

Fator de difusão: uma medida da difusão do conhecimento através das citações

Sônia Regina Zanotto*
Samile Andréa de Souza Vanz
Ida Regina Chittó Stumpf**

Artículo recibido:
8 de diciembre de 2015
Artículo aceptado:
27 de septiembre de 2016

RESUMO

Este artigo aborda a análise das citações recebidas pelas publicações do Instituto Brasileiro de Geografia e Estatística (IBGE) no período de 2001-2010 sob a ótica da teoria de fatores de difusão do conhecimento. Considerando-se que as citações representam uma maneira de medir o quanto as ideias científicas são difundidas, entende-se que quanto maior o número de diferentes autores, instituições e países citantes maior será o grau de difusão. Identifica 3.158 documentos citantes à produção científica do IBGE. Das 1,272

* Instituto Brasileiro de Geografia e Estatística (IBGE). zanotto.sonia@gmail.com.

** Universidade Federal do Rio Grande do Sul. samilevanz@terra.com.br, irstumpf@ufrgs.br.

instituições de filiações dos autores citantes, 748 (58.80%) são brasileiras ou estão sediadas no Brasil, enquanto 518 (40.72%) são estrangeiras. As instituições voltadas para atividades educacionais lideram (47.96%), seguidas pelas instituições de P&D (22.88%) e as demais ocupam os setores públicos. Em relação aos países, das 7.587 ocorrências de países presentes no vínculo institucional dos autores, 6,168 (81.3%) destas se referem ao Brasil e as demais ocorrências estão distribuídas entre 49 países diferentes situados na América do Sul (82.35%), América do Norte (9.71%), Europa (6.93%), e América Central, Ásia, Oceania, África e Oriente Médio (aproximadamente 1% dos autores citantes). Ao aplicar o fator de difusão com base no estudo proposto Rousseau, Liu e Ye (2012) obteve-se o índice de Gini de 0.62 em 2009 e 0.72 em 2001 e 2003 respectivamente. Conclui que o fator de difusão é satisfatório, ou seja, há um grande número de autores, instituições e países diferentes que citam as publicações do IBGE.

Palavra-chave: Fator de difusão; Difusão do conhecimento; Bibliometria. Análise de citações; Índice de Gini; Informação estatística oficial; IBGE.

Diffusion factor: A measure of knowledge dissemination through citations

*Sônia Regina Zanotto, Samile Andréa de Souza Vanz
and Ida Regina Chittó Stumpf*

ABSTRACT

This paper discusses the analysis of the citations received by the publications of the Brazilian Institute of Geography and Statistics (IBGE) in the period of 2001-2010 under perspective of the theory of diffusion factors. Considering that citations represent a way to measure the spreading of scientific ideas, it is presumable that the larger the number of different authors, institutions and countries citing one idea,

the larger its diffusion is. Identifies 3,158 documents citing IBGE's scientific output. Out of 1,272 institutions which citing authors are affiliated with, 748 (58.80%) are Brazilian institutions or based in Brazil, while 518 (40.72%) are from other countries. Institutions focused on education lead (47.96%) the amount, followed by R&D (22.88%), and the rest (29.16%) is lying on public sectors. A number of 6,168 (81.3%) from 7,587 citation occurrences happens to be of authors with affiliations with Brazilian institutions, with the rest being distributed between the other South American countries (82.35%), North America (9.71%), Europe (6.93%), and Central America, Asia, Oceania, Africa and Middle East (sum is about 1%). Applying the diffusion factors proposed by Rousseau, Liu & Ye (2012), the Gini Index obtained varies from 0.62 in 2009 and 0.72 in 2001 and 2003, respectively. Concludes that the diffusion factor is satisfactory because there is a large number of authors, institutions and different countries that cite IBGE's publications.

Keywords: Diffusion factor. Knowledge dissemination. Bibliometrics. Citation analysis. Gini Index. Official statistical information. IBGE.

RESUMEN

Este artículo aborda el análisis de las citas recibidas por las publicaciones del Instituto Brasileiro de Geografia e Estatística (IBGE) en el período de 2001 a 2010 desde la perspectiva de la teoría de los factores de la difusión del conocimiento. Teniendo en cuenta que las citas son una manera de medir cómo se difunden las ideas científicas, entiéndese que, cuanto mayor es el número de diferentes autores, instituciones y países citantes, mayor será el grado de difusión. Identifica 3,158 documentos citantes a la producción científica del IBGE. De las 1,272 instituciones de afiliación de los autores citantes, 748 (58.80 %) son de Brasil o tienen su sede en Brasil, mientras que 518 (40.72%) son extranjeros. Las insti-

tuciones direccionadas para las actividades educativas lideran (47.96%), seguidas por las instituciones de I&D (22.88 %) y las demás ocupan los sectores públicos. En relación con los países, de las 7,587 ocurrencias de países presentes en el vínculo institucional de los autores, 6,168 (81.3%) de éstas se refieren a Brasil y otras se distribuyen entre 49 países ubicados en América del Sur (82.35%), América del Norte (9.71%), Europa (6.93%), América Central, Asia, Oceanía, África y Medio Oriente (aproximadamente 1% de los autores citados). Al aplicar el factor de difusión, con base en el estudio propuesto por Rousseau, Liu y Ye (2012), se obtuvo el coeficiente de Gini de 0.62 en 2009 y de 0.72 en 2001 y 2003, respectivamente. Concluye que el factor de difusión es satisfactorio, es decir, hay un gran número de autores, instituciones y países diferentes que citan las publicaciones del IBGE.

Palabras clave: Factor de difusión, difusión del conocimiento, bibliometría, Análisis de citas, índice de Gini, Información estadística oficial, IBGE.

INTRODUÇÃO

É consenso entre a comunidade científica que os indicadores bibliométricos devem ser analisados com cautela e interpretados dentro de um determinado contexto. Isto porque são incompletos em sua capacidade de mensuração e geralmente apresentam resultados isolados que se referem a um único aspecto do problema analisado. O índice de imediatez, a meia vida e o fator de impacto, por exemplo, são medidas de impacto das citações que vem sendo aplicadas há vários anos para mensurar a forma como o periódico recebe citações para seus artigos ao longo do tempo. Nenhuma dessas medidas, no entanto, se preocupa em saber quais são as fontes citantes, como e qual é a dispersão das citações, sendo que, de certa maneira, a dispersão geográfica das fontes citantes representa a extensão do impacto geográfico daquela informação. Entende-se, portanto, que é necessária uma medida que mostre, além do impacto, a dimensão da recepção de uma informação entre a comunidade (Rowlands, 2002).

O Journal Diffusion Factors (Fator de Difusão de Periódicos) foi introduzido em 2002 por Rowlands para medir a influência na pesquisa científica

e a dimensão da difusão dos periódicos, como uma tentativa de complementar o Fator de Impacto (Rowlands, 2002). Desde então, vários pesquisadores vem desenvolvendo medidas de difusão através de diferentes técnicas de coleta de dados. No Brasil, Rummler (2006) apresentou o Índice de Dispersão Segmentar (IDS), indicador que pode ser aplicado a uma obra, um autor, um periódico ou área de conhecimento, e considerou a possibilidade de dimensionar a extensão de impacto de uma unidade de análise, à medida que seja aplicado à análise de citações.

Os resultados científicos são difundidos de um campo para outro, do laboratório para o artigo, da ciência para a tecnologia, e da tecnologia para a sociedade, através de conversas face-a-face, artigos, conferências científicas entre outros meios (Rousseau, Liu e Ye, 2012), caracterizando o processo de comunicação científica, a transferência de conhecimentos. Neste contexto, Rousseau, Liu e Ye (2012) apresentaram a concepção de que as ideias científicas fluem através de sistemas em camadas (*layered system*) e as citações são as unidades de análises, em conjunto com os autores citantes, instituições e países de afiliação destes autores. Para os autores, a difusão do conhecimento está relacionada à forma como o artigo é recebido pela comunidade científica e a difusão é uma medida de alcance, determinada pelos artigos citantes e operacionalizada através das citações (Liu, Rafols e Rousseau; 2012; Liu, Rousseau e Guns, 2013).

A difusão do conhecimento produzido pelo Instituto Brasileiro de Geografia e Estatística (IBGE) é evidenciada através de artigos e demais publicações científicas que citam as informações estatísticas oficiais produzidas pelo Instituto. Estas citações estão distribuídas em dezenas de periódicos científicos, em diversos idiomas e publicadas em vários continentes. Alguns aspectos da difusão e da distribuição geográfica das informações estatísticas oficiais no Brasil foram estudados por Zanotto (2011) e por Zanotto, Vanz e Stumpf (2011). Nestes estudos evidenciou-se a importância do IBGE como a principal instituição de produção e coordenação do Sistema Nacional de Informações Estatísticas do Brasil, cuja produção segue padrões, metodologias e princípios fundamentais recomendados internacionalmente e entre os quais se destacam a imparcialidade, a ética, a transparência, a igualdade de acesso, a eficiência e a confidencialidade.

O sítio na internet é o principal repositório de informações do IBGE, onde as atualizações do Portal IBGE e dos bancos de dados são realizadas em conjunto com a edição de produtos impressos, como anuários estatísticos, revistas, relatórios técnicos, mapas e demais produtos das áreas de estatísticas e de geociências. O sítio também disponibiliza produtos especiais para o público infantil e para professores, além de panoramas acerca do perfil

socioeconômico dos 5,570 municípios brasileiros e respectivas Unidades da Federação, e ainda, ferramentas de mapas interativos. O sítio hospeda a biblioteca digital onde é possível consultar e realizar downloads de publicações históricas e atualizadas. Além do portal na Internet o IBGE divulga através da Sala de Imprensa todos os releases das publicações, assim como utiliza as redes sociais *Facebook*, *Twitter*, *Instagram* e *Youtube* para divulgar notícias sobre as informações produzidas e vídeos explicativos (IBGE, 2016).

No presente trabalho apresenta-se a teoria e a metodologia para cálculo do fator de difusão proposta por Rousseau, Liu e Ye (2012) e a sua aplicação prática às citações feitas às publicações do IBGE, como as relacionadas acima, de forma que torna aplicável esta metodologia a um caso real.

FATOR DE DIFUSÃO

O Fator de Difusão de Periódicos foi apresentado por Rowlands em 2002 na tentativa de mensurar a transdisciplinaridade e a influência dos periódicos em várias áreas. Pois a dinâmica de como as ideias são transferidas de um autor para outro e de uma disciplina para outra é importante para melhor entender a comunicação científica. Rowlands (2002) comparou metaforicamente as citações recebidas por um periódico ao longo do tempo com o ato de jogar seixos em uma lagoa, afirmando que as medidas bibliométricas padrão utilizadas, a saber, fator de impacto, índice de imediatez e meia vida, apresentam o impacto das citações, ou seja, o splash do seixo na lagoa. O fator de impacto quantifica as citações em um período de tempo (alusivamente, mede o tamanho do seixo); o índice de imediatez representa a curva inicial da citação (ou a aceleração do seixo no ar até encontrar água); e a meia vida indica o declínio da curva (o tempo necessário para que a água da lagoa volte ao estado inicial, completamente parada). Observa-se que nenhuma destas medidas mede o tamanho da recepção do seixo pela lagoa, ou seja, metaforicamente, nenhum indicador mede a recepção de um periódico pela comunidade científica.

Desta forma, Rowlands (2002) propôs uma medida da extensão das ondulações resultantes das citações a uma determinada publicação. Em seus primeiros estudos, o *Journal Diffusion Factors* (nome original em inglês para Fator de Difusão de Periódicos) pôde ser entendido como a medida das características das citações a um periódico e foi introduzido para complementar o Fator de Impacto. O Fator de Difusão de Periódico resultava do número de citações recebidas dividido pelo número de diferentes periódicos citantes em um dado período. A janela de citação (período) pode ser variada, assim como os métodos síncronos e diacrônicos podem ser empregados para análise.

De acordo com a fórmula proposta por Rowlands (2002), o fator de difusão do periódico A, da área de economia, que recebeu 1.000 citações de 100 periódicos diferentes, é 10. Para ter o mesmo fator de difusão, o periódico B, que recebeu 10.000 citações, precisaria recebê-las de 1.000 periódicos diferentes, o que é pouco provável na área de economia em função do tamanho da área e do número de periódicos. Ao fazer tal observação, Frandsen (2004) demonstrou que o Fator de Difusão de Periódicos é fortemente influenciado pelo número de citações que o periódico recebe, levando um periódico altamente citado a obter um baixo fator e por outro lado, um periódico pouco citado a receber um alto fator de difusão. O autor propôs uma alteração na fórmula, substituindo-se o número de citações recebidas pelo número de artigos publicados pelo periódico. Desde então os estudos de fatores de difusão buscam capturar diferentes aspectos do processo de citação, avançando além das tradicionais medidas bibliométricas para os periódicos ou o conjunto de artigos (2006). Em recente artigo, Rousseau, Liu e Ye (2012) consideraram que, mesmo que as ideias científicas sejam transmitidas através de conversas formais e informais, e-mail, blogs e outros meios eletrônicos de comunicação, estas ideias normalmente convergem para um artigo científico e que o conhecimento contido em um artigo é difundido através das citações. A difusão de ideias científicas é hoje um fator a ser considerado na determinação da inovação.

A proposta de Rousseau, Liu e Ye (2012) abrange um sistema de contagem em camadas do conjunto de todas as citações, autores citantes, instituições e países de filiação dos autores. É claro que se um artigo recebe citações oriundas de múltiplos países então o conteúdo desse artigo é difundido mais do que se ele tivesse recebido o mesmo número de citações advindas de um mesmo país. Ou seja, a ideia é que quanto mais autores diferentes, melhor; e quanto mais instituições e países diferentes citantes, melhor. O foco, portanto, não é o número total de citações e sim o número total de diferentes fontes citantes. “Diferente” é a palavra-chave quando o foco é a difusão de ideias contidas em artigos científicos.

A aplicação do fator de difusão pode ser feita para análise de um único artigo, ou então, do conjunto dos artigos publicados por um pesquisador, ou em uma área, ou por um grupo de pesquisa, ou por um país (2012). O método proposto por Rousseau, Liu e Ye (2012), atribui 1 como o crédito máximo de contribuição, e a contagem deve ser fracionada de forma igual. Pode-se ainda utilizar outros métodos de contagem sem alteração da proposta. Quanto à autocitação, os autores preferem mantê-la na contagem, mas não é uma obrigatoriedade. A medida de distribuição de Gini é usada como ferramenta estatística para equacionar o problema, como apresentado na *Figura 1*:

Figura 1. Coeficiente de Gini conforme Rousseau, Liu e Ye (2012)

$$G_e(X) = \frac{2 \sum_{j=1}^N jX_j}{\mu N^2} - \frac{1}{N}$$

Fonte: Rousseau, Liu e Ye (2012).

Nota: $G_e(X)$ = coeficiente de Gini para o período;

$X = (x_j)$ e $j = 1, \dots, n$ (matriz de números não negativos).

jX = Variação das posições do período.

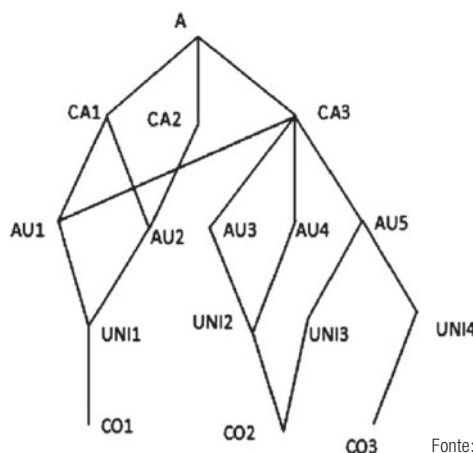
N = Número de camadas analisadas.

μ = Média do conjunto $\{x_j\}$.

O Coeficiente de Gini é uma medida de desigualdade desenvolvida pelo estatístico italiano Corrado Gini, em 1912. É comumente utilizado para calcular a desigualdade de distribuição de renda, entretanto, pode ser usado em outras distribuições. Este índice consiste em um número entre 0 e 1, onde 0 corresponde à completa igualdade de renda (ou seja, existe menor concentração, quando todos têm a mesma renda) e 1 corresponde à completa desigualdade (ou seja, existe maior concentração, quando uma pessoa tem toda a renda, e as demais nada têm). O índice de Gini é o coeficiente expresso em pontos percentuais (é igual ao coeficiente multiplicado por 100) (Sandroni, 1996).

O esquema a seguir representa o método de organização dos indicadores em camadas para análise dos fatores de difusão: citações, autores citantes, instituição de filiação e país de filiação do autor citante.

Figura 2. Esquema do sistema em camadas para cálculo do Fator de Difusão



Fonte: Rousseau, Liu e Ye (2012).

As camadas analisadas correspondem, conforme *Figura 2, a*: Camada 1 (CA): ao conjunto de todas as citações (recebidas pelo artigo A) numa janela de citação dada; Camada 2 (AU): ao conjunto de todos os autores citantes da amostra; Camada 3 (UNI): ao conjunto de todas as instituições a que os autores citantes pertencem; e Camada 4 (CO): ao conjunto de todos os países aos quais pertencem essas instituições. A *Tabela 1* apresenta o método de contagem em camadas para um dado artigo:

Tabela 1. Exemplo da contagem fracionada das camadas de Autor, de Instituição e de País

Autores (Au)	Artigo 1 (CA1)	Artigo 2 (CA2)	Artigo 3 (CA3)	Soma
Autor 1	0,5		0,25	0,75
Autor 2	0,5	1		1
Autor 3			0,25	0,25
Autor 4			0,25	0,25
Autor 5			0,25	0,25
<i>Peso Autor</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>2,50</i>
Universidades (Uni)	Artigo 1 (CA1)	Artigo 2 (CA2)	Artigo 3 (CA3)	Soma
Universidade 1	1	1	0,25	1
Universidade 2			0,5	0,5
Universidade 3			0,125	0,125
Universidade 4			0,125	0,125
<i>Peso Universidade</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1,75</i>
Países (Co)	Artigo 1 (CA1)	Artigo 2 (CA2)	Artigo 3 (CA3)	Soma
País 1	1	1	0,25	1
País 2			0,625	0,625
País 3			0,125	0,125
<i>Peso País</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1,75</i>

Fonte: Rousseau, Liu e Ye (2012).

Nota das autoras: A soma horizontal de cada unidade (Artigo, Instituição, País) tem peso máximo = 1. O termo universidade foi mantido do texto original (University), no entanto, pode-se entender esta unidade como Instituição de filiação do autor.

No exemplo acima um Artigo A é citado por 3 outros artigos (CA), a soma da fração de autoria (AU) é 2.5; levando-se em conta que a contagem de cada artigo citante contribui no total de 1 unidade de autor e se houver mais de um autor, as contribuições de cada um são fracionadas em partes iguais ($1/n$). A soma da fração das instituições (UNI) e países (CO) é respectivamente 1.75, ambas são resultados derivados das contribuições de autores. Então, cada artigo citante contribui com um total de exatamente 1 unidade (universidade/instituto ou país) e se existirem diferentes autores em digamos n instituições, cada autor recebe uma contribuição de ($1/n$) e, se no entanto, um autor tem k endereços institucionais, cada instituto contribui através deste

autor, uma pontuação de $(1/nk)$. O somatório horizontal não pode exceder a um tendo em vista que o peso máximo de cada unidade, por exemplo, um autor, uma instituição ou um país é um. Finalizando a análise, a média da soma destas frações é $\mu = 2,25$, temos $N = 4$; $X = (3; 2.5; 1.75; 1.75)$ e, aplicada a fórmula chega-se ao índice de Gini ponderado como $Ge=0.875$.

METODOLOGIA

Para a aplicação da teoria de fatores de difusão segundo a metodologia de Rousseau, Liu e Ye (2012) foram necessários os procedimentos metodológicos descritos a seguir.

As citações à produção científica do IBGE foram identificadas na Web of Science através da opção de busca avançada das referências citadas (Cited Reference Search). A partir da tela de recuperação, no campo de autor citado (Cited Author), utilizou-se a expressão de busca ao IBGE construída pela agregação das diversas variantes da sigla e do nome completo da Instituição, em inglês e português, com o uso de operador booleano “OR” como apresentado abaixo:

IBGE OR (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA) OR (Brazilian Geography and Statistics Institute) OR (BRAS FUND I BRAS OR BRAS GE OR BRAS GEOGR EST OR BRAS IBGE OR BRAS MIN AGR I BR OR BRASIL F I GEOGR OR BRASIL FUND IBGE OR BRASIL I BRAS GEO OR BRASIL IBGE OR BRAZIL F I BRAS G OR BRAZIL F I GEOGR OR BRAZIL FIBGE OR BRAZIL FIBGE OR BRAZIL FUND I BRA OR BRAZIL FUND I GEO OR BRAZIL FUND IBGE OR BRAZIL I BRAS ES OR BRAZIL I BRAS GEO OR BRAZIL I BRAZ GEO OR BRAZIL I GEOG EST OR BRAZIL IBGE OR CENS DEM OR CENSO OR FUND I OR FUND IBGE OR FUND INST BRAS GE OR FUNDACAO IBGE* OR IBGE INSTITUTO OR PESQ NAC AM DOM OR PESQ NAC POR AM D OR PESQ NAC SAUD NUT OR PESQ PEC MUN OR PNAD).

Ainda na tela de recuperação, no campo de definição do período de busca (Timespan), foi incluído o período de 2001 até 2009 (ano a ano) para a coleta em 2010 e as citações no ano de 2010 para a coleta realizada em junho de 2011. As duas etapas de coleta de dados foram feitas nas três coleções, a saber, Science Citation Index, Social Sciences Citation Index e o Humanities and Arts Citation Index.

Com o objetivo de garantir a qualidade das análises foram necessários procedimentos de limpeza e padronização dos dados coletados. No total, foram excluídos 43 documentos que não correspondiam ao levantamento

correto, por não apresentarem citações ao IBGE ou estarem fora do período de análise. Os procedimentos incluíram a correção e a padronização dos nomes de autores e das respectivas instituições de filiação, que procurou eliminar a duplicidade de entradas causada pelos erros de digitação e grafia diversas das entradas de nomes de pessoas, dos departamentos, faculdades, hospitais, laboratórios, empresas, entidades governamentais, entre outras instituições.

Após, os dados foram analisados com o uso do *software* BibExcel³ e do programa de planilha eletrônica Microsoft Excel 2007. No conjunto dos dados recuperados foram isoladas as informações de autoria contidas no campo AU e contabilizadas ano a ano, e as respectivas informações de instituição e países contidas no campo C1 de endereçamento.

Na *Figura 3*, com objetivo de exemplificar o método de contagem conforme a metodologia de Rousseau, Liu e Ye (2012), foram relacionados três registros bibliográficos importados da base de dados Web of Science. Posteriormente são apresentados os dados fracionados em planilha. Na exemplificação foram mantidos os campos AU que corresponde à autoria; ti ao Título do artigo citante; SO ao Título do periódico citante; C1 ao campo com o endereçamento dos autores, de onde se extraem as informações de instituição e país de filiação; os demais campos de descrição/identificação dos registros bibliográficos originais foram suprimidos.

Figura 3. Exemplo de Documentos Citantes de 2010

Registro 1 de 2010
Au Gioia, Simone M.C.L. Babinski, Marly Weiss, Dominik J. Kerr, Americo A.F.S. TI Insights into the dynamics and sources of atmospheric lead and particulate matter in Sao Paulo, Brazil, from high temporal resolution sampling SO ATMOSPHERIC RESEARCH, VL 98, IS 2-4, BP 478-485, PY 2010 C1 [Gioia, Simone M. C. L.; Babinski, Marly] Univ Sao Paulo, Inst Geociencias, Ctr Pesquisas Geocronol, BR-05508080 Sao Paulo, Brazil. [Weiss, Dominik J.] Univ London Imperial Coll Sci Technol & Med, London SW7 2AZ, England. [Weiss, Dominik J.] Nat Hist Museum, London SW7 5PD, England. [Kerr, Americo A. F. S.] Univ Sao Paulo, Inst Fis, Dept Fis Aplicada, BR-05508080 Sao Paulo, Brazil.

3 BibExcel é um software desenvolvido por Olle Persson (Umeå University, Suécia), que possui um conjunto de ferramentas que auxiliam na análises bibliométricas a partir de dados bibliográficos. Disponível em: <<http://www8.umu.se/inforsk/Bibexcel/>>. Acesso em 10 jun. 2010.

Registro 2 de 2010

Au Michalski, Fernanda
Metzger, Jean Paul
Peres, Carlos A.

TI Rural property size drives patterns of upland and riparian forest retention in a tropical deforestation frontier

SO GLOBAL ENVIRONMENTAL CHANGE-HUMAN AND POLICY DIMENSIONS VL 20 IS 4 BP 705-712, PY 2010

C1 [Michalski, Fernanda; Metzger, Jean Paul] Univ Sao Paulo, Biosci Inst, Dept Ecol, BR-05508900 Sao Paulo, Brazil.
[Michalski, Fernanda] Inst Procarvoros, BR-12940970 Atibaia, SP, Brazil.
[Peres, Carlos A.] Univ E Anglia, Sch Environm Sci, Ctr Ecol Evolut & Conservat, Norwich NR4 7TJ, Norfolk, England.

Registro 3 de 2010

Au Norris, Darren
Michalski, Fernanda
Peres, Carlos A.

TI Habitat patch size modulates terrestrial mammal activity patterns in Amazonian forest fragments

SO JOURNAL OF MAMMALOGY, VL 91 IS 3 BP 551-560 PY 2010

C1 [Norris, Darren; Michalski, Fernanda] Univ Estadual Paulista, Dept Ecol, Lab Biol Conservacao, BR-13506900 Sao Paulo, Brazil.
[Michalski, Fernanda] Inst Procarvoros, BR-12945010 Sao Paulo, Brazil.
[Peres, Carlos A.] Univ E Anglia, Sch Environm Sci, Ctr Ecol Evolut & Conservat, Norwich NR4 7TJ, Norfolk, England.

Legenda:

AU: Campo autoria;

ti: Título do artigo citante;

so: Título do periódico citante;

C1: Campo com o endereçamento dos autores, de onde se extraem as informações de instituição e país de filiação dos autores;

Nota: Por se tratar de um recurso de exemplificação os demais campos de descrição/identificação dos registros bibliográficos originais foram suprimidos.

Os dados observados na *Figura 3* foram contabilizados conforme demonstrado

A partir do exemplo da *Figura 3* procede-se o cálculo fracionado, conforme é apresentado na *Tabela 2*:

Tabela 2. Contagem fracionada das camadas de Autor, de Instituição e de País dos registros 1-3 de 2010

Autores (Au)	Artigo 1 (CA1)	Artigo 2 (CA2)	Artigo 3 (CA3)	Soma
Gioia, Simone M.C.L.	0.25			0.25
Babinski, Marly	0.25			0.25
Weiss, Dominik J.	0.25			0.25
Kerr, Americo A.F.S.	0.25			0.25
Michalski, Fernanda		0.33	0.33	0.66
Metzger, Jean Paul		0.33		0.33
Peres, Carlos A.		0.33	0.33	0.66
Norris, Darren			0.33	0.33
<i>Peso Autor</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>2,98</i>
Universidades/Instituições (uni)	Artigo 1 (CA1)	Artigo 2 (CA2)	Artigo 3 (CA3)	Soma
Universidade de São Paulo	0.5	0.33		0.83
University of London	0.25			0.25
Natural History Museum	0.25			0.25
Instituto Pró-Carnívoros		0.33	0.33	0.66
University of East Anglia		0.33	0.33	0.66
Universidade Estadual Paulista			0.33	0.33
<i>Peso Universidade</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>2,98</i>
Países (co)	Artigo 1 (CA1)	Artigo 2 (CA2)	Artigo 3 (CA3)	Soma
Brazil	0.5	0.66	0.66	1
England	0.5	0.33	0.33	1
<i>Peso País</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>2</i>

Fonte: Dados da pesquisa.

Nota: Cada unidade (Autor, Universidade/instituições, País) tem peso máximo = 1.

Finalizando a exemplificação, um conjunto de artigos teriam sido citados por 3 outros artigos (CA), a soma da fração de autoria citante (AU) é 2.98, e a soma da fração das instituições (UNI) e países citantes (CO) é respectivamente 2.98 e 2. A média da soma destas frações é $\mu = 2.74$, tem-se $N = 4$ e chega-se ao índice de Gini ponderado $Ge=0.93$. Como o resultado está mais próximo a 1 do que a zero, pode-se considerar um ótimo grau de difusão.

RESULTADOS E DISCUSSÃO

No período compreendido entre os anos 2001 a 2010 o IBGE publicou mais de 99 títulos diferentes e recebeu 3.985 citações. Na *Tabela 3* abaixo estão relacionados os 10 produtos informacionais produzidos pelo IBGE mais citados.

Tabela 3. Publicações do IBGE mais citadas no período 2001-2010

Publicação citada	Tema	Frequência	%	Σ %
Censo demográfico	População	887	22.26	22.26
Pesquisa nacional por amostra de domicílios (PNAD)	Trabalho e rendimento	307	7.70	29.96
Pesquisa de orçamentos familiares (POF)	Orçamento familiar	205	5.14	35.11
Levantamento de recursos naturais RADAMBRASIL	Recursos naturais e meio ambiente	170	4.27	39.37
Censo agropecuário	Agropecuária	166	4.17	43.54
Anuário estatístico do Brasil	Informações socioeconômicas e geográficas	155	3.89	47.43
Sistema IBGE de Recuperação Automática (SIDRA)	Informações socioeconômicas e geográficas	138	3.46	50.89
Levantamento sistemático da produção agrícola (LSPA)	Agropecuária	105	2.63	53.53
Produção agrícola municipal (PAM)	Agropecuária	92	2.31	55.83
Contagem da população	População	76	1.91	57.74
Demais publicações	—	1,684	42.26	100.00
<i>Total</i>	—	<i>3,985</i>	<i>100</i>	—

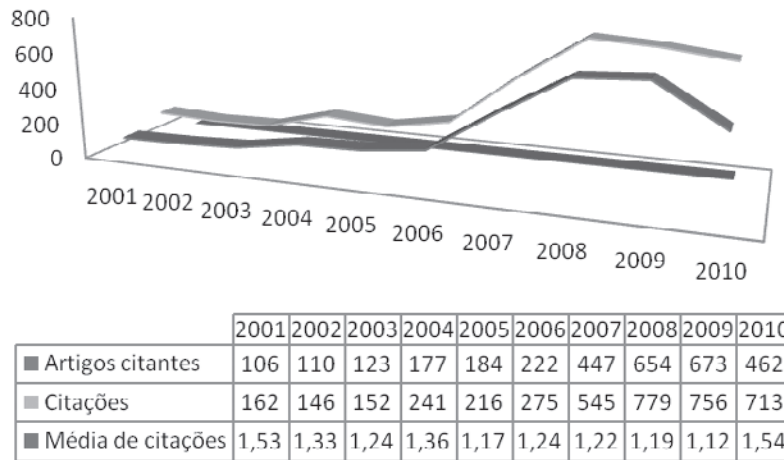
Fonte: Dados da pesquisa.

Observase que os 10 documentos mais citados correspondem a 57.74% das citações recebidas, ficando o restante (42.26) 7% distribuídos nos demais documentos. O primeiro mais citado, que responde por pouco mais de 20% das citações é o Censo Demográfico, uma publicação de reconhecida importância, que apresenta dados populacionais do Brasil desde os primórdios do Estado brasileiro. Os dados do Censo tem periodicidade decenal e são em forma de tabelas sobre idade, sexo, deficiência, cor ou raça, religião, educação, fecundidade, mortalidade, nupcialidade, migração, trabalho e rendimento da população brasileira; sobre o número de componentes e características da família brasileira; dados sobre os domicílios e sua situação

em relação à localização, disponibilidade de serviços de água, esgotamento sanitário, eletricidade, além de enumerar a existência de bens duráveis e características dos responsáveis pelos domicílios (IBGE, 2013). As demais fontes de informação relacionadas na tabela acima, como o SIDRA, por exemplo, perfazem percentuais de participação menores o que poderia ser interpretado como indicador de menor importância entre os documentos produzidos pelo IBGE. Contudo, a evidência de que essas fontes são citadas pelos acadêmicos na produção de novos conhecimentos torna as pesquisas e os levantamentos estatísticos produzidos pelo IBGE de grande relevância.

Dentre as 3,985 citações, identificaram-se 3,158 documentos citantes à produção científica do IBGE. Quando foram analisados os autores destes documentos, encontrou-se 10,707 nomes diferentes, perfazendo uma média de 1.29 citações por autor e frequência que variou de 1 a 19 citações às publicações do IBGE por autor. Na *Figura 4* pode-se observar o movimento crescente dos documentos citantes no decorrer do período, da mesma forma o número de citações e a média de citações por ano.

Figura 4. Artigos e número de citações às Publicações do IBGE na WoS em 2001-2010



No período de 2001 a 2010 observa-se um movimento crescente de documentos citantes ao longo da década, com 106 artigos citantes em 2001 e chegando ao número maior em 2009 com 674. Da mesma forma cresceu o número de citações no período e a média de citações por ano manteve-se entre 1.53 em 2001 e 1.54 em 2010.

Das 1.272 instituições de filiação dos autores citantes classificadas conforme a atividade principal de acordo com a Classificação Nacional de Atividade Econômica (cnae)⁴ apresentadas na *Tabela 3*, 748 (58.80%) são de abrangência nacional, ou seja, são brasileiras ou estão sediadas no Brasil, enquanto 518 (40.72%) são estrangeiras sediadas em outros países. Quando aplica-se o critério de atividade principal para cada uma das instituições, destacam-se as instituições voltadas para atividade de Educação superior (47.96%), seguidas das instituições voltadas para a Pesquisa e desenvolvimento científico (22.88%), seguidos pelos órgãos da Administração pública em geral (13.79%) como as Secretarias municipais de saúde, por empresas de Atividades de atenção à saúde humana (7.78%), como os hospitais e clínicas médicas, por empresas voltadas para atividades de agricultura, pecuária e serviços relacionados (1.73%), por indústrias de transformação (1.57%), por entidades voltadas para atividades de organizações associativas patronais, empresariais e profissionais (1.18%), por companhias de eletricidade e gás (Ex. Eletronorte), da construção (Ex. Águas de Limeira SA), e instituições museológicas, com atividades de exploração de lugares, prédios históricos e atrações similares (Ex.: American Museum of Natural History), atividades financeiras e serviços relacionados (Ex.: bndes), entre outras.

Tabela 4. Instituições de filiação dos autores citantes do IBGE na WoS no período de 2001/2010 segundo atividade econômica principal

Atividade principal	Cód. CNAE*	Número de unidades				Frequência			
		Nacional	Estrangeira	Total	%	Nacional	Estrangeira	Total	%
Educação superior	P 853	294	316	610	47.96	4,780	1.009	5,789	76.49
Pesquisa e desenvolvimento científico	M 72	163	128	291	22.88	900	227	1.127	14.89
Administração pública em geral	O 8411-6	143	26	169	13.29	285	54	339	4.48
Atividades de atenção à saúde humana	Q 86	78	21	99	7.78	132	30	162	2.14

Atividade principal	Cód. CNAE*	Número de unidades				Frequência			
		Nacional	Estrangeira	Total	%	Nacional	Estrangeira	Total	%
Indústria de transformação	Seção C	11	9	20	1.57	12	13	25	0.33
Atividades de organizações associativas patronais, empresariais e profissionais	S 941	9	6	15	1.18	10	9	19	0.25
Organismos internacionais e outras instituições extraterritoriais	U 99	—	—	6	0.47	—	—	22	0.29
Eletricidade e Gás	Seção D	5	0	5	0.39	7	0	7	0.09
Construção	Seção F	5	0	5	0.39	6	0	6	0.08
Atividades de museus e de exploração de lugares e prédios históricos e atrações similares	R 9102-3	2	8	10	0.79	11	11	22	0.29
Atividades Financeiras e serviços relacionados	Seção K	1	0	1	0.08	1	0	1	0.01
Outros	—	2	0	2	0.16	2	0	2	0.03
<i>Total</i>	—	<i>748</i>	<i>518</i>	<i>1,272</i>	<i>100.00</i>	<i>6,188</i>	<i>1,358</i>	<i>7,568</i>	<i>100.00</i>
<i>%</i>	—	<i>58.80</i>	<i>40.72</i>	<i>100.00</i>	—	<i>81.76</i>	<i>17.95</i>	<i>100.00</i>	—

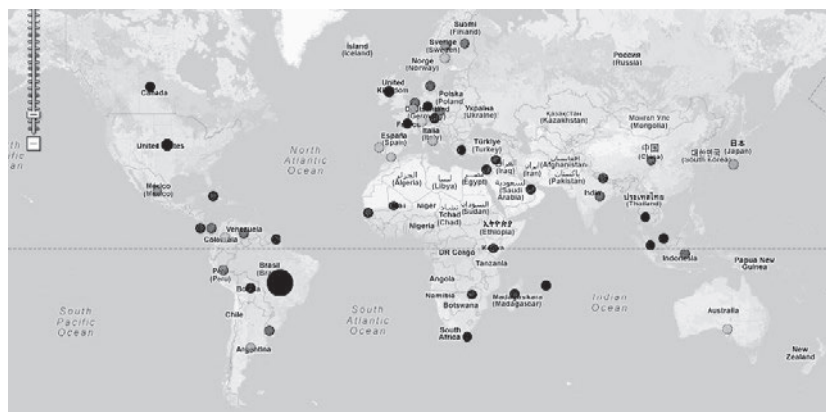
* Nota: As instituições foram agrupadas de seções até subclasses conforme a Classificação cnae.

As instituições de abrangência internacional estão caracterizadas pelos órgãos de desenvolvimento e por seus departamentos e representações regionais, como a Organização das Nações Unidas (ONU), a Comissão Europeia (UE) entre outras e representam menos de 1% das ocorrências. Analisar a atividade principal das instituições citantes revela a importância dos documentos produzidos pelo IBGE para diversos setores da sociedade brasileira

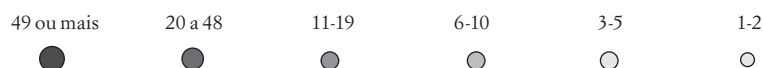
e internacional, ou seja, a informação gerada pelo IBGE é relevante e subsidia a pesquisa científica.

As instituições que citaram a produção científica do IBGE são sediadas em diversos países. Observouse que das 7,587 ocorrências de países presentes no vínculo institucional do campo de localização dos autores, 6,168 (81.30%) destas ocorrências são para o Brasil e as demais estão distribuídas para 49 países diferentes. A distribuição das instituições e autores entre diferentes países favorece a disseminação das informações estatísticas produzidas pelo IBGE. Observou-se que 82.35% das citações são provenientes de 9 países da América do Sul (Argentina, Bolívia, Brasil, Colômbia, Guiana Francesa, Panamá, Peru, Uruguai e Venezuela); 9.71% da América do Norte, representada pelo Canadá e USA; 6.93% das citações foram feitas por países Europeus (Alemanha, Áustria, Bélgica, Dinamarca, Espanha, Finlândia, França, Grécia, Holanda, Itália, Noruega, Portugal, Reino Unido, República Tcheca, Suécia e Suíça). Outros países de diferentes continentes somam juntos aproximadamente 1% das citações: Costa Rica, Cuba e México localizados na América Central; 8 países da Ásia (China, Índia, Indonésia, Japão, Malásia, Nepal, Singapura e Tailândia); Austrália e Nova Zelândia na Oceania; 7 países localizados no continente Africano (África do Sul, Quênia, Madagascar, Mali, Maurício, Senegal, Zimbábue); Israel, Síria e Emirados Árabes no Oriente Médio.

Figura 5. Infográfico da distribuição geográfica dos países de filiação dos autores que citaram a produção científica do IBGE na WoS no período 2001-2010



Legenda: frequência das citações



Os dados apresentados mostram que as publicações do IBGE receberam um grande número de citações no período, advindas de diferentes autores, instituições e países. Com o objetivo de comprovar que estas citações refletem a difusão satisfatória do conhecimento produzido pelo IBGE buscou-se na literatura o fator de difusão equacionado por Rousseau, Liu e Ye (2012). A partir do conjunto das citações recebidas pelas publicações do IBGE no período de 2001-2010, conforme descrito anteriormente, e a aplicação da análise dos fatores de difusão propostos por Rousseau, Liu e Ye (2012), obteve-se os dados da *Tabela 4*:

Tabela 5. Produção científica do IBGE - Artigos Citantes, Autores, Instituições e Países – 2001-2010

Ano	Artigos Citantes(ca)	Soma Frações Autores Citantes(au)	Soma Frações Instituições Citantes(uni)	Soma Frações Países Citantes(co)	G_e
2001	106	105.49	70.02	8.58	0.72
2002	110	109.24	69.39	7.85	0.71
2003	123	121.76	75.53	12.17	0.72
2004	177	176.00	95.51	10.07	0.68
2005	184	182.25	89.74	10.07	0.67
2006	222	221.22	118.88	16.81	0.69
2007	447	444.96	179.66	17.86	0.64
2008	654	638.16	235.75	20.54	0.63
2009	673	671.42	233.50	18.23	0.62
2010	462	460.72	203.66	21.32	0.66

Fonte: Dados da pesquisa.

Após aplicação da fórmula G_e obteve-se resultados entre 0,62 em 2009 e 0,72 em 2001 e 2003, indicando um fator de dispersão satisfatório, ou seja, a existência de um grande número de autores, instituições e países diferentes que citam as publicações do IBGE.

Outra consideração importante é a de que, embora o número absoluto de citações seja variável e crescente no período, como em 2008 e 2009 com 654 e 673 artigos citantes, o índice G_e mantém-se entre 0.63 e 0.62 respectivamente. Sendo que em 2001 e 2003 quando o melhor índice (0.72) foi encontrado tem-se um número de artigos citantes menor, 106 e 123 respectivamente. Desta forma, pode-se dizer que apenas o número absoluto de citações não indica uma boa difusão, mas se analisado em conjunto com os autores citantes, instituições e países citantes é possível determinar o grau de concentração deste conhecimento, conforme estudado por Rousseau, Liu e Ye (2012).

CONSIDERAÇÕES FINAIS

Indicadores bibliométricos são necessários por trazerem novas perspectivas ao entendimento da comunicação científica. O fator de difusão, apesar de não ser completo, é uma medida que pode fornecer informações adicionais sobre a influência e difusão de um artigo, um periódico ou um autor. Partindo da ideia de que as citações representam a difusão do conhecimento e que o coeficiente de Gini padronizado (Ge) reflete a medida desta difusão demonstrou-se que é possível calcular e acompanhar a propagação de um conhecimento ao longo de um período.

Das 1.272 instituições de filiações dos autores citantes, 748 (58.80%) são brasileiras ou estão sediadas no Brasil, enquanto 518 (40.72%) são estrangeiras. As instituições voltadas para atividades educacionais lideram (47.96%), seguidas pelas instituições de P&D (22.88%) e as demais ocupam os setores públicos, como as secretarias de estados, hospitais, empresas de atividade agrícola, pecuária e serviços, além das indústrias, entre outras.

Em relação aos países, das 7,587 ocorrências de países presentes no vínculo institucional dos autores, 6.168 (81.30%) destas se referem ao Brasil, e as demais ocorrências estão distribuídas entre 49 países diferentes. A distribuição por continentes é a seguinte: América do Sul (82.35%); América do Norte (9.71%); Europa (6.93%); e América Central, Ásia, Oceania, África e Oriente Médio (aproximadamente 1% dos autores citantes).

Os resultados apontam para uma distribuição das citações feitas por diversos autores em diversas instituições e países diferentes. Com o intuito de compreender melhor este resultado, aplicou-se um indicador mais sofisticado, o fator de difusão. A relação de que quanto maior a quantidade de autores, instituições e países diferentes melhor será o grau de difusão é bem clara quando analisamos os dados absolutos de citações do IBGE. A aplicação do cálculo do fator de difusão comprovou a ideia aparente nos números absolutos: índice de Gini entre 0,62 em 2009 e 0,72 em 2001 e 2003 indica a desigualdade entre os documentos citantes, o que significa um fator de dispersão satisfatório. Ou seja, a existência de um grande número de autores, instituições e países diferentes que citam as publicações do IBGE é favorável para uma melhor disseminação das informações e do conhecimento gerado pela Instituição.

O presente estudo tem características de experimento compartilhado, sujeito a críticas e observações. A análise bibliográfica realizada apresenta limitações em relação ao período de coleta dos dados, pois está limitada aos dez anos, ficando fora desta análise as citações que ocorreram nos anos precedentes, assim como os documentos indexados após a data de coleta;

também em relação à fonte de dados, visto que a base WoS não cobre de forma exaustiva as publicações científicas, ficando fora desta análise as citações que ocorrem em periódicos não indexados e em documentos primários, como as dissertações, teses, trabalhos de conclusão, entre outros; e em relação à natureza das citações bibliográficas de fontes oficiais, visto que alguns editores de periódicos científicos orientam os autores a que façam referência às informações oficiais apenas como notas de rodapé ou em notas de fontes nas tabelas, não relacionando-as na lista de referências no final dos artigos.

REFERÊNCIAS

- Frandsen, Tove F. 2004. “Journal diffusion factors – a measure of diffusion?”, *Aslib Proceedings*, vol. 56, no. 1: 5-11.
- Frandsen, Tove F., Ronald Rousseau e Ian Rowlands. 2006. “Diffusion factors”, *Journal of Documentation*, vol. 62, no. 1: 58-72.
- Instituto Brasileiro de Geografia e Estatística (IBGE). 2016. Rio de Janeiro. Disponível em <<http://www.ibge.gov.br/home/disseminacao/eventos/missao/instituicao.shtm>>. Acesso em 24 maio 2016.
- Instituto Brasileiro de Geografia e Estatística (IBGE). 2013. *Metodologia do Censo Demográfico 2010*. Rio de Janeiro: IBGE, 713 p. (Série relatórios metodológicos, v. 41). Disponível em <<http://biblioteca.ibge.gov.br/visualizacao/livros/liv81634.pdf>>. Acesso em 24 maio 2016.
- Liu, Yuxian, Ismael Rafols e Ronald Rousseau. 2012. “A framework for knowledge integration and diffusion”, *Journal of Documentation*, vol. 68, no. 1: 31-44.
- Liu, Yuxian; Rousseau, Ronald. 2010. “Knowledge diffusion through publications and citations: a case study using ESI-Fields as unit of diffusion”, *Journal of the American Society for Information Science and Technology*, vol. 61, no. 2: 340-351.
- Liu, Yuxian, Ronald Rousseau e Raf Guns. 2013. “A layered framework to study collaboration as a form of knowledge sharing and diffusion”, *Journal of Informetrics* (Amsterdam), vol. 7, no. 3: 651-664.
- Rousseau, Ronald, Yuxian Liu e Fred Y. Ye. 2012. “A preliminary investigation on diffusion through a layered system”, *Journal of Informetrics* (Amsterdam), vol. 6, no. 2: 177-191.
- Rowlands, Ian. 2020. “Journal diffusion factors: a new approach to measuring research influence”, *Aslib Proceedings*, vol. 54, no. 2: 77-84.
- Rummler, Guido. 2006. “Modelagem de um indicador bibliométrico para análise da dispersão de conhecimentos”, *Ciência da Informação*, Brasília, v. 35, n. 1, p. 63-71, jan./abr. Disponível em <<http://www.scielo.br/pdf/ci/v35n1/v35n1a08.pdf>>. Acessado em 12 março 2013.
- Sandroni, Paulo. 1996. *Dicionário de administração e finanças*. São Paulo: Best Seller.
- Zanotto, Sônia R. 2011. “Informação estatística oficial produzida pelo IBGE: apropriação pela comunidade científica brasileira no período de 2001 a 2009. 2011”. Porto Alegre: Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Comunicação e Informação, Dissertação (Mestrado em Comunicação e Informação), 165 f. Disponível em <<http://www.lume.ufrgs.br/bit>

stream/handle/10183/30278/000777441.pdf?sequence=1. Acessado em: 12 março 2012.

Zanotto, Sônia R.; Samile A. S. Vanz e Ida R. C. Stumpf. 2011. "A informação estatística oficial produzida pelo IBGE e a sua difusão geográfica", In *Anais...* Brasília: XII Encontro Nacional de Ciência da Informação e Biblioteconomia.

Para citar este texto:

Zanotto, Sônia Regina, Souza Vanz, Samile Andréa de, Chittó Stumpf, Ida Regina. 2017. "Fator de difusão: uma medida da difusão do conhecimento através das citações". *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* (Número Especial de Bibliometría): 101-122.

<http://dx.doi.org/10.22201/iibi.24488321xe.2017.nesp1.57887>