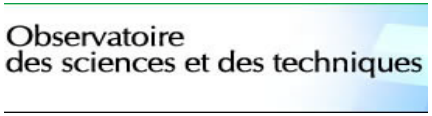
Performance and over-ranking measures on citation profiles

Michel Zitt, Agénor Lahatte

Performance and over-ranking measures on citation profiles

Michel Zitt*/**, Agénor Lahatte**
*INRA Lereco, Nantes, France
**OST, Paris, France



plan

- Introduction : midway between indicators to multicriteria universities ranking: expanding the citation impact family
- Citation profiles : analytical power, flexibility
- Aggregate measures on profiles and focus on over-ranking
- Discussion and perspectives on typology

Conclusion

Context

- Universities rankings are typically based on a multicriteria assessment of indicators covering several missions and dimensions (ARWU, THES, CHE...)

- Here we limit ourselves to a single type of indicator (citations) within a single mission (research) but, within this narrow perspective, we try and show that a « familial » expansion of this indicator is natural, as it needs to reflect various forms of academic competition.

- For this we based ourselves on a flexible instrument, the citation profiles, and observed how summaries can be found.

Relative profile indicators

Recall of Relative Profile indicators – already presented at ISSI-11 Madrid

Compare citation distributions of an actor and a reference (default: world) :

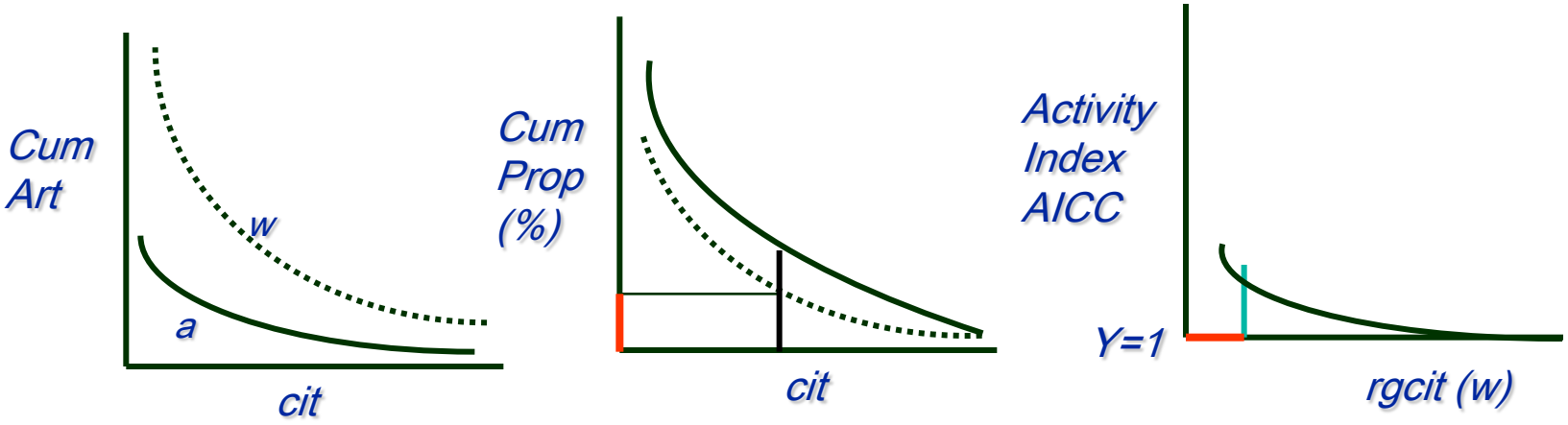
- - ordinal approach (ranks or classes of citation: in abscissas WoS articles are ranked according to some field-normalization criteria : all science ; discipline, etc. (cf. infra). *Zitt et al. 2002, 2004, 2005,* (already presented in ISSI Madrid). See also:
- - cardinal approach based on standardized citations (« impact profiles » *Adams et al.,* 2007)
- - direct comparison of distributions is put forward by economists (*Carayol & Lahatte.,* 2007)

- Principle: rank articles of the database by decreasing citation score, and define quantiles:
- Non cumulative quantiles → **Activity index in citation quantiles**
- AIC= % of actor's publication in quantile q / % of world's publication in quantile q
- Cumulative quantiles → **Activity index in cumulative citation quantiles**
- AICC= % of actor's publication in quantile Q / % of world's publication in quantile Q

- This measure follows a rationale of **performance assessment, akin to the relative impact**. It can be extended to a power indicator: market share of citations by cumulative or non-cumulative quantile.

Relative profile indicator: construction

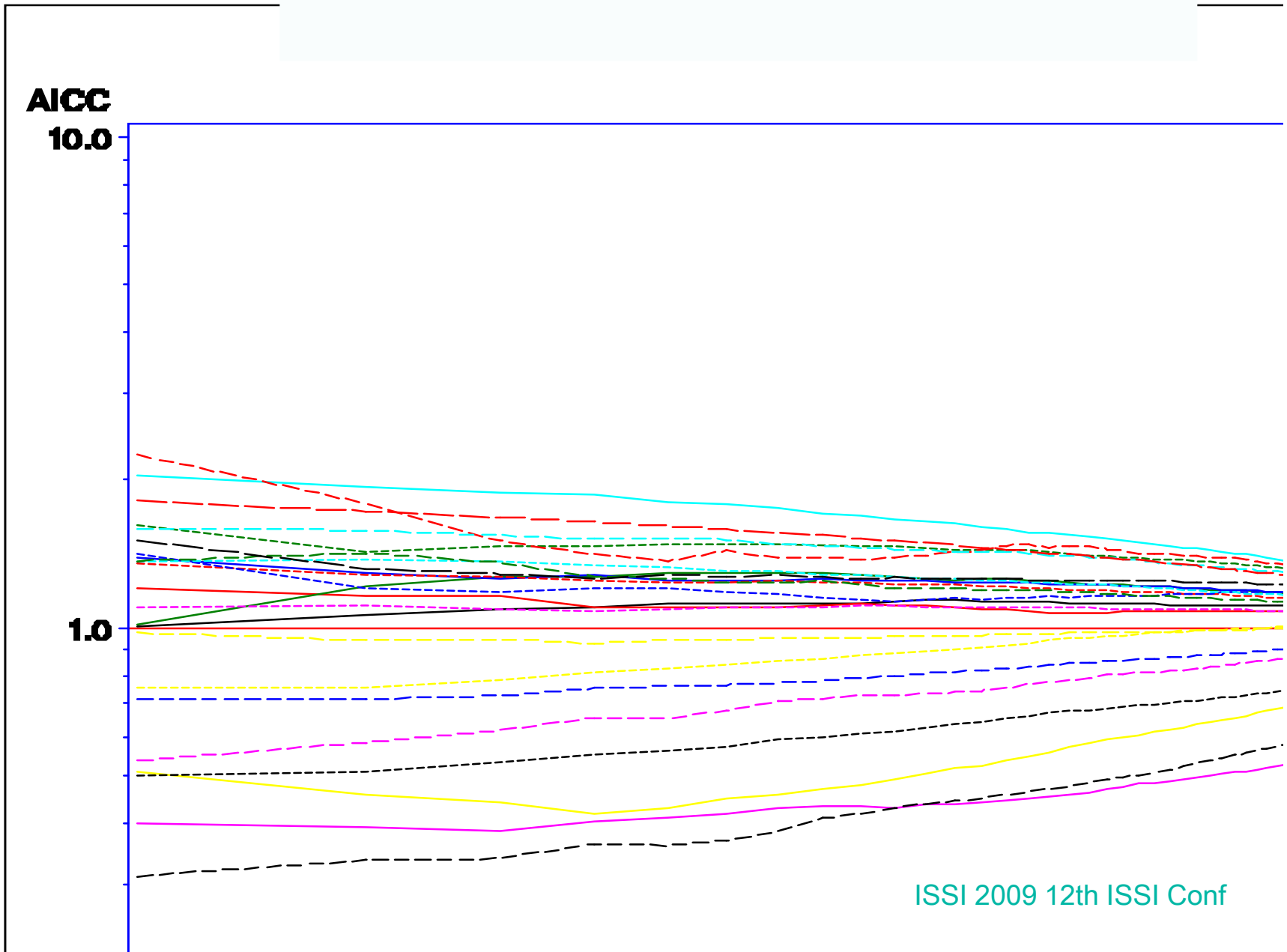
AICC Index on cumulative classes



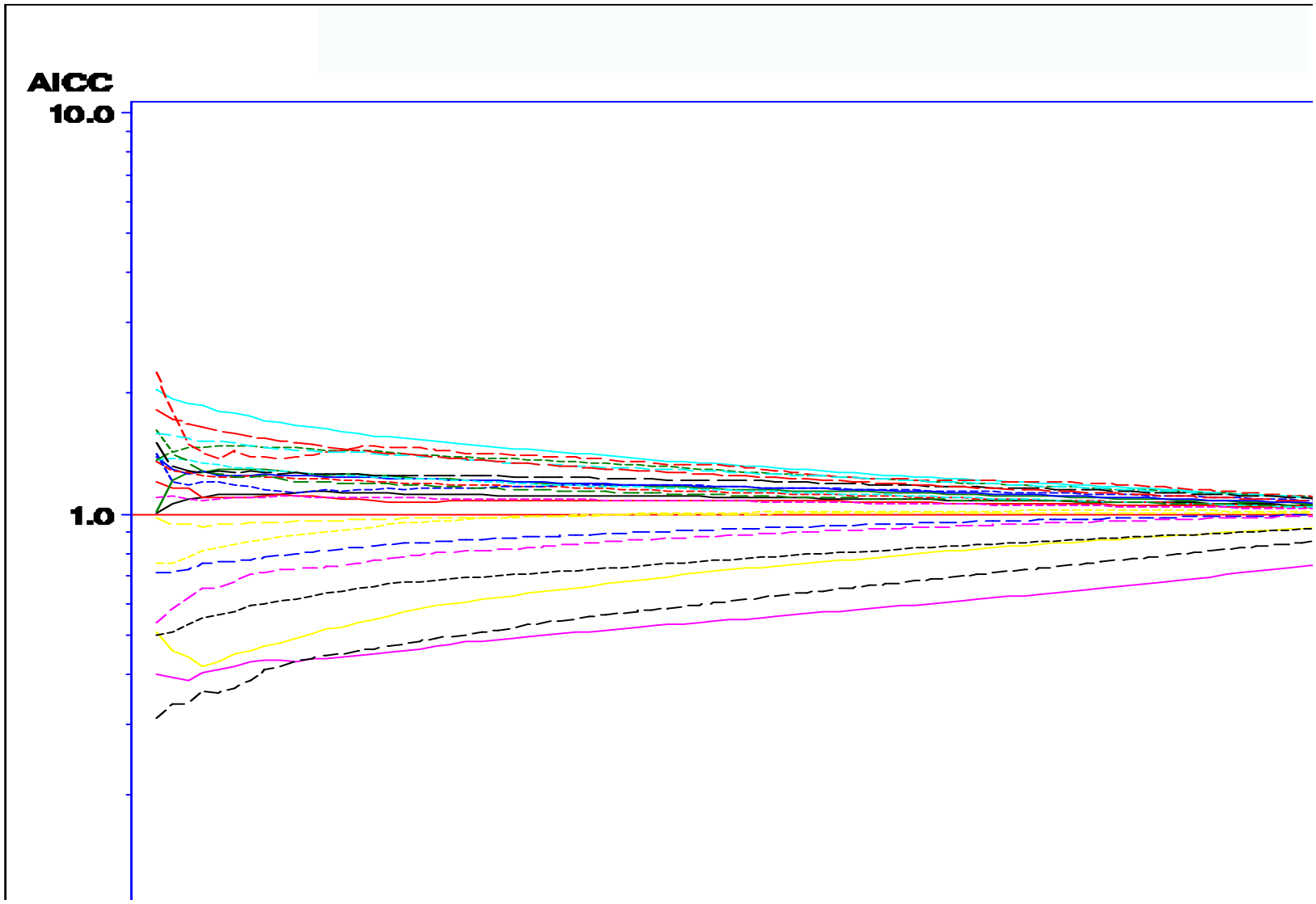
a actor

w World (=default reference)

Example of profiles: national level (AICC) - db log scale



Examples of profiles: national level (AICC): log-lin scale



Profiles fit to simple laws

General shape of profiles (AICC)

- Area of high citations : approximately power law
  Pareto $AICC(r)=K_1 r^{-\beta}$
- Area of low citations : approximately exponential
  Exponential $AICC(r)=K_2 e^{-\alpha r}$

Mixed model : « Truncated power law » (Amaral et al.,2000 ; Achard et al., 2006)

TPL $AICC(r)=K r^{-\beta} e^{-\alpha r}$

Fits are interesting for regular profiles, typically the country level.

Limits are met for small countries, or most actors at the university level.

To deal with irregularites and crossings of profiles, an empirical approach, based on a few measurement points, is preferred. One checks that in the case of regular profiles, the two analyses converge.

Measurement points

Principle:

- 1. Each AICC profile is summarized by series of evenly spaced and weighted measurement points, in order to be , for a given actor, sensitive to any performance in any delineation.
- 2. Several AICC series can be used, each accounting for a level of normalization.
- .

→ 1st principle: reflecting visibility performance in each stratum of the competitive literature

a performance limited to the excellence area (to be performing in the 1% most cited literature, say) or a performance limited to the competitive area (to be performing in the 40% most cited literature) are considered both as valuable. :

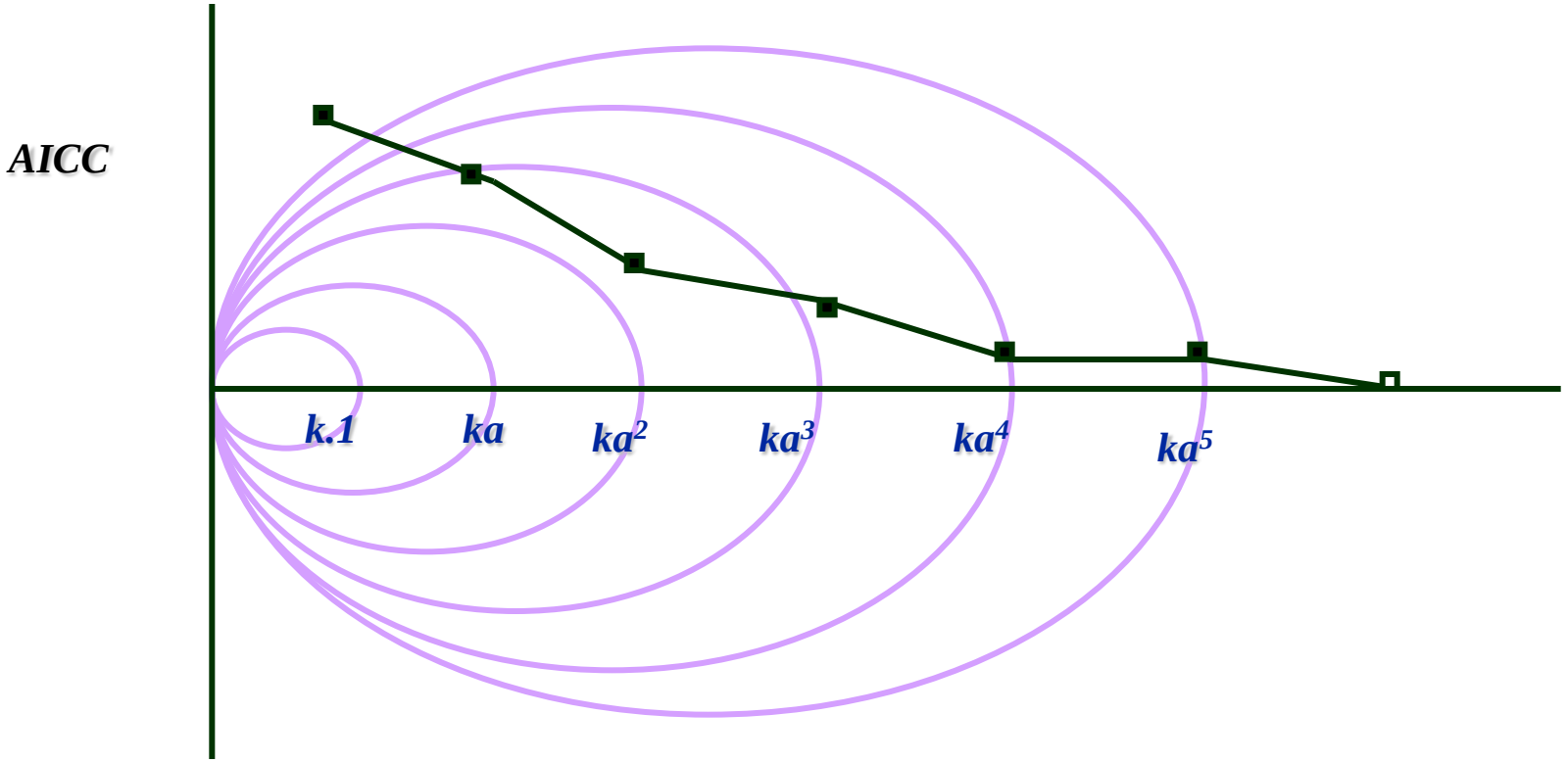
- given the distribution, an exponential scale is chosen and measurement points are equally spaced.

- the last third of the WoS, here, is not considered. This area contains low visibility literature, deliberately for a part of it (secondary communication, including from top authors)

- in practice here, 6+1 measurement points have been picked corresponding to 1, 2, 4, 8, 16, 32 % most cited publications The last point of measure, 64, corresponds to the unit value.

The same process may be extended to other types of indicators, especially market shares. The difference occurs with the last point of measure, which no longer corresponds to the unit value.

Measurement points



Evenly spaced points in an exponential scheme ka^n. In this experiment, k=1, a=2, n=5 (6 meas. pts). The implicit weight of each publication follows the same scheme, the top fraction is then implicitly favored. Its implicit weight could be still enhanced by extending the scale on the left-hand side, at the risk of creating too scarce sets. In the cumulative setting of AICC, each point includes a constant fraction of « new » sources of information (papers) over the predecessor on the left.

Citation profiles and normalization

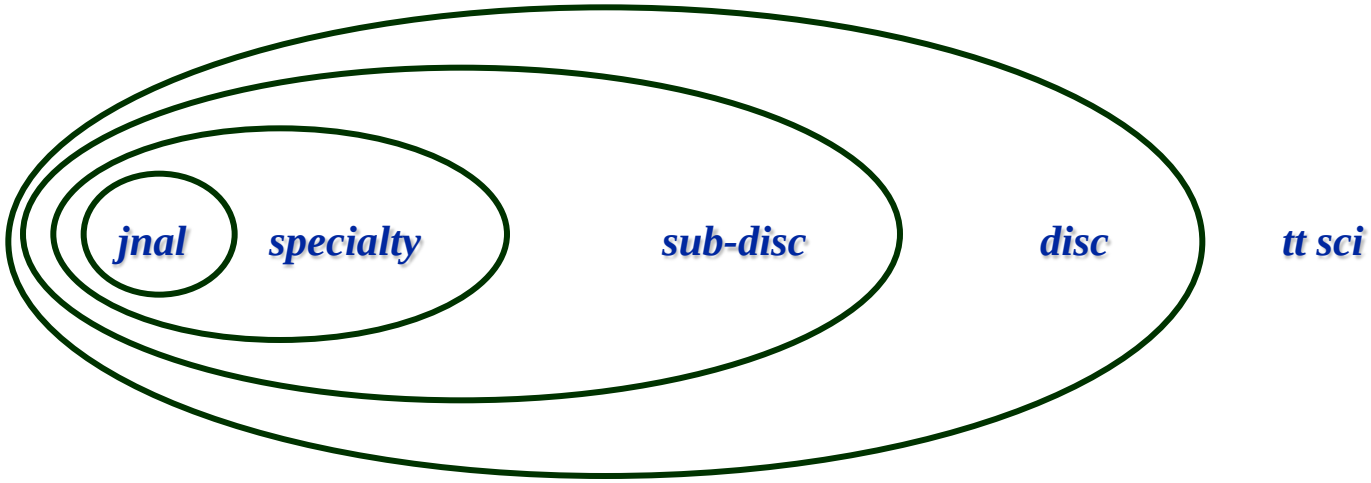
- Field-normalization is among the most classical issues in bibliometrics. It assumes options of field delineation and choice of scale of aggregation. Normalization implies two levels, the level of normalization and the level (superior) of aggregation/ observation. An aggregate profile starting with this lower-level breakdown is simply obtained by aggregating articles from each quantile. This operation implicitly conveys the'weight' (in terms of number of articles in a quantile) of each field – which is the natural choice.

- In the above profiles, the level of normalization is embodied in the scale of abcsissas, simply by changing the set on which the ranking is applied: for all science; or discipline by discipline; or subject cat by subject cat; etc. This allows us to set the « zoom » (MZ, SRR, EB, 2005) and also allows the implementation of structural indexes instead of normalized indexes, such as impact factors. **Levels of normalization typically alter the positions of types of research (basic vs. applied)**.

- Using several levels of normalization implies using several series with n measurement points

- → **2nd principle: use as many series as relevant/legitimate level of normalization**

The scale question



Heterogeneity : a grouping at any level of nomenclature is not in general representative (regarding citation practices) of its embedding higher-level groups.

Over-rankings

Let us consider a close set of actors under scrutiny . We compare pairs of actors i,j and define their over-ranking quite simply.

For a given normalization level (for example « subject categories »)

n measurement points are used. For each point k (1..n), one defines the over-ranking $(ij)_k$

$s(ij)_{k(1..n)}$ =sgn($AICC(i)_k$ - $AICC(j)_k$) [+1, 0, -1]. Aggregation on n points:

$S(ij)=(\sum_{k(1..n)} s_k)/ n$ avec S(ij) [-1;+1] ; S(ji)=-S(ij)

The mesure is positive when actor i over-ranksactor j on a majority of points. Complete over-ranking of i over j corresponds to S(ij)1.

By convention, the average over-ranking vis-à-vis all other actors is :

$S(i.)=\sum_j S(ji) /(m-1)$ where m is the number of actors.

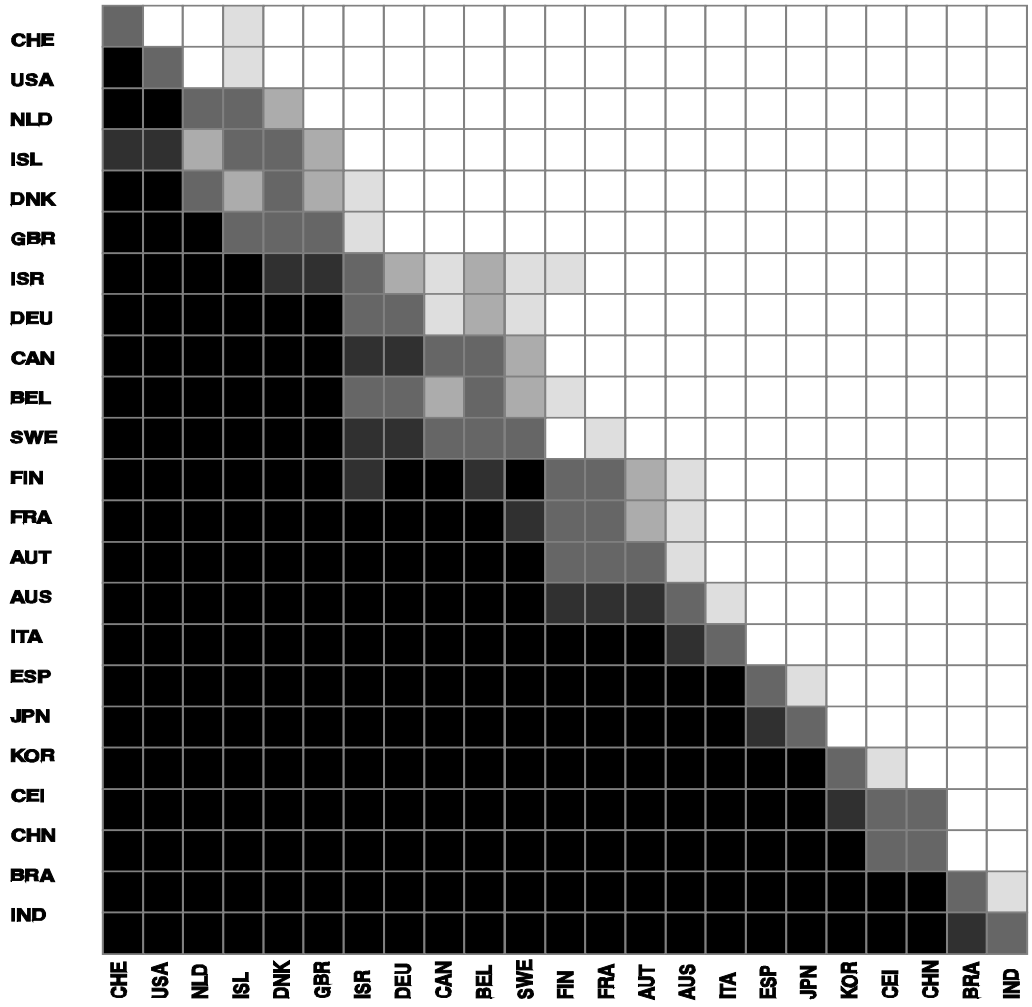
For several levels of normalization

The definition is extended by taking 2n. (for 2 levels) , 3n, etc…

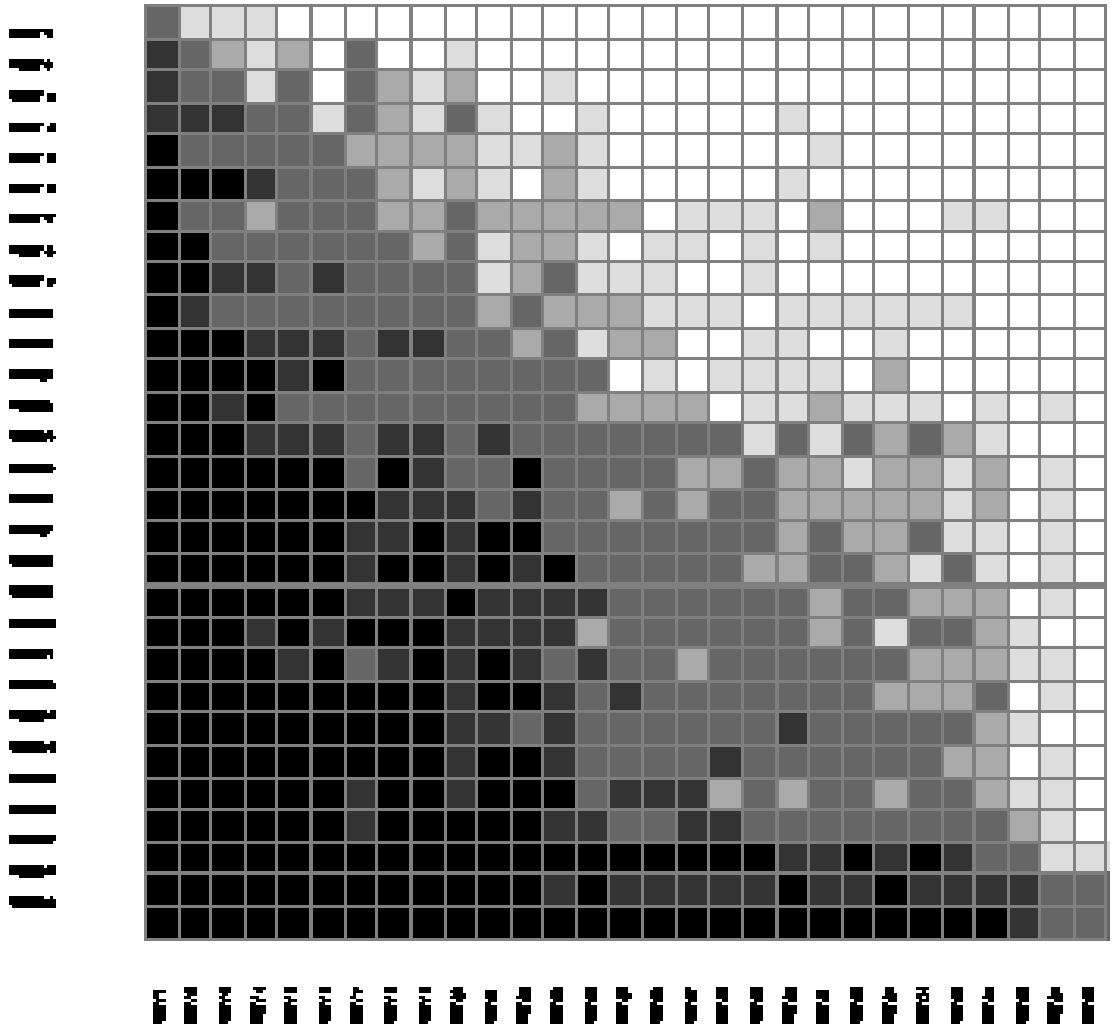
Extension to absolute performance (citation-market share)

Easy extension, but the indicator of maximal rank is no longer unitary.

Matrix over-rkg - countries: AICC, all disc, norm disc + spec



Matrix over-rkg - univ:
AICC, fundamental biology, norm disc



clustering applications

Matrixes of over-ranking suggest clusters that group actors where most pairs of actors are not in a situation of complete over-rankings.

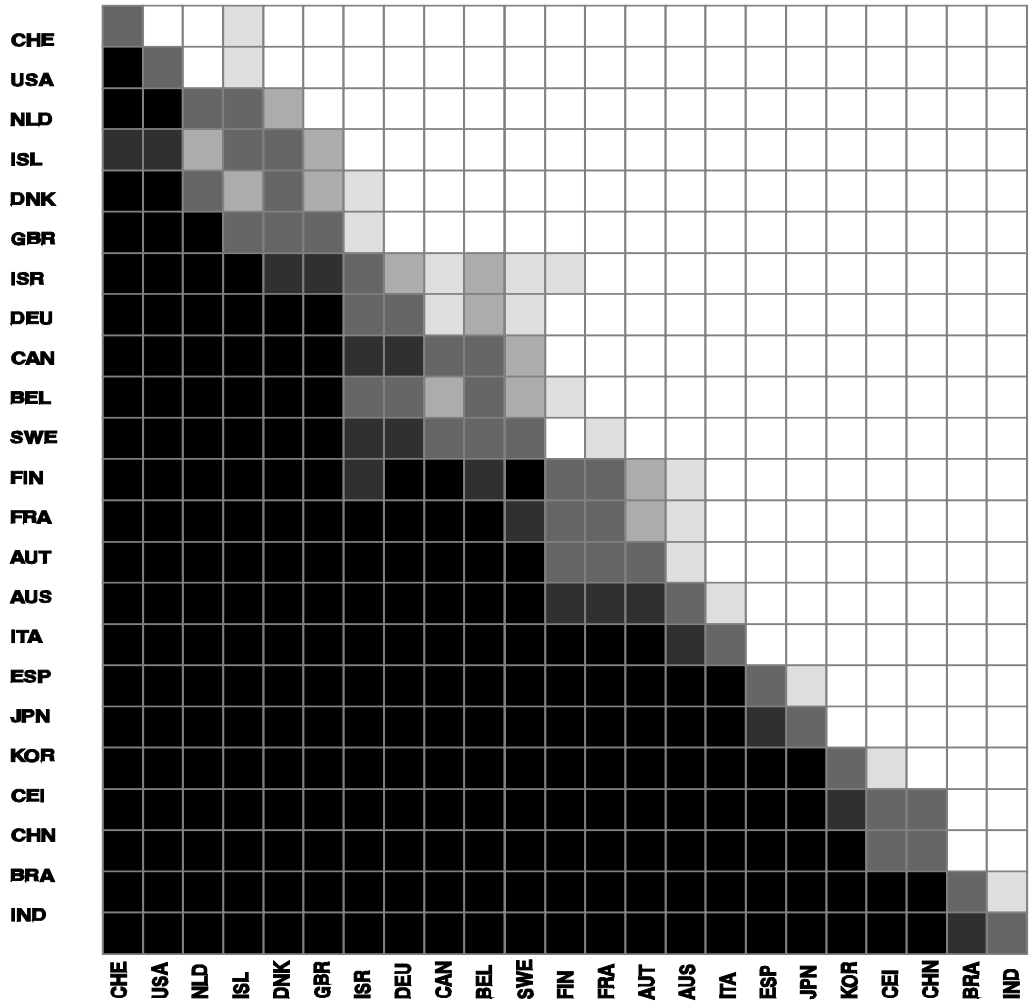
These possible groupings appear in the matrix in the form of quasi-squares in various shades of grey, in other words they tie on some of the measures

Standard clustering methods may be applied to these structures. Clearly, some of the structures associate:

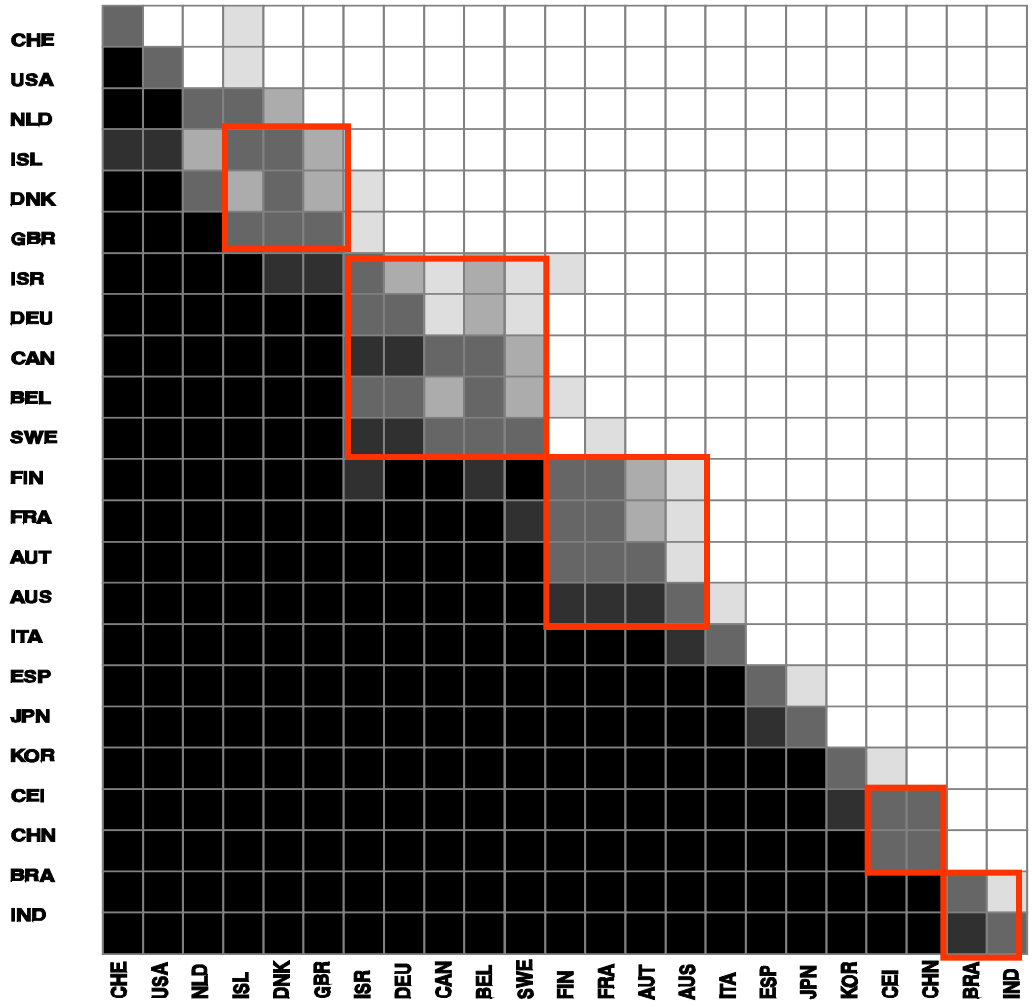
- hierarchies of total over-ranking (grey diagonal without grey neighbors)
- clusters (quasi-squares in shades of grey)

In some cases, clustering give poor results: in absence of natural classes, relative neighborhoods (specific of each actor) are recommended for benchmarking purposes

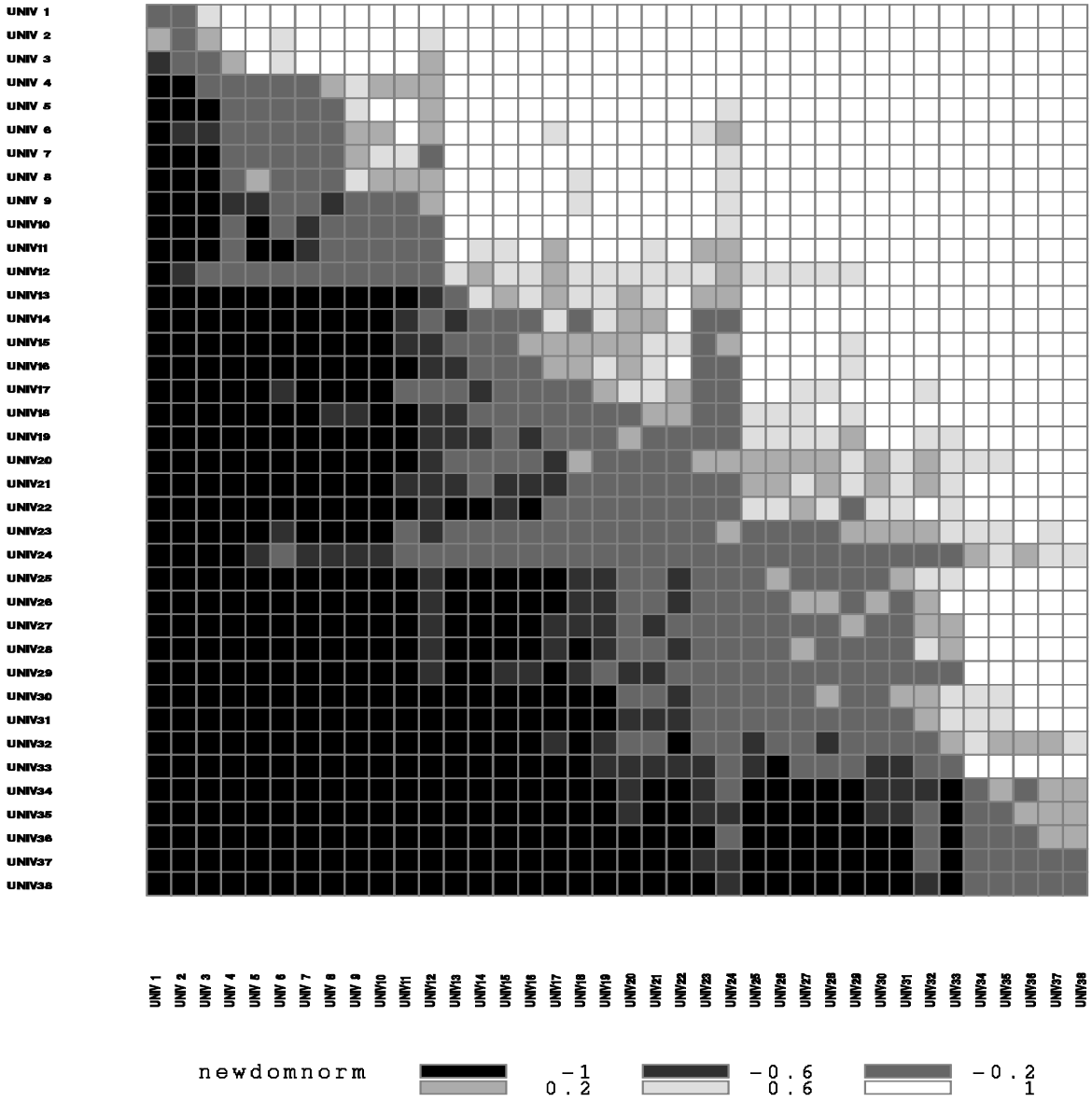
Matrix over-rkg - countries: AICC, all disc, norm disc + spec



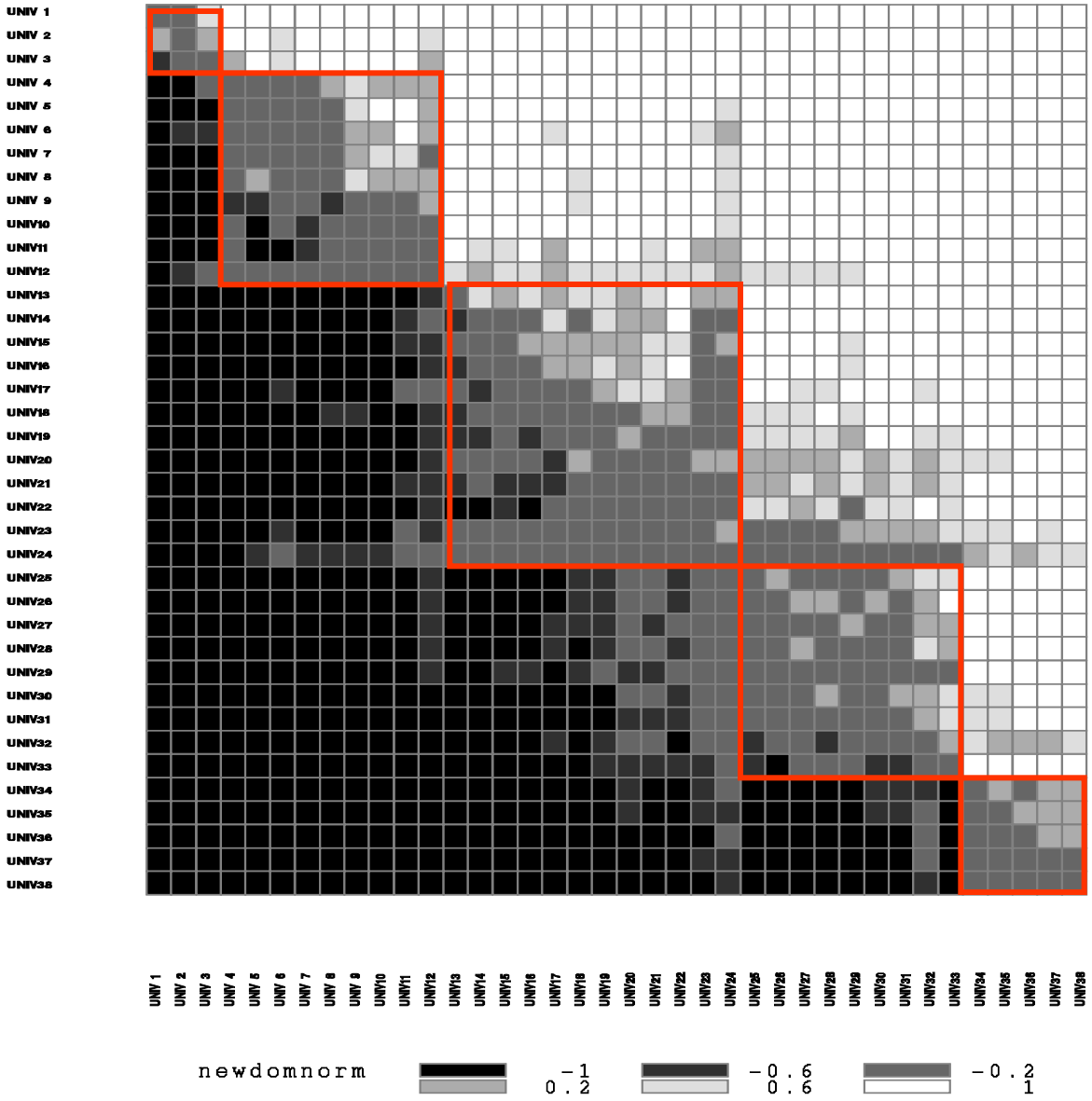
Matrix aver-rkg - countries : AICC, all disc, norm disc + spec



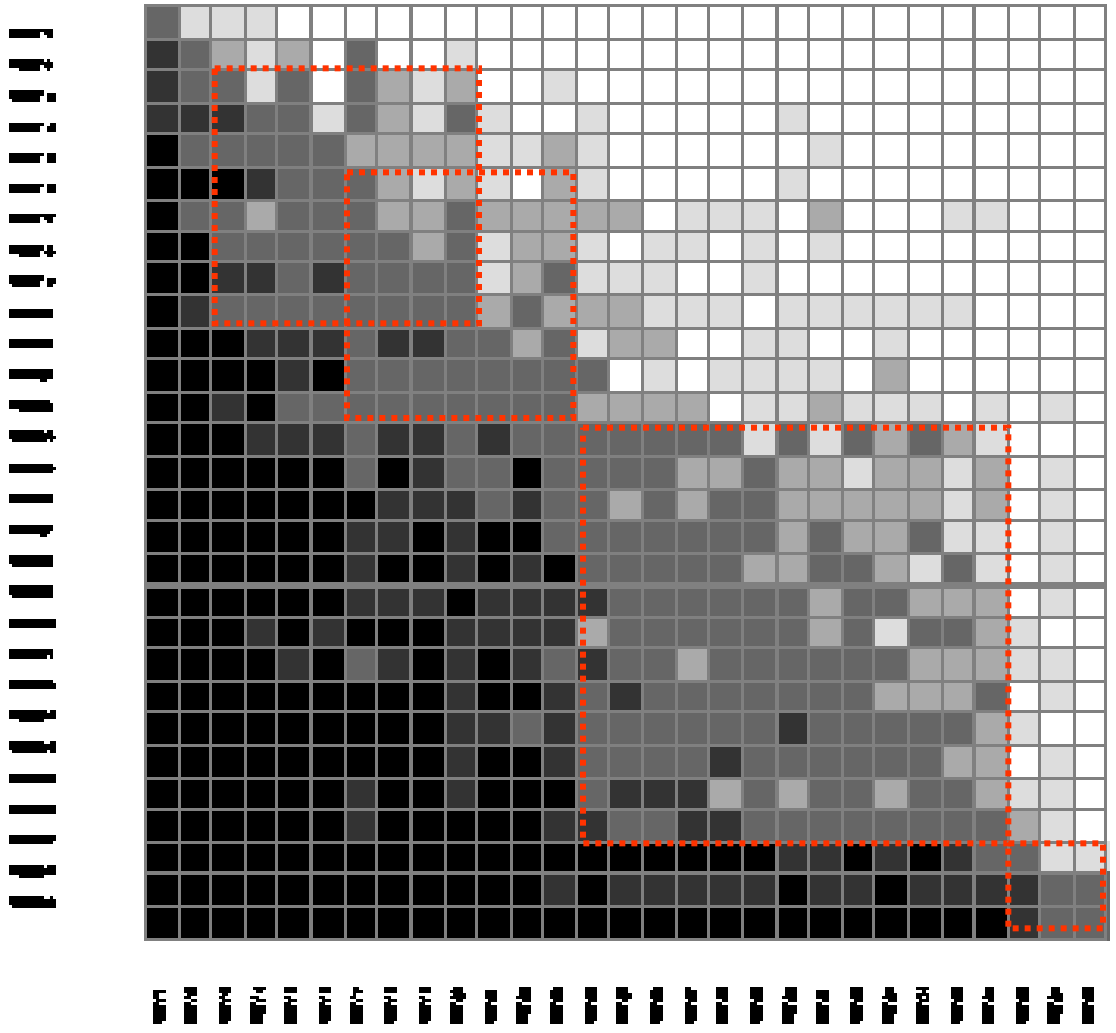
Matrix over-rkg univ:
AICC, physics, norm disc



Matrix over-rkg univ:
AICC, physics, norm disc



Matrix over-rkg univ:
AICC, fundamental biology, norm disc



In summary : synthetic indicators

	Average measure	Avg ranking on a close set of actors
AICC (activity index)	PPI Profile perf indicator	PRI Profile ranking indicator
MSCC (market share of cit)	MPPI Citation market share profile indicator	MPRI Citation market share profile ranking indicator

Discussion

- Analysis of relative citation profiles suggest discussion and perspectives
- Sensitivity/ stability of over-ranking measures
- Direct development: typological applications.
- Theoretical questions
  - - linkage between AICC-Impact relatif
  - - linkage with the theoretical framework of valuation functions (Carayol-Lahatte)
  - - linkage with new normalization methods

Conclusion

Exploration of a class of citation indicators, in a rationale of either relative impacts or citation-market shares, exhibiting the following properties of multiplication of points of views corresponding to areas of competition :

- sensitivity to the performance on at least one of selected subsets of visibility
- sensitivity to the performance on at least one of level of breakdown (specialty, disciplines, etc.)
- scale of axes adapted to the distribution and helping to control large deviations

Over-ranking indicators (in a style close to “dominance indicators” in economics) allow both an analytical and synthetic comparison of actors.

Over-ranking indicators suggest typologies which enlighten both ranked lists of individual actors, and groups tied to some extent.

If we intend, by those indicators, to extend the narrow focus of average citation measure, we do not pretend that they can stand alone, especially for general ranking purposes.

Sources OST

ARTICLE

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, pp. 373-401.*

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.

CONF PAPERS

Michel Zitt, Elise Bassecoulard, Ghislaine Filliatreau, Suzy Ramanana-Rahary (2007), Robust indicators from citation rankings and the question of three literatures, *ISSI 11th Intl Conf., Madrid*

& (2007) Citation rankings profiles: a framework for macro and micro indicators, *IREG-3, Shanghai*

REPORTS

Zitt M., Bauin S., Filliatreau G. Coord.. (2002) Bibliometric indicators of research institutions 1997 (in French) Group Publications of the Cooperative production of inter-institutional indicateurs for science policy, OST/MR, http://www.obs-ost.fr/pub/groupeIpsir.pdf (updated 2004) .

Thanks for your attention

INRA

Observatoire
des sciences et des techniques