



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN
FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y DE LA INFORMACIÓN

Análisis de la producción científica de México en el Web of Science, durante el período 2005-2015, utilizando inteligencia computacional.

TESIS

QUE PARA OPTAR POR EL GRADO DE:

**MAESTRA EN BIBLIOTECOLOGÍA Y
ESTUDIOS DE LA INFORMACIÓN**

PRESENTA:

Lic. Ibis Anette Lozano Díaz

ASESOR: DR. Humberto A. Carrillo Calvet
Facultad de Ciencias y Centro de Ciencias de la Complejidad
Universidad Nacional Autónoma de México

Ciudad de México, Diciembre 2016



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Not everything that counts can be counted.

William Bruce Cameron (1963)

Agradecimientos

Quiero agradecer a mis sinodales, Dra. Jane Russell, Dr. Salvador Gorbea Portal, Dr. Francisco Collazo y Dra. María Luna Morales, **por dedicar parte de su tiempo a la revisión de esta investigación y por sus valiosos comentarios, que sirvieron para enriquecer esta tesis.** A la Dra. Yohannis Martí Lahera, directora de la Biblioteca Central de la Universidad de La Habana, por sus sugerencias sobre los análisis en los indicadores de producción e impacto, tema de su especialidad. Al Dr. Ricardo Arencibia Jorge agradezco su apoyo en el estudio sobre revistas mexicanas. A la Dra. Rosa Lidia Vega Almeida, del Centro de Histoterapia Placentaria de Cuba, por sus valiosos comentarios sobre la redacción de la tesis.

A mi tutor Dr. Humberto Carrillo Calvet por su **visión sobre el tema y por sus sabias sugerencias; su paciencia y exigencia en la conducción de esta investigación.** Además, doy gracias por enseñarme el verdadero rigor de la ciencia, y cómo manejar cada situación que se presenta en el camino de hacer ciencia. A su lado, cada día es una enseñanza.

Agradezco a mis padres **Ibis Díaz y Pedro Lozano**, por todo el amor y apoyo incondicional que recibo cada día de su parte, sin importar la distancia.

A todos los chicos del Laboratorio de Dinámica No Lineal, en especial a **José Luis Jiménez**, **por explicarme cómo funcionan las neuronas del SOM**, contar con su ayuda incondicional durante todo este tiempo y **por convertirse en un verdadero amigo.**

A mi amiga, **Yaniris Rodríguez Sánchez** por darme la luz sobre la **“Ciencia”** y la vida académica.

A los profesores del Posgrado en Bibliotecología y Estudios de la Información, de la Universidad Nacional Autónoma de México por ofrecerme una excelente educación. Por último, agradecer **al CONACYT** su apoyo económico durante el curso 2014-2016.

Resumen

Desde una perspectiva bibliométrica, la presente investigación analiza la producción científica de México, evaluando el desempeño de las revistas y la producción de artículos científicos, tomando en cuenta diversos indicadores independientes de tamaño, basados en citas.

Se considera información contenida en las bases de datos del Sistema Nacional de Investigadores (SNI), del *Web of Science (WoS)* y de *Scopus*, durante el período comprendido entre enero de 2005 y diciembre de 2015.

El método de análisis se basa en el uso de técnicas bibliométricas y de inteligencia computacional. Se construye y se utiliza una batería de indicadores bibliométricos para estimar la eficiencia y la efectividad de la producción. Una parte del análisis se realiza aplicando una tecnología neurocomputacional y técnicas de frontera para la visualización de información, implementadas en el sistema de software LabSOM, desarrollado en el Laboratorio de Dinámica no Lineal de la Facultad de Ciencias.

A la luz de cuatro indicadores de última generación (dos del Journal Metrics de Scopus y dos del *Journal Citation Reports*), se analiza el desempeño de todas las revistas mexicanas (39) indizadas en el WoS. Se identifican las diez revistas con mayor desempeño, de acuerdo a cada uno de estos indicadores y se comparan los resultados.

Por otra parte, se compara la evolución de la producción científica de México en las bases de datos del WoS y de Scopus con la evolución del SNI. Se identifican los campos de investigación en los que la ciencia mexicana tiene mayor producción e influencia. También se propuso un *Índice de Desarrollo Científico*, independiente de tamaño, que es útil para comparar la evolución de la producción de México con la de otros países con un similar nivel de desarrollo científico.

Se comprueba la utilidad de la tecnología neurocomputacional para llevar a cabo análisis multiparamétrico dentro un conjunto de 70 países, tomando en cuenta simultáneamente cuatro indicadores que estiman: productividad, impacto normalizado y dos indicadores de excelencia. Los resultados del análisis neurocomputacional se visualizan en mapas que encriptan el conocimiento obtenido.

Los resultados obtenidos son de potencial utilidad para la evaluación y el diseño de política científica.

Palabras Clave: Minería de datos bibliométrica; visualización de información; redes neuronales; mapas auto-organizados; algoritmo SOM; ViBlioSOM; indicadores; México.

Tabla de Contenido

Introducción	I
1. Exposición de motivos, justificación y objetivos de la investigación.....	I
2. Antecedentes.....	V
Capítulo I: Marco teórico	1
1. Elementos de análisis de la producción académica.....	1
1.1. Indicadores de producción e indicadores basados en citas.....	1
1.2. La eficiencia científica medida a través de indicadores de productividad.....	4
1.3. Consideraciones sobre los indicadores de excelencia.....	5
2. Elementos de análisis del desempeño de las revistas académicas.....	7
2.1. El Factor de Impacto de Garfield.....	9
2.2. Eigenfactor del WoS.....	9
2.3. Source normalized impact per paper.....	10
2.4. SCImago Journal Rank (SJR).....	11
3. Descubrimiento de conocimiento, minería y visualización de información en bases de datos bibliográficas.....	12
3.1. El descubrimiento de conocimiento en bases de datos (KDD).....	12
3.2. La minería de datos como parte de la metodología de descubrimiento de conocimiento en bases de datos.....	14
3.3. Técnicas de la minería de datos.....	16
3.4. Visualización de información.....	17
3.5. KDD, Minería de Datos, visualización de información y la Bibliometría.....	22
3.5.1. El algoritmo SOM y la visualización de información bibliométrica.....	22
Capítulo II: Metodología y fuentes de información	24
2.1. Fuentes de información: los índices de difusión de la ciencia.....	24
2.1.1. El Web of Science (WoS): las bases de datos y herramientas de análisis bibliométrico.....	24
2.1.2. Los recursos de análisis del WoS: Essential Science Indicators y Journal Citation Reports.....	26
2.2. SCOPUS.....	29
2.2.1. Elsevier Journal Metrics.....	30
2.2.2. SCIMAGO Journal and Country Rank.....	30
2.3. Batería de Indicadores bibliométricos empleados en la investigación.....	31
2.3.1. Indicadores de producción y eficiencia.....	31
2.3.2. Indicadores de impacto.....	32
2.3.3. Indicadores del Essential indicators: Highly Cited Paper y Hot Paper.....	35

2.4. Metodología implementada: ViBlioSOM	36
2.4.1. Etapas de la metodología ViBlioSOM	37
Capítulo III: Resultados y discusión.....	41
1. Análisis bibliométrico de las principales revistas mexicanas.....	41
1.1. Núcleo de revistas mexicanas con mayor visibilidad en Web of Science y Scopus 41	
1.2. El top ten de revista mexicanas con mayor visibilidad	45
1.3. Análisis multidimensional de las revistas mexicanas, basada en la técnica de mapas auto-organizados o SOM (Self-Organizing Map)	49
2. Análisis de la producción científica mexicana en el WoS, 2005-2015.....	53
2.1. El Sistema Nacional de Investigadores (SNI)	53
2.2. Análisis comparativo de la evolución del SNI y de la producción científica mexicana en Scopus y el Web of Science	53
2.2.1. Análisis de los artículos científicos mexicanos visibles internacionalmente	56
2.3. Estudio comparativo de la producción mexicana con otros países.	59
2.4. Highly Cited y Hot papers: Comparación de la producción mexicana con otros países	63
2.5. Principales campos de investigación en la ciencia mexicana	71
2.5.1. Los campos de investigación con mayor impacto en la ciencia mexicana	73
2.5.2. Los campos de investigación científica mexicana con mayor efectividad y excelencia	74
2.6. Colaboración científica	78
3. Análisis multiparamétrico del perfil bibliométrico de México y otros países, basado en la técnica de mapas auto-organizados	79
3.1. Ranking de los países de acuerdo al ESI del WoS.....	79
3.2. Ranking de los países de acuerdo al grado de desarrollo científico.....	80
3.3. Aplicación de ViBlioSOM: Análisis multiparamétrico del perfil bibliométrico de México y otros países, considerando el índice de desarrollo científico, impacto y excelencia... ..	85
3.3.1. Análisis usando indicadores dependientes de tamaño	85
3.3.2. Análisis usando indicadores independientes del tamaño.....	89
Consideraciones finales.....	93
Referencias	95
ANEXOS	104

Relación de Figuras

Todas las figuras y tablas de esta tesis son de elaboración propia, con excepción de la Figura 1 y Figura 2

Figura 1. Fases de la metodología para el descubrimiento de conocimiento en bases de datos (KDD) (Fayyad, et al., 1996).....	15
Figura 2. Estructura de una Red Neuronal bajo el algoritmo SOM (Guzmán-Sánchez, 2009)	23
Figura 3. Metodología ViBlioSOM: etapas del descubrimiento, análisis y visualización de información bibliométrica	37
Figura 4. Mapa de componente del indicador: Índice de desarrollo científico (SDI) para países.....	39
Figura 5. Clustering de países considerando cuatro indicadores: Índice de desarrollo científico (SDI), Impacto Normalizado (NI), Highly Cited (HCP) y Hot Papers (Hot)	40
Figura 6. Revistas mexicanas indexadas en SCOPUS y Web of Science, 2005-2015.....	44
Figura 7. Revistas mexicanas con mayores valores en el Source Normalized Impact per Paper y SCImago Journal Rank para el año 2015	48
Figura 8. Revistas mexicanas con mayores valores en el Eigenfactor Score y Impact Factor para el año 2015.....	49
Figura 9. Mapa de Visibilidad por zonas de las revistas mexicanas de acuerdo con: a) Source Normalized Impact per Paper, b) Impact Factor, c) Eigenfactor Score, y d) SCImago Journal Rank. (Self-organized maps desarrollado por ViblioSOM). (Anexos 8 y 9: los mapas aparecen separado.....	52
Figura 10. Crecimiento paralelo del número total de miembros mexicanos en el SNI y la producción científica en Web of Science y Scopus, durante el período 2005-2015.....	55
Figura 11. La eficiencia de la ciencia mexicana en Web of Science y SCOPUS, durante 2005-2015	56
Figura 12. Crecimiento paralelo del número total de miembros mexicanos en el SNI y el número de artículos científicos en Web of Science y Scopus, durante el período 2005-2015	58
Figura 13. El índice de producción científica nacional (National Scientific Productivity) en Web of Science y Scopus, durante el período 2005-2015.....	58
Figura 14. Comparación de la producción mexicana en WoS con la de los países más productivos de la región.....	59
Figura 15. Comparación de la producción científica mexicana en el WoS por cada 100,000 habitantes con países de la región, 2005-2015.....	60
Figura 16. Comparación de la producción mexicana en WoS con países que no pertenecen a la región latinoamericana, con similar producción de acuerdo al ESI, durante el período 2005-2015	61
Figura 17. La producción científica mexicana en el WoS por cada 100,000 habitantes con la de los países con similar producción de acuerdo al ESI, 2005-2015.....	62
Figura 18. Comparativa del número de Top Papers mexicanos en el Web of Science con países de la región y otros países con similar producción, 2005-2015.....	63
Figura 19. Comparativa del número de Highly Cited Papers mexicanos en el Web of Science con países de la región y países de similar producción, 2005-2015.....	64

Figura 20. Comparativa del número de Hot Papers mexicanos en el Web of Science con países de la región y países de similar producción, 2005-2015	65
Figura 21. Comparativa del número de Top Papers mexicanos por cada 1,000 documentos en el Web of Science con países de la región y con países de similar producción, 2005-2015	66
Figura 22. Comparativa del número de Highly Cited Papers mexicanos por cada 1,000 documentos en el Web of Science con países de la región y países de similar producción, 2005-2015	67
Figura 23. Comparativa del número de Hot Papers mexicanos por cada 1,000 documentos en el Web of Science con países de la región y países de similar producción, 2005-2015	68
Figura 24. Comparativa del número de Highly Cited Papers mexicanos por cada 100,000 habitantes, con países de la región y países de similar producción, 2005-2015	69
Figura 25. Comparativa del número de Hot Papers mexicanos por cada 100,000 habitantes, con países de la región y países de similar producción, 2005-2015.....	70
Figura 26. Distribución del número de documentos mexicanos en el WoS, durante 2005-2015, entre 22 campos de investigación	72
Figura 27. Distribución del número de citas entre 22 campos de investigación en el WoS, durante 2005-2015	73
Figura 28. Distribución del promedio de citas por documento a través de 22 campos de investigación en el WoS, durante 2005-2015.....	74
Figura 29. Distribución entre 22 campos de investigación del número de Highly Cited Papers mexicanos en el WoS, durante 2005-2015	76
Figura 30. Distribución de los 9 de 22 campos de investigación del número de Hot Papers mexicanos en el WoS, durante 2005-2015	77
Figura 31. Evolución anual de la colaboración con los países más colaboradores con México en el período 2005-2015. Total de países colaboradores: 193.....	78
Figura 32. Evolución en el tiempo del índice de desarrollo científico, número de artículos publicados en revistas indexadas en WoS por cada 100,000 habitantes, para México y países con similar desempeño.....	81
Figura 33. Países con altos valores de impacto (NI) y el índice de desarrollo científico (SDI) durante el 2005-2015.....	83
Figura 34. Relación entre excelencia Highly Cited Papers (HCP) e impacto normalizado durante el período 2005-2015, para países con similar índice de desarrollo científico que México	84
Figura 35. Hot papers e impacto normalizado para países con similar índice de desarrollo científico que México, durante el período 2005-2015.....	84
Figura 36. Mapa de Cluster identificado con la técnica Self Organized Maps para 70 países que tienen perfiles similares a México, HCP y Hot como indicadores de tamaño dependiente (Anexos 10, 11 y 12, donde cada mapa está separado)	88
Figura 37. Mapa de Clusters identificados con la técnica Self Organized Maps para 70 países que tienen perfiles similares a México, HCP y Hot como indicadores de tamaño independiente (Anexo 13, 14 y 15, donde cada mapa está separado).....	92

Relación de Tablas

Tabla 1. Indicadores basados en citas con mayor uso derivados de Bases de datos como Web of Science y Scopus	3
Tabla 2. Número de documentos recuperados en la base de datos Scopus por cada tipo documental	30
Tabla 3. Número de documentos recuperados en la base de datos Web of Science por cada tipo documental	26
Tabla 4. Número de trabajos mexicanos publicados en fuentes indexadas en el WoS, durante el período 2005-2015.....	41
Tabla 5. Correlación de los indicadores estudiados para una muestra de las 39 revistas mexicanas cubiertas por SCOPUS y WoS.....	45
Tabla 6. Top-ten 2015 de revistas mexicanas en correspondencia con Impact Factor, Source Normalized Impact per Paper, Eigenfactor Score and SCImago Journal Rank	47
Tabla 7. Evolución de la producción científica mexicana en el Web of Science y Scopus, durante 2005-2015.	54
Tabla 8. Evolución del número de artículos mexicanos en el Web of Science y Scopus, durante 2005-2015.	57
Tabla 9. Rank de países según ESI durante el período 2005-2015.	79
Tabla 10. Correlación de los indicadores estudiados para una muestra de 100 países ...	82

Glosario de abreviaturas de países

Países	Siglas	Países	Siglas
ALGERIA	DZA	LEBANON	LBN
ARGENTINA	ARG	LITHUANIA	LTU
ARMENIA	ARM	LUXEMBOURG	LUX
AUSTRALIA	AUS	MALAWI	MWI
AUSTRIA	AUT	MALAYSIA	MYS
BANGLADESH	BGD	MEXICO	MEX
BELARUS	BLR	MOROCCO	MAR
BELGIUM	BEL	NEPAL	NPL
BOLIVIA	BOL	NETHERLANDS	NLD
BRAZIL	BRA	NEW ZEALAND	NZL
BULGARIA	BGR	NIGERIA	NGA
BURKINA FASO	BFA	NORTHERN IRELAND	NIR
CAMEROON	CMR	NORWAY	NOR
CANADA	CAN	OMAN	OMN
CHILE	CHL	PAKISTAN	PAK
CHINA MAINLAND	CHN	PANAMA	PAN
COLOMBIA	COL	PERU	PER
COSTA RICA	CRI	PHILIPPINES	PHL
CROATIA	HRV	POLAND	POL
CUBA	CUB	PORTUGAL	PRT
CYPRUS	CYP	QATAR	QAT
CZECH REPUBLIC	CZE	REPUBLIC OF GEORGIA	GEO
DENMARK	DNK	ROMANIA	ROU
ECUADOR	ECU	RUSSIA	RUS
EGYPT	EGY	SAUDI ARABIA	SAU
ENGLAND	UK	SCOTLAND	ESC
ESTONIA	EST	SENEGAL	SEN
ETHIOPIA	ETH	SERBIA	SRB
FINLAND	FIN	SINGAPORE	SGP
FRANCE	FRA	SLOVAKIA	SVK
GAMBIA	GMB	SLOVENIA	SVN
GERMANY (FED REP GER)	DEU	SOUTH AFRICA	ZAF
GHANA	GHA	SOUTH KOREA	PRK
GREECE	GRC	SPAIN	ESP
HONG KONG	HKG	SRI LANKA	LKA
HUNGARY	HUN	SWEDEN	SWE
ICELAND	ISL	SWITZERLAND	CHE
INDIA	IND	TAIWAN	TW

INDONESIA	IDN	TANZANIA	TZA
IRAN	IRN	THAILAND	THA
IRELAND	IRL	TUNISIA	TUN
ISRAEL	ISR	TURKEY	TUR
ITALY	ITA	UGANDA	UGA
JAPAN	JPN	UKRAINE	UKR
JORDAN	JOR	UNITED ARAB EMIRATES	ARE
KENYA	KEN	URUGUAY	URY
KUWAIT	KWT	USA	USA
LATVIA	LVA	VENEZUELA	VEN
		VIETNAM	VNM
		WALES	WAL
		ZAMBIA	ZMB
		ZIMBABWE	ZWE

Glosario de abreviaturas de revistas mexicanas

Título	Abreviatura
Acta Botánica Mexicana	Act Bot Mex
Agrociencia	Agrociencia
Annals of Hepatology	Ann Hepatol
Atmosfera	Atmosfera
Boletín de la Sociedad Geológica Mexicana	B Soc Geol Mex
Botanical Sciences	Bot Sci
Ciencias Marinas	Cienc Marinas
Cirugía y Cirujanos	Cir Cir
Convergencia	Convergencia
Economía Mexicana, Nueva Época	Econ Mex N E
Gaceta Medica de México	Gac Med Mex
Geofísica International	Geofis Int
Gestión y Política Publica	Gest Polit Pub
Hidrobiológica	Hidrobiologica
Investigación Bibliotecológica	Invest Bibliotecol
Investigación Económica	Invest Econ
Journal of the Mexican Chemical Society	J Mex Chem Soc
Latin American Economic Review	Lat Am Econ Rev
Madera Bosques	Madera Bosques
Papeles de Población	Papeles de Poblacion
Perfiles Latinoamericanos	Perf Lat
Política y Gobierno	Polit y Gob
Revista Chapingo, Serie Ciencias Forestales y del Ambiente	Rev Chap Cienc Forest Amb
Revista de Investigación Clínica	Rev Invest Clin
Revista Fitotecnia Mexicana	Rev Fitot Mex
Revista Internacional de Contaminación Ambiental	Rev Int Cont Amb
Revista Latinoamericana de Investigación en Matemática Educativa	Rev Lat Invest Mat Ed
Revista Mexicana de Astronomía y Astrofísica	Rev Mex Astron Astrof
Revista Mexicana de Biodiversidad	Rev Mex Biodiv
Revista Mexicana de Ciencias Geológicas	Rev Mex Cienc Geol
Revista Mexicana De Ciencias Pecuarias	Rev Mex Cienc Pec
Revista Mexicana de Física	Rev Mex Fis
Revista Mexicana de Ingeniera Química	Rev Mex Ing Quim
Revista Mexicana de Psicología	Rev Mex Psicol
Salud Mental	Salud Mental
Salud Publica de México	Salud Pub Mex
Tecnología y Ciencias del Agua	Tecnol Cienc Agua
Trimestre Económico	Trimest Econ
Veterinaria México	Vet Mex

Glosario de abreviaturas de Indicadores

%Citdoc	Proporción de documentos citados
ANdoc Scopus	Número anual de documentos en Scopus
ANdoc WoS	Número anual de documentos en WoS
ANp Scopus	Número anual de artículos en Scopus
ANp WoS	Número anual de artículos en WoS
Cita/Ndocxtem	Relación eficiencia-eficacia
citas/doc	Relación de citas por documento
Citdoc	Total de documentos citados
EFS	Eigenfactor
HCP	Highly cited papers
HCPxhab	Número de highly cited papers por cada 100,000 habitantes
HCPxWoS	Número de highly cited papers por cada 1,000 documentos en WoS
HOT	Hot papers

HOTxhab	Número de hot papers por cada 100,000 habitantes
HOTxWoS	Número de top papers por cada 1,000 documentos en WoS
IF	Factor de Impacto
NI	Impacto Normalizado
NSP	Índice de productividad nacional (National Scientific Productivity)
RI	Impacto relativo
SDI	Índice de desarrollo científico (Scientific Development Index)
SJR	SCImago Journal Rank
SNIP	Source normalized impact per paperz
TOP	Top paper
TOPxwoS	Número de hot paper por cada 1,000 documentos en WoS

Introducción

1. Exposición de motivos, justificación y objetivos de la investigación

El siglo XXI, marca un antes y un después para los sistemas de la ciencia, es decir, se evidencia la necesidad e importancia de incorporar programas donde la aplicación de herramientas bibliométricas y la minería de datos, analicen de manera consecuente el dinamismo de las funciones y salidas, en cada uno de sus campos de conocimiento. El análisis de la producción científica de un país tiene gran importancia porque constituye un instrumento para monitorear y desarrollar las actividades de Ciencia y Tecnología, con el objetivo de incrementar la visibilidad internacional de las instituciones, y la calidad de los profesores e investigadores.

El avance de las nuevas tecnologías de la información en esta era de constantes cambios tecnológicos, ha propiciado que, la Ciencia de la Información evolucione sus prácticas y modelos en el almacenamiento, tratamiento, gestión y representación de la información. En este contexto, se comienza a fomentar la creación de grandes bases de datos capaces de almacenar e indexar todo el conocimiento científico que genera la humanidad. Por su parte, el auge de Internet, en el último decenio del pasado siglo y lo que va del presente, incrementó la existencia de bases de datos que almacenan grandes volúmenes de información científica en formato digital.

Diferentes países e instituciones a nivel mundial se han dado a la tarea de crear recursos para analizar este gran volumen de datos científicos; a través de técnicas de inteligencia computacional (Inteligencia artificial) y modelos bibliométricos. Un ejemplo de ello es la fundación en 1957 por Eugene Garfield del *Institute for Scientific Information* (ISI) con los índices de citas (Cawkell & Garfield, 2001; Reuters, 2016), actualmente el *Web of Science* (WoS) pertenece al consorcio norteamericano Thomson Reuters, desde el año 2008, se considera como una de las plataformas más versátil y completa sobre la literatura académica con amplia cobertura temática y un gran número de indicadores bibliométricos que permiten realizar estudios métricos de la ciencia.

La humanidad no solo almacena información de carácter científico, sino que también concentra datos relativos a casi todos los asuntos del quehacer humano. En este contexto hoy día, las naciones se enfrentan a un importante desafío, y es poder transformar todo ese capital de conocimiento que se encuentra accesible y concentrado de manera remota, en información útil para la toma de decisiones. De esta manera, emergen disciplinas para solucionar este desafío como, el descubrimiento de conocimiento en bases de datos o *Knowledge Discovery in Databases* (KDD), encargada de descubrir el conocimiento oculto en estas fuentes y la minería de datos se ocupa de extraer el valor de la información implícita, en las bases de datos y al mismo tiempo visualizar nuevo conocimiento.

Esta tesis constituye una investigación bibliométrica, con un enfoque de minería de datos, sobre la producción científica de México, en el *Web of Science* durante el período comprendido entre el año 2005 y 2015, la cual abarca tanto un análisis de la producción y el impacto de artículos producidos, como la evaluación de las principales revistas de acuerdo a indicadores de última generación. Se considera para el estudio la producción de artículos científicos contenidos en la colección principal del WoS a través de los indicadores que ofrece el *Essential Science Indicator* (ESI) y para la evaluación de las revistas se consideran dos indicadores del *Journal Metrics* de Elsevier y dos de Thomson. Se utilizaron indicadores independientes del tamaño para analizar la evolución de la producción científica mexicana y comparar de manera justa la producción de México con otros países.

Se utilizaron técnicas de inteligencia computacional basadas en redes neuronales para llevar a cabo el análisis exploratorio de la información contenida en *Web of Science*. Para esto se usó el software LabSOM, que implementa tecnología neurocomputacional basada en la familia de algoritmos *Self Organizing Maps* (SOM), la cual es útil para el análisis de datos digitales y la generación automática de representaciones visuales en forma de mapas.

El desarrollo de la investigación consideró el período 2005-2015 y estuvo guiado por las siguientes preguntas:

1. ¿Cómo evoluciona comparativamente la producción científica mexicana en el Web of Science y Scopus?
2. ¿Cómo se posicionan las revistas mexicanas, si se evalúan de acuerdo a diferentes indicadores propuestos por la comunidad científica? ¿Qué justificaría posibles diferencias en las evaluaciones?
3. ¿Cuál es el perfil bibliométrico de la producción científica de México, tomando en cuenta indicadores de productividad, impacto y excelencia y cómo se compara con el perfil de otros países con similar índice de desarrollo científico?
4. ¿Qué utilidad pueden tener las tecnologías basadas en inteligencia computacional para el análisis y la visualización del conocimiento que resulte de las respuestas a las preguntas anteriores?

Objetivo General

Analizar desde una perspectiva bibliométrica y utilizando técnicas de inteligencia computacional:

1. La producción científica mexicana durante 2005-2015 en el *Core Collection del Web of Science*;
2. Usar indicadores de última generación (*Scimago Journal Rank, Impact Factor, Eigenfactor Score y Source Normalized Impact per Paper*) para identificar el núcleo principal de revistas mexicanas con mayor visibilidad en *Web of Science y Scopus* y analizar el efecto diferencial que produce la evaluación de estas revistas en términos de estos indicadores.

Objetivos Específicos

Considerar el período 2005-2015 para:

1. Analizar la forma en que ha evolucionado el conjunto de revistas científicas mexicanas incluidas en el Journal Citation Reports de WoS y el Journal Metrics de Scopus.
2. Identificar y representar el núcleo de revistas mexicanas con mayor influencia internacional en el Web of Science y SCOPUS, considerando cuatro indicadores

que miden el impacto de las citas (*Scimago Journal Rank*, *Impact Factor*, *Eigenfactor Score* y *Source Normalized Impact per Paper*). Utilizar la tecnología neurocomputacional del sistema de software ViBlioSOM para analizar las diferencias que resultan al evaluar este núcleo de revistas en términos de estos cuatro indicadores.

3. Analizar la evolución de la producción científica mexicana en la colección principal del *Web of Science* (WoS) y compararla con la evolución de la producción científica en la base de datos de Scopus.
4. Analizar comparativamente la evolución de la producción científica mexicana en el WoS y Scopus, con el desarrollo número de investigadores en el Sistema Nacional de Investigadores.
5. Comparar la producción científica mexicana en la colección principal del WoS con la producción científica de países de América Latina y otros países, con similar volumen de producción a México.
6. Analizar la producción e impacto de México en WoS, en los principales campos de conocimiento de acuerdo a la clasificación del Essential Science Indicator del WoS.
7. Analizar el Impacto Normalizado, indicadores de producción y de impacto independientes de tamaño para evaluar comparativamente el grado de desarrollo de México y otros países con similar índice de desarrollo científico.
8. Utilizar la tecnología neurocomputacional de ViBlioSOM para caracterizar los perfiles de producción, impacto y excelencia de México y otros países con similar grado de desarrollo científico. Elaborar mapas de conocimiento en los que se puedan representar los grupos de países que tienen perfiles bibliométricos similares.

2. Antecedentes

En la revisión bibliográfica identificamos estudios que emplean herramientas bibliométricas para determinar la evolución de dominios de conocimiento en la ciencia mexicana (Bravo Vinaja, 2005; Castillo-Pérez & Muñoz-Valera, 2014; Chinchilla-Rodríguez, 2004; Hernández, Dray, & Russell, 2013; Perez-Santos & Anaya-Ruiz, 2013). También se encuentran trabajos que estudian la región latinoamericana e incluyen el estudio del comportamiento de la producción científica mexicana (Chinchilla-Rodríguez, Zacca-Gonzalez, Vargas-Quesada, & Moya-Aregon, 2015; Macias-Chapula, 2005; Michan, Russell, Pereyra, Cruset, & Beltran, 2008; Romero, Garcia, & Kiwi, 2009).

Aparecen los estudios que realiza el Foro Consultivo Científico y Tecnológico sobre la producción científica mexicana por cada entidad federativa, considerando además indicadores socioeconómicos (Laclette & Zúñiga-Bello, 2012). Este reporte tuvo en consideración los siguientes indicadores: producción total (suma de documentos publicados en cada institución), número de citas que recibieron las publicaciones de cada una de las instituciones, citas por publicación, porcentaje de documentos que recibieron citas, porcentaje de publicaciones en colaboración internacional, porcentaje de publicaciones que aparecieron en revistas de alto impacto y se empleó como fuente de información la base de datos SCOPUS.

Otro antecedente en el tiempo lo constituye el estudio realizado por Kostoff y colaboradores (Kostoff, et al., 2005), el cual tuvo como objetivo principal identificar la estructura de la literatura científica mexicana en la temática Física del estado Sólido en el *Science Citation Index* para el año 2002.

Después de una revisión bibliográfica, diversas son las investigaciones que han empleado técnicas bibliométricas para estudiar la producción científica mexicana, pero la investigación que se propone aporta nuevos indicadores y complementa a los anteriores debido a que emplea técnicas de inteligencia computacional, basada en el algoritmo de mapas auto-organizados (*Self-Organizing Maps* (SOM)), con el objetivo de representar el perfil bibliométrico de la producción científica mexicana con mayor visibilidad e influencia internacional en las bases de

datos del Web of Science y comparar el desempeño de las revistas mexicanas por medio de indicadores de última generación, basados en las citas.

Esta tesis se estructura en cinco capítulos. La primera sección expone elementos introductorios de la investigación: motivos, justificación, objetivos y antecedentes del estudio. El capítulo I aborda aspectos conceptuales sobre los indicadores para describir la ciencia, el descubrimiento de conocimiento en bases de datos, la minería de datos y visualización de información bibliométrica. El capítulo II expone la metodología empleada para el estudio, así como, las fuentes de información, los indicadores y las herramientas bibliométricas empleadas. El capítulo III presenta los resultados obtenidos en la investigación dividido en tres secciones. La primera sección expone un análisis bibliométrico de las principales revistas científicas mexicanas. El análisis de la producción científica mexicana en el Web of Science durante el período 2005-2015 se presenta en la sección dos del capítulo III. La siguiente sección presenta un análisis multifactorial del perfil bibliométrico de México y un grupo de países, con similar índice de desarrollo científico. El análisis multifactorial utiliza la técnica de mapas auto-organizados y toma en cuenta cuatro indicadores: uno de impacto, otro de producción y dos de excelencia. Por último, se ofrecen las consideraciones finales de la investigación.

Capítulo I: Marco teórico

1. Elementos de análisis de la producción académica

1.1. Indicadores de producción e indicadores basados en citas

Los indicadores bibliométricos constituyen una de las herramientas más utilizadas para la medición del producto de la investigación científica, porque la documentación (independientemente del tipo de soporte) es el vehículo más prolífico y exitoso para la transferencia del conocimiento científico, conjuntamente con su transferencia oral por medio de conferencias y comunicaciones personales (Russell, 2002).

En correspondencia con el entorno informacional en el cual se desenvuelve la ciencia actualmente, es necesario conocer nuevos patrones y comportamientos para analizar el contexto científico, por tal razón una de las nuevas tendencias en el campo de las métricas de la información es: el auge de los estudios a partir del cálculo de indicadores producción e indicadores basados en las citas. Estos últimos, se emplean como un criterio más en el análisis de artículos y publicaciones científicas porque a partir del reconocimiento que recibe una fuente de otra, permite reconocer estudios, conceptos, teorías, metodologías y herramientas científicas que inspiraron al investigador a desarrollar nuevas investigaciones o conocimiento; de esta forma se identifican nuevos patrones e influencias en la investigación científica.

El desarrollo de las bases de datos en los años 60 del siglo XX, trajo consigo la creación de una serie de indicadores utilizados en el estudio de la producción o volumen de investigación, así como, la medición del desempeño de las revistas, instituciones, países y grupos de investigación en colaboración a través de indicadores basados en las citas, tradicionalmente extraídos de las bases de datos de citas como *Science Citation Index* (SCI), *Social Science Citation Index* (SSCI) y *Journal Citation Reports* (JCR); y en la actualidad, las bases de datos: Scopus, Web of Science, *Google Scholar*, *Microsoft Academic Search* y *CiteSeerX* se utilizan con frecuencia para investigar el crecimiento y comportamiento científico.

Los resultados de la ciencia (output), se analizan a través de la cuantificación de la producción científica que genera un país, institución o investigador (Abramo, D'Angelo, & Rosati, 2013; Andras, 2011; Arencibia Jorge & Moya-Anegon, 2008;

Barker, 2007; Melkers, 1993). Para lograr obtener una visión multidimensional de la ciencia se considera necesario la combinación de los indicadores existentes y la revisión por pares (*peer review*). De manera general, cuando se persiga analizar aspectos de carácter cualitativo, se impone recurrir a las opiniones de expertos mientras que para determinar aquellos aspectos de corte cuantitativo o si se pretende estudiar y comparar el comportamiento de investigadores, instituciones o países, el uso de indicadores bibliométricos se hace necesario (Piedras, 2009).

El *peer review* se reconoce como el sistema de evaluación de la actividad científica con mayor tradición. Este método se basa en “el conocimiento y experiencia acumulados por determinados especialistas en su campo de investigación, lo que los convierte en poseedores de criterios sólidos y de carácter casi único para poder valorar adecuadamente las características, tendencias y resultados de la actividad de investigación en dicho campo” (Suarez Balseiro & Maura Sardó, 2005).

Los indicadores de producción tienen una función descriptiva, en la cual estos están encaminados a producir información sobre el estado del volumen del sistema científico objeto de estudio, por ejemplo, cuántos artículos publican los investigadores, número artículos por centros de investigación. Específicamente en la presente investigación se determinó incluir los siguientes indicadores de producción: total de artículos publicados por (Ndoc): años, áreas de investigación, países, revistas; y porcentaje del total de artículos publicados (%Ndoc) por: años, áreas de investigación, países, revistas; los cuales miden el volumen de la producción científica mexicana durante el período de estudio.

Los índices de citas ofrecen una visión valorativa de la influencia científica, ofrecen una óptica de la calidad de la investigación diferente al juicio de expertos porque toma en consideración el impacto y repercusión que ha tenido la investigación dentro de la comunidad, esto no significa, que no exista correlación entre las observaciones que derivan de ambos métodos (*peer review* e índices de citas) (Arencibia Jorge, 2010).

Las tres bases de datos más importantes disponibles hoy en día, para la realización de análisis de citas son: WoS, Scopus y Google Académico, las cuales

calculan sus propios índices basados en las citas (Tabla 1). Existen otras bases de datos disponibles, pero cubren un número limitado de campos científicos. Por otra parte, algunas de estas bases de datos no contienen los datos sobre las referencias de las publicaciones, por tanto, no se pueden utilizar para calcular los indicadores basados en las citas.

Tabla 1. Indicadores basados en citas con mayor uso derivados de Bases de datos como Web of Science y Scopus

Algunos Indicadores basados en las citas calculados por bases de datos	Conteo de citas
	Journal Impact Factor
	Índice de Inmediatez
	Índice H
	Vida media de las citas
	Vida media citante
	Autocitas
	Citas/No. de papers no citados
	Article Influence® score
	SJR Rank
	Eigenfactor® score
	Citas por documento

La ciencia, a pesar de sus múltiples aristas, es sensible de ser estudiada bajo tres aspectos fundamentales: productividad, impacto y excelencia científica, por ello en la actualidad la mayoría de las metodologías de indicadores bibliométricos o cuantitativos miden la ciencia en sus múltiples ángulos, y toman en cuenta indicadores de productividad, impacto, visibilidad, y excelencia (Arencibia Jorge, 2010; Callon & Courtial, 1995; González de Dios, Moya, & Mateos Hernández, 1997; Licea de Arenas, 1993; Rodríguez Sánchez, 2012; Sancho, 1990; Schneider, 2006).

Los indicadores basados en las citas ofrecen diferentes interpretaciones: como una medida del impacto científico de la investigación (J. R. Cole, 2000; Van Raan, 2004); de prestigio académico para otorgar premios y honores; también reflejan la calidad, relevancia e impacto de un documento en la obra de otro, para

de esta forma demostrar que, las citas se realizan a los aportes científicos (Lutz Bornmann & Hans-Dieter, 2008).

Mencionar además como, este tipo de indicador se aplica en los análisis de contexto y contenido para, derivar diferentes interpretaciones y significados del comportamiento de las citas en un artículo y por consiguiente medir el impacto científico. Los análisis de contexto, tienen el objetivo de ilustrar la relación entre el documento y las citas, para ello se genera un esquema de clasificación o taxonomía sobre la funcionalidad de las citas; se determina el contexto en el cual fue creada la cita (Cronin, 1984). Por su parte, el análisis de contenido de las citas describe el contenido semántico del documento citado con el fin de caracterizar las obras citadas (Lutz Bornmann & Hans-Dieter, 2008).

1.2. La eficiencia científica medida a través de indicadores de productividad

Los términos de eficiencia y efectividad provienen de las ciencias económicas, pero hoy día se aplican a todos los ámbitos, incluyendo el sector científico. En el caso de la ciencia, la productividad se interpreta como eficiencia y el rendimiento del sistema como efectividad, términos que se emplean en los estudios métricos del grupo de cienciómetras liderado por el Dr. Humberto Carrillo en el Laboratorio de Dinámica no lineal, de la Facultad de Ciencias de la Universidad Nacional Autónoma de México (Villaseñor-García, Arencibia Jorge, & Carrillo-Calvet, 2016). No necesariamente una alta producción se encuentra relacionada con el rendimiento o efectividad (Abramo & D'Angelo, 2016b). La interpretación económica de la ciencia se basa en la productividad de la investigación (eficiencia) y en el rendimiento (efectividad) a través del cálculo de indicadores que tomen en cuenta las medidas de inversión en la investigación (Lutz Bornmann & Haunschild, 2016).

Los metristas emplean diferentes indicadores para medir la eficiencia de la investigación, pero deben ser muy cuidadosos en los índices que calculan para realizar estudios evaluativos u obtener ranking (clasificación de mayor a menor), que se utilizan posteriormente en la toma de decisiones. En este sentido, es necesario prestar atención a los indicadores que incluyen medidas de input (gastos

en investigación, número de investigadores activos, gastos en investigación) en relación con la productividad; otro aspecto es que los estudios a nivel macro, meso o micro no deben considerar los indicadores dependientes del tamaño (número de citas, número de Highly Cited papers) solamente, sino que estos deben ser combinados con otros indicadores para obtener indicadores de tamaño independiente (Glanzel, Thijs, & Debackere, 2016), también los gobiernos e instituciones de investigación deben ofrecer los datos necesarios para calcular este tipo de indicadores, es decir, nombre y afiliación de los científicos, de investigación, rango académico, recursos asignados.

Por último, con el objetivo de evitar distorsiones es necesario clasificar el objeto de estudio (artículo, revista, institución, país) por campo de investigación, a lo que se denomina normalización, proceso que se realiza debido a los diferentes patrones de citación que existe a través de las áreas de la ciencia, (Abramo, D'Angelo, & Di Costa, 2008), la normalización de las citas por campos de investigación permite analizar, comparar y evaluar diferentes artículos, revistas, instituciones, y países de acuerdo a su eficiencia y rendimiento de forma justa.

1.3. Consideraciones sobre los indicadores de excelencia

Los artículos altamente citados (Highly Cited Paper o HCP) se consideran uno de los indicadores más empleados para estudiar la excelencia científica (Lutz Bornmann, Leydesdorff, Walch-Solimena, & Ettl, 2011; Lutz Bornmann, Stefaner, de Moya Anegón, & Mutz, 2016), determinar rankings de centros o facultades con mayor desempeño (Bloch, Schneider, & Sinkjaer, 2016; Lutz Bornmann, Stefaner, de Moya Anegón, & Mutz, 2014; Waltman, et al., 2012), en el otorgamiento de premios (Rodríguez-Navarro, 2011) y para identificar áreas científicas con alta excelencia (Baudoin, Peltier, Graillet-Gak, & Haeffner-Cavaillon, 2004; Lutz Bornmann & Leydesdorff, 2016).

Los HCP se pueden definir de diferentes maneras, en esta investigación se toma el concepto de la herramienta *Essential Science Indicators* producido por *Thomson Reuters*, se define como el número de artículos que recibieron un total de

citas superior al 1% de citas cuando se compara con todos los artículos publicados en el mismo año y en la misma disciplina ("Web of Science," 2016).

Los artículos altamente citados tienen una naturaleza compleja, la cual está determinada por las siguientes características, extensas coautorías de diferentes países con altos niveles de colaboración internacional, se publican en revistas con alto factor de impacto, suelen ser citados por científicos extranjeros, existe una mayor representación de los artículos de revisión como artículos que reciben un alto número de citas con respecto a otro tipo de artículos, y el porcentaje de autocitas es muy bajo para los artículos altamente citados (Dag W Aksnes, 2003; Miyairi & Chang, 2012; Tijssen, Visser, & van Leeuwen, 2001).

El hecho que un artículo reciba alto número de citas y se considere como HCP, lleva implícito las decisiones de muchos investigadores para citar determinado artículo; en este sentido, la calidad y visibilidad juegan un papel fundamental en este proceso (Lutz Bornmann, 2014). La calidad está relacionada con la estructura del conocimiento científico (conocimiento núcleo y de frontera) (S. Cole, 2000); en la actualidad una gran parte de los artículos publicados tienen un corto período de importancia y no trascienden, es decir, un bajo número de trabajos se convierten en literatura núcleo para un campo de la ciencia, lo cual provoca que reciban un alto número de citas.

Por su parte la visibilidad, es relevante porque generalmente los artículos altamente citados son concebidos por un promedio de 8 autores, factor que favorece la visibilidad de la investigación y por tanto obtener un mayor número de citas (Dag W Aksnes, 2003). Además, un artículo con un alto número de citas de otras contribuciones implica que un mayor número de especialistas consideró este trabajo; por tanto, aumenta la visibilidad. En general influyen diversos factores sociales para que un autor cite un documento en particular; y específicamente los científicos tienden a citar las contribuciones que son útiles para su propia investigación (Cronin, 1984). Además, gran parte de lo que se publica son estudios que no se consideran núcleo en un campo científico y existen diferencias en los hábitos de citación para cada área de investigación.

Los trabajos altamente citados son concebidos por autores con altos niveles de citas, pioneros e influyentes en sus campos, además son productivos y pueden obtener una cantidad desproporcionada de los fondos de subvención; por tanto, sus trabajos en virtud de su visibilidad son un factor clave de efectos de la reputación en el mundo académico. Por otra parte, la investigación ofrece un enfoque de enlace o relación entre ciencia-práctica, la industria y la academia, estudios que tienen implicaciones prácticas de la vida real que pueden impactar a nivel individual y social (Dag W. Aksnes & Rip, 2009; Antonakis, Bastardoz, Liu, & Schriesheim, 2014). Importante entender lo que distingue a los artículos altamente citados de los menos conocidos, diferencia que se centra específicamente en la aplicación de los métodos académicos de la investigación, y la validez de las conclusiones del artículo.

2. Elementos de análisis del desempeño de las revistas académicas

Las revistas científicas constituyen el canal de comunicación de los científicos e investigadores, se consideran como un criterio para evaluar el rendimiento de la investigación de un científico. En este contexto, un ranking de revistas constituye una herramienta importante para los investigadores en el momento de elegir la publicación adecuada para canalizar sus resultados de investigación, los bibliotecarios pueden planificar una mejor selección para incrementar las colecciones y los editores conocen el grado de calidad de su revista.

Los rankings de revistas se concibieron con el objetivo de evaluar el impacto de las revistas en campos específicos, los más empleados en la actualidad son los basados en las estadísticas de citas, como el factor de impacto (IF) creado por Eugene Garfield desde 1963 y se reporta por el *Journal Citation Reports* de *Thompson Reuters*; los nuevos índices de Google Académico, *Journal Metrics*, y *SCImago Journal Rank* (SJR). Además, en la literatura se encuentran estudios que han realizado ranking de revistas en diferentes campos de la ciencia con el propósito de determinar, a través del cálculo de indicadores basados en las citas, la calidad de las revistas (Pajic, 2015; Parameswaran & Sebastian, 2006; Seiler & Wohlrabe, 2014; Serenko & Dohan, 2011; Vanclay, 2011).

Clasificar las revistas científicas en un orden, constituye una herramienta importante para las instituciones académicas donde existen esquemas de evaluación y es un método para reducir la incertidumbre en el proceso de selección de la publicación y en la medición/evaluación de los resultados de la investigación (Pajic, 2015; Polonsky, 2004). Proporcionan cierta estabilidad para la realización de estudios evaluativos para un período de tiempo; siempre y cuando no influyan factores como el aumento desproporcional de autocitas y la aparición de hot papers (artículos calientes); es un comportamiento común en revistas de menor rango donde las pequeñas fluctuaciones en el número de citas pueden producir grandes cambios en el ranking. A medida que la calidad y el prestigio de las revistas aumenta, la variabilidad relativa del factor de impacto disminuye, y por lo tanto el ranking de revistas ofrece una tendencia a ser más fiable y estable para su nivel (Pajic, 2015).

Otra de las ventajas de los rankings es como los académicos reciben altos salarios cuando publican en revistas de primer nivel, algunas universidades han desarrollado políticas para recompensar financieramente las facultades que tienen un mayor número de artículos en revistas núcleo. Las bibliotecas pueden utilizar esta clasificación para asignar recursos y establecer prioridades para la suscripción. Los editores pueden saber acerca de la posición relativa de su revista y de esta manera desarrollar campañas de marketing, atraer a los suscriptores y sugerir temas muy exclusivos y actuales (Serenko & Dohan, 2011).

Existen varios enfoques posibles para obtener un ranking de revistas, el método cualitativo que se basa generalmente en las encuestas que registran las percepciones de los encuestados (Peffer & Tang, 2003). Por otro lado, el método cuantitativo que se fundamenta en el factor de impacto de la revista, que es una métrica para medir el impacto de las revistas en varias áreas temáticas; este tipo de análisis de citas se considera la metodología más objetiva para evaluar la calidad de la revista y obtener un ranking.

Existen alternativas al factor de impacto que junto al criterio de expertos permiten validar la influencia de las revistas como; el *Source Normalized Impact per Paper* (SNIP), el *SCImago Journal Ranking* (SJR) y *Eigenfactor* (EFS); estos

indicadores se consideran de importancia para construir ranking de revistas en bases de datos como *Web of Science* y *Scopus* (Falagas, Kouranos, Arencibia-Jorge, & Karageorgopoulos, 2008; Leydesdorff & Opthof, 2010; Henk F. Moed, 2011).

2.1. El Factor de Impacto de Garfield

El *factor de impacto* (IF), mide la importancia de una publicación científica, es calculado anualmente por las publicaciones que se encuentran recogidas en el *Journal Citation Reports* del *Web of Science*, *Thomson-Reuters*. En la literatura se cuestiona el uso del IF como indicador que ofrece información sobre la calidad de las publicaciones. Además, como su período de cálculo base para citas es muy corto, no toma en cuenta que, los artículos clásicos son citados más frecuentemente que los originales, incluso décadas después de haber sido escritos. Finalmente, sus resultados no son comparables en distintas áreas de investigación (Garfield & Sher, 1963).

2.2. Eigenfactor del WoS

Eigenfactor (EFS) tiene sus antecedentes como *eigenvector centrality*, indicador que se introdujo por el sociólogo Phillip Bonacich en 1972 como una forma de cuantificar el estado o la popularidad de un individuo dentro de una red de comunicación. El objetivo de Bonacich era utilizar la estructura de una red para identificar quiénes eran las personas importantes en la red (Bergstrom, West, & Wiseman, 2008).

Luego Carl T. Bergstrom, experto en *Biología de la Información* y *Métrica de la Información Científica* en la Universidad de Washington diseñó en el año 2007 el indicador: *eigenfactor*.

Identificar cuántos nos citan, quiénes y cómo lo hacen, es lo que valora el *Eigenfactor*; cuando este indicador es combinado con otros se obtiene un ranking de revistas científicas más confiables para validar la influencia de las publicaciones (Wesley-Smith, Bergstrom, & West, 2016); en la red de referencias que citan una revista. Además, evalúa la importancia y posición de cada revista en la red; no proporciona directamente el costo de las publicaciones, pero los bibliotecarios

pueden determinar a través de este indicador el costo-efectividad de una colección de revistas.

Cada área de la ciencia tiene diferentes comportamientos de las citas y las escalas para producir/recibir citas también difieren, en ese sentido, el uso de la red de citas, en el cual se basa el algoritmo del cálculo de *Eigenfactor* permite de forma automática identificar estas diferencias y comparar las diferentes áreas de investigación. Es un indicador que para su cálculo se toman las citas recibidas en un período de cinco años (Bergstrom, 2007; West, Bergstrom, & Bergstrom, 2010).

2.3. *Source normalized impact per paper*

El indicador SNIP fue concebido por Henk Moed en el 2010, y los valores son reportados en la base de datos Scopus y en la herramienta *Journal Metrics*. SNIP calcula el impacto de las citas en las revistas científicas utilizando un enfoque normalizado de la fuente “*source normalized approach*”, con el objetivo de rectificar las diferentes prácticas de citas sin recurrir a un sistema de clasificación temática en el que se definen explícitamente los límites de los campos temáticos de las revistas. *Source normalized approach* realiza un conteo de las citas por cada campo de investigación basado en las características de las fuentes que originaron la cita (Waltman, van Eck, van Leeuwen, & Visser, 2013a).

Este indicador mide el impacto de las citas de contexto en la revista conocido en inglés como: *journal's contextual citation impact*. Para ello toma en consideración las características del grupo de artículos que citan una revista. En este enfoque, se calcula la frecuencia que un artículo de una revista cita (citing papers) otros artículos procedentes de otras revistas, a lo que se denomina *journal's subject field* (Moed, 2010).

En términos generales, el SNIP evalúa el impacto de las citas de contexto en la revista científica, escenario que está determinado por las referencias y además por el grado de cubrimiento que la base de datos ofrece para la revista. El SNIP es un indicador que ofrece potencialidades y debilidades (Waltman, et al., 2013b) para evaluar el desempeño de las revistas; pero todavía no existe un único indicador para

validar las publicaciones científicas. El proceso de comunicación científica es un concepto multidimensional que no puede ser expresado en una sola medida.

En términos generales, SNIP se convierte en un indicador con importantes ventajas sobre las tradiciones enfoques de normalización en los sistemas de clasificación de las revistas y los índices basados en las citas para validar el desempeño de las publicaciones (Moed, 2010; H. F. Moed, 2011; Schoepfel & Prost, 2009; Waltman, van Eck, van Leeuwen, & Visser, 2013b) debido a los siguientes puntos:

- Este indicador permite delimitar el conjunto de artículos que citan una revista sin depender de un sistema de clasificación por temática para las revistas, sino que se basa en las relaciones de las citas.
- El nuevo indicador corrige las diferencias en las referencias encontradas entre campos temáticos, específicamente la frecuencia con la que los autores citan otros documentos y la velocidad en la maduración del impacto de las citas.
- Además, corrige las diferencias en la cobertura de la base de datos para diferentes campos de investigación. Ejemplo: un bajo porcentaje de referencias activas conduce a un menor valor del indicador, *database citation potentials*, que tiende a ofrecer valores más altos del SNIP en comparación con el *raw impact per paper indicator*. Variables que intervienen en el cálculo del SNIP y que se definen el capítulo siguiente.
- Permite corregir las diferencias en las categorías temáticas para las revistas multidisciplinarias, como Nature o Science.

2.4. SCImago Journal Rank (SJR)

SCImago Journal Rank (SJR) es un indicador creado por el grupo de investigación SCImago en el año 2009. SJR es un índice de tamaño independiente destinado a medir el prestigio promedio por artículo de una revista que está siendo evaluada. Este indicador tiene la función de representar el impacto científico de una revista, desde una perspectiva cuantitativa y cualitativa.

La idea esencial que subyace en la aplicación de indicador en la evaluación de revistas científicas es la asignación de pesos a las citas bibliográficas en base a la importancia de las revistas que emitieron las citas. Las citas hechas por las

revistas importantes se les asigna mayor peso que las emitidas por revistas menos importantes. Esta "importancia" se calcula de forma recursiva, es decir, las revistas importantes serán los que a su vez reciben muchas citas de otras revistas importantes. Este proceso se determina gracias a la aplicación del algoritmo de PageRank creado por Google en 1998. Por ello, este indicador se considera como un *journal prestige indicator* (González-Pereira, Guerrero-Bote, & Moya-Anegón, 2009).

3. Descubrimiento de conocimiento, minería y visualización de información en bases de datos bibliográficas

3.1. El descubrimiento de conocimiento en bases de datos (KDD)

En la década de los años 80 del siglo pasado, aparecen tecnologías como las computadoras y los lenguajes de cuarta generación. Estos avances tecnológicos provocaron un cambio de pensamiento y acción; el usuario comprendió que los datos no solo se necesitaban para el procesamiento de transacciones online, también los datos facilitaban tomar decisiones a la alta dirección para mantener las instituciones en entornos muy competitivos. A partir de este momento, las bases de datos se utilizaron para el análisis de los datos almacenados y recuperar información. Con el transcurso del tiempo, el volumen de datos se incrementó y las herramientas para gestionar los datos no eran las adecuadas y surgió la necesidad de herramientas de análisis para grandes cantidades de datos.

En la era moderna, la mayoría de la información está disponible a través de documentos en lenguaje natural no estructurados, esto se debe al crecimiento exponencial de la web, las bibliotecas digitales, la documentación en formato digital, etc. Sin embargo, este gran cúmulo de datos lleva intrínsecamente información estratégica, a la cual no es fácil de acceder a través de técnicas clásicas de recuperación de información.

El descubrimiento de conocimiento en bases de datos o en inglés *Knowledge Discovery in Databases* (KDD) fue acuñado en el primer "workshop" de KDD en 1989 y se define como: el conjunto de acciones que se desarrollan para analizar las grandes cantidades de datos almacenadas por las organizaciones, con el fin de

obtener información de interés para mejorar el funcionamiento y rendimiento de las mismas (U. M. Fayyad, Piatetsky-Shapiro, Smyth, & Uthurusamy, 1996).

Bernstein, define KDD como el resultado de un proceso exploratorio que, implica la construcción de modelos a partir de los datos almacenados (Bernstein , Provost, & Hill, 2005).

La extracción de información (IE) es un método con la precisión razonable capaz de transformar los datos no estructurados que están en una base de datos estructurada, en formas intermedias. Las formas intermedias más habituales son: bolsa de palabras, N-grams, palabras clave, multitérminos, conceptos de jerarquía, mapas conceptuales (Sirsat, Chavan, & Deshpande, 2014).

Hoy en día, mucho más allá del reconocimiento de patrones, a este proceso se le ha dado una variedad de nombres, incluyendo: la minería de datos, extracción de conocimiento, el descubrimiento de información, recolección de información, extracción de información (U. Fayyad, G. Piatetsky-Shapiro, & P. Smyth, 1996), entre otras, pero en esta investigación se empleará el término: descubrimiento de conocimiento en bases de datos o en inglés *Knowledge Discovery in Databases* (KDD).

El objetivo principal del KDD es la extracción del conocimiento que se encuentra dentro de grandes cantidades de datos en bruto, identificar los patrones más importantes y significativos, así como, representar los objetivos del usuario. El descubrimiento de conocimiento oculto, es posible gracias a la aplicación de sofisticadas técnicas como la inteligencia artificial, la minería de datos y la visualización de información, las cuales permiten encontrar patrones y relaciones dentro de los datos permitiendo la creación de modelos, y representaciones abstractas de la realidad.

El valor real de los datos reside en la información que se puede extraer de ellos, información que ayude a tomar decisiones o mejorar la comprensión de los fenómenos que se presentan en la realidad. Hoy, más que nunca, los métodos analíticos avanzados son el arma secreta de muchos negocios exitosos. Empleando métodos analíticos avanzados para la explotación de datos, los negocios

incrementan sus ganancias, maximizan la eficiencia operativa, reducen costos y mejoran la satisfacción del cliente (Gamarra, Guerrero, & Montero, 2016; Holzinger, Dehmer, & Jurisica, 2014; Witten & Frank, 2005). El reto no es la extracción de información significativa de los datos no estructurados sino hacer los datos útiles al usuario final.

3.2. La minería de datos como parte de la metodología de descubrimiento de conocimiento en bases de datos

Desde el punto de vista académico, el término Minería de datos en inglés *Data mining* se considera como un proceso dentro de la metodología KDD (Usama Fayyad, Gregory Piatetsky-Shapiro, & Padhraic Smyth, 1996; Gullo, 2014; Holmes, 2013; Holzinger, et al., 2014; Senger, Leach, Ralph, & David, 2008; Wanderley, Tacla, Barthes, & Paraiso, 2015); pero en el ámbito empresarial ambos términos se emplean de manera indistinta. En este trabajo de investigación se emplea el término minería de datos como fase del proceso de KDD.

En este sentido, la metodología KDD involucra varios algoritmos de minería de datos o *data mining*. Esta última se define por Fayyad como la etapa de descubrimiento del KDD que consiste en el uso de algoritmos concretos que generan una enumeración de patrones a partir de los datos pre-procesados (U. Fayyad & Stolorz, 1997).

KDD es una metodología interactiva e iterativa que incluye varios pasos, una de las primeras propuestas fue planteada por Fayyad en 1996 (Figura 1) y en ella el proceso se divide en las siguientes fases: comprensión del problema, selección de la base de datos objetivo del proceso de descubrimiento, limpieza y preprocesado de dicha base de datos, reducción, elección del método que se va a usar para resolver el problema, elección del algoritmo, ejecución del algoritmo de Minería de datos elegido, interpretación de los patrones obtenidos y consolidación del conocimiento obtenido (Usama Fayyad, et al., 1996).

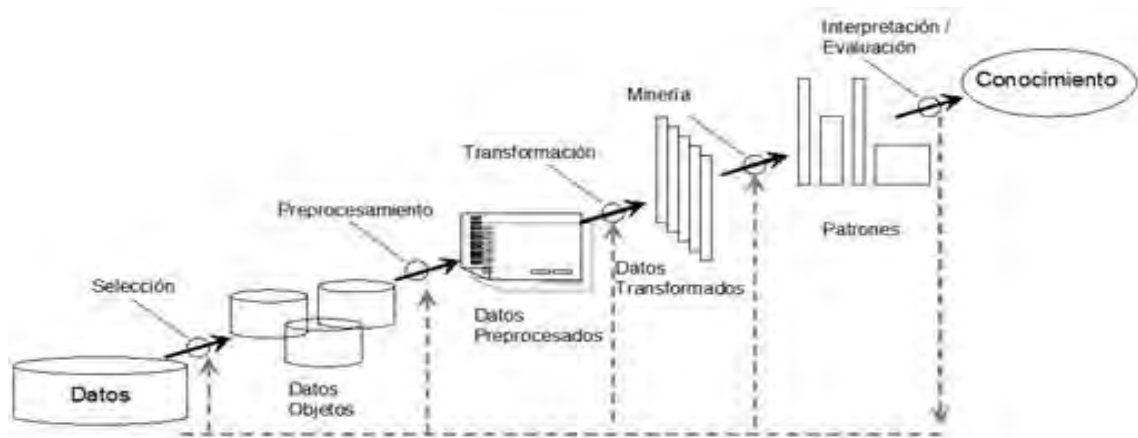


Figura 1. Fases de la metodología para el descubrimiento de conocimiento en bases de datos (KDD) (Fayyad, et al., 1996)

KDD cuenta de los siguientes pasos según el criterio de varios autores desde los que definieron esta herramienta hasta alguna de las más recientes propuestas (Liao, Chu, & Hsiao, 2012):

- Determinar las fuentes de información que pueden ser útiles y dónde conseguirlas.
- Diseñar el esquema del almacén de datos (*Data Warehouse*) para unificar toda la información recuperada.
- Implantación del almacén de datos para facilitar la navegación y visualización previa de sus datos, y determinar los elementos que pueden ser estudiados.
- Selección, limpieza y transformación de los datos que se van a analizar. La selección incluye tanto una criba o fusión horizontal (filas) como vertical (atributos). La limpieza y pre-procesamiento de datos implica el trazado de una estrategia para eliminar los ruidos, textos incompletos, en fin, toda la información que no es necesaria.
- Minería de datos: incluye la selección de la tarea de descubrimiento a realizar, por ejemplo, clasificación, agrupamiento o clustering, regresión, etc. El proceso de minería de datos, busca patrones que puedan expresarse como modelo.
- Estimación, interpretación, transformación y representación de los patrones extraídos.

- Difusión y utilidad del nuevo conocimiento para añadir el conocimiento descubierto al sistema con el propósito de resolver conflictos potenciales con el conocimiento existente, trazar acciones y tomar decisiones.

La Minería de datos es la fase de mayor importancia y tiene como tareas determinar el patrón de descubrimiento, la agrupación, clasificación, y resumen (Fu, 2011), específicamente, se ocupa de la extracción de patrones interesantes por la elección de un método de extracción de datos o tareas específicas (por ejemplo, el resumen, la clasificación, agrupación, la regresión, y así sucesivamente), luego los algoritmos adecuados para llevar a cabo la tarea a la mano, y por último una representación adecuada de los resultados de salida; por ello se considera la fase de mayor importancia en el proceso de KDD (Kurgan & Musilek, 2006).

3.3. Técnicas de la minería de datos

El volumen de datos almacenados y disponibles de forma remota en las bases de datos, ha generado la necesidad de explotar estos datos y transformarlos en información, es decir, los datos constituyen el ingrediente o materia prima de la minería de datos, la cual ha estado bajo desarrollo por décadas, y se ha nutrido de otras áreas de investigación como las estadísticas, inteligencia artificial y el aprendizaje de máquinas. Hoy, la madurez de las técnicas y método de la minería de datos, junto con los motores de las bases de datos relacionales de alto desempeño, hicieron que de la minería de datos sea una práctica común para los entornos de data *warehouse* (almacén de datos).

En el núcleo del proceso de minería de datos se destaca la aplicación de diferentes métodos, que permiten el descubrimiento de patrones de distintos tipos, entre los que se pueden destacar, por mencionar algunos, clasificación, caracterización, asociación, “clustering” y discriminación, los cuales han sido conceptualizados por diferentes autores (Han & Kamber, 2000; Kishida, 2003; Kuncheva, 2004; Kurgan & Musilek, 2006; Leskovec, Kleinberg, & Faloutsos, 2007; Liao, et al., 2012; Pyle, 1999).

Una de las primeras y más completas conceptualizaciones de las técnicas de minería de datos lo presentó Fayyad en su artículo (U. Fayyad & Stolorz, 1997),

además, la fase se compone de varias técnicas las cuales se mencionan a continuación: *Predictive modeling (modelo predictivo)*, *Clustering (agrupación)*, *Data summarization (resumen de datos)*, *Dependency modeling*, *Change and deviation detection (detección de cambio y desviación)*.

Siti Khadijah, en la novena Conferencia Internacional sobre Ciencia Cognitiva (*9th International Conference on Cognitive Science*), planteó que, una de las técnicas de minería de datos más usada es: Clustering o agrupamiento, seguido de la Clasificación, Sequential pattern o patrón secuencial, Predicción, y Reglas de asociación (PhridviRaj & GuruRao, 2014; Shu-Hsien, Pei-Hui, & Pei-Yuan, 2012); en este sentido, la presente investigación empleará la técnica de *Clustering*, para representar la información.

Clustering es la técnica la cual toma un conjunto de entidades representadas por registros de una base de datos, la divide en grupos o clusters de manera tal que los elementos dentro de cada grupo sean similares. El objetivo del *clustering* es reducir la cantidad de datos mediante la caracterización o agrupamiento de datos con características similares. Esta agrupación es acorde con los procesos humanos de información y una de las motivaciones para usar algoritmos *clustering* es proveer herramientas automáticas que ayuden a la construcción de taxonomías. Los métodos pueden también ser usados para minimizar los efectos de los factores humanos que afectan el proceso de clasificación (Kishida, 2003; Kostoff, et al., 2007; Tseng & Tsay, 2013; Villaseñor García, 2016).

3.4. Visualización de información

La teoría de la visualización se inició hace más de 2400 años. El filósofo Platón comprendió que percibimos los objetos a través de los sentidos con el uso de la mente. Por lo cual, nuestra percepción del mundo real es una traducción de la energía física de nuestro entorno en señales neuronales codificados. La mente interpreta continuamente y categóricamente nuestra percepción del entorno que nos rodea. Ejemplo de ellos es: el uso de una computadora como fuente de entrada para las funciones de procesamiento de la mente, las interfaces de texto no sólo reducen la complejidad de la interfaz, y sino

también filtran el uso de las más potentes funciones de procesamiento de información en la mente humana.

La capacidad de visualizar las implicaciones de los datos, es tan antigua como la humanidad misma. Los orígenes de la visualización de la información, se identifican por la búsqueda del *insight* a partir de la entropía que existe en un gran volumen de datos, a partir de dos perspectivas: estructura y cambio. Sus antecedentes se localizan en la historia del desarrollo del conocimiento y del lenguaje visual, estrechamente fusionado a la influencia de los progresivos avances tecnológicos y científicos, con dos propósitos esenciales: descubrir y explicar (C. Chen, 2006).

El término Visualización de la Información se manejó por primera vez en 1989; en esta etapa el término refiere a una propuesta de arquitectura cognitiva para interfaces de usuario interactivas. Luego, el estudio de las representaciones figuras de la información desde distintas disciplinas, y bajo la influencia de los paradigmas cognitivo y tecnológico, la *Human Computer Interaction* (HCI), la fortaleció como área de conocimiento distintiva de investigación, en función de poder desarrollar sus aplicaciones de forma satisfactoria en el manejo y representación gráfica de la información. Entonces se consolida como un área de investigación, con bases interdisciplinarias, y enfocadas en la representación visual de un gran volumen de información, en el uso de técnicas gráficas interactivas para comprender y analizar los datos de forma intuitiva, detectando y explorando sus relaciones y patrones de comportamiento (Card, MacKinlay, & Shneiderman, 1999).

La Visualización de Información surge de la informática (los gráficos computacionales y la interacción persona - computadora), la psicología (los procesos cognitivos y de la psicología de la percepción), de la semiótica (significado de los símbolos), diseño gráfico (para artefactos bien adaptados a su propósito), de la cartografía (para crear las representaciones visuales que son abstracciones cuidadosamente elegidas del mundo), y del arte (que han refinado los métodos para transportar el significado visual), y actualmente cuenta con su propia agenda de investigación (C. Chen, 2002; C. Chen, 2006).

La visualización de información es el proceso de **interiorización del conocimiento** mediante la **percepción** de la información (Dürsteler, 2000). También se puede enfocar sobre los aspectos cognitivos y perceptuales de los seres humanos. La informática en este caso es el medio para obtener las visualizaciones y “traducir” lo cognitivo a un lenguaje informático (Boyack, Wylie, & Davidson, 2002; Guzmán-Sánchez, 2009; Marcus, 2015).

Los diferentes enfoques conceptuales de la visualización de información se enmarcan en tres puntos de vistas (Purchase, Andrienko, Jankun-Kelly, & Ward, 2008), la interpretación de una visualización a través de su forma física; en la exploración y manipulación de la representación exterior por el lector; y la exploración y manipulación del modelo de datos interno por el sistema en orden para descubrir interrelaciones, tendencias y patrones, enfoque en el cual se enmarca las representaciones visuales que se construyen en esta investigación.

Chaomei Chen, define el término de visualización de información, como aquellas representaciones gráficas de información procedentes de una computadora que tienen el objetivo de constituir visualmente espacios y estructuras de información que faciliten una rápida asimilación y comprensión, y la posibilidad de identificar y extraer patrones a partir de una gran cantidad de información, incrementando el valor. (C. Chen, 2010).

En el presente, el crecimiento exponencial y la complejidad de la información que se genera, se torna casi imposible para el ser humano poder extraer conclusiones, tendencias y patrones a partir de un gran volumen de datos; es aquí donde la Visualización de la Información hace su aporte significativo, permitiendo que la mente del ser humano procese con mayor rapidez, reconozca las características esenciales de la información, y desencadene un acto de transferencia de conocimiento, a partir de lo que cada individuo pueda interiorizar con el apoyo de herramientas gráficas generadas por computadoras.

La comunidad científica especializada en visualización de la información ha desarrollado técnicas capaces de amplificar la cognición humana para permitir una visión abstracta en conjuntos de datos entendibles. Existe una gran diversidad de especificaciones de software, aplicaciones e infraestructura cibernética que pueden

estar al servicio de los investigadores para diseñar visualizaciones de información tales como: Análisis de Componentes Principal (ACP), Eigenvalue: Valor propio / Autovector (DVP) (Börner, Chen, & Boyack, 2003), Pathfinder Networks (PFnets) (Schvaneveldt, 1990), Triangulación (Börner, et al., 2003; Guzmán-Sánchez, 2009).

Otros algoritmos de mayor popularidad son: técnicas de *clustering* o análisis de conglomerados (Hartigan, 1975); técnicas de poda como el método de escalamiento de red Pathfinder o PFNETs (Schvaneveldt, 1990), y de posicionamiento de los nodos (Kamada & Kawai, 1989); técnicas estadísticas multivariantes tales como el Escalamiento Multidimensional o MDS (Multidimensional Scaling)(Kruskal & Wish, 1978); y técnicas de redes neuronales, como el modelo de mapas auto-organizativos o SOM (Self-Organizing Map) (T. Kohonen, 1982).

Específicamente, este estudio emplea como técnica de visualización *Self-organizing maps* o *mapas auto-organizados* (SOMs), basada en *Redes Neuronales Artificiales* (RNA); esta última realiza una clasificación de los datos, o sea, la formación de un cluster, a partir de las cercanías o similitudes entre las variables en análisis (Teuvo Kohonen, 1990).

El continuo progreso de la ciencia y la tecnología, muestra una tendencia hacia la integración de diferentes disciplinas científicas, es decir, emerge un proceso de profundización de la investigación, una disciplina se subdivide gradualmente en varias sub-disciplinas, y se desarrolla en una dirección más especializada y sofisticada (Xu, Guo, Yue, Ru, & Fang, 2016). Además, es necesario el conocimiento de varias disciplinas para resolver muchos problemas prácticos; esta necesidad promueve la combinación cruzada de las teorías, métodos y tecnologías entre las diferentes disciplinas, y por consiguiente se forma la interdisciplinariedad.

Hjørland y Albrechtsen en su clásico artículo, "*Towards a new horizon in information science: domain analysis*", mencionan la influencia en la Ciencia de la Información de campos como las ciencias cognitivas, lingüística, psicología, investigación educacional, ciencia de la computación, sociología y filosofía; y explican las relaciones o influencias que se pueden presentar de manera indirecta, demostrando así, tendencias en nuestra disciplina (Hjørland & Albrechtsen, 1995).

Erbacher (Erbacher, 2007) define el entramado interdisciplinar de la visualización de información como la integración de aportes procedentes de múltiples dominios, y por supuesto, la ciencia de la información juega un papel importante en este rubro como:

- Ciencia de la computación, quien aporta a la viabilidad y conveniencia de las técnicas y aplicaciones propuestas, así como del desarrollo e implementación de las técnicas de visualización
- Matemática y estadística, que contribuyen a la transformación visual y a probar y diseñar algoritmos computacionales
- Arte y arquitectura, cuyos expertos intervienen en muchos grupos de investigación para crear interesantes formas visuales
- Psicología cognitiva, que identifica las reglas intrínsecas para diseñar una visualización considerando la percepción humana
- Representación y análisis de información, espacios documentales y dominios de conocimiento (Börner, et al., 2003)
- Estudios métricos y evaluación de la ciencia
- Análisis de tendencias
- Web semántica
- Búsqueda y recuperación de información

La Ciencia de la Información, es un área científica de naturaleza interdisciplinar cambiante (Saracevic, 1995); la propia visualización de la información como campo científico tiene una naturaleza interdisciplinar, por lo que ambas comparten intereses coincidentes relacionados con el conocimiento y su comunicación, la organización de la información a partir del procesamiento de los datos, la diseminación y tratamiento de la información, la relación que se establece entre los usuarios y la información, así como, los métodos de investigación que los apoyen. En este sentido, Zhu y Chen, conceptualizaron la visualización de información como campo de estudio aplicable a cualquier dominio de las personas con la necesidad de extraer información interna dentro de un gran volumen de información (Zhu & Chen, 2005).

3.5. KDD, Minería de Datos, visualización de información y la Bibliometría

3.5.1. El algoritmo SOM y la visualización de información bibliométrica

Teuvo Kohonen, interesado en comprender la clasificación natural que hace el cerebro, ideó el algoritmo SOM, que sirve para la generación de mapas auto-organizados. Se le llama SOM al algoritmo matemático por sus siglas en inglés: *Self-Organizing Maps*. Este algoritmo conduce el proceso de entrenamiento de la red neuronal y produce un mapa topográfico bidimensional para la organización y visualización de información. La distribución topográfica de la información (en este trabajo, información bibliométrica) permite identificar cúmulos o clusters que guardan cierta similitud. En términos matemáticos, el SOM define una función del espacio de entrada (información de entrada) a una red de neuronas en el plano (mapa). En otros términos, ésta función define una proyección del conjunto de datos multidimensionales (invisible) a un espacio perceptible visualmente (normalmente bidimensional). La visualización del conjunto de datos permite que las relaciones de similitud que se presentan entre los datos dentro del espacio multidimensional puedan ser observadas en un despliegue bidimensional denominado "mapa auto-organizado" (T. Kohonen, Kaski, & Lappalainen, 1997).

La denominación mapas auto-organizados surge porque el proceso de entrenamiento de la red neuronal y la generación de los mapas no requieren de la intervención del especialista (aprendizaje no supervisado). Este modelo tiene una capa de entrada que recibirá la señal, la dimensión (n) dependerá del problema (e.g. cantidad de indicadores bibliométricos utilizados). El dato de entrada será un vector, $X = (x_1; x_2; \dots; x_n)$, mientras que la capa de salida (formada por **M** neuronas) es la encargada de procesar la información en paralelo y formar el mapa (figura 2). La idea básica del modelo, es crear una imagen de un espacio multidimensional de entrada en un espacio de salida de menor dimensión (reducir dimensiones) (Guzmán-Sánchez, 2009; T. Kohonen, et al., 1997)

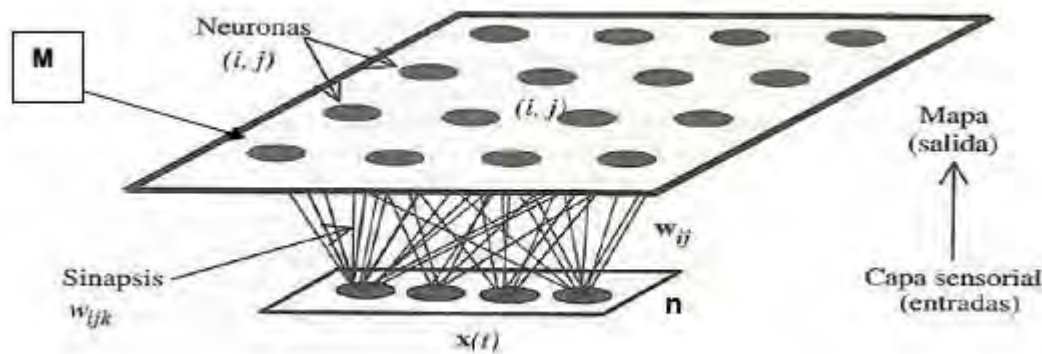


Figura 2. Estructura de una Red Neuronal bajo el algoritmo SOM (Guzmán-Sánchez, 2009)

Una característica importante en los mapas son las relaciones de semejanza (*similarity*) o disimilitud (*dissimilarity*) que se formen entre los datos. Las similitudes que se presentan entre los datos dentro del espacio multidimensional pueden ser observadas en un despliegue bidimensional denominado “mapa”. La semejanza o diferencia se establece cuando se consideran los datos métricos en forma de vectores, por tanto, es posible medir las distancias que se establecen entre ellos a partir de la aplicación de una métrica (fórmula matemática). Se asume que a menor distancia entre dos vectores, existe una mayor similitud entre los respectivos datos métricos (E. Villaseñor, 2004).

La aplicación de la red neuronal SOM para generar mapas auto-organizados en el campo de las disciplinas métricas de la información se inició en la década del siglo XXI. Los pioneros en emplear este tipo de técnica con fines bibliométricos e incluyendo el análisis de la matriz de palabras clave de co-ocurrencia en los dominios biomédicos fueron Sotolongo, Guzmán y Carrillo (Sotolongo-Aguilar, Guzmán-Sánchez, & Carrillo, 2002); en el mapeo de la ciencia y la tecnología (Moya-Anegón, Herrero-Solana, & Jiménez-Contreras, 2006; Polanco, Francois, & Lamirel, 2001), posteriormente, en la investigación de la evolución en el ámbito biomédico (Moya-Anegón, Herrero-Solana, & Jiménez-Contreras, 2006) (Guzmán-Sánchez, 2009); más recientemente para el estudio de patentes (Díaz-Pérez, et al., 2014; Pérez-Arreortúa, Díaz-Pérez, Giráldez-Reyes, & Carrillo-Calvet, 2014).

Diferentes autores han avizorado la importancia de la Minería de Datos y el KDD en los estudios métricos de la información. Siguiendo los métodos y técnicas de estas disciplinas, el análisis bibliométrico se sirve hoy de fórmulas matemáticas y estadísticas (algoritmos), que le permiten analizar grandes conjuntos de datos, para obtener nuevo conocimiento (Gorbea Portal, 2013; Guzmán Sánchez, 2009; Nieminen, Pölonen, & Sipola, 2013; Siguenza-Guzman, Saquicela, Avila-Ordóñez, Vandewalle, & Cattrysse, 2015).

Innovando en esta dirección, Sotolongo, Guzmán y Carrillo desarrollaron la metodología ViblioSOM (Visualización Bibliométrica con el Algoritmo SOM), para el análisis métrico y la visualización de la información bibliográfica digital (Sotolongo-Aguilar, et al., 2002). Esta metodología consta de una cadena secuencial de procesos para el descubrimiento y la visualización de conocimiento. Este método ocupa la red neuronal SOM para llevar a cabo la minería de datos y la visualización de información, siguiendo la secuencia natural del KDD. Esta metodología fue implementada en un sistema de software llamado ViblioSOM (Visualización Bibliométrica con el Algoritmo SOM), desarrollado por el Laboratorio de Dinámica no lineal en la Facultad de Ciencias de la Universidad Nacional Autónoma de México (UNAM) y la Empresa "Tecnologías Inteligentes y Modelación de Sistemas" (TIMS); con esta herramienta, el análisis y la interpretación de los datos a escala multidimensional están automatizados y facilita la visualización. Metodología que se explicará con detenimiento en el próximo capítulo.

Capítulo II: Metodología y fuentes de información

2.1. Fuentes de información: los índices de difusión de la ciencia

2.1.1. El Web of Science (WoS): las bases de datos y herramientas de análisis bibliométrico

Eugene Garfield, funda en 1957 en *Institute for Scientific Information (ISI)*, años más tarde en 1992 esta institución fue adquirida por *Thomson Scientific & Healthcare*, y actualmente se conoce como Thomson Reuters, tras la compra de Reuters por Thomson en 2008. *Thomson Reuters* ofrece un total de 32 productos y servicios, entre ellos el servicio en línea *Web of Science (WOS)* que se compone de las siguientes bases de datos:

- Science Citation Index Expanded (SCI-EXPANDED) –1900 presente
- Social Sciences Citation Index (SSCI) --1956-presente
- Arts & Humanities Citation Index (A&HCI) --1975-presente
- Conference Proceedings Citation Index- Science (CPCI-S) –1990 presente
- Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) --1990-presente
- Book Citation Index– Science (BKCI-S) --2005-presente
- Book Citation Index– Social Sciences & Humanities (BKCI-SSH) --2005-presente

Además, desde el año 2008 el WoS incluye una serie de índices de citas regionales como, por ejemplo: *Chinese Science Citation Database*, producido en colaboración con la Academia China de Ciencias; en el 2013 el *Citation Index SciELO*, que abarca Brasil, España, Portugal, el Caribe y Sudáfrica, y más de 12 países de América Latina; y en el 2014 se incluyó el Índice de Citación de Corea (KCI), con las actualizaciones de la Fundación Nacional de Investigación de Corea del Sur. La presente investigación se centra mayormente en la producción mexicana registrada en la colección principal del Web of Science (WOS), disponible a través del portal de bibliotecas de la Universidad Nacional Autónoma de México (UNAM).

Los motivos que nos han llevado a elegir esta base de datos, como fuente de obtención de información bibliográfica para nuestro estudio, obedecen a distintas argumentaciones que se detallan a continuación:

- Multidisciplinariedad: Todos los campos de la Ciencia están representados
- Selectividad: Las revistas son escogidas sobre la base de criterios cuantitativos (atendiendo a la cantidad de citas que reciben), y definitivamente seleccionadas según juicio de expertos
- Cobertura total de los contenidos de las más de 10, 000 revistas fuente
- Dirección completa de cada uno de los autores de los trabajos registrados en la base de datos
- Referencias bibliográficas: cada una de las referencias de los artículos se procesan, lo cual resulta una de sus principales prestaciones, pues facilita su utilización para el estudio de la base intelectual y construcción de indicadores

Se seleccionaron las bases de datos del Web of Science, además, por ser la fuente de información que publica artículos en revistas de corriente principal, parámetro que se considera como un elemento evaluativo establecido por el Consejo Nacional de Ciencia y Tecnología y el Sistema Nacional de Investigadores (SNI) para sus instituciones e investigadores.

La extracción de los datos para la realización del estudio se llevó a cabo en la versión en línea de las bases de datos del Web of Science, disponible a través del portal de la Biblioteca Electrónica de la Universidad Nacional Autónoma de México. La búsqueda se llevó a cabo en el mes de enero de 2016. Se prefirió esta fecha para la descarga de los artículos con vistas a evitar la pérdida de los registros correspondientes al año 2014, pues usualmente el 10 % de los documentos correspondientes a un año –principalmente los publicados durante la última etapa–, ingresan a las bases de datos del WoS durante los tres o cuatro primeros meses del año siguiente (Chinchilla-Rodríguez, 2004).

La búsqueda en WoS consistió en recuperar todos los artículos con al menos un autor mexicano; por lo que se trazó como estrategia la identificación de la palabra “Mexico” en el campo país, que ocupa la etiqueta CU de COUNTRY en la colección principal del Web of Science, a través de la búsqueda avanzada, al mismo tiempo,

se delimitó un espacio de tiempo del 2005 al año 2015 y refinó la búsqueda por tipo de documento (*article, letter, proceedings paper, reprint, book, review, book chapter, editorial material*). Se recuperaron un total de 136,620 documentos. En la tabla 3 aparece a continuación la tipología documental de la muestra y el número de registros recuperados.

Tabla 2. Número de documentos recuperados en la base de datos Web of Science por cada tipo documental

Tipo de documento WoS	Número de documentos WoS
ARTICLE	111,081
PROCEEDINGS PAPER	21,866
REVIEW	4,288
BOOK CHAPTER	4,244
EDITORIAL MATERIAL	2,724
LETTER	1,391
REPRINT	5
BOOK	4
Total	136,620

2.1.2. Los recursos de análisis del WoS: *Essential Science Indicators* y *Journal Citation Reports*

Se utilizó la interfaz web de la herramienta *Essential Science Indicators* (ESI), es un recurso que forma parte de la plataforma *InCites*, de *Thomson Reuters* para identificar los artículos altamente citados. Este producto contiene las cifras de las citas para los países, organizaciones, revistas y científicos. Calcula diferentes índices para más de 10,000 revistas en ciencia, tecnología, medicina y ciencias sociales. Incluye dos tipos documentales: artículos científicos y revisiones; se omiten las cartas al editor, erratas, artículos biográficos.

Essential Science Indicators, permite realizar complejos análisis de la literatura científica, como por ejemplo analizar el rendimiento de la investigación de las empresas, instituciones, naciones y revistas; identificar tendencias significativas

en las ciencias y las ciencias sociales. Ofrece Top Ranking por países, revistas, científicos, y las instituciones a través de 22 campos de investigación.

ESI es una herramienta integrada al WoS, en ella existen tres pestañas, cada una destinada a ofrecer diferentes tipos de información bibliométrica:

- “*Indicators*”, donde se pueden comparar países, instituciones, autores, revistas, temáticas, campos de investigación y frentes de investigación emergentes de acuerdo al número de Top, Highly cited y Hot papers.
- “*Field Baseline*”, incluye a su vez tres pestañas: (*Citation Rates*) donde se muestra el promedio de citas por artículo y año para cada área de investigación; (*Percentiles*) en ella se definen los niveles de citas para cada una de los 22 campos de investigación; es decir, la cantidad mínima de citas necesarias para que un artículo se encuentre el en 50%, 20%, 10%, 1%, 0.1%, y 0.01% de una misma disciplina y la pestaña “*Field Rankings*”, muestra un ranking de las 22 áreas de investigación en ESI en un período de 10 años y agrega además los conteos del número de highly cited papers.
- “*Citation Thresholds*” (Umbrales de citas): muestra el umbral de citas, es decir, el número mínimo de citas recibidas, cuando se obtiene un ranking de artículos para cada campo de investigación en orden descendente de citas, luego se selecciona la fracción superior de artículos.

El Umbral ESI revela el número de citas recibidas por el top 1% de los autores y de las instituciones; en el caso de los países y las revistas considera los que quedaron en el top 50% superior de una misma disciplina, en un período de 10 años.

Las revistas en ESI se clasifican en 22 amplios campos de la investigación, cada una es asignada a solo un campo y en el caso de revistas (como *Science*, *Nature* o Actas de la Academia Nacional de Ciencias [PNAS]) forman la categoría número 22, Multidisciplinarias. Los artículos de revistas multidisciplinares son reclasificados en campos específicos a través de un procedimiento automatizado.

Esta herramienta nos permitió realizar el análisis comparativo entre países, porque ofrece los datos del número de Highly cited y Hot papers por países. Además, de la pestaña “*Field Baseline*” se extrajo el dato (promedio de citas) para

poder calcular manualmente el impacto normalizado por cada país para un período comprendido entre 2005-2015 y para las 22 áreas de investigación.

Este estudio nos permitió calcular el *Impacto normalizado* con esta herramienta sin necesidad de obtener ese dato calculado automáticamente, pues el *Impacto normalizado* es un indicador que lo calcula la herramienta de análisis bibliométrico denominada *InCites*, de la cual todavía no se tiene acceso a través del WoS en el portal de bibliotecas de la UNAM.

Una limitante que ofrece esta herramienta (ESI), es el cómputo del impacto normalizado a nivel de institución por cada una de las 22 áreas de investigación. Este cálculo no se pudo realizar porque solo se puede determinar a través de un análisis de cada artículo en la pestaña *Documents* para cada área de investigación en la herramienta, es decir, calcular el promedio de citas por institución en cada una de las 22 área de excelencia solo es posible si se analiza artículo por artículo para cada temática; aspecto que constituyó una limitante para este estudio.

Otra de las fuentes de la información para la realización de esta tesis fue el *Journal Citation Reports (JCR)*, el cual forma parte de la plataforma *InCites*. *Journal Citation Reports* permite evaluar críticamente las revistas de corriente principal, a través de un grupo de indicadores, basados en datos de citas. JCR es una herramienta que ayuda a medir la influencia de la investigación y el impacto de las revistas por categoría temática.

Es reconocida como una herramienta importante para, evaluar el desempeño de las revistas, *Journal Citation Reports* proporciona el contexto para entender el grado de impacto e influencia de una revista dentro de la literatura académica a nivel mundial. Aspecto que se da gracias al cálculo del Factor de impacto, *Eigenfactor*, y los conteos de citas de manera general. Cubre un total de 10,800 revistas de más de 2.550 editores en aproximadamente 232 disciplinas, procedente de 83 países.

El JCR ofreció los datos necesarios de indicadores como el Factor de impacto y *Eigenfactor Score*, para las revistas mexicanas en el año 2015, y de esta manera comparar el desempeño de las mismas.

2.2. SCOPUS

Scopus es una base multidisciplinar de resúmenes y referencias bibliográficas de la literatura científica, incluye más de 18,000 títulos procedentes 5,000 editoriales internacionales. Scopus permite una visión multidisciplinaria de la ciencia e integra todas las fuentes relevantes para la investigación básica, aplicada e innovación tecnológica a través de patentes, fuentes de Internet de contenido científico, revistas científicas de acceso abierto, memorias de congresos y conferencias. Se actualiza diariamente y contiene los *Articles in Press* de más de 3,000 revistas.

Scopus es una herramienta que emplean los metristas de la información para realizar estudios bibliométricos y evaluaciones de producción científica, por su amplia cobertura temática, sino también porque ofrece acceso a herramientas que ofrecen importantes datos, tales como: perfil de autor, perfil de institución, rastreador de citas, índice h y analizador de revistas científicas.

La presente investigación incluyó también una pequeña comparación del volumen de producción en WoS con Scopus. En el caso de la base de datos Scopus, la búsqueda se efectuó en enero de 2016. La recuperación consistió en especificar por la etiqueta *AFFILCOUNTRY* la palabra Mexico, así como, limitar el período de tiempo entre 2005 y el año 2015; por último, limitar por el tipo de documento con la etiqueta de *DOCTYPE* y especificar los mismos tipos documentales que se delimitaron en el WoS (*article, letter, proceedings paper, reprint, book, review, book chapter, editorial material*). La tabla 2 ofrece el total de 172,707 documentos y el número de registros por tipo de documento.

Tabla 3. Número de documentos recuperados en la base de datos Scopus por cada tipo documental

Tipo de documento WoS	Número de documentos Scopus
ARTICLE	124,146
PROCEEDINGS PAPER	32,262
REVIEW	7,892
BOOK CHAPTER	4,237
EDITORIAL MATERIAL	1,885
LETTER	1,628
REPRINT	292
BOOK	365
Total	172,707

2.2.1. Elsevier Journal Metrics

El análisis de las revistas mexicanas requirió no solo de los indicadores propuestos por el *Journal Citation Reports* del WoS, además, se tuvo en consideración *Source Normalized Impact per Paper* (SNIP) y *SCImago Journal Rank* (SJR) de la herramienta *Journal Metrics* del Elsevier. *Journal Metrics* incluye más de 20,000 revistas, actas y series de libros. Es una herramienta que se actualiza una vez al año, calcula para el conjunto de revistas que incluye: *Source Normalized Impact per Paper* (SNIP), *SCImago Journal Rank* (SJR) y el *Impact per Publication* (IPP), indicadores que permiten comparar revistas de diferentes campos temáticos (Elsevier, 2016).

2.2.2. SCIMAGO Journal and Country Rank

Otro de los índices empleados en la investigación fue el *SCImago Journal & Country Rank*, portal que ofrece un Rank de revistas científicas y países a partir de la información contenida en la base de datos Scopus (Elsevier). Calcula un grupo de indicadores para evaluar y analizar las publicaciones científicas. La plataforma debe su nombre al trabajo desarrollado por el Grupo *SCImago* que desarrolló su métrica científica, el *SCImago Journal Rank*. SJR proporciona un **índice de calidad relativo** de las revistas incluidas en la base de datos Scopus a partir de 1996.

Cubre 26 áreas, 305 categorías, la búsqueda principal permite combinar la selección de áreas y categorías con la de países. El orden se puede establecer por SJR (impacto), o por otros parámetros como H-index, cantidad de citas, citas por documento, etc. La información de una revista nos proporciona todos los datos de publicación: el H-index de la revista, un breve resumen y el enlace a la publicación. Por último, nos permite comparar revistas ("SCImago Journal & Country Rank," 2016).

2.3. Batería de Indicadores bibliométricos empleados en la investigación

Los indicadores que se utilizan en la presente investigación pretenden caracterizar la producción científica mexicana y desempeño de las revistas mexicanas con mayor impacto. Constituyen una batería de 22 indicadores en total. Cuatro indicadores de producción para capturar la dimensión cuantitativa de la ciencia mexicana; seis indicadores bibliométricos destinados a caracterizar la influencia científica de la producción; ocho indicadores de excelencia científica y cuatro indicadores para evaluar el desempeño de las revistas mexicanas.

2.3.1 Indicadores de producción y eficiencia

Específicamente en la presente investigación se determinó incluir los siguientes indicadores de producción:

Ndoc: Total de artículos publicados por: años, áreas de investigación, países, revistas;

%Ndoc: Porcentaje del total de artículos publicados por: años, áreas de investigación, países, revistas; los cuales miden el volumen de la producción científica mexicana durante el período de estudio.

Para medir la eficiencia de la productividad mexicana, el presente estudio tomó en consideración dos indicadores de eficiencia científica (independientes de tamaño), con el objetivo de realizar comparaciones más justas.

- NSP: Índice de productividad nacional (*National Scientific Productivity*). Ndoc Wos/SNI.

Es el cociente del número de documentos indexados en WoS por autores mexicanos, entre el número de investigadores en el Sistema Nacional de Investigadores (SNI).

Es importante señalar que, el indicador es una estimación de la productividad nacional, porque hay artículos científicos producidos en México por investigadores que no pertenecen al Sistema Nacional de Investigadores. Sin embargo, el número de artículos cuyo autor no es miembro de SNI, constituyen un pequeño porcentaje de la producción total del país.

SDI: Índice de desarrollo científico (*Scientific Development Index*) de un país.
(Ndoc WoS/Número de habitantes) * 100,000.

La división del número documentos publicados en WoS entre el número de habitantes del país. El total de habitantes por cada país fue extraído del Banco de datos procedente del Banco Mundial (Banco Mundial, 2015)

El indicador SDI es también calculado por varias organizaciones, e.g. la Red de Indicadores de Ciencia y Tecnología Iberoamericana e Interamericana (RICYT); la UNESCO (UNESCO, 2015) y la OECD (OECD, 2016).

2.3.2. Indicadores de impacto

La presente investigación tomó en consideración algunos de los diferentes indicadores basados en las citas que ofrece el Web of Science para estudiar, analizar y comparar la producción científica mexicana con otros países. En este sentido, se listan los diferentes indicadores que permitieron analizar el impacto de la ciencia mexicana:

- Citdoc: Total de documentos citados es el total de artículos que recibieron al menos una cita durante el período analizado.
- %Citdoc: Proporción de documentos citados: % de artículos citados entre el total de documentos publicados.
- Citas/doc: Relación de citas por documento: número de citas en relación al total de documentos
- Cita/Ndocxtem: Relación eficiencia-eficacia: Relación entre el total de citas recibidas y el número de documentos por temáticas.

- RI: Impacto relativo, es el (valor promedio de citas que los artículos de un país, publicado en el año (y), e indexados en WoS, se han acumulado durante el intervalo de tiempo T) / (número medio de citas que artículos publicados por el mundo entero en el año (y) e indexadas en WoS, han acumulado durante T).
- NI: Impacto Normalizado, es el cociente del promedio de citas por país por cada una de las 22 áreas de investigación de ESI y el rendimiento medio del mundo. Existen diferencias en las tasas de citas para los diversos campos de investigación y períodos de tiempo, por tanto, el cálculo se realiza por separado.

Para medir el impacto de la revista mexicanas se utilizaron cuatro indicadores, dos del *Journal Citation Report* y dos del *Journal Metrics*:

IF: Factor de impacto para una determinada publicación en un año concreto resulta de dividir el número de veces que son citados los artículos publicados en dicha revista durante los dos años previos entre el número de artículos que publicó dicha revista durante ese mismo período de tiempo (Garfield, 2006).

La fórmula para calcular el Factor de Impacto es:

Factor de Impacto en el año $x = C_x / P_{x-1-2}$

Donde:

C_x = Citas en el año x a artículos publicados durante los dos años precedentes; y

P_{x-1-2} = Número de artículos publicados en los años $x-1$ y $x-2$

EFS: *Eigenfactor Score*, mide la repercusión global de las revistas a partir del cálculo iterativo del nivel de citación recibida por una revista, según las citas procedan a su vez de revistas más o menos citadas y, por consiguiente, más o menos influyentes (Jacso, 2010) y aplica algoritmos de relevancia tipo *PageRank* de Google.

SNIP: *Source normalized impact per paper* se define como la división entre dos indicadores (Moed, 2010), el Raw impact per paper (RIP) y el Journal's database citation potential (DCP) (Waltman, et al., 2013a).

SNIP= RIP/DCP; donde:

RIP de un artículo es igual al promedio de citas que recibe un artículo en los tres años precedentes al año de análisis.

$$DCP = \frac{1}{3} \times \frac{n}{\frac{1}{p_1 r_1} + \frac{1}{p_2 r_2} + \dots + \frac{1}{p_n r_n}}$$

Donde, n denota el número de artículos que citan determinada revista; r_i , es el número de referencias activas en un artículo del conjunto de publicaciones que citan determinada revista; y p_i es igual a la proporción entre h artículos en el grupo de artículos que citan una revista y el total de publicaciones (artículos, review, proceedings) que aparecieron en la misma revista, para el mismo año de la publicación seleccionada.

SJR: *SCImago Journal Rank*. El cálculo de este índice se realiza contabilizando el número de citas recibidas, ponderando la importancia o prestigio de las revistas de las que proceden dichas citas. El cálculo de las citas recibidas por las revistas se realiza en un período de tres años, otorgando un peso mayor a las citas procedentes de revistas de alto prestigio (aquellas con altas tasas de citación y baja autocitación) utilizando para ello el algoritmo de *Google PageRank*. Este algoritmo clasifica la importancia de los sitios web examinando la estructura de hipervínculo de la *World Wide Web*, en este caso de las revistas clasifica las citas (Page, Brin, Motwani, & Winograd, 1998).

La fórmula que sigue este algoritmo es la siguiente:

$$SJR_i = c * \frac{PSJR_i}{Art_i}$$

Donde:

PSJR_i – Prestigio de Scimago Journal Rank de la revista *i*.

Art_i – Número de documentos primarios (artículos, reviews, y conference papers) de la revista *i*.

Para obtener **PSJR_i** se aplica la siguiente fórmula:

$$PSJR_i = \frac{(1 - d - e)}{N} + e * \frac{Art_i}{\sum_{j=1}^N Art_j} + d * \left[\sum_{j=1}^N C_{ji} * \frac{PSJR_j}{C_j} * CF + \frac{Art_i}{\sum_{j=1}^N Art_j} * \sum_{k \in DN} PSJR_k \right]$$

Donde,

PSJR_i – Prestigio de Scimago Journal Rank de la revista *i*.

C_{ji} – Las referencias de la revista *j* a la revista *i*.

C_j – Número de referencias de la revista *j*.

d – Constante igual a 0.9.

e – Constante igual a 0.0999.

N – Total de revistas en la base de datos.

Art_j – Número de documentos primarios (artículos, reviews, y conference papers) de la revista *i*.

2.3.3. Indicadores del Essential indicators: Highly Cited Paper y Hot Paper

ESI es una herramienta que puede brindar respuesta a varias interrogantes cuando estamos realizando un estudio de corte métrico, por ejemplo: conocer, cuál es el artículo más citado en inmunología; qué país presenta mayor impacto en matemáticas; qué lugar ocupa determinado país de acuerdo al número de highly cited paper. Para ello ESI definió los siguientes indicadores:

TOP: Total de Top papers: es la suma de los *Highly Cited Papers* y *Hot Papers*. Se define como, el número de artículos en el 1% durante los últimos diez años, más el número de artículos en los últimos dos años que recibieron un alto número de citas en los dos últimos meses. Un documento que es a la vez *Highly Cited Paper* y *Hot Paper* se cuenta una sola vez ("Web of Science," 2016).

HCP: Total de Highly Cited Papers es el número de artículos que recibieron un total de citas superior al 1% de citas cuando se compara con todos los artículos

publicados en el mismo año y en la misma disciplina (L. Bornmann, 2014; Martínez, Herrera, Contreras, Ruiz, & Herrera-Viedma, 2015; Web of Science," 2016).

HOT: Total de Hot papers es el número de artículos en los dos últimos años que recibieron una cantidad de citas suficientes en los dos últimos meses para ubicarse en el 0,1% superior en comparación con otros artículos pares del mismo período y disciplina ("Web of Science," 2016).

Con este grupo de indicadores definidos por ESI, los cuales se consideran dependientes del tamaño, fue necesario definir otros indicadores independientes del tamaño, que permitieran realizar comparaciones justas entre México y otros países:

- HOTxWoS: Número de top papers por cada 1000 documentos en WoS. Número de hot papers entre el total de artículos publicados en WoS.
- TOPxwoS: Número de hot papers por cada 1000 documentos en WoS. Número de top papers entre el total de artículos publicados en WoS.
- HCPxWoS: Número de highly cited papers por cada 1000 documentos en WoS. Número de highly cited papers entre el total de artículos publicados en WoS.
- HCPxhab: Número de highly cited papers por cada 100,000 habitantes
- HOTxhab: Número de hot papers por cada 100,000 habitantes

2.4. Metodología implementada: ViBlioSOM

La metodología ViblioSOM (Visualización Bibliométrica con el Algoritmo SOM) empleada en esta investigación está integrada a un sistema de software llamado ViblioSOM. Este último usa para el análisis de información científica y reconocimiento de patrones desde dos perspectivas, la primera en el análisis multidimensional, es decir, encontrar el grado de relación entre variables de diferentes categorías o de la misma categoría pero cruzadas, y la segunda como herramienta para el análisis de información científica y reconocimiento de patrones, perspectivas que se aplican en este estudio (E. Villaseñor, Carrillo, de la Escalera, & Millán, 2008; Elio Villaseñor, Martínez de la Escalera, Carrillo Calvet, & Cruz, 2008).

ViblioSOM se usa para llevar a cabo el análisis del desempeño de las principales revistas científicas mexicanas en términos de indicadores de impacto y se usa también para caracterizar el perfil de desempeño bibliométrico de México y de un grupo de países que tienen un tamaño comparable al de México (volumen semejante de producción científica o un número semejante de habitantes). La Figura 3 ilustra las principales etapas de la metodología implementada en el software ViBlioSOM.

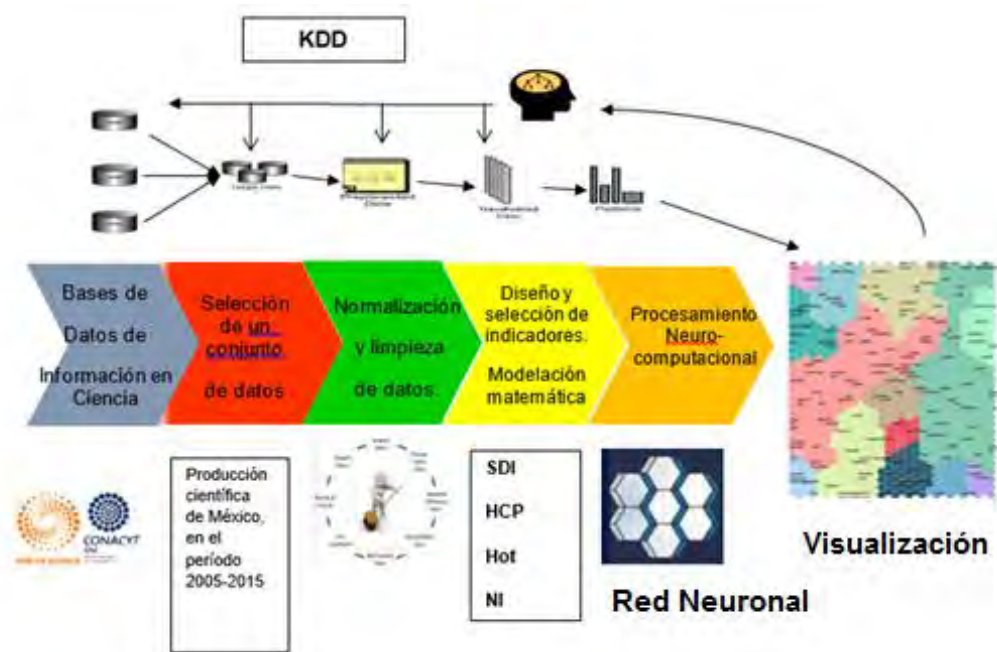


Figura 3. Metodología ViBlioSOM: etapas del descubrimiento, análisis y visualización de información bibliométrica

2.4.1. Etapas de la metodología ViBlioSOM

1. En un primer momento se determinaron las fuentes de información de donde se extrajeron los datos, en este caso: la colección principal del Web of Science, *Journal Citation Reports* (JCR), el *Essential Science Indicators* herramienta analítica del WoS; además la base de datos del Sistema Nacional de Investigadores (SNI), y el *Journal Metrics* de Elsevier.

2. Se conformó una tabla con los indicadores como columnas y las entidades (países) como renglones. En el ámbito del SOM, la tabla es una matriz y los

renglones son vectores. Con el propósito que los indicadores tengan el mismo peso durante la identificación de cúmulos que hace el SOM, es necesario eliminar rangos de variabilidad dispares entre los indicadores con un preprocesamiento, para que los indicadores que tienen valores grandes, no predominen durante el entrenamiento y, por tanto, no eclipsen a los más chicos. Este preprocesamiento consiste en dividir cada indicador (toda la columna) por el mayor de los indicadores en dicha columna. Después de esta operación, todos los indicadores tienen un rango de variación entre cero y uno (considerando que todos los indicadores son positivos o igual a cero).

3. Una vez que se tiene la matriz, se utiliza la herramienta de software para correr el algoritmo SOM. Los parámetros que se configuran son el tamaño del mapa (lado x ancho) y el número de iteraciones, el resto de los parámetros del algoritmo se conservan en su valor por defecto.

4. Al término del entrenamiento, se analizan dos tipos de visualizaciones: mapas de componentes y mapas de clustering (Villaseñor García, 2016). Cada componente es un indicador; de esta manera, se obtuvo un mapa de componente por cada indicador. La interpretación de estos mapas, permite analizar posibles correlaciones entre los indicadores. También son útiles para visualizar la distribución del indicador e identificar visualmente aquellas unidades que tienen un valor alto o bajo en el indicador. En ambas visualizaciones se mapean las etiquetas que identifican a las unidades de interés (e.g países). La siguiente figura 4 muestra un ejemplo de un mapa de componente. La escala de colores que se usa va de verde a rojo, pasando por el amarillo y el naranja. Las zonas rojas corresponden a valores altos del indicador y las zonas verdes a zonas con valores bajos en el indicador.

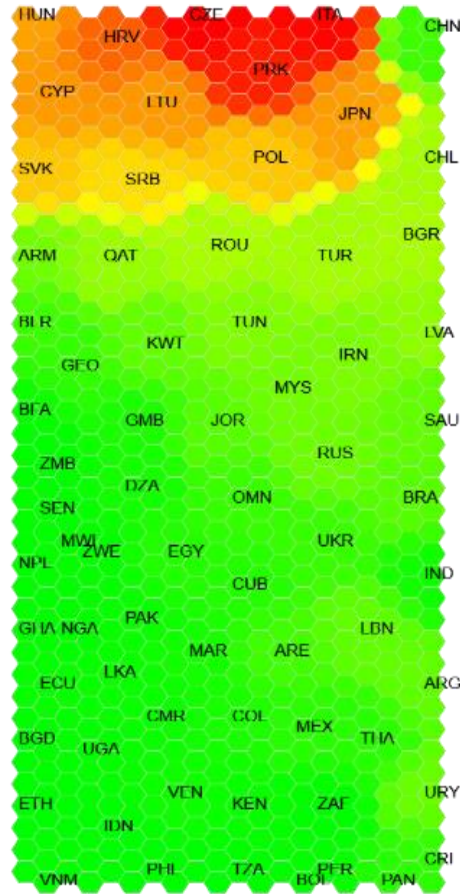


Figura 4. Mapa de componente del indicador: Índice de desarrollo científico (SDI) para países

La otra visualización que se utiliza en este trabajo es el mapa de *clustering*. Este mapa permite agrupar a las unidades de acuerdo a su parecido en los indicadores. Así, dos países, por ejemplo, están cercanos en el mapa porque tienen valores similares en los indicadores. Para facilitar la interpretación de los mapas, ponemos junto al mapa de *clustering* los mapas de indicadores (Figura 5).



Figura 5. Clustering de países considerando cuatro indicadores: Índice de desarrollo científico (SDI), Impacto Normalizado (NI), Highly Cited (HCP) y Hot Papers (Hot)

Capítulo III: Resultados y discusión

1. Análisis bibliométrico de las principales revistas mexicanas

1.1. Núcleo de revistas mexicanas con mayor visibilidad en Web of Science y Scopus

Este capítulo muestra un análisis extendiendo y actualizado del desempeño de las revistas mexicanas indexadas en *Web of Science* y *Scopus*, considerando indicadores de última generación basados en las citas, para el año 2015. Previamente se publicó un artículo sobre el desempeño de las revistas mexicanas en estas mismas fuentes de información para el año 2013 (Arencibia-Jorge, Atenógenes Villaseñor, Lozano-Díaz, & Carrillo Calvet, 2016).

Durante el período 2005-2015, el volumen de producción científica mexicana en el WoS fue de un total 136,620 documentos, de los cuales 111,081 artículos científicos fueron publicados en un total 9,327 títulos de revistas. A continuación, la tabla 4 muestra las 50 revistas que publicaron un número mayor a 216 artículos con al menos un autor de origen mexicano. En esta muestra donde hay revistas mexicanas y extranjeras, un total de 20 revistas son mexicanas y publican un total de 8,762 artículos, el 7% de la producción de artículos. En este sentido, es preciso identificar las revistas mexicanas de corriente principal con mayor repercusión internacional, tanto en WoS como en Scopus a partir de los indicadores bibliométricos que ofrecen estos recursos, para validar el impacto de las revistas científicas. En esta muestra de 50 revistas con mayor número de trabajos publicados por autores mexicanos demuestra su preferencia por publicar en revistas que se encuentran en el cuartil 4.

Tabla 4. Número de trabajos mexicanos publicados en fuentes indexadas en el WoS, durante el período 2005-2015

No.	Títulos de fuentes	Ndoc. WoS	%Ndoc	País	Quartil
1	Aip Conference Proceedings	1,373	1.24	USA	Q4
2	Revista Mexicana de Física	980	0.88	México	Q4
3	Proceedings of Spie	976	0.88	USA	Q2
4	Plos One	965	0.87	USA	Q1

5	Lecture Notes in Computer Science	910	0.82	USA	Q4
6	Revista Mexicana de Biodiversidad	881	0.79	México	Q4
7	Physical Review D	761	0.69	USA	Q1
8	Salud Publica de México	740	0.67	México	Q4
9	Gaceta Medica de México	692	0.62	México	Q4
10	Cirugía y Cirujanos	666	0.60	México	Q4
11	Astrophysical Journal	661	0.60	USA	Q1
12	Agrociencia	658	0.59	México	Q4
13	Monthly Notices of the Royal Astronomical Society	636	0.57	Reino Unido	Q1
14	Journal of Physics Conference Series	630	0.57	Reino Unido	Q4
15	Lecture Notes in Artificial Intelligence	594	0.53	Alemania	Q4
16	Revista de Investigación Clínica	582	0.52	México	Q4
17	Salud Mental	494	0.44	México	Q4
18	Proceedings of The Society of Photo Optical Instrumentation Engineers Spie	489	0.44	USA	Q4
19	Astronomy Astrophysics	480	0.43	Francia	Q1
20	Zootaxa	465	0.42	Nueva Zelanda	Q3
21	Physical Review Letters	444	0.40	USA	Q1
22	Interciencia	418	0.38	Venezuela	Q4
23	Revista Fitotecnia Mexicana	411	0.37	México	Q4
24	Revista de Biología Tropical	410	0.37	Costa Rica	Q4
25	Physics Letters B	405	0.36	Holanda	Q1
26	Revista Mexicana De Ingeniería Química	395	0.36	México	Q4
27	Physical Review E	388	0.35	USA	Q2
28	Journal Of Chemical Physics	367	0.33	USA	Q1
29	Acta Horticulturae	322	0.29	Bélgica	Q3
30	Revista Mexicana De Ciencias Geológicas	307	0.28	México	Q4
31	Physical Review B	303	0.27	USA	Q1
32	Archives Of Medical Research	299	0.27	USA	Q2
33	Studies In Computational Intelligence	292	0.26	Alemania	Q4
34	Revista Mexicana de Astronomía y Astrofísica Serie de Conferencias	287	0.26	México	

35	Journal of High Energy Physics	286	0.26	Italia	Q1
36	Optics Express	282	0.25	USA	Q1
37	Industrial Engineering Chemistry Research	276	0.25	USA	Q1
38	Hidrobiologica	276	0.25	México	Q4
39	International Journal of Electrochemical Science	269	0.24	Serbia	Q3
40	Electronics Robotics and Automotive Mechanics Conference	271	0.24	México	
41	Journal of The Mexican Chemical Society	268	0.24	México	Q4
42	Journal Of Applied Physics	263	0.24	USA	Q2
43	Southwestern Entomologist	252	0.23	USA	Q4
44	Physical Review A	248	0.22	USA	Q1
45	Catalysis Today	239	0.22	Holanda	Q1
46	Revista Mexicana de Astronomía y Astrofísica	230	0.21	México	Q2
47	Ciencias Marinas	229	0.21	México	Q4
48	Acta Botánica Mexicana	228	0.21	México	Q4
48	Journal of Applied Research and Technology	222	0.20	México	Q4
50	Tecnología y Ciencias del Agua	216	0.19	México	Q4

La evolución de la cobertura de las revistas mexicanas en Web of Science y Scopus durante el período 2005-2015 se muestra en la figur1, el cual muestra como la producción mexicana en estas dos importantes fuentes de conocimiento ha tenido un crecimiento importante durante los últimos once años. Se destaca una mayor inclusión de títulos de revistas mexicanas en Scopus. En este sentido, es importante conocer aquellas revistas que coinciden en ambas bases de datos y además analizar el desempeño de este grupo de revistas mexicanas de acuerdo a los indicadores calculados por WoS: Factor de impacto (IF), Eigenfactor Score (EFS); y Source Normalized Impact per Paper (SNIP) y SCImago Journal Rank (SJR) de Scopus.

En el año 2005, solo 13 revistas mexicanas estaban listadas por el Journal Citation Reports (JCR), sin embargo, 49 revistas mexicanas fueron incluidas en Scopus ese mismo año y para el año 2015 cuenta con 139. El número de revistas

mexicanas indexadas en JCR ha aumentado con el tiempo, pero se mantiene a partir del año 2010 hasta el 2015, lo que demuestra que las revistas mexicanas indexadas en esta base de datos mantienen los mejores valores del grupo de revistas mexicanas. La figura 6 denota como la cobertura en Scopus fue más inclusiva, a partir del año 2009, pues en este momento cuando la comunidad científica e incluso la OCDE comienza a utilizar las estadísticas basadas en los informes bibliométrico de Scopus. La cobertura de revistas en el JCR aumentó de 18 en el año 2008 a 33 en 2009, debido a que en ese período fue la inclusión de algunas revistas Latinoamérica. De manera general, la cobertura de las revistas mexicanas en estas bases de datos todavía es muy limitada.

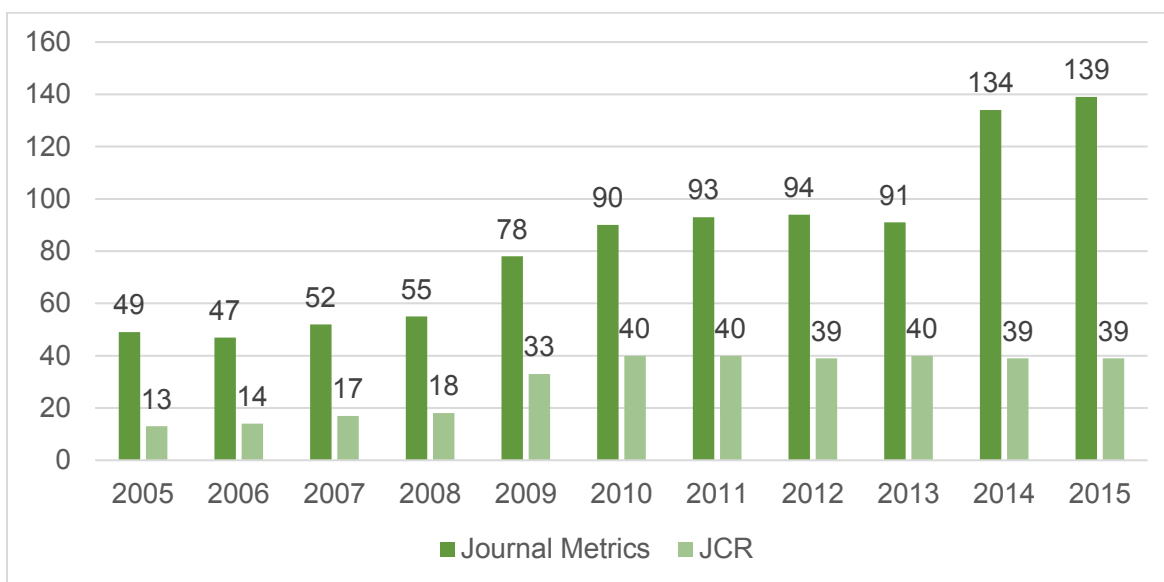


Figura 6. Revistas mexicanas indexadas en SCOPUS y Web of Science, 2005-2015

Se identificaron 39 revistas mexicanas que coinciden, en ambas bases de datos con los mejores valores según los indicadores y para este grupo de revistas: Factor de impacto (IF), Eigenfactor Score (EFS), Source Normalized Impact per Paper (SNIP), y SCImago Journal Rank (SJR).

La batería de indicadores seleccionados para caracterizar las 39 revistas mexicanas más visibles (Ver glosario de títulos de revistas), incluye dos indicadores del Web of Science y otros dos de la herramienta Journal Metrics de Scopus. La tabla 5, muestra el grado de correlación entre estos indicadores. Las parejas (IF,

SJR) y (EF, IF) exhiben los valores de correlación más altos, por su parte (SNIP, EF) y (SNIP, IF) muestran los inferiores. La independencia del indicador SNIP produce una diferencia notable, que se observa claramente.

Tabla 5. Correlación de los indicadores estudiados para una muestra de las 39 revistas mexicanas cubiertas por SCOPUS y WoS

	IF	SJR	SNIP	EF
IF	1	0.93	0.49	0.81
SJR		1	0.49	0.78
SNIP			1	0.56
EF				1

Existe alta correlación observada entre el factor de impacto y Scimago Journal Rank. Sin embargo, las diferencias de cobertura entre Scopus y Web of Science apoyan la posición del SJR y SNIP como alternativas del IF en entornos bibliométricos; y la normalización temática involucrada en el cálculo SNIP permite identificar este indicador como una alternativa para IF en los ejercicios de evaluación de la investigación.

1.2. El top ten de revista mexicanas con mayor visibilidad

La siguiente tabla muestra el top-ten de revistas mexicanas según los indicadores seleccionados: IF, EFS, SJR y el SNIP. La existencia de un grado de correlación entre los indicadores, implica que cuatro revistas mexicanas coincidan en los cuatro top-ten (Annals of Hepatology, Salud Pública de México, Revista Mexicana de Astronomía y Astrofísica y Revista Mexicana de Ciencias Geológicas), y una de ellas (Revista Mexicana de Astronomía y Astrofísica) es líder de acuerdo con: IF y SJR. Se destaca la revista Annals of Hepatology la cual, es líder en el ranking según el EFS; además ocupa el segundo lugar en el resto de los indicadores calculados.

Importante señalar algunos aspectos que se destacan con el indicador SNIP:

1. Hay fuentes como la revista *Acta Botánica Mexicana* y *Annals of Hepatology*; son las más visibles en Scopus de acuerdo al SNIP
2. Dos revistas de Ciencias Sociales están incluidas en el SNIP entre las diez primeras (*Gestión y Política Pública* e *Investigación Económica*), que no se observan en el núcleo de revistas más visibles según IF, EFS y SJR. (Tabla 6)

Tabla 6. Top-ten 2015 de revistas mexicanas en correspondencia con Impact Factor, Source Normalized Impact per Paper, Eigenfactor Score and SCImago Journal Rank

Título de Revista	2015 SNIP	Título de Revista	2015 EFS
Acta Botánica Mexicana	1.057	Annals of Hepatology	0.00362
Annals of Hepatology	0.874	Salud Publica de México	0.00245
Salud Publica de México	0.841	Revista Mexicana de Astronomía y Astrofísica	0.00213
Investigación Económica	0.686	Revista Mexicana de Biodiversidad	0.00179
Revista Mexicana de Astronomía y Astrofísica	0.637	Revista Mexicana de Ciencias Geológicas	0.00085
Gestión y Política Pública	0.607	Revista Mexicana de Física	0.00073
Revista Mexicana de Biodiversidad	0.598	Journal of the Mexican Chemical Society	0.00066
Revista Mexicana de Ciencias Geológicas	0.573	Ciencias Marinas	0.00062
Revista Mexicana De Ciencias Pecuarias	0.569	Revista Mexicana De Ciencias Pecuarias	0.00058
Revista Latinoamericana de Investigación en Matemática Educativa	0.563	Agrociencia	0.00057
Título de Revista	2015 SJR	Título de Revista	2015 IF
Revista Mexicana de Astronomía y Astrofísica	1.292	Revista Mexicana de Astronomía y Astrofísica	2.364
Annals of Hepatology	0.701	Annals of Hepatology	1.811
Salud Publica de México	0.608	Salud Publica de México	1.107
Atmosfera	0.377	Revista Mexicana de Ingeniera Química	0.924
Ciencias Marinas	0.348	Ciencias Marinas	0.75
Revista Mexicana de Ciencias Geológicas	0.318	Revista Mexicana de Ciencias Geológicas	0.736
Revista Mexicana de Biodiversidad	0.316	Journal of the Mexican Chemical Society	0.722
Geofísica International	0.3	Boletín de la Sociedad Geológica Mexicana	0.679
Revista Mexicana de Ingeniera Química	0.273	Botanical Sciences	0.624
Botanical Sciences	0.265	Revista Mexicana De Ciencias Pecuarias	0.577

Al comparar los valores de los indicadores que ofrece la herramienta Journal Metrics de Elsevier, se observa: la revista Acta Botánica Mexicana y Annals of Hepatology, presenta altos valores de Source Normalized Impact per Paper. Por su parte, la Revista Mexicana de Astronomía y Astrofísica, presenta un alto valor Scimago Journal Rank, las revistas mexicanas con mayor impacto / influencia pertenecen al campo de la física, ciencias médicas y ciencias sociales (Figura 7).

La Revista Mexicana de Astronomía y Astrofísica, durante 20 años su factor de impacto ha estado fluctuando entre 1y 3, para ubicarla en el Q2 según el JCR. Esta revista desarrolló una de las mejores políticas editoriales en América Latina, para incrementar el factor de impacto sin recurrir a un editor comercial. Las revistas de América Latina y el Caribe que alcanzan IF = 1, que son contadas, se pueden considerar con IF altos, teniendo en consideración el umbral (frontera) de las revistas regionales, han permanecido históricamente por debajo de 1.

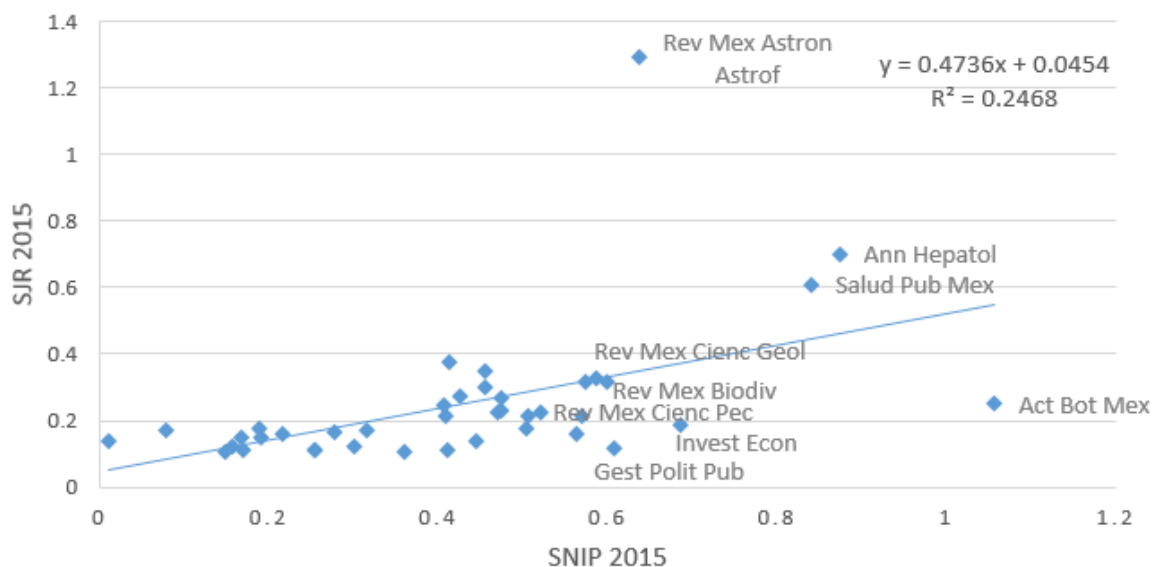


Figura 7. Revistas mexicanas con mayores valores en el Source Normalized Impact per Paper y SCImago Journal Rank para el año 2015

Cuando se contrasta, el Impact Factor y Eigenfactor Score, indicadores que ofrece el Journal Citation Reports del WoS, las revistas que alcanzan mejores valores al comparar este grupo de revistas mexicanas son: Annals of Hepatology,

Salud Pública de México y Revista Mexicana de Astronomía (Figura 8), las revistas mexicanas con mayor influencia pertenecen a la física y las ciencias médicas, áreas de investigación con mayor cubrimiento en el WoS. En el caso de Revista Mexicana de Astronomía y Astrofísica fue noticia en 1984-85 cuando alcanzó un factor de impacto igual a 3; la primera en lograrlo en América Latina y el Caribe.

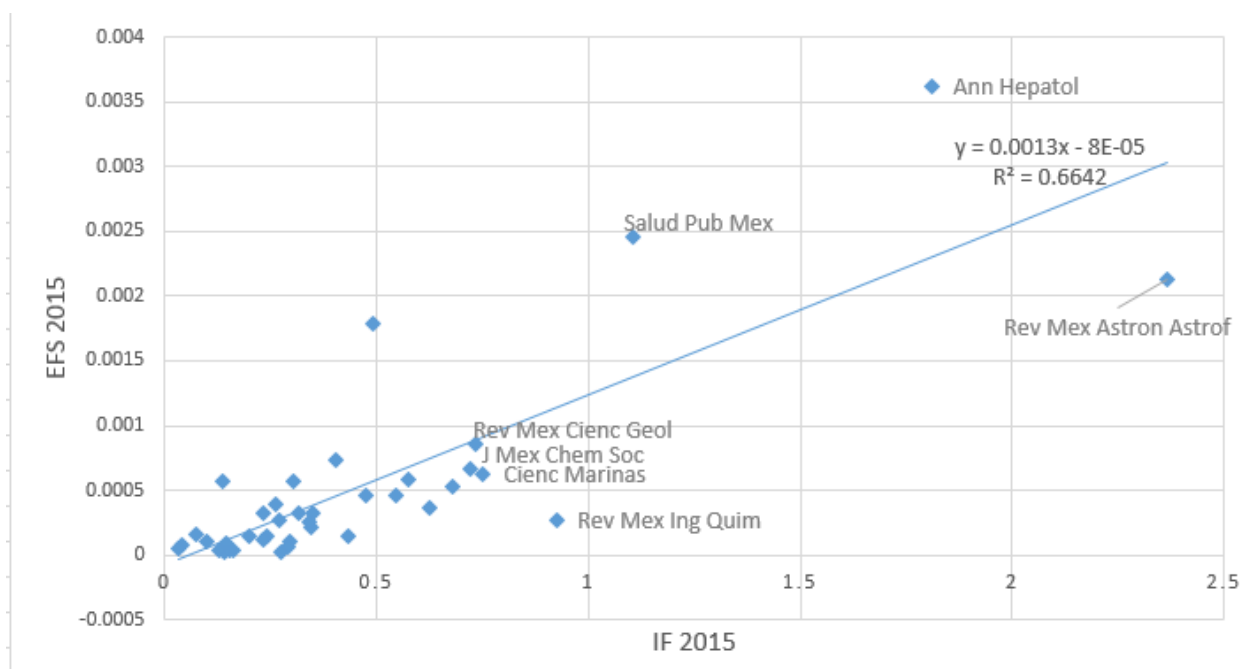


Figura 8. Revistas mexicanas con mayores valores en el Eigenfactor Score y Impact Factor para el año 2015

1.3. Análisis multidimensional de las revistas mexicanas, basada en la técnica de mapas auto-organizados o SOM (Self-Organizing Map)

El análisis multifactorial del conjunto de 39 revistas mexicanas se realizó con el método de inteligencia artificial, basado en redes neuronales artificiales de Mapas auto-organizados. Con esta técnica cada revista toma un punto en el espacio de cuatro dimensiones euclidiana, cuyas coordenadas vienen dadas por los valores de los cuatro indicadores: FI, EFS, SNIP y SJR. Nuestra técnica de visualización incluye cuatro mapas de componentes para exponer la comparación y clasificación realizada por la Red neuronal. Estos mapas se dibujan sobre una cuadrícula

hexagonal, en el que cada hexágono representa una neurona. En este caso se utilizó una rejilla de 20x20 neuronas y se realizaron 500 iteraciones.

Las ventajas de los indicadores propuestos por el Web of Science y Scopus para estimar el impacto e influencia de las revistas, se observaron a través de mapas bibliométricos basados en la técnica de auto-organización de redes neuronales artificiales (Figura 9) (Anexos 8 y 9). Las zonas de mayor impacto presentado por los colores rojo y amarillo, y la zona de bajo impacto es de color verde. En el mapa de componentes del indicador *Source Normalized Impact per Paper* (SNIP) exhibe un grupo mayor de revistas con mejores medidas de impacto.

En los cuatro mapas de componentes, las revistas Salud Pública, Revista Mexicana de Astronomía y Astrofísica, y Annals of Hepatology se ubican en la zona de color rojo, porque alcanzan los mejores valores de SNIP, IF, EFS y SJR (Figura 9).

En el mapa de componente inciso a) SNIP 2015 de la figura 9, demuestra como un grupo de 15 revistas mexicanas alcanzan mejores valores de impacto con el cálculo del SNIP. Las revistas: Acta Botánica Mexicana, Revista Mexicana de Biodiversidad, Revista Mexicana de Ciencias Geológicas, Investigación Económica, Revista Mexicana de Física, Agrociencia, Investigación Bibliotecológica, Madera y Bosques, Revista Mexicana de Ciencias Pecuarias entre otras, se ubican en la zona de color rojo y amarillo de acuerdo a los valores del SNIP. Sin embargo, para el resto de los indicadores, este conjunto de revistas aparece en la zona de baja intensidad (Color verde).

Este grupo de revistas mexicanas tienen mejores calificaciones con el SNIP debido a que, esta nueva métrica para evaluar el impacto de la revista, toma en cuenta las características del dominio de conocimiento sino también las diferentes prácticas de citar entre campos temáticos. Para ello, utiliza un enfoque normalizado de la fuente "*source normalized approach*", con este enfoque las citas a la revista deben ser normalizadas basadas en la lista de referencias de la fuente citante; y el campo de una revista se determina por los artículos que citan a la revista.

La ventana de citas en el SNIP es igual a tres años, un año mayor que en el caso del IF del Journal Citation Reports (JCR), es decir, ofrece un tiempo promedio para que el impacto de la revista madure. Particularmente, favorece aquellas disciplinas donde el impacto de la cita es lento como, por ejemplo: las matemáticas, las ciencias sociales y humanidades; y parte de las ingenierías. Fenómeno que se presenta por un retraso en la publicación, es decir, el período de tiempo entre la fecha de envío de un documento a una revista y su fecha de publicación oficial (Moed, 2010; H. F. Moed, 2011; Schoepfel & Prost, 2009; Waltman, et al., 2013b). Por ello, las revistas mexicanas, ofrecen mejores índices de influencia con el SNIP, que con el factor de impacto.

Además, queda demostrado con la aplicación de la metodología, como los mapas visualizan, la correlación numérica que se calculó para conocer el grado de correlación que existe entre los cuatro indicadores calculados (Tabla 5).

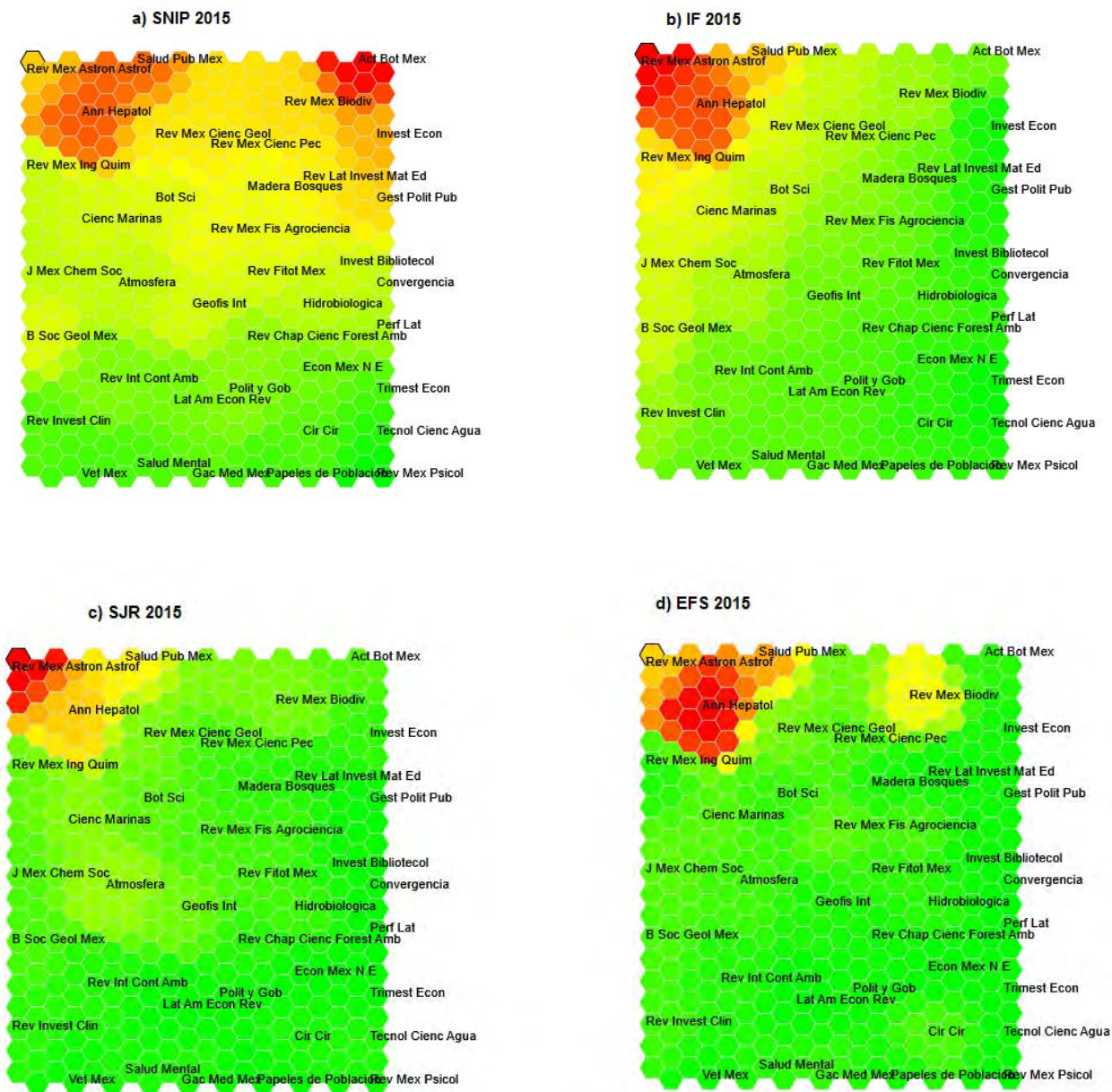


Figura 9. Mapa de Visibilidad por zonas de las revistas mexicanas de acuerdo con: a) Source Normalized Impact per Paper, b) Impact Factor, c) Eigenfactor Score, y d) SCImago Journal Rank. (Self-organized maps desarrollado por ViblioSOM). (Anexos 8 y 9: los mapas aparecen separado)

2. Análisis de la producción científica mexicana en el WoS, 2005-2015

2.1. El Sistema Nacional de Investigadores (SNI)

Otra de las fuentes de información utilizadas en la presente investigación son los datos que ofrece el Sistema Nacional de Investigadores; instancia que fue creada por Acuerdo Presidencial el 26 de julio de 1984. El objetivo de esta entidad es reconocer la labor de las personas dedicadas a producir conocimiento científico y tecnología en México. El Sistema Nacional de Investigación (SNI) constituye el principal recurso del país para generar nuevo conocimiento científico. Por tanto, se espera que el volumen de producción de la nación evolucione en proporción directa al número de investigadores SNI. Específicamente en este capítulo esta base de datos fue de gran importancia.

2.2. Análisis comparativo de la evolución del SNI y de la producción científica mexicana en Scopus y el Web of Science

Desde el año 2005 hasta el 2015 la producción científica de México ha aumentado tanto en Web of Science (WoS) y Scopus, con tasas de crecimiento promedio anual de 6,85% y 7,11%, respectivamente. En consecuencia, el indicador de producción para el 2015 en la base de datos Scopus aumentó con respecto al año 2005, y en WoS, aunque más pequeño, tiene un comportamiento similar. Durante el período 2011-2015 en la base de datos Scopus el país produjo 92,989 registros, lo que representa el 53,8 % del total de documentos (172,707) en el período de once años 2005-2015. La producción en la colección principal de Web of Science durante el mismo período fue de 136,620 registros, de los cuales 73,142 (53, 5%) de ellos se produjeron durante el 2011-2015 (Tabla 7).

Tabla 7. Evolución de la producción científica mexicana en el Web of Science y Scopus, durante 2005-2015.

Años	Ndoc WoS	Ndoc Scopus	SNI	NSP WoS	NSP Scopus
Año 2005	8,624	10,824	10,859	0.79	0.99
Año 2006	9,130	12,109	12,426	0.73	0.97
Año 2007	10,409	12,478	13,437	0.77	0.92
Año 2008	11,468	13,881	14,559	0.78	0.95
Año 2009	11,753	14,756	15,481	0.75	0.95
Año 2010	12,094	15,670	16,519	0.73	0.94
Año 2011	12,826	16,838	17,568	0.73	0.95
Año 2012	13,845	17,916	18,476	0.74	0.97
Año 2013	14,998	18,908	19,655	0.76	0.96
Año 2014	15,579	20,027	22,408	0.69	0.89
Año 2015	15,894	19,300	23,032	0.69	0.83
	136,620	172,707			

Leyenda: Número anual de documentos en Wos (ANdoc WoS), Número Anual de documentos en Scopus (ANdoc Scopus); el número de miembros en el Sistema Nacional de Investigadores (SNI) y el índice de productividad nacional (National Scientific Productivity, NSP Scopus, NSP WoS)

El número de documentos producidos anualmente en Scopus es cercano al número de investigadores en el país; y el volumen de producción en WoS, está un poco más alejada al número de investigadores en el SNI. La cantidad de miembros inscritos en el Sistema Nacional de Investigadores (SNI) está por encima de las curvas de Scopus y WoS, además existe un aumento en el incremento del número de SNI específicamente del año 2013-2014, el crecimiento en el número de científicos mexicanos se debió al programa de desarrollo de recursos humanos sostenido por el Consejo Nacional de Ciencia y Tecnología (CONACYT) de México (Tabla 7 y Figura 10).

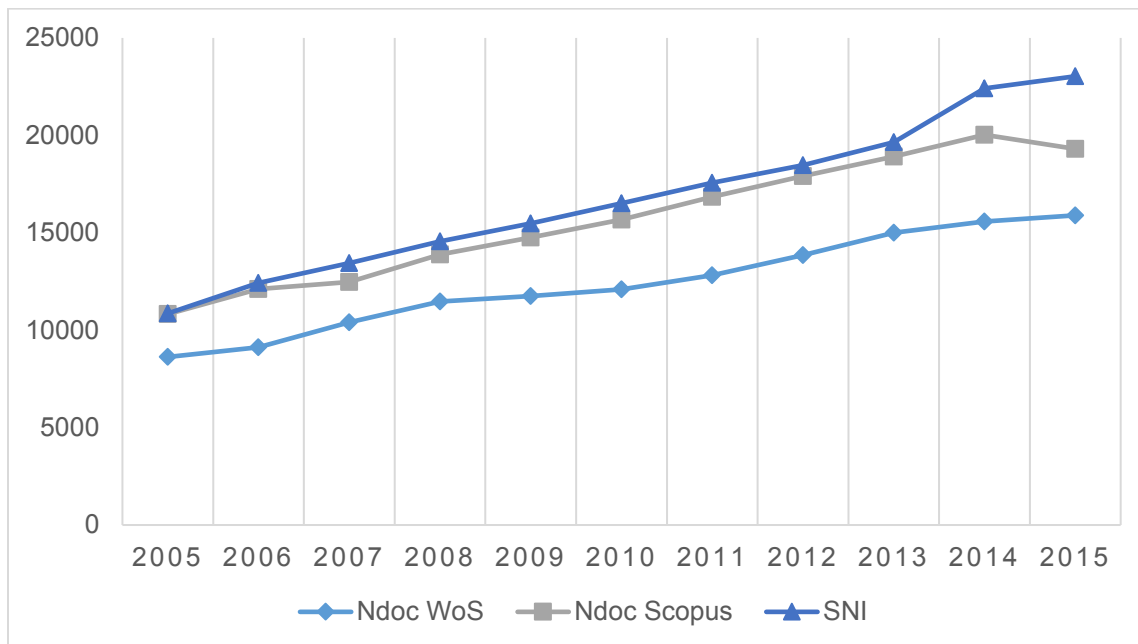


Figura 10. Crecimiento paralelo del número total de miembros mexicanos en el SNI y la producción científica en Web of Science y Scopus, durante el período 2005-2015

El crecimiento de la producción científica en ambas bases de datos se correlaciona con el crecimiento del total de investigadores del Sistema Nacional de Investigación que aumentó durante el período (Figura 10 y 11). La tasa media de crecimiento anual del número de investigadores adscritos al SNI es igual a 8% durante el mismo período, pasando aproximadamente de un investigador por cada 10,000 habitantes en 2005 a 2 en el 2015.

En promedio, México produce aproximadamente un documento en cada una de estas dos bases de datos; ya sea en Scopus o en el Web of Science (WoS) según lo muestran los valores de la tasa anual de productividad científica nacional (*National Scientific Productivity*) (Tabla 7 y Figura 11). La figura 11 denota una caída para los años 2013 – 2014 aunque, existe un aumento del número de miembros en el SNI para este período, este comportamiento está dado por un decrecimiento de la producción científica.

Además, para el análisis de producción científica mexicana es necesario tener en consideración que, los tipos documentales incluidos en este estudio (*article, letter, proceedings paper, reprint, book, review, book chapter* y *editorial material*) no son el único producto válido de investigadores nacionales, sino, patentes, modelos

de utilidad, reportes de proyectos, los cuales no se incluyen en este estudio, así como tampoco se tuvo en cuenta las bases de datos nacionales.

Importante señalar que, el año 2015 no incluye toda la producción científica, porque alrededor de un 10% de los trabajos de cada año, son incluidos en el Web of Science (WoS) al año siguiente en los primeros meses, debido a los retrasos en la inclusión de trabajos en las bases de datos Web of Science (Chinchilla-Rodriguez, 2004).

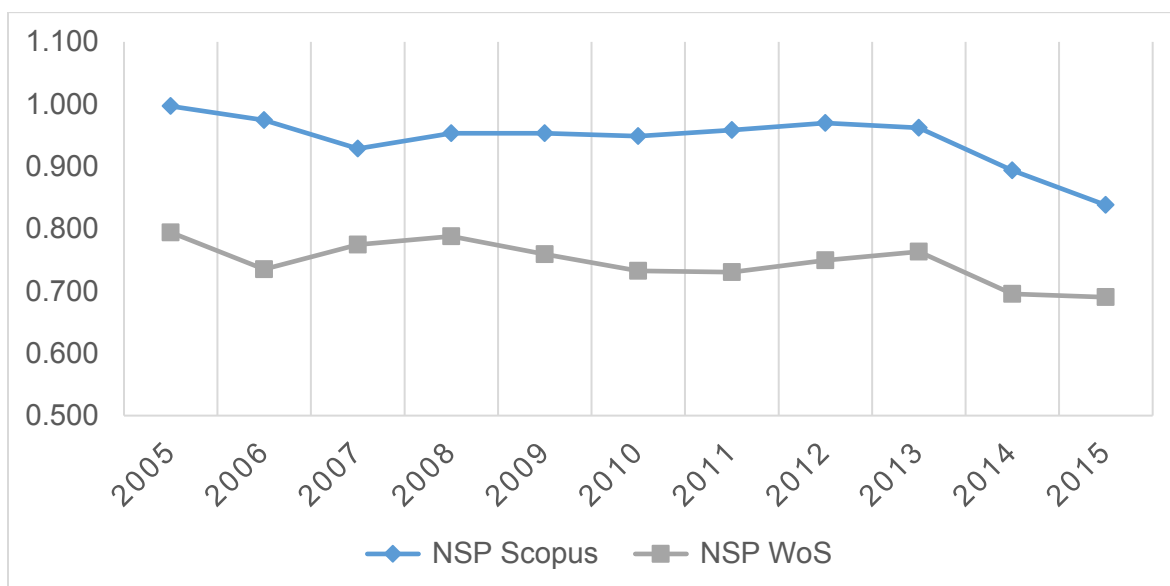


Figura 11. La eficiencia de la ciencia mexicana en Web of Science y SCOPUS, durante 2005-2015

2.2.1. Análisis de los artículos científicos mexicanos visibles internacionalmente

En un análisis específico con los artículos científicos publicados en Web of Science y Scopus, muestra como la producción de artículos científicos y el número total de miembros del SNI, han aumentado constantemente durante el período 2005-2015, en WoS con un total de 111,081 y Scopus con 124,146 artículos (Tabla 8 y figura 12).

Tabla 8. Evolución del número de artículos mexicanos en el Web of Science y Scopus, durante 2005-2015.

Años	ANp WoS	ANp Scopus	SNI	NSP WoS	NSP Scopus
Año 2005	7,128	7,166	10,859	0.66	0.66
Año 2006	7,183	7,986	12,426	0.58	0.64
Año 2007	8,259	8,664	13,437	0.61	0.64
Año 2008	8,979	9,893	14,559	0.62	0.68
Año 2009	9,301	10,749	15,481	0.60	0.69
Año 2010	9,881	10,803	16,519	0.60	0.65
Año 2011	10,604	11,813	17,568	0.60	0.67
Año 2012	11,497	12,974	18,476	0.62	0.70
Año 2013	12,215	14,061	19,655	0.62	0.72
Año 2014	12,711	15,016	22,408	0.57	0.67
Año 2015	13,323	15,021	23,032	0.58	0.65
	111,081	124,146			

Leyenda: número anual de artículos en WoS (ANp WoS), número anual de artículos en Scopus (ANp Scopus); el número de miembros en el Sistema Nacional de Investigadores (SNI) y el índice de productividad nacional (National Scientific Productivity, NSP Scopus, NSP WoS)

En la figura 12 se observan las diferentes tendencias en las tasas de crecimiento, es decir, han impactado de manera decreciente en los indicadores nacionales de productividad científica (NSP Scopus y NSP WoS) en la última parte de este período. La disminución de los valores en estos indicadores podría interpretarse como una pérdida de eficiencia en la producción científica.

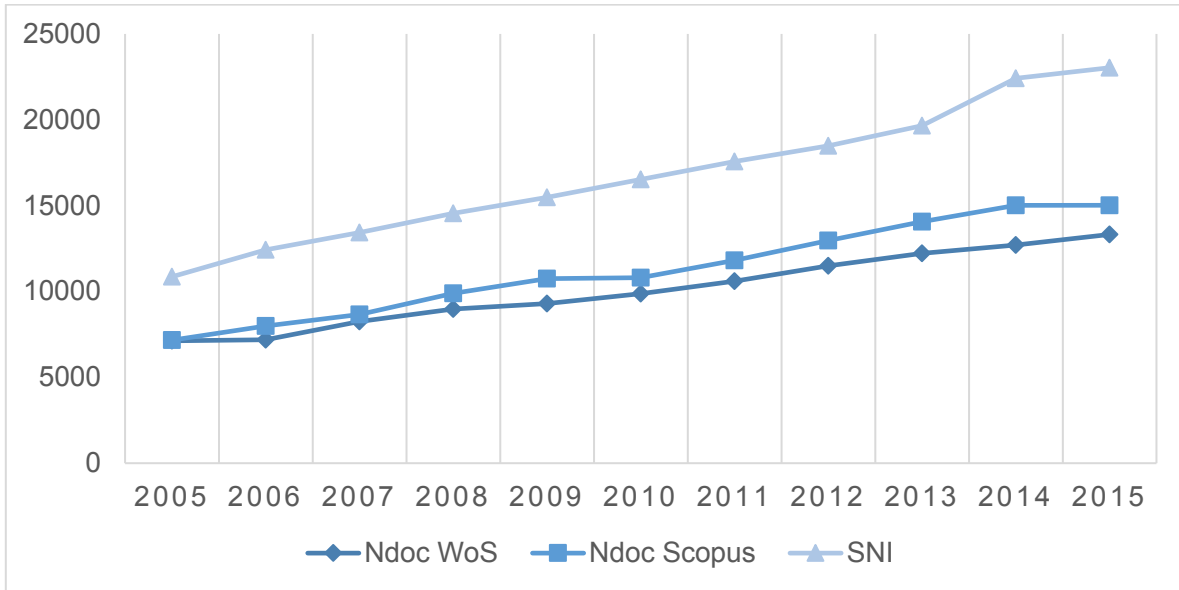


Figura 12. Crecimiento paralelo del número total de miembros mexicanos en el SNI y el número de artículos científicos en Web of Science y Scopus, durante el período 2005-2015

La figura 13 muestra como el indicador de productividad científica nacional en WoS ha oscilado alrededor de un valor medio de 0.61 durante el 2005-2013, pero a partir del año 2013, disminuyó su valor a 0.57. El índice de NSP en Scopus también fluctúa su valor, pasando de 0.66 en el año 2005 a 0.72 en 2013, y existe una disminución a partir de este año hasta el año 2015.

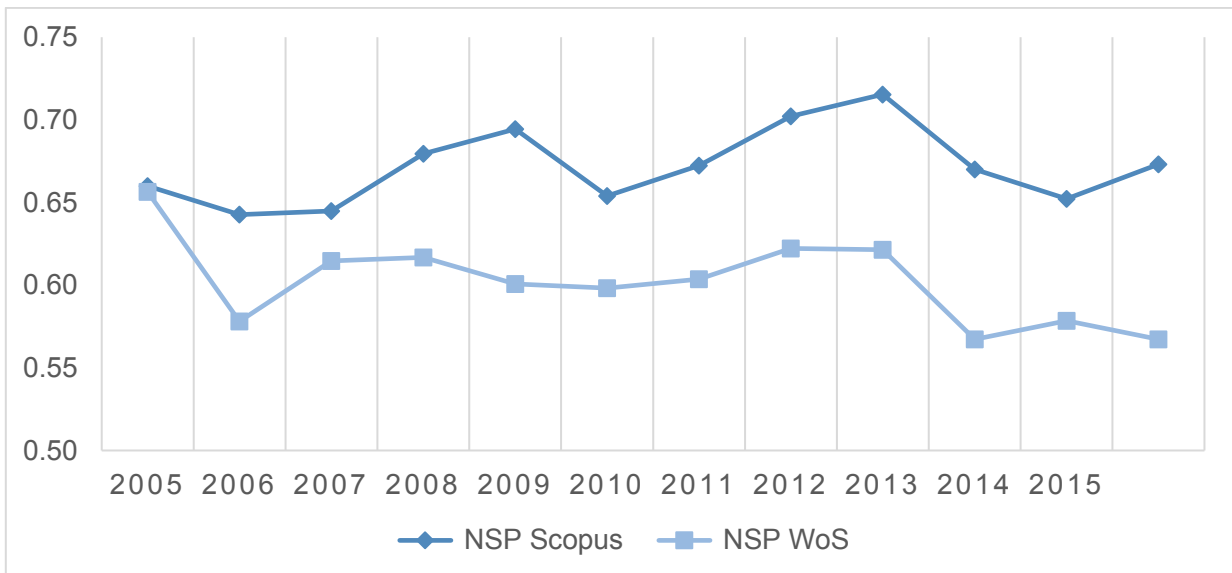


Figura 13. El índice de producción científica nacional (National Scientific Productivity) en Web of Science y Scopus, durante el período 2005-2015

2.3. Estudio comparativo de la producción mexicana con otros países.

En el análisis comparativo de la producción científica mexicana se seleccionaron países con similar producción científica de acuerdo con el ranking del *Essential Science Indicator* para el período 2005-2015. En la comparación con países de la región de producción científica similar, el número de documentos en el WoS de México se encuentra en segundo lugar superado por Brasil, como muestra la figura 14. La producción de conocimiento en América Latina está encabezada por países como Brasil, México, Argentina y Chile; Brasil es el país de la región que más recursos invierte en ciencia y tecnología por lo que su producción científica alcanza valores que superan a México (UNESCO, 2015).

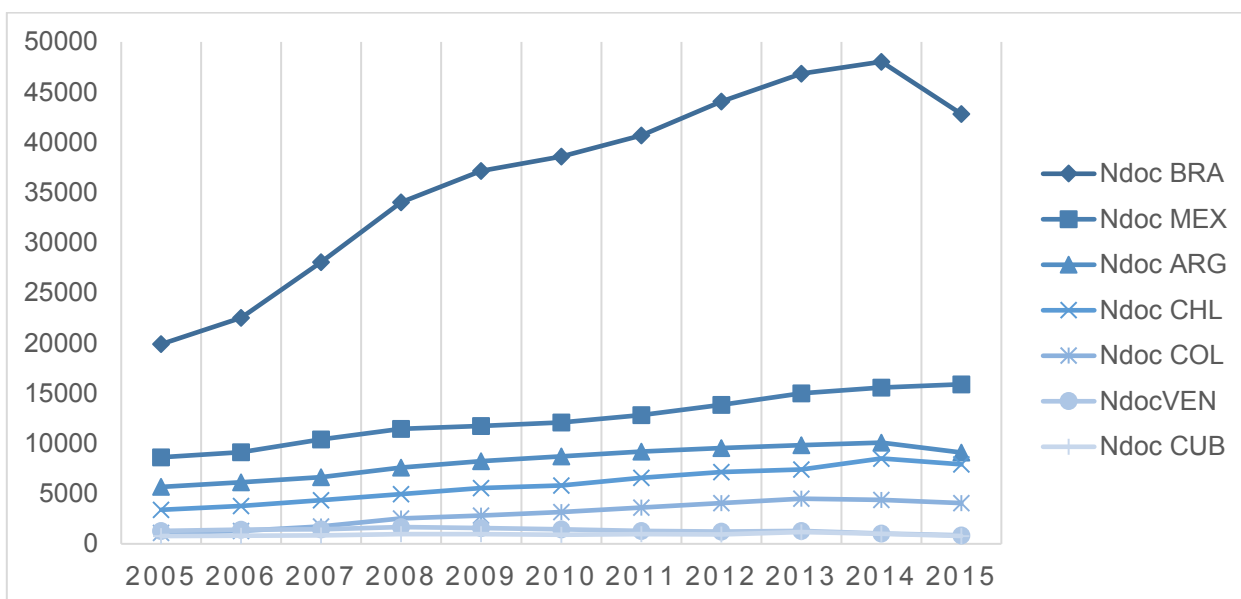


Figura 14. Comparación de la producción mexicana en WoS con la de los países más productivos de la región

El comportamiento de la producción científica de México, Brasil, Argentina, Colombia, Venezuela y Cuba en relación con el número de habitantes en el período 2005-2015, denota como México ocupa el cuarto lugar (figura 15). La disminución notable en el último año de la muestra seleccionada, se debe a que la descarga de los registros de la base de datos finalizó en enero de 2016, con lo cual es probable que un importante número de artículos no se pudieran encontrar en el momento de

la descarga de los registros de la base de datos fuente. Es reconocido en la literatura científica, que alrededor de un 10% de los trabajos de cada año, son incluidos en el Web of Science (WoS) al año siguiente en los primeros meses, debido a los retrasos en la inclusión de trabajos en las bases de datos Web of Science. Por lo que autores de reconocida experiencia en los estudios realizados a partir de información extraída del WoS, no incluyen el año correspondiente a la carga de la base de datos, sino el año de publicación en la revista (Chinchilla-Rodriguez, 2004).

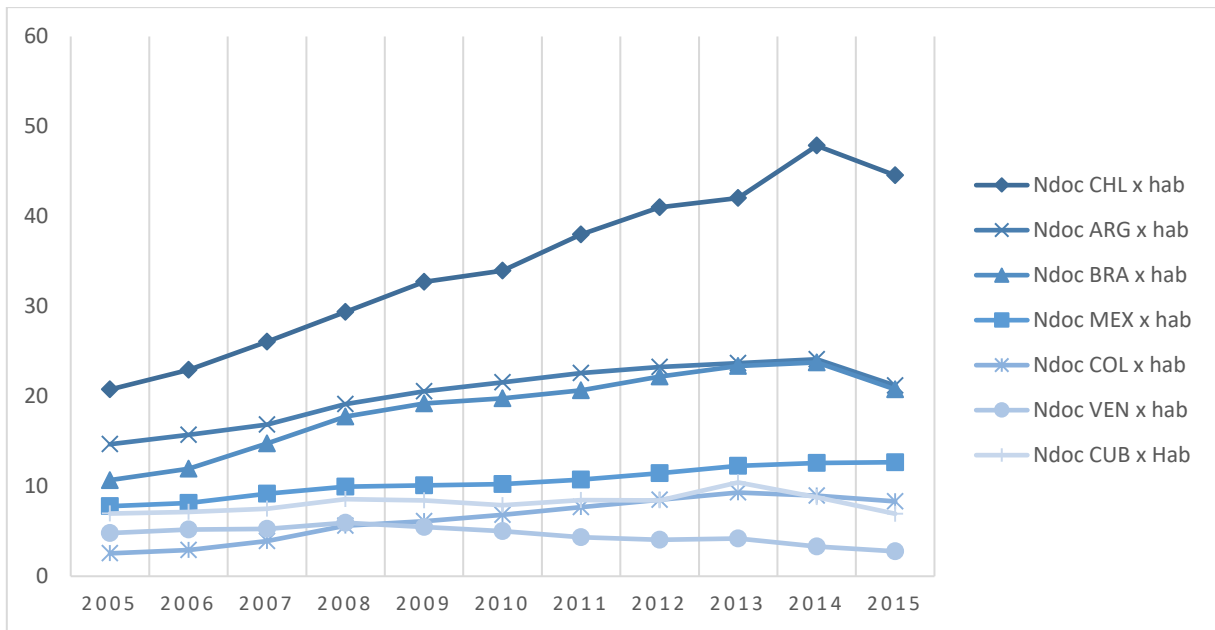


Figura 15. Comparación de la producción científica mexicana en el WoS por cada 100,000 habitantes con países de la región, 2005-2015

En una comparativa con países que no pertenecen a la región latinoamericana, la mayor producción de artículos científicos se concentra en las naciones más desarrolladas: Estados Unidos, China, Reino Unido, Alemania, Japón, Francia, Canadá, pero en este estudio se determinó comparar a México con países de similar producción científica. En la figura 16, México ocupa la posición número tres, superada por República Checa y Portugal, se denota el desarrollo que ha alcanzado el país en materia de generación de conocimiento científico, aspecto

que se favorece gracias al plan que tiene el CONACYT para hacer avanzar la ciencia mexicana.

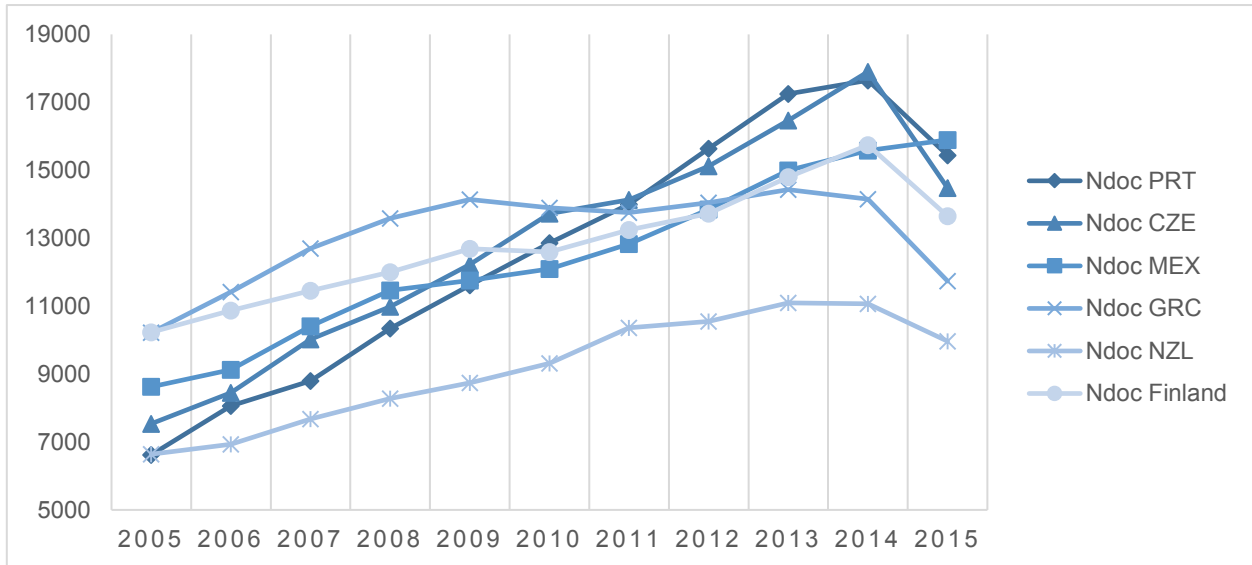


Figura 16. Comparación de la producción mexicana en WoS con países que no pertenecen a la región latinoamericana, con similar producción de acuerdo al ESI, durante el período 2005-2015

Los investigadores desempeñan un papel importante para promover la innovación y por tanto aumentar el número de documentos científicos publicados en revistas indexadas en Web of Science, pero México está rezagado en este rubro, como lo muestra la figura 16 donde el país queda en último lugar (figura 16).

Los investigadores generan conocimientos y estos pueden ser usados por las empresas como un medio para aumentar la productividad mediante procesos perfeccionados, introducir nuevas tecnologías, crear mejores programas organizativos, desarrollar nuevos productos y publicación de artículos científicos. La posibilidad de que la actuación de los investigadores pueda concretarse en beneficios económicos depende en parte de los vínculos con la especialización regional y con las industrias.

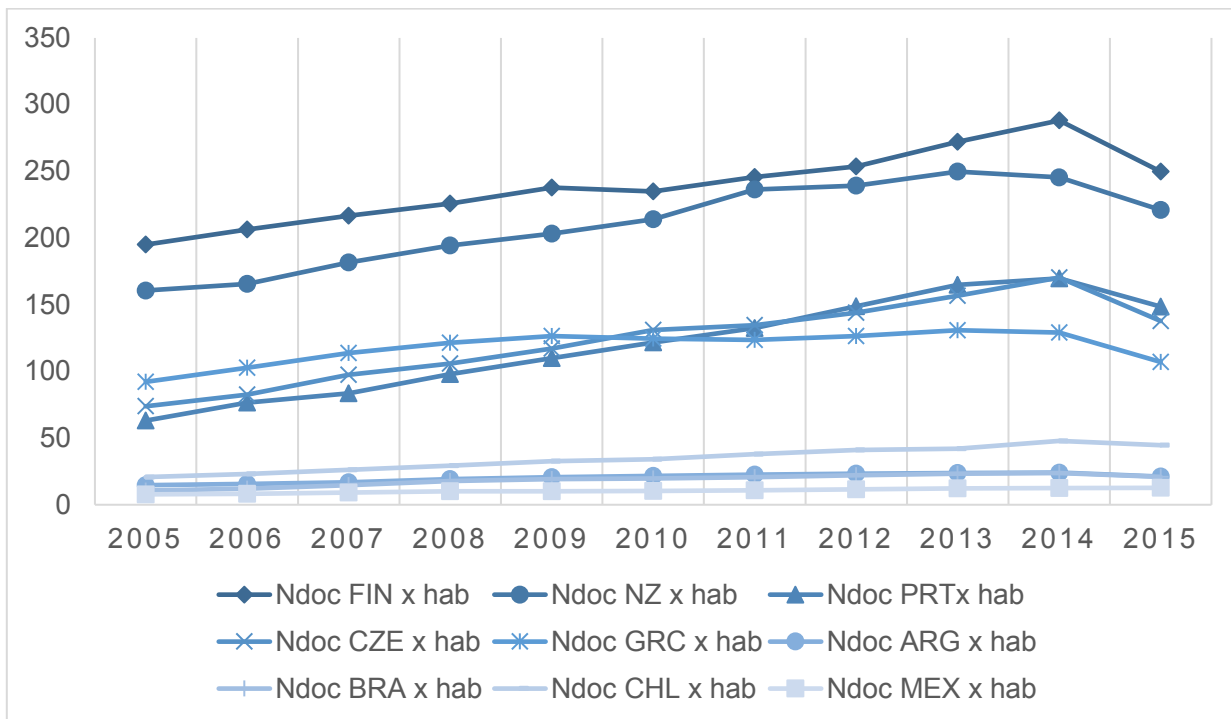


Figura 17. La producción científica mexicana en el WoS por cada 100,000 habitantes con la de los países con similar producción de acuerdo al ESI, 2005-2015

La ciencia y la investigación proporcionan las bases para un crecimiento económico sostenible, mediante la creación y el mantenimiento de las bases de conocimientos y tecnologías que sustentan la competitividad de un país. En las economías avanzadas, la innovación en todas sus formas es el principal motor de crecimiento económico a largo plazo. Muchos de los avances científicos pueden utilizarse rápidamente para desarrollar nuevos productos y servicios para mejorar progresivamente los existentes.

Los activos intangibles tales como nuevas formas de hacer las cosas que resultan en innovaciones, mejoras organizacionales y de procesos en la eficacia resultan de la ciencia y la investigación. En todos estos casos, la posición internacional de la salida de la investigación como indica la calidad de las publicaciones indica el potencial para la resolución de problemas, así como la capacidad técnica, humana y de organización del país.

2.4. Highly Cited y Hot papers: Comparación de la producción mexicana con otros países

La figura 18 ofrece la posición de México en comparación con países de la región y países fuera de la región de acuerdo con el número de Top Papers. México ocupa el segundo lugar después de Brasil en la producción de Top Papers, sin embargo, al comparar con otros países queda en último lugar. Es decir, países más avanzados como Finlandia, Portugal y República Checa, que en su totalidad producen un número similar de artículos en el WoS como México, la cantidad de artículos altamente citados supera la mitad de los artículos altamente citados mexicanos (Anexo 1).

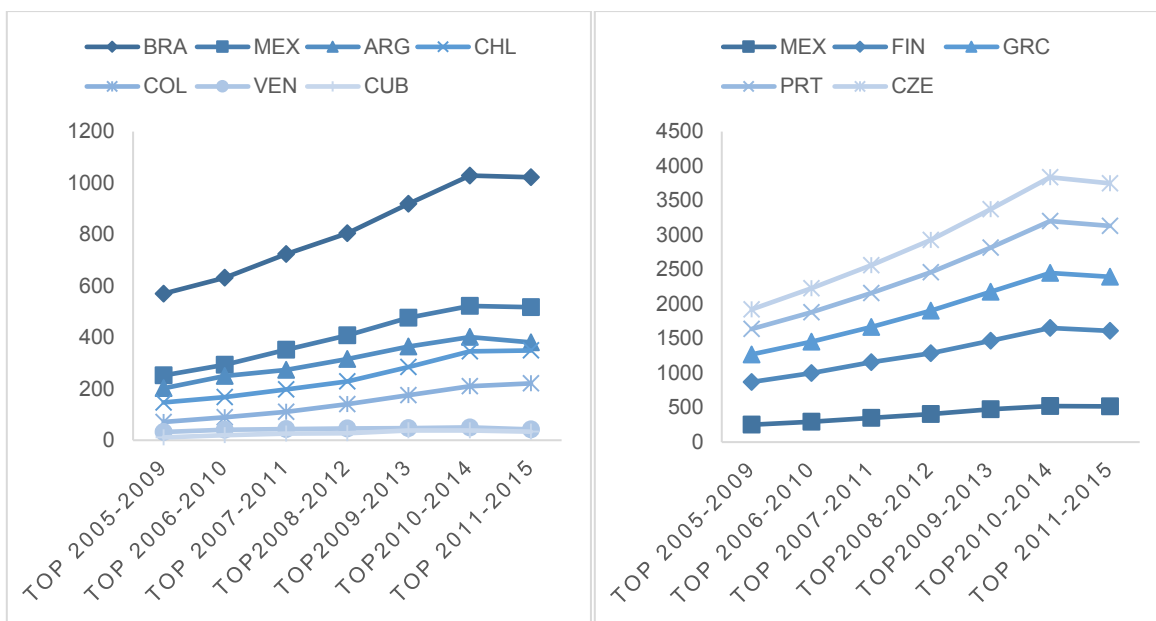


Figura 18. Comparativa del número de Top Papers mexicanos en el Web of Science con países de la región y otros países con similar producción, 2005-2015

La figura 19 muestra el número de Highly Cited Papers mexicanos en el Web of Science con países de la región y países de producción similar, 2005-2015. Este indicador proporciona la excelencia científica por país, de tal manera México ocupa el segundo lugar cuando se compara con países de la región, después de Brasil, y en último lugar cuando se compara con países como: Finlandia, Nueva Zelanda, Grecia y Portugal.

Los Highly Cited Papers permiten conocer la excelencia en la investigación y la influencia reconocida que está teniendo en el debate académico en todo el mundo la ciencia mexicana. Por otro lado, estos resultados facilitan tener una visión de la base de investigación con excelencia, importante elemento para atraer la inversión extranjera a la I+D+I en México (Anexo 1).

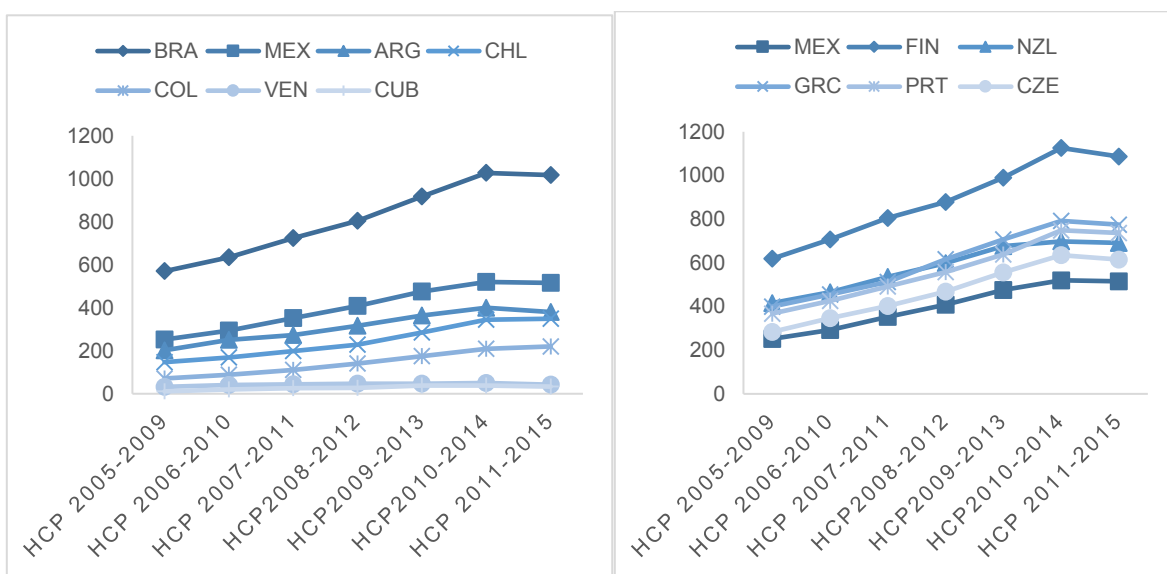


Figura 19. Comparativa del número de Highly Cited Papers mexicanos en el Web of Science con países de la región y países de similar producción, 2005-2015

La figura 20 muestra, durante el período 2005-2015, la tendencia de la investigación mexicana teniendo en consideración el número de Hot papers publicados. Durante los períodos 2005-2009, 2006-2010, y 2007-2011, México y los países comparables, tanto de la región como los otros, ninguno de sus artículos, logró recibir citas en los dos últimos años y en los últimos dos meses para ubicarse en el 0.1% superior en comparación con los trabajos publicados en la misma disciplina y de esta manera ser considerados Hot paper (Reuters, 2016). Comportamiento que, se incrementa a partir del período 2008-2012 para todos los países incluyendo México que publicó 23 hot papers para el período 2010-2014 y 36 hot papers para los años 2011-2015 (Anexo 1).

En la figura se observa cómo, el impacto de la investigación mexicana solo es superada por Brasil, sin embargo, en relación con otros países fuera de la región latinoamericana, la investigación mexicana no recibe un alto reconocimiento, quiere decir, que la ciencia procedente de estos países desarrollados recibe un alto y rápido número de citas, hay un uso más amplio de los conocimientos generados por Nueva Zelanda, República Checa y Portugal. En este sentido, México debe incrementar la inversión en I+D.

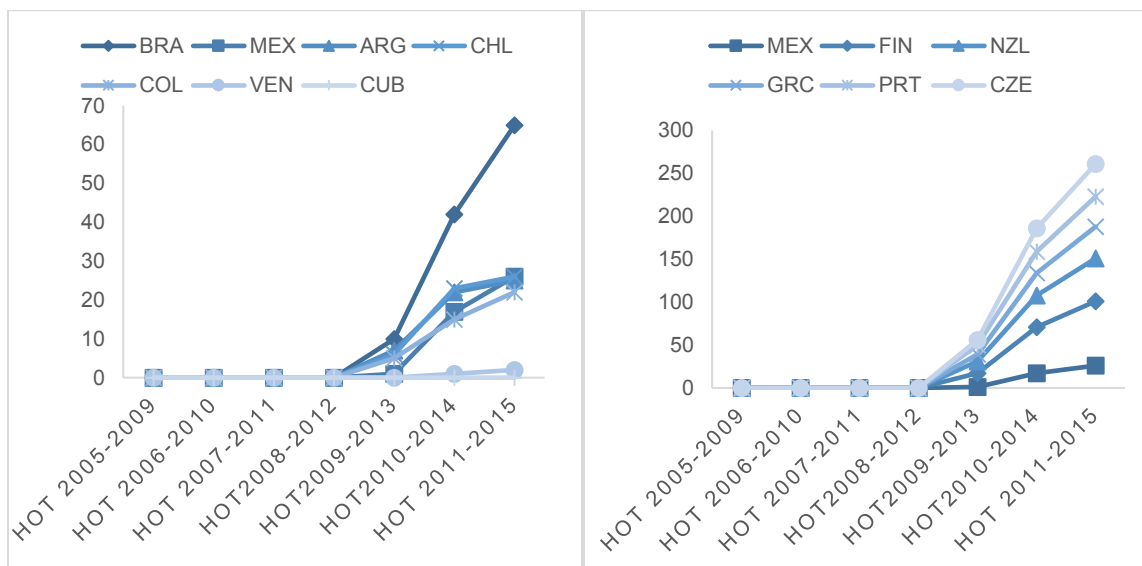


Figura 20. Comparativa del número de Hot Papers mexicanos en el Web of Science con países de la región y países de similar producción, 2005-2015

Los valores más altos del número de Top Papers por cada 1,000 documentos en WoS, fueron obtenidos por países como: Colombia, Chile y Argentina, los cuales tienen valores relativamente bajos de productividad con respecto a México (figura 21) (Anexos 2 y 3), un elemento importante a señalar es el comportamiento de Cuba para el período de estudio, la cual ha tenido una tasa de crecimiento mayor al resto de los países de esta figura y Brasil ocupa el último lugar en la misma. Sin embargo, al comparar México con Finlandia, Nueva Zelanda, Portugal, Grecia y República Checa porque tienen similar producción científica, México ocupa el último lugar, es decir, México país en vías de desarrollo tiene niveles de producción científica a la par de países desarrollados, pero sus resultados de investigación no reciben un alto impacto, es decir, un alto número de citas, como reciben los países antes mencionados.

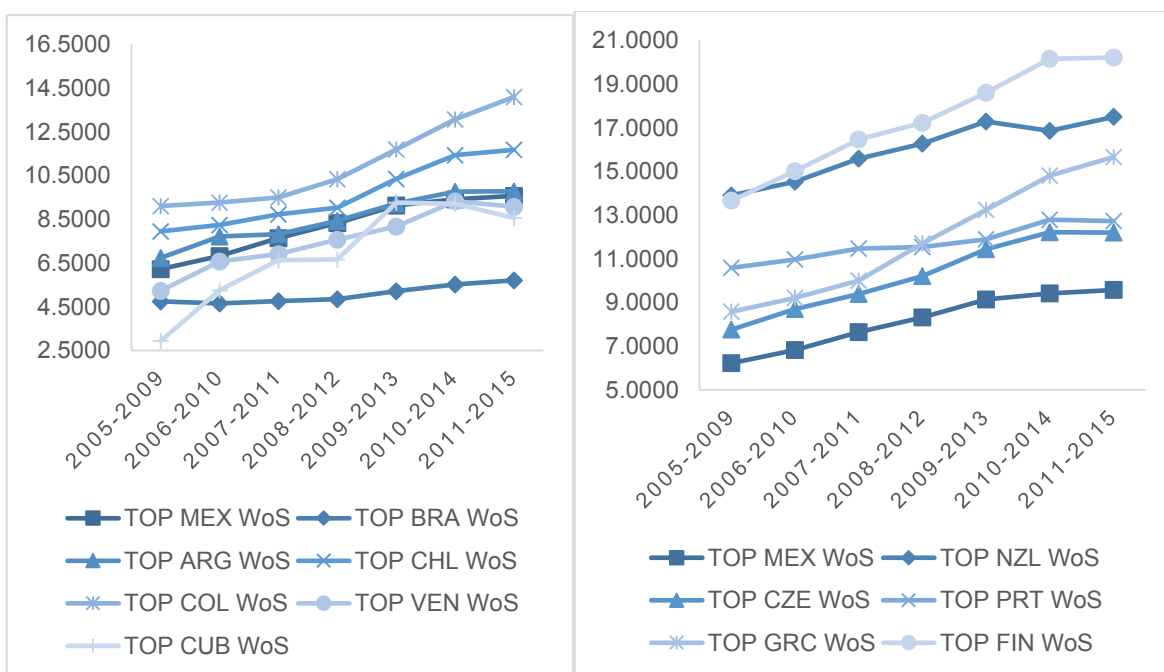


Figura 21. Comparativa del número de Top Papers mexicanos por cada 1,000 documentos en el Web of Science con países de la región y con países de similar producción, 2005-2015

La figura 22, muestra el número de Highly Cited Papers mexicanos por cada 1,000 documentos en WoS, en comparación con países de la región para el período 2005-2015, Colombia, Chile y Argentina superan a México, es decir, durante el período de estudio se denota la estabilidad de México, sin embargo, Brasil queda en último lugar, porque a pesar de producir un mayor número de documentos el número de Highly Cited Papers brasileños por cada 1,000 documentos en WoS es bajo. Contrastando México con otros países que no pertenecen a la región, los valores más altos fueron obtenidos por Finlandia, Nueva Zelanda y Grecia (Anexos 2 y 3).

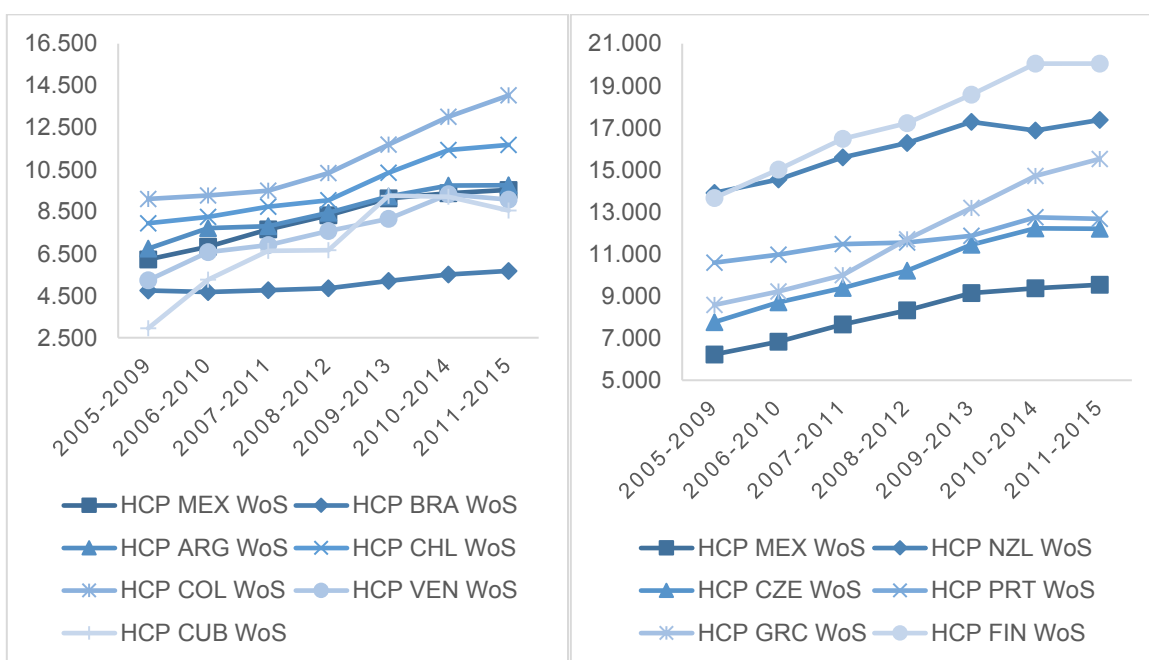


Figura 22. Comparativa del número de Highly Cited Papers mexicanos por cada 1,000 documentos en el Web of Science con países de la región y países de similar producción, 2005-2015

Colombia, Chile y Argentina presentan el número mayor de Hot Papers por cada 1,000 documentos en WoS, por encima de México, sin embargo, Brasil ocupa el último lugar, a pesar de que este país es el mayor productor de la región y además el gobierno brasileño destina mayores recursos económicos a la I+D que México (figura 23). Además, al comparar México con otros países, Finlandia, Nueva Zelanda y Grecia presentan un número mayor de Hot Papers por cada 1,000 documentos en WoS. En este sentido, es preciso mencionar que son países comparables con México porque tienen similar producción científica (Anexos 2 y 3).

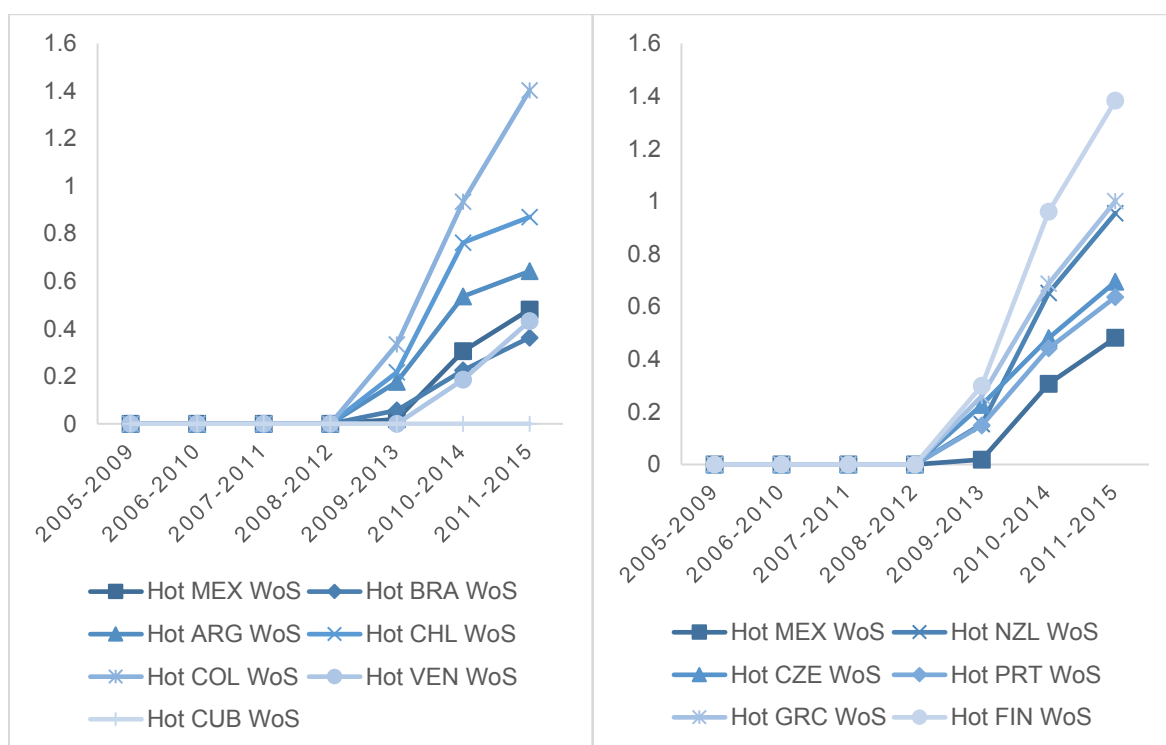


Figura 23. Comparativa del número de Hot Papers mexicanos por cada 1,000 documentos en el Web of Science con países de la región y países de similar producción, 2005-2015

En una comparativa del número de Highly Cited Papers mexicanos por cada 100,000 habitantes, con países de la región y países de similar producción, se denota como, en este caso, Chile y Argentina presentan un porcentaje mayor de Highly Cited Papers, en cambio, Brasil, que se considera como el país de la región con mayor prestigio, presenta un comportamiento similar al de México. En comparación con otros países México sigue ocupando el último lugar (figura 24) (Anexos 4y 5).

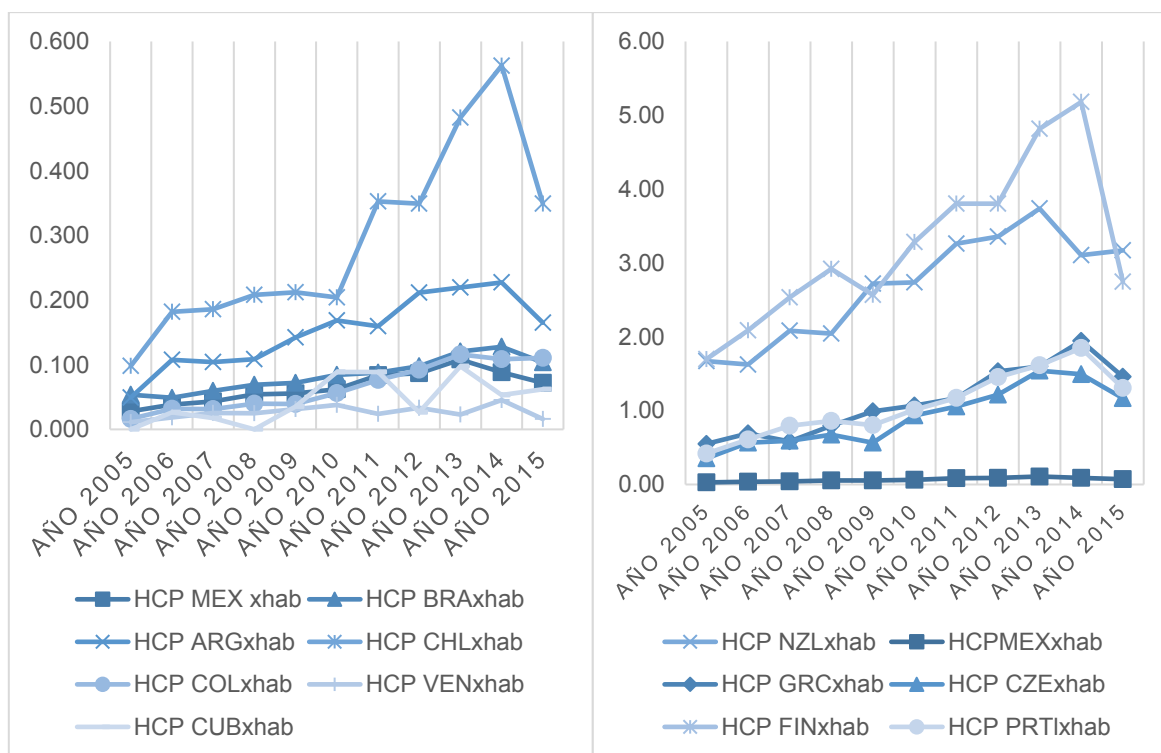


Figura 24. Comparativa del número de Highly Cited Papers mexicanos por cada 100,000 habitantes, con países de la región y países de similar producción, 2005-2015

La figura 25 muestra del período 2013-2015, los años a partir de los cuales tanto México y los países comparables, comenzaron a publicar artículos que, se consideran Hot paper, es decir, artículos publicados en los dos últimos años, de acuerdo, con citas recibidas en los últimos dos meses se hallan en el 0.1% superior en comparación con los trabajos publicados en la misma disciplina. Desde el año 2005 hasta el año 2012, no se publicó al menos un artículo que se considera Hot paper (Reuters, 2016). Además, Chile, Argentina y Colombia presentan un número mayor de Hot paper por cada 100,000 habitantes, por encima de México y Brasil. En comparación con países desarrollados México sigue ocupando el último lugar (Anexos 4 y 5).

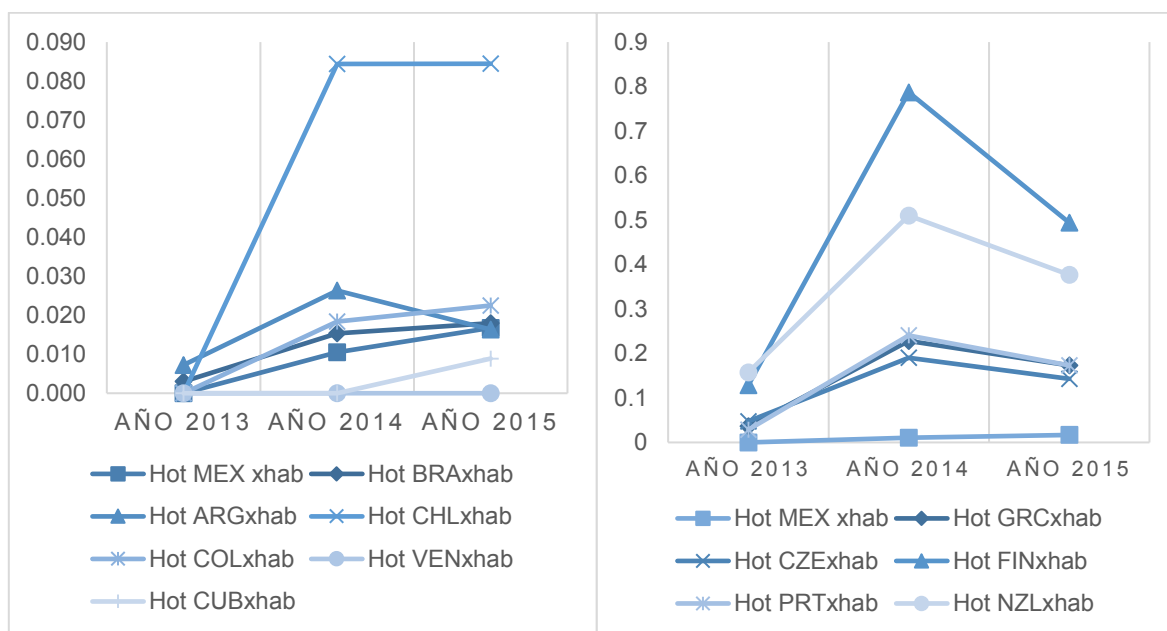


Figura 25. Comparativa del número de Hot Papers mexicanos por cada 100,000 habitantes, con países de la región y países de similar producción, 2005-2015

2.5. Principales campos de investigación en la ciencia mexicana

El análisis de las temáticas o campos de investigación de la ciencia mexicana en el WoS durante el período 2005-2015, implicó la necesidad de establecer una distribución temática. Para la presente investigación se determinó la clasificación temática que ofrece el Essential Science Indicators porque esta herramienta asigna una sola revista por cada una de las 22 categorías o campos de investigación, y en el caso de revistas multidisciplinarias como Science y Nature, las categorías se asignan por artículo, de esta manera utiliza una clasificación temática más general, y la ventaja principal es que no permite solapamientos; aspecto que sí ocurre en la clasificación de la colección principal del Web of Science, donde algunas revistas pertenecen a varias categorías.

Se observa que la mayor producción de documentos mexicanos en el Web of Science para el período 2005-2015, corresponde a PLANT & ANIMAL SCIENCE con 42,518 documentos publicados sobre este campo de investigación, le sigue PHYSICS (38,463), CLINICAL MEDICINE (33,108) y CHEMISTRY (32,938) (figura 26). Es importante mencionar como SPACE SCIENCE en esta figura no aparece como uno de los campos de investigación con mayor número de documentos, pero su producción representa el 7% del total de documentos producidos a nivel mundial en ese campo de la ciencia, le sigue PLANT & ANIMAL SCIENCE (6%), ENVIRONMENT/ECOLOGY (6%) y AGRICULTURAL SCIENCES con 5%, la ciencia mexicana en estos campos de investigación reúne porcentajes muy bajos.

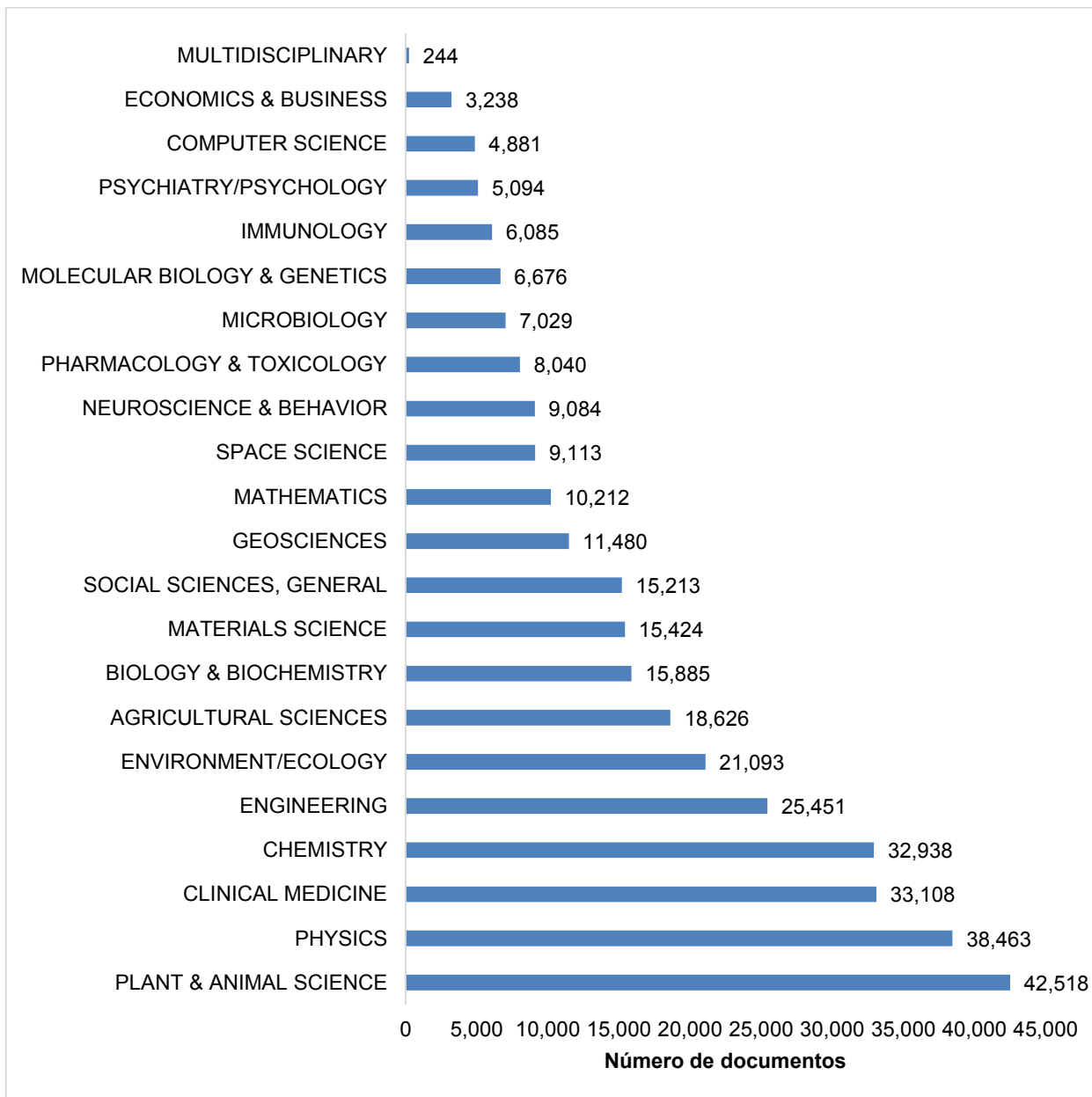


Figura 26. Distribución del número de documentos mexicanos en el WoS, durante 2005-2015, entre 22 campos de investigación

2.5.1. Los campos de investigación con mayor impacto en la ciencia mexicana

La figura 27 muestra los campos de investigación mexicanos con mayor impacto, PHYSICS, CLINICAL MEDICINE, CHEMISTRY, PLANT & ANIMAL SCIENCE, BIOLOGY & BIOCHEMISTRY y ENVIRONMENT/ECOLOGY son los campos de investigación que reciben un mayor número de citas y, por tanto, un mayor reconocimiento por parte de la comunidad científica. Los campos de investigación con mayor impacto durante el período de estudio, coinciden con las áreas de conocimiento estratégico propuestas por el CONACYT para fomentar la I+D+I y por tanto detectar nuevas oportunidades de desarrollo económico, en donde la ciencia y la innovación juegan un papel importante.

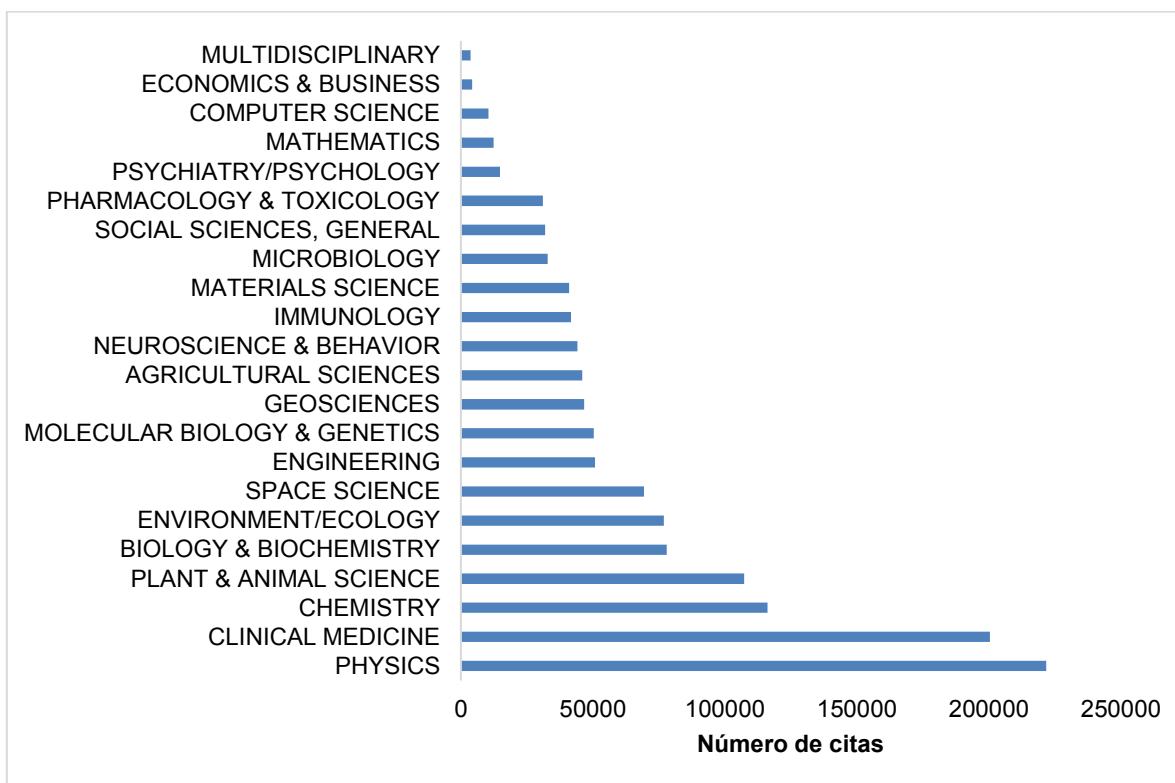


Figura 27. Distribución del número de citas entre 22 campos de investigación en el WoS, durante 2005-2015

2.5.2. Los campos de investigación científica mexicana con mayor efectividad y excelencia

Cuando se analiza la efectividad de la ciencia mexicana entre 22 campos de investigación, es decir, en una relación entre el total de citas y el total de documentos, áreas como: Multidisciplinary, Space Science, Molecular Biology & Genetics, e Immunology muestran los mejores valores porque aunque son áreas con un bajo volumen de documentos si reciben un mayor número de citas que otras áreas por el impacto que las investigaciones en estos campos para la comunidad científica internacional.

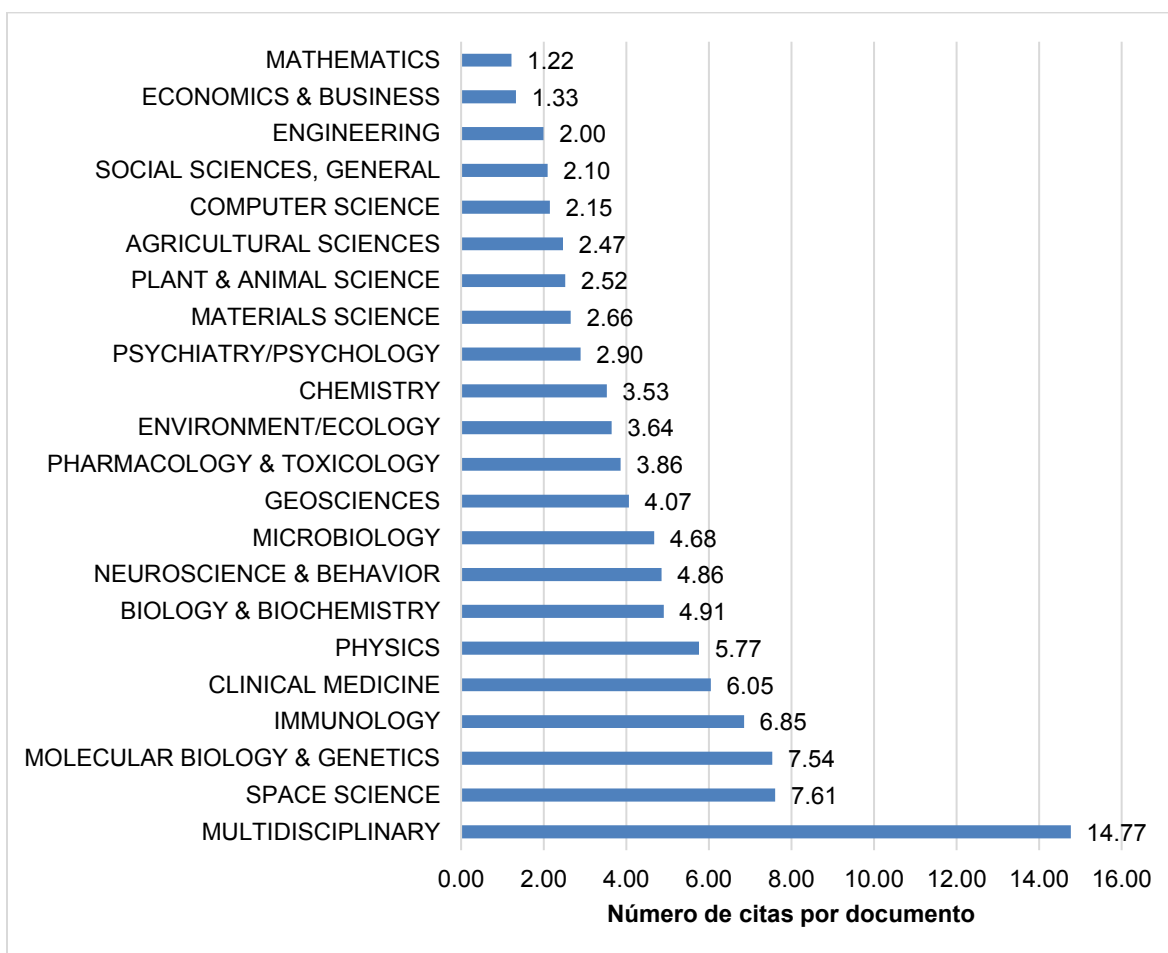


Figura 28. Distribución del promedio de citas por documento a través de 22 campos de investigación en el WoS, durante 2005-2015

La efectividad de la ciencia mexicana prevalece en campos de investigación como: CLINICAL MEDICINE, PHYSICS y PLANT & ANIMAL SCIENCE, esto no solo se interpreta como una coincidencia en los temas que los investigadores publican con mayor frecuencia, sino que durante el período existe un mayor número de Highly Cited Papers en estas temáticas; por tanto, se denota un aumento del número de citas recibidas y al mismo tiempo hay un mayor impacto y reconocimiento de la ciencia mexicana en estas áreas por parte de la comunidad científica internacional.

Este comportamiento en estos campos de investigación se debe en parte a los documentos relacionados con CLINICAL MEDICINE los cuales generan aportes en medicina general interna, genética y oncología, es decir, temas de gran importancia por el impacto que tienen estas contribuciones para la población mexicana; en el caso de PHYSICS los investigadores mexicanos formaron parte de proyectos en temáticas como la astrofísica, la física nuclear y de partículas; y en PLANT & ANIMAL SCIENCE, los aportes refieren a aspectos relacionados con la agronomía, horticultura y tecnología de los alimentos, líneas que se consideran de importancia para el desarrollo agropecuario del país. (figura 29)

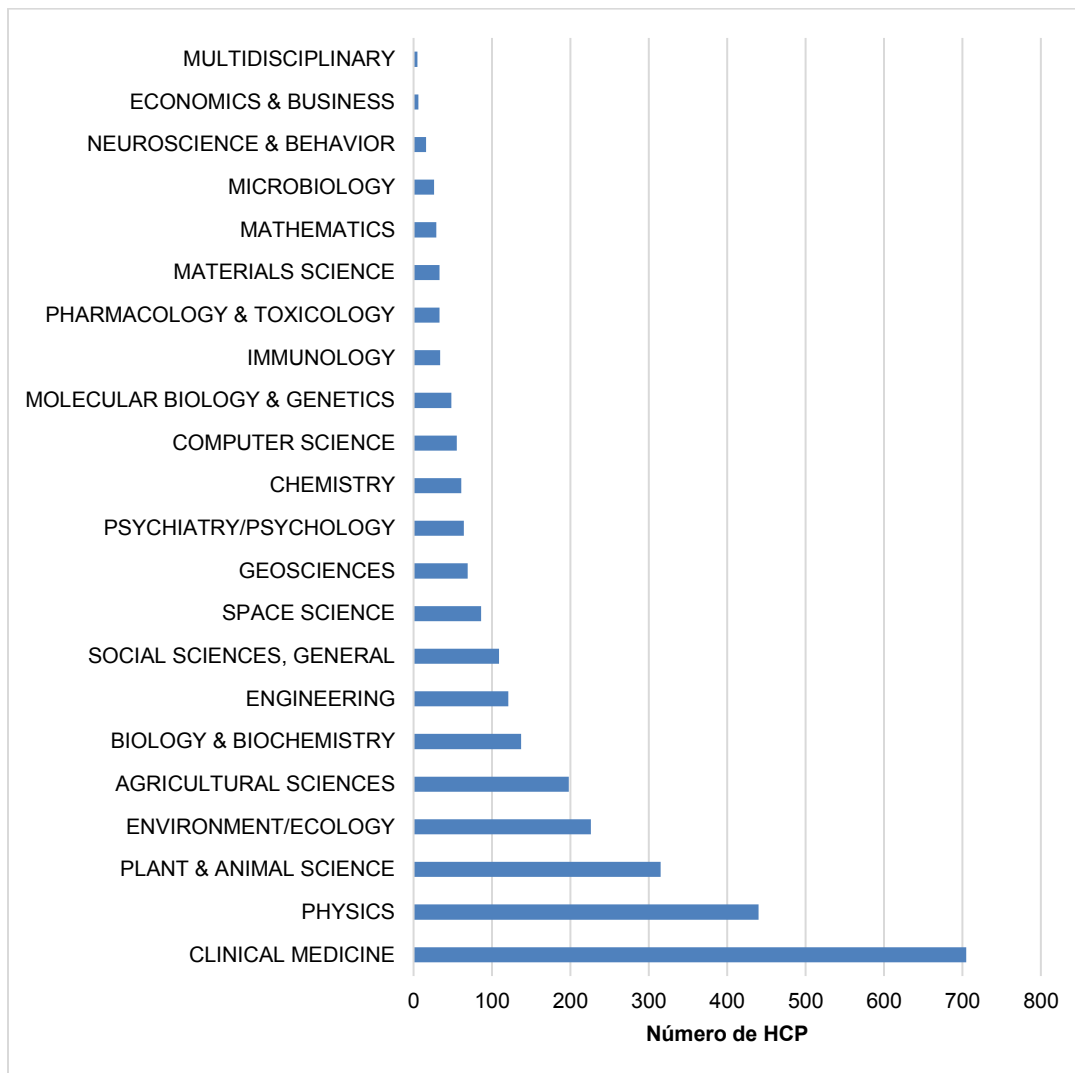


Figura 29. Distribución entre 22 campos de investigación del número de Highly Cited Papers mexicanos en el WoS, durante 2005-2015

La figura 30 muestra el número de Hot Papers mexicanos en solo 9 campos de investigación los otros 13 campos de investigación no se representan porque no se publicaron artículos que de acuerdo con las citas recibidas en los últimos meses se hallan en el 0.1% superior con respecto a otros artículos publicados en las misma disciplina en los dos últimos año, es decir, CLINICAL MEDICINE, PLANT & ANIMAL SCIENCE, PHYSICS y ENVIRONMENT/ECOLOGY constituyen las áreas de la ciencia mexicana que recibieron y están recibiendo un reconocimiento muy rápido, lo cual puede constituir una señal de una nueva e importante disposición en la investigación mexicana para consolidar proyectos de I+D+I.

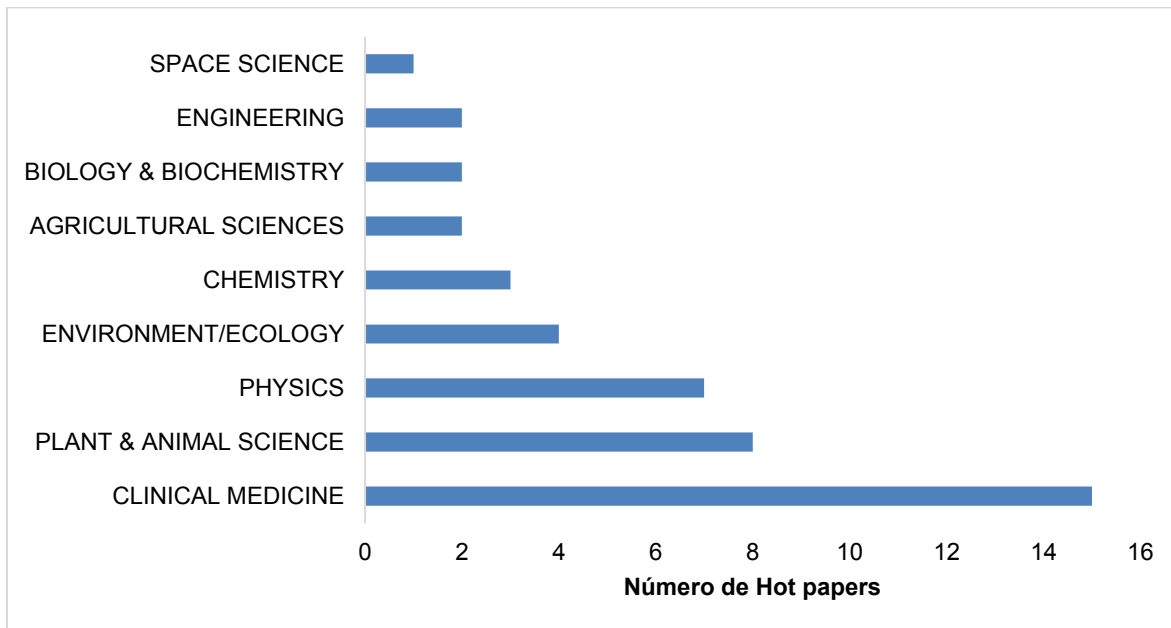


Figura 30. Distribución de los 9 de 22 campos de investigación del número de Hot Papers mexicanos en el WoS, durante 2005-2015

2.6. Colaboración científica

En la actualidad, la colaboración es una actividad inherente al proceso investigador y no es únicamente importante porque se difunden los resultados de una investigación conjunta, sino porque permiten: a) generar conocimiento dentro de comunidades científicas distintas y sin barreras geográficas; b) determinar el nivel de integración en la comunidad científica; c) poner de manifiesto el grado de internacionalidad o aislamiento, su grado de colaboración nacional, institucional, regional, las diferencias entre áreas y el efecto de los resultados en términos de visibilidad; d) la detección de los denominados “colegios invisibles”, y además f) son un indicador de la movilidad de los investigadores y de la globalización de la investigación científica en su conjunto (Chinchilla-Rodríguez, 2004)

La figura 31 muestra el comportamiento de la colaboración internacional por año y por países, siendo USA el país más colaborador, con una tendencia al aumento durante los últimos años del período. En contraste, la colaboración con España, Francia, Alemania creció en la última etapa analizada. Se denota además como países de la región como Brasil, Colombia y Argentina mantiene fuertes lazos de colaboración con México.

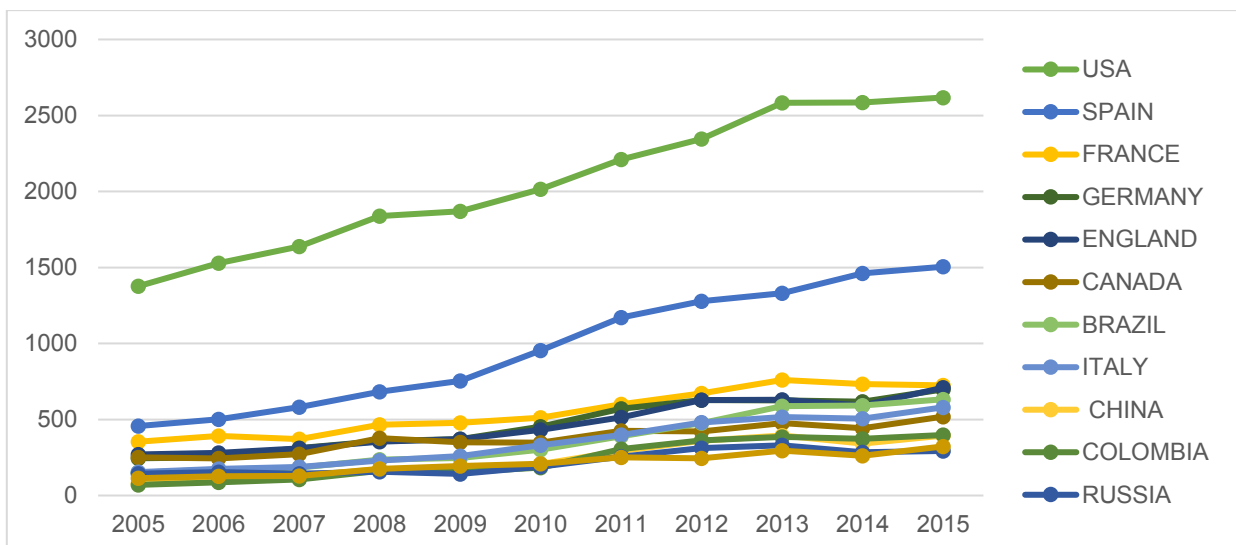


Figura 31. Evolución anual de la colaboración con los países más colaboradores con México en el período 2005-2015. **Total de países colaboradores: 193**

3. Análisis multiparamétrico del perfil bibliométrico de México y otros países, basado en la técnica de mapas auto-organizados

3.1. Ranking de los países de acuerdo al ESI del WoS

México, se encuentra en el ranking número 36 de los 155 países con mayor número de citas, según el Essential Science Indicators (ESI) de WoS. En este ranking, los primeros lugares lo ocupan países altamente desarrollados como USA, Alemania, Inglaterra, China Francia y Japón. (Tabla 9 y Anexo 6).

Tabla 9. Rank de países según ESI durante el período 2005-2015.

	Países	No. Citas	Citas/doc	SDI	HCP	Hot	RI	IN
1	USA	66790551	17.7	11.83	64464	1420	1.54	1.47
2	GERMANY (FED REP GER)	15704549	15.89	12.22	15291	430	1.39	1.30
3	ENGLAND	15594264	17.5	13.81	17349	497	1.53	1.51
4	CHINA MAINLAND	13923842	8.4	1.21	15863	574	0.73	0.74
5	FRANCE	10505229	15.07	10.53	10144	309	1.31	1.27
6	JAPAN	10094917	11.95	6.64	6203	167	1.04	0.95
7	CANADA	9514790	15.72	17.03	9951	272	1.37	1.33
8	ITALY	8317827	14.28	9.5	7740	258	1.24	1.11
9	AUSTRALIA	6548488	14.19	19.65	7863	291	1.24	1.24
10	NETHERLANDS	6488596	18.74	20.55	7571	233	1.63	1.52
11	SPAIN	6472534	13.03	10.7	6409	228	1.14	1.10
12	SWITZERLAND	5004142	19.97	30.59	6126	215	1.74	1.64
13	SOUTH KOREA	4139433	9.16	8.97	3469	135	0.8	0.83
14	SWEDEN	3905546	16.95	23.79	4063	139	1.48	1.38
15	INDIA	3644391	7.76	0.37	2529	70	0.68	0.64
16	BELGIUM	3211131	16.86	16.97	3720	115	1.47	1.36
17	BRAZIL	2654948	7.66	1.72	1812	78	0.67	0.68
18	SCOTLAND	2590044	18.85	29.79	3070	110	1.64	1.58
19	DENMARK	2536006	18.16	24.76	3184	125	1.58	1.46
20	TAIWAN	2439899	9.42	11.21	1606	52	0.82	0.87
21	AUSTRIA	2029723	15.64	15.2	2372	76	1.36	1.29
22	ISRAEL	1981133	14.95	16.13	1830	68	1.3	1.24
23	RUSSIA	1821034	5.93	2.14	1345	65	0.52	0.52
24	FINLAND	1778844	15.61	20.86	1917	63	1.36	1.33
25	POLAND	1778534	8.07	5.8	1585	67	0.7	0.66
26	NORWAY	1604353	15.09	20.7	1864	69	1.32	1.28

27	TURKEY	1582711	6.67	3.13	1162	39	0.58	0.62
28	SINGAPORE	1514563	14.79	18.73	2190	86	1.29	1.21
29	HONG KONG	1466883	13.14	15.41	1820	57	1.15	1.24
30	GREECE	1356016	12.26	10.09	1371	52	1.07	1.02
31	PORTUGAL	1259786	11.76	10.3	1251	51	1.03	1.05
32	IRAN	1152290	6.06	2.42	1029	48	0.53	0.53
33	NEW ZEALAND	1089433	13.58	17.79	1241	38	1.18	1.18
34	IRELAND	1080339	15.54	11.22	1320	53	1.35	1.51
35	CZECH REPUBLIC	1040485	10.35	9.57	1087	38	0.9	0.85
36	MEXICO	939491	8.66	0.88	847	36	0.76	0.78

Leyenda: Número de citas (No. citas), promedio de citas por documentos (citas/doc), índice de desarrollo científico (SDI), *Highly Cited Papers* (HCP), *Hot papers* (Hot), Impacto relativo (RI) e Impacto normalizado (NI)

3.2. Ranking de los países de acuerdo al grado de desarrollo científico

El Índice de desarrollo científico (SDI), es un indicador elaborado en el marco de esta investigación, y se define como la razón del número documentos publicados en WoS por habitante para cada país. Este indicador permite comparar países de manera más justa; por lo que una vez que fue calculado este indicador para los 155 países, se propuso un nuevo ranking de países de acuerdo al SDI, de esta manera se seleccionan los primeros 100 países con mayores valores de SDI para realizar una comparación y determinar sus perfiles de acuerdo a cuatro indicadores bibliométricos.

Luego con el cálculo del Índice de desarrollo científico por país, se propone un nuevo ranking; los países que ocuparon los primeros lugares son: Suiza, Escocia, Islandia, Dinamarca, Suecia y Finlandia, por su parte México en este ranking de países ocupa el número 67 (Anexo 7). Después se determinó que del total de la población de países (100) a analizar se excluyeron los primeros 30 países, y se analizaron 70 países que muestran un desempeño similar a México de acuerdo el índice de desarrollo científico (Anexo 7).

A partir de este ranking, resulta interesante observar cómo ha evolucionado este indicador para el período de estudio y comparar a México con países que quedaron muy cerca según el ranking (tres arriba y tres debajo de México).

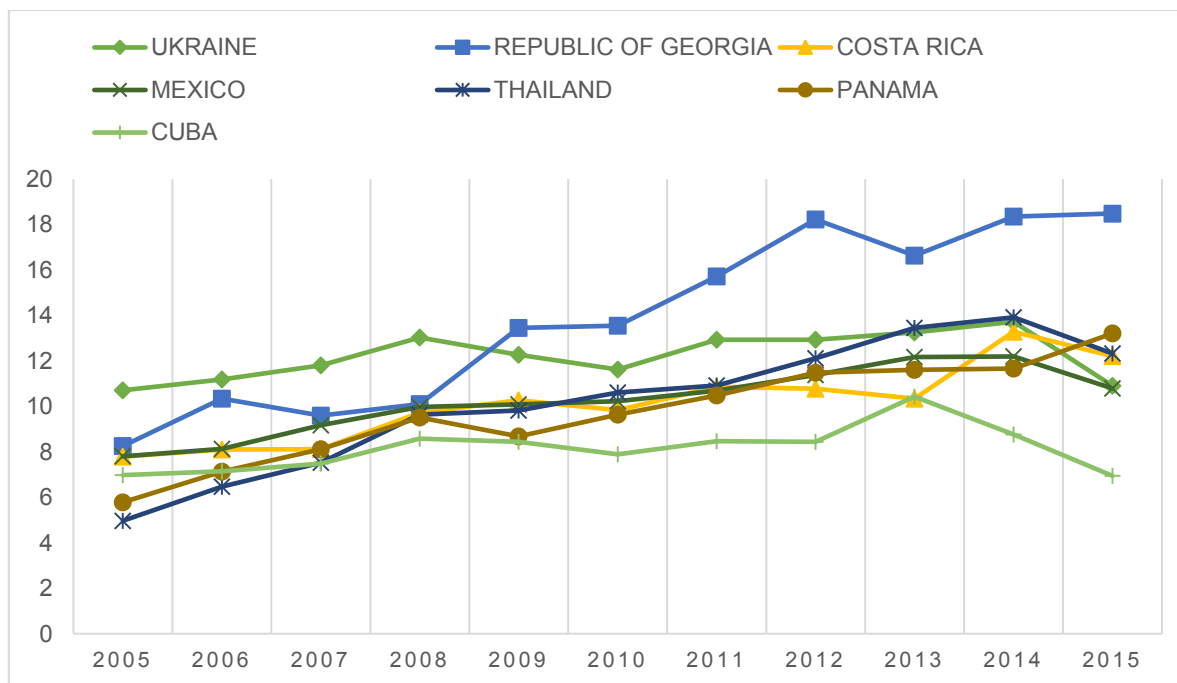


Figura 32. Evolución en el tiempo del índice de desarrollo científico, número de artículos publicados en revistas indexadas en WoS por cada 100,000 habitantes, para México y países con similar desempeño

En la figura 32 se observa la evolución del índice de desarrollo científico, donde México mantiene un comportamiento estable a través del período, superado por Ucrania, Georgia y Tailandia, países partir del año 2008, su índice de desarrollo científico comienza a aumentar, es decir, aumenta el número de artículos publicados en revistas indexadas en WoS por cada 100,000 habitantes. Importante mencionar como Panamá y Costa Rica, en el último año del período muestran un ascenso en su desarrollo científico.

Para contrastar este grupo de países, se seleccionó además del índice de desarrollo científico, el Impacto Normalizado, así como, el número de HCP y Hot, estos últimos procedentes del ESI herramienta de WoS. Antes de proceder a comparar la población de 100 países, es necesario observar el grado de correlación entre estos indicadores, la tabla 10 muestra lo anterior. Las parejas (HCP;Hot) y (NI;

SDI) exhiben los valores de correlación más altos, por su parte (HCP, SDI) y (Hot, SDI) muestran valores inferiores.

En todas estas figuras quedaron excluidos los países que ocupan en el ranking del número 1 al 30, con el fin de obtener una idea más clara de la conducta más típica de cada país y eliminar los comportamientos atípicos (outliers) porque presentan altos valores para los indicadores calculados. Este análisis abarca el período 2005-2015.

Tabla 10. Correlación de los indicadores estudiados para una muestra de 100 países

	SDI	HCP	Hot	NI
SDI	1	0.29	0.27	0.30
HCP		1	0.99	0.24
Hot			1	0.22
IN				1

A partir de este análisis de correlación se realizan los diagramas de dispersión con las parejas de indicadores que ofrecen un bajo grado de correlación. En las figuras 33 y 34 se muestran los datos de los 70 países desde dos perspectivas diferentes: en primer lugar, se identifican los países que poseen un mejor rendimiento bibliométrico al comparar en un diagrama de dispersión, el impacto normalizado y la eficiencia (SDI) (figura 33); en segundo lugar, se contrastan el número de HCP paper y el impacto normalizado (figura 34). Complementa el análisis una comparación entre Hot y el impacto normalizado.

La figura 33 muestra el conjunto de países con mejor rendimiento bibliométrico desde una perspectiva basada en la eficiencia y el impacto: se contrastan indicadores que toman en cuenta el volumen de la producción científica en relación al número de habitantes, y el impacto normalizado, estimado por el cociente del promedio de citas por país por cada una de las 22 área de investigación de ESI y el rendimiento medio del mundo. Existen diferencias en las tasas de citas para los diversos campos de investigación y períodos de tiempo, por tanto, el cálculo se realiza por separado.

El hecho de que, los países se organizan un tanto alejado de la línea recta en el gráfico (con un coeficiente de correlación de 0.098), demuestra que existe un bajo grado de correlación entre estos dos indicadores: los valores del NI no crecen en proporción directa a los valores del índice de desarrollo científico (SDI). Se destacan países como Italia, Korea del Sur (PKR), Hungría, Japón, Croasia (HRV), Jordania, Lituania, Luxemburgo, México, Panamá, Colombia y Cuba; demostrando que el impacto científico de México es comparable con países de la región y países con economías superiores a la nuestra.

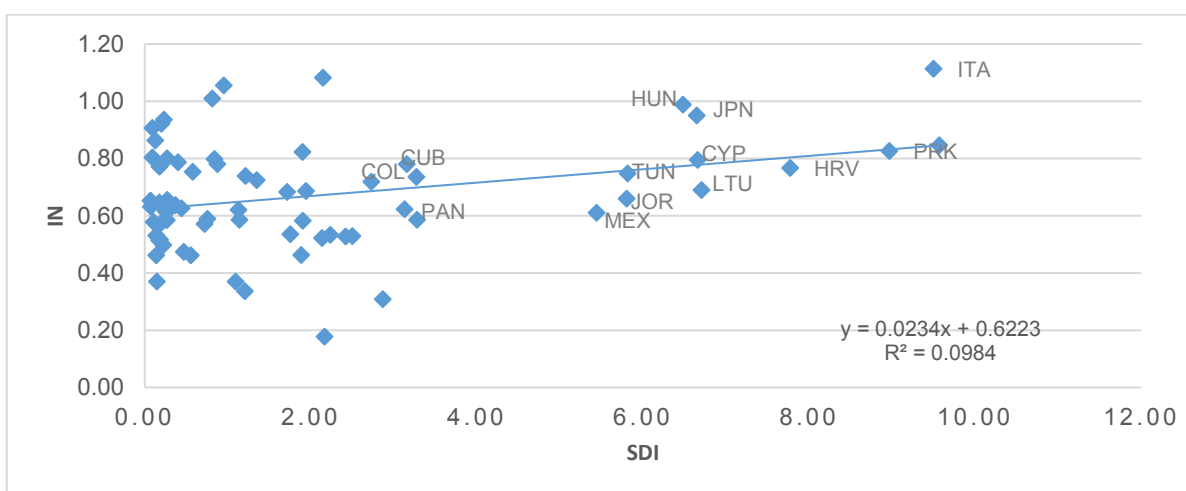


Figura 33. Países con altos valores de impacto (NI) y el índice de desarrollo científico (SDI) durante el 2005-2015

El diagrama de dispersión (figura 34), el cual compara el conjunto de 70 países desde una perspectiva basada en el número de Highly Cited Papers y el impacto normalizado (NI). El índice de correlación (R^2) demuestra el bajo grado de correlación entre estos indicadores. China, Korea del Sur, Argentina, Japón, México, Italia, Brasil y Chile, presentan en este orden los mejores valores del número de HCP e Impacto normalizado.

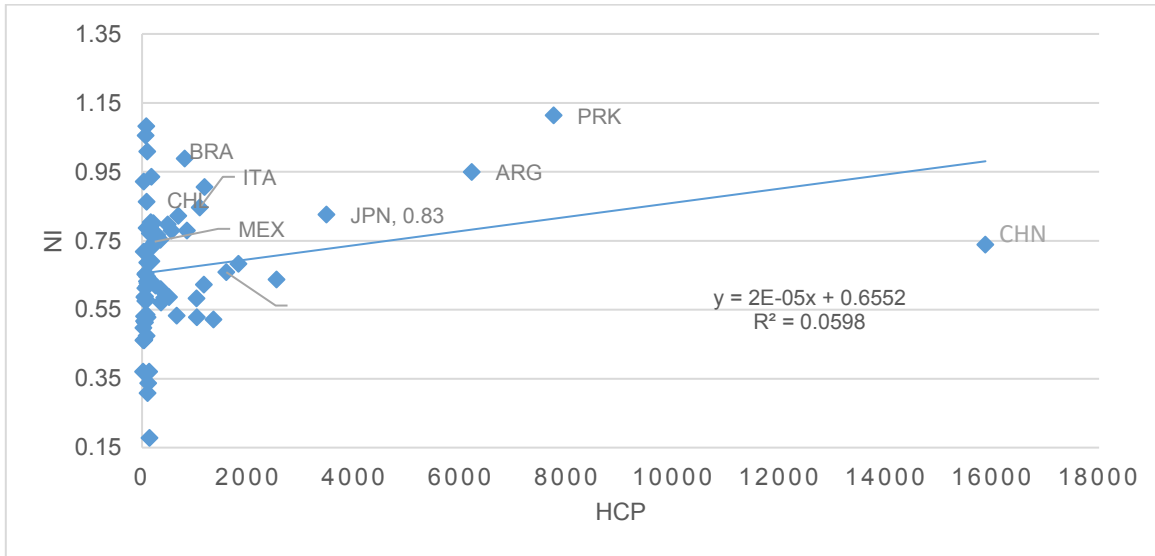


Figura 34. Relación entre excelencia Highly Cited Papers (HCP) e impacto normalizado durante el período 2005-2015, para países con similar índice de desarrollo científico que México

La figura 35, exhibe la relación biparamétrica entre el indicador número de Hot e Impacto Normalizado (NI), se denota que existe una baja correlación entre estos indicadores (coeficiente de correlación = 0.046). De acuerdo con estos dos indicadores los mejores valores fueron alcanzados por: Italia, Korea del Sur, México, Argentina, Brasil, Panamá y Cuba.

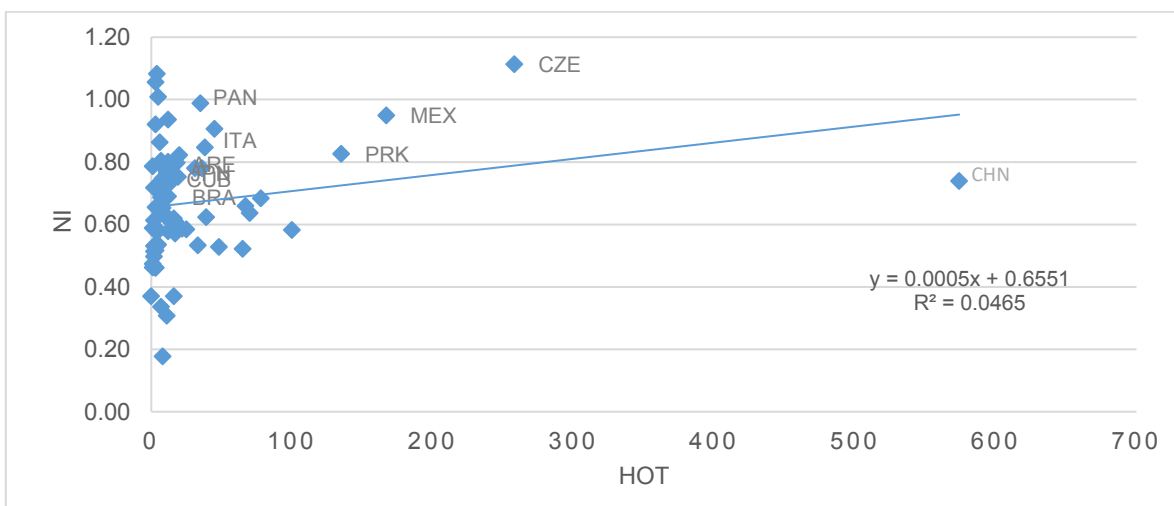


Figura 35. Hot papers e impacto normalizado para países con similar índice de desarrollo científico que México, durante el período 2005-2015

Para obtener un comportamiento más abarcador de la relación multivariada entre estos indicadores, los cuales presentan una baja correlación para el conjunto de 70 países. La siguiente sección aborda desde el enfoque de la inteligencia computacional, basada en técnicas de *mapas auto-organizados* o *SOM (Self-Organizing Map)*, una caracterización multidimensional y más amplia del grupo de países **con similar índice de desempeño científico** que México para estos cuatro indicadores (SDI, HCP, Hot, NI).

3.3. Aplicación de ViBlioSOM: Análisis multiparamétrico del perfil bibliométrico de México y otros países, considerando el índice de desarrollo científico, impacto y excelencia

3.3.1. Análisis usando indicadores dependientes de tamaño

A manera de demostración, en esta sección presentamos un análisis llevado a cabo antes de normalizar algunos de los indicadores, para demostrar el tipo de dificultad con la que uno se enfrenta en este escenario. En contraste, la siguiente sección nos muestra la importante diferencia que hace en los resultados el empleo de indicadores independientes de tamaño (normalizados).

La figura 36 muestra el análisis multifactorial sobre el rendimiento de 70 países, de acuerdo con los indicadores: Índice de desarrollo científico (SDI), *Highly Cited Papers* (HCP), *Hot paper* (Hot) e Impacto Normalizado (NI). En lugar de obtener cuatro clasificaciones diferentes para cada uno de estos indicadores, o analizar el conjunto de datos desde la perspectiva de dos indicadores por medio de diagramas de dispersión, aquí se presenta una comparación del desempeño de este grupo de países con el uso de indicadores de eficiencia, impacto y excelencia simultáneamente.

Cada país toma un punto en un espacio de cuatro dimensiones euclidiana, cuyas coordenadas vienen dadas por los valores de los cuatro indicadores. Nuestra técnica de visualización incluye cinco mapas (cartografías) para exponer la comparación y clasificación realizada por la Red neuronal. Estos mapas se dibujan sobre una cuadrícula hexagonal, en el que cada hexágono representa una neurona. En este caso se utilizó una rejilla de 40x20 neuronas y se realizaron 500 iteraciones.

Una vez que la red neural es entrenada por medio de un procedimiento de adaptación y sin supervisión, cada país se representa en el mapa con una proyección no lineal desde el espacio de 4D a la red neural. La proyección agrupa en un mismo cluster del mapa a los países con perfiles similares. Cada cluster de países representa un patrón bibliométrico de desempeño, esta tecnología nos permite localizar visualmente diferentes perfiles de desarrollo científico, impacto y excelencia, así como identificar los grupos de países que tienen patrones similares a México.

En el mapa de la figura 36 se observan cinco visualizaciones, la primera imagen denominada mapa de cluster o clustering, integra los cuatro indicadores calculados para los 70 países. Este mapa muestra 27 regiones que se distinguen por colores. Los países que quedaron dentro de una misma región tienen patrones de rendimiento bibliométricos similares, de acuerdo con los cuatro indicadores seleccionados. Los comportamientos atípicos son representados por un solo país en un cluster como el caso de: China (CHN), Italia (ITA), Sur Korea (PRK), República Checa (CZE), Hungría (HUN), Japón (JPN) y Polonia (POL) porque son países que presentan un alto desempeño con respecto al índice de desarrollo científico, impacto normalizado.

Los otros cuatro mapas que le siguen al clustering (figura 36) se denominan mapas de componentes. En este caso existe un mapa para cada uno de los cuatro indicadores, SDI, HCP, Hot y NI. Estos mapas son fundamentales para la interpretación de los resultados. La escala cromática permite observar la distribución de los valores en cada indicador. El rango de variación de cada componente (indicador) se asocia con una barra cromática, de tal manera que: el valor máximo del indicador corresponde al rojo más oscuro, el valor medio es asociado con el tono amarillo, y el valor mínimo se exhibe con el color verde claro. Los mapas de componentes permiten interpretar el significado de la distribución de datos en la red neuronal.

Los mapas de componentes de los indicadores Highly Cited Papers (HCP) y Hot paper (Hot), no ofrecen suficiente información para poder interpretar en el mapa de clustering la zona de excelencia. En los mapas de HCP y Hot solo destaca China

como el país con mayor número de HCP y Hot. México y el resto de los países quedaron en el área verde, lo cual se interpreta como que este grupo de países no tiene artículos de excelencia, y realmente si tienen artículos que se consideran de HCP y Hot.

Con la representación e interpretación de la figura 36, esta investigación corroboró que, es necesario aplicar la metodología ViBlioSOM nuevamente, para obtener una visión más completa de los perfiles de impacto, excelencia y producción del conjunto de 70 países. Se identificaron dos elementos que propiciaron la generación de un nuevo mapa.

Primero, el comportamiento atípico de China, este país tiene altos valores para los cuatro indicadores estudiados (SDI= 1.21; HCP= 15,863; Hot= 574; NI= 0.74). Segundo, los indicadores: número total de HCP y Hot por país, cumplen con la variante de ser indicadores de tamaño dependiente, es decir, se obtiene al contar el número absoluto de artículos HCP y Hot de un país, por tal razón, la comparación y construcción de mapas con la metodología ViBlioSOM, para los indicadores seleccionados, no muestra patrones de comportamiento aceptables

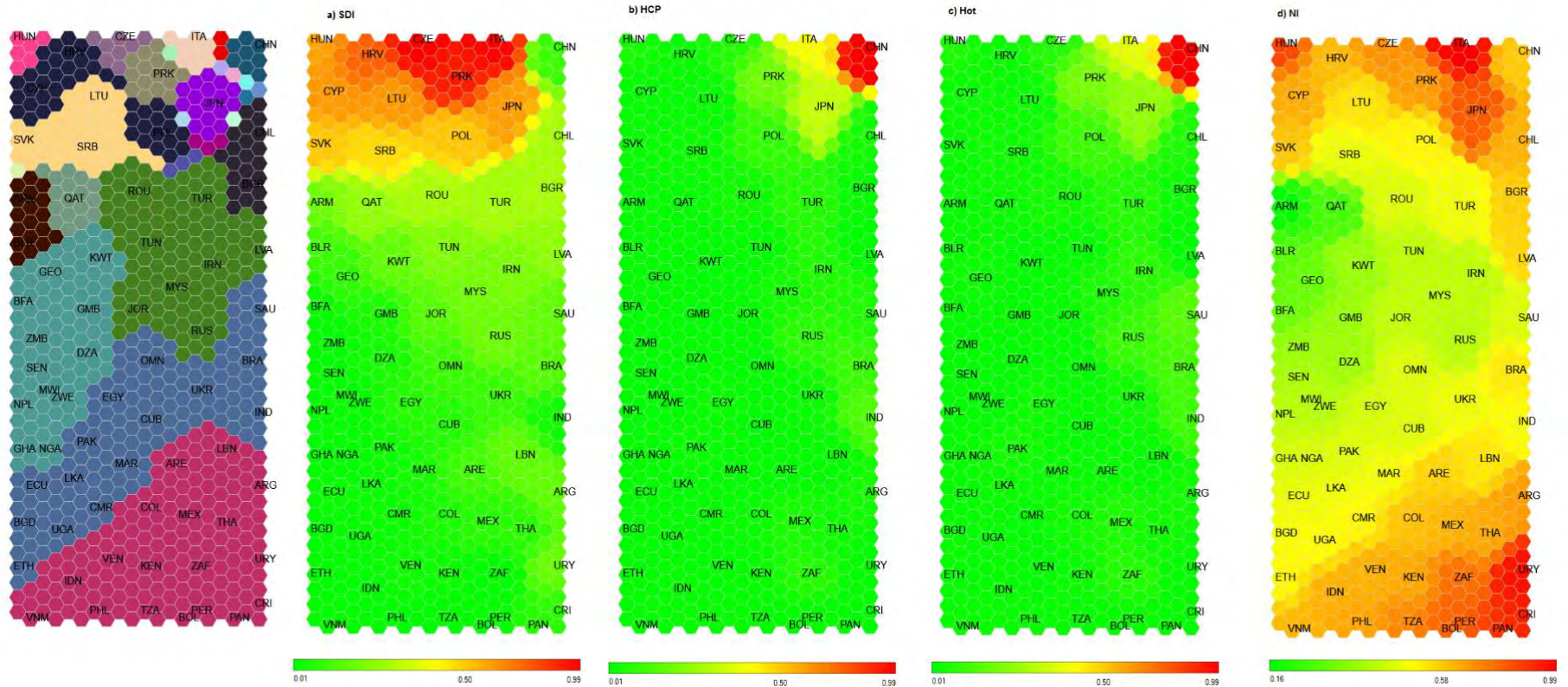


Figura 36. Mapa de Cluster identificado con la técnica Self Organized Maps para 70 países que tienen perfiles similares a México, HCP y Hot como indicadores de tamaño dependiente (Anexos 10, 11 y 12, donde cada mapa está separado)

3.3.2. Análisis usando indicadores independientes del tamaño

La comparación de la ciencia mexicana con otros países, y considerando indicadores de producción, impacto y excelencia, se puede realizar de manera justa al convertir los indicadores de excelencia, dependientes del tamaño a indicadores independientes del tamaño.

Los indicadores de tamaño dependiente (Size-dependent indicators), tienen como objetivo proporcionar una medida general del desempeño científico, ejemplo de ello es: el conteo de citas, el número de artículos altamente citados (highly cited paper), el índice H, hot paper; y los indicadores independientes del tamaño (size-independent indicators) como el promedio de citas por artículo, factor de impacto, tienen por objetivo proporcionar una medida del rendimiento promedio por artículo, revista, institución o país, según sea el objeto de estudio (Waltman, 2016).

El impacto de la investigación científica de un país puede ser comparado con otros países utilizando el impacto promedio por artículo y la normalización por campo de conocimiento, como corrección de los diferentes niveles de citación; conocido como impacto normalizado. Esto tiene la ventaja de generar indicadores que se clasifican como “independientes del tamaño” (size-independent indicator) (Abramo & DAngelo, 2016a) de manera que las comparaciones se vuelven justas. También, esta información puede ser presentada en relación a los recuentos totales de la población (como se presenta en esta tesis), el Producto Interno Bruto (PIB) o número de investigadores activos por cada país.

Para nuestro caso se determinó dividir el número de HCP y Hot entre el número total de documentos publicados en WoS durante el período. De esta manera, los indicadores HCP y Hot se convierten en indicadores independientes del tamaño. En adelante referiremos a ambos indicadores por HCP y Hot en el entendido de que han sido normalizados.

La figura 37 muestra el análisis multifactorial sobre el rendimiento de 69 países sin China, de acuerdo con los **indicadores de tamaño independiente**: Índice de desarrollo científico (SDI), Highly Cited Papers (HCP), Hot paper (Hot) e Impacto Normalizado (NI). Cada país toma un punto en un espacio euclidiano de

cuatro dimensiones, cuyas coordenadas vienen dadas por los valores de los cuatro indicadores. Nuestra técnica de visualización incluye cinco mapas (cartografías) para exponer la comparación y clasificación realizada por la Red neuronal. Estos mapas se dibujan sobre una cuadrícula hexagonal, en el que cada hexágono representa una neurona. En este caso se utilizó una rejilla de 40x20 neuronas y se realizaron 500 iteraciones para entrenar la red neuronal.

En el mapa de la figura 37 se observan cinco visualizaciones, la primera llamada mapa de cluster o clustering, el cual integra los cuatro indicadores calculados para los 69 países. Este mapa muestra 27 regiones que se distinguen por colores. Los países que quedaron dentro de una misma región ofrecen patrones de rendimiento bibliométricos similares, de acuerdo con los cuatro indicadores seleccionados.

Los comportamientos atípicos son representados por un solo país en un cluster como el caso de: Italia (ITA), Perú (PER), Panamá (PAN), Uruguay (URY), y Zimbabwe (ZWE) porque son países que cumplen con un alto desempeño con respecto a los indicadores estudiados y el grupo de 70 países. Señalamos el caso de Italia, país que resalta con respecto al resto, porque presenta mayor desempeño de acuerdo al índice de desarrollo científico, excelencia e impacto, y por eso ocupa un solo cluster.

Países como Sri Lanka (LKA), Georgia (GEO) y Gambia (GBM), ocupan una misma región en el mapa de clustering (región de color violeta oscuro), pero, aunque presentan valores de índice de desarrollo científico e impacto normalizado bajo, sus investigaciones se consideran de excelencia; esto implica que aunque estos países no tienen un alto índice de desarrollo científico, en ellos existen grupos de investigación que producen artículos que son altamente citados, es decir, investigaciones de excelencia.

Sur Korea (PRK), República Checa (CZE), Hungría (HUN), Japón (JPN) y Polonia (POL), muestran perfiles semejantes en cuanto al índice de desarrollo científico e impacto normalizado; pero tiene valores bajos de HCP y Hot.

En la región de color gris México (MEX), China (CHN), Argentina (ARG), Chile (CHL), Bulgaria (BGR), Turquía (TUR) tienen perfiles de producción, impacto

y excelencia similares, es decir, el índice de desarrollo para estos países es medio bajo al igual que los valores de HCP y Hot; pero su investigación tiene un alto impacto normalizado. En esta región, México como país queda ubicado por encima de Brasil, pues presenta mejores valores de eficiencia, impacto y excelencia.

En el mapa de clustering se denota que la parte inferior derecha (azul, violeta claro y rosado oscuro) del mapa puede considerarse como una zona muy eficiente, la parte inferior izquierda (color rojo, café oscuro, violeta y azul) comprende los países que sus investigaciones tienen una alta excelencia y la zona media la conforman países con alto impacto normalizado, es decir, una zona que se puede clasificar como zona de efectividad.

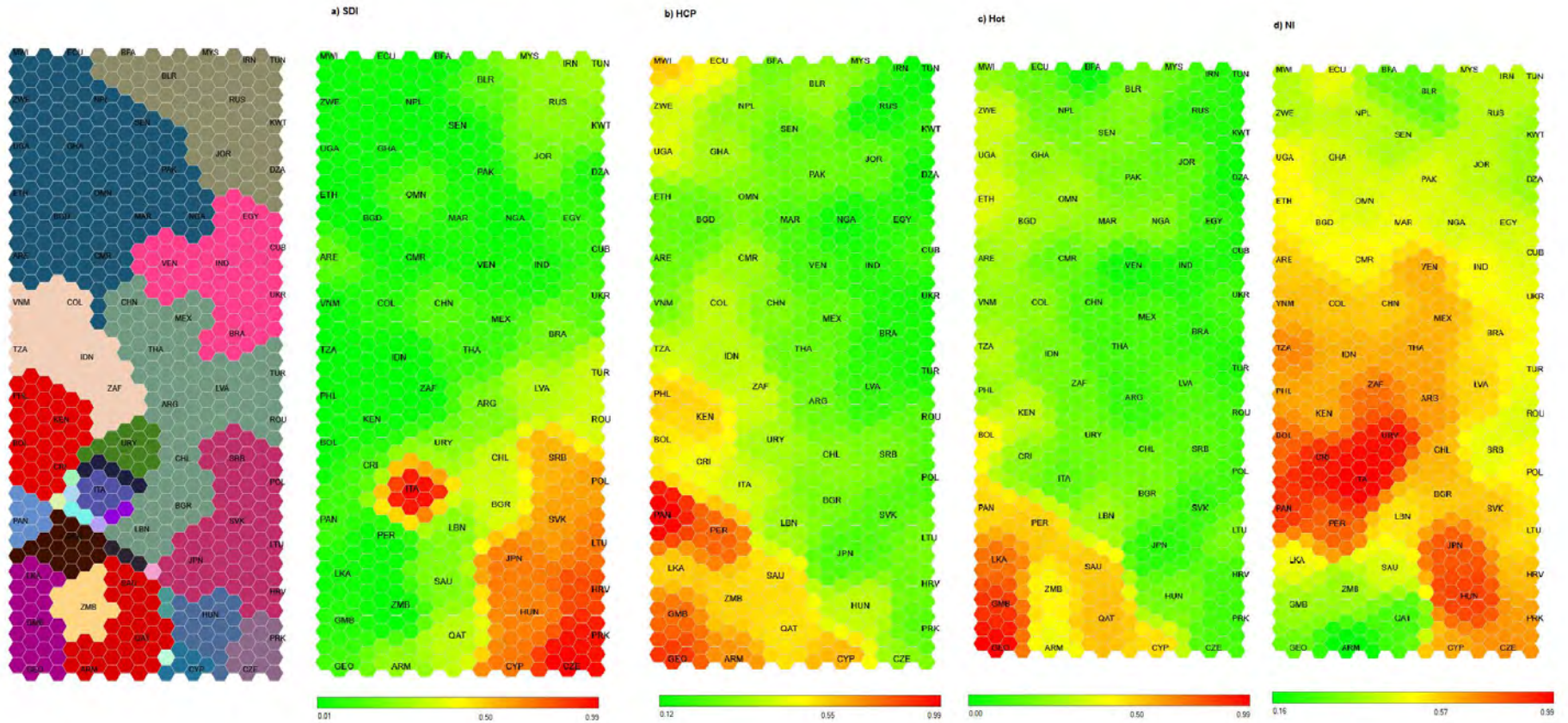


Figura 37. Mapa de Clusters identificados con la técnica Self Organized Maps para 70 países que tienen perfiles similares a México, HCP y Hot como indicadores de tamaño independiente (Anexo 13, 14 y 15, donde cada mapa está separado)

Consideraciones finales

El análisis de las principales revistas mexicanas nos permitió identificar y observar gráficamente las diferencias que surgen al evaluarlas considerando los cuatro indicadores de última generación, seleccionados en nuestra investigación (IF, EFS y SJR). De este análisis se concluye que, aunque existe un alto grado de coincidencias, existen fuertes diferencias al evaluar a algunas revistas, asunto que debe tomarse en cuenta cuando se lleve a cabo una práctica de evaluación. La diferencia más marcada surgió al considerar al Source Normalized Impact per Paper. Varias revistas mexicanas quedan bien calificadas respecto al SNIP, aunque obtienen muy baja puntuación respecto a los otros tres indicadores.

La metodología ViBlioSOM fue muy útil para obtener estos resultados y los que se obtuvieron para el análisis multifactorial de la producción científica de los países que tienen un grado de desarrollo científico comparable con el de México. Esta metodología está implementada en el sistema de software llamado LabSOM, desarrollado por el Laboratorio de Dinámica no Lineal de la Facultad de Ciencias de la Universidad Autónoma Nacional Universidad de México (UNAM) y la Compañía "Tecnologías Inteligentes y Modelación" de Sistemas". Con esta herramienta, el análisis, la visualización e interpretación de los datos se realizó de manera automatizada. Así el conocimiento obtenido quedó representado visualmente en los mapas realizados con la tecnología LabSOM.

El análisis comparativo de la evolución de la producción científica mexicana en el Web of Science y Scopus durante el período 2005-2015, revela que ambas producciones progresan de manera ascendente y lo mismo ocurre a la población de investigadores nacionales. La tasa de crecimiento de estos indicadores es muy semejante, pero el número de investigadores, SNI, supera a Ndoc Scopus y los valores de este indicador siempre superan a Ndoc WoS. Esto significa que la productividad nacional de México es de menos de un artículo por año, por investigador nacional, de hecho, esta fue en el año 2015 de 0.73 en Scopus y de 0.59 en WoS.

En todo este trabajo hemos hecho nuestro mejor esfuerzo por crear y utilizar indicadores independientes de tamaño. La consideración del indicador SDI

(Scientific Development Index) nos permitió identificar un grupo de 70 países que de manera más justa pueden ser comparados con México. Este indicador normaliza el volumen total de la producción de los países dividiéndola por el número de habitantes. Considerando indicadores que miden la eficiencia, impacto y excelencia y utilizando otra vez la metodología ViBlioSOM y el software LabSOM, pudimos identificar automáticamente los grupos de países (clusters) que tienen similar perfil bibliométrico, así como aquellos que tienen perfiles muy peculiares (outliers) dentro del conjunto de países considerados. Este estudio rindió interesantes conclusiones que no podrían ser anticipadas, varios países con perfiles peculiares, entre los que destacan Italia, Perú y Panamá; un desempeño medio-alto para México en cuanto al impacto normalizado, aunque relativamente bajo en los indicadores de excelencia. También llama la atención que algunos países que tienen dentro de este grupo de 70 países los más bajos índices de desarrollo científico, tienen sin embargo alta grado de excelencia (alta puntuación en top papers): Perú panamá, Gambia Georgia, Sri Lanka, Sambia, Armenia, Zimbaue y Qatar.

Finalmente comentamos que se identificaron aquellos campos de investigación en los que México tiene mayor producción e impacto durante el período de estudio: Física, Medicina Clínica, Ciencias de los animales y plantas, Química, Biología-Bioquímica y Medioambiente/Ecología. Resultó que son también estos campos de investigación aquellos que destacan por el número de artículos altamente citados.

Referencias

- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2008). Assessment of sectoral aggregation distortion in research productivity measurements. *Research Evaluation*, 17(2), 111-121.
- Abramo, G., D'Angelo, C. A., & Rosati, F. (2013). The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences. [Article]. *Journal of Informetrics*, 7(1), 198-208.
- Abramo, G., & D'Angelo, C. A. (2016a). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics*, 10(2), 646-651.
- Abramo, G., & D'Angelo, C. A. (2016b). A farewell to the MNCS and like size-independent indicators: Rejoinder. *Journal of Informetrics*, 10(2), 679-683.
- Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3).
- Aksnes, D. W., & Rip, A. (2009). Researchers perceptions of citations. *Research Policy*, 38, 895–905.
- Andras, P. (2011). Research: metrics, quality, and management implications. [Article]. *Research Evaluation*, 20(2), 90-106.
- Antonakis, J., Bastardoz, N., Liu, Y., & Schriesheim, C. A. (2014). What makes articles highly cited? *The Leadership Quarterly*, 25, 152–179.
- Arencibia-Jorge, R., Atenógenes Villaseñor, E., Lozano-Díaz, I. A., & Carrillo Calvet, H. (2016). Elsevier's journal metrics for the identification of a mainstream journals core. *LIBRES*, 26(1).
- Arencibia Jorge, R. (2010). *Visibilidad Internacional de la Ciencia y Educación Superior Cubanas: desafíos del estudio de la producción científica*. Univeridad de Granada, Granada.
- Arencibia Jorge, R., & Moya-Anegón, F. (2008). La evaluación de la investigación científica: una aproximación teórica desde la cienciometría. *Acimed*, 17(4).
- Barker, K. (2007). The UK Research Assessment Exercise: the evolution of a national research evaluation system. [Article; Proceedings Paper]. *Research Evaluation*, 16(1), 3-12.
- Baudoin, L., Peltier, C., Graillet-Gak, C., & Haeffner-Cavaillon, N. (2004). Zoom on the excellence of the biomedical research in France. *M S-Medecine Sciences*, 20(12), 1149-1155.
- Bergstrom, C. T. (2007). Eigenfactor. Measuring the value and prestige of scholarly journals *College & Research Libraries News*, 68(5), 314-316
- Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The Eigenfactor (TM) Metrics. *Journal of Neuroscience*, 28(45), 11433-11434.
- Bernstein, A., Provost, F., & Hill, S. (2005). Toward intelligent assistance for a data mining process: an ontology-based approach for cost-sensitive classification. *IEEE Trans Knowl Data Eng*, 17.
- Bloch, C., Schneider, J. W., & Sinkjaer, T. (2016). Size, Accumulation and Performance for Research Grants: Examining the Role of Size for Centres of Excellence. *Plos One*, 11(2).

- Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing Knowledge Domains. *Annual Review of Information Science & Technology*, 37.
- Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation*, 23(2).
- Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. [Article]. *Research Evaluation*, 23(2), 166-173.
- Bornmann, L., & Hans-Dieter, D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45-80.
- Bornmann, L., & Haunschild, R. (2016). Efficiency of research performance and the glass researcher. *Journal of Informetrics*, 10(2), 652-654.
- Bornmann, L., & Leydesdorff, L. (2016). Which are the best performing regions in information science in terms of highly cited papers? Some improvements of our previous mapping approaches. *Journal of Informetrics*, 6(2), 336-345.
- Bornmann, L., Leydesdorff, L., Walch-Solimena, C., & Ettl, C. (2011). Mapping excellence in the geography of science: An approach based on Scopus data. *Journal of Informetrics*, 5(4), 537-546.
- Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014). Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers A visualisation of results from multi-level models. *Online Information Review*, 38(1), 43-58.
- Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2016). Excellence networks in science: A Web-based application based on Bayesian multilevel logistic regression (BMLR) for the identification of institutions collaborating successfully. *Journal of Informetrics*, 10(1), 312-327.
- Boyack, K. W., Wylie, B. N., & Davidson, G. S. (2002). Information Visualization, Human-Computer Interaction, and Cognitive Psychology: Domain Visualizations. In K. Börner & C. Chen (Eds.), *Visual Interfaces to Digital Libraries* (pp. 145-158). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bravo Vinaja, A. (2005). *Análisis bibliométrico de la producción científica de México en ciencias agrícolas a través de las bases de datos internacionales: agrícola, agris, cab abstracts, science citation index, social science citation index y tropag & rural, en el período 1983-2002* Universidad Carlos III, Madrid.
- Callon, M., J.P. , & Courtial, H. (1995). *Cienciometría*. Gijón: ediciones TREA.
- Card, S. K., MacKinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. San Francisco: Morgan Kaufmann.
- Castillo-Pérez, J. J., & Muñoz-Valera, L. (2014). Producción científica mexicana sobre influenza, 2000-2012. *Salud Pública de México*, 56(5), 424-425.
- Cawkell, T., & Garfield, E. (2001). Institute for Scientific Information. In E. H. Fredriksson (Ed.), *A Century of Science Publishing* (pp. 149-160).
- Cole, J. R. (2000). A short history of the use of citations as a measure of the impact of scientific and scholarly work. In B. Cronin & H. B. Atkins (Eds.), *The Web of Knowledge. A Festschrift in Honor of Eugene Garfield, Information Today*: Medford, NJ.

- Cole, S. (2000). The role of journals in the growth of scientific knowledge. In B. Cronin & H. B. Atkins (Eds.), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Medford: ASIS.
- Cronin, B. (1984). *The Citation Process. The Role and Significance of Citations in Scientific Communication*. Oxford: Taylor Graham.
- Chen, C. (2002). Information visualization. *Information Visualization*, 1, 1-4.
- Chen, C. (2006). *Information visualization: Beyond the horizon*. London Springer.
- Chen, C. (2010). Information visualization. *Wiley Interdisciplinary Review: Computational Statistics*, 2(4), 387-403.
- Chinchilla-Rodríguez, Z. (2004). *Análisis del dominio científico Español:1995-2002 Web of Science*. Universidad de Granada, Granada, España.
- Chinchilla-Rodríguez, Z., Zacca-Gonzalez, G., Vargas-Quesada, B., & Moya-Anegón, F. (2015). Latin American scientific output in Public Health: combined analysis using bibliometric, socioeconomic and health indicators. [Article]. *Scientometrics*, 102(1), 609-628.
- Díaz-Pérez, M., Giráldez-Reyes, R., Carrillo-Calvet, H. A., Armas Peña, D., Rodríguez-Font, R. J., & Villaseñor-García, E. A. (2014). Tecnologías constituidas, innovaciones en proceso y tecnologías introducidas en el mercado internacional: caso de estudio. *TransInformação*, 23(3).
- Dürsteler, J. C. (2000). *Visualización de información*. Barcelona: Gestión Elsevier (2016). Journal Metrics, from <https://www.journalmetrics.com/contact-us.php>
- Erbacher, R. F. (2007). *Exemplifying the inter-disciplinary nature of visualization research*. Paper presented at the 1th International Conference Information Visualization, London.
- Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. *Faseb Journal*, 22(8), 2623-2628.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine* 17(3).
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27-34.
- Fayyad, U., & Stolorz, P. (1997). Data mining and KDD: Promise and challenges. *Future Generation Computer Systems*, 13, 99-115.
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). Advances in knowledge discovery and data mining.
- Fu, T. (2011). A review on time series data mining. *Eng Appl Artif Intell*, 24.
- Gamarra, C., Guerrero, J. M., & Montero, E. (2016). A knowledge discovery in databases approach for industrial microgrid planning. *Renewable and Sustainable Energy Reviews*, 60, 615-630.
- Garfield, E. (2006). The history and meaning of the journal impact factor. *Jama-Journal of the American Medical Association*, 295(1), 90-93.
- Garfield, E., & Sher, I. H. (1963). New factors in evaluation of scientific literature through citation indexing. *American Documentation*, 14(3), 195-201.
- Glanzel, W., Thijs, B., & Debackere, K. (2016). Productivity, performance, efficiency, impact "What do we measure anyway?: Some comments on the paper, A

- farewell to the MNCS and like size-independent indicators by Abramo and D'Angelo. *Journal of Informetrics*, 10(2), 658-660.
- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2009). The SJR indicator: A new indicator of journals' scientific prestige. *Journal of Informetrics*, 4(3).
- González de Dios, J., Moya, M., & Mateos Hernández, M. A. (1997). Indicadores bibliométricos: Características y limitaciones en el análisis de la actividad científica *Anales Español Pediatría*, 47(3).
- Gorbea Portal, S. (2013). Tendencias transdisciplinarias en los estudios métricos de la información y su relación con la gestión de la información y del conocimiento. *Perspectivas em Gestão & Conhecimento*, 3(1), 13-27.
- Gullo, F. (2014). From Patterns in Data to Knowledge Discovery: What Data Mining Can Do. *Physics Procedia*, 62, 18-22.
- Guzmán-Sánchez, M. V. (2009). *ViBlioSOM: Metodología para la visualización de información métrica con mapas auto-organizados*. Universidad de La Habana, La Habana, Cuba.
- Guzmán Sánchez, M. V. (2009). *ViBlioSOM: Metodología para la Visualización de Información métrica con Mapas Auto-organizados*. Universidad de La Habana, La Habana.
- Han, J., & Kamber, M. (2000). *Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems)*: Morgan Kaufmann.
- Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.
- Hernández, Y. H., Dray, M. K., & Russell, J. M. (2013). Methodological approaches to identify and describe Mexican chemistry research in bibliographic databases. [Article]. *Investigacion Bibliotecologica*, 27(59), 35-66.
- Hjørland, B., & Albrechtsen, H. (1995). Towards a new horizon in information science: domain analysis. *Journal of the American Society for Information Science and Technology*, 46(6), 400-425.
- Holmes, J. H. (2013). Knowledge Discovery in Biomedical Data: Theory and Methods *Methods in Biomedical Informatics: A Pragmatic Approach* (pp. 170-240). Philadelphia, USA: University of Pennsylvania
- Holzinger, A., Dehmer, M., & Jurisica, I. (2014). Knowledge Discovery and interactive Data Mining in Bioinformatics - State-of-the-Art, future challenges and research directions. *BMC Bioinformatics*, 15(6).
- Jacso, P. (2010). Eigenfactor and article influence scores in the Journal Citation Reports. *Online Information Review*, 34(2), 339-348.
- Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31, 7-15.
- Kishida, K. (2003). Techniques of document clustering: A review. [Review]. *Library and Information Science*(49), 33-75.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.
- Kohonen, T. (1990). The Self-organizing Map. *Proceedings of the IEEE*, 78(9).
- Kohonen, T., Kaski, S., & Lappalainen, H. (1997). Self-organized formation of various invariant-feature filters in the adaptive-subspace SOM. *Neural Computation*, 9, 1321-1344.

- Kostoff, R. N., del Rio, J. A., Cortes, H. D., Smith, C., Smith, A., Wagner, C., et al. (2005). The structure and infrastructure of Mexico's science and technology. [Article]. *Technological Forecasting and Social Change*, 72(7), 798-814.
- Kostoff, R. N., del Rio, J. A., Cortes, H. D., Smith, C., Smith, A., Wagner, C., et al. (2007). Clustering methodologies for identifying country core competencies. [Article]. *Journal of Information Science*, 33(1), 21-40.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling (quantitative applications in the Social Sciences)*. Newbury Park: SAGE.
- Kuncheva, L. (2004). *Combining Pattern Classifiers: Methods and Algorithms*: Wiley-Interscience.
- Kurgan, L., & Musilek, P. (2006). A survey of knowledge discovery and data mining process models. *Knowl Eng Rev*, 21(1).
- Laclette, J. P., & Zúñiga-Bello, P. (2012). *Ranking Nacional de Ciencia, Tecnología e Innovación*. México, Distrito Federal: Foro Consultivo Científico y Tecnológico, AC.
- Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007). Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, 1(1), 2.
- Leydesdorff, L., & Opthof, T. (2010). Scopus's Source Normalized Impact per Paper (SNIP) Versus a Journal Impact Factor Based on Fractional Counting of Citations. *Journal of the American Society for Information Science and Technology*, 61(11), 2365-2369.
- Liao, S.-H., Chu, P.-H., & Hsiao, P.-Y. (2012). Data mining techniques and applications: A decade review from 2000 to 2011. *Expert Systems with Applications*, 39, 11303-11311.
- Licea de Arenas, J. (1993). Indicadores de la actividad científica *Selección de lecturas de Estudios Métricos de la Información*. La Habana: Félix Varela.
- Macias-Chapula, C. A. (2005). Toward a model of communications in public health in Latin America and the Caribbean. [Article]. *Revista Panamericana De Salud Publica-Pan American Journal of Public Health*, 18(6), 427-438.
- Marcus, A. (2015). Visualizing the Future of Information Visualization *HCI and User-Experience Design: Fast-Forward to the Past, Present, and Future* (pp. 199-204). London: Springer London.
- Martinez, M., Herrera, M., Contreras, E., Ruiz, A., & Herrera-Viedma, E. (2015). Characterizing highly cited papers in Social Work through H-Classics. [Article]. *Scientometrics*, 102(2), 1713-1729.
- Melkers, J. (1993). Bibliometrics as a Tool for Analysis of R&D Impacts. In B. Bozeman & J. Melkers (Eds.), *Evaluating R&D Impacts: Methods and Practice* (pp. 43-61). Boston, MA: Springer US.
- Michan, L., Russell, J. M., Pereyra, A. S., Cruset, A. L., & Beltran, C. L. (2008). Analysis of the state of systematics in latin america. [Article]. *Interciencia*, 33(10), 754-761.
- Miyairi, N., & Chang, H.-W. (2012). Bibliometric characteristics of highly cited papers from Taiwan, 2000-2009. *Scientometrics*, 92(1), 197-205.
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265-277.

- Moed, H. F. (2011). *SNIP and Beyond*. Paper presented at the *Proceedings of Issi 2011: The 13th Conference of the International Society for Scientometrics and Informetrics*.
- Moed, H. F. (2011). *SNIP and Beyond*. In E. Noyons, P. Ngulube & J. Leta (Eds.), *Proceedings of Issi 2011: The 13th Conference of the International Society for Scientometrics and Informetrics, Vols 1 and 2* (pp. 584-594).
- Moya-Anegón, F., Herrero-Solana, V., & Jiménez-Contreras, E. (2006). A connectionist and multivariate approach to science maps: SOM, clustering and MDS applied to Library & information science research. *Journal of Information Science*, 32(1), 61-75.
- Nieminen, P., Pölonen, I., & Sipola, T. (2013). Research literature clustering using diffusion maps. *Journal of Informetrics*, 7, 874 - 886.
- OECD (2016). *Main Science and Technology Indicators, Volume 2016 Issue 1*: OECD Publishing.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). *The pagerank citation ranking: bringing order to the web. Technical report*. Stanford Digital Library Technologies Project.
- Pajic, D. (2015). On the stability of citation-based journal rankings. *Journal of Informetrics*, 9, 990–1006.
- Parameswaran, A., & Sebastian, R. (2006). The Value of South and Southeast Asian Studies Journal Rankings. *Serials Review*, 32(3), 154-163.
- Peffer, K., & Tang, Y. (2003). Identifying and evaluating the universe of outlets for information systems research: Ranking the journals. *Journal of Information Technology Theory and Application*, 5, 63-84.
- Pérez-Arreortúa, N., Díaz-Pérez, M., Giráldez-Reyes, R., & Carrillo-Calvet, H. A. (2014). Análisis de contenido del dominio tecnológico vegetable oil combustion. *TransInformação*, 26(3).
- Perez-Santos, J. L. M., & Anaya-Ruiz, M. (2013). Mexican Breast Cancer Research Output, 2003-2012. [Article]. *Asian Pacific Journal of Cancer Prevention*, 14(10), 5921-5923.
- PhridviRaj, M. S. B., & GuruRao, C. V. (2014). Data mining past, present and future a typical survey on data streams. *Procedia Technology*, 12, 255 - 263.
- Piedras, Y. (2009). *La producción científica sobre comunicación: análisis desde una perspectiva de dominio en el web de la ciencia. período 2000-2007*. Universidad de La Habana, La Habana.
- Polanco, X., Francois, C., & Lamirel, J. C. (2001). Using artificial neural networks for mapping of science and technology: A multi-self-organizing-maps approach. *Scientometrics*, 51(1), 267-292.
- Polonsky, M. J. (2004). Journal Rankings: Does One Size Fit All? *Australasian Marketing Journal*, 12(2).
- Purchase, H. C., Andrienko, N., Jankun-Kelly, T. J., & Ward, M. (2008). Theoretical Foundations of Information Visualization. In A. Kerren, J. T. Stasko, J.-D. Fekete & C. North (Eds.), *Information Visualization: Human-Centered Issues and Perspectives* (pp. 46-64). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Pyle, D. (1999). *Data Preparation for Data Mining (The Morgan Kaufmann Series in Data Management Systems)*: {Morgan Kaufmann}.

- Reuters, T. (2016). SCI Timeline, from <http://wokinfo.com/sci-anniversary.html#timeline>
- Reuters, T. (2016). Web of Science
- Rodriguez-Navarro, A. (2011). Measuring research excellence Number of Nobel Prize achievements versus conventional bibliometric indicators. *Journal of Documentation*, 67(4), 582-600.
- Rodríguez Sánchez, Y. (2012). *Metodología Bibliométrica para la Evaluación de la Actividad Científica*. Instituto Superior Politécnico José Antonio Echevarría, La Habana.
- Romero, A. H., Garcia, A., & Kiwi, M. (2009). Evaluation of the scientific impact, productivity and biological age based upon the h-index in three Latin American countries: the materials science case. [Article]. *Annalen Der Physik*, 18(4), 198-205.
- Sancho, R. (1990). Indicadores bibliométricos utilizados en la evaluación de la Ciencia y la Tecnología. *Revista española de Documentación Científica*, 13(842 - 865).
- Saracevic, T. (1995). Interdisciplinary nature of information science. *Ciência da Informação*, 24(1).
- SCImago Journal & Country Rank (2016). from <http://www.scimagojr.com/>
- Schneider, J. W. (2006). Concept symbols revisited: Naming clusters by parsing and filtering of noun phrases from citation contexts of concept symbols. *Scientometrics*, 68(3), 573-593.
- Schoepfel, J., & Prost, H. (2009). Comparison of SCImago Journal Rank Indicator (SJR) with JCR journal impact factor (IF) for French journals. *Psychologie Francaise*, 54(4), 287-305.
- Schvaneveldt, R. (1990). *Pathfinder Associative Networks: Studies in Knowledge Organization*. Norwood NJ Ablex.
- Seiler, C., & Wohlrabe, K. (2014). How robust are journal rankings based on the impact factor? Evidence from the economic sciences. *Journal of Informetrics*, 8(4), 904-911.
- Senger, S., Leach, A. R., Ralph, A. W., & David, C. S. (2008). Chapter 11 - SAR Knowledge Bases in Drug Discovery *Annual Reports in Computational Chemistry* (Vol. Volume 4, pp. 203-216): Elsevier.
- Serenko, A., & Dohan, M. (2011). Comparing the expert survey and citation impact journal ranking methods: Example from the field of Artificial Intelligence. *Journal of Informetrics*, 5(4), 629-648.
- Shu-Hsien, L., Pei-Hui, C., & Pei-Yuan, H. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications*, 39(11303-11311).
- Siguenza-Guzman, L., Saquicela, V., Avila-Ordóñez, E., Vandewalle, J., & Cattrysse, D. (2015). Literature Review of Data Mining Applications in Academic Libraries. *The Journal of Academic Librarianship*, 45, 499 –510.
- Sirsat, S. R., Chavan, V., & Deshpande, S. P. (2014). Mining knowledge from text repositories using information extraction: A review. [Review]. *Sadhana-Academy Proceedings in Engineering Sciences*, 39(1), 53-62.

- Sotolongo-Aguilar, G., Guzmán-Sánchez, M. V., & Carrillo, H. (2002). ViBlioSOM: Visualización de información bibliométrica mediante el mapeo autoorganizado. *Revista Española de Documentación Científica* 25(4).
- Suarez Balseiro, C., & Maura Sardó, M. (2005). *Bibliometría evaluativa. Seguimiento de la actividad científica y tecnológica (SACT)*. Paper presented at the Foro sobre Estudios Métricos de la Información.
- Tijssen, R. J. W., Visser, M. S., & van Leeuwen, T. N. (2001). *Searching for scientific excellence: Scientometric measurements and citation analyses of national research systems*.
- Tseng, Y. H., & Tsay, M. Y. (2013). Journal clustering of library and information science for subfield delineation using the bibliometric analysis toolkit: CATAR. [Article]. *Scientometrics*, 95(2), 503-528.
- UNESCO (2015). *Informe de la UNESCO sobre la ciencia hacia 2030*.
- Van Raan, A. F. J. (2004). Measuring science. Capita selecta of current main issues. In H. F. Moed, W. Glanzel & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht: Kluwer Academic Publishers.
- Vanclay, J. K. (2011). An evaluation of the Australian Research Council's journal ranking. [Article]. *Journal of Informetrics*, 5(2), 265-274.
- Villaseñor-García, E.-A., Arencibia Jorge, R., & Carrillo-Calvet, H. (2016). Multiparametric characterization of scientometric performance profiles assisted by neural networks: a study of Mexican higher education institutions. *Scientometrics*.
- Villaseñor, E. (2004). *Análisis inteligente de datos con redes neuronales artificiales*. Universidad Nacional Autónoma de México, México.
- Villaseñor, E., Carrillo, H., de la Escalera, N., & Millán, V. (2008). The use of Weighted Metric SOM Algorithm as a Visualization Tool for Demographic Studies. *Advances in Artificial Intelligence: Algorithms and Applications*, 13.
- Villaseñor, E., Martínez de la Escalera, N., Carrillo Calvet, H., & Cruz, N. (2008). Sistemas dinámicos y visualización informétrica: una aplicación de la red neuronal SOM. *Aportaciones Matemáticas* 38.
- Villaseñor García, E. A. (2016). *Aportaciones al análisis y visualización informétrica con la red neuronal SOM*. Universidad Nacional Autónoma de México, Distrito Federal, México.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365-391.
- Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., van Eck, N. J., et al. (2012). The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, 63(12), 2419-2432.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., & Visser, M. S. (2013a). Some modifications to the SNIP journal impact indicator. *Journal of Informetrics*, 7(2), 272-285.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., & Visser, M. S. (2013b). Some modifications to the SNIP journal impact indicator. *Journal of Informetrics*, 7(2), 272-285.

- Wanderley, G. M. P., Tacla, C. A., Barthes, J.-P. A., & Paraiso, E. C. (2015). Knowledge discovery in task-oriented dialogue. *Expert Systems with Applications*, 42, 6807-6818.
- Web of Science (2016). Thomson Reuters.
- Wesley-Smith, I., Bergstrom, C. T., & West, J. D. (2016). *Static ranking of scholarly papers using article-level Eigenfactor (ALEF)*. Paper presented at the The 9th ACM International Conference on Web Search and Data Mining, San Francisco, CA, USA.
- West, J. D., Bergstrom, T. C., & Bergstrom, C. T. (2010). The Eigenfactor Metrics (TM): A Network Approach to Assessing Scholarly Journals. *College & Research Libraries*, 71(3), 236-244.
- Witten, I., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)*: Morgan Kaufmann.
- Xu, H., Guo, T., Yue, Z., Ru, L., & Fang, S. (2016). Interdisciplinary topics of information science: a study based on the terms interdisciplinarity index series. [journal article]. *Scientometrics*, 106(2), 583-601.
- Zhu, B., & Chen, H. (2005). Information Visualization. *Annual Review of Information Science and Technology*, 39, 139-177.

ANEXOS

Anexo 1. Número total de artículos Top, Highly Cited Papers y Hot papers por países, en WoS, durante 2005-2015

Países	Top 2005-2009	Top 2006-2010	Top 2007-2011	Top2008-2012	Top2009-2013	Top2010-2014	Top 2011-2015
BRA	570	632	724	805	919	1029	1022
MEX	252	293	352	408	476	522	517
ARG	202	250	273	316	364	401	381
CHL	147	168	198	228	285	345	349
COL	71	89	110	140	175	210	221
VEN	32	41	43	46	47	50	42
CUB	11	20	26	27	38	38	33
FIN	619	707	805	878	992	1,132	1096
GRC	400	455	512	617	710	797	782
PRT	368	426	492	558	641	752	740
CZE	284	347	402	468	556	635	615
Países	HCP 2005-2009	HCP 2006-2010	HCP 2007-2011	HCP2008-2012	HCP2009-2013	HCP2010-2014	HCP 2011-2015
BRA	570	635	724	805	918	1027	1017
MEX	252	293	352	408	476	520	515
ARG	202	250	273	316	364	400	380
CHL	147	168	198	228	285	345	349
COL	71	89	110	140	175	209	220
VEN	32	41	43	46	47	50	42
CUB	11	20	26	27	38	38	33

FIN	619	707	805	878	990	1,126	1087
NZL	416	465	536	599	676	698	691
GRC	400	455	512	617	707	792	775
PRT	368	426	492	558	639	749	736
CZE	284	347	402	468	556	635	615
Países	Hot 2005-2009	Hot 2006-2010	Hot 2007-2011	Hot2008-2012	Hot2009-2013	Hot2010-2014	Hot 2011-2015
BRA	0	0	0	0	10	42	65
MEX	0	0	0	0	1	17	26
ARG	0	0	0	0	7	22	25
CHL	0	0	0	0	6	23	26
COL	0	0	0	0	5	15	22
VEN	0	0	0	0	0	1	2
CUB	0	0	0	0	0	0	0
FIN	0	0	0	0	16	54	75
GRC	0	0	0	0	14	37	50
PRT	0	0	0	0	8	26	37
CZE	0	0	0	0	11	25	35
NZL	0	0	0	0	6	27	38

Anexo 2. Número de Top, Highly Cited Papers y Hot papers mexicanos por cada 1000 documentos en WoS, durante 2005-2015 entre México y países de la región latinoamericana

Años	Ndoc MEX	TOP MEX	HCP MEX	Hot MEX	TOP MEX WoS	HCP MEX WoS	Hot MEX WoS
2005-2009	40,472	252	252	0	6.2265	6.227	0
2006-2010	42,886	293	293	0	6.8321	6.832	0
2007-2011	46,005	352	352	0	7.6513	7.651	0
2008-2012	49,005	408	408	0	8.3257	8.326	0
2009-2013	52,094	476	476	1	9.1373	9.137	0.019
2010-2014	55,445	522	520	17	9.4147	9.379	0.307
2011-2015	53,988	517	515	26	9.5762	9.539	0.482
Años	Ndoc BRA	TOP BRA	HCP BRA	Hot BRA	TOP BRA WoS	HCP BRA WoS	Hot BRA WoS
2005-2009	120,223	570	570	0	4.7412	4.7412	0
2006-2010	135,794	632	635	0	4.6541	4.6762	0
2007-2011	151,956	724	724	0	4.7645	4.7645	0
2008-2012	165,934	805	805	0	4.8513	4.8513	0
2009-2013	176,412	919	918	10	5.2094	5.2037	0.0567
2010-2014	186,203	1029	1027	42	5.5262	5.5155	0.2256
2011-2015	178,976	1022	1017	65	5.7103	5.6823	0.3632
Años	Ndoc ARG	TOP ARG	HCP ARG	Hot ARG	TOP ARG WoS	HCP ARG WoS	Hot ARG WoS
2005-2009	29,965	202	202	0	6.741	6.741	0
2006-2010	32,402	250	250	0	7.716	7.716	0
2007-2011	34,929	273	273	0	7.816	7.816	0

2008-2012	37,429	316	316	0	8.443	8.443	0
2009-2013	39,443	364	364	7	9.229	9.229	0.177
2010-2014	41,035	401	400	22	9.772	9.748	0.536
2011-2015	38,933	381	380	25	9.786	9.760	0.642
Años	Ndoc CHL	TOP CHL	HCP CHL	Hot CHL	TOP CHL WoS	HCP CHL WoS	Hot CHL WoS
2005-2009	18,488	147	147	0	7.951	7.951	0
2006-2010	20,367	168	168	0	8.249	8.249	0
2007-2011	22,675	198	198	0	8.732	8.732	0
2008-2012	25,215	228	228	0	9.042	9.042	0
2009-2013	27,539	285	285	6	10.349	10.349	0.218
2010-2014	30,178	345	345	23	11.432	11.432	0.762
2011-2015	29,876	349	349	26	11.682	11.682	0.870
Años	Ndoc COL	TOP COL	HCP COL	Hot COL	TOP COL WoS	HCP COL WoS	Hot COL WoS
2005-2009	7,794	71	71	0	9.110	9.110	0
2006-2010	9,602	89	89	0	9.269	9.269	0
2007-2011	11,578	110	110	0	9.501	9.501	0
2008-2012	13,543	140	140	0	10.337	10.337	0
2009-2013	14,958	175	175	5	11.699	11.699	0.334
2010-2014	16,054	210	209	15	13.081	13.019	0.934
2011-2015	15,672	221	220	22	14.102	14.038	1.404
Años	Ndoc VEN	TOP VEN	HCP VEN	Hot VEN	%TOP VEN	%HCP VEN	%Hot VEN
2005-2009	6,111	32	32	0	5.236	5.236	0
2006-2010	6,235	41	41	0	6.576	6.576	0

2007-2011	6,218	43	43	0	6.915	6.915	0
2008-2012	6,070	46	46	0	7.578	7.578	0
2009-2013	5,756	47	47	0	8.165	8.165	0.000
2010-2014	5,361	50	50	1	9.327	9.327	0.187
2011-2015	4,625	42	42	2	9.081	9.081	0.432
Años	Ndoc CUB	TOP CUB	HCP CUB	Hot CUB	TOP CUB WoS	HCP CUB WoS	Hot CUB WoS
2005-2009	3,730	11	11	0	2.949	2.949	0
2006-2010	3,798	20	20	0	5.266	5.266	0
2007-2011	3,921	26	26	0	6.631	6.631	0
2008-2012	4,054	27	27	0	6.660	6.660	0
2009-2013	4,102	38	38	0	9.264	9.264	0
2010-2014	4,127	38	38	0	9.208	9.208	0
2011-2015	3,854	33	33	0	8.563	8.563	0

Anexo 3. Número de Top, Highly Cited Papers y Hot papers mexicanos por cada 1000 en WoS, durante 2005-2015 entre México y otros países

Años	Ndoc NZL	TOP NZL	HCP NZL	Hot NZL	TOP NZL WoS	HCP NZL WoS	Hot NZL WoS
2005-2009	29,906	416	416	0	13.910	13.910	0
2006-2010	31,986	465	465	0	14.538	14.538	0
2007-2011	34,369	536	536	0	15.595	15.595	0
2008-2012	36,812	599	599	0	16.272	16.272	0
2009-2013	39,106	676	676	6	17.286	17.286	0.153
2010-2014	41,379	698	698	27	16.868	16.868	0.653
2011-2015	39,770	696	691	38	17.501	17.375	0.955
Años	Ndoc CZE	TOP CZE	HCP CZE	Hot CZE	TOP CZE WoS	HCP CZE WoS	Hot CZE WoS
2005-2009	36,593	284	284	0	7.761	7.761	0
2006-2010	39,849	347	347	0	8.708	8.708	0
2007-2011	42,823	402	402	0	9.387	9.387	0
2008-2012	45,853	468	468	0	10.207	10.207	0
2009-2013	48,592	556	556	11	11.442	11.442	0.226
2010-2014	51,956	635	635	25	12.222	12.222	0.481
2011-2015	50,378	615	615	35	12.208	12.208	0.695
Años	Ndoc PRT	TOP PRT	HCP PRT	Hot PRT	TOP PRT WoS	HCP PRT WoS	Hot PRT WoS
2005-2009	34,747	368	368	0	10.591	10.591	0

2006-2010	38,818	426	426	0	10.974	10.974	0
2007-2011	42,860	492	492	0	11.479	11.479	0
2008-2012	48,348	558	558	0	11.541	11.541	0
2009-2013	53,877	641	639	8	11.897	11.860	0.148
2010-2014	58,770	752	749	26	12.796	12.745	0.442
2011-2015	58,105	740	736	37	12.736	12.667	0.637
Años	Ndoc GRC	TOP GRC	HCP GRC	Hot GRC	TOP GRC WoS	HCP GRC WoS	Hot GRC WoS
2005-2009	46,627	400	400	0	8.579	8.579	0
2006-2010	49,353	455	455	0	9.219	9.219	0
2007-2011	51,192	512	512	0	10.002	10.002	0
2008-2012	52,705	617	617	0	11.707	11.707	0
2009-2013	53,585	710	707	14	13.250	13.194	0.261
2010-2014	53,829	797	792	37	14.806	14.713	0.687
2011-2015	49,922	782	775	50	15.664	15.524	1.002
Años	Ndoc FIN	TOP FIN	HCP FIN	Hot FIN	TOP FIN WoS	HCP FIN WoS	Hot FIN WoS
2005-2009	45,290	619	619	0	13.667	13.667	0
2006-2010	47,078	707	707	0	15.018	15.018	0
2007-2011	48,866	805	805	0	16.474	16.474	0
2008-2012	50,964	878	878	0	17.228	17.228	0
2009-2013	53,297	992	990	16	18.613	18.575	0.300
2010-2014	56,157	1,132	1,126	54	20.158	20.051	0.962
2011-2015	54,203	1096	1087	75	20.220	20.054	1.384

Anexo 4. Número de Top, Highly Cited Papers y Hot papers por cada 100,000 habitantes, en WoS, durante 2005-2015 para México y países latinoamericanos

Años	Total de hab MEX	HCP MEX	Hot MEX	HCP MEX x hab	Hot MEX x hab
Año 2005	110731826	31	0	0.028	0.000
Año 2006	112116694	43	0	0.038	0.000
Año 2007	113529819	49	0	0.043	0.000
Año 2008	114968039	62	0	0.054	0.000
Año 2009	116422752	65	0	0.056	0.000
Año 2010	117886404	73	0	0.062	0.000
Año 2011	119361233	100	0	0.084	0.000
Año 2012	120847477	105	0	0.087	0.000
Año 2013	122332399	132	0	0.108	0.000
Año 2014	123799215	109	13	0.088	0.011
Año 2015	125385833	90	21	0.072	0.017
Años	Total de hab BRA	HCP BRA	Hot BRA	HCP BRA x hab	Hot BRA x hab
Año 2005	186142403	100	0	0.054	0.000
Año 2006	188134315	92	0	0.049	0.000
Año 2007	189996976	113	0	0.059	0.000
Año 2008	191765567	132	0	0.069	0.000
Año 2009	193490922	139	0	0.072	0.000
Año 2010	195210154	165	0	0.085	0.000
Año 2011	196935134	173	0	0.088	0.000
Año 2012	198656019	194	0	0.098	0.000
Año 2013	200361925	241	6	0.120	0.003

Año 2014	202033670	258	31	0.128	0.015
Año 2015	206077898	214	37	0.104	0.018
Años	Total de hab ARG	HCP ARG	Hot ARG	HCP ARG x hab	Hot ARG x hab
Año 2005	38647854	19	0	0.049	0
Año 2006	38988923	42	0	0.108	0
Año 2007	39331357	41	0	0.104	0
Año 2008	39676083	43	0	0.108	0
Año 2009	40023641	57	0	0.142	0
Año 2010	40374224	68	0	0.168	0
Año 2011	40728738	65	0	0.160	0
Año 2012	41086927	87	0	0.212	0
Año 2013	41446246	91	3	0.220	0.007
Año 2014	41803125	95	11	0.227	0.026
Año 2015	42980026	71	7	0.165	0.016
Años	Total de hab CHL	HCP CHL	Hot CHL	HCP CHL x hab	Hot CHL x hab
Año 2005	16337749	16	0	0.098	0.000
Año 2006	16504530	30	0	0.182	0.000
Año 2007	16668892	31	0	0.186	0.000
Año 2008	16831184	35	0	0.208	0.000
Año 2009	16991729	36	0	0.212	0.000
Año 2010	17150760	35	0	0.204	0.000
Año 2011	17308449	61	0	0.352	0.000
Año 2012	17464814	61	0	0.349	0.000

Año 2013	17619708	85	0	0.482	0.000
Año 2014	17772871	100	15	0.563	0.084
Año 2015	17762647	62	15	0.349	0.084
Años	Total de hab COL	HCP COL	Hot COL	HCP COL x hab	Hot COL x hab
Año 2005	43184026	7	0	0.016	0.000
Año 2006	43841370	14	0	0.032	0.000
Año 2007	44498390	14	0	0.031	0.000
Año 2008	45153037	18	0	0.040	0.000
Año 2009	45802561	18	0	0.039	0.000
Año 2010	46444798	26	0	0.056	0.000
Año 2011	47078792	36	0	0.076	0.000
Año 2012	47704427	44	0	0.092	0.000
Año 2013	48321405	56	0	0.116	0.000
Año 2014	48929706	53	9	0.108	0.018
Año 2015	48929706	54	11	0.110	0.022
Años	Total de hab VEN	HCP VEN	Hot VEN	HCP VEN x hab	Hot VEN x hab
Año 2005	26725897	3	0	0.011	0
Año 2006	27190882	5	0	0.018	0
Año 2007	27655937	7	0	0.025	0
Año 2008	28120312	7	0	0.025	0
Año 2009	28583040	9	0	0.031	0
Año 2010	29043283	11	0	0.038	0
Año 2011	29500625	7	0	0.024	0

Año 2012	29954782	10	0	0.033	0
Año 2013	30405207	7	0	0.023	0
Año 2014	30851343	14	0	0.045	0
Año 2015	30851343	5	0	0.016	0
Años	Total de hab CUB	HCP CUB	Hot CUB	HCP CUB x hab	Hot CUB x hab
Año 2005	11292078	0	0	0.000	0
Año 2006	11301100	3	0	0.027	0
Año 2007	11301674	2	0	0.018	0
Año 2008	11296355	0	0	0.000	0
Año 2009	11288826	4	0	0.035	0
Año 2010	11281768	10	0	0.089	0
Año 2011	11276053	10	0	0.089	0
Año 2012	11270957	3	0	0.027	0
Año 2013	11265629	11	0	0.098	0
Año 2014	11258597	6	0	0.053	0
Año 2015	11258597	7	1	0.062	0.009

Anexo 5. Número de Top, Highly Cited Papers y Hot papers por cada 100,000 habitantes, en WoS, durante 2005-2015 para México y otros países

Años	Total de hab GRC	HCP GRC	Hot GRC	HCP GRC x hab	Hot GRC x hab
Año 2005	11092913	61	0	0.55	0
Año 2006	11127947	77	0	0.69	0
Año 2007	11163002	65	0	0.58	0
Año 2008	11186439	89	0	0.80	0
Año 2009	11187085	111	0	0.99	0
Año 2010	11153454	119	0	1.07	0
Año 2011	11123213	130	0	1.17	0
Año 2012	11092771	170	0	1.53	0
Año 2013	11027549	176	4	1.60	0.036
Año 2014	10957740	213	25	1.94	0.228
Año 2015	10957740	160	19	1.46	0.173
Años	Total de hab CZE	HCP CZE	Hot CZE	HCP CZE x hab	Hot CZE x hab
Año 2005	10211216	36	0	0.35	0
Año 2006	10238905	58	0	0.57	0
Año 2007	10298828	61	0	0.59	0
Año 2008	10384603	70	0	0.67	0
Año 2009	10443936	59	0	0.56	0
Año 2010	10474410	98	0	0.94	0
Año 2011	10496088	111	0	1.06	0
Año 2012	10510785	128	0	1.22	0
Año 2013	10514272	162	5	1.54	0.048

Año 2014	10510566	157	20	1.49	0.190
Año 2015	10510566	123	15	1.17	0.143
Años	Total de hab FIN	HCP FIN	Hot FIN	HCP FIN x hab	Hot FIN x hab
Año 2005	5246096	89	0	1.70	0
Año 2006	5266268	110	0	2.09	0
Año 2007	5288720	134	0	2.53	0
Año 2008	5313399	155	0	2.92	0
Año 2009	5338871	137	0	2.57	0
Año 2010	5363352	176	0	3.28	0
Año 2011	5388272	205	0	3.80	0
Año 2012	5413971	206	0	3.80	0
Año 2013	5438972	262	7	4.82	0.129
Año 2014	5463596	283	43	5.18	0.787
Año 2015	5463596	150	27	2.75	0.494
Años	Total de hab PRT	HCP PRT	Hot PRT	HCP PRT x hab	Hot PRT x hab
Año 2005	10503330	44	0	0.42	0
Año 2006	10522288	64	0	0.61	0
Año 2007	10542964	84	0	0.80	0
Año 2008	10558177	91	0	0.86	0
Año 2009	10568247	85	0	0.80	0
Año 2010	10573100	107	0	1.01	0
Año 2011	10557560	124	0	1.17	0
Año 2012	10514844	153	0	1.46	0
Año 2013	10457295	169	3	1.62	0.029
Año 2014	10397393	192	25	1.85	0.240

Año 2015	10397393	136	18	1.31	0.173
Años	Total de habitantes NZL	HCP NZL	Hot NZL	HCP NZL x hab	Hot NZL x hab
Año 2005	4133900	69	0	1.67	0
Año 2006	4184600	68	0	1.63	0
Año 2007	4223800	88	0	2.08	0
Año 2008	4259800	87	0	2.04	0
Año 2009	4302600	117	0	2.72	0
Año 2010	4350700	119	0	2.74	0
Año 2011	4384000	143	0	3.26	0
Año 2012	4408100	148	0	3.36	0
Año 2013	4442100	166	7	3.74	0.158
Año 2014	4509700	140	23	3.10	0.510
Año 2015	4509700	143	17	3.17	0.377

Anexo 6. Ranking de 155 países con mayor excelencia científica según el ESI durante el período 2005-2015

	Países	Web of Science Documents	No. Habitantes	Citas	Citas/doc	Top Papers	SDI	HCP	Hot	RI
1	USA	3773427	318857056	66790551	17.7	69608	11.83	64464	1420	1.54
2	GERMANY (FED REP GER)	988153	80889505	15704549	15.89	16301	12.22	15291	430	1.39
3	ENGLAND	891184	64510376	15594264	17.5	18387	13.81	17349	497	1.53
4	CHINA MAINLAND	1657047	1364270000	13923842	8.4	16073	1.21	15863	574	0.73
5	FRANCE	697105	66201365	10505229	15.07	10775	10.53	10144	309	1.31
6	JAPAN	844743	127131800	10094917	11.95	6727	6.64	6203	167	1.04
7	CANADA	605129	35540419	9514790	15.72	10512	17.03	9951	272	1.37
8	ITALY	582608	61336387	8317827	14.28	8244	9.50	7740	258	1.24
9	AUSTRALIA	461499	23490736	6548488	14.19	8079	19.65	7863	291	1.24
10	NETHERLANDS	346329	16854183	6488596	18.74	7940	20.55	7571	233	1.63
11	SPAIN	496756	46404602	6472534	13.03	6679	10.70	6409	228	1.14
12	SWITZERLAND	250553	8190229	5004142	19.97	6467	30.59	6126	215	1.74
13	SOUTH KOREA	452102	50423955	4139433	9.16	3605	8.97	3469	135	0.80
14	SWEDEN	230482	9689555	3905546	16.95	4285	23.79	4063	139	1.48
15	INDIA	469564	1267401849	3644391	7.76	2608	0.37	2529	70	0.68
16	BELGIUM	190464	11225207	3211131	16.86	3900	16.97	3720	115	1.47
17	BRAZIL	346807	202033670	2654948	7.66	1871	1.72	1812	78	0.67
18	SCOTLAND	137408	4612719	2590044	18.85	3254	29.79	3070	110	1.64
19	DENMARK	139659	5639565	2536006	18.16	3317	24.76	3184	125	1.58
20	TAIWAN	259124	23113990	2439899	9.42	1689	11.21	1606	52	0.82
21	AUSTRIA	129750	8534492	2029723	15.64	2480	15.20	2372	76	1.36

22	ISRAEL	132545	8215300	1981133	14.95	1979	16.13	1830	68	1.30
23	RUSSIA	307293	143819569	1821034	5.93	1450	2.14	1345	65	0.52
24	FINLAND	113983	5463596	1778844	15.61	1981	20.86	1917	63	1.36
25	POLAND	220413	37995529	1778534	8.07	1656	5.80	1585	67	0.70
26	NORWAY	106321	5136475	1604353	15.09	1977	20.70	1864	69	1.32
27	TURKEY	237216	75837020	1582711	6.67	1222	3.13	1162	39	0.58
28	SINGAPORE	102426	5469700	1514563	14.79	2246	18.73	2190	86	1.29
29	HONG KONG	111609	7241700	1466883	13.14	1932	15.41	1820	57	1.15
30	GREECE	110602	10957740	1356016	12.26	1411	10.09	1371	52	1.07
31	PORTUGAL	107104	10397393	1259786	11.76	1282	10.30	1251	51	1.03
32	IRAN	190008	78470222	1152290	6.06	1047	2.42	1029	48	0.53
33	NEW ZEALAND	80218	4509700	1089433	13.58	1289	17.79	1241	38	1.18
34	IRELAND	69508	6197100	1080339	15.54	1368	11.22	1320	53	1.35
35	CZECH REPUBLIC	100559	10510566	1040485	10.35	1101	9.57	1087	38	0.90
36	MEXICO	108523	123799215	939491	8.66	877	0.88	847	36	0.76
37	SOUTH AFRICA	89247	961463609	931350	10.44	1218	0.09	1174	45	0.91
38	ARGENTINA	79453	41803125	816565	10.28	701	1.90	684	20	0.90
39	WALES	47027	3063456	767812	16.33	934	15.35	900	33	1.42
40	HUNGARY	63925	9861673	762843	11.93	839	6.48	802	35	1.04
41	CHILE	56080	17762647	567692	10.12	569	3.16	559	31	0.88
42	THAILAND	56952	67725979	563302	9.89	493	0.84	481	18	0.86
43	EGYPT	64696	89579670	431269	6.67	363	0.72	359	17	0.58
44	MALAYSIA	66880	29901997	398449	5.96	629	2.24	651	33	0.52
45	ROMANIA	65321	19910995	388497	5.95	531	3.28	507	22	0.52
46	SAUDI ARABIA	58836	30886545	376379	6.4	1004	1.90	1026	100	0.56
47	SLOVENIA	35442	2062218	339740	9.59	357	17.19	351	15	0.84

48	NORTHERN IRELAND	22976	1840498	339023	14.76	361	12.48	357	16	1.29
49	PAKISTAN	49165	185044286	296871	6.04	410	0.27	415	25	0.53
50	UKRAINE	51262	45362900	287805	5.61	262	1.13	253	16	0.49
51	SLOVAKIA	31506	5417000	269758	8.56	248	5.82	238	8	0.75
52	CROATIA	32945	4238389	263014	7.98	287	7.77	285	11	0.70
53	COLOMBIA	27594	47791393	237342	8.6	349	0.58	345	19	0.75
54	SERBIA	38797	7129428	221820	5.72	341	5.44	350	17	0.50
55	BULGARIA	23641	7223938	207853	8.79	207	3.27	211	14	0.77
56	ESTONIA	14386	1313645	191108	13.28	330	10.95	322	22	1.16
57	ICELAND	8270	327589	169864	20.54	262	25.25	249	16	1.79
58	TUNISIA	27508	10996600	166777	6.06	115	2.50	103	3	0.53
59	KENYA	12136	44863583	160548	13.23	214	0.27	215	12	1.15
60	LITHUANIA	19642	2929323	135487	6.9	177	6.71	180	12	0.60
61	VIETNAM	16040	90730000	121266	7.56	154	0.18	147	12	0.66
62	INDONESIA	12294	254454778	115433	9.39	150	0.05	152	7	0.82
63	NIGERIA	19276	177475986	113396	5.88	84	0.11	89	12	0.51
64	PHILIPPINES	9215	99138690	107845	11.7	171	0.09	168	7	1.02
65	VENEZUELA	12280	30693827	104828	8.54	88	0.40	84	1	0.74
66	MOROCCO	15007	33921203	104249	6.95	110	0.44	107	10	0.61
67	ALGERIA	18330	38934334	100778	5.5	83	0.47	84	1	0.48
68	PERU	7216	30973148	98410	13.64	177	0.23	177	12	1.19
69	BANGLADESH	11694	159077513	92524	7.91	96	0.07	93	10	0.69
70	UNITED ARAB EMIRATES	12260	9086139	89131	7.27	108	1.35	106	10	0.63
71	UGANDA	6708	37782971	87503	13.04	90	0.18	96	7	1.14
72	TANZANIA	6615	51822621	83077	12.56	89	0.13	86	6	1.10
73	URUGUAY	7338	3419516	82247	11.21	87	2.15	82	4	0.98

74	BELARUS	11406	9470000	79695	6.99	119	1.20	119	7	0.61
75	LEBANON	8824	4546774	77661	8.8	99	1.94	97	7	0.77
76	CYPRUS	7683	1153658	76038	9.9	154	6.66	150	12	0.86
77	JORDAN	11087	6,321,000	73581	6.64	93	1.75	87	5	0.58
78	ARMENIA	6512	3006154	70273	10.79	143	2.17	140	8	0.94
79	LUXEMBOURG	6393	556074	69163	10.82	101	11.50	99	5	0.94
80	CUBA	8593	11379111	68641	7.99	60	0.76	58	1	0.70
81	PANAMA	3147	3867535	63203	20.08	98	0.81	98	5	1.75
82	COSTA RICA	4531	4757606	61316	13.53	71	0.95	68	3	1.18
83	REPUBLIC OF GEORGIA	4931	4504100	53561	10.86	131	1.09	136	16	0.95
84	GHANA	5333	26786598	51502	9.66	56	0.20	60	4	0.84
85	ETHIOPIA	6842	96958732	50676	7.41	53	0.07	60	8	0.65
86	SRI LANKA	5194	20639000	49196	9.47	85	0.25	93	12	0.83
87	KUWAIT	7068	3753121	48086	6.8	29	1.88	44	1	0.59
88	LATVIA	5432	1990351	47402	8.73	67	2.73	33	2	0.76
89	CAMEROON	6133	22773014	47155	7.69	41	0.27	67	3	0.67
90	ECUADOR	4173	15902916	44519	10.67	64	0.26	63	2	0.93
91	QATAR	6229	2172065	39465	6.34	101	2.87	107	11	0.55
92	MALAWI	2903	16695253	38179	13.15	52	0.17	54	2	1.15
93	NEPAL	3870	28174724	35286	9.12	38	0.14	39	2	0.80
94	OMAN	4839	4236057	33317	6.89	43	1.14	45	4	0.60
95	ZIMBABWE	2815	15245855	32254	11.46	38	0.18	38	3	1.00
96	SENEGAL	3301	14672557	30833	9.34	24	0.22	27	2	0.81
97	BOLIVIA	2104	10561887	29312	13.93	33	0.20	35	3	1.21
98	ZAMBIA	2183	15721343	28283	12.96	39	0.14	39	3	1.13
99	BURKINA FASO	2581	17589198	27077	10.49	23	0.15	23	0	0.91
100	GAMBIA	1072	1928201	25610	23.89	26	0.56	26	3	2.08

101	AZERBAIJAN	4225	9537823	25048	5.93	49	0.44	45	4	0.52
102	SERBIA & MONTENEGRO	1961	7129428	24092	12.29	8	0.28	6	0	1.07
103	BOTSWANA	2082	3817554	23748	11.41	24	0.55	21	3	0.99
104	MOZAMBIQUE	1415	27216276	22855	16.15	27	0.05	29	0	1.41
105	CAMBODIA	1628	15328136	22252	13.67	42	0.11	39	2	1.19
106	SUDAN	2579	39350274	21773	8.44	16	0.07	18	2	0.74
107	MOLDOVA	2565	4068000	21377	8.33	28	0.63	26	2	0.73
108	SYRIA	2723	22,087,048	21029	7.72	21	0.12	23	3	0.67
109	COTE IVOIRE	2112	22157107	20155	9.54	18	0.10	19	4	0.83
110	MADAGASCAR	1875	23571713	19799	10.56	24	0.08	24	0	0.92
111	UZBEKISTAN	3521	30757700	19611	5.57	10	0.11	9	1	0.49
112	MALI	1408	17086022	19067	13.54	18	0.08	21	2	1.18
113	MACEDONIA	2706	2081545	18342	6.78	25	1.30	26	3	0.59
114	NEW CALEDONIA	1506	266000	17552	11.65	17	5.66	17	3	1.02
116	BENIN	2160	10598482	17050	7.89	12	0.20	12	4	0.69
117	IRAQ	4668	34812326	17034	3.65	30	0.13	27	5	0.32
118	BOSNIA & HERZEGOVINA	3617	3817554	16746	4.63	12	0.95	12	1	0.40
119	KAZAKHSTAN	4106	17289111	16526	4.02	15	0.24	16	0	0.35
120	MALTA	1671	427404	15953	9.55	16	3.91	17	0	0.83
121	JAMAICA	1908	2721252	15857	8.31	20	0.70	19	4	0.72
122	GABON	1097	1687673	14987	13.66	15	0.65	15	0	1.19
123	CONGO DEMOCRATIC REPUBLIC	1389	74877030	14380	10.35	31	0.02	29	4	0.90
124	MONGOLIA	1703	2909871	14161	8.32	11	0.59	12	2	0.73

125	TRINIDAD & TOBAGO	1864	1354483	13653	7.32	11	1.38	11	0	0.64
126	NAMIBIA	1121	2402858	13526	12.07	15	0.47	17	0	1.05
127	MONACO	814	37623	12988	15.96	15	21.64	15	0	1.39
128	PAPUA NEW GUINEA	964	12275527	12750	13.23	23	0.08	22	3	1.15
129	GUATEMALA	1132	16015494	11949	10.56	18	0.07	19	2	0.92
130	GUADELOUPE	1289	403,750	11316	8.78	13	3.19	14	0	0.77
131	LAOS	1098	6689300	10927	9.95	14	0.16	13	2	0.87
132	FIJI	1046	886450	9750	9.32	16	1.18	13	4	0.81
133	NICARAGUA	685	6013913	9307	13.59	12	0.11	12	0	1.18
134	FRENCH GUIANA	652	259109	9022	13.84	14	2.52	13	1	1.21
135	GREENLAND	710	55,847	8993	12.67	16	12.71	11	0	1.10
136	NIGER	878	19113728	8678	9.88	9	0.05	9	0	0.86
137	BAHRAIN	1509	1403289	8640	5.73	9	1.08	10	6	0.50
138	LIBYA	1493	6258984	8553	5.73	9	0.24	8	2	0.50
139	RWANDA	954	11341544	8523	8.93	18	0.08	20	4	0.78
140	FRENCH POLYNESIA	778	279781	8376	10.77	9	2.78	9	1	0.94
141	YEMEN	1373	26183676	8246	6.01	9	0.05	10	2	0.52
142	BARBADOS	692	283380	7464	10.79	10	2.44	9	2	0.94
143	BERMUDA	325	65181	7216	22.2	13	4.99	12	0	1.94
144	YUGOSLAVIA	484	1219000	6770	13.99	5	0.40			1.22
145	HONDURAS	477	7961680	6755	14.16	8	0.06	8	0	1.23
146	CONGO PEOPLES REP	675	25026772	6629	9.82	7	0.03	6	0	0.86
147	LIECHTENSTEIN	548	37286	6474	11.81	1	14.70	1	0	1.03
148	MAURITIUS	849	3969625	6311	7.43	4	0.21	6	0	0.65

149	MONTENEGRO	1369	621800	5956	4.35	5	2.20	5	0	0.38
150	MACAU	1712	647000	5786	3.38	25	2.65	28	5	0.29
151	SEYCHELLES	299	91526	5237	17.52	11	3.27	10	2	1.53
152	KYRGYZSTAN	752	5834200	4642	6.17	4	0.13	5	0	0.54
153	GUINEA BISSAU	289	1800513	4249	14.7	2	0.16			1.28
154	VATICAN	134	258883	2457	18.34	3	0.52	3	0	1.60
155	GIBRALTAR	26	32230	496	19.08	1	0.81	0	0	1.66

Anexo 7. Ranking de 100 países según el índice de desarrollo científico (SDI) durante el 2005-2015

	Países	Ndoc WoS	No. Habitantes	Citas	Citas/doc	Top Papers	SDI	HCP	Hot	RI	NI
1	SWITZERLAND	250553	8190229	5004142	19.97	6467	30.59	6126	215	1.741	1.64
2	SCOTLAND	137408	4612719	2590044	18.85	3254	29.79	3070	110	1.643	1.58
3	ICELAND	8270	327589	169864	20.54	262	25.25	249	16	1.791	1.54
4	DENMARK	139659	5639565	2536006	18.16	3317	24.76	3184	125	1.583	1.46
5	SWEDEN	230482	9689555	3905546	16.95	4285	23.79	4063	139	1.478	1.38
6	FINLAND	113983	5463596	1778844	15.61	1981	20.86	1917	63	1.361	1.33
7	NORWAY	106321	5136475	1604353	15.09	1977	20.70	1864	69	1.316	1.28
8	NETHERLANDS	346329	16854183	6488596	18.74	7940	20.55	7571	233	1.634	1.52
9	AUSTRALIA	461499	23490736	6548488	14.19	8079	19.65	7863	291	1.237	1.24
10	SINGAPORE	102426	5469700	1514563	14.79	2246	18.73	2190	86	1.289	1.21
11	NEW ZEALAND	80218	4509700	1089433	13.58	1289	17.79	1241	38	1.184	1.18
12	SLOVENIA	35442	2062218	339740	9.59	357	17.19	351	15	0.836	0.97
13	CANADA	605129	35540419	9514790	15.72	10512	17.03	9951	272	1.371	1.33
14	BELGIUM	190464	11225207	3211131	16.86	3900	16.97	3720	115	1.470	1.36
15	ISRAEL	132545	8215300	1981133	14.95	1979	16.13	1830	68	1.303	1.24
16	HONG KONG	111609	7241700	1466883	13.14	1932	15.41	1820	57	1.146	1.24
17	WALES	47027	3063456	767812	16.33	934	15.35	900	33	1.424	1.42
18	AUSTRIA	129750	8534492	2029723	15.64	2480	15.20	2372	76	1.364	1.29
19	ENGLAND	891184	64510376	15594264	17.5	18387	13.81	17349	497	1.526	1.51
20	NORTHERN IRELAND	22976	1840498	339023	14.76	361	12.48	357	16	1.287	1.21
21	GERMANY (FED REP GER)	988153	80889505	15704549	15.89	16301	12.22	15291	430	1.385	1.30
22	USA	3773427	318857056	66790551	17.7	69608	11.83	64464	1420	1.543	1.47

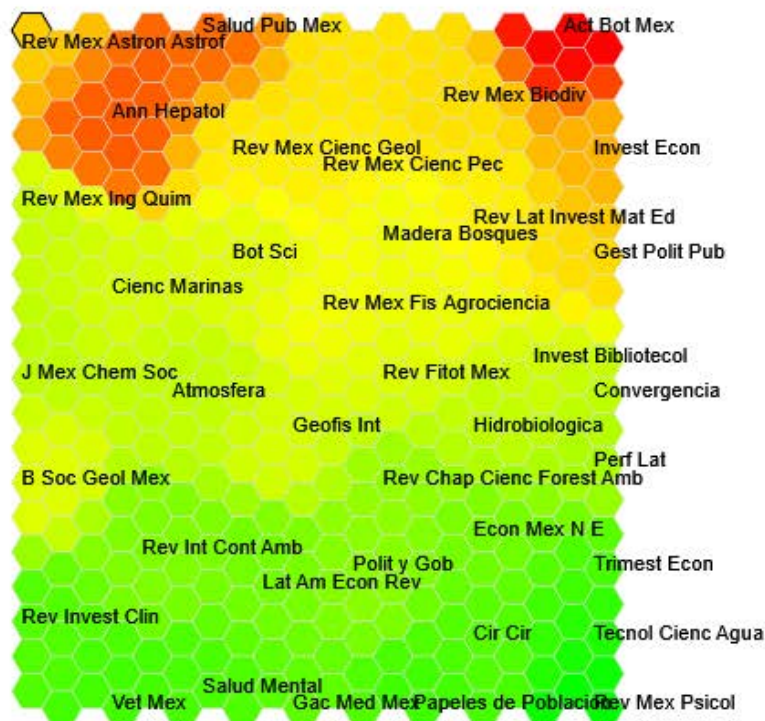
23	LUXEMBOURG	6393	556074	69163	10.82	101	11.50	99	5	0.943	0.89
24	IRELAND	69508	6197100	1080339	15.54	1368	11.22	1320	53	1.355	1.51
25	TAIWAN	259124	23113990	2439899	9.42	1689	11.21	1606	52	0.821	0.87
26	ESTONIA	14386	1313645	191108	13.28	330	10.95	322	22	1.158	1.08
27	SPAIN	496756	46404602	6472534	13.03	6679	10.70	6409	228	1.136	1.10
28	FRANCE	697105	66201365	10505229	15.07	10775	10.53	10144	309	1.314	1.27
29	PORTUGAL	107104	10397393	1259786	11.76	1282	10.30	1251	51	1.025	1.05
30	GREECE	110602	10957740	1356016	12.26	1411	10.09	1371	52	1.069	1.02
31	CZECH REPUBLIC	100559	10510566	1040485	10.35	1101	9.57	1087	38	0.902	0.85
32	ITALY	582608	61336387	8317827	14.28	8244	9.50	7740	258	1.245	1.11
33	SOUTH KOREA	452102	50423955	4139433	9.16	3605	8.97	3469	135	0.799	0.83
34	CROATIA	32945	4238389	263014	7.98	287	7.77	285	11	0.696	0.77
35	LITHUANIA	19642	2929323	135487	6.9	177	6.71	180	12	0.602	0.69
36	CYPRUS	7683	1153658	76038	9.9	154	6.66	150	12	0.863	0.79
37	JAPAN	844743	127131800	10094917	11.95	6727	6.64	6203	167	1.042	0.95
38	HUNGARY	63925	9861673	762843	11.93	839	6.48	802	35	1.040	0.99
39	SLOVAKIA	31506	5417000	269758	8.56	248	5.82	238	8	0.746	0.75
40	POLAND	220413	37995529	1778534	8.07	1656	5.80	1585	67	0.704	0.66
41	SERBIA	38797	7129428	221820	5.72	341	5.44	350	17	0.499	0.61
42	ROMANIA	65321	19910995	388497	5.95	531	3.28	507	22	0.519	0.59
43	BULGARIA	23641	7223938	207853	8.79	207	3.27	211	14	0.766	0.74
44	CHILE	56080	17762647	567692	10.12	569	3.16	559	31	0.882	0.78
45	TURKEY	237216	75837020	1582711	6.67	1222	3.13	1162	39	0.582	0.62
46	QATAR	6229	2172065	39465	6.34	101	2.87	107	11	0.553	0.31
47	LATVIA	5432	1990351	47402	8.73	67	2.73	33	2	0.761	0.72
48	TUNISIA	27508	10996600	166777	6.06	115	2.50	103	3	0.528	0.53
49	IRAN	190008	78470222	1152290	6.06	1047	2.42	1029	48	0.528	0.53
50	MALAYSIA	66880	29901997	398449	5.96	629	2.24	651	33	0.520	0.53

51	ARMENIA	6512	3006154	70273	10.79	143	2.17	140	8	0.941	0.18
52	URUGUAY	7338	3419516	82247	11.21	87	2.15	82	4	0.977	1.08
53	RUSSIA	307293	143819569	1821034	5.93	1450	2.14	1345	65	0.517	0.52
54	LEBANON	8824	4546774	77661	8.8	99	1.94	97	7	0.767	0.69
55	SAUDI ARABIA	58836	30886545	376379	6.4	1004	1.90	1026	100	0.558	0.58
56	ARGENTINA	79453	41803125	816565	10.28	701	1.90	684	20	0.896	0.82
57	KUWAIT	7068	3753121	48086	6.8	29	1.88	44	1	0.593	0.46
58	JORDAN	11087	6,321,000	73581	6.64	93	1.75	87	5	0.579	0.54
59	BRAZIL	346807	202033670	2654948	7.66	1871	1.72	1812	78	0.668	0.68
60	UNITED ARAB EMIRATES	12260	9086139	89131	7.27	108	1.35	106	10	0.634	0.72
61	CHINA MAINLAND	1657047	1364270000	13923842	8.4	16073	1.21	15863	574	0.732	0.74
62	BELARUS	11406	9470000	79695	6.99	119	1.20	119	7	0.609	0.34
63	OMAN	4839	4236057	33317	6.89	43	1.14	45	4	0.601	0.59
64	UKRAINE	51262	45362900	287805	5.61	262	1.13	253	16	0.489	0.62
65	REPUBLIC OF GEORGIA	4931	4504100	53561	10.86	131	1.09	136	16	0.947	0.37
66	COSTA RICA	4531	4757606	61316	13.53	71	0.95	68	3	1.180	1.06
67	MEXICO	108523	123799215	939491	8.66	877	0.88	847	36	0.755	0.78
68	THAILAND	56952	67725979	563302	9.89	493	0.84	481	18	0.862	0.80
69	PANAMA	3147	3867535	63203	20.08	98	0.81	98	5	1.751	1.01
70	CUBA	8593	11379111	68641	7.99	60	0.76	58	1	0.697	0.59
71	EGYPT	64696	89579670	431269	6.67	363	0.72	359	17	0.582	0.57
72	COLOMBIA	27594	47791393	237342	8.6	349	0.58	345	19	0.750	0.75
73	GAMBIA	1072	1928201	25610	23.89	26	0.56	26	3	2.083	0.46
74	ALGERIA	18330	38934334	100778	5.5	83	0.47	84	1	0.480	0.47
75	MOROCCO	15007	33921203	104249	6.95	110	0.44	107	10	0.606	0.63
76	VENEZUELA	12280	30693827	104828	8.54	88	0.40	84	1	0.745	0.79

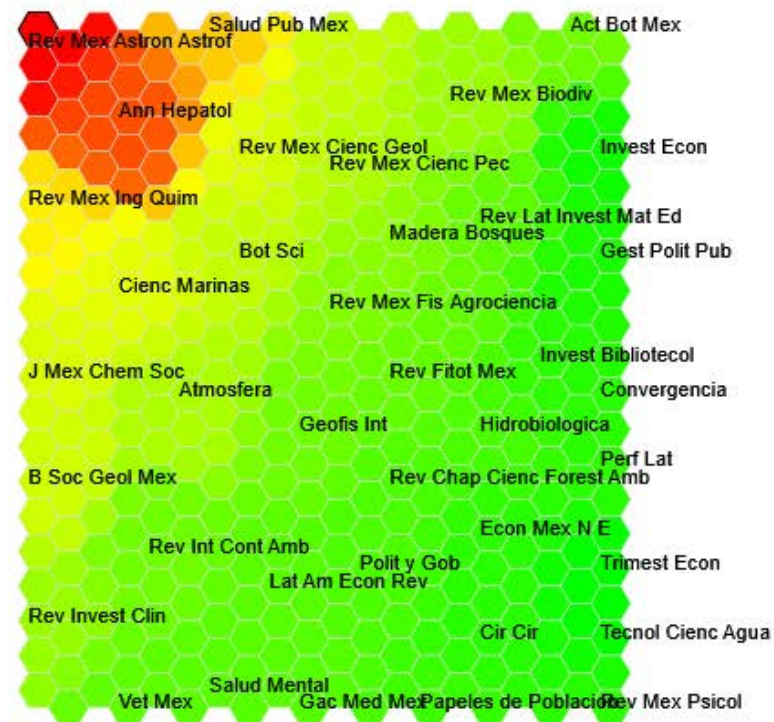
77	INDIA	469564	1267401849	3644391	7.76	2608	0.37	2529	70	0.677	0.64
78	KENYA	12136	44863583	160548	13.23	214	0.27	215	12	1.153	0.80
79	CAMEROON	6133	22773014	47155	7.69	41	0.27	67	3	0.670	0.66
80	PAKISTAN	49165	185044286	296871	6.04	410	0.27	415	25	0.527	0.58
81	ECUADOR	4173	15902916	44519	10.67	64	0.26	63	2	0.930	0.61
82	SRI LANKA	5194	20639000	49196	9.47	85	0.25	93	12	0.826	0.62
83	PERU	7216	30973148	98410	13.64	177	0.23	177	12	1.189	0.94
84	SENEGAL	3301	14672557	30833	9.34	24	0.22	27	2	0.814	0.50
85	BOLIVIA	2104	10561887	29312	13.93	33	0.20	35	3	1.214	0.92
86	GHANA	5333	26786598	51502	9.66	56	0.20	60	4	0.842	0.58
87	ZIMBABWE	2815	15245855	32254	11.46	38	0.18	38	3	0.999	0.52
88	UGANDA	6708	37782971	87503	13.04	90	0.18	96	7	1.137	0.65
89	VIETNAM	16040	90730000	121266	7.56	154	0.18	147	12	0.659	0.77
90	MALAWI	2903	16695253	38179	13.15	52	0.17	54	2	1.146	0.51
91	BURKINA FASO	2581	17589198	27077	10.49	23	0.15	23	0	0.915	0.37
92	ZAMBIA	2183	15721343	28283	12.96	39	0.14	39	3	1.130	0.46
93	NEPAL	3870	28174724	35286	9.12	38	0.14	39	2	0.795	0.53
94	TANZANIA	6615	51822621	83077	12.56	89	0.13	86	6	1.095	0.86
95	NIGERIA	19276	177475986	113396	5.88	84	0.11	89	12	0.513	0.58
96	PHILIPPINES	9215	99138690	107845	11.7	171	0.09	168	7	1.020	0.80
97	SOUTH AFRICA	89247	961463609	931350	10.44	1218	0.09	1174	45	0.910	0.91
98	BANGLADESH	11694	159077513	92524	7.91	96	0.07	93	10	0.690	0.63
99	ETHIOPIA	6842	96958732	50676	7.41	53	0.07	60	8	0.646	0.65
100	INDONESIA	12294	254454778	115433	9.39	150	0.05	152	7	0.819	0.79

Anexo 8. Mapa de 39 revistas mexicanas con mejores valores de IF y SNIP para el año 2015

a) SNIP 2015

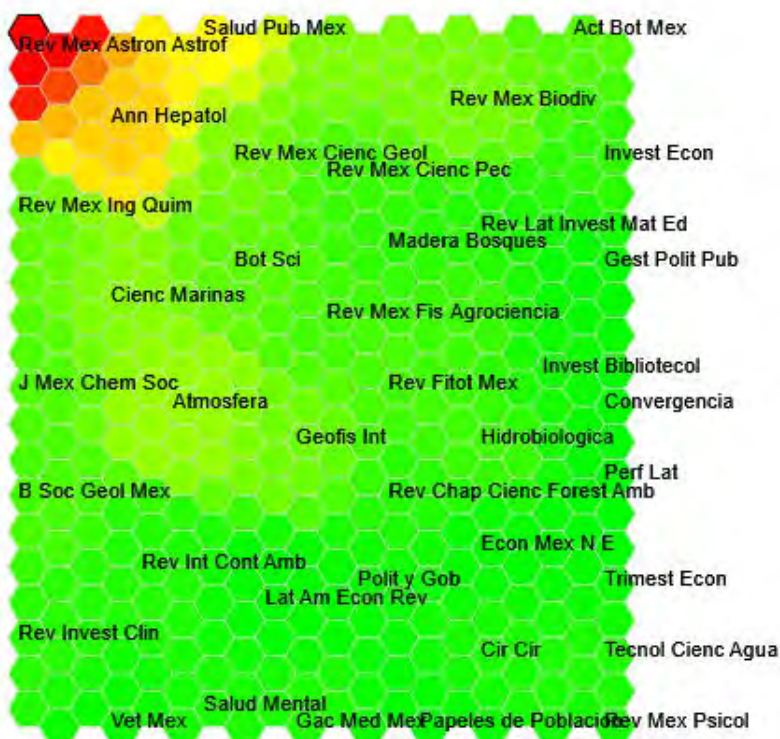


b) IF 2015

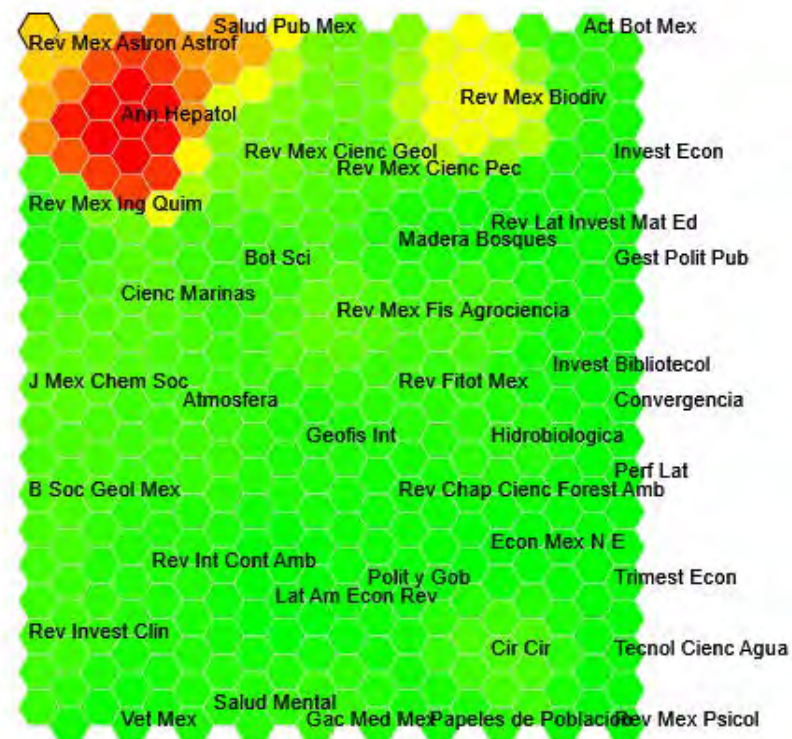


Anexo 9. Mapa de 39 revistas mexicanas con mejores valores de EFS y SJR, para el año 2015

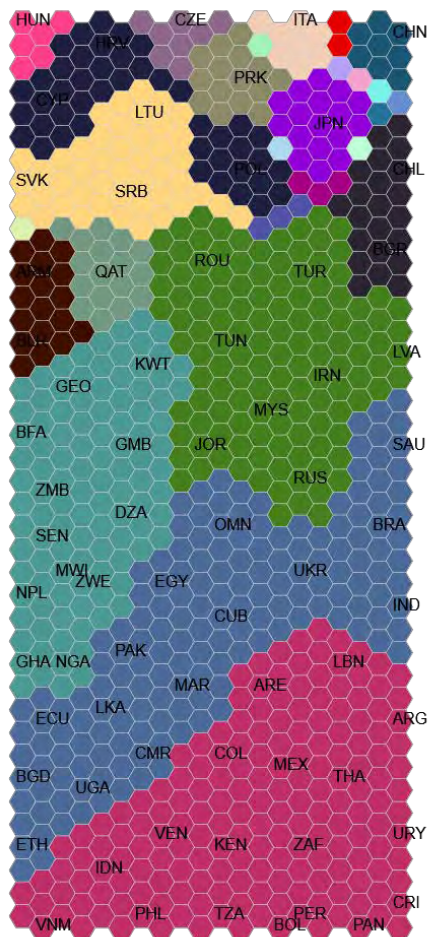
c) SJR 2015



d) EFS 2015

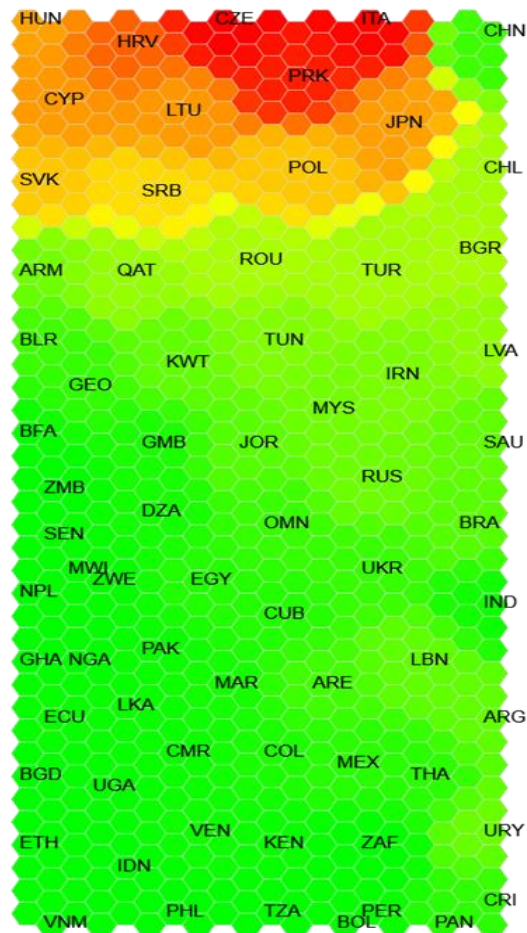


Anexo 10. Clustering de países donde se analiza multiparamétricamente el Índice de desarrollo científico (SDI), Highly Cited Papers (HCP), Hot paper (Hot) e Impacto Normalizado (NI), donde HCP y Hot se consideran como indicadores de tamaño dependiente

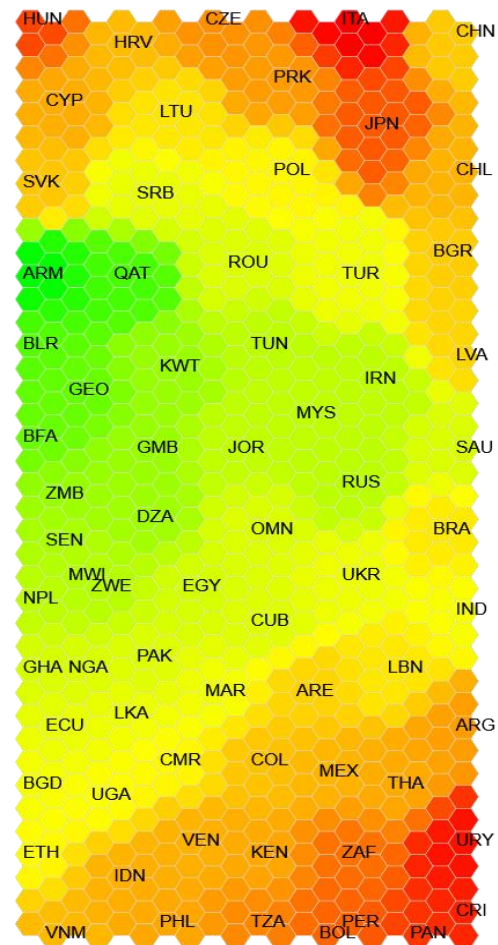


Anexo 11. Mapas de componentes de países para los indicadores Índice de desarrollo científico (SDI) e Impacto Normalizado (NI) como indicadores de tamaño dependiente

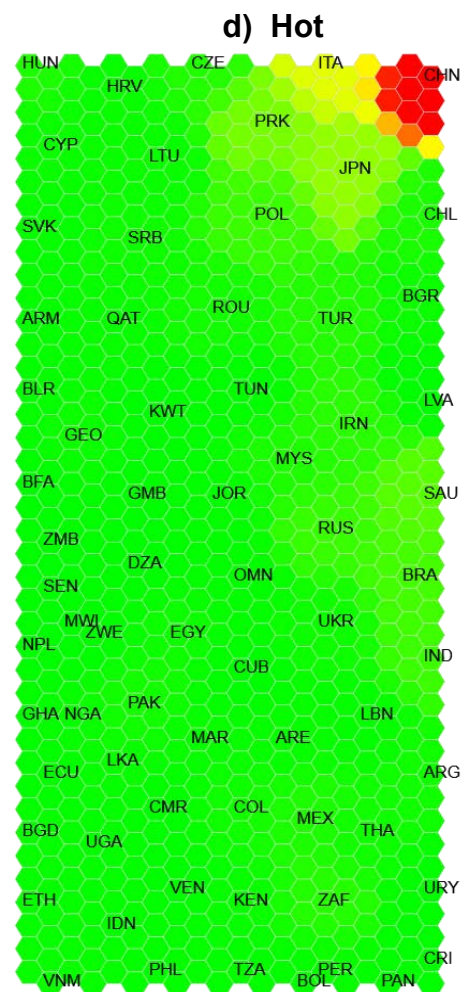
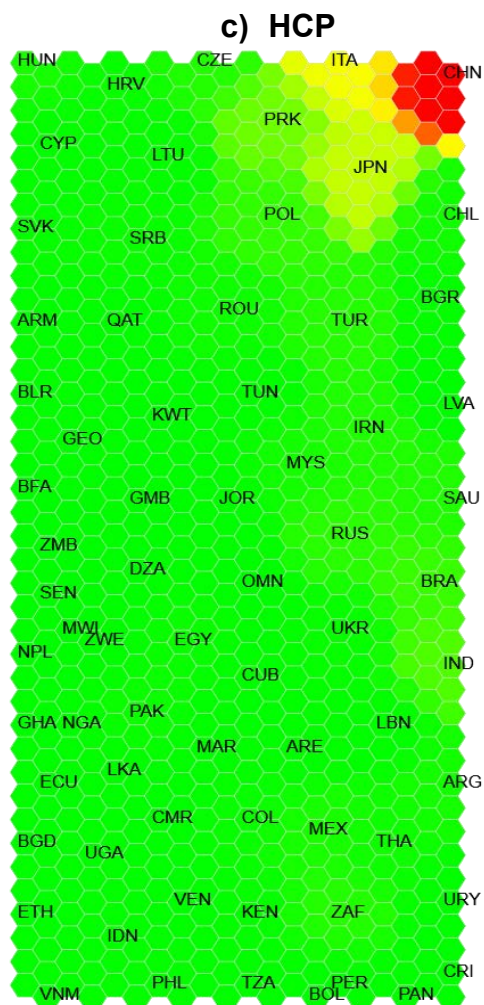
a) SDI



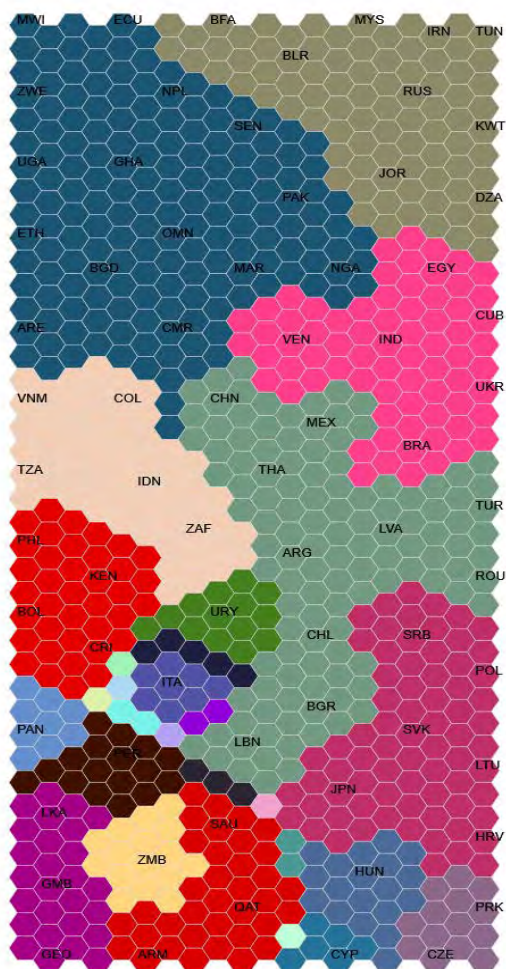
b) NI



Anexo 12. Mapas de componentes de países para los indicadores Highly Cited Papers (HCP) y Hot paper (Hot) como indicadores de tamaño dependiente

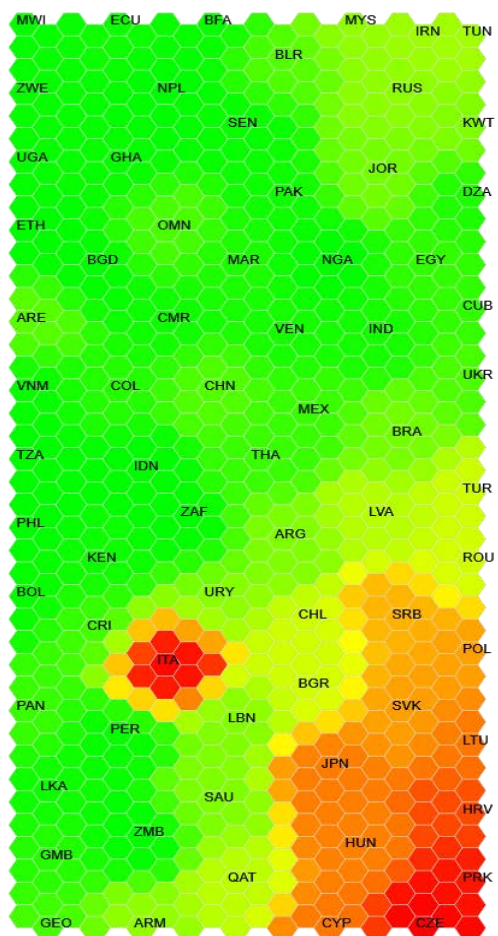


Anexo 13. Clustering de países donde se analiza multiparamétricamente el Índice de desarrollo científico (SDI), Highly Cited Papers (HCP), Hot paper (Hot) e Impacto Normalizado (NI), donde HCP y Hot se consideran como indicadores de tamaño independiente

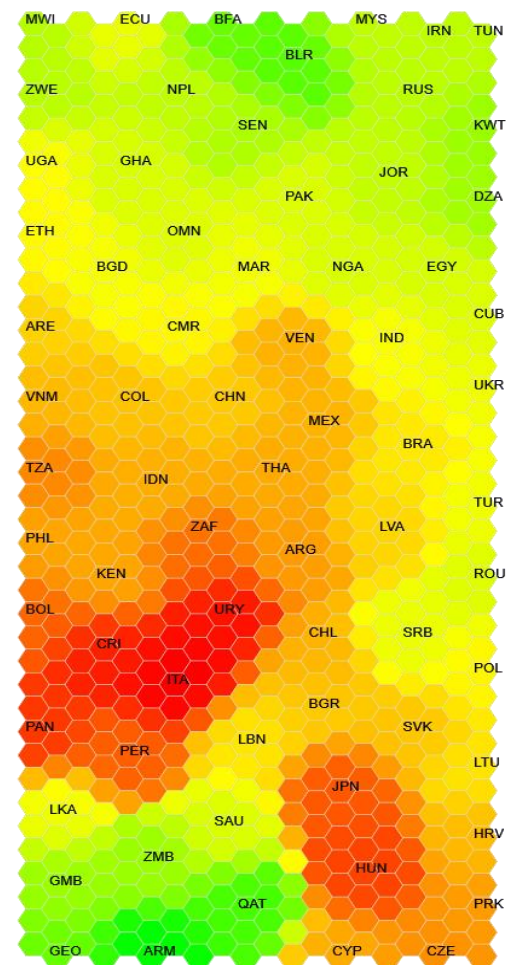


Anexo 14. Mapas de componentes de países para los indicadores Índice de desarrollo científico (SDI) e Impacto Normalizado (NI) como indicadores de tamaño independiente

a) SDI

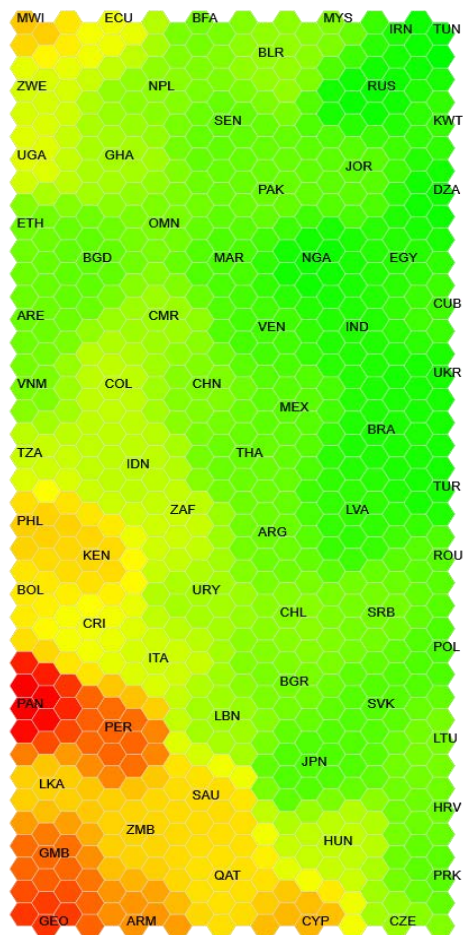


b) NI



Anexo 15. Mapas de componentes de países para los indicadores Highly Cited Papers (HCP) y Hot paper (Hot) como indicadores de tamaño independiente

c) HCP



d) Hot

