BIBLIOMETRICS FOR RESEARCH EVALUATION: AN INTRODUCTION

Wolfgang Glänzel

Centre for R&D Monitoring and Dept MSI, KU Leuven, Belgium

ECOOM

Structure of presentation

Research evaluation is one **important, however, not the only** field of application of bibliometric methods.

Bibliometrics depicts essential aspects of scientific activities by *quantitative* and *statistical* methods, and its output proved to be a valuable supplement to qualitative methods such as peer reviews.

Bibliometrics has developed tools to quantify that part of research output, which is documented in the framework of scholarly communication.

Measures of different aspects of research output are called "indicators". Most basic indicators are determined for

- publication output (as measure of productivity),
- co-authorship (as measure of collaboration) and
- citation rates (as measure of impact)

or the combination thereof.

Most indicators are derived from simple counts of items extracted from various bibliographies and databases.

Advanced measures are "network indicator" derived from the analysis of co-authorship and citation networks.

The broad issue of co-authorship is covered by a separate lecture.

Measures of different aspects of research output are called “indicators”. Most basic indicators are determined for

- publication output (as measure of productivity),
- co-authorship (as measure of collaboration) and
- citation rates (as measure of impact)

or the combination thereof.

Most indicators are derived from simple counts of items extracted from various bibliographies and databases.

Advanced measures are “network indicator” derived from the analysis of co-authorship and citation networks.

The broad issue of co-authorship is covered by a separate lecture.

Measures of different aspects of research output are called “indicators”.
Most basic indicators are determined for

- publication output (as measure of productivity),
- co-authorship (as measure of collaboration) and
- citation rates (as measure of impact)

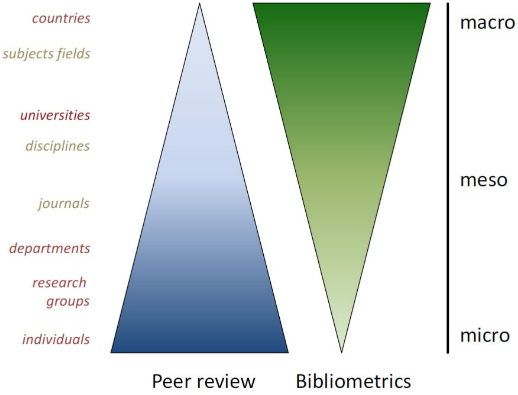
or the combination thereof.

Most indicators are derived from simple counts of items extracted from various bibliographies and databases.

Advanced measures are “network indicator” derived from the analysis of co-authorship and citation networks.

The broad issue of co-authorship is covered by a separate lecture.

The weight of qualitative (peer evaluation) and quantitative (bibliometrics) methods as function of the aggregation level



Source: GLÄNZEL, 2011

Theoretical foundation of indicators

Although a deterministic approach might be sufficient to build systems of indicators, a stochastic approach has several advantages:

- Provides mathematical interpretations beside the bibliometric ones.
- Helps understand complex structures such as communication networks.
- Provides information about statistical reliability, random errors and confidence intervals of indicators.
- Provides information about lower bounds for indicator applications and when reference standards can be taken at face values.
- Allows predictions of the expectation and probability of future events.

Source of bibliometric studies are usually papers published in periodicals and serials.

Only conveyors of original scientific information are included. These documents are considered citable items.

Citable items comprise: *research articles*, *short communications* and *notes*, *letters*, *reviews*, and *proceedings papers*.

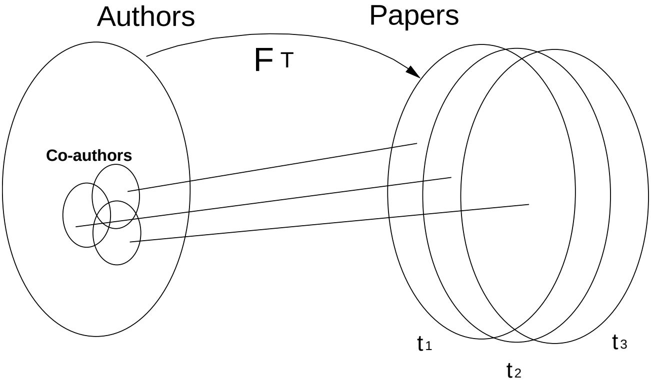
For instance, book reviews, editorials, corrections/errata, meeting abstracts and reprints are not considered original research output.

National publication counts and percentage of publications in each document-type for the 11 most active countries and the world total (2007)

Country	All papers	A	L	R	B	E	M	Rest
					Percentage			
USA	392,488	66.5	2.4	5.0	0.6	5.4	19.4	0.7
UK	104,561	65.9	4.6	5.9	1.2	5.5	16.2	0.6
Germany	95,892	72.3	1.7	4.6	0.1	3.2	17.4	0.6
China PR	95,231	92.0	0.6	1.5	0.0	0.7	4.8	0.3
Japan	89,575	78.8	1.5	2.6	0.0	1.2	15.4	0.4
France	63,656	77.6	2.0	4.4	0.1	2.8	12.5	0.6
Canada	57,500	71.7	2.1	4.9	0.5	3.9	16.2	0.6
Italy	55,223	72.7	3.5	4.6	0.1	2.6	16.0	0.5
Spain	41,274	75.9	3.2	4.0	0.1	2.5	13.8	0.5
Australia	35,327	72.4	3.3	5.9	0.6	4.0	13.3	0.5
India	32,842	86.4	3.5	2.6	0.0	2.1	4.7	0.6
World total	1,299,678	68.6	2.8	3.7	0.5	4.8	16.6	2.9

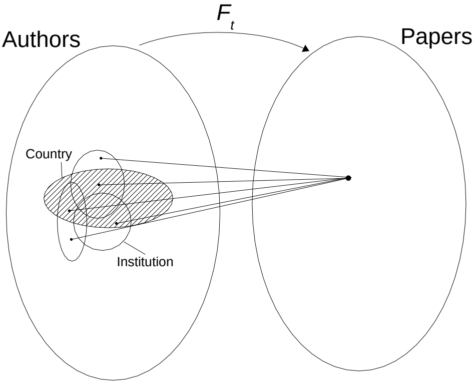
Source: ZHANG ET AL., *JASIST*, 2011

The publication process from the bibliometric viewpoint at time t and in the period $T = [s, t]$



Source: GLÄNZEL, *Bibliometrics as a Research Field*, 2003

Example for different levels of aggregations represented by overlapping systems of subsets



Source: GLÄNZEL, *Bibliometrics as a Research Field*, 2003

Counting schemes for publication activity

Counting schemes are the method according to which publications are to be assigned to the contributing units.

1. The fractional counting scheme: if n units (authors, institutions, countries, etc.) have contributed to the paper in question, each contributing unit takes the value $1/n$ for this paper (partially additive)
2. The first address count: a paper is assigned to one unit only, on the basis of the first address in the address list of a paper (additive)
3. The full or integer counting scheme assigns a co-publication fully to each contributing unit (non-additive)

☞ From the mathematical viewpoint, the credit distribution for the contribution of i authors (or units) can be represented by some proper weights (a_i). For the above three cases we have then (1) $a_i = 1/n$ for $\forall i$, (2) $a_1 = 1$ and $a_i = 0$ if $i > 1$, and (3) $a_i = 1$ for $\forall i$.

Counting schemes for publication activity

Counting schemes are the method according to which publications are to be assigned to the contributing units.

1. The fractional counting scheme: if n units (authors, institutions, countries, etc.) have contributed to the paper in question, each contributing unit takes the value $1/n$ for this paper (partially additive)
2. The first address count: a paper is assigned to one unit only, on the basis of the first address in the address list of a paper (additive)
3. The full or integer counting scheme assigns a co-publication fully to each contributing unit (non-additive)

☞ From the mathematical viewpoint, the credit distribution for the contribution of i authors (or units) can be represented by some proper weights (a_i). For the above three cases we have then (1) $a_i = 1/n$ for $\forall i$, (2) $a_1 = 1$ and $a_i = 0$ if $i > 1$, and (3) $a_i = 1$ for $\forall i$.

Counting schemes for publication activity

Nowadays only fractional and full counting are in use.

The first-address count is obsolete; it has historical roots in the architecture of bibliographic databases. In the past most bibliographic databases recorded only one address (for reprint requests).

This counting scheme should not be applied any more; databases with incomplete address recording are not appropriate for bibliometric use.

Problems in counting publications.

- Contentual issue: Finding correct credit distribution for authors, institutions and countries. – *Unsolved*.
- Technical issue: Consistency of counting over different levels of aggregations.

First-address and fractional counts (with restrictions) can be summed up to the total.

⇒ Full counts have to be determined for each level separately.

Figures based on fractional counts cannot be summed up *among* different levels of aggregation and may not be used out of their context.

Example: The US contribution to a paper with authors from France, Germany and USA amounts to 1/3 if fractional counting is based on countries, to 1/2 if it is used for comparison between the US and the EU.

Problems in counting publications.

- Contentual issue: Finding correct credit distribution for authors, institutions and countries. – *Unsolved*.
- Technical issue: Consistency of counting over different levels of aggregations.

First-address and fractional counts (with restrictions) can be summed up to the total.

$\Rightarrow$ Full counts have to be determined for each level separately.

Figures based on fractional counts cannot be summed up *among* different levels of aggregation and may not be used out of their context.

Example: The US contribution to a paper with authors from France, Germany and USA amounts to 1/3 if fractional counting is based on countries, to 1/2 if it is used for comparison between the US and the EU.

Example for different fractional counting according to different levels of aggregation

SCI CDE with Abstracts (Jan 93 - Jul 93) (D4.0)

Authors:	Prassides-K Kroto-HW Taylor-R Walton-DRM David-WIF Tomkinson-J Haddon-RC Rosseinsky-MJ Murphy-DW
Title:	Fullerenes and Fullerides in the Solid-State - Neutron-Scattering Studies
Full source:	CARBON 1992, Vol 30, Iss 8, pp 1277-1286
Addresses:	UNIV-SUSSEX, SCH CHEM & MOLEC SCI, BRIGHTON BN1-9QJ, E-SUSSEX, ENGLAND RUTHERFORD-APPLETON-LAB, DIDCOT OX11-0QX, OXON, ENGLAND AT&T-BELL-LABS, MURRAY-HILL, NJ07974, USA

Source: Glänzel, *Bibliometrics as a Research Field*, 2003

Example for different fractional counting according to different levels of aggregation (cont'd)

This paper has 9 co-authors working at 3 different institutions that are located in 2 different countries.

Applying fractional counting to the example, each co-author contributes with a "share" of 0.111, each institution involved with 0.333 and each country with 0.500.

Fractional counts cannot be summed up among the different levels of aggregation.

Example for different fractional counting according to different levels of aggregation (cont'd)

	Full count	Fractional count
Prassides-K	1	0.111
Kroto-HW	1	0.111
…	…	…
Murphy-DW	1	0.111
Univ-Sussex	1	0.333
RUTHERFORD-APPLETON-LAB	1	0.333
AT&T-BELL-LABS	1	0.333
United Kingdom	1	0.500
USA	1	0.500

Watch out!

Fractional counts are only valid within the same aggregation structure (e.g., institutional, national, supra-national, etc.). Otherwise fractional counting might become inconsistent.

Example: A paper with three corporate addresses, say, France, Germany and USA counts 1/3 for each contributing country.
If the aggregation level is changed, so that instead of individual EU member states the EU (as a supra-national region) is the unit of analysis, the paper counts 1/2 each for the EU and the US.
The US contribution to the same paper has thus different weights (1/3 and 1/2) according the two aggregation levels.
Fractional counting at different levels must not be mixed in the same analysis.

Factors influencing publication activity

Most important (measurable) factors influencing publication activity

1. the subject matter
2. the author's age
3. the author's social status
4. document type
5. the observation period

At higher level of aggregations (e.g., at institutional or national level), the influence of the factors age and social status superpose since populations at this level are rather heterogeneous.

Factors influencing publication activity

Most important (measurable) factors influencing publication activity

1. the subject matter
2. the author's age
3. the author's social status
4. document type
5. the observation period

At higher level of aggregations (e.g., at institutional or national level), the influence of the factors age and social status superpose since populations at this level are rather heterogeneous.

Problems of subject assignment

Specialised databases provide hierarchical subject classification at document level.
These schemes allow retrieval at very low levels of classification, for very specialised topics.

For instance, the database *Mathematical Reviews* uses the three-level Mathematics Subject Classification system (MSC).

Large multidisciplinary journal databases often provide subject assignment through journals.
Papers are indirectly assigned to subjects through the journals in which they have been published.

Problems of subject assignment

Specialised databases provide hierarchical subject classification at document level.
These schemes allow retrieval at very low levels of classification, for very specialised topics.

For instance, the database *Mathematical Reviews* uses the three-level Mathematics Subject Classification system (MSC).

Large multidisciplinary journal databases often provide subject assignment through journals.
Papers are indirectly assigned to subjects through the journals in which they have been published.

This system can also be used for evaluative purposes and to build further hierarchical structures, e.g., with 2 to 4 different levels.

Narin, *Evaluative Bibliometrics*, 1976

Since assignment is not unique in both solutions, bibliometric indicators are not additive over subject categories, sub-fields, fields or most other units of aggregation.

Example: Hierarchical structure of the MSC code system (subject: Probability theory)

```
51-xx Geometry
52-xx Convex and discrete geometry
53-xx Differential geometry
54-xx General topology
55-xx Algebraic topology
57-xx Manifolds and cell complexes
58-xx Global analysis, analysis on manifolds
60-xx Probability theory and stochastic processes
62-xx Statistics
65-xx Numerical analysis
68-xx Computer science
```

```
60Axx Foundations of probability theory
60Bxx Probability theory on algebraic and topological structures
60C05 Combinatorial probability
60D05 Geometric probability, stochastic geometry, random sets
60Exx Distribution theory
60Fxx Limit theorems
60Gxx Stochastic processes
60Hxx Stochastic analysis
60Jxx Markov processes
60Kxx Special processes
```

```
60E05 Distributions: general theory
60E07 Infinitely divisible distributions; stable distributions
60E10 Characteristic functions; other transforms
60E15 Inequalities; stochastic orderings
60E99 None of the above, but in this section
```

Source: Glänzel, *Bibliometrics as a Research Field*, 2003

ISI Classification schemes

ISI/Thomson Reuters classification is based on journal assignment.

☞ Two subject classification systems by TR are commonly used.

1. ISI Subject Categories (part of the citation indexes and the JCR)
   - Fine grained
   - Forms a fuzzy system with multiple assignments
2. ESI Fields (part of the Essential Science Indicators)
   - Coarse classification
   - Forms a partition with unique assignment

Building a new hierarchically structured classification scheme

I) The “cognitive” approach (setting the categories):
An initial scheme was elaborated on the basis of both the experience of bibliometricians and external experts.

II) The “pragmatic” approach (journal classification):
The journal set extracted from the WoS was classified into the preset subfields. The scheme has been adjusted according to co-heading frequency to keep multiple assignments within reasonable limits.

III) The “bibliometric” approach (article classification):
Articles published in core journals can be unambiguously classified into the subfield of the given journals. Articles of ambiguously assignable journals are classified individually using the analysis of references.

Building a new hierarchically structured classification scheme

I) The “cognitive” approach (setting the categories):
An initial scheme was elaborated on the basis of both the experience of bibliometricians and external experts.

II) The “pragmatic” approach (journal classification):
The journal set extracted from the WoS was classified into the preset subfields. The scheme has been adjusted according to co-heading frequency to keep multiple assignments within reasonable limits.

III) The “bibliometric” approach (article classification):
Articles published in core journals can be unambiguously classified into the subfield of the given journals. Articles of ambiguously assignable journals are classified individually using the analysis of references.

Building a new hierarchically structured classification scheme

I) The "cognitive" approach (setting the categories):
An initial scheme was elaborated on the basis of both the experience of bibliometricians and external experts.

II) The "pragmatic" approach (journal classification):
The journal set extracted from the WoS was classified into the preset subfields. The scheme has been adjusted according to co-heading frequency to keep multiple assignments within reasonable limits.

III) The "bibliometric" approach (article classification):
Articles published in core journals can be unambiguously classified into the subfield of the given journals. Articles of ambiguously assignable journals are classified individually using the analysis of references.

The hierarchical structure of the ECOOM scheme

1. level: 15 major fields
2. level: 65 sub-fields
3. level: about 250 subject categories

The ECOOM classification scheme

Hierarchical structure of a scheme based on ISI categories (subject: non-internal medicine)

1. AGRICULTURE & ENVIRONMENT
2. BIOLOGY (ORGANISMIC & SUPRAORGANISMIC LEVEL)
3. BIOSCIENCES (GENERAL, CELLULAR & SUBCELLULAR BIOLOGY; GENETICS)
4. BIOMEDICAL RESEARCH
5. CLINICAL AND EXPERIMENTAL MEDICINE I (GENERAL & INTERNAL MEDICINE)
6. CLINICAL AND EXPERIMENTAL MEDICINE II (NON-INTERNAL MEDICINE SPECIALTIES)
7. NEUROSCIENCE & BEHAVIOR

M1 age & gender related medicine
M2 dentistry
M3 dermatology/urogenital system
M4 ophthalmology/otolaryngology
M5 paramedicine
M6 psychiatry & neurology
M7 radiology & nuclear medicine
M8 rheumatology/orthopedics
M9 surgery

AZ ANDROLOGY
LI GERIATRICS & GERONTOLOGY
LJ GERONTOLOGY
SD OBSTETRICS & GYNECOLOGY
TQ PEDIATRICS

Source: GLÄNZEL, *Bibliometrics as a Research Field*, 2003

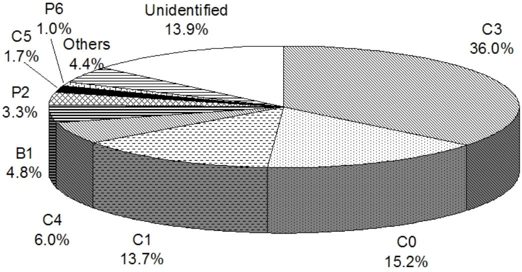
The subject categories after step II

1. **AGRICULTURE & ENVIRONMENT**
   - A1 Agricultural Science & Technology
   - A2 Plant & Soil Science & Technology
   - A3 Environmental Science & Technology
   - A4 Food & Animal Science & Technology
2. **BIOLOGY (ORGANISMIC & SUPRAORGANISMIC LEVEL)**
   - Z1 Animal Sciences
   - Z2 Aquatic Sciences
   - Z3 Microbiology
   - Z4 Plant Sciences
   - Z5 Pure & Applied Ecology
   - Z6 Veterinary Sciences
3. **BIOSCIENCES (GENERAL, CELLULAR & SUBCELLULAR BIOLOGY; GENETICS)**
   - B0 Multidisciplinary Biology
   - B1 Biochemistry/Biophysics/Molecular Biology
   - B2 Cell Biology
   - B3 Genetics & Developmental Biology
4. **BIOMEDICAL RESEARCH**
   - R1 Anatomy & Pathology
   - R2 Biomaterials & Bioengineering
   - R3 Experimental/Laboratory Medicine
   - R4 Pharmacology & Toxicology
   - R5 Physiology
5. **CLINICAL AND EXPERIMENTAL MEDICINE I (GENERAL & INTERNAL MEDICINE)**
   - I1 Cardiovascular & Respiratory Medicine
   - I2 Endocrinology & Metabolism
   - I3 General & Internal Medicine
   - I4 Hematology & Oncology
   - I5 Immunology
6. **CLINICAL AND EXPERIMENTAL MEDICINE II (NON-INTERNAL MEDICINE SPECIALTIES)**
   - M1 Age & Gender Related Medicine
   - M2 Dentistry
   - M3 Dermatology/Urogenital System
   - M4 Ophthalmology/Otolaryngology
   - M5 Paramedicine
   - M6 Psychiatry & Neurology
   - M7 Radiology & Nuclear Medicine
   - M8 Rheumatology/Orthopedics
   - M9 Surgery
7. **NEUROSCIENCE & BEHAVIOR**
   - N1 Neurosciences & Psychopharmacology
   - N2 Psychology & Behavioral Sciences
8. **CHEMISTRY**
   - C0 Multidisciplinary Chemistry
   - C1 Analytical, Inorganic & Nuclear Chemistry
   - C2 Applied Chemistry & Chemical Engineering
   - C3 Organic & Medicinal Chemistry
   - C4 Physical Chemistry
   - C5 Polymer Science
   - C6 Materials Science
9. **PHYSICS**
   - P0 Multidisciplinary Physics
   - P1 Applied Physics
   - P2 Atomic, Molecular & Chemical Physics
   - P3 Classical Physics
   - P4 Mathematical & Theoretical Physics
   - P5 Particle & Nuclear Physics
   - P6 Physics of Solids, Fluids And Plasmas
10. **GEOSCIENCES & SPACE SCIENCES**
    - G1 Astronomy & Astrophysics
    - G2 Geosciences & Technology
    - G3 Hydrology/Oceanography
    - G4 Meteorology/Atmospheric & Aerospace Science & Technology
    - G5 Mineralogy & Petrology
11. **ENGINEERING**
    - E1 Computer Science/Information Technology
    - E2 Electrical & Electronic Engineering
    - E3 Energy & Fuels
    - E4 General & Traditional Engineering
12. **MATHEMATICS**
    - H1 Applied Mathematics
    - H2 Pure Mathematics
13. **SOCIAL SCIENCES I (GENERAL, REGIONAL & COMMUNITY ISSUES)**
    - S1 Education & Information
    - S2 General, Regional & Community Issues
14. **SOCIAL SCIENCES II (ECONOMICAL & POLITICAL ISSUES)**
    - O1 Economics, Business & Management
    - O2 History, Politics & Law
15. **ARTS & HUMANITIES**
    - U1 Arts & Literature
    - U2 Language & Culture
    - U3 Philosophy & Religion

The ECOOM classification scheme

Journal classification after step III (GLÄNZEL & SCHUBERT, 2003)

[assignment of papers in ‘Angewandte Chemie – International Edition’ (1993)]



Data source: Thomson Reuters Web of Knowledge

Subject classification plays an important part in data retrieval, domain studies, comparative analysis (different communication behaviour in individual disciplines) and the determination of specialisation and publication profiles of institution and countries.

Two indicators of specialisation are popular, the *Activity Index* and its derivative, the *Relative Specialisation Index*.

Subject classification plays an important part in data retrieval, domain studies, comparative analysis (different communication behaviour in individual disciplines) and the determination of specialisation and publication profiles of institution and countries.

Two indicators of specialisation are popular, the *Activity Index* and its derivative, the *Relative Specialisation Index*.

The Activity Index (AI) is a version of the economists' Revealed Comparative Advantage (RCA).

$$AI = \frac{\text{the world share of the given country (region) in publications in the given field}}{\text{the overall world share of the given country (region) in publications}}$$

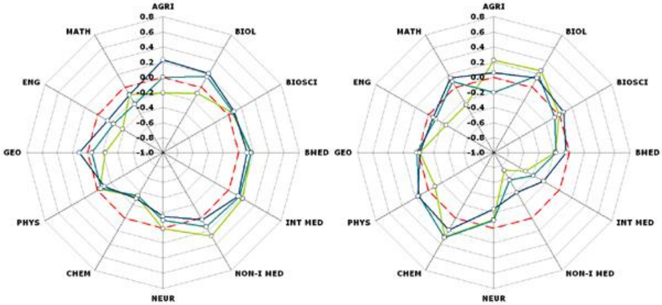
- Balassa, *The Manchester School*, 1965
- Frame, *Interciencia*, 1977
- Schubert & Braun, *Scientometrics*, 1986

The Relative Specialisation Index (RSI) is defined as

$$RSI = (AI - 1)/(AI + 1)$$

.

Example: Relative Specialisation Index of Denmark (left) and Czech Republic (right) in 1983/1993/2003



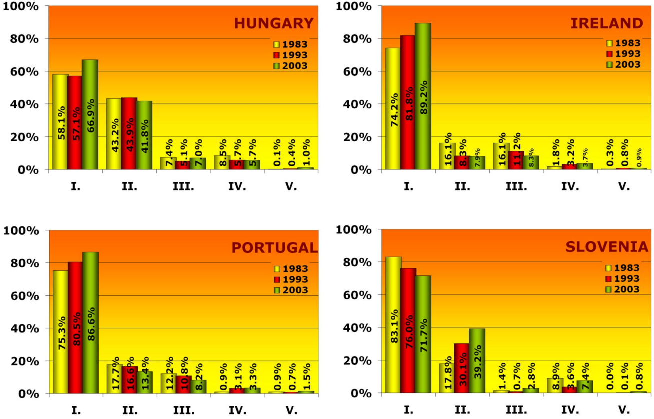
Source: GLÄNZEL ET AL., *STI Leiden*, 2004

Publication profiles by sectors

Further distinction can be made, for instance, between academic and industrial research, or, within the academic sector, between university and non-university research.

The analysis of publication activity in different sectors is a rather delicate question since industry research is less visible through publications in scientific journals than academic research and the identification of independent or associated institutions proved to be difficult.

Evolution of the distribution of SCIE-publications by sectors (1983/1993/2003)



I higher education, II public institution or government, III hospital, IV private institution, V others

Source: Glänzel & Schlemmer, *ISSI Newsletter*, 2009

Citation as a measure of reception

Basically two notions of citations have become prevalent in bibliometrics,

(1) the information science related and
(2) the sociological approach.

According to the first notion, citation is "*one important form of use of scientific information within the framework of documented science communication*".
GLÄNZEL & SCHOEPFLIN, *Information Processing & Management*, 1999

Sociology of science considers citations part of the reward system in science, atoms of peer recognition.
MERTON, *Science*, 1968.

Holmes & Oppenheim found that citations are not primarily a measure of quality, though they significantly correlate with other quality measures.
HOLMES & OPPENHEIM, *(Information Research*, 2001)

Citation as a measure of reception

Basically two notions of citations have become prevalent in bibliometrics,

(1) the information science related and
(2) the sociological approach.

According to the first notion, citation is "*one important form of use of scientific information within the framework of documented science communication*".
Glänzel & Schoepflin, *Information Processing & Management*, 1999

Sociology of science considers citations part of the reward system in science, atoms of peer recognition.
Merton, *Science*, 1968.

Holmes & Oppenheim found that citations are not primarily a measure of quality, though they significantly correlate with other quality measures.
Holmes & Oppenheim, *(Information Research*, 2001)

15 different reasons for giving citations to others' work
(GARFIELD, *Current Contents*, 1970)

1. Paying homage to pioneers
2. Giving credit for related work (homage to peer)
3. Identifying methodology, equipment, etc.
4. Providing background reading
5. Correcting one's own work
6. Correcting the work of others
7. Criticising previous work
8. Substantiating claims
9. Alerting to forthcoming work
10. Providing leads to poorly disseminated, poorly indexed, or uncited work
11. Authenticating data and classes of facts – physical constants, etc.
12. Identifying original publications in which an idea or concept was discussed
13. Identifying original publications or other work describing an eponymic concept or term
14. Disclaiming work or ideas of others (negative claim)
15. Disputing priority claims of others (negative homage)

According to the bibliometricians' view (Braun et al., *Scientometric Indicators*, 1985),

> *"if a paper receives 5 or 10 citations a year throughout several years after its publication, it is very likely that its content will become integrated into the body of knowledge of the respective subject field; if, on the other hand, no reference is made at all to the paper during 5 to 10 years after publication, it is likely that the results involved do not contribute essentially to the contemporary scientific paradigm system of the subject field in question."*

This might serve as groundwork for science-policy relevant application of citation analysis to the evaluation of research.

1. Author self-citations

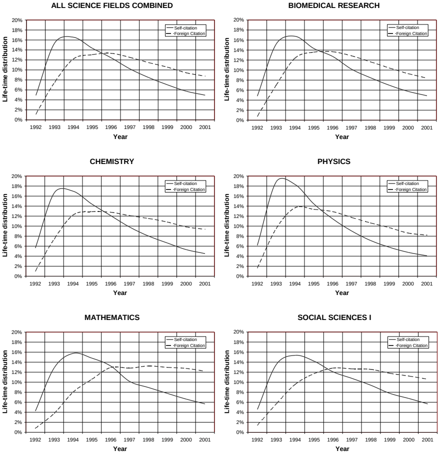
Author self-citations are an inevitable part of scholarly communication: The almost absolute lack of self-citations over a longer period is just as pathological as an always-overwhelming share.

The first phenomenon may indicate lack of originality in research, whilst the latter symptom might indicate isolation and lacking communication.

Author self-citations help to avoid redundancies and "self-plagiarism", notably in follow-up publications. A healthy share of self-citations might just reflect 'normal' use of information. However, if citations are considered part of the reward system, the interpretation might radically change.

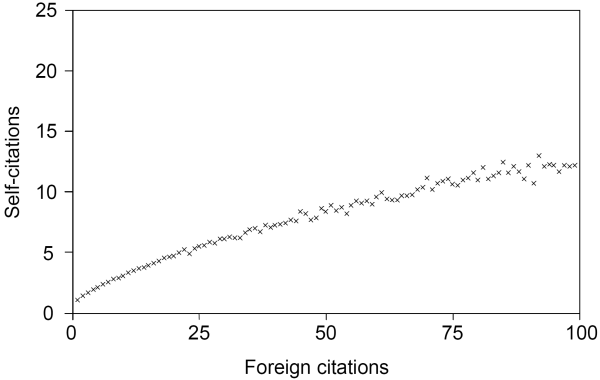
The role of self-citations

Distribution of author self-citations and foreign citations over time



GLÄNZEL ET AL., *Scientometrics*, 2004

Square-root law for author self-citations
(Plot of expected number of self-citations over number of foreign citations)



GLÄNZEL ET AL., *Scientometrics*, 2004

2. Journal self-citations

A large share of journal self-citations allows the conclusion that the journal in question is highly specialised, a low share of self-references (e.g., < 10 %) is, for example, characteristic for review journals (Schubert & Braun, *Scientometrics*, 1993).

Increasing extent of journal self-citation has been reported in the context of possible manipulation of journal Impact Factors.

- Smith, *BMJ*, 1997
- Weingart, *Scientometrics*, 2005

Factors influencing citation impact

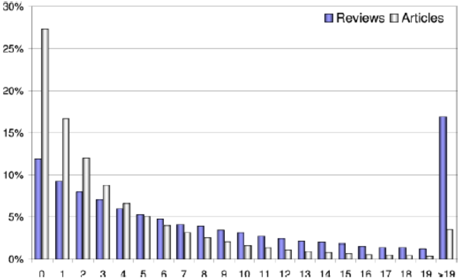
Most important measurable factors influencing citation impact

1. the subject matter and within the subject, the "level of abstraction"
2. the paper's age
3. the paper's "social status" (through the co-author(s) and the journal)
4. the document type
5. the observation period

Similarly to the publication patterns, separation of factors is almost impossible since the effect of factors often superpose.

Examples for the influence of subject and document type

Subject field	Mean citation rate
Mechanical, civil and other engineering	1.12
Mathematics	1.46
Analytical chemistry	3.00
Solid state physics	3.06
Neurosciences	4.54



Source: Glänzel & Moed, *Scientometrics*, 2002 (top); Glänzel, *ISSI Newsletter*, 2008 (bottom)

The citation windows

As a rule of thumb, the larger the citation windows the more reliable results are obtained. On the other hand, science policy is interested in the evaluation of the most recent results.

Taking into account that the reviewing and publication process and database indexing considerably contribute to the gap between research and data access of the published results, the standard solution can only be a compromise.

Citation windows ranging between three and five years have successfully been used at different levels of aggregation.

Moed, *Scientometrics*, 1996
Glänzel, *Scientometrics*, 1997 van Raan, *JASIST*, 2006

The citation windows

As a rule of thumb, the larger the citation windows the more reliable results are obtained. On the other hand, science policy is interested in the evaluation of the most recent results.

Taking into account that the reviewing and publication process and database indexing considerably contribute to the gap between research and data access of the published results, the standard solution can only be a compromise.

Citation windows ranging between three and five years have successfully been used at different levels of aggregation.

MOED, *Scientometrics*, 1996

GLÄNZEL, *Scientometrics*, 1997 VAN RAAN, *JASIST*, 2006

The citation windows

As a rule of thumb, the larger the citation windows the more reliable results are obtained. On the other hand, science policy is interested in the evaluation of the most recent results.

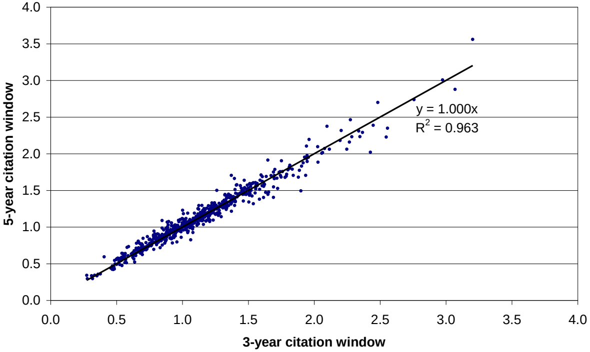
Taking into account that the reviewing and publication process and database indexing considerably contribute to the gap between research and data access of the published results, the standard solution can only be a compromise.

Citation windows ranging between three and five years have successfully been used at different levels of aggregation.

MOED, *Scientometrics*, 1996

GLÄNZEL, *Scientometrics*, 1997 van RAAN, *JASIST*, 2006

Plot of NMCR based on 5-year citation window vs. 3-year window for 676 European universities and research institutions



Source: Glänzel et al., *Arch. Immunol. Ther. Exp.*, 2008

Standard citation indicators

We use the following notations.

- c_i number of citations to paper i.
- n number of publications.
- x_i impact of journal J_i where the paper is published; $x_i = E(X|J_i)$
- f_i impact of subject F_i where the paper belongs to; $f_i = E(X|F_i)$

Literature

- Braun et al., *Scientometric Indicators*, 1985
- Braun & Glänzel, *Scientometrics*, 1990
- Moed et al., *Scientometrics*, 1995

Standard citation indicators at ISSRU/ECOOM

- **Observed citation rates**
  - Share of uncited papers: $p_0 = \sum_{i=1}^{n} \chi(c_i = 0)/n$
  - Mean Observed Citation Rate: $MOCR = \sum_{i=1}^{n} c_i/n$

  Note that these indicators are approximately normally distributed if $n \approx 40$ or greater.
- **Expected citation rates**
  - Mean Expected Observed Citation Rate: $MECR = \sum_{i=1}^{n} x_i/n$
  - Field Expected Citation Rate: $FECR = \sum_{i=1}^{n} f_i/n$
- **Relative citation rates**
  - "Publication Strategy" $MECR/FECR$
  - Normalised Mean Citation Rate: $NMCR = MOCR/FECR$
  - Relative Citation Rate: $RCR = MOCR/MECR$

☛ Here the Budapest/Leuven notation was used just for introduction. At CWTS in Leiden analogous indicators are defined: CPP, JCSm, FCSm and their ratios CPP/JCSm and CPP/FCSm. Other institutes are using similar indicators as well.
MOED ET AL., *Scientometrics*, 1995

- **Observed citation rates**
  - Share of uncited papers: $p_0 = \sum_{i=1}^{n} \chi(c_i = 0)/n$
  - Mean Observed Citation Rate: $MOCR = \sum_{i=1}^{n} c_i/n$

Note that these indicators are approximately normally distributed if $n \approx 40$ or greater.

- **Expected citation rates**
  - Mean Expected Observed Citation Rate: $MECR = \sum_{i=1}^{n} x_i/n$
  - Field Expected Citation Rate: $FECR = \sum_{i=1}^{n} f_i/n$
- **Relative citation rates**
  - "Publication Strategy" $MECR/FECR$
  - Normalised Mean Citation Rate: $NMCR = MOCR/FECR$
  - Relative Citation Rate: $RCR = MOCR/MECR$

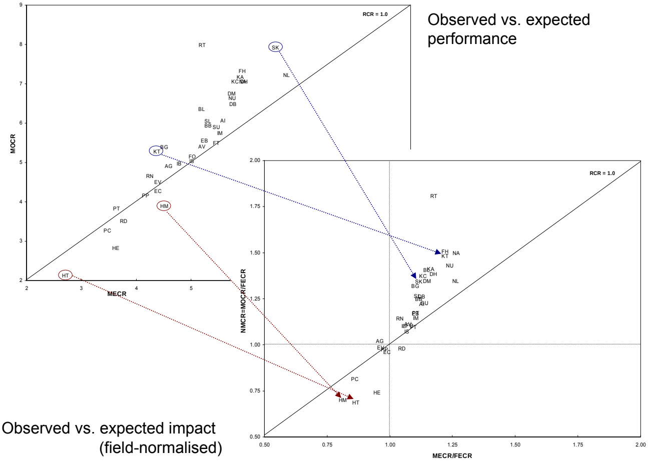
☛ Here the Budapest/Leuven notation was used just for introduction. At CWTS in Leiden analogous indicators are defined: CPP, JCSm, FCSm and their ratios CPP/JCSm and CPP/FCSm. Other institutes are using similar indicators as well.

Moed et al., *Scientometrics*, 1995

NB. The publication year(s) and citation windows underlying all indicators used for the comparison must be identical. Results obtained from comparisons between different publication periods and/or different citation windows must be necessarily invalid.

NB. Indicators should be based on sufficiently large publication sets with ideally $n \geq 30$. In this case, shares and averages are approximately normally distributed even if the underlying citation rates have skewed and non-negative integer-valued distributions. If sets are much smaller, citation indicators just remain insignificant numbers the policy application of which were an unreliable and irresponsible endeavour.

Relative vs.‘absolute’ indicators



Observed vs. expected performance

Observed vs. expected impact
(field-normalised)

Data source: Thomson Reuters Web of Knowledge

Improvement of bibliometric indicators for the evaluation of research

Publication-activity and citation-impact statistics are influenced by various factors (subject, age, time, status, communication form, etc.) and need therefore to be normalised.

Two paradigmatic approaches are under discussion.

- *A posteriori normalisation*: mathematical manipulation of (standard) indicators
- *A priori normalisation*: fractional counting prior to indicator calculation

Improvement of bibliometric indicators

Improvement of bibliometric indicators for the evaluation of research

Publication-activity and citation-impact statistics are influenced by various factors (subject, age, time, status, communication form, etc.) and need therefore to be normalised.

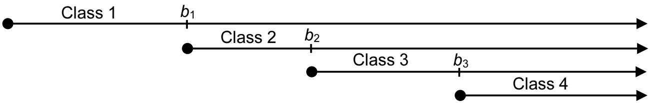
Two paradigmatic approaches are under discussion.

- *A posteriori normalisation*: mathematical manipulation of (standard) indicators
- *A priori normalisation*: fractional counting prior to indicator calculation

An a posteriori normalisation for bibliometric inicators (Example)

Characteristic scores are originated from iteratively truncating samples at their mean value and recalculating the mean of the truncated sample until the procedure is stopped or no new scores are obtained.

Visualisation of characteristic scores and scales for four classes



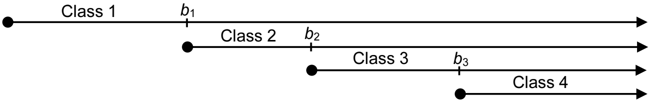
Glänzel, *Journal of Informetrics*, 2007

Advantage: Self-adjusting (no arbitrary thresholds) and no tie problems.

An a posteriori normalisation for bibliometric inicators (Example)

Characteristic scores are originated from iteratively truncating samples at their mean value and recalculating the mean of the truncated sample until the procedure is stopped or no new scores are obtained.

Visualisation of characteristic scores and scales for four classes



GLÄNZEL, *Journal of Informetrics*, 2007

Advantage: Self-adjusting (no arbitrary thresholds) and no tie problems.

$(\beta_2 - \beta_1)$ can be considered a proxy for the scale parameter of the underlying distribution, where β_i are the empirical values of the corresponding scores b_i.
The transformation suggested by SCHUBERT ET AL. (1989) can then be applied for scale normalisation.

$$u^* = \frac{x}{\beta_2 - \beta_1},$$

where x represents the actual citation statistic.

This procedure is repeated until no new scores or a predefined number of scores k are obtained.

Now we define the following classes.

$[b_0, b_1)$ is the class of ‘poorly cited’ papers,
$[b_1, b_2)$ contains ‘fairly cited’ papers,
$[b_2, b_3)$ contains ‘remarkably cited’ papers and
$[b_3, \infty)$ is the class of ‘outstandingly cited’ papers.

The values $k = 2$ and $k = 3$ are often used to identify *highly cited* papers.

This procedure is repeated until no new scores or a predefined number of scores k are obtained.

Now we define the following classes.

$[b_0, b_1)$ is the class of 'poorly cited' papers,

$[b_1, b_2)$ contains 'fairly cited' papers,

$[b_2, b_3)$ contains 'remarkably cited' papers and

$[b_3, \infty)$ is the class of 'outstandingly cited' papers.

The values $k = 2$ and $k = 3$ are often used to identify *highly cited* papers.

This procedure is repeated until no new scores or a predefined number of scores k are obtained.

Now we define the following classes.

$[b_0, b_1)$ is the class of 'poorly cited' papers,
$[b_1, b_2)$ contains 'fairly cited' papers,
$[b_2, b_3)$ contains 'remarkably cited' papers and
$[b_3, \infty)$ is the class of 'outstandingly cited' papers.

The values $k = 2$ and $k = 3$ are often used to identify *highly cited* papers.

Threshold for "highly" cited papers (β_3) based on *Characteristic Scores and Scales* and their (a posteriori) normalised versions (β_3^*) according to GLÄNZEL (2011)
[B1: biochemistry/biophysics/molecular biology; H1: applied mathematics]

	B1		H1	
	1980-2000	2006-2008	1980-2000	2006-2008
β_3	196.55	22.69	49.66	4.47
β_3^*	3.51	3.53	3.28	3.46

Data source: Thomson Reuters Web of Knowledge

Threshold for "highly" cited papers (β_3) based on *Characteristic Scores and Scales* and their (a posteriori) normalised versions (β_3^*) according to GLÄNZEL (2011)
[B1: biochemistry/biophysics/molecular biology; H1: applied mathematics]

	B1		H1	
	1980-2000	2006-2008	1980-2000	2006-2008
β_3	196.55	22.69	49.66	4.47
β_3^*	3.51	3.53	3.28	3.46

Data source: Thomson Reuters Web of Knowledge

A posteriori normalisation – pros and cons

- Works with “standard” measures available from several sources (TR, Elsevier, SCIMago, etc.)
- Easy to calculate and robust
- Can be applied to longitudinal studies and to the analysis of time series
- The same item takes different values in the case of multiple subject assignment

A priori normalisation – pros and cons

- Each individual item takes only one value (independently of assignment)
- Less sensitive to arbitrary subject assignments
- Requires rather complicated calculations or algorithms
- Works only for cited items

A posteriori normalisation – pros and cons

- Works with “standard” measures available from several sources (TR, Elsevier, SCIMago, etc.)
- Easy to calculate and robust
- Can be applied to longitudinal studies and to the analysis of time series
- The same item takes different values in the case of multiple subject assignment

A priori normalisation – pros and cons

- Each individual item takes only one value (independently of assignment)
- Less sensitive to arbitrary subject assignments
- Requires rather complicated calculations or algorithms
- Works only for cited items

The application of CSS classes is shown using the example of papers published in the subfield “cardiovascular & respiratory medicine” in the publication years 2007 and 2009. Citations are counted for three-year citation windows.

The scores for the world total are used as the reference standard. The procedure was stopped at $k = 3$.

The countries’ citation rates are assigned to these classes and the national distributions are then compared with the “standard” citation distribution according to the world total.

If a country’s distribution is a true “mirror” of the world standard, its distribution over classes is expected to coincide with that of the world.

Characteristic scores of publications in 2007 and 2009 for 20 selected subfields according to the Leuven-Budapest scheme

Subfield*	2007 (5-year citation window)			2009 (3-year citation window)		
	b_1	b_2	b_3	b_1	b_2	b_3
A2	6.43	13.80	21.97	2.68	6.01	10.68
B1	16.75	39.24	79.61	8.21	19.96	38.24
B2	23.05	58.33	116.72	11.34	28.96	56.28
C1	9.37	22.04	40.48	5.13	12.37	21.68
C3	11.22	24.68	42.04	5.84	12.24	20.83
C6	8.21	23.67	51.24	4.56	12.71	26.50
E1	5.04	14.75	29.83	2.37	6.64	12.60
E2	4.71	11.90	21.97	2.27	6.15	11.54
E3	6.57	17.82	34.00	4.19	11.19	21.10
G1	15.55	38.35	74.51	8.75	20.82	39.17
H1	5.21	14.36	29.83	2.41	6.66	12.88
I1	13.52	34.87	69.24	6.01	15.92	29.58
I5	16.24	41.52	84.74	7.96	19.26	39.49
M6	11.50	28.31	51.81	5.27	13.51	24.88
N1	15.28	35.38	64.73	7.18	16.92	29.77
P4	7.25	17.71	32.75	3.09	8.12	15.13
P6	7.27	20.05	43.89	4.30	12.15	26.54
R2	10.60	23.99	42.54	4.82	10.64	18.37
R4	11.42	26.19	48.62	5.49	12.65	22.50
Z3	12.80	29.48	54.96	6.36	15.25	28.88

Source: Thomson Reuters – Web of Knowledge

Legend: A2: plant & soil science & technology; B1: biochemistry/biophysics/molecular biology; B2: cell biology; C1: analytical, inorganic & nuclear chemistry; C3: organic & medicinal chemistry; C6: materials science; E1: computer science/information technology; E2: electrical & electronic engineering; E3: energy & fuels; G1: astronomy & astrophysics; H1: applied mathematics; I1: cardiovascular & respiratory medicine; I5: immunology; M6: psychiatry & neurology; N1: neurosciences & psychopharmacology; P4: mathematical & theoretical physics; P6: physics of solids; R2: biomaterials & bioengineering; R4: pharmacology & toxicology; Z3: microbiology

CSS-class shares of publications in 2007 and 2009 for 20 selected subfields according to the Leuven-Budapest scheme

Subfield*	2007 (5-year citation window)				2009 (3-year citation window)			
	Class 1	Class 2	Class 3	Class 4	Class 1	Class 2	Class 3	Class 4
A2	65.2%	22.6%	8.1%	4.2%	63.3%	26.0%	7.1%	3.6%
B1	69.4%	22.5%	6.0%	2.1%	70.6%	21.0%	6.3%	2.2%
B2	72.0%	20.2%	5.6%	2.2%	71.6%	20.1%	5.8%	2.4%
C1	68.2%	22.5%	6.6%	2.7%	69.2%	21.3%	6.4%	3.0%
C3	67.4%	22.2%	7.5%	3.0%	63.6%	24.9%	7.7%	3.9%
C6	73.5%	19.5%	5.3%	1.8%	71.6%	20.5%	5.8%	2.1%
E1	73.7%	18.8%	5.5%	2.0%	71.4%	19.9%	6.2%	2.4%
E2	68.2%	21.7%	7.0%	3.1%	70.8%	20.9%	5.7%	2.5%
E3	70.7%	20.2%	6.3%	2.9%	70.9%	20.6%	6.1%	2.4%
G1	70.1%	21.4%	6.3%	2.2%	68.1%	22.4%	7.2%	2.4%
H1	72.3%	20.3%	5.4%	1.9%	71.0%	20.4%	6.2%	2.4%
I1	70.2%	21.3%	6.2%	2.3%	71.2%	20.0%	6.1%	2.7%
I5	71.9%	20.4%	5.4%	2.2%	68.7%	22.8%	6.1%	2.3%
M6	68.9%	21.6%	6.5%	3.0%	69.9%	20.9%	6.3%	2.9%
N1	69.1%	21.7%	6.4%	2.8%	69.1%	21.1%	6.8%	3.0%
P4	69.6%	21.2%	6.7%	2.4%	71.2%	20.8%	5.7%	2.3%
P6	72.4%	20.7%	5.3%	1.7%	72.8%	20.4%	5.2%	1.6%
R2	72.4%	20.7%	5.3%	1.7%	64.7%	23.7%	7.8%	3.8%
R4	68.4%	22.5%	6.4%	2.7%	67.3%	22.5%	7.1%	3.0%
Z3	68.2%	22.3%	6.8%	2.6%	69.3%	22.1%	6.2%	2.5%

Source: Thomson Reuters – Web of Knowledge

Legend: A2: plant & soil science & technology; B1: biochemistry/biophysics/molecular biology; B2: cell biology; C1: analytical, inorganic & nuclear chemistry; C3: organic & medicinal chemistry; C6: materials science; E1: computer science/information technology; E2: electrical & electronic engineering; E3: energy & fuels; G1: astronomy & astrophysics; H1: applied mathematics; I1: cardiovascular & respiratory medicine; I5: immunology; M6: psychiatry & neurology; N1: neurosciences & psychopharmacology; P4: mathematical & theoretical physics; P6: physics of solids; R2: biomaterials & bioengineering; R4: pharmacology & toxicology; Z3: microbiology

CSS in cross-national comparison

National shares of publications in the reference CSS classes in 2007 and 2009 for subfield I1 (cardiovascular & respiratory medicine) in alphabetic order

Country	2007 (5-year citation window)				2009 (3-year citation window)			
	Class 1	Class 2	Class 3	Class 4	Class 1	Class 2	Class 3	Class 4
BEL	61.5%	24.1%	9.8%	4.6%	61.2%	24.0%	9.3%	5.5%
BRA	73.5%	19.8%	4.7%	2.0%	87.0%	8.8%	3.1%	1.2%
CAN	61.8%	25.3%	9.2%	3.7%	59.4%	26.6%	8.7%	5.3%
CHE	60.8%	25.2%	10.7%	3.3%	61.7%	23.0%	9.5%	5.8%
CHN	68.7%	24.4%	5.6%	1.3%	72.8%	21.0%	4.8%	1.4%
DEU	62.5%	24.5%	8.9%	4.1%	63.0%	23.7%	8.6%	4.7%
ESP	73.8%	17.8%	5.2%	3.2%	72.9%	17.0%	6.7%	3.4%
FRA	71.3%	17.8%	7.5%	3.4%	66.4%	20.9%	7.9%	4.8%
GBR	61.0%	26.2%	8.5%	4.3%	62.1%	24.0%	8.9%	5.0%
GRC	74.8%	19.4%	4.2%	1.6%	75.6%	17.8%	4.6%	2.0%
ITA	70.8%	20.0%	6.3%	3.0%	66.9%	21.7%	7.3%	4.0%
JPN	73.2%	19.9%	5.3%	1.5%	71.6%	21.3%	5.2%	1.8%
KOR	74.2%	18.2%	5.2%	2.3%	65.4%	25.1%	7.6%	1.9%
NLD	56.4%	28.9%	9.9%	4.8%	57.7%	28.0%	9.9%	4.4%
POL	71.4%	20.6%	4.2%	3.8%	82.4%	10.3%	3.8%	3.5%
SWE	59.1%	27.7%	10.0%	3.2%	60.2%	24.3%	9.9%	5.6%
TUR	92.7%	6.3%	0.9%	0.0%	93.8%	4.7%	1.1%	0.4%
TWN	78.6%	17.4%	2.6%	1.4%	76.4%	16.8%	5.0%	1.7%
USA	61.0%	26.4%	9.0%	3.6%	61.8%	25.0%	8.9%	4.3%
Total	70.2%	21.3%	6.2%	2.3%	71.2%	20.0%	6.1%	2.7%

Source: Thomson Reuters – Web of Knowledge

These shares should be calculated for each subject separately.

However, shares might be combined over subjects but *not* over classes.

- ☛ One should keep in mind that the results for large subject fields and for all fields combined calculated in this way might be affected by biases caused by deviating publication profiles of different units.

One precondition for the application of CSS to combined subjects is the *unique assignment of papers* to performance classes.

Example:

A paper is assigned to two subjects, here denoted by S1 and S2.

The paper might then be assigned, e.g., to Class 3 in subject S1 and to Class 4 in S2 because its citation rate does not exceed b_3 in S1 but it is greater than the corresponding threshold b_3 in S2.

A direct combination can, therefore, not provide an acceptable solution.

These shares should be calculated for each subject separately.

However, shares might be combined over subjects but *not* over classes.

- ☛ One should keep in mind that the results for large subject fields and for all fields combined calculated in this way might be affected by biases caused by deviating publication profiles of different units.

One precondition for the application of CSS to combined subjects is the *unique assignment of papers* to performance classes.

Example:

A paper is assigned to two subjects, here denoted by S1 and S2.

The paper might then be assigned, e.g., to Class 3 in subject S1 and to Class 4 in S2 because its citation rate does not exceed b_3 in S1 but it is greater than the corresponding threshold b_3 in S2.

A direct combination can, therefore, not provide an acceptable solution.

These shares should be calculated for each subject separately.

However, shares might be combined over subjects but *not* over classes.

☛ One should keep in mind that the results for large subject fields and for all fields combined calculated in this way might be affected by biases caused by deviating publication profiles of different units.

One precondition for the application of CSS to combined subjects is the *unique assignment of papers* to performance classes.

Example:

A paper is assigned to two subjects, here denoted by S1 and S2.

The paper might then be assigned, e.g., to Class 3 in subject S1 and to Class 4 in S2 because its citation rate does not exceed b_3 in S1 but it is greater than the corresponding threshold b_3 in S2.

A direct combination can, therefore, not provide an acceptable solution.

Solution:

- A proper subject-based fractionation must be applied such that each publication is gauged against only one individual threshold value.
- Fractionation by subjects and thus the calculation of proper weights for the corresponding individual subject-expected citation rates is necessary.
- This results in an 'implicit' classification without calculating any common thresholds b_k.
- The first step is identical with the procedure of calculating subfield-expected citation rates (Glänzel et al., 2009).
  - A fractionation is applied when the citation means of subfields is determined (on the basis of the respective number of subfields to which a publication is assigned). Both publications and citations are fractionated.

Solution (contd.):

- Individual expectations are calculated for each paper, which is the mean value of the fractionated subfield standards.
- In the following step of the iteration, all papers, that have received less citations than their individual expectation, are removed.
- The above procedure is repeated on the remaining set. This is done three times in total to obtain the individual characteristics scores b_k^* $(k = 1, 2, 3)$ for each paper.
- At the end, all papers can now uniquely be assigned to one of the four classes.

☛ *If the underlying paper set comprises only publications from one single subfield, the individual thresholds are identical with the common characteristic scores of the subfield.*

CSS classes in all fields combined in 2007 and 2009

Class	Share (in %)	
	2007 (5-year cites)	2009 (3-year cites)
1	69.8	69.7
2	21.5	21.4
3	6.3	6.4
4	2.4	2.5

Distribution of publications over major fields in 2007 and 2009 according to the Leuven-Budapest scheme

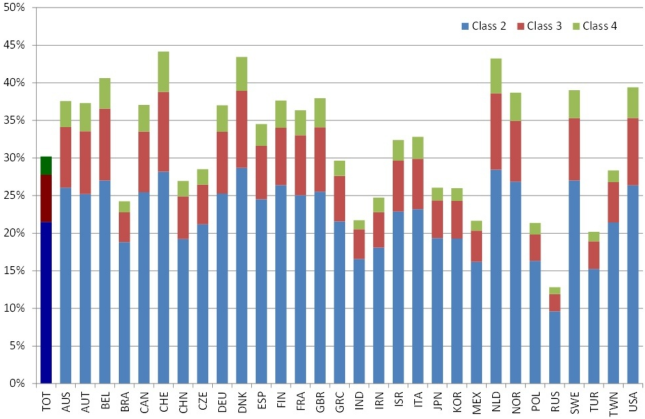
Field	2007 (5-year citations) WoS	2007 (5-year citations) Class 4	2009 (3-year citations) WoS	2009 (3-year citations) Class 4
A	7.00%	8.20%	7.50%	8.50%
B	10.10%	10.10%	9.30%	9.30%
C	20.20%	19.80%	20.00%	21.70%
E	11.20%	8.50%	11.80%	9.10%
G	5.70%	6.90%	5.80%	6.70%
H	4.50%	4.10%	5.00%	4.10%
I	12.20%	11.00%	12.00%	10.50%
M	18.40%	18.30%	18.70%	18.30%
N	5.70%	6.80%	5.60%	6.70%
P	15.00%	13.60%	14.30%	13.20%
R	7.20%	6.40%	7.20%	6.80%
Z	10.30%	9.60%	10.00%	9.80%

Source: Thomson Reuters – Web of Knowledge

Legend: A: Agriculture & environment; B: Biosciences (General, cellular & subcellular biology; genetics); C: Chemistry; E: Engineering; G: Geosciences & space sciences; H: Mathematics I: Clinical and experimental medicine I (General & internal medicine); M: Clinical and experimental medicine II (Non-internal medicine specialties); N: Neuroscience & behavior; P: Physics; R: Biomedical research; Z: Biology (Organismic & supraorganismic level)

CSS in all fields combined

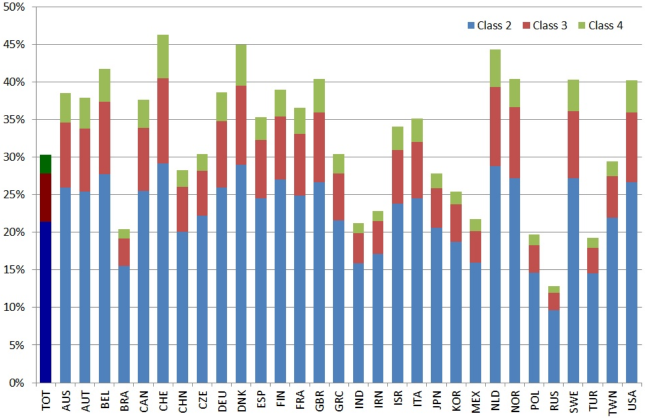
The world standard (left-most column) and national shares of publications in the upper three CSS classes in all fields combined in 2007 (5-year citation window)



Source: Thomson Reuters – Web of Knowledge

CSS in all fields combined

The world standard (left-most column) and national shares of publications in the upper three CSS classes in all fields combined in 2009 (3-year citation window)

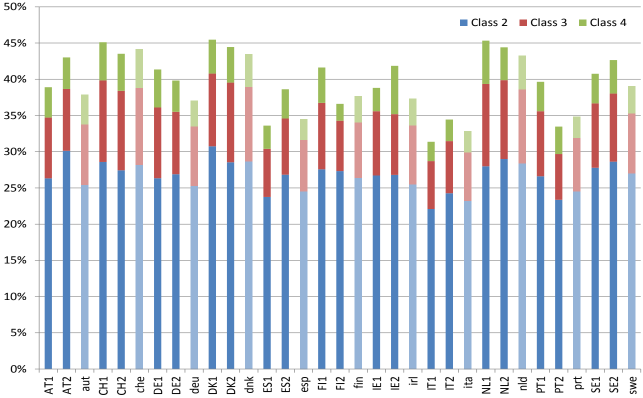


Source: Thomson Reuters – Web of Knowledge

CSS performance classes in institutional research assessment

- For the cross-institutional comparison of class profiles two universities each from eleven European countries were selected.
- Although the universities' profiles mostly mirror the national patterns, also distinctly more favourable situation than in the national standards could be found.
- This is contrasted by a less favourable situation for the several South-European universities IT1, PT2, ES1 as well as for FI2 and CH2.
- The selected Danish and Dutch universities represent an high standard.
- DK1 and PT1 are technical universities while SE1 stands for a medical university.
    - This again substantiates the subject-independence of the method.

Shares of publications of selected universities and countries in the upper three CSS classes in all fields combined in 2007 (5-year citation window)



Source: Thomson Reuters – Web of Knowledge

Acknowledgement

Parts of this lecture are based on a joint presentation with Sybille Hinze (IFQ, Berlin) prepared for the *European Summer School for Scientometrics* (esss).

Thank you very much for your attention.
Vielen Dank für Ihre Aufmerksamkeit!
Hartelijk dank voor uw aandacht!
¡Muchísimas gracias por su atención!
Köszönöm szépen a figyelmüket!
Molte grazie per la vostra attenzione.
Muito obrigado pela vossa atenção.