

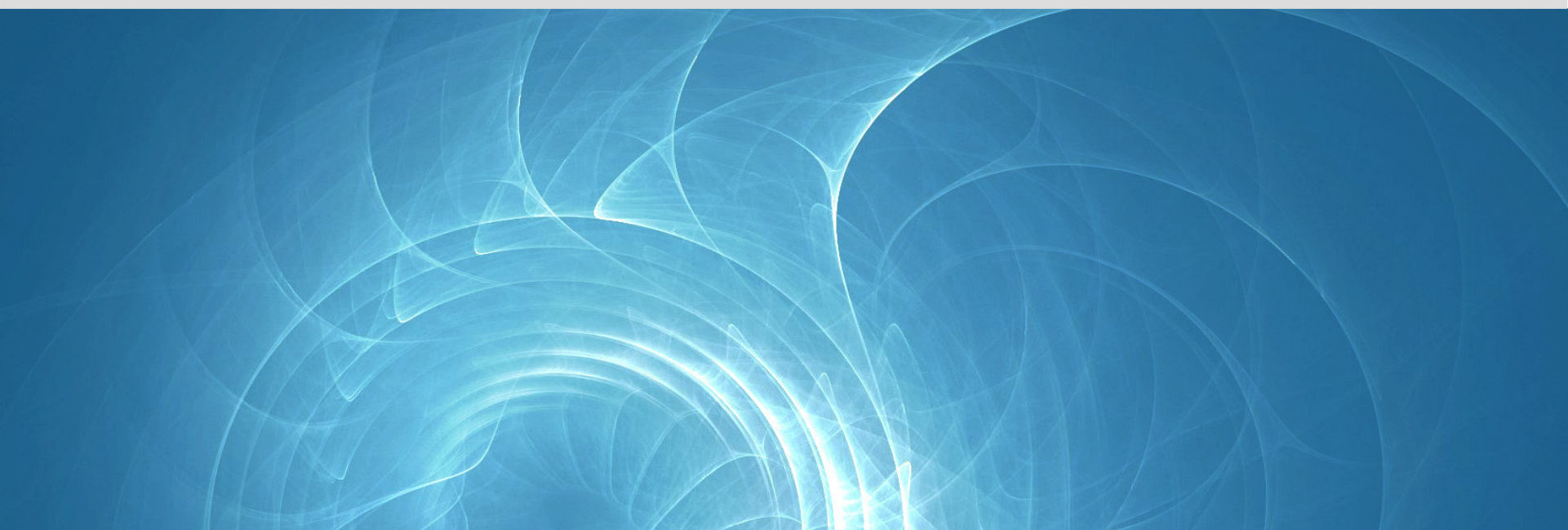


Science-Metrix

**Bibliometrics as a performance measurement tool
for the evaluation of research:**

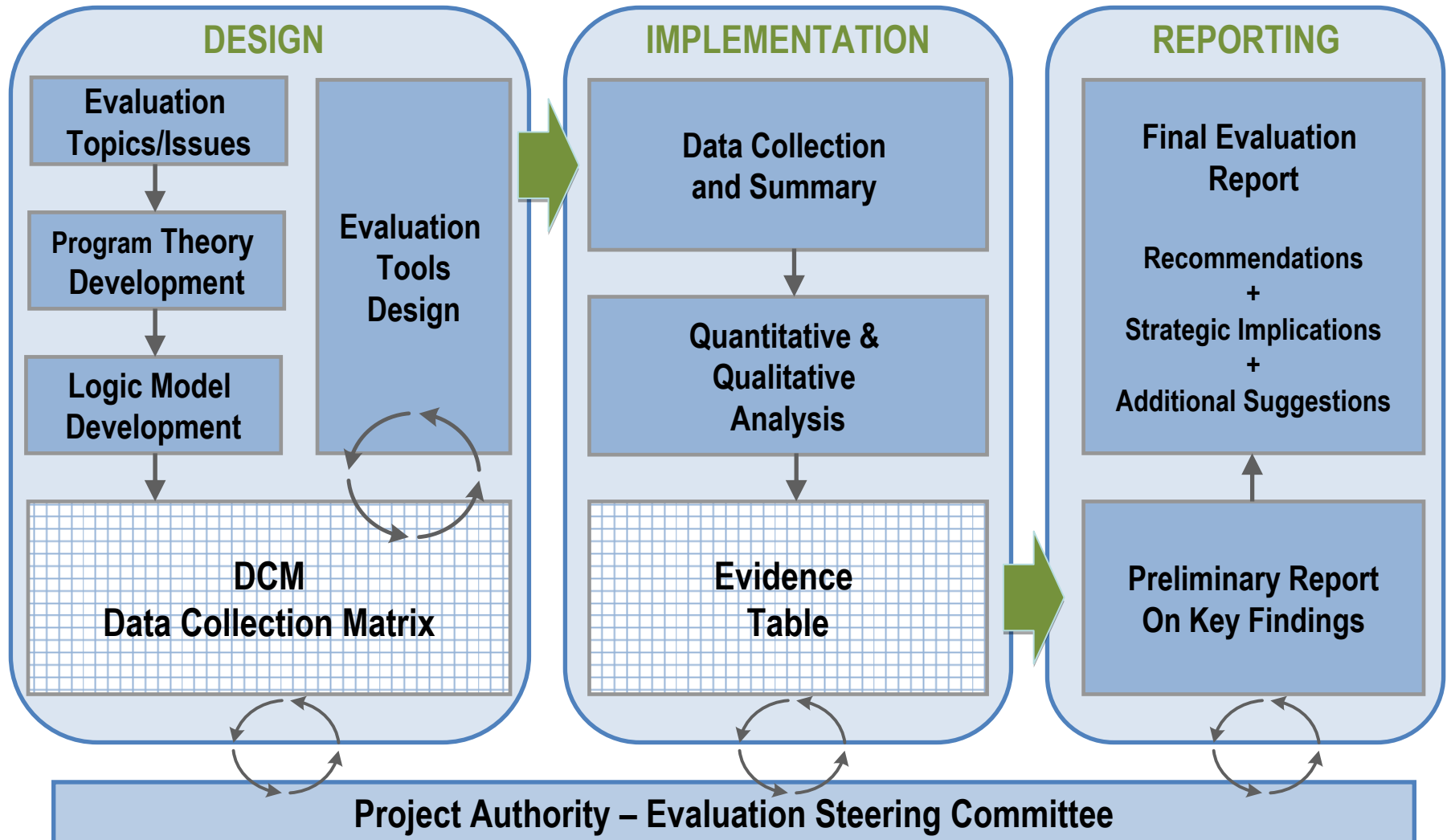
The cases of the National Cancer Institute of Canada and
the Canadian Forest Service

2009 Annual CES Conference | June 1st, 2009



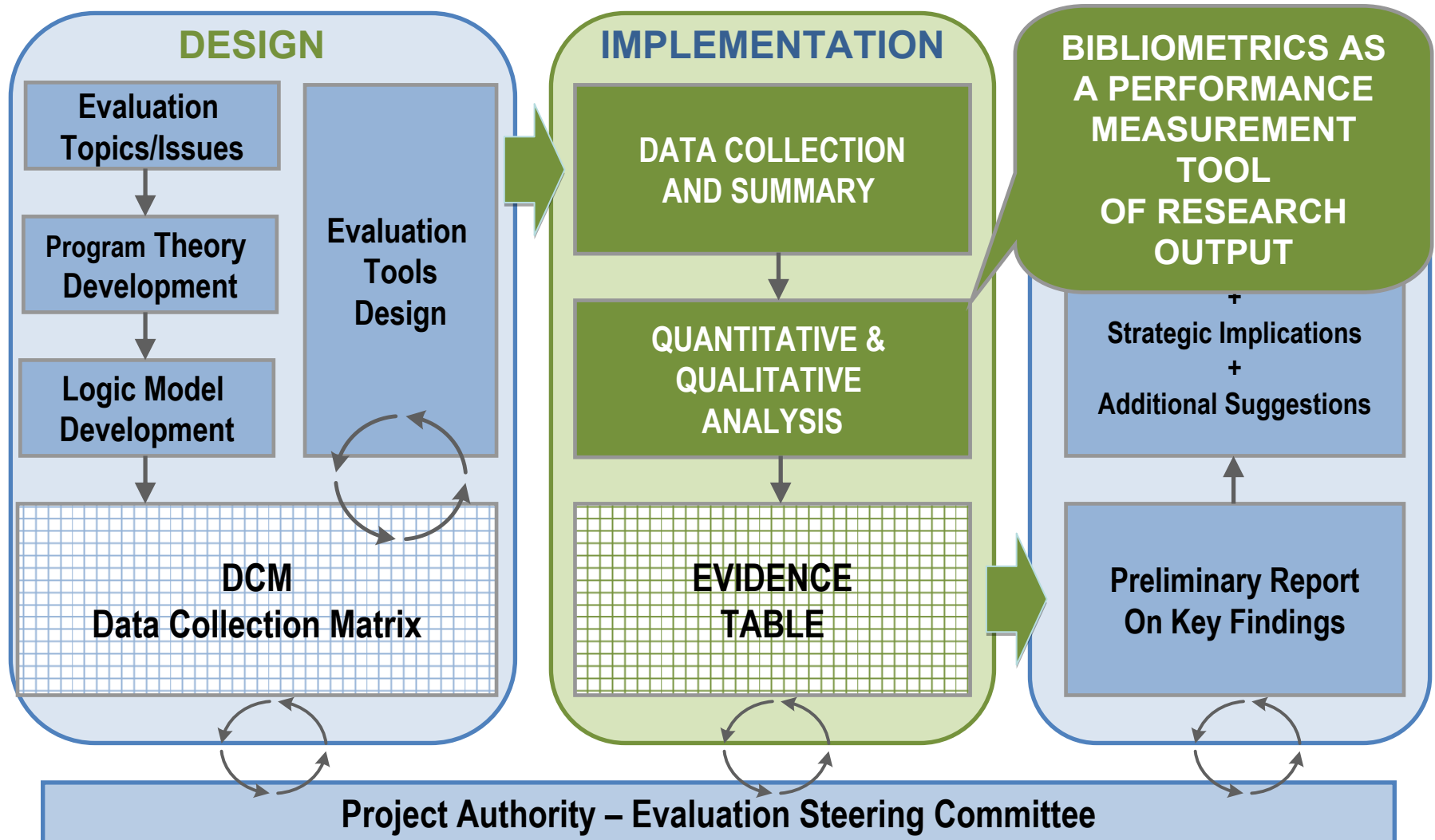


Integrated Program Evaluation Tools: From Logic Model to Final Report





Integrated Program Evaluation Tools: From Logic Model to Final Report





Bibliometrics as a Performance Measurement Tool for the Evaluation of Research

- In the NSE, the results of scientific research are mainly disseminated through the publication of peer-reviewed papers in scholarly journals.
- Thus, the scientific performance of researchers in the NSE is best assessed through the quantification of peer-reviewed papers.
- **Definition of bibliometrics:** A set of methods used in the quantification of bibliographic records of peer-reviewed publications and/or patents.
- Developed by information scientists, bibliometrics, **in combination with peer-review**, is now being applied more frequently in Canada and abroad to the field of research evaluation. This is attributable to the fact that:
 - Governments ➔ recognized the importance of accountability for public spending in research thereby ➔ the need to balance the more traditional assessment of peer review with **more objective methods like bibliometrics**.
 - Bibliometric indicators are considered to be the **most objective and reliable measures of research output** as they are based on a set of internationally recognized standards, in addition to being cost effective in their usage.

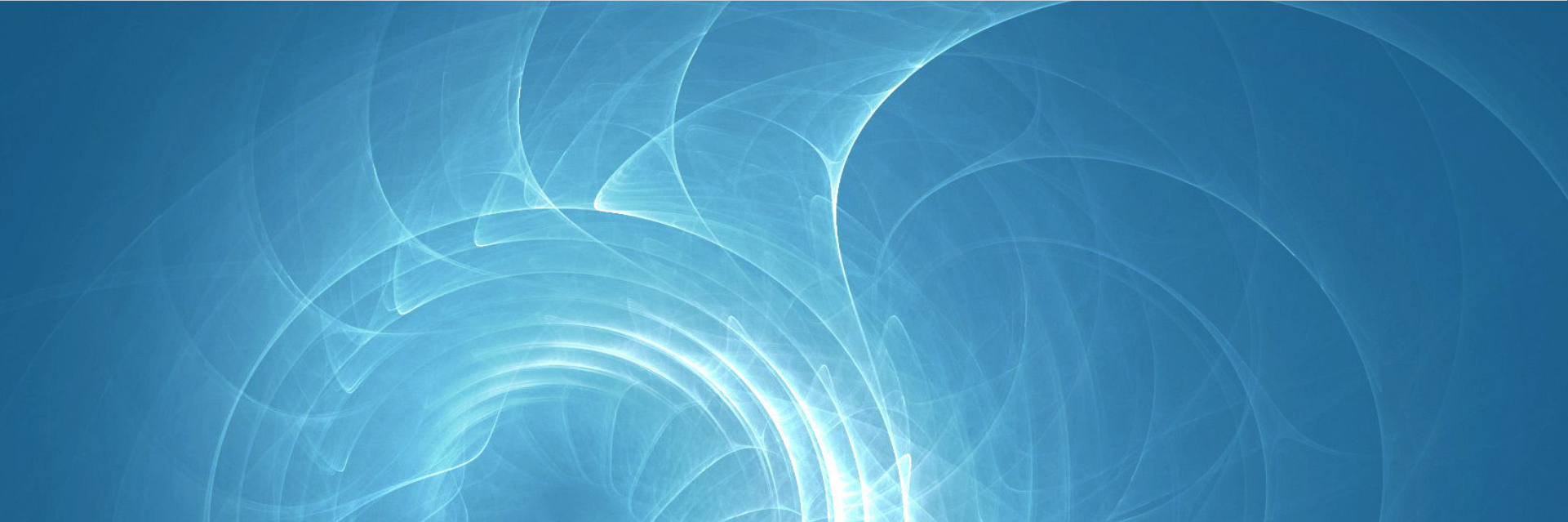


Bibliometrics as a Performance Measurement Tool for the Evaluation of Research

- This study's objective is to show how bibliometrics can be used as a performance measurement tool for the evaluation of research in the context of:
 - Evaluating the selection process and effects of a funding program: the case of the National Cancer Institute of Canada
 - Monitoring the performance of a research organization: the case of the Canadian Forest Service
- It should be noted that **Research Evaluation** makes use of a variety of indicators to draw the most complete picture possible of the complex aspects that account for the performance of research organizations.
- This study is focusing on only one aspect of Performance Measurement in the context of research evaluation, namely **Bibliometrics**.

Case One

The National Cancer Institute of Canada (NCIC):
Evaluating the selection process and effects of a funding program





Case One (NCIC): Objectives

- Funds awarded by the NCIC are raised exclusively through donations to the Canadian Cancer Society and the Terry Fox Foundation.
- As such, the NCIC is accountable to the public for the efficient and effective use of donor dollars for the purpose of funding research.
- In 2005, the NCIC launched an evaluation of its Operating and Program Project Grants.
- Within the context of this evaluation, an assessment of the scientific research funded by the NCIC was performed to address three questions:
 - 1) Has the NCIC peer-review process been successful in selecting outstanding Canadian cancer researchers in terms of scientific impact?
 - 2) Has NCIC funding contributed to increasing the scientific performance, in terms of scientific impact, of the researchers it supported?
 - 3) How do NCIC-supported researchers compare, in terms of scientific output, to their US counterparts who received support from the NCI?



Case One (NCIC): Methods

- **Selection of bibliographic database:** Web of Science (Thomson Reuters)
- **Three datasets:**
 - 1) NCIC-funded applicants (core dataset)
 - 2) US NCI-funded applicants (benchmark dataset)
 - 3) Canada without NCIC-funded applicants
- **Construction of datasets:**
 - 1) For the core and benchmark datasets, lists of funded applicants from comparable funding programs were provided by the NCIC and US NCI
 - 2) Automatic retrieval of funded-applicants' papers from the database which maximizes recall
 - 3) Cleaning of publication portfolios to maximize precision
 - 4) For the core dataset, papers were tagged as being published while not receiving support from the NCIC ("**Non-supported papers**") or as being published with NCIC support ("**NCIC-supported papers**") based on grant periods
 - 5) For the Canadian dataset, papers with researchers other than NCIC-funded applicants that indicated a Canadian address were retrieved from the database (all papers and papers specific to cancer research)



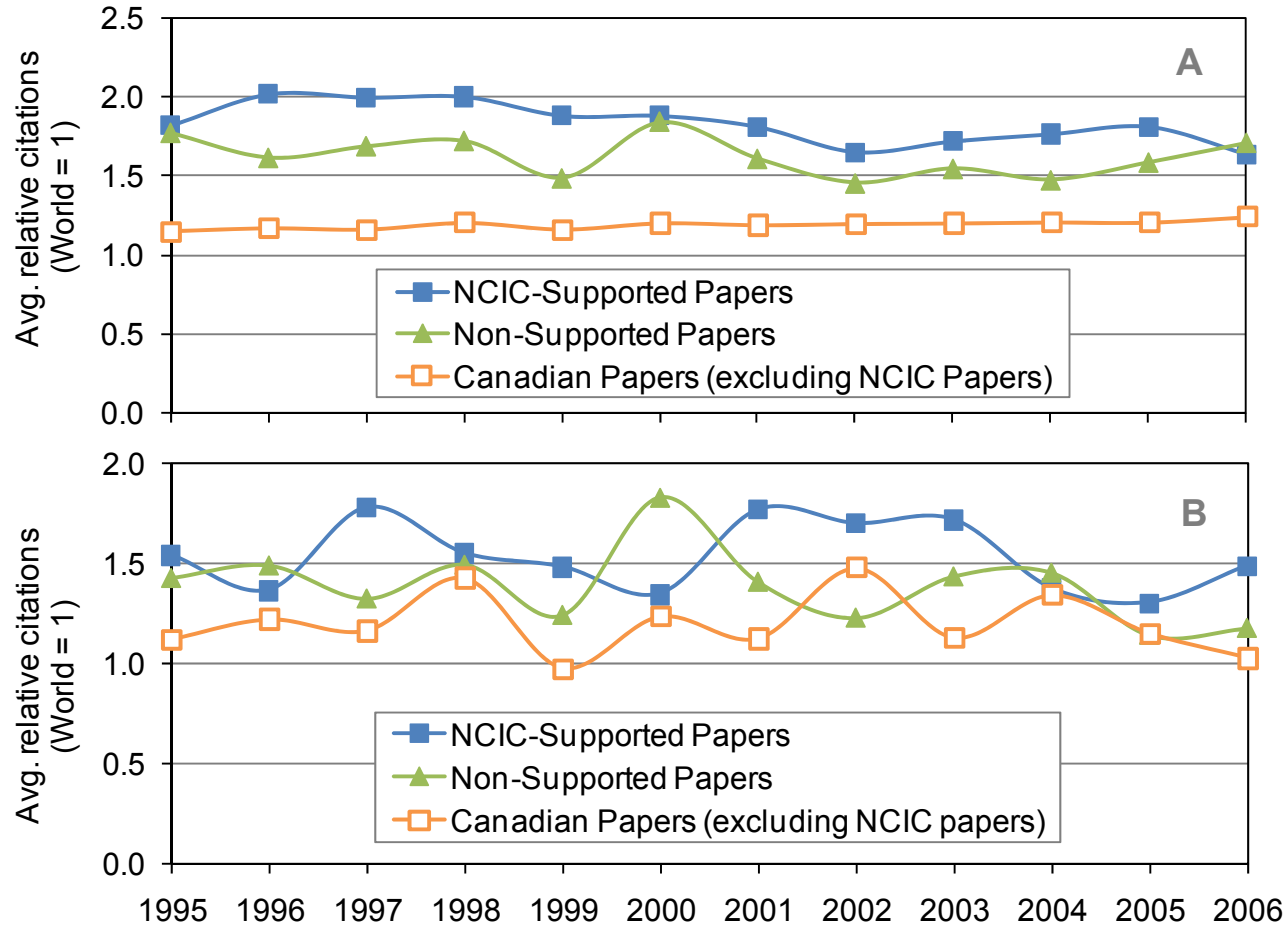
Case One (NCIC): Methods

- **Bibliometric indicators:**
 - 1) **Average of relative citations (ARC):** Provides a *direct* measure of scientific impact based on paper citation counts. Self-citations are excluded. The indicator accounts for different citation patterns across fields and subfields of science and for difference in the age of papers.
 - 2) **Average of relative impact factors (ARIF):** Provides an *indirect* measure of scientific impact based on the impact factors of journals in which papers are published. The indicator accounts for different citation patterns across fields and subfields of science and for difference in the age of papers.
 - 3) **Most cited papers:** Percentage of papers that are in the 5% of most cited papers. Self-citations are excluded. The indicator accounts for different citation patterns across fields and subfields of science and for difference in the age of papers.
- All indicators provided similar data. Only the ARC is presented.



Has the NCIC peer-review process been successful in selecting outstanding Canadian cancer researchers?

Average of relative citations of papers by NCIC-supported and non-supported Canadian researchers overall (A) and in cancer research (B)



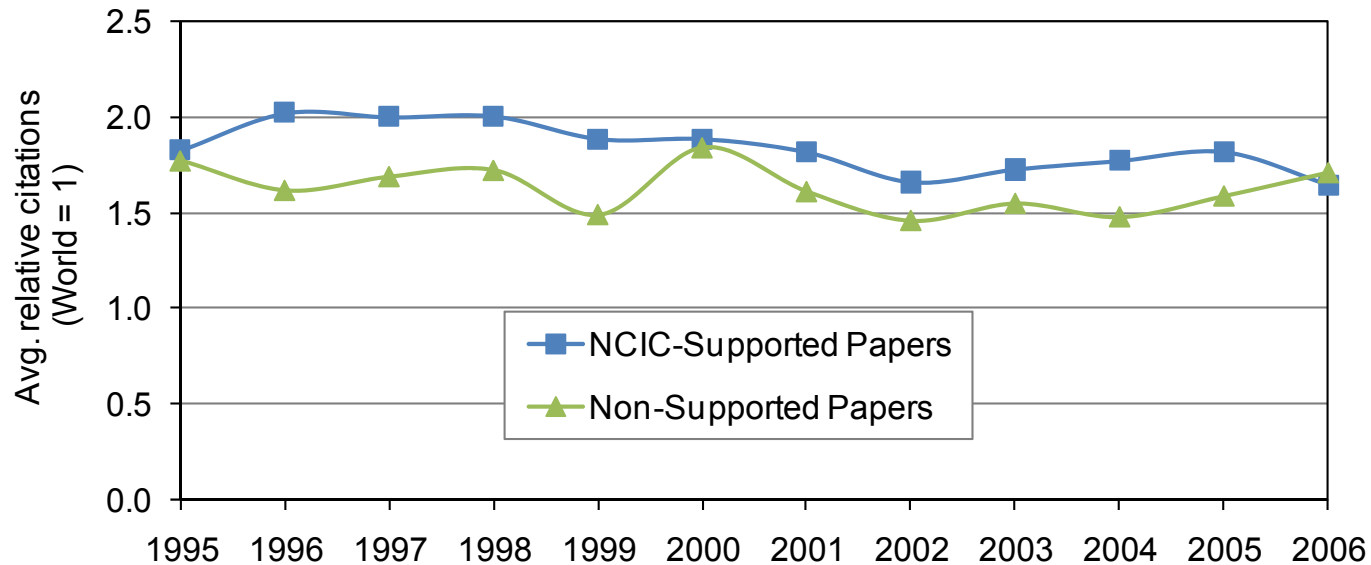
Source: Calculated by OST and Science-Matrix from Thomson Reuters' WoS

- NCIC-funded scientists (n = 679)
- 12,244 out of 22,793 papers by NCIC-funded researchers were published with support from the NCIC between 1995 and 2006.
- For both A and B, NCIC-supported researchers had significantly higher scientific impact than other Canadian researchers.
- NCIC effectively selected outstanding Canadian researchers in cancer research.



Has NCIC funding contributed to increasing the scientific performance of the researchers it supported?

Average of relative citations of NCIC-supported papers and non-supported papers

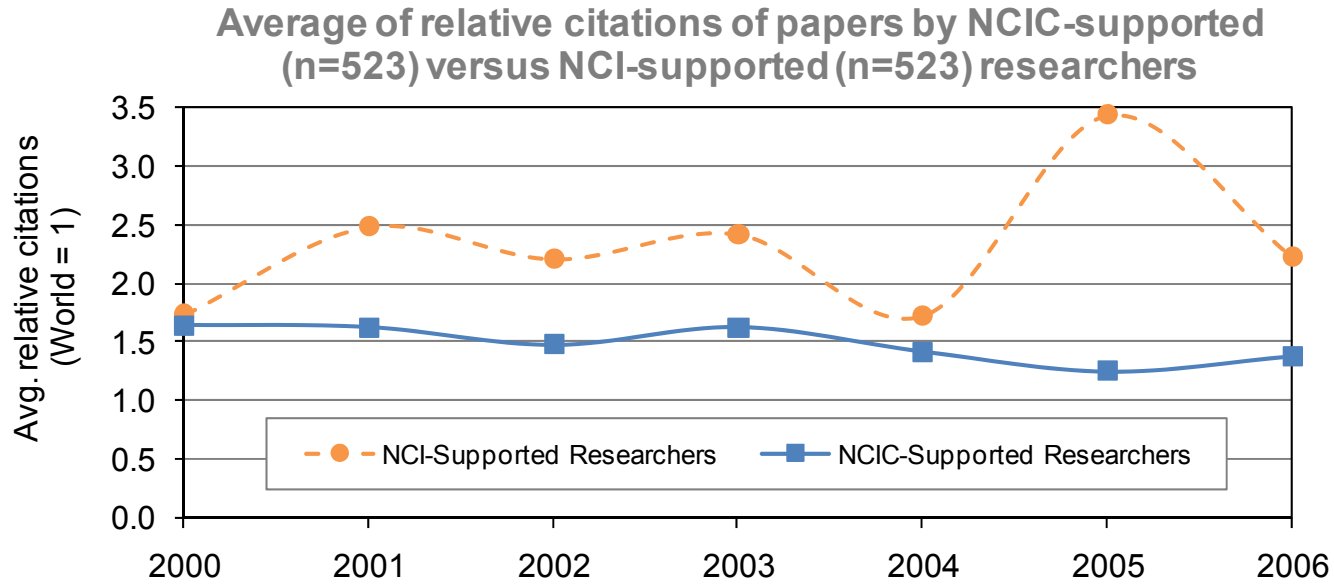


Source: Calculated by OST and Science-Matrix from Thomson Reuters' WoS

- The scientific impact of papers authored by NCIC-funded researchers while they were receiving funds from the NCIC is significantly higher than that of their papers authored while not supported by the institute.
- This finding supports the hypothesis that the funds provided by the NCIC contributed to increasing the scientific performance of successful applicants.



How do NCIC-supported researchers compare to their US counterparts who received support from the NCI?

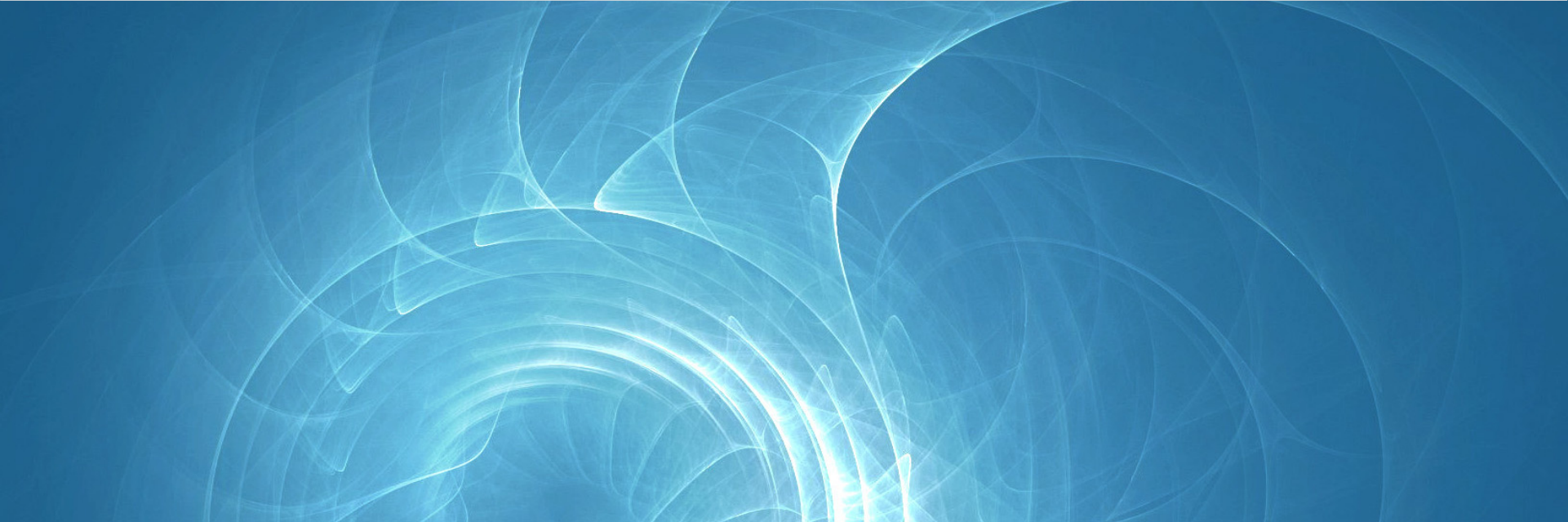


Source: Calculated by OST and Science-Matrix from Thomson Reuters' WoS

- NCIC-funded scientists published nearly as many papers as did NCI-supported scientists (11,019 versus 11,794), despite a level of funding that is nearly five times smaller (considering only financial support from the NCI and the NCIC).
- However, the scientific impact of papers authored by NCIC-funded researchers is significantly lower than that of papers authored by NCI-funded researchers.
- Access to greater financial resources may not directly translate into more scientific publications, but might instead result in projects of higher scientific impact that could not have been accomplished otherwise.

Case Two

The Canadian Forest Service (CFS): Monitoring research performance



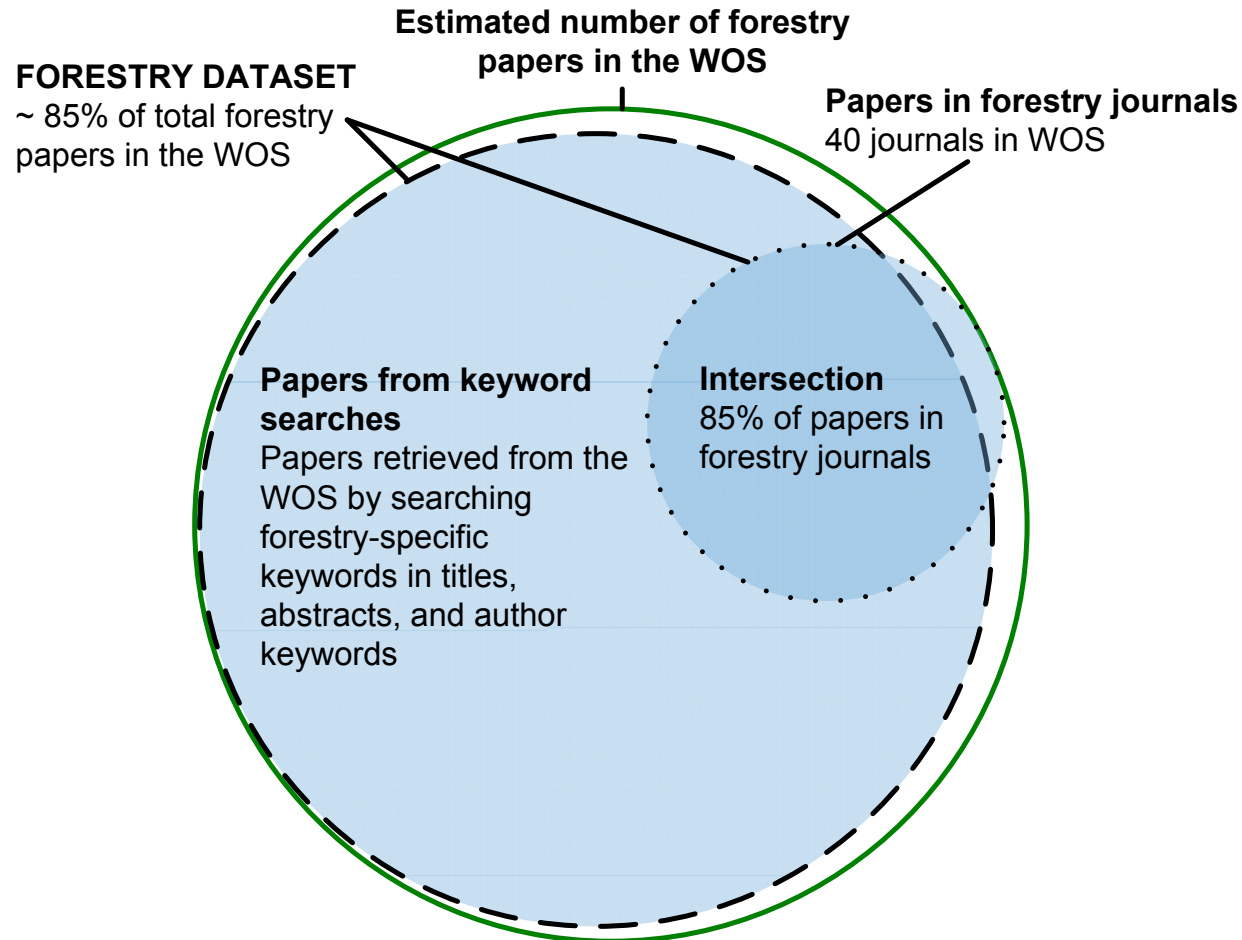


Case Two (CFS): Objectives

- Bibliometrics can be used in a broader context to monitor the research performance of an organization compared to its competitors, its country and the world.
- Among other things, bibliometric indicators provide information on the level of research activity, the relative intensity of research performed in a specific field relative to the world (i.e., within a specific specialization), the scientific impact of research and the level of scientific collaboration.
- CFS has a long history as a research organization where publication (internal and external) has played a major role in disseminating knowledge gained through research.
- Since the mid-1990s, program review emphases have shifted to:
 - Journal publications
 - Collaboration with universities
- This study provides a basis for objective CFS internal, national, and international comparisons.
- It is not perfect (e.g., it does not cover internal publications) and needs historical and policy background for interpretation.

Case Two (CFS): Methods

- Selection of bibliographic database: Web of Science (Thomson Reuters)
- Construction of datasets:



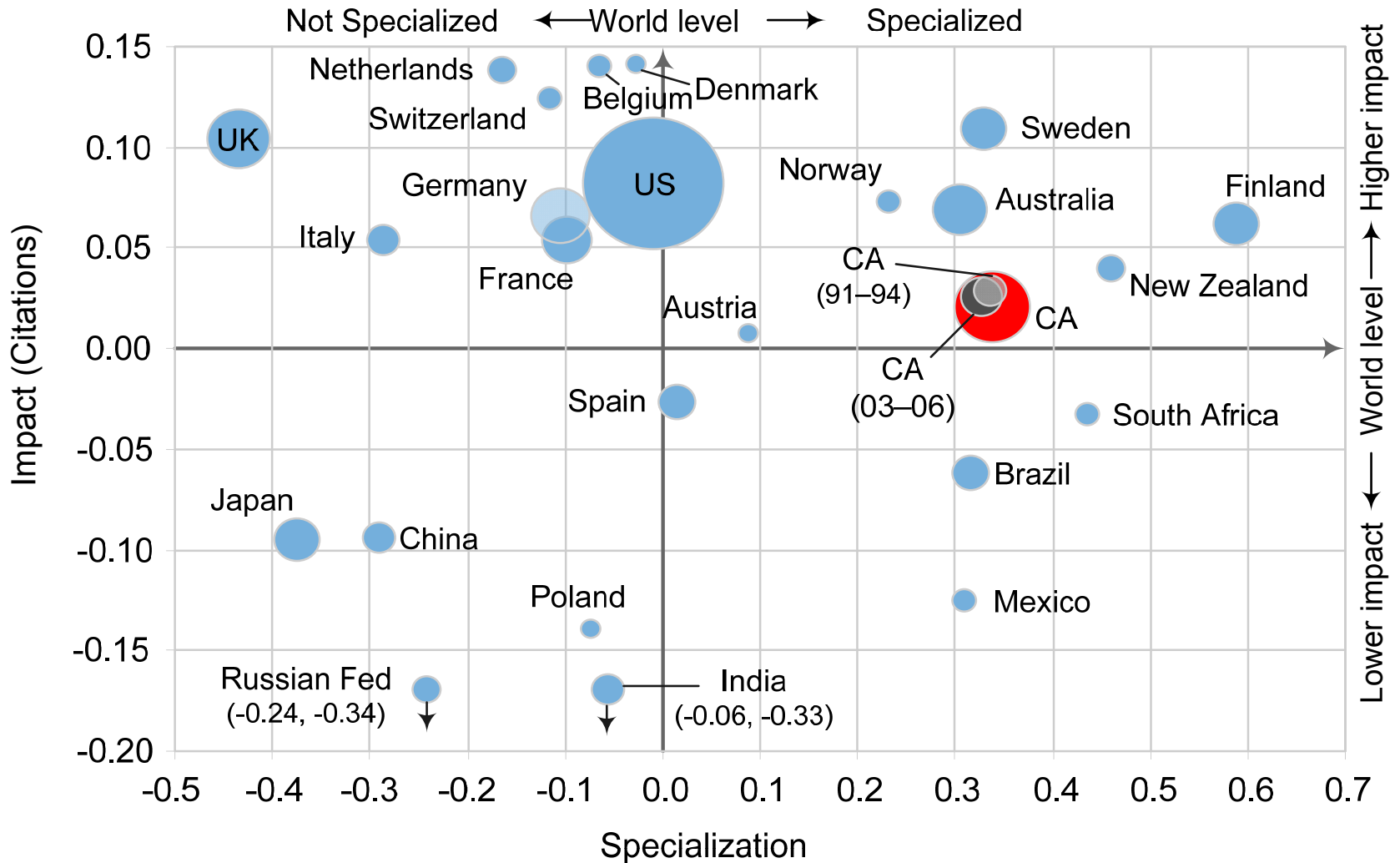


Case Two (CFS): Methods

- **Bibliometric indicators:**

- 1) **Number of papers:** The number of scientific papers by institution and by country, based on author addresses in a specific dataset
- 2) **Papers per capita:** The number of papers at the country level, weighted per capita using population statistics
- 3) **Specialization index (SI):** An indicator of the intensity of research of a given entity in a research area relative to the intensity of the world in the same research area
- 4) **Average of relative citations (ARC):** Provides a *direct* measure of scientific impact based on citation counts to papers.
- 5) **Average of relative impact factors (ARIF):** Provides an *indirect* measure of scientific impact based on journals' impact factors
- 6) **Collaboration rate (national, international & total):** This is an indicator of the relative intensity of collaboration between entities (e.g., countries, institutions)

Positional analysis of leading countries in forestry (number of papers, SI, and ARC), 1991–2006





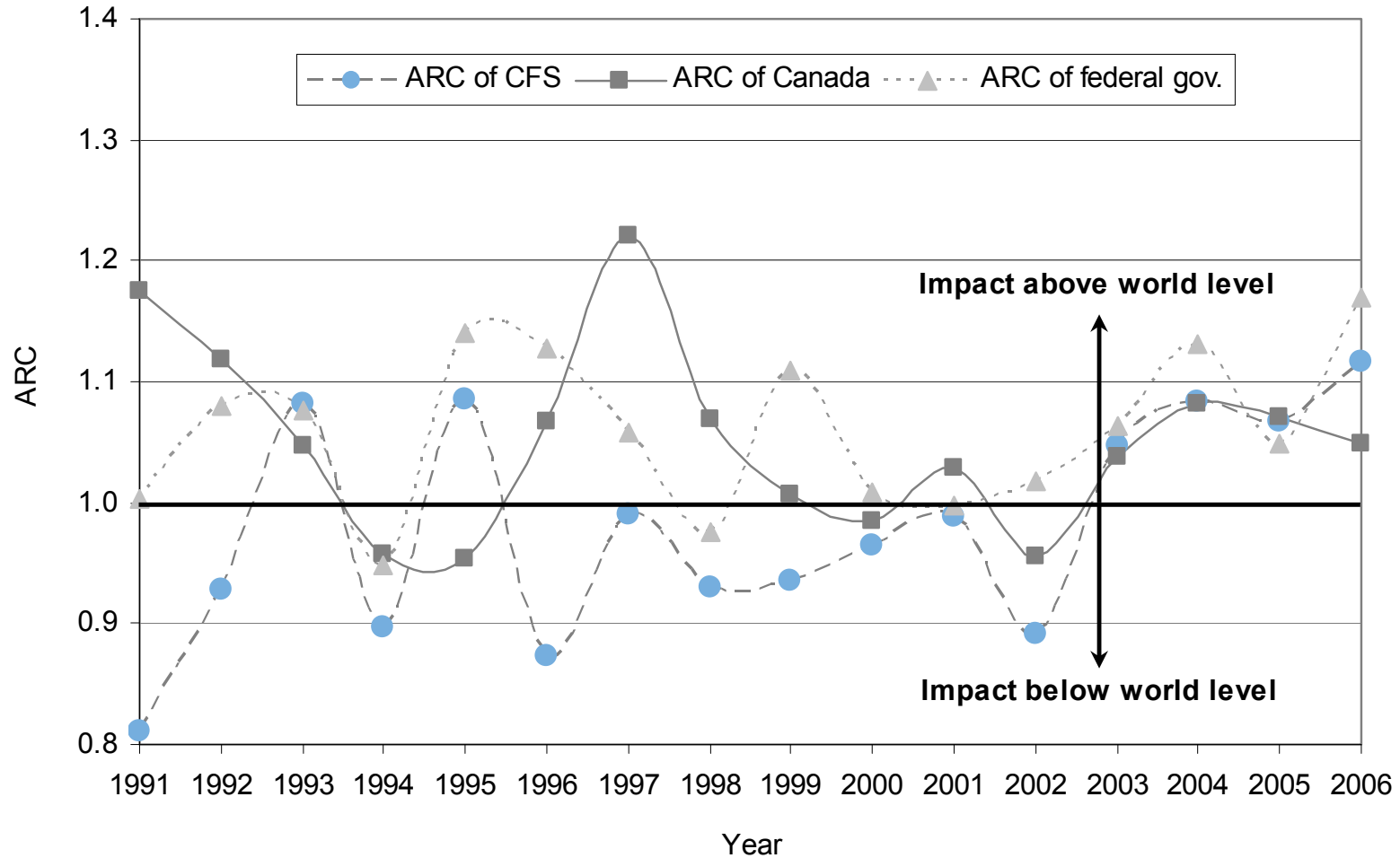
Strength of countries in forest research, 1991–2006

(only the top 25 countries in absolute output presented)

Country	Papers	Per Capita Output	SI	ARIF	ARC	Overall Rank
Sweden	4,649	32.9				
Finland	4,024	49.0				
US	37,917	8.6				
Australia	6,172	20.5				
Canada	11,361	23.1				
UK	7,471	7.9				
New Zealand	1,799	30.1				
Switzerland	1,571	13.6				
Belgium	1,273	7.8				
France	5,325	5.6				
Netherlands	1,967	7.8				
Norway	1,255	17.7				
Germany	7,142	5.5				
Denmark	1,058	12.5				
Spain	3,257	5.1				
Italy	2,342	2.5				
Austria	1,171	9.0				
Brazil	2,874	1.0				
South Africa	1,519	2.3				
Mexico	1,318	0.8				
Japan	4,455	2.2				
India	2,273	0.1				
China	2,275	0.1				
Poland	1,137	1.8				
Russian Fed	1,823	0.8				



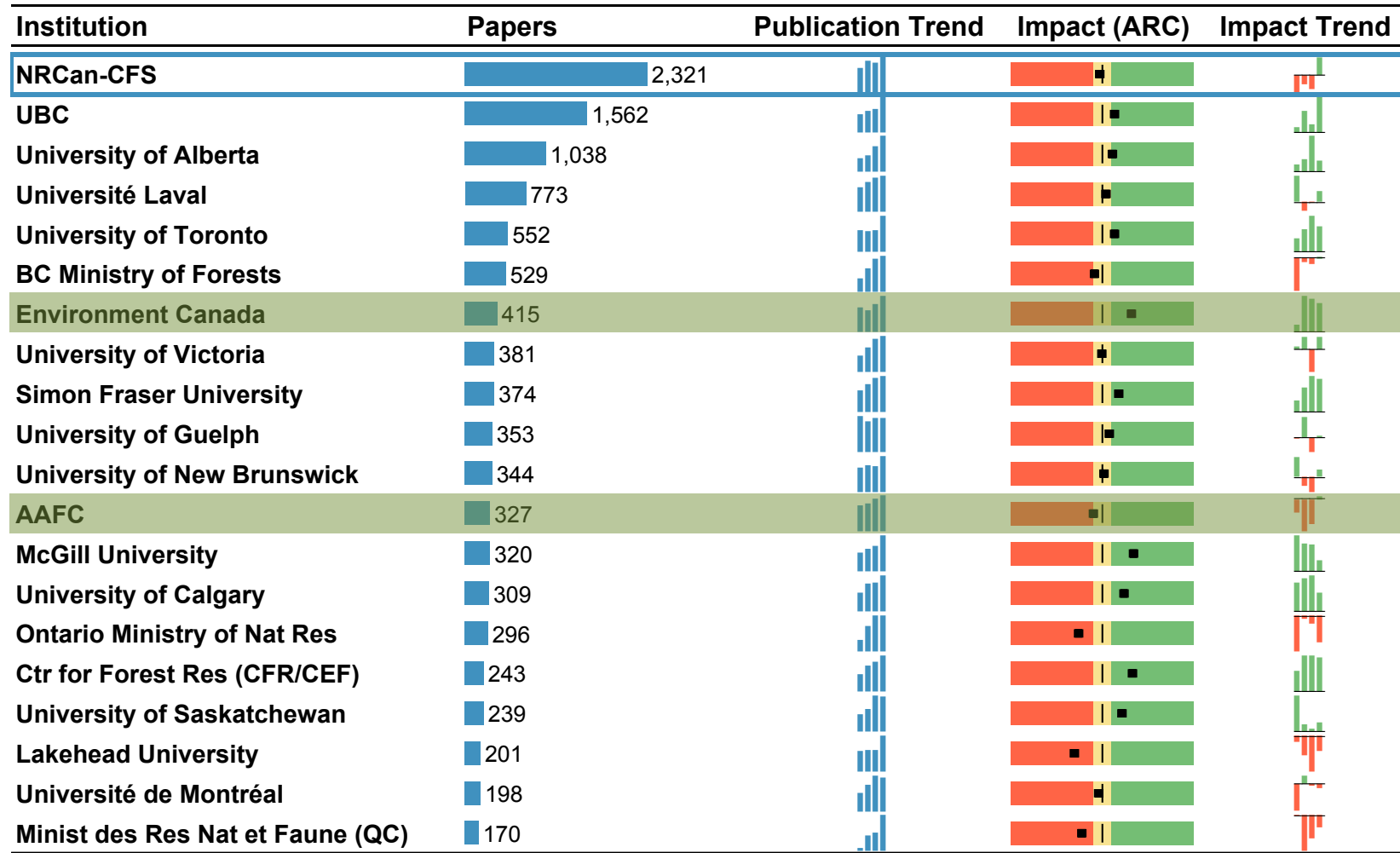
CFS's scientific impact in forestry compared to the Canadian federal government and Canada, ARC 1991–2006





Strength of leading Canadian institutions in forestry, 1991–2006

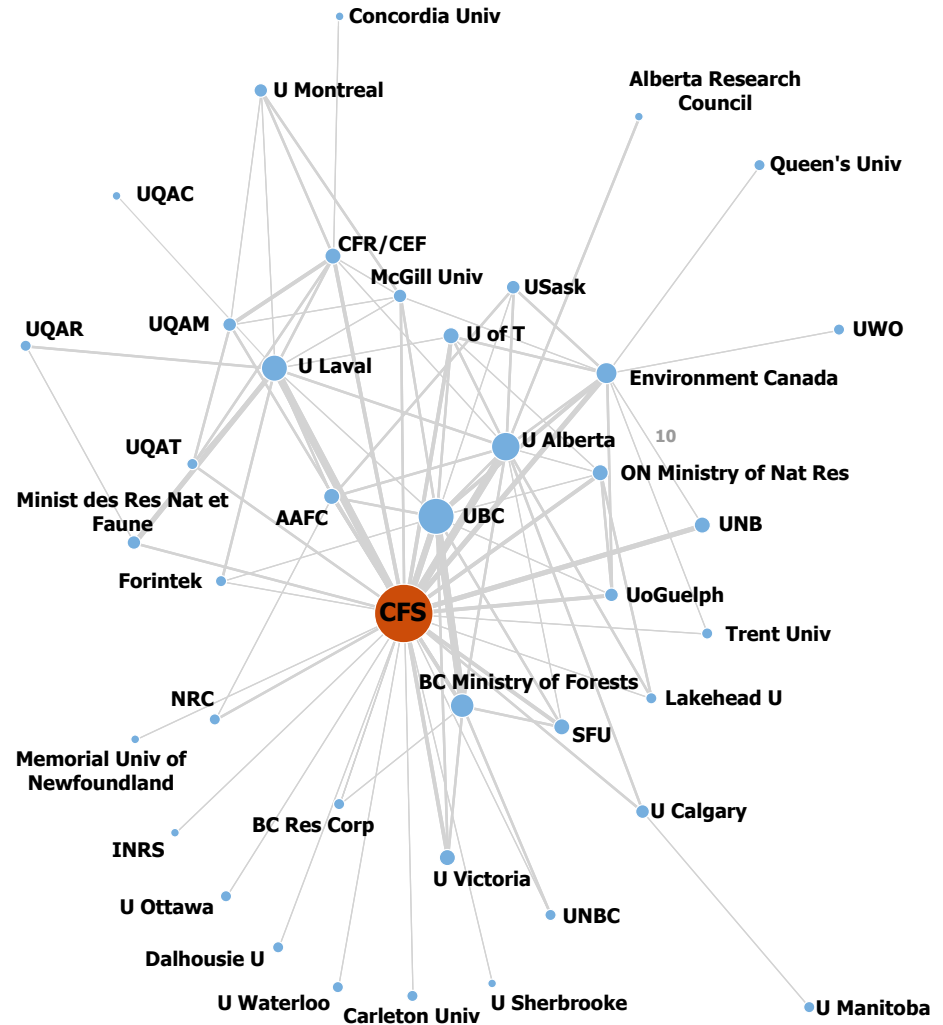
(only the top 20 institutions in absolute output presented)





Collaboration network of leading Canadian institutions in forestry *(only links representing at least 10 collaborations for 1991–2006 plotted)*

- CFS has the highest number of papers authored in national collaborations
- It is followed by UBC and the University of Alberta
- CFS is the centre of a Canadian collaboration network, with links to 30 institutions
- CFS's national collaboration rate increased from ~42% to ~62%
- CFS's international collaboration rate increased from ~14% to ~22%





Monitoring CFS Research: Key findings

- CFS was the most prolific institution in Canada in terms of the number of papers published in forestry. It ranked 3rd internationally behind the USDA (including USFS) and the Swedish University of Agricultural Sciences.
- The impact of CFS's publications fluctuated at or below the world level from 1991 to 2002 and was well below that of many leading Canadian and foreign institutions.
- However, during the 2003–2006 period, the impact of CFS increased to above world level. It even increased above Canada's overall impact in 2006.
- CFS is a major collaborator, ranking 3rd in number of collaborations within the network of the top 50 world institutions, and it acts as the central hub within the network of leading Canadian institutions.



Conclusion

- Bibliometric indicators do provide objective answers to questions raised by the evaluation of research:
 - Is public money being well spent for the purpose of funding research?
 - Is the funding helping researchers achieve the expected level of scientific excellence?
 - Are research organizations (and/or countries) keeping pace globally?
 - Are government scientists collaborating in knowledge production with industry and universities?
- When used on their own, bibliometric indicators do have limitations.
- They are better used in combination with other indicators (e.g., peer ratings), i.e., using the ***Method of Converging Partial Indicators***.
- Bibliometrics is a complex science. Amateur and off-the-shelf use of it should be avoided as it often leads to errors and sustained misconceptions. Proper training of practitioners, users, evaluators and managers is crucial.
- Bibliometrics should be part of a ***practice-driven*** research evaluation continuum.



Contact information



David Campbell, MSc

david.campbell@science-metrix.com

Frédéric Bertrand, MSc

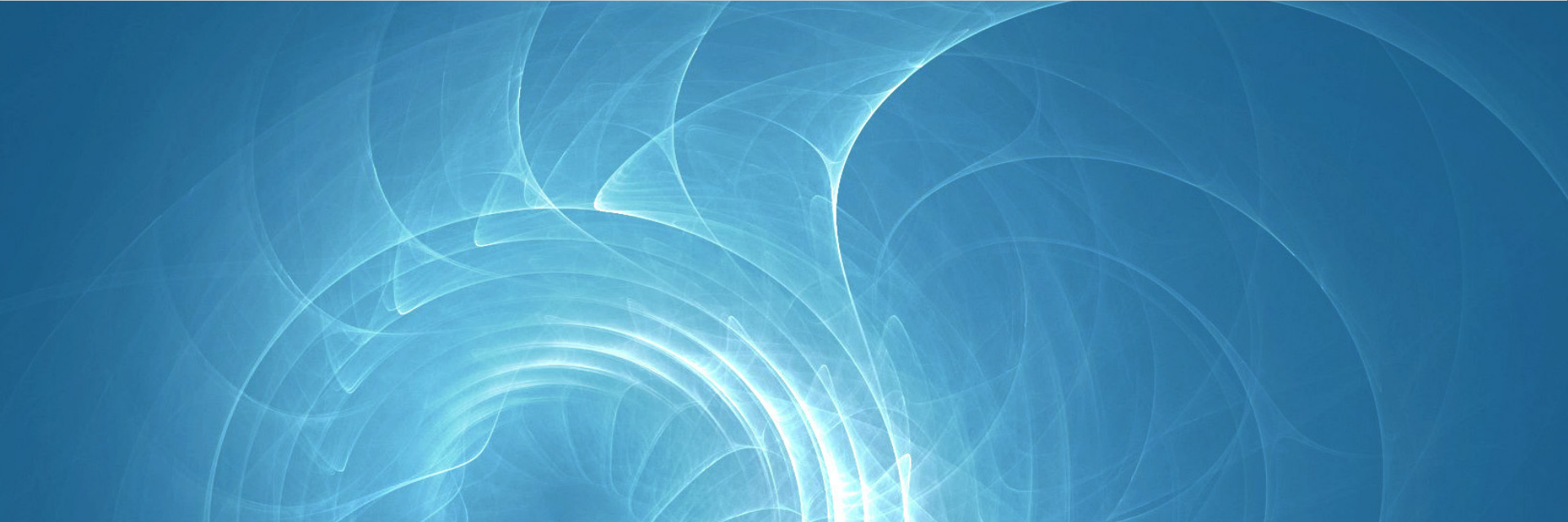
frederic.bertrand@science-metrix.com

www.science-metrix.com

Thank you for your time!

Annexes

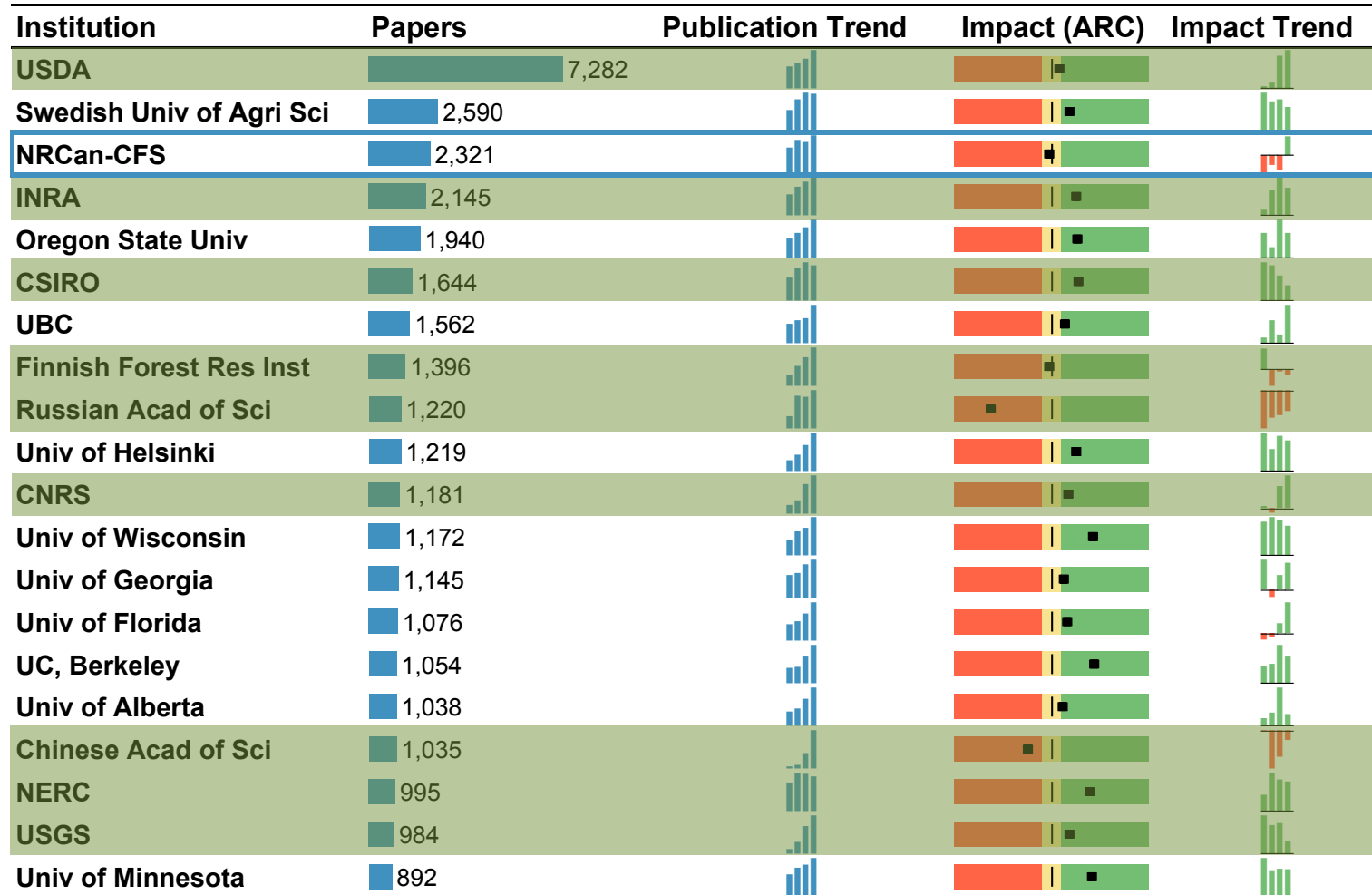
1. Supplementary data
2. Limitations of bibliometrics





Strength of world leading institutions in forestry, 1991–2006

(only the top 20 institutions in absolute output presented)

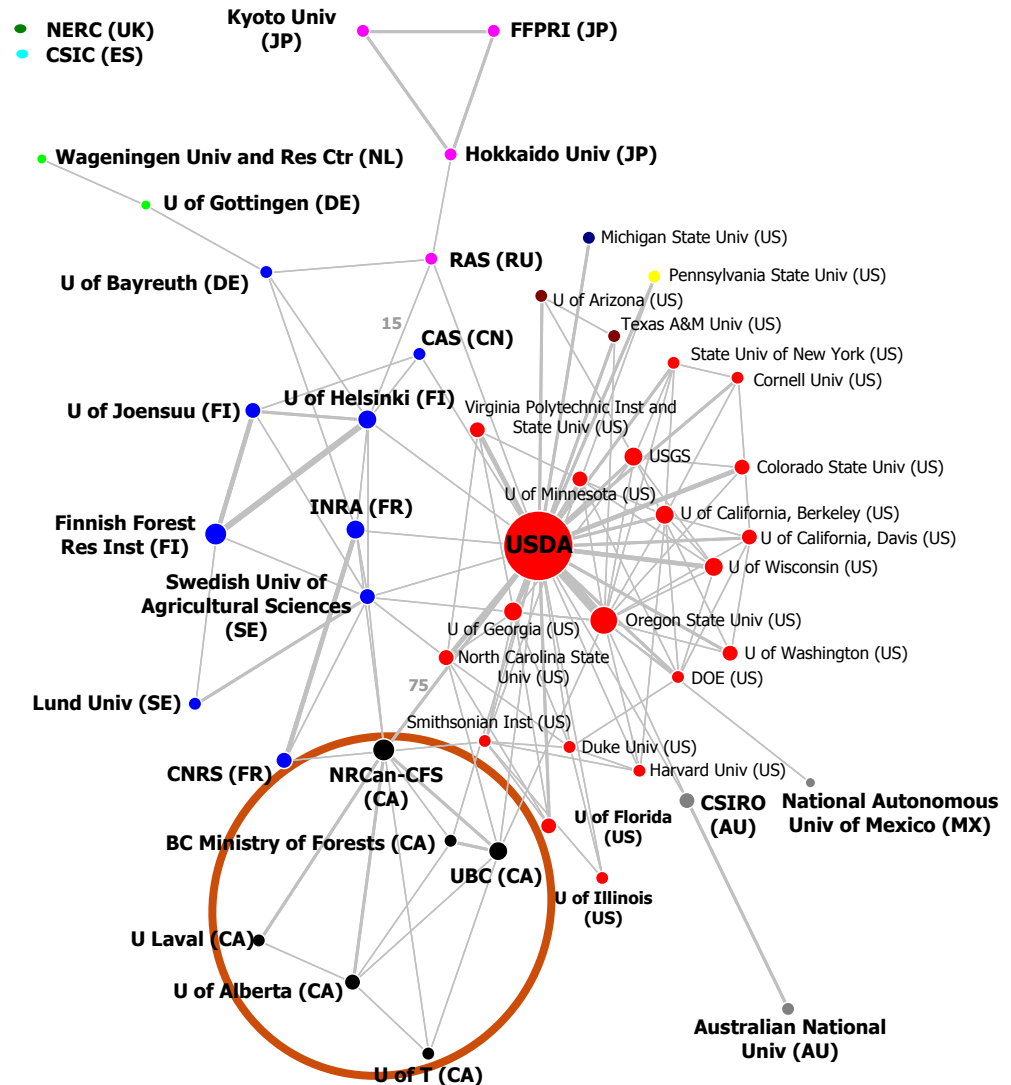




Collaboration network of world leading institutions

(only links representing at least 15 collaborations for 1991–2006 plotted)

- The main node in the network is the USDA (including USFS), with links to all groups
- Circled in red is the Canadian grouping
- Of all Canadian institutions, CFS has the highest number of papers in collaboration
- CFS is also the only Canadian institution linked with both Europe and the US at the threshold used (15)





Limitations of bibliometrics

Internationally, bibliometrics is the most widely accepted method for measuring the outputs of scientific activity. As bibliometrics is often used in the context of performance assessment and management, it is not uncommon to find that comparative approaches using different time periods, organizations, and countries are favoured. The key to comparability is to use a bibliographic database with extensive coverage of the scientific literature over time, countries and scientific domains. In this respect, it is worth mentioning that the WoS, which is used in this study, has some well documented weaknesses:

- It has a slight bias for countries that publish in English-language journals. Thus, for countries whose researchers would tend to publish more in other languages, their scientific production is underestimated in the context of international comparisons. In the NCIC-versus-NCI comparison, the impact is limited, given that researchers of both groups are from Anglo-Saxon countries, namely Canada and the US.
- Another factor affecting publication counts is the difference in publication and citation practices between disciplinary fields. For instance, it is well known that mathematicians publish and cite less than biomedical researchers. Hence, one should not directly compare publications and citation counts between fields.



Limitations of bibliometrics

- Due to differential coverage of sources in Thomson Reuters' scientific databases, bibliometrics indicators are quite reliable for natural sciences, engineering, and health fields but are much less so in social sciences and humanities fields (Archambault et al., 2006; Glänzel and Schoepflin, 1999; Hicks, 2004; Moed, Luwel and Nederhof, 2002; van Raan, 2005). For example, documents (e.g., articles, reviews) published in refereed journals are covered extensively, while books are not. Therefore, it is not surprising to see that publication counts of professors from social sciences and humanities are smaller than those from the natural sciences. In the current study, this limitation has little effect because the main subfields of activity of NCIC- and NCI-supported researchers fall within the fields of clinical medicine and biomedical research, in which refereed journals constitute the core medium for knowledge dissemination.
- Errors in counting the number of papers or citations of an entity (e.g., institution, country) could occur in the WoS (as it could in any other database) due to indexing errors arising from different ways of citing the name of an institution (e.g., Can For Serv, CFS, Canadian Forest Service, Can Forest Service) or to historical changes in an institution's name (e.g., Forestry Canada became the Canadian Forest Service). To limit these types of errors, Science-Metrix analysts spent an appreciable amount of time harmonizing the name of researchers, institutions, and countries in the database.



Limitations of bibliometrics

- Errors in counting the number of papers or citations of an entity (e.g., institution, country) could occur in the WoS (as it could in any other database) due to indexing errors arising from different ways of citing the name of an institution (e.g., Can For Serv, CFS, Canadian Forest Service, Can Forest Service) or to historical changes in an institution's name (e.g., Forestry Canada became the Canadian Forest Service). To limit these types of errors, Science Metrix analysts spent an appreciable amount of time harmonizing the name of researchers, institutions, and countries in the database.

The simplest bibliometric indicator is the number of papers published by an entity (i.e., a researcher, an institution, a country). However, because entities can differ substantially with respect to their levels of resources (e.g., funding, number of researchers, equipment) available to them as they conduct their research, this approach cannot be used to compare the efficiency with which these entities have produced their papers. King (1987) reviewed a number of objections to the use of publication counts for performance assessment in scientific research:

- Social and political pressures, such as those favouring researchers with the highest number of publications in grant or tenure competitions, might affect the emphasis that different entities (i.e., researchers, institutions, countries) will put on publishing results and, therefore, could affect data comparability. These pressures could also lead to undesirable publication practices, such as fragmentation of results in many papers to obtain “least publishable units”.
- Along with the increasing number of multi-authored papers as a result of increased collaboration, it seems that the “gratuitous conferring” of co-authorship is becoming more common. This could lead to what appears to be a better performance by those institutions whose researchers adopt this strategy.



Limitations of bibliometrics

Because all bibliometric indicators have some weaknesses when considered individually, they are better used as a set of indicators for the assessment of scientific performance. When all of the indicators point in the same direction, the results are regarded as being more reliable than those based on a single indicator (King, 1987). It is also important to recognize that bibliometric indicators do not reveal which, among comparables, is the most efficient entity at performing research (or the best at converting research inputs into research outputs), as none of these indicators relates research inputs to research outputs. Compared to data on research outputs, it is very difficult to find comparable data of research inputs.

References:

- Archambault, E., *et al.* (2006). *Scientometrics*, 68(3): 329-342.
- Glänzel, W. & Schoepflin, U. (1999). *Information Processing and Management*. 35: 31-44.
- Hicks, D. (2004). In Moed, H.F., Glänzel, W., Schmoch, U. (Eds.) *Handbook of Quantitative Science and Technology Research* (pp. 476-496). Dordrecht: Kluwer Academic.
- King, J. (1987). *Journal of Information Science*, 13: 261-271.
- Moed, H.F., *et al.* (2002). *Library Trends*. 50: 498-520.
- Van Ran, A. F. J. (2005). *Scientometrics*, 62(1): 133-143.

© 2009 Science-Matrix Inc.