Performance and Over-Ranking Measures on Citation Profiles

Michel Zitt[1,2] and Agénor Lahatte[2]

[1]zitt@nantes.inra.fr
Lereco, INRA, Nantes (France) & Observatoire des Sciences et des Techniques (OST), Paris (France)

[2]lahatte@obs-ost.fr
Observatoire des Sciences et des Techniques (OST), Paris (France)

Abstract
We present two families of indicators based on measures of actors' activity in classes of citation scores, namely quantile classes or ranks of citations defined by a global reference - by default, the world. Those "activity index in citation classes" (direct: AIC; cumulative: AICC) are versatile indicators which can emulate relative citation measures in a variety of ways (structural vs. normalized indexes, levels of normalization). Analytical scrutiny of the profile visualization at any level (nations, institutions) confirms the necessity of discarding the low-cited tail for evaluative indicators. Global profile characterization may use parametric or non-parametric techniques. Our focus here is on measures belonging to the latter family. One of these is an average position in highly cited classes, another is a score based on over-rankings comparisons between actors in the upper part of the profile. Examples are given at the national and institutional level, primarily on "performance indicators" based on AICC, akin to relative impact, and also including an extension to a "power indicator" (size-depending): the "market share" of citations.

Introduction
Bibliometric indicators are increasingly used in the evaluation of scientific performance and the positioning of research players at various levels: ranking of scientific journals, academic departments, assessment of individuals, institutions, regions, nations. Pitfalls have been recurrently underlined, especially at the level of individuals. Bibliometric measures are data-dependent and field-dependent, as shown in numerous classic studies (Schubert and Braun,1996; Murugesan and Moravcsik, 1978) and one of the lasting issues is the delineation of reliable sets for establishing reference values (Zitt et al., 2005). Similar efforts are made by bibliometric research teams to achieve some truncation of literature either in a mild way, removing the low tail, or a strong way, eliciting excellence. Composite measures of the h-index family (Hirsch, 2005), robust towards changes in the low tail in individual distributions - extended to aggregate levels in a variety of ways - are another way to achieve this goal, with normalization issues however.
Here we start from a vantage point of *relative indicators*, the comparison of an actor's distribution of citations to a reference distribution of world articles in a field, with some level of aggregation. We establish the relative profiles of actors in various citation classes and then try to summarize the features of the profile. Here, we focus on non-parametric approaches, appropriate for universities and small actors, where poor fit is expected from parametric models. The focus is on "average citation performance", in a rationale closed to relative impact, but the profile analysis can be extended, with appropriate settings, to "power measures" such as market shares of citations, as shown in one example.
Section I presents methods for exploring the "relative citation profiles". There are two results sections, one devoted to an application at the macro level - countries (Section II), the second one to an application at the micro level, a comparison of French universities (Section III), before concluding.

Methods
Since Price (1976), an abundant literature has been devoted to citations distributions, either in the power-law framework (Rousseau, 1990, Bookstein, 1990) or other models (Burrell, 2001,

Proceedings of ISSI 2009, pages 40–51.
Edited by B. Larsen and J. Leta.

Van Raan, 2001). In this work we do not consider actors' and world citation distributions per se, but rather the relation expressed by their ratio, at any citation level. We define this basic tool with versatile applications, and then consider the two global measures of actors' performance which can be derived there from: one parametric, using a basic bibliometric model, the other based on point measures, among them over-rankings between pairs of acgtors. Finally we envision the extension of these indicators to "power" measures, such as citation market share.

Relative profiles on citation rankings

If we wish to compare the actor's activity with the world (or any other reference), we may either use classes naturally defined on the citation levels on the x axis (see e.g. Adams, 2007; another approach can be found in Glaenzel & Schubert, 1988) or use ranks/ quantiles. Here we picked the latter solution, presented in Zitt (2002, 2005, 2007). Let us recall the main points of this approach.

The x axis is rescaled using the world reference correspondence between frequency and ranks. The ranks/quantiles are those defined for the world reference. On the basis of various options, detailed below, the subset of articles of a particular actor is distributed amongst classes defined for a reference, here all world actors. The AIC is the ratio of the actor's presence to the expectation in a particular (non-cumulative) quantile class $q(r_1, r_2)$ with bounds r_1 and r_2, in a rationale of "activity index" relating an observed to an expected level of activity:

$$AIC(q) = (n_{aq}/N_a)/(n_{wq}/N_w)$$

where n_{aq} and n_{wq} represents the publications number of the actor and the world respectively, corresponding to the q quantile of the world distribution; N_a and N_w denote the total publications (all classes) of the actor and the world. The expectation is usually the nominal value[2] of the class, typically 5% for the top-cited half-decile of "excellence", 10% for the first decile, etc. If the actor places 7.5% of its production in this top 5% class (if nominal), the ratio AIC is 1.5. Rather than AIC - useful to single out particular classes - we use the indicator on cumulative classes, noted AICC, applicable to simple ranks as well as classes (quantiles), starting from the top one. AICC gives the over/under-representation of the actor for successive embedded sets of literature starting with the top cited class. For example:

$$AICC(r) = (n_{ar}/N_a)/(n_{wr}/N_w) = (n_{aq}/N_a)/(n_{wq}/N_w)$$

where the bounds of q are 1 and r. A flat profile identifying with the x axis indicates an alignment on the world average, a "descending profile", from left to right, indicates a good performance. The higher the steepness, the better the overall performance.

The AIC-AICC indicators depend on a few options:

- the usual prerequisites in citation analysis: time-window; type of counting: distinct integer count is a natural option here.
- the choice of scale matters for the indicators being built. A logarithmic transformation of ranks is recommended with respect to the distributions encountered and the abundance of ties in the lower tail of citation distributions. Evenly spaced classes on this scale implicitly value papers in highly cited classes, which are given equal weights in the profile, though containing small numbers of papers (a point discussed below). Citation data in a similar rationale have been made available from the ESI product for a few years (Thomson-Reuters, 2008).

[2] However, the presence of ties may result in quite a difference between expected values and nominal values, especially in the low-citations classes, implying that real frequencies, not nominal ones, are used as expected values in the denominator of the ratio.

- the choice of option for treating ties, a large number of ties being met in the low tail. Here we chose an even distribution of ties, without creation of void classes.
- the choice of ordering clause for articles in the database, allowing to define relative ranks/quantile classes, e.g. by field, subfield, etc. . For example, choosing the content of the 1% top cited class as the 1% top cited class in the WoS or the 1% top cited class field by field or the 1% top cited class journal by journal can be considered as options of normalization. As shown by Zitt et al. (2005), field-normalized indicators such as relative impacts are quite unstable when the level of normalization (the "calibration of the zoom") varies. Citation profiles embody field normalization choices by changing the ranking option for the x axis.
- classical impact indicators may be decomposed into a product of "normalized" and "structural" indicators, for example impact can be seen as the product of a relative citation ratio RCR (Schubert & Braun, 1986) and an impact factor. Not only "normalized indicators" can be represented using an appropriate scaling of the x axis, but also "structural indicators", such as impact factors based on journals where the actor publishes. The latter scale is obtained by grouping articles by descending impact of their journal.

A collection of country level profiles in double-log scale is shown in Fig.1 (see end of paper).

Summarizing the profile: Profile Performance Indicators and Profile Ranking Indicators

The scrutiny of the whole profiles is interesting for detailed assessment of the competitive position of actors. If we attempt to measure the overall performance, a restriction of the scope to "competitive research" should be considered. A result from above-mentioned works and ongoing projects is the necessity to truncate bibliometric series for benchmarking. The low tail of citation distribution, either for articles or journals, should not be deemed "mediocre" literature. If in some cases it reflects the low performance of articles, it also mirrors different modes of communication, represented by some nationally-oriented journals or "transfer" literature where prominent scientists also occasionally publish. The important point is that the competitive pressure is lower than in standard international literature. Because it is a different form of communication, we chose to discard it from these particular benchmarking exercises. Further comments about the removal of the low tail, at the journal level, may be found in Zitt et al. (1999) and Van Leeuwen et al. (2001). One of the motivations for the h-index (Hirsch, 2005) was to build an index which would not be sensitive to the bottom fraction of citation distribution at the individual level. A focus on elite papers also inspired Vinkler (2008). Whatever the method, truncation - either in terms of journals or in terms of papers - should be considered for evaluative purposes.

The scrutiny of the profiles in various scales suggests that the transition starts in the range 55% - 70% of the literature covered by the WoS in its present extension. The phenomenon then gathers strength, especially ca. the point 80, also with the massive effect of nationally-oriented journals. In the 20% less cited, mainstream countries are strongly under-represented, and conversely for BRICKs. Eventually the truncation at about 65% seems sensible, but the discussion on establishing firmly the threshold remains open. We eventually set the threshold at 64%, at the level of breakdown considered: fields, large disciplines, etc. The 36% less cited documents were discarded from analyses.

The upper part of the profiles also raises questions. Excellence indicators attract an increasing amount of attention in research evaluation (Tijssen and al., 2002). The general idea is that the smaller the considered elite (high thresholds), the stronger the unbalance between leading and lagging actors. Extending the scale on the side of exceedingly small elites induces strong fluctuations. We limited this effect by not extending the scale of excellence beyond 1% more cited. On this basis, three types of indicators are examined.

Parametric approach: PPI

For memory sake, this approach explores modeling/fitting of AICC series, typically for series on countries or big institutions which exhibit a good regularity. Fitting AICC series were first tested on few models reducible to a linear form by simple transformation, namely power law and logarithmic, with the constraint AICC(r_{max})=1. In this case a single parameter, the slope in log or double-log presentation, stands as the profile performance indicator (PPI). If both distributions, for the actor and the reference, followed a Pareto-I model, one can show analytically that in this configuration AICC(r) follows a hyperbolic scheme and the corresponding "relative impact" RI(r), a quite classical indicator, is constant and then a quasi-perfect indicator, reflecting the ratio of the slope parameters of the two distributions for a and w. However, power-law is not a very good approximation for the whole range of AICC, in accordance we found that empirical RI(r) for an actor is far from constant. A more satisfactory model will be presented elsewhere.

A basic indicator on measurement points

We argued (op.cit. 2007) that empirical measures based on smoothed AICC emulate the outcomes of regression models. Moreover, these measures are more adapted in the case of irregular profiles, met for example among small actors, especially universities. The profile fluctuations may hinder the goodness of fit of PPI models. In such cases, a rough estimate consists in retaining the observation of AICC on some conventional points of measures. The measurement points picked were exponentially spaced: in the present work, we used the series 1, 2, 4, 8, 16, 32, 64% of the most cited articles. The rest of the distribution was ignored (see below). A side-effect of exponential scale is the fact that each measurement point, from left to right, reflects a fixed proportion of new papers compared to the previous point in the scale. In the scale chosen in this particular experiment, this proportion is 1/2. We assume that the last reference point (64) is reasonably robust, taking a large part of the publications of most actors, and not yet jeopardized by the strong tail effects. Averaging could be practiced as well for this particular point.

Then the first six points retain all the information. The average measure of AICC on these points, noted PPI (profile performance indicator) in the followings, is expected to be strongly correlated with parametric estimates of the "slope" of regular profiles using the corresponding scale, and a reliable substitutes in other cases. We used a non-weighted mean of the log measures on the six points, giving the same importance to the fact that an actor performs well in each of the cumulative classes defined. This does not mean that papers in each class are given the same weight, since the number of papers by class follows an exponential scheme. Then a paper in the first class has a very strong implicit weight, all the more so, that it remains present in the successive cumulative classes. However, compared with relative impact measure, this approach is less sensitive to deviant papers with extreme citation values, especially when the first class is set to 1% rather than smaller elite classes.

Over-rankings on measurement points

Basing ourselves on the same measurement points, we investigated a second non parametric approach, by pair of actors, with an appropriate visualization. Let us define a profile ranking indicator [PRI(ij)] which considers over-rankings between pairs of actors (ij) within a close set of actors, for example universities from a particular country. Complying with the citation profile approach, namely the cumulative form AICC, the PRI takes a few points of measurement in the profiles. The series 1 - 64 was used again. PRI(ij) reflects the bilateral comparison of actors i and j on each point of measurement k=1 .. n. The comparison C(ij) on a point k is defined by

$C(ij)_k = -1$ if $AICC(i)_k < AICC(j)_k$; $C(ij)_k = 0$ if $AICC(i)_k = AICC(j)_k$; $C(ij)_k = 1$ if $AICC(i)_k > AICC(j)_k$

PRI(ij) is then determined as the average of C(ij) over the n points of measurement, here n=6. By construction, $-1 \leq PRI(ij) \leq 1$ and PRI(ji) = -PRI(ij). For example, if university A over-performs university B in 5 of the selected points, out of six, the normalized score is 5/6 for A/B and -5/6 for B/A. If a couple (ij) exhibits a value of 1, we can speak of "complete dominance". Other situations reflect "partial dominance" or "equality". The dominance in the "excellence" or the "competitive" areas could be singled out, using respectively the first and last half of the series of points. More sophisticated weighting schemes might be envisioned, depending on how citations are valuated.
Dominance PRI(ij) for a pair of actors is defined after a set of ranking options for articles in the database, here the WoS. The next indicator, PRI(i), assumes a set of reference, here the set of universities being compared. It denotes the average of PRI(ij) over all actors j compared with i. It can be considered as a global measure of performance, but depends on the reference set: it will exhibit quite different values if the reference chosen for comparing French universities is the set of all universities in this country, as the example shown here, or the set of top performing 100 world universities.
A theoretical study of dominance in relation with valuation function of scientific production is found in Carayol & Lahatte (2008) comparing pairs of actors. These authors study how generic properties of such valuation function do transform into various dominance relations and especially rule dominances when two profiles are crossing. Their work is based on number of citations, a "power indicator", in contrast with PRI which is a "performance" indicator primarily independent from size. One may notice that both the scale used for defining classes, in the present work, and the degree of extension towards smallest quantiles of excellence, in PRI calculation, may be interpreted in the light of valuation functions. For example, instead of starting with the 1% most cited class, we might use a further decomposition of the excellence range by adding levels such as 0.5%, 0.25%, 0.12%, giving the super-excellent articles a still larger weight, at some risk of interpretation.

Extension to multi-scale profiles

As mentioned above ("the choice of an ordering clause"), the performance may be observed at several scales, for example the journal level, the specialty level, the discipline level. Any number of levels, provided that each is judged legitimate, may be addressed, for example by giving each level the same weights. If we use 6 measurements points by level as above, we may for example simply multiply the number of points, with equal weighting: say 18 points for three levels together: large discipline, specialty, journal. We limit ourselves to give results on a single-level approach.

Extension from performance to "power" indicators

AIC-AICC and related PPI belong to the family of " performance indicators", primarily independent from the institution's size, in contrast with the number or market share of publications and citations which mix performance and scale effects and can be considered as "power indicators". The simplest extension of the h-index (Hirsch, 2005) to aggregates, by considering the institution as a macro-individual, is also a power indicator. A power indicator can be designed by weighting the citation profiles by some function of the actor's size, the simplest option picked here is using the direct number of publications. The ratio AIC for a class is $(n_a/N_a)/(n_w/N_w)$. The simplest measure of power is the multiplication of AIC or AICC by the total market share (N_a/N_w), the modified AIC or AICC is then reduced to the actor's market share by class $MSC=(n_a/n_w)$, respectively cumulative class (MSCC). Different valuation of the size criterion might be used - at the expense of simple interpretation - giving a variety of composite indicators.

In contrast with the cumulative AICC, converging to the unit value for the last class of citation, the corresponding market share of an actor reaches its final value on the last point, for the selected delineation of the database, original or adjusted to a threshold (here 64%). As mentioned, the scheme by Carayol & Lahatte follows a "power indicator" rationale.

The global profile is characterized by MPPI and MPRI, corresponding respectively to PPI and PRI. The comparison of homologous indicators, respectively based on AICC and MSCC, is appealing, in relation with the issue of increasing returns to size.

Sources

Primary data come from the Thomson-Reuters database offline version contracted by the *Observatoire des Sciences et Techniques* (OST Paris). For the country level treatment, types of documents considered citable, including reviews, were used. Citation quantiles are based on OST indicators. The publication year 2001 and a 5-years citations window are used at the macro level. The micro-level approach is based on the IPERU project data, concerning 101 French universities and higher education schools *Grandes Ecoles* which are associated to the French ministry of research and higher Education. This project embodies an elaborate unification process, achieved by self-identification of actors on the French scene. In a first round data, used here, social sciences and humanities were not considered and if most universities carried out this self-identification, the data are not complete for the others: the results should then be considered as partial, especially for the "market shares" of publications. The indicators were calculated by OST on publication year 2004 data, with citations cumulated from 2004 through 2006. For the micro-level study, only the "article" type of document was considered. At the project IPERU is on-going, we had at this stage to keep the results anonymous, universities are identified by conventional numbers.

Results at the macro-level: countries

Fig. 1 presents selected country profiles, in double-log scale. A representation in linear scale, on earlier data, is found in Zitt et al, 2007 (see end of this paper for all figures).

Average performance on point measures

The average measure on selected points of AICC is reported, together with over-ranking measures. It correlates rather well with measures on appropriate parametric models (not shown), and with global over-ranking scores (Table 1).

Table 1. PPI and PRI for selected countries

Countries	PPI (Ave log AICC)- 6pts	PRI (Ave dominance)	Countries	PPI (Ave log AICC)- 6pts	PRI (Ave dominance)	Countries	PPI (Ave log AICC)- 6pts	PRI (Ave dominance)	Countries	PPI (Ave log AICC)- 6pts	PRI (Ave dominance)
CHE	0.40	0,94	ISR	0.17	0,33	AUT	0.06	-0,12	KOR	-0.35	-0,65
USA	0.33	0,83	DEU	0.15	0,23	FRA	0.05	-0,12	CHN	-0.40	-0,72
ISL	0.29	0,70	CAN	0.14	0,25	AUS	0.01	-0,17	RUS	-0.44	-0,75
NLD	0.25	0,67	SWE	0.13	0,20	ITA	-0.05	-0,32	BRA	-0.54	-0,83
DNK	0.23	0,71	BEL	0.09	0,25	ESP	-0.18	-0,43	IND	-0.66	-0,96
GBR	0.19	0,52	FIN	0.08	-0,03	JPN	-0.23	-0,52			

Over-ranking on point measures

Figure 2 shows the dominance matrix between countries in the "all disciplines" case, based on AICC series, on a close set of 23 countries, selected both amongst mainstream and emerging countries. Countries are ranked by their global dominance index in the set. The relation ij between country in row i and country in column j is coded in white if i over-ranks j completely (i.e. on every measurement point), in black if j over-ranks i completely. Shades of grey denote partial over-rankings. This matrix turns out to be a simple and practical tool, which can also suggest groupings into several clusters, each one built by cross-links of partial dominance (large squares with shades of grey). Groups clearly emerge in Fig. 2 and 3.

Extension to citation "market shares"

Unsurprisingly, the domination scheme is more clear-cut when over-rankings on market shares of publication are represented instead of AICC (Fig. 3, all disciplines).Two indicators were calculated, as for pure performance, but based on MSCC: MPPI, based on averaging on the measurement points; MPRI, based on over-rankings. Atypical situations are found for small countries such as Iceland, which is never totally dominated because of a top performance in a highly-cited fraction. Switzerland is ranked first with an average score close to 1, which means that it is better in every measurement point against the other countries. On the other hand, the BRICK group (Brazil, Russia, India, China, Korea) is still totally dominated, being consistently over-ranked on the selected points; shades of grey on the bottom right of matrix emphasize nevertheless a competition within BRICKs in various extensions of cited sets.

Results at the micro-level: experiment on French universities

The current globalization context leads high education or research institutions, universities in particular, to adapt to an increasing competition. Each university has to play several roles in the knowledge economy, in terms of education, research and innovation. The rankings of universities based on an evaluation of their performance or activities, both on national and international levels, become crucial for visibility, reputation and management of institutions. Many ranking exercises have been proposed in the recent years, with unequal levels of emphasis on bibliometric indicators: Shanghai JTU ranking (2007, update) by Liu & Cheng (2005); Times Higher Education ranking (THE-QS, 2008, update); CWTS Leiden (CWTS, 2008; Moed, 2006); Thomson-Reuters ESI (2007, update); CHE various rankings, e.g. excellence ranking of European graduate programs (2007, update). Department level rankings have also developed, for example for economics departments (Coupé, 2003) and Lubrano et al. (2003) among others. All rankings convey explicit and implicit limitations and biases, and the multiplicity of published rankings, an illustration of the relativity of points of view, should warn against naive conclusions. The present ranking is particularly limited in scope, since only based on citations measures.
In terms of systems size (organization and governance is another question) there is some overlap between micro and macro-level entities: some universities have a much larger scientific output than small countries. On average however, statistical irregularities in profiles, expected from small actors, are more likely to be found at the institutional level, making parametric approaches of citation profiles less appropriate. Only measures on the six measurement points were used.

Settings

In this section we focus on an experimental comparison of 101 French universities and higher education institutes. PROs as such were not included in the present work, but the major part of publications of some PROs is covered because of massive overlaps with university, mainly in the form of joint labs. This is the case for CNRS. Other PROs overlap with university at a lesser but significant extent (for example INSERM in medical science and fundamental biology, CEA mostly in physics, INRA in applied and fundamental biology).

The analysis was carried out at the "large discipline" level (8 academic disciplines + multidisciplinary). In this implementation, we used the Thomson Reuters re-assignation of most articles in Nature, Science and PNAS to subject categories. Normalization then is based on the "large discipline" level, just for example sake.
In contrast with size-dependent power measures (market shares, institutional h-index) which alleviate significance issues by marginalizing small actors, average performance measures such as PRI or relative impacts, can be trapped by small samples/populations configurations, making thresholds necessary. For simplicity sake, absolute thresholds were used: a university was taken in a discipline-level table if it published at least 40 articles, except for mathematics and 'multidisciplinary' where the threshold was set to 20. The size-level selection is more effective for higher education institutes than for universities, generally bigger. Results are given here for the discipline of "fundamental biology" where 29 actors are above the threshold

Results

<u>Performance profile and over-ranking matrix</u>

The correlations indexes of

(a) PPI: average value of AICC on the six measurement points, and

(b) PRI: the average dominance from over-rankings on the 29 actors is found in Table 2

Similarly, (c) denotes MPPI and (d) MPRI, established for citation shares (see below).
(a) and (b) are highly correlated as expected, especially on Spearman ranks, same for (c) and (d).

Table 2 - Performance and Market share Indicators: Pearson and Spearman-rank correlations

	PEARSON Correlation N = 29 ***& Prob > |r| under H0: Rho=0***				**SPEARMAN RANK Correlation N= 29** ***& Prob > |r| under H0: Rho=0***			
	(a)PPI: log AICCAve	(c)MPPI MSharAve	(b)PRI AICCrkg	(d)MPRI MShareRkg	(a)PPI: log AICCAve	(c)MPPI MSharAve	(b)PRI AICCrkg	(d)MPRI MShareRkg
(a)PPI	1.000	0.404 *0.030*	0.918 *<.0001*	0.389 *0.037*	1.000	0.414 *0.026*	0.936 *<.0001*	0.386 *0.039*
(c)MPPI		1.000	0.540 *0.003*	0.984 *<.0001*		1.000	0.511 *0.005*	0.988 *<.0001*
(b)PRI			1.000	0.530 *0.003*			1.000	0.509 *0.005*
(d)MPRI				1.000				1.000

The over-ranking of pairs of actors is presented Fig 4 in a particular discipline (fund. biology). The matrix is coded in grey-shades as the previous ones. Odd rendering may appear due to irregular profiles, such as a concentration of publications in a few classes of citations, more likely to happen for small actors. In such a case, long wakes may appear along rows/columns, the actor avoiding to be completely dominated by excellent actors, or to be completely dominant over less performing ones. As for countries, configurations may suggest groupings by actor linked by partial dominance relations.

Profile on a power measure: citation "market shares"
The average dominance obtained for a university respectively on the AICC and the MSCC (not shown) revealed, as expected, quite different. The correlation between PRI and its homologous MPRI on over-ranking, and the correlation between PPI and MPPI on point measures averages, suggest a weak effect of increasing returns to size, especially considering that fundamental biology is a rather favorable case. This topic is analyzed in parallel projects.

Discussion and Conclusion

The citation profiles help to visualize phenomena in the successive classes of visibility, including classes of excellence and non-competitive literature. They also suggest various types of indicators in the rationale of relative measures, versatile towards normalization issues and suitable for measure of average "pure performance" as well as "market shares of citation" and composite indicators. We focused here on two types of indicators, non-parametric and applicable whatever the regularity of the profile: the first one is a rough approximation of profile "slope" using measurement points corresponding to successive selections of cited sets, arranged in a specific scale; the second one is based on over-ranking within pairs of actors, using the same measurement points. The latter two approaches are more appropriate at the university level. The results show a high level of agreement between cardinal and rank approaches. Performance scores and citation world shares scores reveal themselves lowly correlated, even in the discipline of fundamental biology, which suggests, for French universities, a rather low effect of size, a finding rejoining other ongoing works. It should be noted that these results on French universities are purely exploratory and need further improvements, especially for the identification of actors - self-unification of actors was achieved for most universities, but not all of them, at the date of the study - and the choice of normalization level(s). The extension to European universities, and more generally to world-class universities, may be envisioned.
Let us conclude on the specific interest of these indicators on relative profiles. In contrast with indicators of the h-index family, they are relative measures, depending on field references, and are not built as composites, although intermediaries between their two main instances, pure performance and market shares, may be thought of; on the other hand, they share with the h-index family the idea of ignoring information on the tail of bibliometric series, although in a quite different way. Now if we compare the profile indicators to classical performance measures such as the relative impact, they share the mastering of field-normalization but, starting from the profile analysis, they are more appropriate for controlling truncations and extreme values effects, and likely to be modeled in more concise way. Other features: they are immediately generalized to multi-scale measures, in a particularly simple way if one considers that basic levels such as journals, specialties, disciplines, are equally legitimate; depending on the configurations obtained, the over-ranking matrixes can also suggest dominance-based grouping of actors.

References

Adams J., Gurney K., Marshall S. (2007) Profiling citation impacts: a new methodology, Scientometrics 72, 2, 325-344

Bookstein, A. (1990). Informetric distributions, part II: resilience to ambiguity. *Journal of the American Society for Information Science, 41*(5), 376-386.

Burrell Q.L. (2001). "Ambiguity" and scientometric measurement: A dissenting view, Journal of the American Society for Information Science and Technology, vol 52, n°12, pp. 1075-1080.

Coupé, T. (2003). Revealed Performances: Worldwide Rankings of Economists and Economics Departments, 1990–2000, *Journal of the European Economic Association, 1(6), 1309-1345*.

CWTS, Leiden University (2008). *The Leiden Ranking 2008*. Retrieved December 15, 2008 from: http://www.cwts.nl/ranking/LeidenRankingWebsite.html. .

CHE Excellence ranking of European graduate programs (2007, update) Retrieved January 17, 2009 from http://www.daad.de/deutschland/hochschulen/hochschulranking/06543.en.html.

Carayol N., Lahatte A. (2008) Dominance Relations and Universities Ranking, Research in progress communication, 2nd ENID-PRIME Indicators Conference on STI Indicators for Policy, Oslo, May 2008.

Thomson-Reuters, Essential Indicators: Retrieved January 17, 2009 from http://www.thomsonreuters.com/products_services/scientific/Essential_Science_Indicators.

Glänzel, W., Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, 14(2), 123-127

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569-16572.

Institute of Higher Education of Shanghai Jiao Tong University. (2007). *Academic Ranking of World Universities*. Retrieved on May 20, 2008 from: http://www.arwu.org/rank/2007/ranking2007.htm.

Liu, N.C., & Cheng, Y. (2005). The Academic Ranking of World Universities. *Higher Education in Europe*, 30(2), 127-136.

Lubrano M., Bauwens L., Kirman A. and Protopopescu C (2003). Ranking Economics Departments in Europe: A Statistical Approach. *Journal of the European Economic Association, 1(6), 1367-1401.*

Moed, H. F. (2006). *Bibliometric Rankings of World Universities*. Retrieved June 25, 2007 from: http://www.cwts.nl/hm/bibl_rnk_wrld_univ_full.pdf.

Murugesan, P., & Moravcsik, M. J. (1978). Variation of the nature of citation measures with journal and scientific specialties. *Journal of the American Society for Information Science, 29*, 141-155.

Price, D. J. d. S. (1976). A General theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science, 27*, 292-306.

Rousseau R. (1990), Relations between continuous versions of bibliometric laws, *Journal of the American Society for Information Science, 4,* 197-203

Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics, 9*(5-6), 281-291.

Schubert, A., & Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics, 36*(3), 311-324.

Shanghai Jiao Tong University, Institute of Higher Education (2007). *Academic Ranking of World Universities*. Retrieved on May 20, 2008 from: http://www.arwu.org/rank/2007/ranking2007.htm.

Tijssen, R. J. W., Visser, M. S., & van Leeuwen, T. N. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics, 54*(3), 381-397.

THE-QS (2008). *World University Rankings-The Methodology: Citations per Faculty*. Retrieved December 15, 2008, from http://www.topuniversities.com/worlduniversityrankings/methodology/citations_per_faculty

Van Leeuwen, T. N., Moed H.F., Tijssen R.J.W., Visser M.S. , Van Raan A.F.J (2001), Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance, Scientometrics 51, 1, 335-346

van Raan, A. F. J. (2001). Competition amongst scientists for publication status: Toward a model of scientific publication and citation distributions. *Scientometrics, 51*(1), 347-357.

Vinkler P. (2008). The pi-index. A new indicator for assessing scientific impact (submitted to Journal Information Science).

Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics, 62*(1), 117-131.

Zitt M., Bauin S., Filliatreau G. Co-ords. (2002) Bibliometric indicators of public research institutions - Indicateurs bibliométriques des institutions publiques de recherche (1997 - hors sciences humaines et sociales), Production Coopérative d'indicateurs inter-institutionnels de politique scientifique, OST/MR, http://www.obs-ost.fr. (in French) study updated in 2004

Zitt M., Ramanana-Rahary S., Bassecoulard E. (2005). Relativity of citation performance and excellence measures: from cross-field to cross-scale effects of field-normalisation, *Scientometrics*, *vol 63, n°2, 373-401.*

Zitt M., Ramanana-Rahary S., Bassecoulard E. (2003) Correcting Glasses help Fair Comparisons in International Landscape: Country Indicators as a Function of ISI Database Delineation. *Scientometrics*, *vol 56, n°2, 259-282*
Zitt M., Bassecoulard E., Filliatreau G., Ramanana-Rahary S. (2007) Revisiting country and institution indicators from citation spotlighting competitive literature. In: Torres-Salinas D. and Moed H.F. (eds). *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and informetrics. CSIC, Madrid, Spain. June 25-27, 2007* Madrid (SPA): CINDOC-CSIC, pp. 797-802

Figures

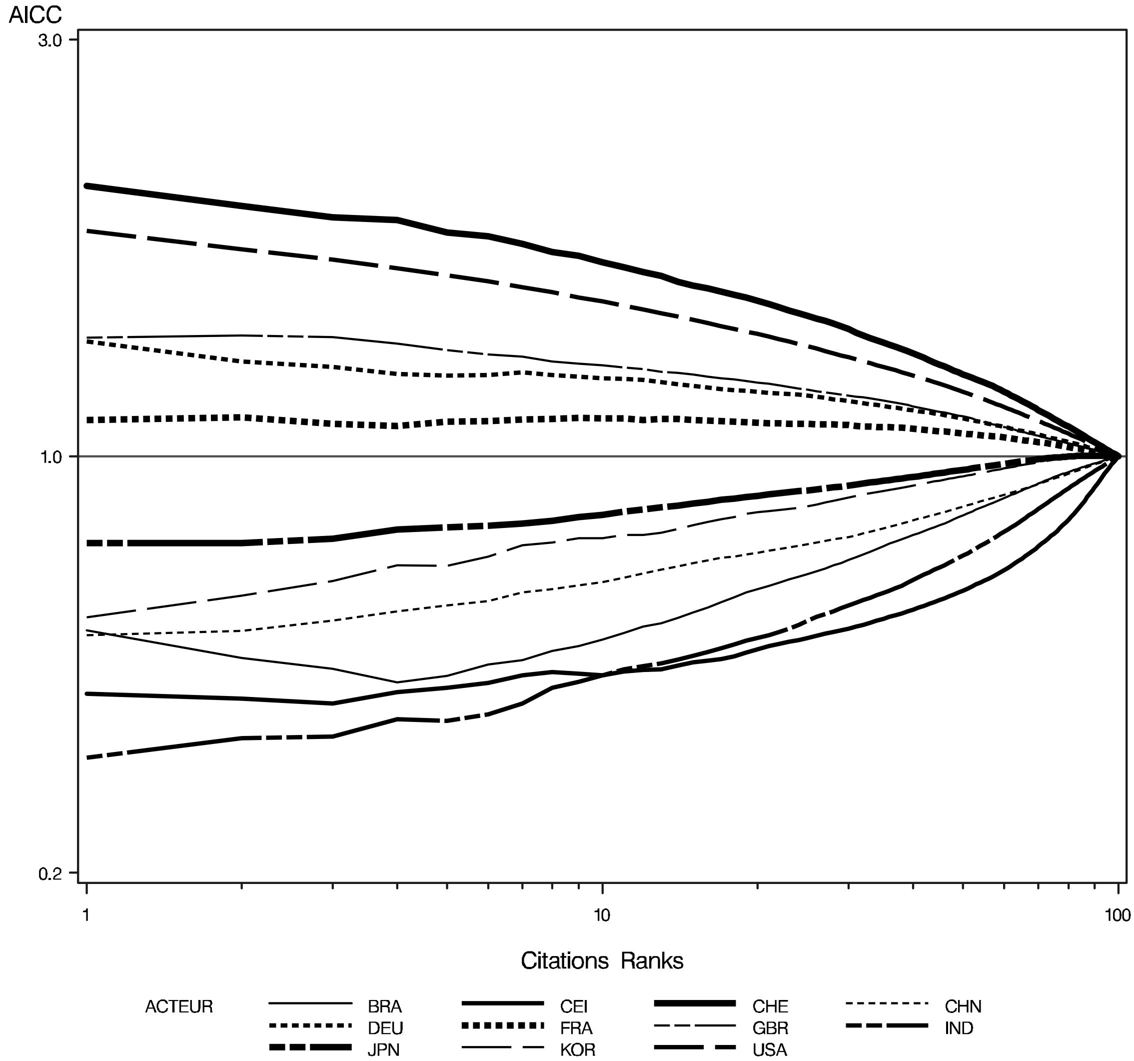

Figure 1. Selected country profiles : double-log plot of the cumulative activity index (AICC) vs. citation levels

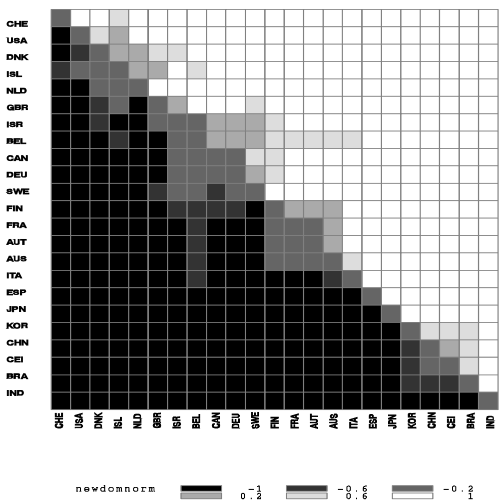

Figure 2. Over-ranking matrix (performance) – countries

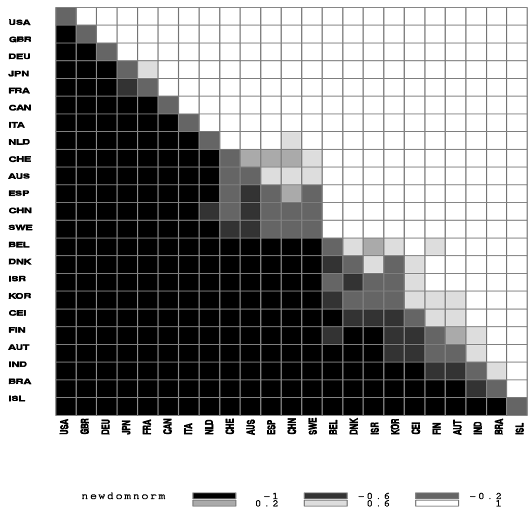

Figure 3. Over-ranking Matrix (citation shares) – countries

Row-wise reading: the actor in row totally dominates the actor in column if the cell is white, and is totally dominated if the cell is black. Grey shades depict partial dominance. Actors are ranked by total average dominance (same ranking row-wise and column-wise); more dominant actors are in the top rows and the left columns. All information is present in a triangular matrix.

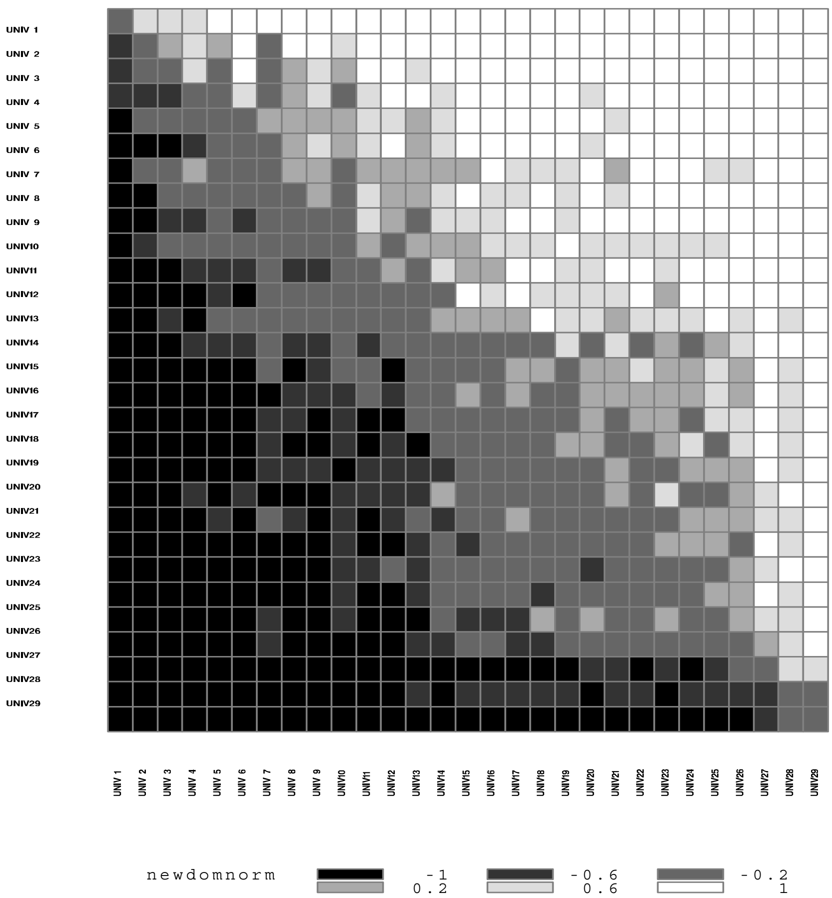

Figure 4. Over-ranking Matrix - performance - French universities (fundamental biology)