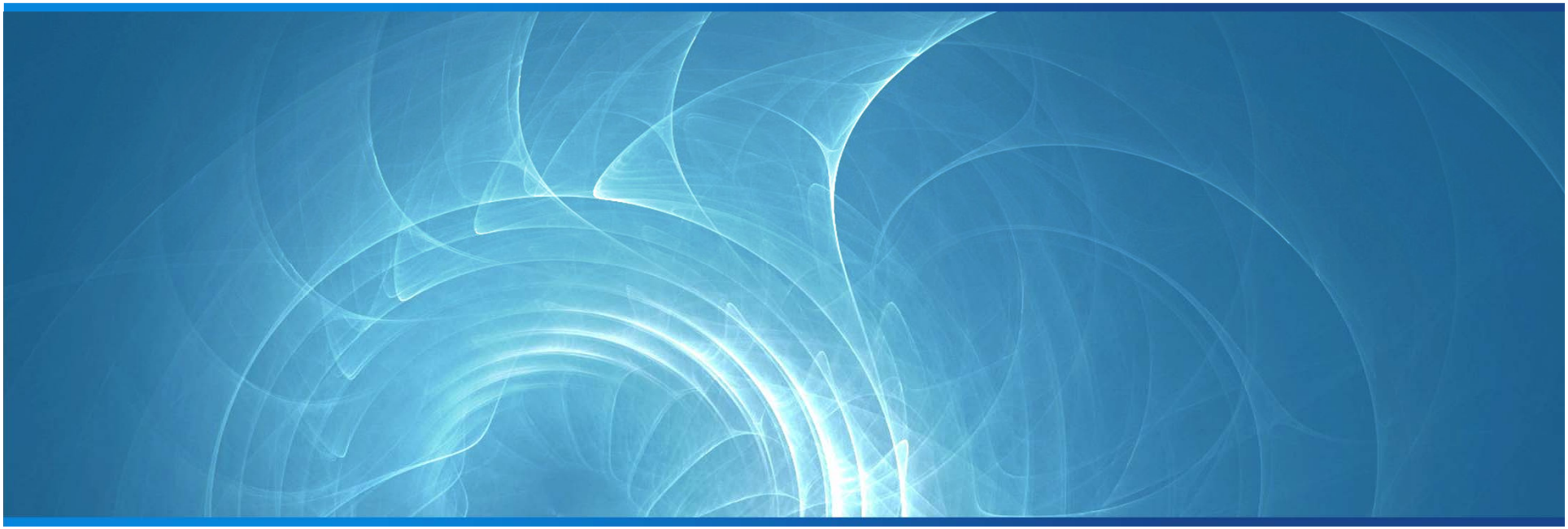
Science-Metrix

Practical Applications of Bibliometrics:
What Makes Sense in Different Contexts?



Evaluation 2010 | AEA Conference

Panel Session 231 | Analysis and Evaluation of Research Portfolios Using Quantitative Science Metrics: Practice

Sponsored by R, T&D TIG | Texas D | Thursday, November 11, 2010 | 9:15 AM to 10:45 AM

Outline

- **Overview:** **Bibliometrics in Research Evaluation**
  - Bibliometrics for Research Evaluation
  - Integration of Bibliometric Methods within Evaluation Frameworks
  - Different Contexts and Uses
- **Case One:** **Evaluating the Effects of Funding**
  - Alberta Ingenuity Funds
- **Case Two:** **Evaluating the Research Performance of Science-Based Governmental Departments**
  - Study 1: Natural Resources Canada (Minerals and Metals Sector)
  - Study 2: Environment Canada (Environmental Research)
- **Conclusion**

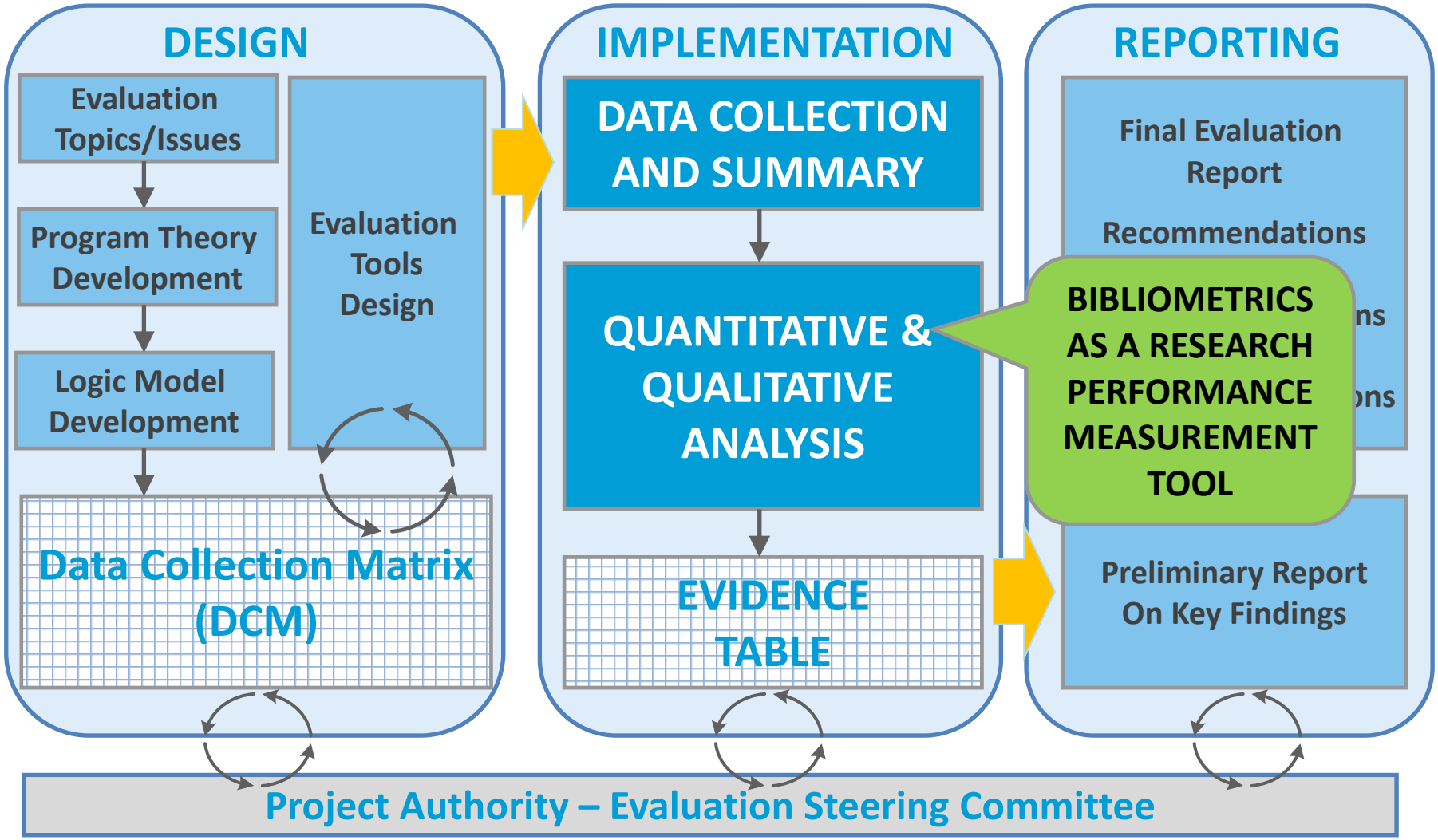
Bibliometrics in Research Evaluation

- Quantitative methods for research portfolio and program evaluation are used by "evaluators, program managers and policy-makers to assess the impacts of their research portfolios and programs, and make informed decisions on future research investments" (Wagner & Jordan, 2009).
- Federal agencies have been under increasing pressure to develop quantitative indicators for the assessment of their programs (Hicks et al., 2004).
- In academic institutions, the increased use of quantitative methods of assessment reflects pressures to provide measures of productivity, quality/impact (Holden et al., 2005).
- Bibliometrics provides information to managers about the size, composition, diversity and other essential features of an existing research portfolio (Srivastava et al., 2007).
- Although "the bibliometric toolbox" of appropriate standard indicators has been available from the beginning, the "correct application of these tools to different levels of aggregation is still a challenge" (Glänzel et al., 2008).
- Bibliometrics in the evaluation of individuals, groups, and institutions becomes more appropriate the more it is: formal; open; supplemented with expert and background knowledge; carried out in a clear policy context; and how they were operationalized and weighted; and enlightening rather than formulaic Moed (2009).

Bibliometrics: Advantages/Limitations

Advantages	Limitations
Inexpensive and economical (time-saving, exhaustive, relatively simple to implement)	Significant disciplinary differences / Variability in types of outputs and citations
Ideal for assessing research production	Difficult to compare research productivity Limited indication of quality and outcomes
Updatable and useful for ongoing monitoring	Databases represent only a proportion of the global literature (Single type of publication)
Can allow for inter-temporal or international comparisons	Bias of databases in favour of english literature
Scalable, from micro (researcher, institute) to macro (country, world) level	Scale effect (non linearity in the underlying data)
Independent, objective and reproducible	Limited understanding / Vulnerable to manipulation
Powerful when linked with other data/indicators	Better when multiple indicators are used

Integration of Bibliometric Methods within Evaluation Frameworks



Different Contexts and Applications

Application		Context: Funding Agency	Context: Science-based Governmental Organization
	Effects of funding	Case one	
	Assessment of selection process (including peer-review)	Case one	
	Research Performance Evaluation		Case two

Case One: Effect of Funding

Application		Context: Funding Agency	Context: Science-based Governmental Organization
	Effects of funding	**Case one**	**Alberta Ingenuity Fund**
	Assessment of selection process (including peer-review)	**Case one**	
	Research Performance Evaluation		Case two

Case One: Context

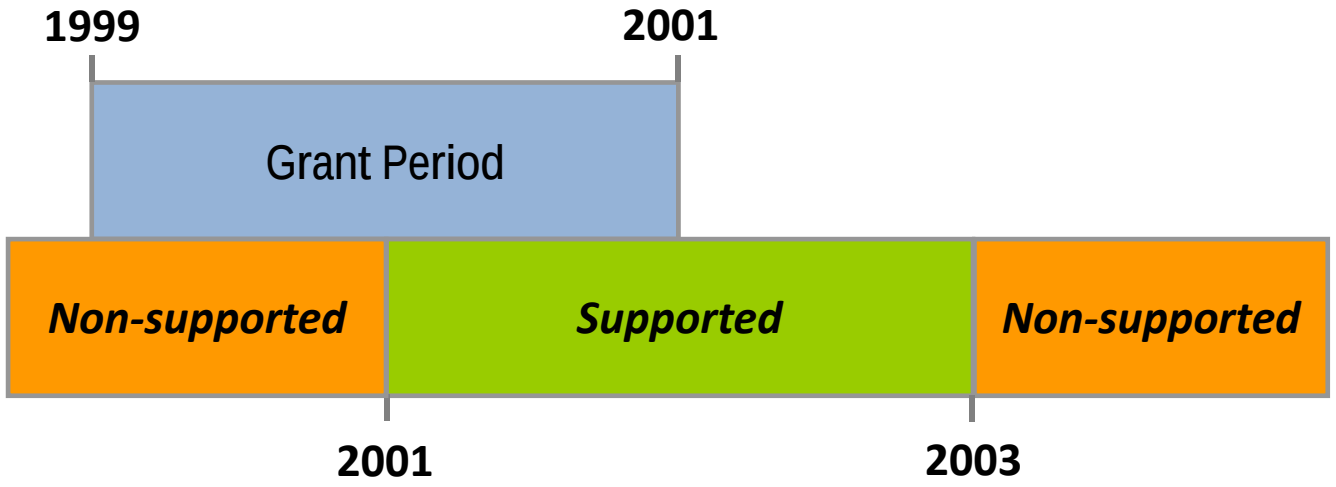
- **Alberta Ingenuity Fund (AIF) was established in 2000:**
  - It aims to nurture knowledge discovery and promote its application in areas (mostly in the NSE and the HS) with lasting impacts for the benefit of Albertans and the world.
  - It achieves its mission by supporting high calibre research performed by outstanding researchers (various programs) & students (Graduate Student Scholarships [GSS] program).
  - It evaluates applications using a rigorous peer-review system.
  - It is accountable to Albertans and reports to the Government of Alberta.

  - **The first dimension of this study pertains to assessing the grant adjudication process**

  - **The second dimension of this study pertains to assessing the effects of AIF funding**

Case One: Methods

- **Selection of a bibliographic database: Scopus (Elsevier)**
- **Three datasets:**
  - AIF-funded applicants (core dataset) & unsuccessful applicants (benchmark dataset): build publication portfolios for a random sample of 300 students in each group
  - See: Bibliometric Analysis of Individuals Supported by Alberta Ingenuity Fund (http://www.albertaingenuity.ca/files/Science_Metrix_Bibliometrics_June8_09.pdf)
  - Alberta, Canada, and the world (benchmark dataset): papers retrieved based on author addresses.
- **Supported & unsupported papers for AIF-funded applicants & unsuccessful applicants:**



Case One: Methods

- **Bibliometric indicators:**
  - **Number of papers:** Whole counting of the number of scientific papers written by authors associated with a funding organization (i.e., AIF) based on author names, or by authors associated with a region (i.e., Alberta, Canada) based on author addresses.
  - **Average of Relative Citations (ARC):** A direct measure of scientific impact based on paper citation counts (2-year citation window). Normalized by subfields of science (NSF classification) and year.
  - **Average of Relative Impact Factors (ARIF):** A proxy for the 'quality' of the journal in which papers are published. It is based on symmetric impact factors of journals. Normalized by subfields of science (NSF classification) and year.
- **Statistical analyses:**
  - Because data on scientific output (i.e., number of papers, ARC, ARIF) is not normally distributed, non-parametric tests were used.
  - The Mann-Whitney U test was used to compare pairs of independent samples and the Wilcoxon signed rank test was used to compare two related samples.

Case One: Results

Comparison of the scientific impact (based on ARC and ARIF) of the papers produced by AIF-supported students and by unsuccessful GSS applicants (prior to application only) with those of papers from researchers from Alberta, Canada and the world, 1996–2008

Group	ARC*	*p* -value	ARIF**	*p* -value
AIF-supported students	1.5	–	1.3	–
Alberta without AIF-supported students	1.2	< 0.001	1.1	< 0.001
Canada without AIF-supported students	1.2	< 0.001	1.1	< 0.001
World without AIF-supported students	1.0	–	1.0	–

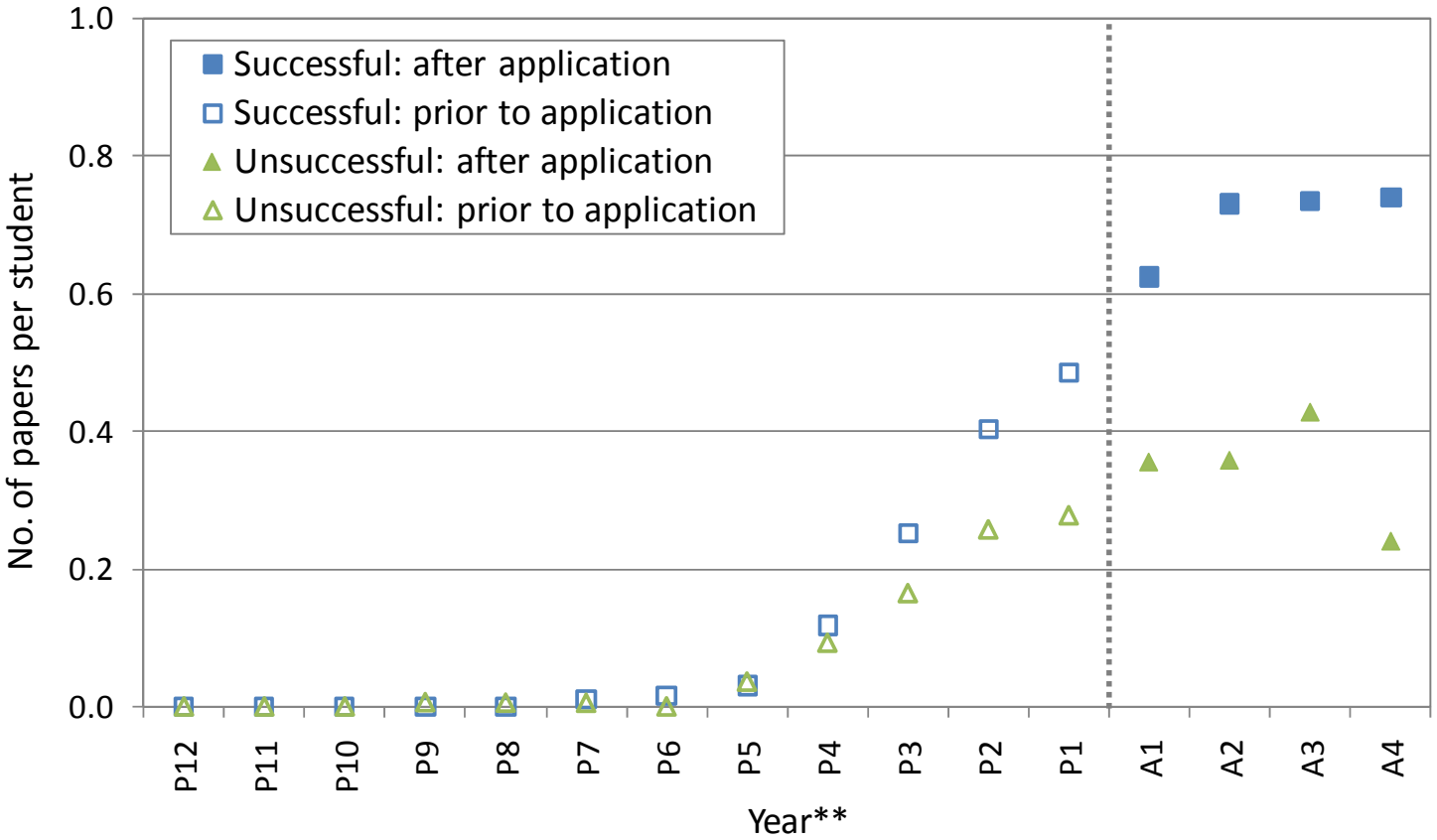
Group	ARC*	*p* -value	ARIF**	*p* -value
Unsuccessful GSS applicants	1.0	–	1.0	–
Alberta without unsuccessful GSS applicants	1.2	> 0.05	1.1	> 0.05
Canada without unsuccessful GSS applicants	1.2	> 0.05	1.1	< 0.05
World without unsuccessful GSS applicants	1.0	–	1.0	–

Note: * Relative citation counts are unreliable at this analytical level for the most recent years (2007 and 2008); therefore, papers published in those years were not included in computation of the ARC indicator. ** The impact factors of journals prior to 1998 cannot be calculated in Scopus; therefore, papers published in 1996 and 1997 were not included in computation of the ARIF indicator.

Source: Calculated by Science-Metrix from Scopus

Case One: Results

Trends in the number of published papers per AIF-supported student (N = 194) and per unsuccessful GSS applicant (N = 194)*, up to 12 years prior to application and up to 4 years after application, 1996–2008



Note: * The sample size for successful applicants was reduced from 300 to 194 in order to compare the periods before and after application. ** Letters refer to one of two periods: P = prior to receiving support or prior to application, A = after application. Sample size (i.e., no. of students) decreases across the period after application. Source: Calculated by Science-Metrix from Scopus

Case One: Results

Comparison of the scientific impact (based on ARC and ARIF) of the papers produced by AIF-supported students with those of papers by unsuccessful GSS applicants, 1996–2008

Group	Successful applicants	Unsuccessful applicants	*p* -value
ARC*			
Before application papers	1.1	1.1	> 0.05
After application papers	1.3	1.0	> 0.05
All papers (before and after application)	1.4	1.0	< 0.001
ARIF**			
Before application papers	1.2	1.1	> 0.05
After application papers	1.2	1.2	> 0.05
All papers (before and after application)	1.3	1.1	< 0.001

Note: The sample size for both successful and unsuccessful applicants was reduced from 300 to 194 to enable comparisons between the periods before and after application. * Relative citation counts are unreliable at this analytical level for the most recent years (2007 and 2008); therefore, papers published in those years were not included in computation of the ARC indicator. ** The impact factors of journals prior to 1998 cannot be calculated in Scopus; therefore, papers published in 1996 and 1997 were not included in computation of the ARIF indicator.

Source: Calculated by Science-Metrix from Scopus

Case One: Pros & Cons of Bibliometrics

Strengths	Weaknesses
Effective to assess whether the peer-review process successfully led to the selection of outstanding researchers (or students) in terms of scientific impact & production	Care must be taken when applied to the SSH due to: • The more local orientation of research in the SSH (need to cover different languages) • Other types of outputs should be considered
Effective to measure the effect of funding on the scientific performance of researchers: • One of the most potent application of bibliometrics to the SSH	• Other sources of funding might blur the signal • Uncertainty regarding which papers are supported by the funding source of interest • Important delays before an effect can be detected due to the combined effect of publication time-lags following reception of funding and citation time-lags • In the SSH, the method assumes that the funding did not lead to a shift in the type of outputs

Case Two: Research Performance Evaluation

Application	Context: Funding Agency	Context: Science-based Governmental Organization
Effects of funding	Case one	**A. Natural Resources Canada**
Assessment of selection process (including peer-review)	Case one	**B. Environment Canada**
Research Performance Evaluation		**Case two**

Case Two (A): Context

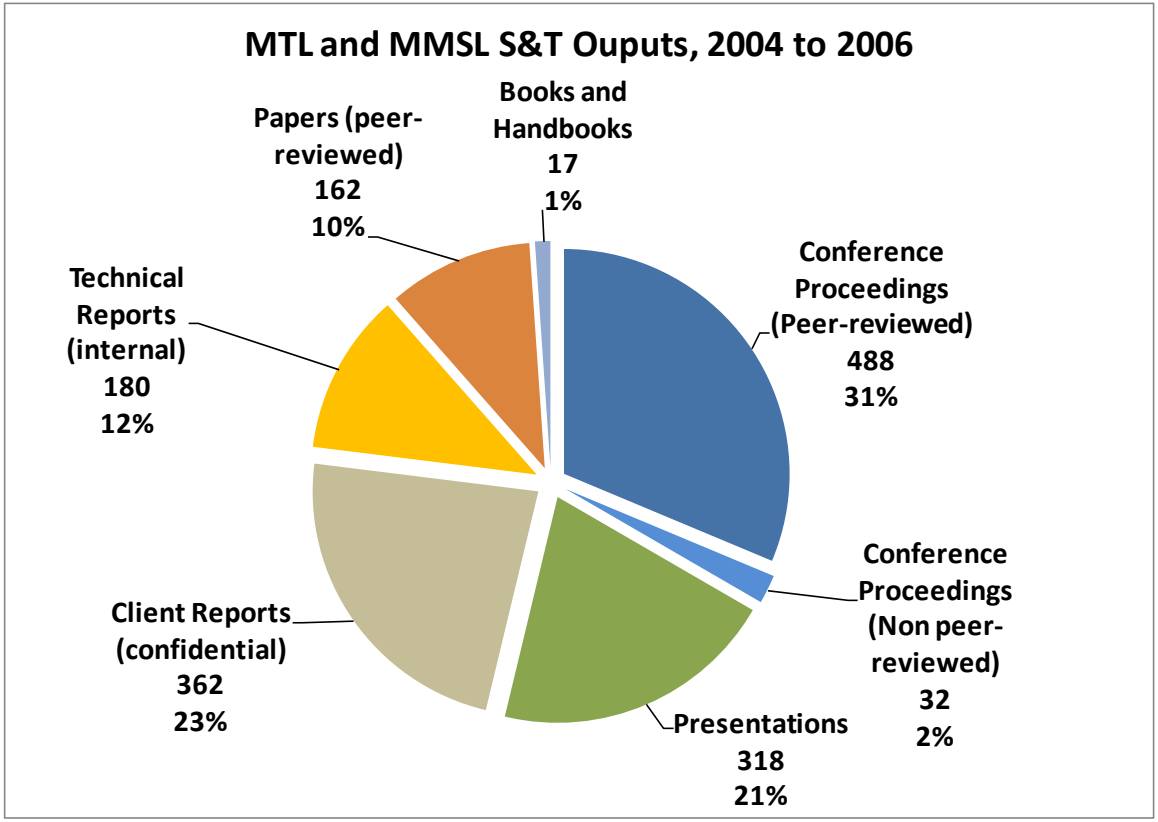
- **Science-Metrix conducted an evaluation of the science and technology (S&T) activities of the Minerals and Metals Sector (MMS) of Natural Resources Canada (NRCan)**
- See: Minerals and Metals Sector (MMS) Science and Technology (S&T) Evaluation (http://www.nrcan.gc.ca/evaluation/reprap/2009/e20090612-eng.php)
- **Both laboratories commonly maintain a balance between S&T activities that have very different ends:**
  - to support policies/standards and regulations (internally funded projects)
  - to enhance core competencies (internally funded projects)
  - to answer client/partner needs (revenue generating projects)
- **Focused on S&T activities of two laboratories:**
  - Material Technology Laboratory (MTL)
  - Mineral and Mining Science Laboratories (MMSL)
  - MTL and MMSL had 240 FTEs (divided almost equally)
  - 40% of MMS FTE's: 75% S&T and 25% support and regulation
  - 175 to 200 active S&T projects at any time
  - 80% of S&T projects are revenue generating in nature

Case Two (A): Methods

- **Use of a multi-faceted evaluation approach:**
  - Client satisfaction, impact and outreach web surveys
  - Interviews with 1) key laboratory individuals and 2) with clients, partners and other stakeholder organizations
  - Case studies (review of S&T projects)
  - **A review of all S&T outputs including**:
    1. A compilation of documented S&T output (performance tracking system)
    2. A bibliometric analysis of scientific output using Scopus Database
- **Why use bibliometrics?**
  - To validate the review documented S&T output with an independent method
  - To provide an estimation of the S&T contribution of labs to their Sector and to the Department (NRCan)
  - Obtain impact measures (and comparison with the world average in the same disciplines)

Case Two (A): Results

- **According to the review of all S&T outputs, MMS produced and disseminated science mainly by other means than peer-review publications**
  - Most of the output were technical/client reports (35%) and conference proceedings (33%)
  - **Peer-reviewed papers accounted only for 10% of total S&T outputs**
  - Cross-validation with bibliometric analysis: 95% coverage for peer-reviewed outputs



Source: Compiled by Science-Metrix from data provided by MMS-NRCan

Case Two (A): Results

- MMS is responsible of 9% of total departmental output
- MMS increased their level of output during the evaluation period
- MMS have a scientific impact comparable to the world average

MMS scientific output: number of papers and citedness, 1997-2006

	Number of papers*				Average of relative citations (ARC)**		
Institution	**1997-01**	**2002-06**	**1997-06**	**%**	**1997-2001**	**2002-2006**	**1997-2006**
Geological Survey of Canada	1,204	1,158	2,362	38%			
Canadian Forest Service	961	1,154	2,115	34%			
CANMET-CETC	221	372	593	10%			
Geomatics Canada	232	355	587	10%			
MMS	**187**	**341**	**528**	**9%**			
CANMET-MTL	***109***	***199***	***308***	***5%***			
CANMET-MMSL	***71***	***122***	***193***	***3%***			
CERL	*9*	*23*	*32*	*1%*	*n.s.*	*n.s.*	*n.s.*
Total	**2,814**	**3,358**	**6,172**	**100%**			

Source: Compiled by Science-Metrix from Scopus Database

Case Two: Context and Methods

- **Study 1: To inform science policy by measuring and comparing the scientific output of countries and international institutions in environmental research:**
  - Based on a bibliometric definition of the Environmental Research domain
  - Domain and sub-fields delineated using a specific journal classification (434 scientific journals; 580,500 papers between 1980 and 2004)
  - Aimed a positioning Canada and EC specifically in Environmental Research
  - See: 25 Years of Canadian Environmental Research - A Scientometric Analysis (http://www.science-metrix.com/pdf/SM_2006_001_EC_Scientometrics_Environment_Full_Report.pdf)
- **Study 2: To compile performance indicators on Environment Canada's (EC) scientific production at the level of: 1) S&T Branch/Directorates, 2) other Branches and 3) EC as a whole:**
  - Based on the counting of all EC papers in bibliographic database
  - Aimed at capturing the total output of EC in ***all*** S&T domains by lab/organizational unit
  - In order to inform S&T portfolio managers and organizational S&T strategy (alignment with mandate and priorities)
  - See: Measuring Environment Canada's Research & Development Performance (http://www.ec.gc.ca/doc/scitech/mecrdp_e.html)

Case Two: Results

- Canada's position as a global leader over the last 25 years is likely to change in the coming years if additional research efforts are not deployed
- EC ranks 7th at the world level and represents 14% of Canada's output in the field
- EC's scientific impact is one of the highest among other leading Canadian institutions (14% higher than the world level)

The most productive institutions in environmental research (1995–2004)

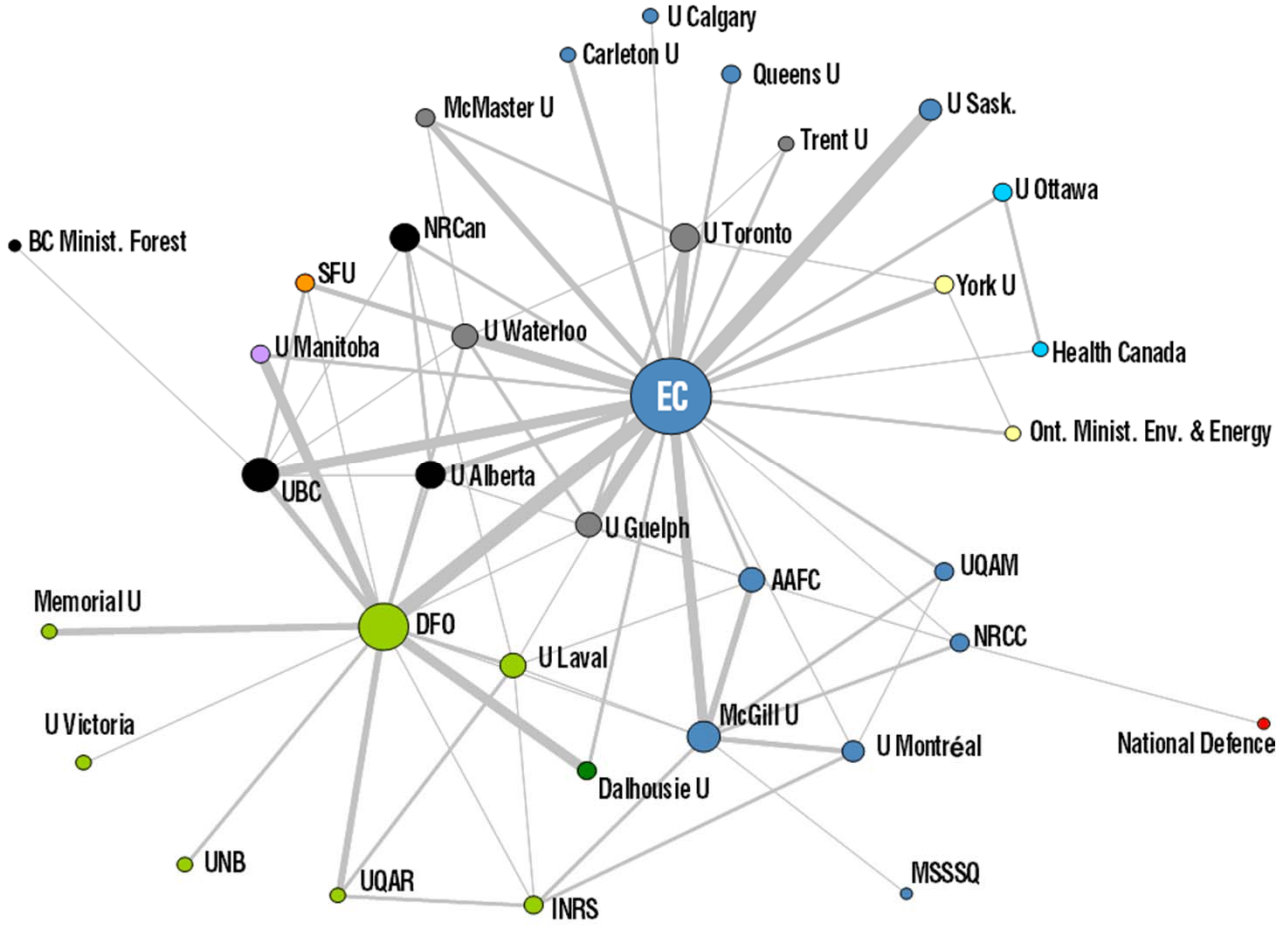
Rk	Institution	Country	Papers	ARIF	Nat't collabo.	Int'l collabo.	Rk 25 Yrs
1	NOAA - National Oceanic & Atmospheric Admin.	US	4,476	1.14	69.5%	22.2%	*(1)*
2	USEPA- US Environmental Protection Agency	US	3,505	1.17	69.2%	13.4%	*(2)*
3	USDA ARS- Agricultural Research Service	US	3,472	1.10	50.0%	14.6%	*(3)*
4	NASA	US	3,279	1.21	79.0%	32.3%	*(4)*
5	USGS - US Geological Survey	US	3,275	1.18	61.2%	12.5%	*(9)*
6	University of Washington	US	3,142	1.18	52.0%	22.4%	*(5)*
7	**Environment Canada**	**Canada**	**3,033**	**1.14**	**54.7%**	**33.9%**	***(7)***

CDN rank	World rank	Institution	Papers 95-04	ARIF 95-04	Int'l collabo	Nat't collabo	*Rank (25 yrs)*
1	**7**	**Environment Canada**	**3 033**	**1,14**	**33,9%**	**54,7%**	***(1)***
2	**25**	**Fisheries and Oceans Canada**	**1 826**	**1,16**	**28,7%**	**53,4%**	***(2)***
3	28	University of British Columbia	1 761	1,11	34,9%	34,0%	*(3)*
4	47	University of Toronto	1 327	1,18	32,7%	35,9%	*(4)*
5	51	McGill University	1 262	1,09	32,6%	43,3%	*(5)*

Case Two: Results

- Over 80% of EC papers were produced in collaboration with external-to-EC scientists
- EC is the primary collaborator for almost all of the 14 most productive institutions
- Collaboration rate has risen significantly over the past 25 years
- EC is the principal hub of environmental research in Canada

Network of the most collaborative Canadian institutions in environmental research (1995–2004)

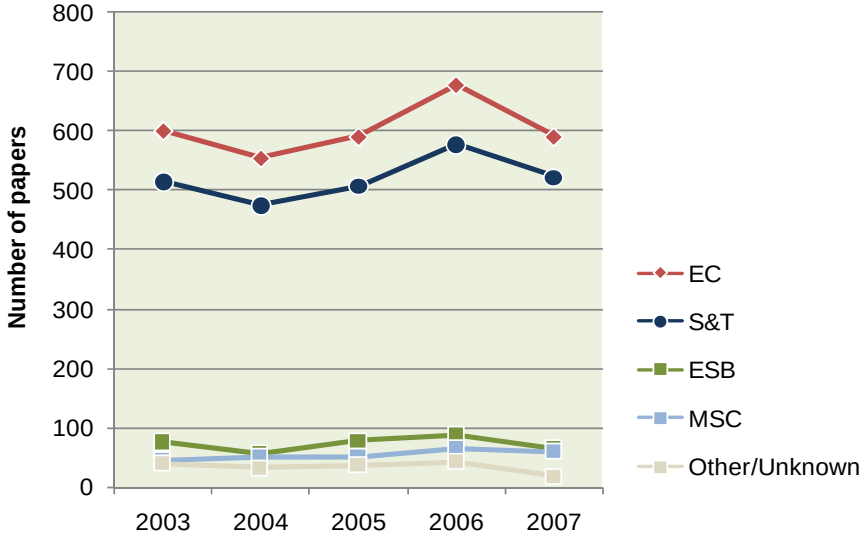


Source: Compiled by Science-Metrix from Thomson-Scientific data

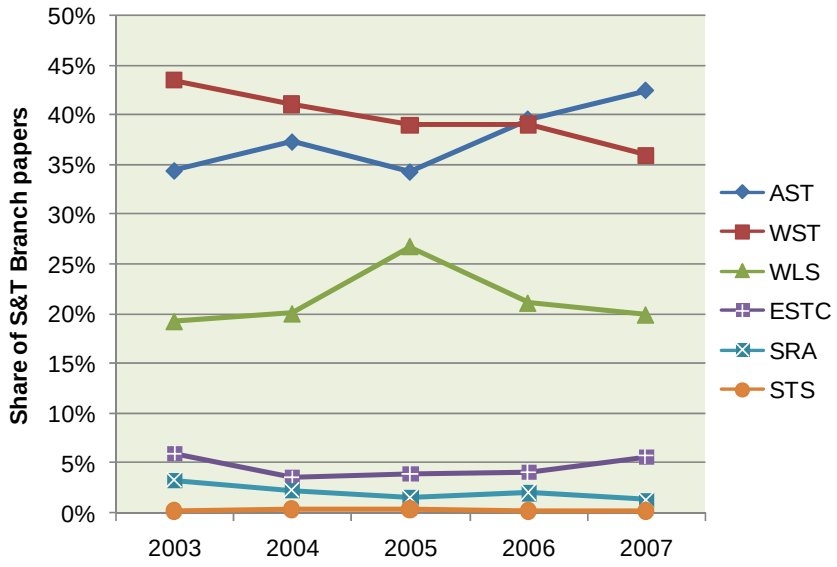
Case Two: Results

- EC produced approx. 600 peer-reviewed papers per year on average in journals indexed in Scopus
- The S&T Branch is the main contributor, 520 papers per year on average (or 85%)
- The Directorates that have contributed the most: Atmospheric S&T (AST), Water S&T (WST) and Wildlife and Landscape Science (WLS)

Total number of papers produced by EC and by Branch



Share (%) of S&T Branch by Directorate



Source: Compiled by Science-Metrix from Thomson-Scientific data

Case Two: Pros & Cons of Bibliometrics

Strengths	Weaknesses
Effective to characterize the **level of scientific output** of organization and units	• Publication in peer-reviewed journals is not a priority in this context (S&T in support of policy, regulations and programs) • Do not capture the whole range of outputs of science-based governmental organizations
Effective method to measure **scientific uptake** (citations)	• Bibliometrics focus on scientific community impacts • Strategic or expected S&T outcomes are not captured/measured
Allow **comparative analysis of the scientific performance** (nationally & internationally)	• Difficult to account for organizational S&T mandate, priorities and levels of resources ($, infrastructure and FTEs)
Effective to characterize scientific **collaboration** trends	• Do not capture the collaboration in other scientific and technical activities

Conclusion

- Using traditional bibliometric indicators, the presented applications illustrate the strengths and limitations of bibliometrics for research portfolio evaluation
- Bibliometrics should ideally be designed and used as a complimentary line of evidence as opposed to an isolated method to evaluate research
- In particular, bibliometric assessments of research organizations with diverse S&T outputs should be complemented by other evaluation methods
- Bibliometrics is a scientific field in itself. New indicators are currently being developed to improve the usefulness of bibliometrics in different contexts
- The limitations of bibliometrics must be adequately managed by experts in the field
- The limitations of bibliometrics, both at the conceptual and methodological levels, need to be reported and communicated transparently to ensure proper and balanced analysis/management responses

Thank you for your time and feedback

CONTACT INFO

AUTHORS

David Campbell, M.Sc.
Senior Bibliometrician/Evaluator | Science-Metrix
david.campbell@science-metrix.com

Frédéric Bertrand M.Sc.
Vice-President, Evaluation | Science-Metrix
frederic.bertrand@science-metrix.com

WEB SITE

www.science-metrix.com

OFFICES

CANADA

Science-Metrix Inc.

Address	1335, Mont-Royal E. Montreal, Quebec Canada H2J 1Y6
Toll-free	1.800.299.8061
Phone	1.514.495.6505
Email	info@science-metrix.com

UNITED STATES

Science-Metrix Corp.

Address	2331, Mill Road, Suite 100 Alexandria, Virginia 22314 USA
Toll-free	1.888.388.8023
Phone	1.703.224.8822
Email	smc@science-metrix.com