

# Avaliação cienciométrica da produção científica mundial de 2013 a 2017 a partir da base Scopus<sup>®</sup>

Lara B. Silveira<sup>1†</sup>, Eric B. Ferreira<sup>2</sup>

<sup>1</sup>*Mestranda do Programa de Pós-graduação em Estatística Aplicada e Biometria. Universidade Federal de Alfenas.*

<sup>2</sup>*Professor do Departamento de Estatística. Universidade Federal de Alfenas. E-mail: [eric.ferreira@unifal-mg.edu.br](mailto:eric.ferreira@unifal-mg.edu.br).*

**Resumo:** *Uma vez que o conhecimento científico se difunde por meio de artigos científicos e que, atualmente, os artigos se encontram em formato eletrônico, o monitoramento de publicações e citações pode ser facilmente feito via internet. O portal SCImago Journal & Country Rank é um repositório público que contém dados da base Scopus<sup>®</sup> sobre a publicação de diversos países, regiões e áreas do conhecimento, ao longo do de vários anos. Este trabalho traz uma análise dos indicadores de quantidade e qualidade de publicação no mundo de 2013 a 2017, e compreende que a evolução da produção científica tem se dado em grande velocidade, preocupa-se cada vez mais com qualidade e o Brasil ainda tem um grande caminho para trilhar para alcançar melhores indicadores.*

**Palavras-chave:** Bibliometria; Scopus; SCImago.

**Abstract:** *Since scientific knowledge diffuses through scientific articles and the articles are now in electronic format, the monitoring of publications and citations can be easily done via the internet. The SCImago Journal & Country Rank portal is a public repository that contains data from the Scopus<sup>®</sup> database on the publication of several countries, regions and areas of knowledge over the course of several years. This work presents an analysis of the indicators of quantity and quality of publication in the world from 2013 to 2017 and understands that the evolution of scientific production has taken place at great speed, is increasingly concerned with quality and Brazil still has a great way to reach better indicators.*

**Keywords:** Bibliometry; Scopus; SCImago.

## Introdução

O termo *Cienciométria* surgiu na antiga União Soviética e na Europa Oriental, sendo empregado especialmente na Hungria (Vanti, 2002). Ainda para a mesma autora, a cienciométria refere-se à aplicação de métodos quantitativos para o estudo da história da ciência e do progresso científico e tecnológico, que se dá a partir da análise de patentes, teses e dissertações, entre outros tipos de produtos da ciência.

As análises bibliométricas utilizam indicadores de produção, citação, mostrando além da produtividade, a relevância e impacto de autores, periódicos, instituições, grupos ou países, nas diferentes áreas do conhecimento (OLIVEIRA; GRACIO, 2011).

---

<sup>†</sup>Autora correspondente: [larabs27@gmail.com](mailto:larabs27@gmail.com).

Bufrem e Prates (2005) afirmam que quando os métodos quantitativos são utilizados para estudar as atividades científicas ou técnicas, do ponto de vista de sua produção ou comunicação, costuma-se denominá-los cientometria. O termo cientometria é considerado um sinônimo de cienciométrica. Ademais, Price (1963), chamou a cienciométrica de “ciência das ciências”, por estudar a evolução e o impacto social das ciências, englobando o sistema de pesquisa como um todo.

A produção de indicadores bibliométricos mais representativos só se tornou uma realidade concreta nas últimas décadas do século XX, em função da criação, manutenção e informatização de bases de dados para armazenamento e consulta de informação científica. (MUGNAINI; JANNUZZI; QUONIAM, 2004).

Existem diversas plataformas eletrônicas nos dias atuais que cumprem o papel de disponibilizar dados bibliométricos e cientométricos para a população. É claro que nem todos esses dados são de acesso público, mas vários deles são. Dentre essas bases de dados destacam-se o *Web of Science* e o *Scopus*.

Archambault e colaboradores (2009), em uma comparação dos bancos de dados das plataformas *Web of Science* e *Scopus* entre os anos de 1996 e 2007, afirmam que as 25 primeiras posições do ranking de publicações de trabalho em ambas as plataformas é o mesmo, tendo uma diferença de uma ou duas posições no ranking entre os países.

O uso de ambas as plataformas é livre. Além disso, tanto o *Web of Science* quanto *Scopus* podem gerar resultados como 1º autor, citações, relevância, qual o periódico está buscando nas guias de acordo com a necessidade do leitor (CHADEGANI et al., 2013). Os autores ainda afirmam que na plataforma *Scopus* os dados são apresentados por autor, ano, assunto e tipo de documento (citável ou não-citável). Em contrapartida, plataforma *Web of Science* não faz essa apresentação dos dados. Nas duas plataformas de dados, os índices h são muito próximos.

Diante disso, optou-se por utilizar a plataforma *Scopus* como base para as análises. Os dados utilizados para este trabalho foram extraídos do portal *SCImago Journal & Country Rank*.

A partir disso, a proposta deste trabalho foi efetuar uma evolução temporal da avaliação e quantitativa da produção científica, em 239 países do mundo, considerando a base de dados do portal *SCImago Journal & Country Rank*.

## **SCImago Journal & Country Rank**

O *SCImago* é um grupo de investigação do Consejo Superior de Investigaciones Científicas (CSIC), University of Granada, Extremadura, Carlos III (Madrid) e Alcalá de Henares, dedicado à análise, representação e recuperação da informação por meio de técnicas de visualização.

A plataforma hospedada no sítio <https://www.scimagojr.com> (Figura 1) leva o seu nome a partir do indicador *SCImago Journal Rank* (SJR), desenvolvido pelo *SCImago* a partir do algoritmo amplamente conhecido *Google PageRank*<sup>TM</sup>. Este indicador mostra a visibilidade dos periódicos contidos no banco de dados *Scopus*<sup>®</sup> a partir de 1996.

O *SCImago Journal Rank* (SJR) é um indicador bibliométrico que mede a influência de um determinado periódico pela média do número de citações recebidas nos últimos 3 anos que antecedem ao ano analisado e que são ponderadas dependendo da área do conhecimento e do prestígio do periódico. A diferença para o *Journal Citation Report* (JCR) é que o SJR leva em consideração para atribuição da métrica, fatores como: origem



Figura 1: Plataforma do Scimago Journal Rank no sítio <https://www.scimagojr.com>.

da citação e o comportamento de citação de cada área do conhecimento.

Além disso, o SCImago Journal & Country Rank é um portal livre que inclui periódicos e os indicadores científicos dos países gerados a partir das informações contidas no banco de dados Scopus (Elsevier B.V.).

Country	↓ Documents	Citable documents	Citations	Self-Citations	Citations per Document	H index
1  Nigeria	6130	6067	48813	11259	7.96	65
2  South Africa	5621	5503	67115	16120	11.94	83
3  Kenya	2718	2653	48355	9299	17.79	85
4  Tunisia	1955	1923	24486	4637	12.52	60
5  Ethiopia	1793	1750	18407	4502	10.27	53

Figura 2: Página de buscas da plataforma do Scimago Journal Rank, na aba *Country Rankings*.

Nesse portal, na aba *Country Rankings*, é possível fazer pesquisas por diversos filtros, como área do conhecimento, categorias de assunto, regiões ou países, tipo de material seriado e ano (desde 1996). Além disso, é possível pesquisar apenas periódicos que publiquem em Acesso Aberto, que façam parte da coleção de periódicos indexados na Scielo ou revistas que também sejam indexadas na base de dados Web of Science.

No tocante as regiões, o portal agrupa países segundo apresentado na Figura 3. Nela o Brasil aparece destacado dos demais países da América Latina, pois foi analisado separadamente. 1.

**Sigmae**, Alfenas, v.8, n,2, p. 8-18, 2019.

64<sup>a</sup> Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18<sup>o</sup> Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

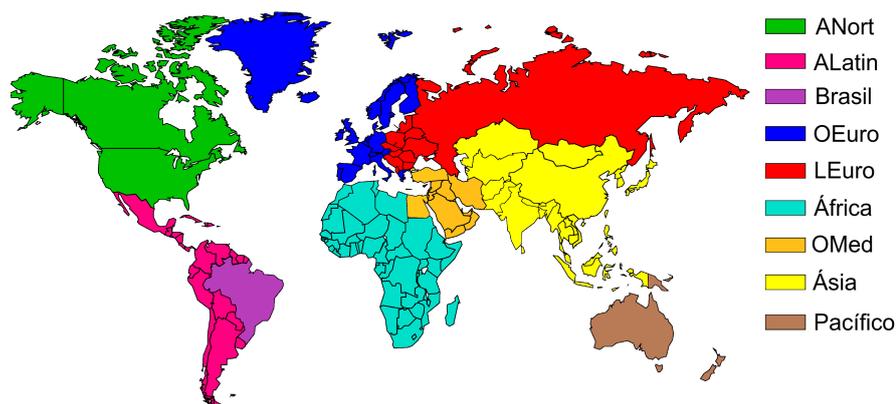


Figura 3: Divisão do mundo em regiões, considerada pelo SCImago.

Os trabalhos também estão estratificados em áreas temáticas principais (27 áreas), categoria de assunto (313 categorias de assuntos específicos) e por país. Os dados de citações são extraídos de mais de 34.100 títulos de mais de 5.000 editores internacionais e métricas de desempenho de 239 países em todo o mundo. As variáveis disponíveis no portal SCImago são as seguintes:

- Documentos ( $D$ ): Número de trabalhos publicados durante o ano selecionado. Geralmente é chamado de produção científica do país. Engloba todos os tipos de trabalhos publicados, incluindo os citáveis e os não-citáveis.
- Documentos citáveis ( $DC$ ): Trabalhos citáveis publicados em um país, em um dado ano. São considerados exclusivamente artigos, revisões e trabalhos publicados em anais de eventos.
- Citações ( $C$ ): Número de citações recebidas pelos trabalhos publicados no ano de origem. Por exemplo, citações nos anos  $X, X + 1, X + 2, X + 3...$  para trabalhos publicados durante o ano  $X$ .
- Citações por Documento ( $CpD$ ): Média de citações por trabalho publicado durante o ano de origem, por exemplo, citações nos anos  $X, X + 1, X + 2, X + 3...$  para trabalhos publicados durante o ano  $X$ .

A variável  $CpD$  é calculada pela seguinte expressão:

$$CpD = \frac{C}{D}$$

- Índice  $h$  ( $h$ ): O índice  $h$  (HIRSCH, 2005) busca aliar quantidade e qualidade, ou seja, o impacto de um conjunto de publicações, que pode ser um país, uma região, um pesquisador, um periódico, etc. Ele é calculado a partir da organização do número de citações que cada artigo científico de um país (p.ex.) possui, em ordem

**Sigmae**, Alfenas, v.8, n,2, p. 8-18, 2019.

64<sup>a</sup> Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18<sup>o</sup> Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

decrecente. Isso feito, diz-se que o índice  $h$  de um país é igual a  $x$  se os seus  $x$  artigos mais citados possuírem pelo menos  $x$  citações.

O índice  $h$  é calculado pela seguinte expressão:

$$h = \max(\min(f(i), i)), \quad i = 1, 2, \dots, n.$$

- Autocitação ( $A_c$ ): Número absoluto de autocitações dos artigos de uma área temática, em um dado país, em um dado ano.

As autocitações apresentam uma dualidade que merece nota. Ao mesmo tempo que existe um lado negativo que envolve auto-promoção de um pesquisador (ou grupo de pesquisadores) e, portanto, estabelecimento de um viés científico, por outro lado trabalhos afins precisam ser citados e, muitas vezes, os pesquisadores de uma linha de pesquisa precisam citar seus trabalhos por serem realmente relevantes (HYLAND, 2003).

Embora essas métricas sejam interessantes para inferir sobre a produtividade científica dos países, regiões e áreas do conhecimento, serão apresentadas na próxima seção indicadores derivados, que buscam ser mais fáceis de interpretar e inferir.

## Material e Métodos

Neste trabalho, foram consideradas todas as regiões do globo, conforme divisão fornecidas pela plataforma. Contudo, foi destacado o Brasil, por motivos óbvios, em todas as análises pertinentes, e as regiões são tratadas por siglas, que podem ser observadas na Tabela 1.

Tabela 1: Regiões em que os trabalhos científicos foram divididos e siglas utilizadas.

Região	Sigla
América do Norte	ANort
América Latina	ALatin
Brasil	Brasil
Oeste Europeu	OEuro
Leste Europeu	LEuro
África	África
Oriente Médio	OMed
Ásia	Ásia
Região do Pacífico	Pacífico

As áreas do conhecimento (áreas temáticas) também são expressas nesse trabalho por meio de siglas. A Tabela 2 representa as áreas temáticas e as siglas utilizadas para cada uma das 27 áreas.

Além das variáveis disponíveis no portal, novas variáveis foram obtidas para a realização deste estudo. Essa necessidade adveio do acúmulo das estatísticas ao longo de vários países para representar uma região e de várias áreas do conhecimento para formar um país, por exemplo. Dessa forma, as estatísticas acumuladas, embora corretas, perdiam em interpretabilidade. Dessa forma, foram propostas as seguintes variáveis relativizadas:

**Sigmae**, Alfenas, v.8, n,2, p. 8-18, 2019.

64<sup>a</sup> Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18<sup>o</sup> Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

Tabela 2: Principais áreas temáticas consideradas neste estudo e suas siglas.

Temática	Sigla	Área Temática	Sigla
Agricultura e Ciências Biológicas	ACB	Farmacologia, Toxicologia e Farmacêutica	FTF
Artes e Humanidades	AH	Profissões da Saúde	PS
Bioquímica, Genética e Biologia Molecular	BGBM	Imunologia e Microbiologia	IM
Ciências Ambientais	CA	Matemática	MAT
Ciência da Computação	CC	Medicina	MED
Ciência da Decisão	CD	Multidisciplinar	MUL
Ciência dos Materiais	CM	Neurociência	NEU
Ciências Sociais	CS	Negócios, Gestão e Contabilidade	NGC
Economia, Econometria e Finanças	EEF	Odontologia	Odonto
Energia	Energia	Psicologia	PSI
Enfermagem	ENF	Química	Quimica
Engenharia	Engenharia	Terra e Ciências Planetárias	TCP
Engenharia Química	EQ	Veterinária	VET
Física e Astronomia	FA		

- Área do Conhecimento ( $AC$ ): Divide as 27 áreas temáticas (Tabela 2) em 3 grandes áreas do conhecimento: Biológicas, Humanas e Exatas.
- Autocitação por citação ( $AcC$ ): Frequência de relativa de autocitações por área por ano em relação ao número de citações por área em um dado ano.

$$AcC = \frac{Ac}{C}$$

- Índice h relativo ( $hDC$ ): considera o índice h relativo ao número de documentos citáveis por país.

$$hDC = \frac{h}{DC}$$

O estudo caracterizou-se por uma coleta de dados secundários na base de dados da plataforma SCImago utilizando-se, do método quantitativo e do estudo descritivo por meio de técnicas de estatística descritiva e dispositivos gráficos. O conjunto de dados contém publicações de 239 países divididos por regiões no período de 2013 a 2017.

Para a análise dos dados foram utilizados os softwares R (R CORE TEAM, 2019), RStudio v1.1.456 (RSTUDIO TEAM, 2016) e o pacote `Plotly` (SIEVERT, 2018).

## Resultados e Discussão

Primeiramente, abordou-se a investigação a quantidade e a qualidade da produção de ciência por uma visão global. Nesse sentido, a Figura 4 apresenta a evolução temporal da quantidade e da qualidade de documentos publicados no mundo.

É possível notar que há um aumento no número de artigos publicados no mundo nos últimos anos, mas isso se dá em ritmo muito mais lento do que se dá a queda na qualidade. O número de citações despensa (Figura 4a), assim como o fator de impacto (número de citações por artigo) (Figura 4b).

Também, o número de autocitações cai de forma absoluta, mas em ritmo menor que o número de citações (Figura 4a). Relativamente, o proporção de autocitações por citação sofre ligeiro aumento (Figura 4b).

*Sigmae*, Alfenas, v.8, n,2, p. 8-18, 2019.

64<sup>a</sup> Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18<sup>o</sup> Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO).

Esses resultados sugerem que o mundo ainda está pouco preocupado, em geral, com o aumento da qualidade. Valoriza-se o número de artigos publicados muito mais do que sua capacidade de produzir conhecimento ou mesmo de gerar tecnologia que se converta em melhoria da qualidade de vida das pessoas.

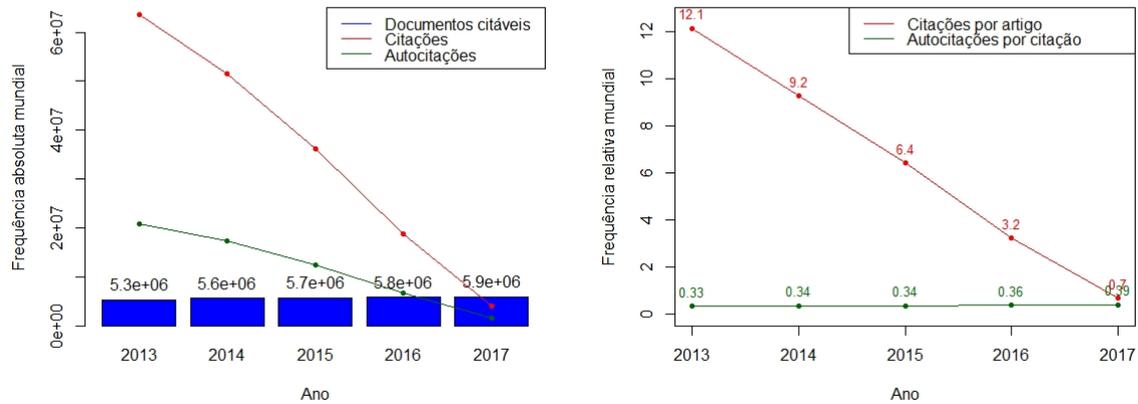


Figura 4: Evolução temporal do número de documentos publicados no mundo.

De 2013 a 2017, o Brasil caiu uma posição no ranking dos países que mais publicaram no mundo, passando da 13<sup>a</sup> posição para a 14<sup>a</sup> posição. Dentre os países da América Latina, o Brasil manteve-se na 1<sup>a</sup> posição em todo o período analisado.

A Figura 5 apresenta o índice h médio por região. A região da América do Norte apresenta o maior índice h médio, seguido do Oeste Europeu, Ásia, Oriente Médio, Leste Europeu, América Latina, Pacífico e, por fim, a África. Apesar do Brasil ser o país que apresenta maior o índice h médio na região da América Latina, ainda há uma discrepância com relação à América do Norte.

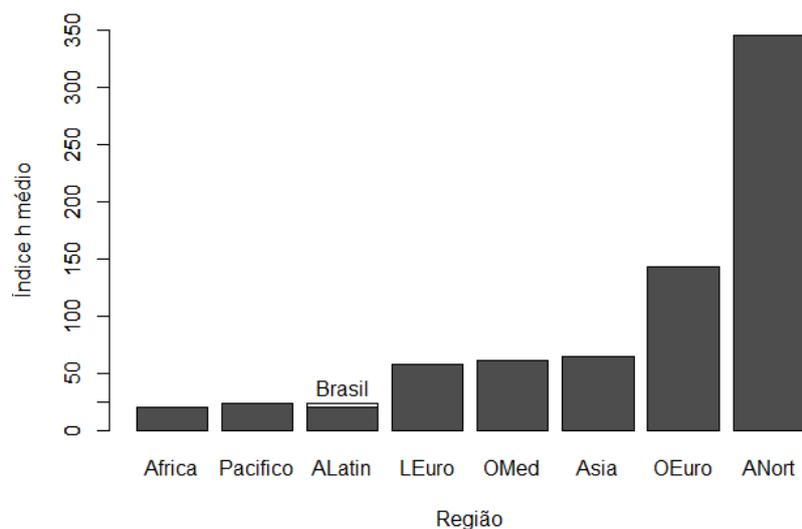


Figura 5: Distribuição do índice h médio por região com destaque para o Brasil em cinza.

Os gráficos de setores apresentam as 3 grandes do conhecimento: exatas, biológicas e humanas. A Figura 6a representa as porcentagens de índice h médio por Área do Conhecimento entre os anos de 2013 e 2017 em todas as regiões do globo. Nota-se que quase 44,82% do índice h é correspondente à área de Biológicas. Já a Figura 6b representa o percentual de trabalhos publicados no mesmo período em todas as regiões do globo. Nesse caso, a área de exatas corresponde a quase metade das publicações no período. Em ambos os casos a área de humanas corresponde à menor porcentagem nos gráficos.

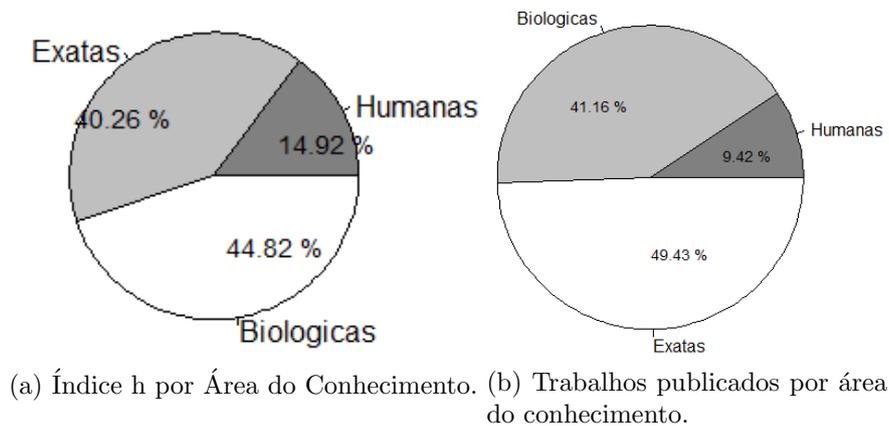


Figura 6: Frequências percentuais do índice h e trabalhos publicados por área do conhecimento.

Com relação às publicações brasileiras, na Figura 7a 54,18% das publicações referem-se à área de Biológicas, seguido de 37,35% da área de Exatas, e por fim, 8,47% da área de Humanas, no período de 2013 a 2017. A Figura 7b apresenta as publicações brasileiras divididas nas 27 áreas temáticas. A grande fatia está explicitamente com a área de Medicina, seguido das áreas Agricultura e Ciências Biológicas, Engenharia e Bioquímica, Genética e Biologia Molecular. Essa grande quantidade de publicações na área de Medicina está atrelada à maior porcentagem de publicações no país estarem na área de Ciências Biológicas.

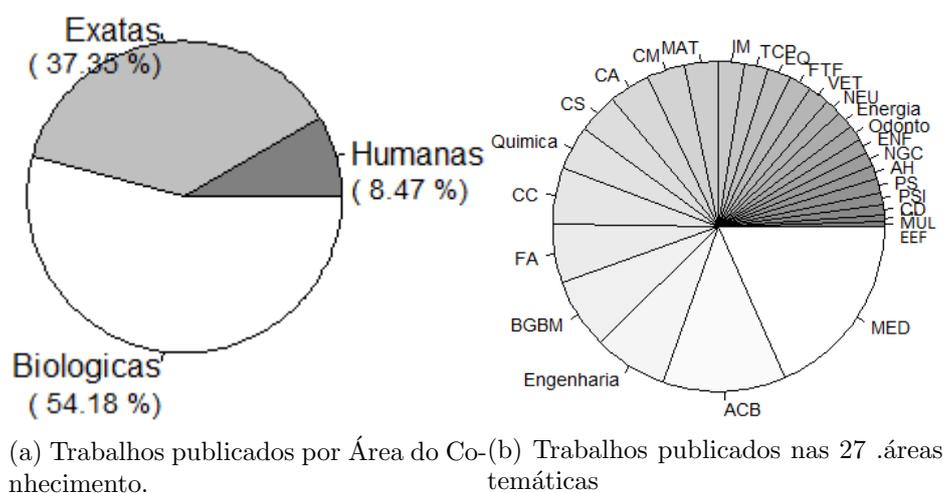


Figura 7: Distribuição percentual dos trabalhos publicados no Brasil por área temática e por área do conhecimento.

Os gráficos de bolhas apresentados na Figura 8 representam a evolução temporal da pesquisa no mundo. O eixo X representa a variável índice h relativo e o eixo Y a variável citações por documento. O diâmetro das bolhas representa a variável autocitação por citação para cada uma das áreas temáticas.

Nota-se que as áreas de Veterinária, Matemática e Ciências Sociais no ano de 2013 são as que obtiveram o maior número de autocitações em relação às outras áreas temáticas. No decorrer dos anos, houve uma pequena variação no volume de autocitações e do índice h relativo. Entretanto, o volume de Citações por Documentos cresceu no período analisado. A área Multidisciplinar, em todo o período analisado, apresentou os maiores índice h relativo e citações por documentos.

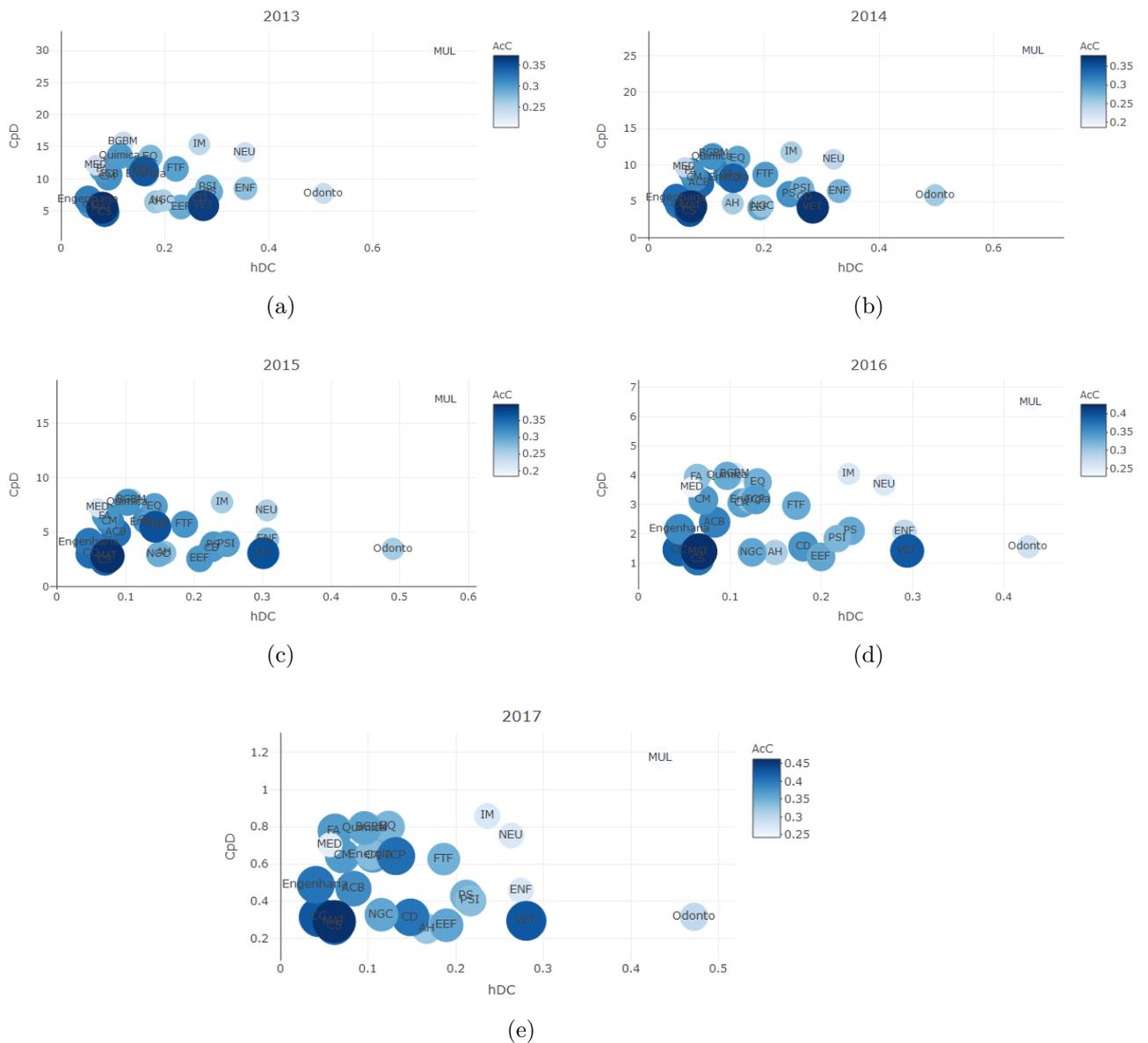


Figura 8: Evolução Temporal das áreas do conhecimento quanto ao índice h relativo (eixo X), às citações por documento (eixo Y) e autocitação por citação (diâmetro da bolha).

No ambiente acadêmico, a autocitação muitas vezes é vista como busca por auto-

promoção, mas em muitos casos, os motivos de praticar a autocitação são diversos, envolvendo fatores psicológicos relacionados à autoconfiança ou para fazer uma ligação de trabalhos novos a trabalhos anteriores (HYLAND,2003).

A área Multidisciplinar foi a que obteve o maior índice h relativo em todo o período. Seguido das áreas Odontologia, Veterinária e Enfermagem. Dessas, somente a área de Veterinária apresentou um alto número de autocitações. No decorrer do período, houve um aumento na quantidade de citações por documentos em todas as áreas.

## Conclusões

De acordo com o banco de dados do SCImago, há um aumento na quantidade de publicações e de citações por documentos ao longo do tempo, embora não haja uma tendência clara de aumento do indicador de qualidade índice h relativo.

Em caráter mundial, a área de Multidisciplinar em todo o período é a que apresenta maior índice h relativo. O Brasil, assim como o mundo, obteve mais publicações na área de Ciências Biológicas.

## Agradecimentos

Agradecemos à Universidade Federal de Alfenas (Unifal-MG) e ao Programa de Pós-Graduação em Estatística Aplicada e Biometria (PPGEAB) pela estrutura e apoio disponibilizados para a realização desse trabalho.

## Referências Bibliográficas

ARCHAMBAULT, É.; CAMPBELL, D.; GINGRAS, Y.; LARIVIÈRE, V. Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, 60(7), 1320–1326. doi: 10.1002/asi.21062, 2009.

BUFREM. L.; PRATES, Y. O saber científico registrado e as práticas de mensuração da informação. *Ciência da Informação*, Brasília, v.34, n.2, p.9-25, maio/ago, 2005.

CHADEGANI, A.A.; SALEHI, H.; YUNUS, M. M.; FARHADI, H.; FOOLADI, M.; FARHADI, M.; et al. A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases. *Asian Social Science*, 9(5), 18-26, 2013.

HIRSCH, J.E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences (PNAS)*. v.102, n.46. p. 16569–16572. 2005.

HYLAND K. Self-citation and self-reference: credibility and promotion in academic publication. *J Am Soc Inf Sci Technol*.54(3):251–9, 2003.

**Sigmae**, Alfenas, v.8, n,2, p. 8-18, 2019.

64<sup>a</sup> Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18<sup>o</sup> Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

MUGNAINI R.; JANUZZI P.; QUONIAM, L. Indicadores Bibliométricos da Produção Brasileira: uma análise a partir da base Pascal. *Ciência da Informação*, Brasília, v.33, n.2, p.123-131, 2004.

OLIVEIRA, E. F. T. de; GRÁCIO, M. C. C. Indicadores bibliométricos em ciência da informação: análise dos pesquisadores mais produtivos no tem “estudos métricos” na base Scopus. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 16, n. 4, p. 16-28, 2011.

PRICE, J. D. S. *Little science, big science*. New Haven, CT : Yale University, 1963.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2019. URL <https://www.R-project.org/>.

RSTUDIO TEAM. *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. 2016. <http://www.rstudio.com/>.

SCImago, (n.d.). SJR — SCImago Journal & Country Rank [Portal]. Retrieved Date you Retrieve. Disponível em <http://www.scimagojr.com>. Acesso em 22 mar. 2019.

SIEVERT, C. *Plotly for R*. 2018. <https://plotly-book.cpsievert.me>. R package version 4.8.0.

VANTI, N.A.P. From Bibliometry to Webometry: A Conceptual Exploration of Several Forms of Measuring Information and Knowledge. *Ciência da Informação*, 31, 369-379. 2002.

**Sigmae**, Alfenas, v.8, n,2, p. 8-18, 2019.

64<sup>a</sup> Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18<sup>o</sup> Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).