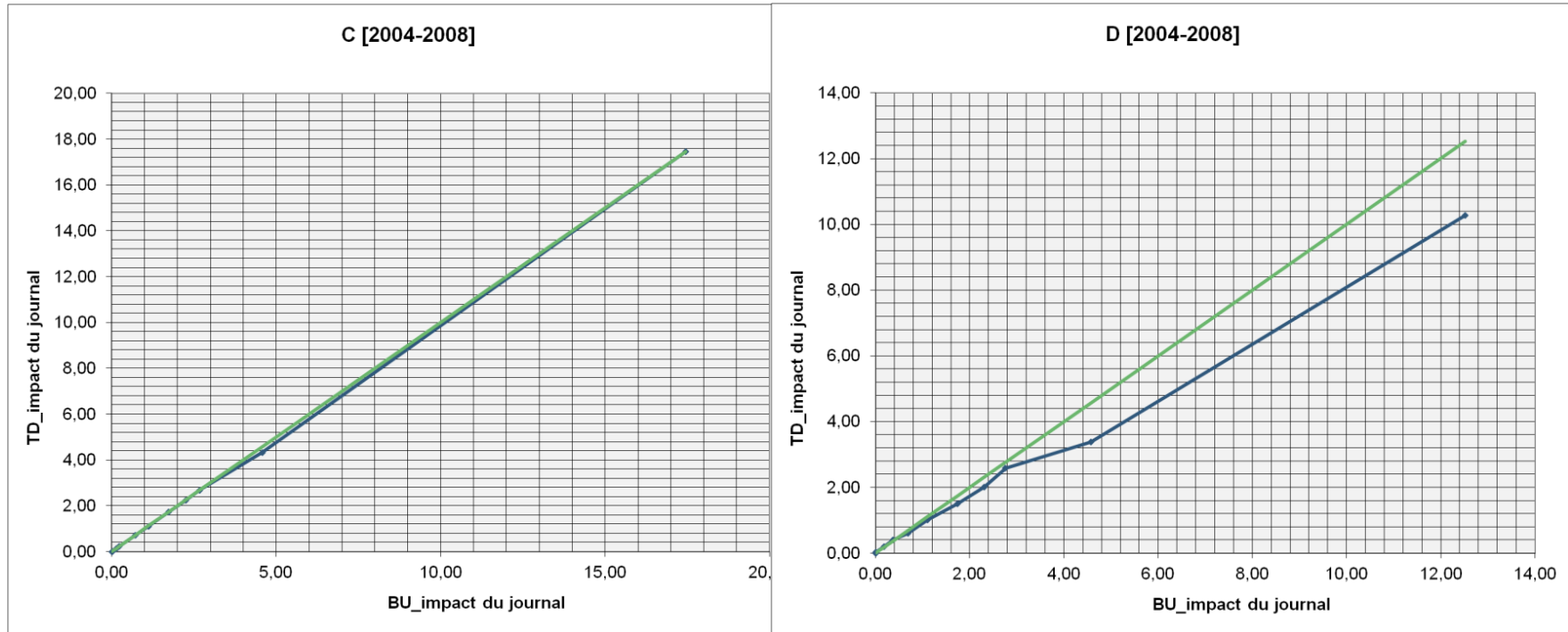
- Efficacité du repérage automatique probablement liée aux caractéristiques de l'établissement mais pas à sa taille
- Stabilité de la part du ratio TD/BU dans le temps

Evolution de la part du repérage TD/BU



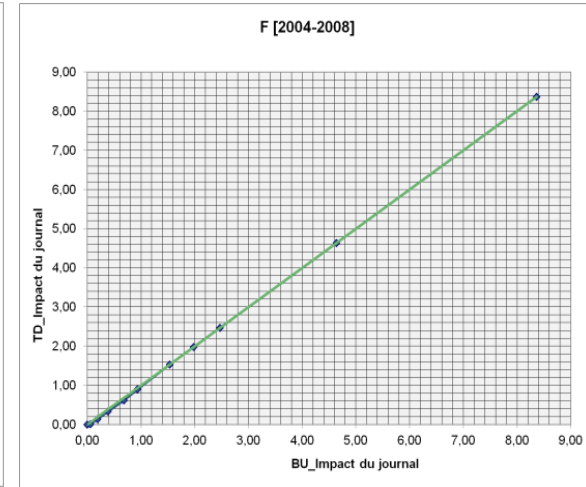
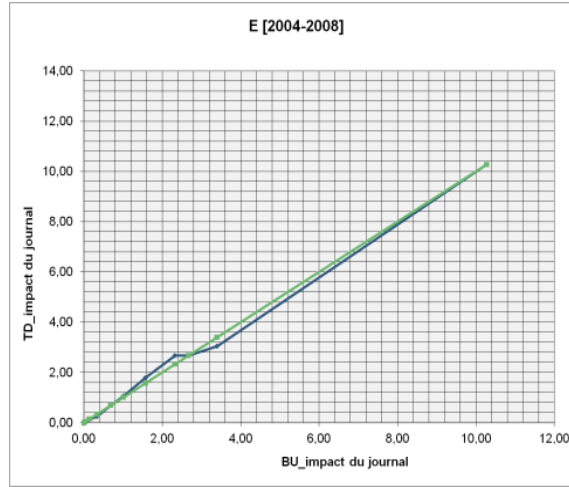
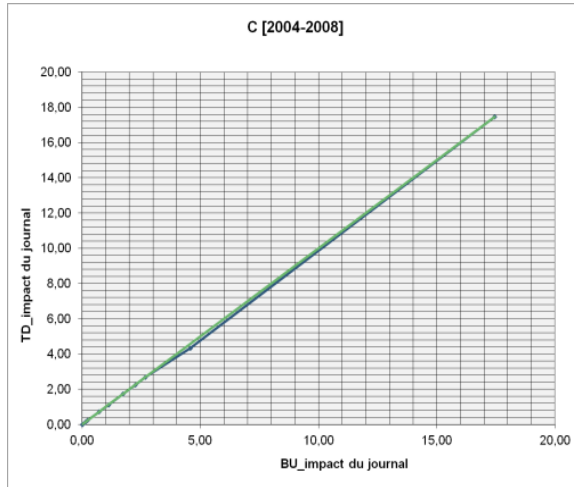
Visibilité en termes de citations :

Comparaison de la distribution de l'impact des journaux des publications de chaque établissement pour les ensembles TD (bleu) et BU (vert) (Q-Q plot test)

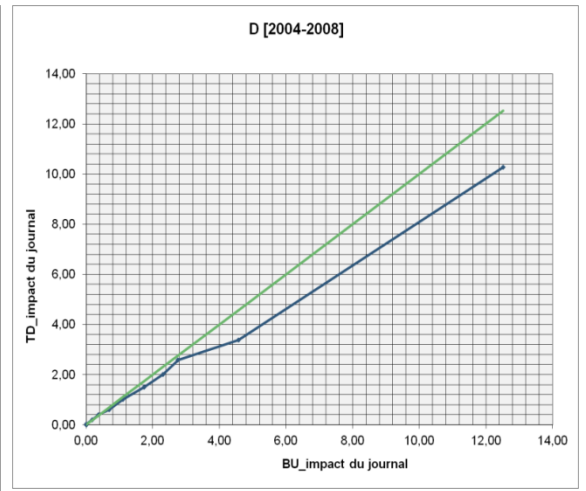
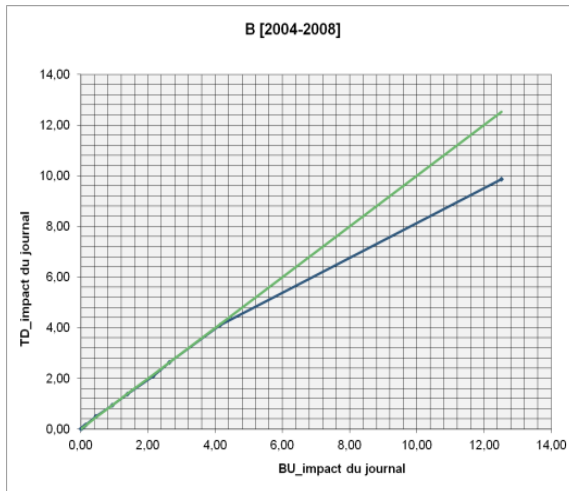
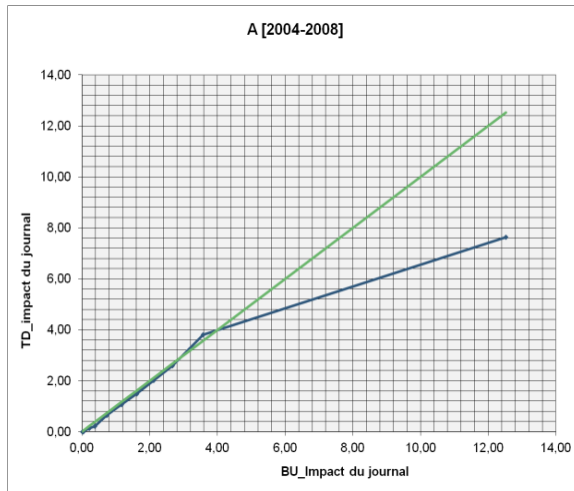


- Plus la différence est grande entre les deux courbes, plus grande est la différence dans les distributions de la visibilité des journaux entre le repérage automatique et le repérage par l'établissement.

Visibilité en termes de citations (suite) :



- une part élevée de publications du TD/BU n'est pas suffisante pour être représentative – (ex EtabE : 73% Etab. B : 76%)



Visibilité en termes de citations (suite) :

Distribution des publications selon leur visibilité (nombre de citations)

Part des publications de l'établissement dans la classe des publications peu ou pas citées

Part des citations de l'établissement dans le top 10 % des publications les plus citées

Établissements		repérage auto (TD)	repérage étab. (BU)	TD/BU
A	part de pub. peu ou pas citées	40,6	36,7	1,11
	part de pub. dans top 10	60,2	57,8	1,04
B	part de pub. peu ou pas citées	34,2	34,7	0,99
	part de pub. dans top 10	57,7	59,2	0,98
C	part de pub. peu ou pas citées	34,5	33,9	1,02
	part de pub. dans top 10	46,4	47,5	0,98
D	part de pub. peu ou pas citées	44,4	48,0	0,93
	part de pub. dans top 10	59,7	56,8	1,05
E	part de pub. peu ou pas citées	39,4	40,0	0,99
	part de pub. dans top 10	57,1	55,6	1,03
F	part de pub. peu ou pas citées	62,5	62,0	1,01
	part de pub. dans top 10	75,0	75,2	1,00

- variété des situations : parts quasi identiques pour les établissements B, C et F, mais très différentes pour les établissements A, D et E

Aspect qualitatif : analyse des adresses manquantes dans le TD

Pour 5 des établissements et sur l'année 2009

- 3 types d'adresses ne sont pas repérées automatiquement :
 - type I : adresses de mauvaise qualité (nom de l'établissement manquant, mal orthographié)
 - type II : adresses avec uniquement le nom d'un service , d'un département, d'un institut ou institution rattaché à l'établissement dont les CHU et hôpitaux (type II b)
 - type III : adresse comportant soit uniquement le nom d'une structure mixte (type UMR) soit le nom d'une structure regroupant l'institution dans une organisation en réseau (PRES = type IIIB)

Etablissements	Publications 2009 - (% des adresses repérées dans BU et manquantes dans TD)				
	Type I	Type II	Type II b	Type III	Type III b
A	2%	32%		58%	7%
B	1%	28%	26%	35%	10%
D	1%	73%		26%	
E	7%	42%	16%	35%	
F	8%	67%		25%	

REPERAGE AUTOMATIQUE

- le moins coûteux, mais pas systématiquement représentatif de la production scientifique d'un établissement et de sa visibilité

REPERAGE PAR LES ETABLISSEMENTS

- le plus coûteux (en temps et en maintenance d'un outil WEB)

Mais permet :

- une représentativité satisfaisante de la production scientifique de l'établissement
- un apprentissage sur la structure des adresses utilisée par les chercheurs de l'établissement
- un possible enrichissement des adresses

Enrichissement des adresses, souplesse des périmètres

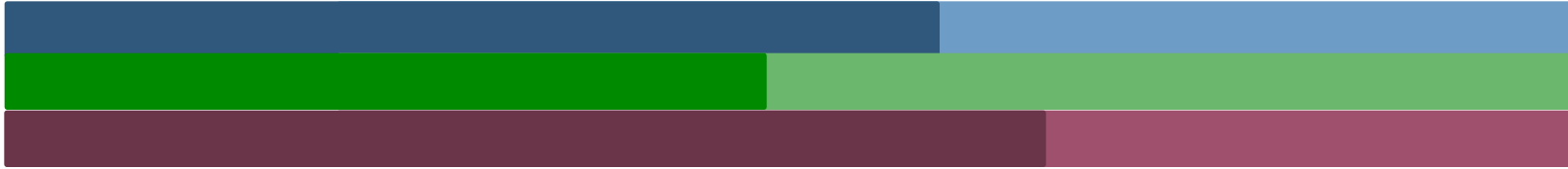
- L'OST a la capacité de combiner n'importe quelle dimension institutionnelle et géographique de la base
- Changements rapides du cadre contextuel national (organisation en réseau, fusion/association, programmes d'excellence....) mais aussi européen et mondial dans lequel évoluent les établissements nécessitent de pouvoir analyser la production scientifique et technologique des établissements dans des périmètres très flexibles: site, PRES, IDEX.....
- Etude de l'intérêt de cet enrichissement pour l'établissement lui-même avec l'université de Bourgogne et avec Grenoble INP qui le pratique depuis le début de leur repérage IPERU

Enrichissement lors du repérage dans l'interface

adrCleost	adresse	Statut	D-SHS	Reperage	Champ_libre 1	Champ_libre 2
I00040863	Univ Dijon, Lab SPMS, F-21065 Dijon, France	O	SHS	O	UB	EA4180
I00041141	INRA, UMR FLAVIC 1129, F-21000 Dijon, France	O	D	O	UB-INRA	UMR1129
I00041142	ENESAD, UMR FLAVIC 1129, F-21000 Dijon, France	O	D	O	UB-INRA	UMR1129
I00041184	Univ Bourgogne, Inst Carnot Bourgogne, Dijon, France	O	D	O	UB-CNRS	UMR5209
I00041321	Univ Bourgogne, CNRS, Ctr Rech Climatol, UMR5210, Fac Sci, F-21000 Dijon, France	O	D	O	UB-CNRS	UMR5210
I00042224	INRA, UMR Microbiol & Geochim Sols, F-21065 Dijon, France	O	D	O	UB-INRA	UMRE1229
I00043081	Univ Dijon, UMR 5594, ARTeHIS, F-21004 Dijon, France	O	SHS	O	UB-CNRS	UMR5594

Complément d'enrichissement à partir d'informations données par l'établissement

Champ_libre 1	Champ_libre 2	contrat concerné	LABEL REPERE (année de publication)	Nom ou sigle (année de publication)	DS	Rattachement FR/IFR (contrat actuel)	Rattachement ED (contrat actuel)	autres tutelles
UB	EA4180	2007 à 2011	EA4180	SPMS	6	MSH (UMS2739)	LISIT (ED491)	
UB-INRA	UMR1129	2007 à 2011	UMR1129	FLAVIC	10	Buffon (FR4101)	E2S (ED490)	INRA
UB-INRA	UMR1129	2007 à 2011	UMR1129	FLAVIC	10	Buffon (FR4101)	E2S (ED490)	INRA
UB-CNRS	UMR5209	2007 à 2011	UMR5209	ICB	2	3M (FR2604) (FR2604)	Carnot (ED149)	CNRS
UB-CNRS	UMR5210	2007 à 2011	UMR5210	CRC	6	Buffon (FR4101)	E2S (ED490)	CNRS
UB-INRA	UMRE1229	2007 à 2011	UMRE1229	MSE	10	Buffon (FR4101)	E2S (ED490)	INRA
UB-CNRS	UMR5594	2007 à 2011	UMR5594	ARTEHIS	6	MSH (UMS2739)	LISIT (ED491)	CNRS



OST : www.obs-ost.fr