



# **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**

**Evolución del perfil cuantitativo, en el *Web of Science*, de la producción científica mexicana en matemática, durante el período 2000-2017**

**TESIS**

**QUE PARA OPTAR POR EL GRADO DE:**

**MAESTRO EN BIBLIOTECOLOGÍA Y**

**ESTUDIOS DE LA INFORMACIÓN**

**PRESENTA:**

***Lic. Rubén Sánchez Perdomo***

**ASESOR: DR. Humberto A. Carrillo Calvet**

***Facultad de Ciencias y Centro de Ciencias de la Complejidad***

***Universidad Nacional Autónoma de México***

**Ciudad de México, noviembre 2018**



Universidad Nacional  
Autónoma de México



**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.

## **Agradecimientos**

Agradecer a los sinodales: Dr. Salvador Gorbea Portal, Dra. Judith Licea Ayala, Dr. Francisco Collazo Reyes y Dra. Maricela Piña Pozas, por sus valiosos comentarios que sirvieron para enriquecer esta tesis.

A mi tutor Dr. Humberto Carrillo Calvet por su confianza y sus sabios consejos para dirigir esta investigación.

Agradezco a mi madre y mi hermana, que a pesar de la distancia siempre recibí su apoyo para lograr este sueño.

Agradezco a mi esposa por el amor, el apoyo y la comprensión en cada una de las etapas del desarrollo de esta investigación. A mi niña por entender que papi no podía jugar con ella como siempre.

A todos los chicos del Laboratorio de Dinámica No Lineal, en especial a José Luis Jiménez por su apoyo en la comprensión de las neuronas del SOM y a Ibis Lozano por su ayuda en la utilización de las herramientas del WoS.

A los profesores del Posgrado en Bibliotecología y Estudios de la Información, de la Universidad Nacional Autónoma de México por sus conocimientos impartidos durante los cursos y seminarios.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por su apoyo económico durante el período 2016-2018.

## Resumen

La presente investigación analiza el perfil cuantitativo de la producción científica de México en el campo de la matemática. La información se extrajo de la base de datos *Web of Science*, durante el período 2000-2017 (18 años). La metodología utilizada se basa en el uso de técnicas bibliométricas y las fortalezas que brinda la inteligencia artificial. El sistema de indicadores son una combinación de indicadores normalizados por categoría e independientes de tamaño, además estos cubren las dimensiones de producción, impacto, alto desempeño, excelencia y colaboración científica. Los resultados analizan la evolución del campo a nivel internacional y regional en relación con el desempeño de la producción científica de México. A través de la inteligencia neurocomputacional se identificaron los perfiles de desempeño cuantitativo de las instituciones mexicanas que producen matemática. Utilizando esta misma técnica se identificó el perfil de las comunidades científicas asociadas a cada categoría temática del campo de la matemática en el *Web of Science*. Los resultados de la investigación pretenden contribuir como instrumento estratégico para los decisores y gestores de políticas científicas.

**Palabras claves:** campo de la matemática, descubrimiento de conocimiento, minería de datos, software ViBlioSOM.

# Tabla de Contenido

<b>Introducción .....</b>	<b>I</b>
1. Exposición de motivos.....	I
2. Objetivos de la investigación .....	II
3. Hipótesis .....	III
4. Antecedentes .....	III
4.1 La matemática en México .....	V
<b>Capítulo 1. Marco Teórico .....</b>	<b>1</b>
1.1 Bibliometría y cienciometría.....	1
1.1.1 Indicadores de producción e impacto científico .....	3
1.1.2 Indicadores de productividad y excelencia científica .....	5
1.2 Descubrimiento de conocimiento en bases de datos (KDD).....	6
1.2.1 Minería de datos.....	9
1.3 Visualización de la información.....	12
1.4 Redes Neuronales Artificiales .....	15
1.4.1 Estructura de una Red Neuronal Artificial .....	16
1.4.2 Algoritmo SOM .....	18
<b>Capítulo 2. Fuentes y Métodos .....</b>	<b>19</b>
2.1 Web of Science.....	19
2.1.1 Herramientas de análisis del WoS: <i>Essential Science Indicators</i> e <i>InCites</i> .....	21
2.1.2 Esquemas de clasificación disponibles en <i>InCites</i> y el WoS .....	22
2.2 Variables e Indicadores bibliométricos de la investigación .....	26
2.3 Metodología ViBlioSOM .....	28
2.3.1 Etapas de la metodología ViBlioSOM.....	29
<b>Capítulo 3. Análisis y Discusión de Resultados.....</b>	<b>33</b>
3.1 Evolución del campo de la matemática a nivel internacional.....	33
3.1.1 Países con mayor volumen de publicaciones a nivel internacional.....	34
3.1.2 Producción en las mejores revistas .....	37
3.1.3 Evolución del volumen de citas .....	38
3.1.4 Países con mayor volumen de citas a nivel internacional.....	39
3.1.5 Colaboración internacional .....	41
3.2 Desarrollo de América Latina en el campo de las matemáticas.....	42
3.2.1 Evolución de la producción científica.....	42

3.2.2 Países con mayor volumen de publicaciones en la región .....	43
3.2.3 Evolución del indicador de impacto.....	44
3.2.4 Países con mayor impacto científico en la región .....	45
3.2.5 Colaboración internacional .....	46
3.3 Evolución de la producción matemática de México .....	47
3.3.1 Volumen de producción científica por entidad federativa.....	47
3.3.2 Evolución del volumen de <i>highly cited papers</i> .....	48
3.3.3 Evolución de la excelencia científica y alto desempeño.....	50
3.3.4 Análisis de coautoría .....	51
3.4 Instituciones mexicanas que producen matemática .....	52
3.4.1 Análisis del impacto de las instituciones .....	52
3.4.2 Porcentaje de artículos publicados en cuartil uno versus volumen de citas por documento.....	54
3.4.3 <i>Category Normalized Citation Impact</i> (CNCI) y porcentaje en Q1 .....	55
3.4.4 Porcentaje de documentos citados y CNCI .....	57
3.4.5 Análisis de la colaboración internacional y el alto desempeño.....	58
3.4.6 Análisis multiparamétrico del perfil de las instituciones mexicanas.....	59
3.5 Categorías temáticas.....	64
3.5.1 Número de documentos y citas recibidas .....	64
3.5.2 Visibilidad esperada y real de las categorías temáticas .....	65
3.5.3 Impacto en las categorías temáticas .....	66
3.5.4 Producción de alto desempeño .....	67
3.5.5 Excelencia científica e impacto normalizado.....	68
3.5.6 Análisis multiparamétrico de las categorías temáticas .....	69
<b>Consideraciones finales .....</b>	<b>75</b>
<b>Referencias.....</b>	<b>78</b>
<b>Anexos .....</b>	<b>94</b>

## Lista de figuras

Figura 1. Descripción general de los procesos que componen la metodología del KDD (Fayyad & et al., 1996).....	9
Figura 2. Crecimiento de las publicaciones sobre visualización de información en la década de los años 90 (Chen, 2002).....	13
Figura 3. Estructura de una neurona (Kornilov, 1997).....	17
Figura 4. Procesos fundamentales de la metodología ViBlioSOM. ....	29
Figura 5. Mapa de componente del indicador: porcentaje de documentos en Q1, análisis de las instituciones. ....	31
Figura 6. Clustering del análisis de las instituciones utilizando cinco indicadores: (%Q1, CNCI, %CD, %HP, %IC).....	32
Figura 7. Evolución del número de documentos publicados en el WoS.....	34
Figura 8. Países con mayor índice de desarrollo científico y su producción anual en el campo de la matemática. ....	36
Figura 9. Evolución del porcentaje de documentos en Q1.....	37
Figura 10. Evolución del impacto del Mundo y México.....	38
Figura 11. Evolución del porcentaje de documentos en colaboración internacional.....	41
Figura 12. Evolución del número de documentos publicados en el WoS para América Latina y México. ....	42
Figura 13. Países de la región con mayor número de publicaciones anuales y su índice de desarrollo científico (SDI).....	43
Figura 14. Evolución de las citas por documento del mundo, América Latina, Chile, Brasil y México. ....	44
Figura 15. Evolución del porcentaje de documentos en colaboración internacional.....	46
Figura 16. Distribución de la producción matemática de los Estados de la República Mexicana.....	47
Figura 17. Evolución del número de highly cited papers en el período. ....	49
Figura 18. Evolución de los indicadores de excelencia (%Exc) y de alto desempeño (%HP).....	50
Figura 19. Número de autores por documento publicado. ....	51

Figura 20. Análisis del porcentaje de documentos en cuartil uno (%Q1) y el volumen de citas por documento (Cr/Ndoc).....	54
Figura 21. Porcentaje de artículos en cuartil uno (%Q1) y CNCI. ....	56
Figura 22. Análisis del porcentaje de documentos citados (%CD) y el CNCI. ....	57
Figura 23. Análisis del desempeño científico (%HP) y la colaboración internacional (%IC).....	58
Figura 24. Análisis multiparamétrico de las instituciones que producen matemática.....	60
Figura 25. Número de documentos publicados y citas recibidas. ....	64
Figura 26. Porcentaje de documentos en cuartil uno (%Q1) y el alto desempeño científico (%HP) .....	65
Figura 27. Análisis del indicador CNCI y las citas por documento. ....	66
Figura 28. Desempeño científico (%HP) y CNCI de las categorías temáticas. ....	67
Figura 29. Excelencia científica (%Exc) y CNCI de las categorías temáticas. ....	68
Figura 30. Análisis multiparamétrico de las categorías temáticas que componen el campo de la matemática .....	70

## **Lista de tablas**

Tabla 1. Indicadores de la investigación. ....	26
Tabla 2. Países con mayor volumen de publicaciones a nivel internacional. .	35
Tabla 3. Países con mayor número de citas recibidas en el campo de la matemática en WoS. ....	39
Tabla 4. Impacto de los países de la región. ....	45
Tabla 5. Umbral de citas para ser considerado un artículo altamente citado.	49
Tabla 6. Instituciones mexicanas con mayor volumen de producción e impacto. .....	53

## Lista de acrónimos de instituciones mexicanas

Universidad Nacional Autónoma de México	UNAM
Instituto Politécnico Nacional	IPN
Universidad Autónoma Metropolitana	UAM
Centro de Investigación y de Estudios Avanzados	CINVESTAV
Centro de Investigación en Matemáticas	CIMAT
Instituto Tecnológico Autónomo de México	ITAM
Benemérita Universidad Autónoma de Puebla	BUAP
Universidad Autónoma de San Luis Potosí	UASLP
Universidad Autónoma del Estado de Morelos	UAEM
Universidad Michoacana de San Nicolás de Hidalgo	UMSNH
Tecnológico de Monterrey	ITESM
Universidad Autónoma de Aguascalientes	UAA
Universidad Autónoma del Estado de Hidalgo	UAEH
Universidad de Sonora	USON
Universidad de Guanajuato	UG
Universidad Autónoma de Yucatán	UADY
Universidad Américas Puebla	UDLAP
Universidad Autónoma de Nuevo León	UANL
Instituto Nacional de Astrofísica; Óptica y Electrónica	INAOE
Universidad de Colima	UCOL
Colegio de Postgraduados	COLPOS
Instituto Potosino de Investigación Científica y Tecnológica	IPICYT
Universidad de Guadalajara	UDG
Universidad Autónoma del Estado de México	UAEMex
Universidad Juárez Autónoma de Tabasco	UJAT
Universidad Autónoma de Querétaro	UAQ
Universidad Autónoma de Zacatecas	UAZ

Instituto Mexicano del Petróleo	IMP
Universidad Autónoma de Chiapas	UNACH
Centro Investigación Científica y Educación Superior de Ensenada	CICESE
Universidad Autónoma de Baja California	UABC
Universidad Autónoma de Tamaulipas	UAT
Centro Internacional de Mejoramiento de Maíz y Trigo	CIMMYT
Centro de Investigaciones en Óptica	CIO
Universidad Autónoma de Sinaloa	UAS
Universidad Autónoma de Ciudad Juárez	UACJ
Universidad Autónoma de Coahuila	UAdeC
Universidad de Anahuac	UA
Instituto Nacional de Salud Pública	INSP
Hospital General de México	HGM
Universidad Panamericana	UNPA
Instituto de Ecología	IEUNAM
Centro de Investigación y Docencia Económica	CIDE
El Colegio de la Frontera Sur	ECOSUR
Universidad Juárez del Estado de Durango	UJED
Universidad del Papaloapan	UP
Universidad Autónoma Benito Juárez de Oaxaca	UABJO
Universidad Autónoma del Carmen	UNACAR
Universidad de Quintana Roo	UQROO
Instituto Mexicano del Seguro Social	IMSS
Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán	INCMNSZ
Centro de Investigación Científica de Yucatán	CICY
Instituto Nacional de Cardiología	INCAR
Centro de Ingeniería y Desarrollo Industrial	CIDESI
Centro de Investigación y Desarrollo Tecnológico en Electroquímica	CIDETEQ

## **Lista de abreviaturas de las categorías temáticas**

MATHEMATICS

MATH

MATHEMATICS; APPLIED

MATH APP

PHYSICS; MATHEMATICAL

PHY MATH

MATHEMATICS; INTERDISCIPLINARY APPLICATIONS

MATH INTER APP

STATISTICS & PROBABILITY

STA PROB

MATHEMATICAL & COMPUTATIONAL BIOLOGY

MATH COMP BIO

SOCIAL SCIENCES; MATHEMATICAL METHODS

SS MATH METH

LOGIC

LOG

PSYCHOLOGY; MATHEMATICAL

PSY MATH

## Introducción

### 1. Exposición de motivos

La ciencia y la tecnología tienen un papel esencial en el desarrollo económico una vez que tiene un impacto directo en la calidad de vida de una sociedad. En este sentido, los gobiernos deben dictaminar políticas públicas que establezcan las prioridades en las investigaciones, en dependencia de las necesidades y capacidades del país. Generalmente los estudios cuantitativos se centran en el análisis de la actividad científica de un país o un dominio de conocimiento científico. Esta práctica relaciona salidas de investigación con indicadores cuantitativos coherentes con las inversiones para el desarrollo científico-técnico de la sociedad. A partir de este análisis se puede visualizar no sólo el estado de la ciencia de un país; sino la evolución que puede alcanzar en el tiempo en dependencia de las diferentes políticas científicas implementadas por los gobiernos.

En el diseño de estas políticas debe contarse con datos precisos que muestran el alcance de la actividad científica y tecnológica en la sociedad; teniendo en cuenta indicadores bibliométricos y cuantitativos. En la actualidad es muy común utilizar los recursos de información que ofrecen las grandes bases de datos como; *Web of Science*, *Scopus*, *Google Scholar*, entre otras. Los grandes volúmenes de información en estas bases de datos hacen complejo el análisis y comprensión de sus contenidos. Pero con el desarrollo tecnológico de las bases de datos y la evolución de las técnicas de aprendizaje inteligente desarrolladas por la comunidad de inteligencia artificial, se hizo posible la recopilación masiva de datos y su interpretación; reto que asumió la metodología, *Knowledge Discovery in Databases* (KDD).

Uno de los modelos más utilizados para el descubrimiento de conocimiento en bases de datos son los referidos a las Redes Neuronales Artificiales (RNA), las cuales son capaces de agrupar elementos por patrones de similitud. Uno de estos modelos que cobra especial interés por su capacidad

de clasificación es: el modelo basado en el algoritmo SOM (*Self Organizing Maps*).

En esta investigación, se aprovecha el potencial que tiene la inteligencia artificial para hacer minería de datos y el potencial que tienen los indicadores de la Bibliometría Evaluativa para descubrir conocimiento (por medios automáticos) relativo a la evolución del campo de la matemática en México. El estudio se orienta a partir de las siguientes preguntas de investigación:

1. ¿Cómo ha evolucionado el perfil cientimétrico de la Matemática mexicana; en el *Web of Science* durante el período 2000-2017?
2. ¿Cómo se compara el perfil cientimétrico de México en el ámbito regional e internacional?
3. ¿Cómo se compara el perfil cientimétrico de las instituciones mexicanas que producen investigaciones en el campo de la Matemática?
4. ¿Cómo se comparan desde el enfoque cientimétrico las principales categorías temáticas que estructuran el campo de la Matemática?

## 2. Objetivos de la investigación

### Objetivo general

Determinar el comportamiento cientimétrico de la evolución de la producción científica mexicana en el campo de la matemática, durante el período 2000-2017, utilizando los datos y los recursos bibliométricos del *Web of Science*.

### Objetivos específicos

1. Caracterizar el perfil cientimétrico de la producción científica mexicana en el campo de la matemática, así como el comportamiento de su desarrollo durante el período 2000-2017.
2. Comparar la producción científica mexicana en el campo de la matemática con la producción del mundo y con otros países de la región latinoamericana.
3. Identificar las instituciones mexicanas con mayor liderazgo en el campo de la matemática.

4. Describir el perfil cuantitativo de las principales categorías temáticas del campo de la matemática en México.

### 3. Hipótesis

Los datos y los recursos del WoS, los indicadores de la Bibliometría Evaluativa y la minería de datos con inteligencia computacional puede ser usadas para revelar aspectos significativos del desarrollo del campo de la Matemática en México, durante el período 2000-2017.

### 4. Antecedentes

En la revisión bibliográfica se encontraron diversos trabajos bibliométricos que analizan el comportamiento de la ciencia mexicana de manera general: (Arvanitis & et al., 1996; Collazo-Reyes & et al., 2011; Laclette & Zúñiga-Bello, 2011; Luna-Morales, 2012; Acosta-Ramírez & et al., 2015; Restrepo-Arango & Urbizagastegui-Alvarado, 2016; Lopez-Olmedo & et al., 2017). Otros investigadores han estudiado el desarrollo científico de las instituciones y universidades (Delgado & Russell, 1992; Lima & et al., 2005; Russell & et al., 2012; Luna-Morales & et al., 2012; Tarango & et al., 2015; Tarango & Machin-Mastromatteo, 2016; UNAM, 2017).

Otros estudios se enfocan en el análisis de las revistas de la corriente principal donde aparecen las publicaciones de los investigadores mexicanos (Luna-Morales & et al., 2009; Arencibia-Jorge & et al., 2016). Otra investigación importante para el gremio científico mexicano es el Atlas de la Ciencia Mexicana (CINVESTAV, 2017) que analiza la producción científica del país a partir de las bases de datos del Sistema Nacional de Investigadores, el *Web of Science* y *Scopus*.

También se han realizado estudios regionales que describen el comportamiento bibliométrico de algunas características de la producción científica de México (Koljatic & Silva, 2001; Macías-Chapula, 2002; Alfaraz & Calvino, 2004; Garreton & et al., 2005; Collazo-Reyes & et al., 2008; Michán & et al., 2008; Romero & et al., 2009; Alonso-Gamboa & Russell, 2012; Murphy, 2013; Collazo-Reyes, 2014; Chinchilla-Rodríguez & et al., 2015; Pinto & et al.,

2015; Schadl & Todeschini, 2015; Sánchez-Perdomo & et al., 2017; Collazo-Reyes & et al., 2017).

Existen otros trabajos que se enfocan en el análisis de las diferentes áreas del conocimiento en México. En el área de las Ciencias Naturales e Ingenierías (Mariscal, 2002; Collazo-Reyes & et al., 2004; Kostoff & et al., 2005; Sierra-Flores & Russell, 2009; Sierra-Flores & et al., 2009; Collazo-Reyes & et al., 2010; Rojas-Sola & Jorda-Albiñana, 2011; Hernández & et al., 2013; Collazo-Reyes & et al., 2014; Uddin & et al., 2015; Hernandez-Garcia & et al., 2016; Russell & et al., 2016); en el área de las Ciencias de la Vida y de la Tierra (Lena, 1997; Licea de Arenas & et al., 2003; Bravo-Vinaja & Sanz-Casado, 2008; Solano & Orihuela, 2010; Michán, 2010; Longar & Ríos, 2014); en el área de las Ciencias Sociales (Licea de Arenas & et al., 2000; Torres, 2009; Martínez Arellano & Ruíz Vaca, 2011; Bonilla & et al., 2015, Restrepo, 2015; Sánchez & et al., 2015) y en el área de las Ciencias Médicas (Licea de Arenas & et al., 2002; García-Silberman & et al., 2004; Macías-Chapula & et al., 2007; Macías-Chapula, 2013; Michán & Muñoz-Velasco, 2013; Castillo-Pérez & et al., 2015).

La principal contribución que persigue el presente estudio es descubrir un nuevo conocimiento a partir de la utilización de la minería de datos y la inteligencia computacional, basándose en el algoritmo del modelo SOM (*Self Organizing Maps*) de redes neuronales (Kohonen, 2013).

Entre las aplicaciones pioneras del modelo SOM en la Bibliometría se encuentran los trabajos resultantes de la colaboración cubano-mexicana del grupo de los profesores Gilberto Sotolongo y María Victoria Guzmán con el grupo del Dr. Humberto Carrillo en la UNAM: (Sotolongo-Aguilar & et al., 2001, 2002) y para la cartografía de la ciencia y la tecnología (Polanco & et al., 2001; Moya-Anegón & et al., 2006). Estudios más recientes se han aplicado en la evolución de vacunas contra la Tuberculosis (Guzmán & et al., 2010) y el mapeo de la estructura temática de las ciencias médicas (Skupin & et al. 2013).

Particularmente en México se han realizado estudios para la identificación del núcleo de revistas mexicanas (Arencibia-Jorge & et al., 2016) y el análisis multiparamétrico del perfil cientométrico en las instituciones de educación superior mexicanas (Villaseñor & et al., 2016). Todos los estudios

anteriormente mencionados utilizaron diferentes técnicas bibliométricas para analizar la producción científica mexicana, pero la presente investigación utiliza nuevos indicadores de alto impacto y aprovecha las fortalezas que ofrece la inteligencia neuro-computacional. Con el objetivo de representar el perfil cienciométrico de la producción científica mexicana en el campo de la matemática en el *Web of Science*.

#### **4.1 La matemática en México**

La comunidad de investigadores en matemáticas se destaca por integrar profesionales altamente especializados. El pasado siglo XX fue muy importante en cuanto al crecimiento del número de investigadores matemáticos, esto se debe en gran medida a la institucionalización de la carrera de matemática en muchas universidades del mundo.

Particularmente en México la institucionalización del campo de la matemática fue en el año 1637, cuando se funda la Real y Pontificia Universidad de México, donde se crea la cátedra de Astrología y Matemáticas. A inicios del siglo XX las instituciones impulsoras del desarrollo del campo de la matemática en México fueron la Escuela Nacional de Altos Estudios y la Universidad Nacional de México, ambas fundadas en el año 1910. Una década después de haber adquirido la Universidad Nacional de México su autonomía, en el año 1939 se crea la Facultad de Ciencias de la UNAM y años más tarde, en el 1942, esta institución funda el Instituto de Matemáticas y organiza ese mismo año el 1er Congreso Nacional de Matemáticas, un evento sumamente importante para reunir a la comunidad académica de las matemáticas en México. Este evento fomenta en el año 1943 la creación de la Sociedad Matemática Mexicana. Desde entonces han sido muchas las universidades que han abierto programas de pregrado y posgrado en el país, con el objetivo de formar jóvenes profesores e investigadores en el campo.

A la par del desarrollo institucional del campo de la matemática en México, en el ámbito mundial se evidencia un crecimiento en las investigaciones sobre el campo, distintos han sido los estudios con carácter bibliométrico para identificar características propias del campo. Un estudio bibliométrico muy interesante que analiza el campo de la matemática en el

ámbito mundial en un período de 140 años (1868-2008) evidencia este comportamiento, sus autores identifican la tasa de crecimiento de las publicaciones; mediante el modelo matemático de Bradford y el crecimiento de la comunidad de autores que publican en el campo de la matemática, basándose en el modelo matemático de Lotka (Behrens & Luksch, 2011).

Otras investigaciones que identificaron el comportamiento de la matemática en algunos países son: la cooperación internacional de la investigación matemática de China (Dang & Zhang, 2003; Zhou & Tian, 2014); la investigación matemática de la India (Arunachalam, 2001); la producción científica brasileña en el campo de las matemáticas (Castanha & Gracio, 2013); el efecto de la guerra en la producción científica de la matemática en Croacia (Dravec, 2012); el análisis bibliométrico del impacto de tres revistas chinas en matemáticas (Zhou & Leydesdorff, 2007) y la evaluación del prestigio de la investigación en matemática a partir del factor de impacto (Ferrer-Sapena & et al., 2016).

Existen otras investigaciones que analizan un estrato más especializado del campo, como es el caso del estudio de la producción científica de los autores en lógica matemática en el ámbito mundial (Huber & Wagner-Döbler, 2001) y el análisis de las citas que reciben los matemáticos galardonados, mediante la concesión de importantes premios: becas Guggenheim (2001-2010), becas de Sloan (2001-2005) y las becas CAREER de la *National Science Foundation* durante 2001-2005 (Smolinsky & Lercher, 2012).

En el caso particular del campo de la matemática en México se destacan algunas contribuciones que describen muy bien el desarrollo del campo desde inicios del siglo pasado y las perspectivas futuras del desarrollo del campo, considerando estos años como la etapa de formación de la carrera en el país (Adem, 1991; Gorostiza, 1991; Lomnitz & Cházaro, 1999). Otro estudio más reciente analiza el comportamiento en la búsqueda de información de la comunidad de los matemáticos de México, utilizan como instrumento de investigación el cuestionario y manejan variables como: nivel de estudios, edad, antigüedad laboral, categoría y pertenencia al SNI contra la fuente y el recurso más utilizado por los investigadores (Guevara-Villanueva, 2007). Otro

interesante estudio analiza el desempeño científico desarrollado por 64 mexicanos que obtuvieron su Ph.D. en universidades americanas a partir del año 1980 a 1998. Como resultado fundamental demostraron que los matemáticos que trabajaban en el extranjero tuvieron un mejor desempeño que sus colegas en México (González & et al., 2003).

Al considerar los antecedentes planteados, esta investigación pretende formar parte del acervo documental de los estudios cuantitativos aplicados al análisis del comportamiento de la actividad científica mexicana, utilizando la minería de datos y la inteligencia computacional, pudiendo contribuir como instrumento estratégico para los decisores y gestores de políticas científicas.

La investigación está estructurada en tres capítulos y una sección introductoria que aporta elementos como: exposición de motivos, objetivos, hipótesis y antecedentes de la investigación. El Capítulo 1 aborda elementos conceptuales sobre el análisis bibliométrico, el descubrimiento de conocimiento en bases de datos, la visualización de la información y las redes neuronales artificiales. El Capítulo 2 describe la metodología utilizada en el estudio y elementos como: las fuentes de información, las herramientas bibliométricas, las variables y la batería de indicadores utilizados. El Capítulo 3 presenta los resultados obtenidos en la investigación dividido en cinco secciones. Las dos primeras secciones muestran la evolución del campo de la matemática a nivel internacional y el desarrollo de América Latina. Luego le sigue la sección donde se identifica la distribución de la producción matemática en México. La cuarta sección está dedicada a las instituciones mexicanas que producen matemática. Y en la última sección se muestra el desarrollo de las categorías temáticas que componen el campo de la matemática en el WoS. Por último, se exponen los aspectos más relevantes identificados en el estudio en forma de consideraciones finales.

## Capítulo 1. Marco Teórico

### 1.1 Bibliometría y cienciometría

La ciencia contemporánea se ha caracterizado en gran medida por la penetración de métodos y modelos matemáticos en todas las esferas del conocimiento. Este fenómeno se le ha denominado como la matematización del conocimiento científico (Gorbea-Portal, 1998). La aplicación de estos métodos matemáticos a las colecciones y fondos de las bibliotecas trajo consigo el surgimiento de varias especialidades métricas de la información.

La Bibliometría o *Bibliometrics*, por ejemplo, es definida por Alan Pritchard como la aplicación de los métodos matemáticos y estadísticos a los procesos de comunicación escrita y a la naturaleza en curso de un desarrollo dentro de una disciplina. (Pritchard 1969). Se reconoce en la literatura que anterior a esta definición Paul Otlet, en el año 1934, en su obra el “*Traité de documentation*”, definió la “*bibliometrie*” como la parte definida de la bibliología que se ocupa de la medida o cantidad aplicada a los libros (Otlet, 1934); esta definición no tuvo mucha repercusión en la comunidad científica.

Otra especialidad métrica como la cienciometría, alcanzó un alto grado de especialización por el año 1979, su origen está asociado por la creación en ese año de la revista *Scientometrics* y por la convergencia de dos movimientos científicos: uno americano, centrado en el estudio de la ciencia de la ciencia; y el otro en la entonces Unión Soviética (URSS). Este último movimiento denominado como *Naukovodemia* (Cienciología), es dirigido por dos escuelas soviéticas bajo las figuras de Nalimov y Dobrov. El primero en colaboración con Mulchenko definen a la cienciometría (*Naukometriia*) como la aplicación de los métodos cuantitativos a la investigación sobre el desarrollo de la ciencia como un proceso de información (Nalimov & Mulchenko, 1969). El movimiento norteamericano estuvo impulsado por la figura de Derek John de Solla Price, quien en su libro “*Little Science, Big Science*” (Pequeña Ciencia, Gran Ciencia) expone por primera vez un modelo de crecimiento exponencial de la ciencia, el

cual debe ser tomado en cuenta como premisa para el estudio de su desarrollo (Price, 1963).

A finales del pasado siglo se produjeron varios desarrollos, primero, las bases de datos de citas como el *Web of Science*, *Scopus* y *Google Scholar* comenzaron a ampliar enormemente su cobertura. Si bien esta extensión de la cobertura es valiosa, también conduce a problemas de comparación con resultados bastante diferentes que aparecen en función de las bases de datos utilizadas (Mingers & Leydesdorff, 2015). En segundo lugar, surgen nuevas métricas basados en indicadores de citas y evaluación de revistas (Cole, 2000; Van Raan, 2004). El tercer desarrollo técnico va a estar enfocado en el mapeo y visualización de las redes bibliométricas (Börner, 2003; Chen, 2017).

La evaluación de los resultados de investigación puede ser llevada a cabo de dos formas distintas: la primera, la evaluación por pares como medio de la valoración científica, medido en concursos para investigadores, proyectos de investigación y publicación de artículos en revistas científicas; la segunda, evaluación de la comunidad científica a partir de los resultados de investigación publicados (Van Raan, 1996). El proceso de evaluación con mayor tradición en la comunidad científica internacional es (*peer review*) la revisión por pares (Suárez-Balseiro & Maura-Sardó, 2005). Sin embargo, desde la segunda mitad del pasado siglo XX, los estudios bibliométricos y de evaluación de la ciencia han cobrado especial interés en los países desarrollados (Kostoff, 1995; Warner, 2000; Bence & Oppenheim, 2004). Lo anterior y otros factores ha favorecido el desarrollo de las especialidades métricas de la información y ha propiciado la evolución de los indicadores.

Los indicadores constituyen una de las herramientas más utilizadas para la medición del producto de la investigación científica, ya que la documentación (independientemente del tipo de soporte) es el vehículo más prolífico y exitoso para la transferencia del conocimiento científico (Russell, 2004). Son parámetros que se utilizan en el análisis estadístico y evaluativo de los rasgos cuantificables de la literatura científica, han demostrado ser herramientas útiles en la evaluación del rendimiento de la investigación en muchos campos científicos (Sancho, 1992; Korevaar & Moed, 1996; Guzmán-Sánchez &

Sotolongo-Aguilar, 2002; Maltrás-Barba, 2003). Se ha debatido mucho en la literatura sobre las ventajas y limitaciones que poseen los indicadores (Debackere & Glanzel, 2004; Geisler, 2005; Kostoff, 2001; Nederhof, 2006; Snizek, 1995), sin embargo, constituyen herramientas clave en la gestión de la política científica y tecnológica, y en los procesos de toma de decisiones estratégicas (Arencibia, 2010).

Los indicadores bibliométricos según la opinión de varios autores pueden dividirse en dos grandes grupos: los que miden la calidad y el impacto de las publicaciones científicas, considerados igualmente como indicadores de actividad. Estos proporcionan información sobre el volumen y el impacto de las actividades de investigación mediante recuentos de variables bibliográficas (tales como autores, artículos, palabras clave, patentes, citas, entre otros) (Spinak, 2001; Vinkler, 2006; Arencibia, 2010). El otro grupo, son los indicadores que miden el impacto de las relaciones entre las publicaciones científicas, considerados también como indicadores relacionales de primera, segunda y tercera generación. Los indicadores relacionales se proponen conocer los vínculos y las interacciones entre las diferentes variables bibliográficas mediante los conceptos de cocitación y coocurrencia, intentando describir el contenido de las actividades y su evolución (Spinak, 2001; Guzmán-Sánchez & Sotolongo-Aguilar, 2002; Bailón-Moreno & et al., 2005; Vinkler, 2006; Arencibia, 2010). Partiendo de estas propuestas y basado en la batería de indicadores seleccionados en la investigación en el siguiente acápite se analizan los indicadores bibliométricos desde dos perspectivas distintas: indicadores de producción e impacto científico y los indicadores de productividad y excelencia científica.

### **1.1.1 Indicadores de producción e impacto científico**

Los indicadores de producción científica tienen una función descriptiva y son empleados fundamentalmente para el análisis de aspectos cuantitativos del volumen del sistema científico objeto de estudio (Martí, 2011). Su objetivo es el cómputo de las publicaciones de un agente, considerándose como publicaciones los documentos propagados a través de canales formales y públicos (Sancho, 2001). Por ejemplo, cuántos artículos publicados tiene

México en el campo de la matemática, cuántos artículos tienen publicados los centros que producen matemática en el país.

El volumen de la producción científica es un indicador dependiente de tamaño; en qué sentido, en el caso de una institución, mientras mayor sea la comunidad de investigadores es de esperar que mayor sea su producción científica. Igual sucede con los países, mientras mayor sea la comunidad científica de ese país o el número de habitantes, mayor será su producción científica, es por ese motivo que para valorar estos aspectos en la investigación se utilizaron en algunos casos indicadores independientes de tamaño que fortalece y hace más justo los análisis.

Los índices de citas del *Institute for Scientific Information* (ISI), impulsados por el trabajo de Eugene Garfield (1925-2017), fueron el camino para estudiar a fondo la comunicación científica, utilizando las más diversas perspectivas, y generando líneas de investigación que hoy constituyen pilares de la Ciencia de la Información (Arencibia, 2010). Los índices de citas ofrecen una dimensión de la calidad de la investigación diferente a la que tradicionalmente brinda el juicio de un experto (*peer review*); sin embargo, no implica que no exista correlación entre los análisis que resultan de ambos métodos, de hecho, esta correlación ha sido demostrada por numerosos autores en el ámbito internacional (Aksnes & Taxt, 2004; Rinia & et al., 1998; So, 1998).

Se ha debatido mucho en la literatura sobre las motivaciones para citar (Ahmed & et al., 2004; Baldi & Hargens, 1995; Case & Higgins, 2000; Cosijn & Ingwersen, 2000; Kapseon, 2004; Moravcsik & Murugesan, 1975; Vinkler, 1987; White, 2004; Nicolaisen, 2007; Cañedo-Andalia, 1999). Algunos autores consideran las citas como moneda de pago (Kaplan, 1965), como proceso cognitivo (Harter, 1992), como acto persuasivo (Latour & Woolgar, 1986) o como concepto simbólico (Small, 1978). La única razón del uso de los conteos de citas para la evaluación de investigadores es que brinda una medida de la utilidad y el impacto del trabajo científico (Garfield, 1979). Este concepto de Eugene Garfield es adecuado para los fines de esta investigación, ya que los indicadores de citas utilizados muestran el impacto real que producen los

trabajos de los investigadores mexicanos en la comunidad científica internacional.

### 1.1.2 Indicadores de productividad y excelencia científica

En los estudios bibliométricos cuando se utilizan indicadores para describir una unidad de análisis, generalmente se utilizan indicadores que miden el volumen o tamaño de su desempeño científico (indicadores de producción científica). Pero no se consideran factores como cuán eficiente o eficaz es ese volumen de producción científica. En este sentido, en el Laboratorio de Dinámica No Lineal de la Facultad de Ciencias de la UNAM, existe un grupo de cientiométricos liderados por el Dr. Carrillo que comienzan a introducir este tipo de factores independientes de tamaño en sus investigaciones (Villaseñor & et al., 2016; Lozano, 2016). Esta investigación utiliza estos indicadores de productividad buscando medir la eficiencia y la eficacia de la producción científica, poniendo en la balanza la producción frente a la capacidad de producción.

Para el desarrollo de estudios cientiométricos, es deseable poder cuantificar de alguna manera el grado de excelencia que corresponde a las investigaciones realizadas. Los indicadores de excelencia científica son utilizados en algunos casos para construir *rankings* sobre las instituciones (Waltman & et al., 2012; Bornmann & et al., 2014), para evaluar las áreas científicas con mayor excelencia (Bornmann & Leydesdorff, 2012) y para otorgar premios (Rodríguez-Navarro, 2011). El impacto de los documentos de excelencia científica da una idea de la calidad de la investigación y cómo esta repercute en la comunidad científica internacional (Bornmann, 2014). Si es cierto que generalmente los artículos de excelencia científica se escriben en colaboración múltiple y es un factor que favorece la visibilidad de la investigación (Aksnes, 2003). En esta investigación se utilizan con el objetivo de determinar la élite del campo de la matemática mexicana, tanto de las instituciones que producen matemática, como de las comunidades científicas por cada una de las categorías temáticas que identifican el campo en el WoS.

## 1.2 Descubrimiento de conocimiento en bases de datos (KDD)

En la década de los 80 las tecnologías de las bases de datos tuvieron gran desarrollo y evolución. A la par de esto, se fue registrando un gran crecimiento en los volúmenes de datos digitales, creando una sobrecarga de información. Este fenómeno trajo consigo un cambio en el enfoque que tenían los datos, dejaron de ser objeto de almacenamiento, búsqueda y recuperación de información, para ser concebidos como una valiosa fuente de conocimiento e inteligencia (Gorbea-Portal, 2013).

En consecuencia, a este crecimiento y el nuevo valor que adquieren los datos se crean muchos algoritmos y reglas matemáticas para relacionar y analizar los grandes volúmenes de datos. Es precisamente en esta nueva forma de análisis que surge la metodología denominada como el descubrimiento de conocimiento en bases de datos (*Knowledge Discovery in Databases*), o *KDD* en sus siglas en inglés, su denominación y concepto fueron acuñados por primera vez en el primer taller de *KDD* en el año 1989, haciendo énfasis en que el conocimiento es el producto final de un descubrimiento basado en los datos (Matheus & et al., 1993).

El descubrimiento de patrones útiles en grandes volúmenes de datos en la literatura ha sido descrito con diferentes nombres: extracción de conocimiento, descubrimiento de información, recolección de información, arqueología de datos, procesamiento de patrones de datos, análisis exploratorio de datos y análisis inteligente de datos (Tukey, 1977; Fayyad & et al., 1996; Berthold & Hand; 2000). Para los efectos de esta investigación la denominación que se considerará será “descubrimiento de conocimiento o *KDD*”.

La definición más aceptada por la comunidad científica es la propuesta por (Fayyad & et al., 1996) que definen el *KDD* como: *procesos no comunes de identificación de patrones válidos, novedoso, potencialmente útiles y finalmente comprensibles en los datos*. En correspondencia con esta definición otros autores también consideran el *KDD* como un proceso (Fayyad, 1997; Sang & Siau, 2001; Han & Kamber, 2006). En esta investigación se adopta la postura anterior, definida por (Fayyad & et al., 1996) considerando el *KDD* como una

metodología conformada por varios procesos, como, por ejemplo, la limpieza y normalización de los datos, la construcción del *Data Warehouse*, la minería de datos, entre otros que sirven de guía para la elaboración de síntesis, ideas y conocimientos. Se alimenta de grandes volúmenes de datos multidimensionales para lograr un conocimiento adicional. Es un conjunto de procesos complejos que requiere la aplicación de técnicas computacionales de alto desempeño, principalmente desarrolladas dentro de los ámbitos del aprendizaje automático (*machine learning*). Una de sus características fundamentales es que tiene que poseer interactividad e iteración. Es decir, que cada proceso se puede repetir varias veces, e interactivo porque en cada paso o proceso (se interactúa con el resto de los procesos) puedes modificar los parámetros (Matheus & et al., 1993; Fayyad & et al., 1996; Dunkel & et al., 1997; Sang & Siau, 2001; Villaseñor, 2004).

A su vez, la metodología *KDD*, es un campo de la inteligencia artificial de rápido crecimiento, que combina técnicas del aprendizaje de máquina, reconocimiento de patrones, estadística, bases de datos, y visualización de información para de forma automática extraer conocimiento (Fayyad, 1997). En esencia, la metodología del *KDD* se utiliza para la interpretación de grandes cantidades de datos y la identificación de relaciones o patrones de conocimiento significativo en los datos. Para lograr este objetivo hacen falta técnicas de aprendizaje automático (Michalski & et al., 1998), estadística (Mitchell, 1997), bases de datos, técnicas de representación del conocimiento, razonamiento basado en casos, adquisición de conocimiento, redes neuronales artificiales (Kaski, 1997) y visualización de datos (Börner, 2005).

En un intento para describir la metodología del *KDD* los autores (Fayyad & et al., 1996; Han & Kamber, 2006) proponen las etapas o procesos siguientes:

- (1) Comprensión del dominio de aplicación: desarrollar una comprensión del dominio de la aplicación y un conocimiento previo pertinente, luego identificar el objetivo de la metodología del *KDD*, a partir de lo que se quiere lograr.
- (2) Selección, limpieza y preprocesamiento de los datos: seleccionar un conjunto de datos, o centrarse en un subconjunto de variables o muestras

de datos, en el que se realizará el descubrimiento. En la limpieza y preprocesamiento de datos se hace necesario eliminar el ruido si corresponde, recopilar la información necesaria para modelar o dar cuenta del ruido, decidir estrategias para manejar los campos de datos faltantes y tomar en cuenta la información de la secuencia temporal y los cambios conocidos.

- (3) Reducción y proyección de datos: con la reducción de la dimensionalidad o los métodos de transformación, se puede reducir el número efectivo de variables consideradas, o se pueden encontrar representaciones invariables para los datos.
- (4) Selección del método de minería de datos: según los objetivos propuestos en la etapa uno, se debe decidir qué método de minería de datos utilizar. Por ejemplo, el resumen, la clasificación, la regresión, la agrupación, entre otros.
- (5) Selección y ejecución del algoritmo: se debe decidir qué modelos y parámetros pueden ser apropiados (por ejemplo, los modelos de datos categóricos son diferentes a los modelos de vectores sobre los reales) y hacer coincidir un método de minería de datos particular con los criterios generales de la metodología del *KDD*.
- (6) Minería de datos: es la búsqueda de patrones de interés en una forma de representación particular o un conjunto de tales representaciones, incluyendo reglas de clasificación o árboles, regresión y agrupamiento. La etapa o proceso fundamental dentro de la metodología del *KDD* es la minería de datos (Hätönen & et al., 1996; Sang & Siau, 2001).
- (7) Interpretación de patrones minados: este paso puede implicar la visualización de los patrones y modelos extraídos o la visualización de los datos dados los modelos extraídos. Es la etapa final, donde se interpretan los resultados obtenidos y se vislumbra un nuevo conocimiento.

En la figura 1 se muestran las etapas o procesos que comprende la metodología del *KDD*, cabe mencionar que no es un ciclo lineal, ya que su capacidad de iteración e interacción permite repetir los procesos y modificar varios parámetros.

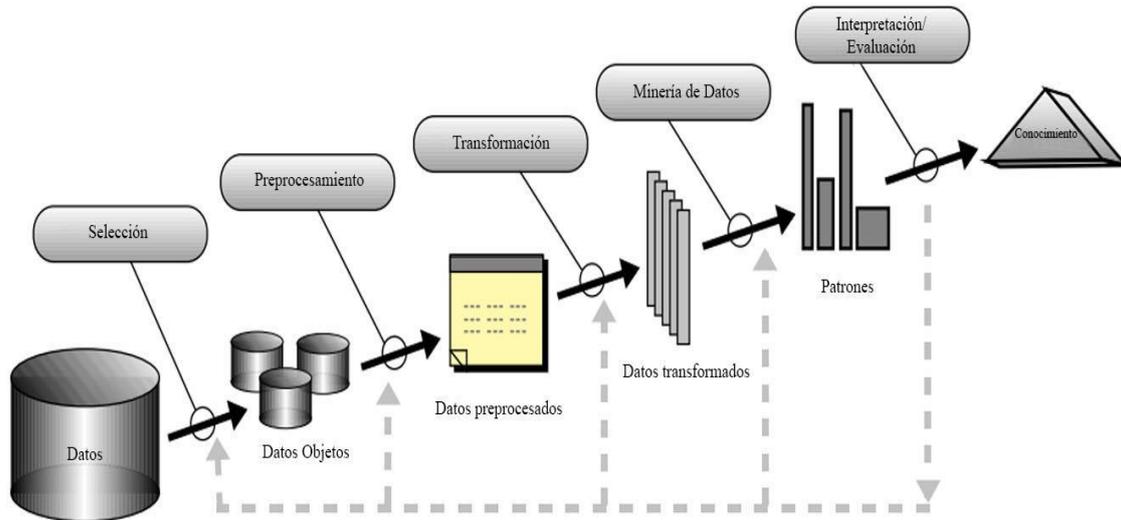


Figura 1. Descripción general de los procesos que componen la metodología del *KDD* (Fayyad & et al., 1996)

Cada una de las etapas del proceso *KDD* son esenciales para garantizar que el conocimiento útil se derive de los datos, desde la fase de preparación y selección de los datos, hasta la fase de minería de datos e interpretación de resultados. En este conjunto de procesos interviene la creación de una estructura de datos, el cual se denominada en la literatura como *Data Warehouse* (almacén de datos). Este comprende un repositorio de datos que han sido separados y desheredados de las bases de datos operacionales de un sistema y contiene una versión limpia, normalizada y anónima de los datos formateados específicamente para su análisis (Gorbea-Portal, 2013).

### 1.2.1 Minería de datos

Como se pudo identificar anteriormente uno de los procesos más relevante de la metodología del *KDD* es la minería de datos. El término minería de datos ha sido utilizado principalmente por los estadísticos, los analistas de datos y las comunidades de sistemas de información gerencial (*MIS*, siglas en inglés). Las bases para el desarrollo de la minería de datos emergen de áreas como: la Estadística, los Sistemas y la Estructura de Base de Datos, el Aprendizaje Automático, la Visualización de Datos, la Búsqueda y Recuperación de la Información, los Modelos de Redes Neuronales, el

Reconocimiento de Patrones y el Análisis de Datos Espaciales (Han & Lamber, 2006).

La minería de datos es un proceso de gran importancia en la actualidad para la planeación estratégica y la toma de decisiones. Los avances tecnológicos y el desarrollo del mundo digital (aumento en la capacidad de memoria y almacenamiento), han propiciado una evolución constante en las capacidades de cómputo y procesamiento de datos. En consecuencia, los análisis pueden contener cada vez más un mayor volumen de datos y operar en espacios de una dimensión alta (Villaseñor, 2004). Las grandes colecciones de datos poseen información valiosa, el proceso de búsqueda y análisis de los datos es complejo, más si el análisis se realiza en un espacio de más de tres dimensiones. (Keim & Kriegel, 1994). Con la integración de métodos estadísticos con métodos de aprendizaje de máquina y la utilización de redes neuronales, es posible realizar un análisis exploratorio de datos más complejo, trabajando en un espacio de más de tres dimensiones (Mannila, 1996).

Dentro de la metodología del *KDD*, en el proceso de la minería de datos se realizan distintas consultas y extracción de información previamente desconocida, válida y procesable desde grandes bases de datos. Para la exploración y análisis de los datos se utilizan medios automáticos y semiautomáticos para descubrir patrones, reglas significativas y tendencias previamente desconocidas. Su aplicación puede hacerse con un carácter retrospectivo (histórico), prospectivo (proyecciones) o comprensivo (entender lo que ocurre) y su objetivo final es aportar resultados para la toma de decisiones (Berry & Linoff, 1997; Cabena & et al., 1998; Thuraisingham, 1999). En este proceso se aplican diferentes técnicas para extraer los distintos patrones de datos, los cuales se analizan en la fase final del proceso.

#### **1.2.1.1 Técnicas de la minería de datos**

Las técnicas de la minería de datos persiguen dos objetivos fundamentales, la predicción y la descripción. La predicción implica el uso de algunas variables en la base de datos para predecir valores desconocidos o futuros de otras variables de interés. Sin embargo, la descripción se centra en encontrar patrones interpretables por humanos que describen los datos

(Fayyad & et al., 1996; Valcárcel, 2004). Este último objetivo, la descripción, es el que se aplica en esta investigación, identificando los patrones comunes que posee el campo de la matemática en México.

Existen disímiles técnicas para la aplicación de la minería de datos, entre estas podemos mencionar, la clasificación (*classification*), la regresión (*regression*), el modelo predictivo (*predictive modeling*), el análisis de secuencias (*analysis of sequences*), el agrupamiento (*clustering*), entre otras. Para esta investigación la técnica utilizada fue el agrupamiento, más conocido en la literatura por *clustering*.

Esta técnica es muy utilizada para la minería de datos, la cual tiene por objetivo identificar y clasificar grupos en un conjunto de datos, de tal manera que elementos asignados a un mismo grupo o clase tengan características similares entre sí, mientras que elementos pertenecientes a grupos distintos sean disímiles. Otro objetivo que persigue esta técnica es, reducir la cantidad de datos mediante la caracterización o agrupamiento de datos con características similares. Esta agrupación está acorde con los procesos humanos de información y una de las motivaciones para usar algoritmos *clustering* es proveer herramientas automáticas que ayudan a la construcción de taxonomías (Villaseñor, 2004; Kostoff & et al., 2007; Villaseñor, 2016).

De manera general, el principal problema que enfrentan los métodos de *clustering* es la interpretación de los *clusters*. La mayoría de los algoritmos de *clustering* prefieren cierto tipo de *cluster* y los algoritmos siempre asignan los datos a estos tipos de *cluster*, aun cuando los datos dentro del *clustering* no tengan la similitud esperada. Por lo tanto, se advierte esta situación a la hora de realizar inferencias e interpretaciones acerca de la estructura del *cluster* (Villaseñor, 2004). Una vez obtenido los *clusters* podemos identificar los perfiles de agrupamiento que presentaron los datos, según el contexto de aplicación se pueden identificar grupos de: investigadores e instituciones con similar producción científica, usuarios, clientes, pacientes de hospitales, etc. (Maimon & Rokach, 2010).

### 1.3 Visualización de la información

La investigación en el campo de la visualización de la información tiene su génesis a mediados de la década del 80, con los primeros estudios sobre interfaces de usuario y factores humanos. Unos años más tarde en los años 90 comienzan a publicarse trabajos sobre visualización de datos. A partir del año 1999 comienzan a interesar temas como la interactividad y la jerarquía, siendo esos años donde se fortalece la visualización de la información como campo distintivo.

Su desarrollo está impulsado en gran medida por los avances de la Informática, la Inteligencia Artificial y la *Human Computer Interaction (HCI)*. El principio fundamental de la visualización de la información es contribuir como una disciplina instrumental para identificar patrones no evidentes, correlaciones o agrupamientos de un gran volumen de información compleja (Zhu & Chen, 2005; Torres, 2009). También se utiliza para revelar características de un fenómeno abstracto al transformar datos en formas visuales (Brachman & Anand, 1996). La intención en todo momento es optimizar el uso de nuestra percepción y la habilidad de nuestro pensamiento para tratar con fenómenos que por sí solos no pueden ser representados visualmente en un espacio bidimensional (Bertin, 1999).

La investigación alrededor de la visualización de la información es amplia en la medida en que se ha convertido en un campo de estudio que ha crecido exponencialmente en los últimos años (Chen, 2002). Ejemplo de la anterior afirmación es una gráfica publicada en su artículo, donde se evidencia un elevado crecimiento en los artículos sobre el tema de la visualización de la información a partir de los años 90.

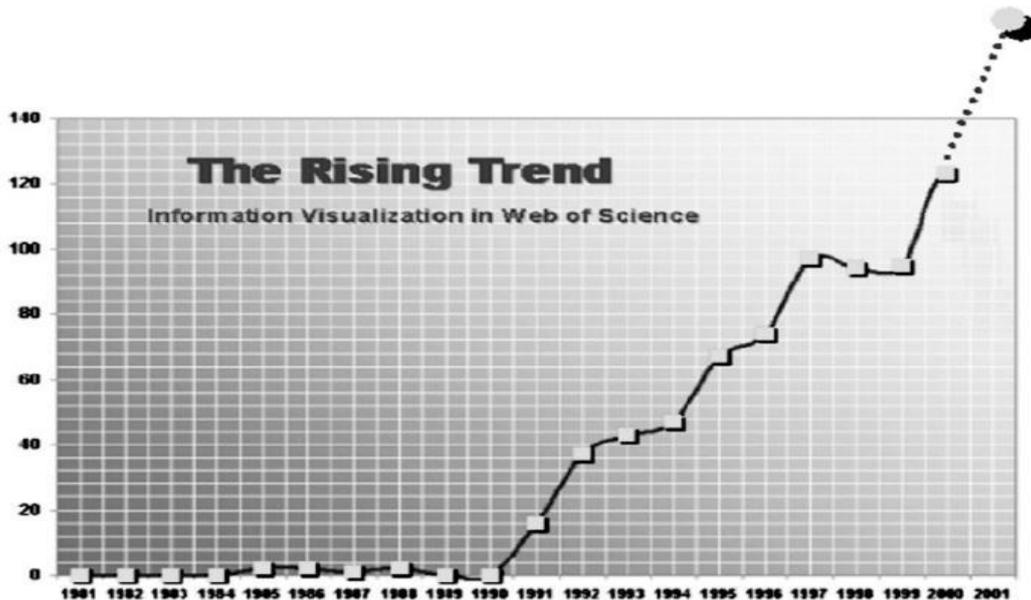


Figura 2. Crecimiento de las publicaciones sobre visualización de información en la década de los años 90 (Chen, 2002)

En la literatura algunos consideran a la visualización de la información como un proceso (Gershon & et al., 1998; Card & et al., 1999; Chen, 1999; Spence, 2000), otros autores consideran que es una herramienta (Robertson & et al., 1993), para otros es una disciplina (Catarci & Cruz, 1996; Zhu & Chen, 2005; Torres, 2009). Sin embargo, muchos de estos autores a pesar de sus disímiles propuestas poseen coincidencias en cuanto al propósito que le atribuyen a la visualización de la información, que no es más que: la representación de datos complejos u abstracciones de forma gráfica, el análisis de los datos y el reconocimiento de patrones, la obtención de un nuevo conocimiento y la simplificación del proceso de cognición, el enfrentamiento a la complejidad de los datos y la comprensión cualitativa de la información.

Los problemas generales que hoy en día afronta la visualización de la información son:

- Lograr que la cantidad de datos procesados no sean un límite para la percepción.
- Integrar la visualización con otros recursos multimedia como la voz.
- Crear representaciones visuales con mayor resolución espacial evitando las nubes de datos o etiquetas de datos.

- Obtener interfaces visuales que permitan la interacción temporal con el usuario (Chen, 2005).

Las técnicas aplicadas para la visualización de la información se enfocan en tres aspectos fundamentales: la estructura, la representación, y la interacción de los datos, prestando especial atención a las metáforas visuales, los algoritmos de clasificación y distribución de los datos, y la transformación interactiva de la información visualizada (Torres, 2009). Existen varios algoritmos de clasificación y distribución visual para diseñar visualizaciones de información, los más populares son: *clustering* o análisis de conglomerados (Hartigan, 1975); técnicas estadísticas multivariantes tales como el escalamiento multidimensional o MDS (*Multidimensional Scaling*) (Kruskal & Wish, 1978); técnicas de poda como el método de posicionamiento de los nodos (Kamada & Kawai, 1989) y el método de escalamiento de red *Pathfinder* o PFNETs (Schvaneveldt, 1990); y técnicas de redes neuronales, como el modelo de mapas auto-organizativos o SOM (*Self-Organizing Map*) (Kohonen, 1993).

Cada uno de los métodos anteriores facilita la reducción de estructuras de datos multidimensionales a formas de representación en 2 o 3 dimensiones, creando la posibilidad de una mejor interpretación de los datos. Pero al reducir las dimensiones de la realidad estructural de los datos se debe producir una pérdida de información. Para preservar la estructura original de los datos estas técnicas son capaces de solucionar el problema: por ejemplo; el *clustering* se centra en las relaciones de agrupación, el MDS se orienta a las distancias entre elementos, el *Pathfinder* a las relaciones locales más fuertes y el algoritmo SOM a las relaciones de vecindad. Este último algoritmo, se ha aplicado en estudios métricos, fundamentalmente para la clasificación de información, es decir, para la formación de *cluster* y su representación en mapas conceptuales (Honkela & et al., 1996, Börner & et al., 2003). En el siguiente acápite se habla con más detalle sobre el tema.

## 1.4 Redes Neuronales Artificiales

Las redes neuronales artificiales (RNA) han demostrado ser de gran utilidad para resolver problemas de minería de datos. Estas han sido útiles particularmente en la organización creativa de información, el descubrimiento de conocimiento y la visualización de información (Sotolongo-Aguilar & et al., 2002). Las RNA se nutren de modelos matemáticos que permiten hacer computación inteligente (Jain & et al., 1996) y se utilizan fundamentalmente para el reconocimiento de patrones, el control adaptivo, la predicción de series de tiempo, la clasificación de señales y *clustering*, entre otras (Guzmán & et al., 2004). Una de las grandes ventajas que ofrece la utilización de una red neuronal, es su capacidad de cómputo en paralelo, donde cada elemento de procesamiento (neuronas) de la red se representa como un nodo, recibiendo múltiples entradas de información y enviando salidas simultáneamente (Kohonen, 1993). El procesamiento en paralelo y distribuido en la red neuronal hace posible trabajar con muchos datos en una escala masiva.

Por ello las conexiones entre las neuronas se establecen con una estructura jerárquica tratando de emular la fisiología del cerebro, siempre en busca de nuevos modelos de procesamiento para solucionar problemas concretos del mundo real (Freeman & Skapura, 1993). Las redes neuronales artificiales cuentan con cinco principios fundamentales:

- (1) Aprendizaje adaptativo: esta característica se destaca porque es la capacidad que posee una red neuronal artificial para modificarse constantemente con el fin de adaptarse a nuevas condiciones de trabajo. Por lo tanto, no es necesario elaborar un modelo a priori, ni establecer funciones probabilísticas.
- (2) Autoorganización: la autoorganización significa generalización, consiste en la modificación de la red completa cuando recibe datos nuevos. De esta forma la red neuronal responde a datos o situaciones que no ha experimentado antes, hace inferencias a partir de la base de su entrenamiento.

- (3) Tolerancia a fallos: los algoritmos de computación tradicional almacenan la información en espacios únicos, localizados y direccionables, lo que trae consigo una cierta vulnerabilidad a la pérdida de datos y a la inutilización del sistema. En cambio, las redes neuronales artificiales poseen la capacidad de almacenar la información de forma distribuida y con un alto grado de redundancia. Esto permite que las redes sean capaces de reconocer patrones de información con ruido, distorsión o incompletos y a pesar de eso continuar trabajando.
- (4) Operación en tiempo real: el procesamiento en paralelo de las redes neuronales permite realizar el reconocimiento de patrones en tiempo real, siendo uno de los métodos existentes más indicado para esto. Resulta importante señalar que esta característica sólo se aprecia cuando se implementan redes con hardware, especialmente diseñados para el procesamiento paralelo.
- (5) Fácil inserción en la tecnología existente: Es relativamente sencillo obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilita la integración modular en los sistemas existentes (Guzmán & et al., 2004).

#### **1.4.1 Estructura de una Red Neuronal Artificial**

Las Redes Neuronales Artificiales se componen de una gran cantidad de neuronas, estas suelen llamarse nodos o unidades de procesamiento. La cantidad de nodos que componen la red neuronal generalmente es definida por el usuario, en virtud del análisis que vaya a realizarse y el volumen de datos. Un nodo o neurona cuenta con una cantidad variable de entradas que provienen del exterior ( $X_1, X_2, \dots, X_m$ ). Sin embargo, dispone de una sola salida ( $X_j$ ), la cual transmite la información hacia otras neuronas o al exterior. Cada señal de entrada tiene asociada una magnitud llamada peso ( $W_{j0}, W_{j+1}, \dots, W_{j+m}$ ). El peso responde a la intensidad de los enlaces sinápticos entre las neuronas. En la figura 3 se puede observar el esquema de una neurona.

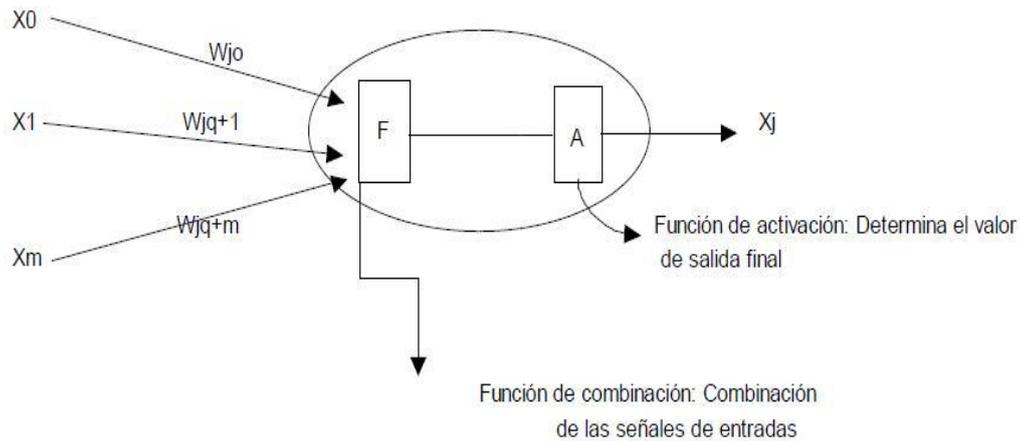


Figura 3. Estructura de una neurona (Kornilov, 1997)

La red neuronal posee variaciones de topologías, estas se clasifican utilizando estos criterios (Sotolongo-Aguilar & Guzmán-Sánchez, 2001):

- 1) Número de niveles o capas.
- 2) Número de neuronas por nivel.
- 3) Formas de conexión.

También suelen clasificarse por el tipo de algoritmo de aprendizaje, refiriéndose a redes neuronales supervisadas y redes neuronales no supervisadas (González, 1996). La selección y utilización de una topología va a ser determinada según el objetivo que se trace y el problema que se quiera solucionar.

Por último, se puede añadir que las Redes Neuronales Artificiales se han revelado como un instrumento útil para obtener información a partir de grandes volúmenes de datos. Su procesamiento en paralelo es ideal para prevenir fallas o errores en los datos. En la actualidad existen disímiles modelos de redes neuronales, creados para resolver situaciones y problemas determinados. Particularmente en la minería de datos básicamente se han utilizado dos tipos de redes neuronales: los perceptrones multicapa y el modelo de los mapas auto-organizados (*Self-Organizing Map, SOM*), más conocido como las redes de Kohonen.

### 1.4.2 Algoritmo SOM

Los mapas basados en el algoritmo *SOM* (*Self Organizing Maps*) están inspirados en las propias funciones de la corteza cerebral y cobra especial interés por su capacidad de clasificación. Este modelo es un eficiente algoritmo neuronal (no supervisado) que permite la proyección de datos que habitan en un espacio multidimensional, a una retícula bidimensional denominada mapa, preservando cualitativamente la organización (topología) del conjunto original (Guzmán & et al., 2004). Las redes *SOM* fueron presentadas por Teuvo Kohonen (1982), quien le denomina mapas auto-organizados, porque no requiere un entrenamiento supervisado, lo cual hace que este modelo sea muy útil para descubrir la estructura y el conocimiento contenido en una base de datos.

Una red auto-organizada es capaz de descubrir rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e incorporarlos a su estructura interna de conexiones (los pesos o vectores de referencia). Las neuronas deben auto-organizarse en función de los estímulos (datos) procedentes del exterior y en la mayoría de los casos se debe recurrir a un número considerable de patrones de entrada. En el proceso de aprendizaje, las neuronas compiten unas con otras con el fin de llevar a cabo una tarea dada. Se pretende que cuando se presente a la red un patrón de entrada, sólo una de las neuronas de salida (o un grupo de vecinas) se active. Por tanto, las neuronas compiten por activarse, quedando finalmente una como neurona vencedora y anuladas el resto, que son forzadas a sus valores de respuesta mínimos.

El objetivo de este aprendizaje es categorizar los datos que se introducen en la red. El algoritmo *SOM* es considerado muy eficiente para visualizar grandes volúmenes de datos multidimensionales (Kohonen, 1997; Kaski, 1997; Börner & et al., 2003; Moya-Anegón & et al., 2006) y también es una de las contribuciones más importantes a la visualización de la información (Börner, 2005).

## Capítulo 2. Fuentes y Métodos

### 2.1 Web of Science

En esta investigación se seleccionó como fuente documental la base de datos “*Web of Science*” (WoS). Esta fuente de información tiene una larga tradición, originalmente fundada en el año 1957 por Eugene Garfield en el *Institute for Scientific Information* (ISI). En el año 1992 esta institución es adquirida por la *Thomson Reuters* hasta el año 2016, donde fue comprada por la empresa transnacional *Clarivate Analytics*. Esta base de datos captura meticulosamente sus metadatos y conexiones de citas de una literatura de investigación vinculada a un núcleo rigurosamente seleccionado de publicaciones periódicas (Clarivate Analytics, 2018).

La plataforma del WoS conecta su colección principal (*Core Collection*) con índices regionales de citas, datos de patentes, índices de temas especializados y un índice de conjuntos de datos de investigación, con un total de más de 33,000 revistas. La colección principal del WoS está compuesta por las bases de datos siguientes:

- *Science Citation Index Expanded* (1900-presente)
- *Social Sciences Citation Index* (1900-presente)
- *Arts & Humanities Citation Index* (1975-presente)
- *Conference Proceedings Citation Index- Science* (1990-presente)
- *Conference Proceedings Citation Index- Social Science & Humanities* (1990-presente)
- *Book Citation Index– Science* (2005-presente)
- *Book Citation Index– Social Sciences & Humanities* (2005-presente)
- *Emerging Sources Citation Index* (2015-presente)

En estas bases de datos se puede acceder a literatura académica líder de todo el mundo en materia de ciencias naturales y exactas, ciencias sociales, artes y humanidades. Además, se tiene acceso a actas de congresos internacionales, conferencias, simposios, seminarios, coloquios, talleres y convenciones.

Los índices regionales de citas que ofrece la plataforma del WoS son:

- *SciELO Citation Index* (1997-presente): ofrece acceso a literatura académica en materia de ciencias, ciencias sociales, arte y humanidades publicada en las principales revistas de acceso abierto de América Latina, Portugal, España y Sudáfrica.
- *Russian Science Citation Index* (2005-presente): brinda el acceso a información bibliográfica y a las citas de artículos académicos más importantes de los investigadores rusos en más de 500 revistas de educación, medicina, tecnología y ciencia. La Biblioteca Electrónica Científica en Línea (*Scientific Electronic Library, eLIBRARY.RU*)
- *KCI - Korean Journal Database* (1980-presente): proporciona acceso a artículos de revistas multidisciplinarias más influyentes. Está administrado por la Fundación Nacional de Investigación de Corea.

Para este estudio se seleccionó el WoS como fuente de información basado en las fortalezas de la base de datos expuestas anteriormente y otros argumentos expuestos en la literatura por algunos autores que responden a los intereses y objetivos de la investigación:

- Mayor número de publicaciones: los investigadores mexicanos tienen un mayor número de publicaciones en la base de datos *Web of Science*, en comparación con las bases de datos regionales (Hernández & et al., 2013; Vélez-Cuartas & et al., 2016). Influye mucho en este sentido los criterios de evaluación que exigen el Consejo Nacional de Ciencia y Tecnología (CONACyT) y el Sistema Nacional de Investigadores (SNI), donde sus instituciones e investigadores deben de publicar sus artículos de investigación en revistas de la corriente principal.

- Afiliación institucional: el WoS provee información sobre la afiliación institucional de cada uno de los autores que publican en las fuentes que contempla esta valiosa base de datos (Clarivate Analytics, 2018).
- Multidisciplinariedad: todas las disciplinas están cubiertas (Leydesdorff & Bornmann, 2016).
- Alcance internacional: utilizar la base de datos del WoS permite obtener una visión más clara e internacional de la producción científica de un país, un campo de conocimiento, una institución e incluso hasta de un investigador (Archambault & et al., 2009).
- Herramientas de análisis: la plataforma del WoS ofrece a los investigadores herramientas de análisis bibliométrico; *InCites*, *Journal Citation Reports* y *Essential Science Indicators*, estas herramientas se explican en el siguiente acápite.

### **2.1.1 Herramientas de análisis del WoS: *Essential Science Indicators* e *InCites***

- *Essential Science Indicators*

*Essential Science Indicators* es una herramienta que revela las tendencias científicas emergentes, así como los investigadores, instituciones y revistas más influyentes en su campo de investigación. Los datos se derivan de más de 12 millones de artículos en más de 12,000 revistas indexadas en la Colección Principal del WoS, durante un período de los últimos 10 años. Las estadísticas integrales de rendimiento científico hacen de este un recurso analítico esencial para los responsables de la formulación de políticas científicas, administradores, analistas y especialistas en información de agencias gubernamentales, universidades, corporaciones, laboratorios privados, compañías editoriales y fundaciones.

Cada revista está asignada a uno de los 22 campos de investigación que componen el sistema de clasificación de esta herramienta. En el *Essential Science Indicators* una revista sólo puede estar asignada a un campo. Revistas como *Science* y *Nature* están categorizadas como multidisciplinarias, ya que publican investigaciones en muchos campos diferentes. Como resultado, los trabajos publicados en estas revistas

multidisciplinarias se asignan a un campo basado en la representación de las revistas citadas. Por ejemplo, si la mayoría de las referencias citadas en el documento son para revistas de neurociencia, el artículo se categoriza como neurociencia.

- *InCites*

*InCites* es una herramienta basada en citas para que los investigadores, académicos y administradores gubernamentales analicen la productividad institucional e individual en un contexto nacional o internacional. El origen de los datos de *InCites* provienen de la colección principal de WoS, su plataforma *Research Analytics* permite que los indicadores de *Essential Science Indicators*, *Journal Citation Reports* e *Institution Profiles* estén disponibles y explorables desde *InCites*. Su herramienta analítica permite crear tablas dinámicas y gráficos, también combina varias opciones para dirigir el enfoque de su análisis por diferentes entidades como personas, organizaciones, países y regiones, libros, revistas y actas de congresos; así como áreas de investigación. Sobre esta última opción de análisis conviene dedicar un apartado debido a que el WoS posee varios esquemas de clasificación de revistas, a consecuencia de su carácter multidisciplinario y estos esquemas están presentes en *InCites*.

### 2.1.2 Esquemas de clasificación disponibles en *InCites* y el WoS

En la literatura se ha debatido mucho sobre los sistemas de clasificación de las bases de datos (Glänzel & Schubert, 2003; Abrizah & et al., 2013; Leydesdorff & Bornmann, 2016; Wang & Waltman, 2016). Esta investigación parte de las fortalezas que presenta la clasificación temática de la propia fuente de datos WoS.

Las categorías temáticas agregadas por un indexador son muy útiles para fines de recuperación de información, pero no para generar relaciones entre las revistas, ya que difumina las distinciones analíticas de un campo de investigación. En este sentido, el uso de las categorías temáticas de la clasificación de las revistas se ha aceptado como la "mejor práctica" entre los investigadores bibliométricos (Leydesdorff & Bornmann, 2016). El sistema de

clasificación del WoS no rompe con la estructura científica de los campos de conocimiento, si no que respeta la interrelación que se establece entre las revistas y los documentos individuales. Las categorías temáticas del WoS han evolucionado cada vez más desde un esquema de clasificación para la recuperación a un estándar para normalizaciones en evaluaciones bibliométricas (Leydesdorff & Bornmann, 2016).

Para la obtención de un mejor contexto en las interpretaciones bibliométricas es importante conocer los esquemas de clasificación por áreas temáticas que brindan las bases de datos donde se realizan los análisis. Se hace necesario comprender el desempeño dentro del contexto de las áreas temáticas debido a que las tasas de publicación y el comportamiento de citas pueden variar considerablemente de una disciplina a otra, a un tipo de documento y a lo largo del tiempo. Tal es el caso del propio campo de la matemática que su tasa de citación es relativamente baja, pero sus citas pueden persistir durante un largo período de tiempo. Este comportamiento no se evidencia en los documentos de otras áreas como la biología molecular que tiene un crecimiento muy rápido y recibe una mayor cantidad de citas, pero estas no perduran en el tiempo.

En la plataforma de *InCites* se encuentra disponible un total de 17 esquemas de área de investigación, tres de estos esquemas son exclusivos del WoS. Los otros 14 esquemas se basan en el mapeo de datos del WoS a sistemas externos de clasificación de materias. Estos esquemas están diseñados para permitir el uso de indicadores bibliométricos en el contexto de un programa de evaluación de investigación regional. Los esquemas propios del WoS son los siguientes:

- *Web of Science*
- *Essential Science Indicators*
- *GIPP*

La selección del esquema de clasificación dependerá del objetivo del análisis. Si el objetivo es analizar un conjunto de datos pequeños como el resultado de una evaluación de un autor o un departamento, se aconseja utilizar la mayor precisión de una clasificación de materia estrecha. Sin embargo, si el objetivo es analizar un campo de conocimiento e incluso un país o región se recomienda utilizar un esquema más amplio. Los dos primeros esquemas de clasificación exclusivos del WoS utilizados en el estudio se explican a continuación:

- *Web of Science* (clasificación de materia estrecha)

Este esquema es considerado la categorización más estrecha de la base de datos con un total de 252 categorías temáticas asignadas en ciencias naturales y exactas, ciencias sociales, artes y humanidades. El esquema se crea asignando cada revista a una o más categorías de temas. Por ejemplo, las disciplinas amplias como la matemática se representan en distintos subcampos más pequeños. De las 252 categorías temáticas del WoS, nueve categorías están explícitamente etiquetadas por su contenido matemático:

- MATHEMATICS: topología, álgebra, análisis funcional, teoría combinatoria, geometría diferencial y teoría de números.
- STATISTICS & PROBABILITY: estructuras y construcciones matemáticas usadas para analizar la probabilidad de un conjunto dado de eventos de una familia de resultados.
- LOGIC: estudio de la lógica en todas sus formas, incluida la lógica matemática (pura y aplicada), la lógica filosófica y la lógica computacional (o lógica en informática, teórica y aplicada).
- MATHEMATICS, APPLIED: ecuaciones diferenciales, análisis numérico, no linealidad, control, software, análisis de sistemas, matemática computacional y modelado matemático.
- PHYSICS, MATHEMATICAL: lógica, teoría de conjuntos, álgebra, teoría de grupos, teoría de funciones, análisis, geometría, topología y teoría de la probabilidad que tienen aplicaciones en física.

- MATHEMATICAL & COMPUTATIONAL BIOLOGY: bioestadística, la bioinformática, la biometría, el modelado de sistemas biológicos y la biología computacional.
- PSYCHOLOGY, MATHEMATICAL: metodología e instrumentación experimental, métodos multivariados, manipulación estadística y estrategia de investigación.
- SOCIAL SCIENCES, MATHEMATICAL METHODS: metodologías cuantitativas utilizadas para la investigación en ciencias sociales, tales como modelos matemáticos y técnicas estadísticas para la evaluación de datos psicológicos, sociológicos y económicos.
- MATHEMATICS, INTERDISCIPLINARY APPLICATIONS: métodos matemáticos cuyo enfoque principal es en una disciplina específica no matemática, excepto psicología, historia, economía, etc.

Este esquema de clasificación se considera el mejor para realizar análisis bibliométrico detallado. En esta investigación se utilizó este esquema para identificar el desarrollo de las comunidades matemáticas mexicanas.

- *Essential Science Indicators* (clasificación de materia amplia)

Este esquema es de categorización amplia y comprende un total de 22 áreas temáticas en ciencias y ciencias sociales y se basa en asignaciones de revistas. Las revistas de *Arts & Humanities* no están incluidas. Cada revista se encuentra sólo en una de las 22 áreas temáticas y no hay solapamiento entre las categorías, además este esquema cubre a totalidad el esquema del WoS, lo que con un mayor grado de amplitud. En esta investigación se consideró este esquema de clasificación, ya que dentro de las 22 áreas temáticas está incluido el campo de la matemática.

## 2.2 Variables e Indicadores bibliométricos de la investigación

Tabla 1. Indicadores de la investigación.

Variables	Indicadores y descripción	Símbolo
Producción Científica	Total de documentos publicados en WoS por: países, instituciones y comunidades científicas.	Ndoc
	Número anual de documentos publicados en WoS.	ANdoc
	Porcentaje de documentos en el Q1: total de documentos publicados en revistas de mayor factor de impacto.	% Q1
	Porcentaje de excelencia científica: porcentaje de documentos que acumulan un volumen de citas superior al 1 % de los más citados por cada campo de investigación, año y tipo de documento.	% Exc
	Alto desempeño ( <i>High Performance</i> ): porcentaje de documentos que acumulan un volumen de citas superior al 10% de los más citados por cada campo de investigación, año y tipo de documento.	% HP
	Índice de desarrollo científico de un país ( <i>Scientific Development Index</i> ). Es el Ndoc en WoS/Número de habitantes) * 1,000,000. Este indicador es utilizado por varios organismos internacionales, la UNESCO (UNESCO, 2015); la OECD (OECD, 2016) y la Red de indicadores de Ciencia y Tecnología Iberoamericana e Interamericana (RICYT, 2018). Los datos del volumen de habitantes de cada país se obtuvieron del Banco Mundial (Banco Mundial, 2018).	SDI

Impacto/ Visibilidad real	Citas recibidas: es el total de citas recibidas de los documentos publicados en WoS por: años, países, instituciones y comunidades científicas.	Cr
	Documentos citados: es el total de documentos del WoS que recibieron citas por: años, países, instituciones y comunidades científicas.	CD
	Porcentaje de documentos citados: es el porcentaje de documentos del WoS que recibieron al menos una cita.	% CD
	Relación de citas por documento publicado: es el cociente de las citas recibidas entre el número de documentos publicados: por países, instituciones y comunidades científicas.	Cr/Ndoc
	Relación de citas por documento citado: es el cociente de las citas recibidas entre el número de documentos citados: por países, instituciones y comunidades científicas.	Cr/Dc
	Artículos altamente citados ( <i>Highly Cited Papers</i> ): es el número de documentos 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 el mismo campo científico.	HCP
	Porcentaje de excelencia científica: porcentaje de documentos que acumulan un volumen de citas superior al 1 % de los más citados por cada campo de investigación, año y tipo de documento	% Exc
	<i>Category Normalized Citation Impact</i> : Normaliza el impacto, tomando en cuenta la categoría temática, el tipo de documento y la fecha de publicación.	CNCI

Colaboración internacional	Colaboración Internacional: Número de documentos con uno o más autores internacionales.	NdocIC
	Por ciento de documentos en colaboración internacional es: el porcentaje de documentos publicados en colaboración internacional.	% IC

### 2.3 Metodología ViBlioSOM

La metodología ViBlioSOM (Visualización Bibliométrica usando la Red Neuronal SOM) (Guzmán, 2009), ha sido parcialmente implementada en un sistema de software, llamado LabSOM, desarrollado por un grupo multidisciplinario de tecnólogos e investigadores del Laboratorio de Dinámica No Lineal, en la Facultad de Ciencias de la Universidad Nacional Autónoma de México (UNAM); esta herramienta es muy útil para visualizar e interpretar datos multidimensionales.

En esta investigación la metodología fue aplicada para identificar los perfiles de desempeño cientométrico de las instituciones mexicanas que producen investigación científica en el campo de la matemática. También se aplicó para identificar los perfiles de las categorías temáticas que componen el campo de la matemática según la clasificación general de la colección principal del WoS.

La metodología ViBlioSOM guarda una estrecha relación con los procesos de la metodología del *KDD*, las cuales son:

1. Definición del objetivo.
2. Adquisición y selección de los datos.
3. Preprocesamiento.
4. Minería de datos.
5. Visualización e interpretación de resultados.
6. Difusión del conocimiento obtenido a partir de los resultados.

### 2.3.1 Etapas de la metodología ViBlioSOM

En la siguiente figura se muestra las principales etapas de la metodología ViBlioSOM. El software LabSOM implementa la etapa de procesamiento neuro-computacional.

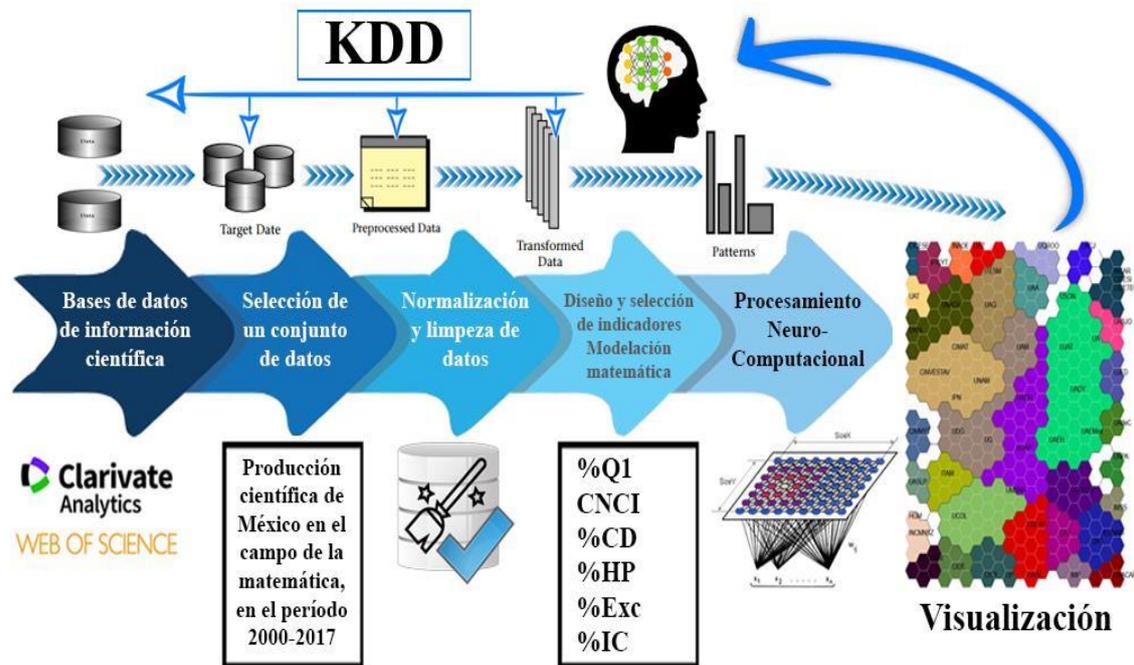


Figura 4. Procesos fundamentales de la metodología ViBlioSOM.

1. En primera instancia se seleccionaron las fuentes de información para la obtención de los datos, se utilizó la herramienta bibliométrica del *InCites* que provee todos los registros bibliográficos incluidos en la Colección Principal del WoS y el *Essential Science Indicators*.
2. Una vez obtenidos los datos se construyó una matriz, esta tabla relaciona por las columnas los valores para cada indicador y por las filas las entidades de análisis (vectores de referencia). Ahora bien, muchos de los indicadores utilizados para el análisis poseen rangos de variabilidad dispares. Por lo tanto, para que los indicadores con valores más grandes no predominen durante el entrenamiento se aplicó una técnica de preprocesamiento denominada MIN\_MAX. Esta técnica está incorporada a las herramientas que ofrece el LabSOM y consiste en dividir el valor de cada indicador (en la columna) entre el mayor valor de la columna que corresponda a ese

indicador. De manera tal que el mayor valor del indicador al dividirse por sí mismo obtiene un valor de uno y el resto de los valores será inferior a este, obteniéndose entonces un rango de variación de todos los indicadores entre cero y uno.

3. Al tener todos los valores normalizados en una misma escala numérica, se procede a realizar el entrenamiento de la red neuronal con los valores de la matriz. Para correr el entrenamiento de la red neuronal se consideran varios factores como son: el tamaño de la red (lado x ancho), el tamaño va a determinar el número de neuronas que van a participar en el entrenamiento. Para el caso del análisis de las instituciones se utilizó una red neuronal de 20 x 20, es decir, 400 neuronas y en el análisis por las categorías temáticas se utilizó una red de 10 x 10, un total de 100 neuronas. Otro de los aspectos a tener en cuenta para el entrenamiento es el número de iteraciones, en el caso de las instituciones se utilizaron un total de 4,000 iteraciones y para las categorías temáticas un total de 100 iteraciones.
4. Al concluir el entrenamiento se analizan dos tipos de visualizaciones: los mapas de componentes y el mapa de *clustering*. Cada uno de los mapas de componentes representa los valores obtenidos para cada indicador. En la figura 5 se muestra un ejemplo de un mapa de componente. En cada mapa de componente, el valor máximo del indicador corresponde al color rojo más oscuro, el valor medio está asociado al tono amarillo y el valor mínimo al verde más claro. El amarillo intenso se asocia con valores cercanos a la media de cada indicador. Estos mapas se dibujan sobre una retícula hexagonal, donde cada hexágono representa una neurona.

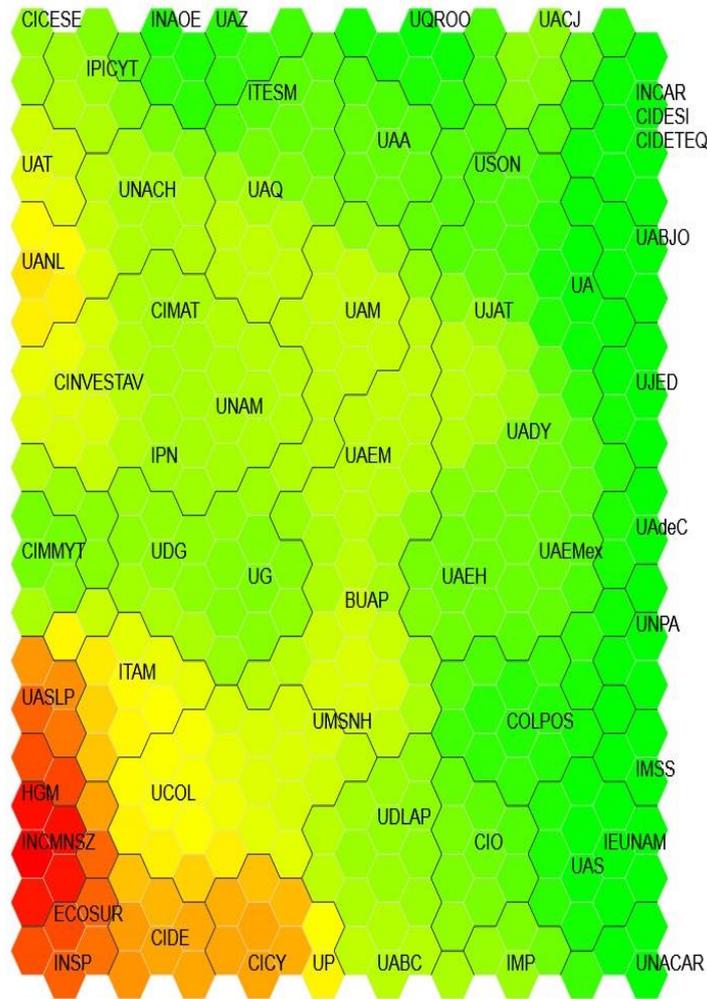


Figura 5. Mapa de componente del indicador: porcentaje de documentos en Q1, análisis de las instituciones.

Así como este mapa de la figura 5, se obtiene uno por cada indicador utilizado en el análisis. La otra visualización que se obtiene es el mapa de *clustering*, donde se observa el agrupamiento por similitud que presentan las entidades analizadas (instituciones y categorías temáticas). Para enriquecer la interpretación de cada grupo, se incorporan los mapas de componentes de cada uno de los indicadores (figura 24). El análisis de estos grupos nos permite caracterizar los perfiles cuantitativos, tanto de las instituciones como de las categorías temáticas. La interpretación de varios indicadores en el mapa de *clustering* es considerado un análisis multiparamétrico, ya que muestra las agrupaciones a partir de las similitudes que poseen las entidades de análisis en los distintos indicadores aplicados. En la figura 6 se muestra el mapa de *clustering* obtenido en el análisis realizado a las instituciones.

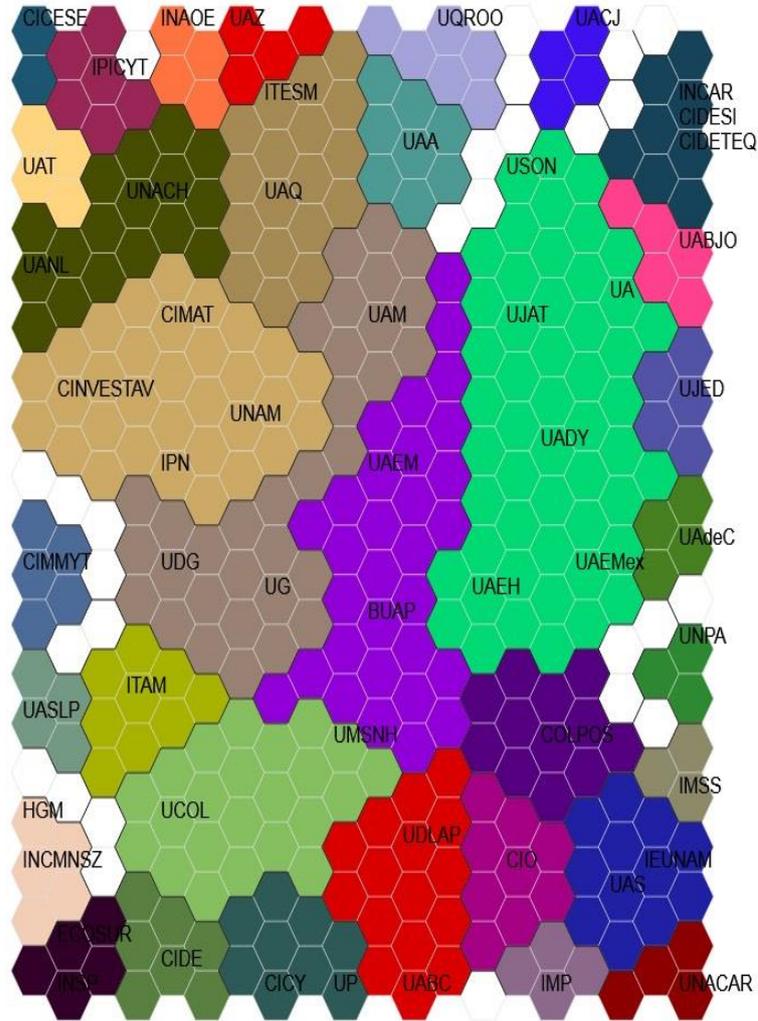


Figura 6. Clustering del análisis de las instituciones utilizando cinco indicadores: porcentaje de artículos publicados en revistas del cuartil uno de impacto (%Q1), Category Normalized Citation Impact (CNCI), porcentaje de Documentos Citados (%CD), porcentaje de *High Performance* (%HP), porcentaje de artículos en Colaboración Internacional (%IC).

5. Una vez obtenidas las visualizaciones, estas se guardan en formato JPG y se comienza con el proceso de interpretación de los resultados; buscando en todo momento el conocimiento que esta tecnología es capaz de aportar.

## **Capítulo 3. Análisis y Discusión de Resultados**

### **3.1 Evolución del campo de la matemática a nivel internacional**

Diversos autores han estudiado el desarrollo de la producción científica en el campo de la matemática en períodos precedentes (1868-2008), las fuentes documentales de estos estudios fueron las bases de datos (ZMATH y MathSciNet). Estos estudios demostraron que en el campo de la Matemática existe un crecimiento de sus publicaciones y de la comunidad de autores, además de una tendencia hacia la publicación en co-autoría múltiple. (Dang & Zhang, 2003; Behrens & Luksch, 2011).

Por otra parte, una investigación hace más de 50 años identificó que la producción científica a nivel internacional en el campo de la Matemática se duplica en aproximadamente 28 años (May, 1966). Otro estudio un poco más reciente determinó que el período de duplicación de la literatura fue en 21 años (Behrens & Luksch, 2011). Sin embargo, en el presente estudio internacionalmente se logra duplicar la productividad en 13 años e inclusive México duplicó su producción científica en 10 años. Todas estas investigaciones nos llevan a determinar que el crecimiento anual en el campo de la matemática es cada vez mayor y por lo tanto el período de duplicación de la literatura es cada vez más corto.

En la figura 7 se observa que tanto la producción científica mexicana como la internacional en el campo de la matemática, ha tenido un crecimiento sostenido.

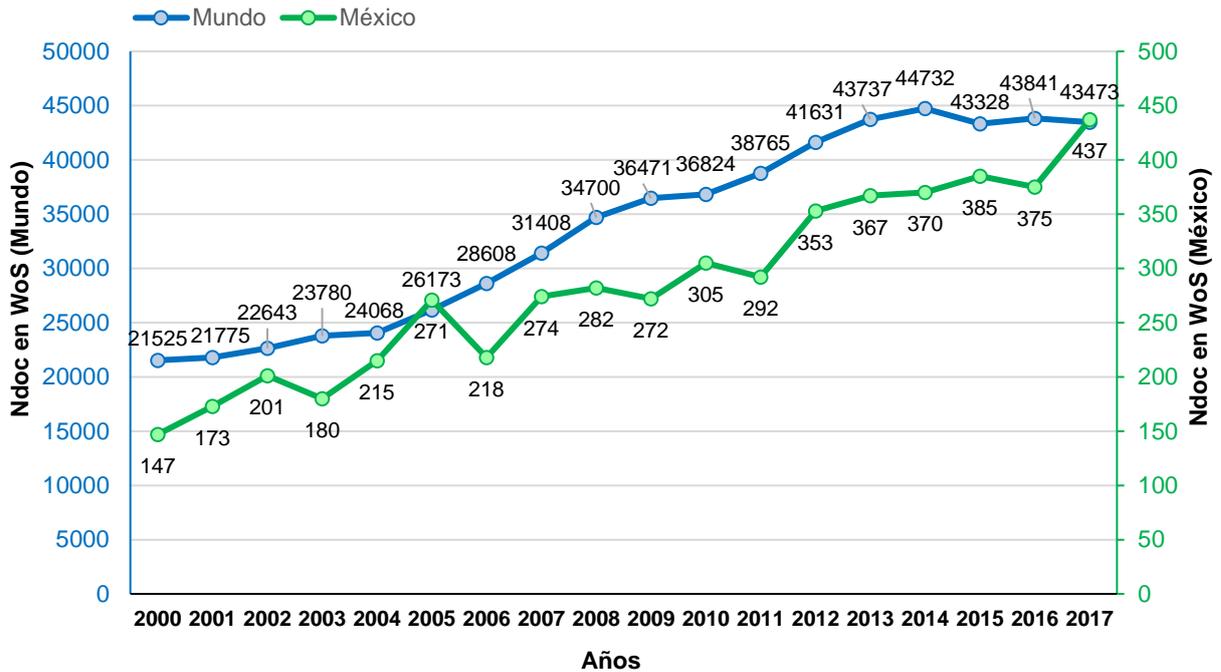


Figura 7. Evolución del número de documentos publicados en el WoS, en el campo de la matemática a nivel internacional y México.

Se debe destacar que en los últimos seis años la producción científica de México aumentó en un 24%; mientras que el mundo sólo logró un aumento de un cuatro por ciento. Esto significa que los científicos del gremio de las matemáticas en México han aumentado su producción científica en el campo, en comparación con el promedio mundial. En este punto se debe enfatizar los años 2005 y 2017; los cuales marcaron en sus respectivas décadas picos de máxima producción científica, a diferencia del comportamiento a nivel internacional donde los artículos publicados aumentaron de manera paulatina y a partir del año 2014 comienzan a disminuir.

### 3.1.1 Países con mayor volumen de publicaciones a nivel internacional

Durante el período de estudio 2000-2017, el volumen de producción científica en el WoS en el campo de la matemática a nivel mundial fue de 607,482 documentos publicados en un total de 1,040 revistas. En la tabla 2 se muestran los países de mayor volumen de producción científica; Estados Unidos y China son los países con mayor volumen de producción científica en el campo con 146,677 y 100,023 documentos publicados, respectivamente. Le siguen Francia y Canadá con 50,176 y 25,723 artículos publicados,

respectivamente. En el anexo 1 se muestran todos los países (127) con una producción científica superior a 18 artículos en el período.

Tabla 2. Países con mayor volumen de publicaciones a nivel internacional.

No.	Países	Ndoc	% M*	No.	Países	Ndoc	% M*
1	Estados Unidos	146,677	24.15	16	Irán	11,924	1.96
2	China	100,023	16.47	17	Turquía	11,062	1.82
3	Francia	50,176	8.26	18	Israel	10,078	1.66
4	Alemania	41,985	6.91	19	Taiwán	9,415	1.55
5	Reino Unido	33,537	5.52	20	Rumania	8,651	1.42
6	Italia	31,553	5.19	21	Suiza	7,438	1.22
7	Rusia	27,472	4.52	22	Holanda	7,168	1.18
8	Japón	27,397	4.51	23	Bélgica	7,078	1.17
9	España	25,777	4.24	24	Austria	6,719	1.11
10	Canadá	25,723	4.23	25	República Checa	6,531	1.08
11	India	15,180	2.50	26	Suecia	6,396	1.05
12	Corea del Sur	14,884	2.45	27	Arabia Saudita	6,295	1.04
13	Polonia	14,576	2.40	28	Portugal	5,925	0.98
14	Australia	12,571	2.07	29	Hungría	5,765	0.95
15	Brasil	12,239	2.01	30	México	5,117	0.84

% M\* Es el porcentaje del volumen de la producción científica a nivel mundial = 607,482

La producción científica de México en el período de estudio es de 5,117 documentos publicados y representa el 0.84% de la producción científica internacional en el campo.

En la figura 8 se observa el promedio anual del Índice de Desarrollo Científico, este indicador define la cantidad de artículos publicados al año por cada millón ( $10^6$ ) de habitantes.

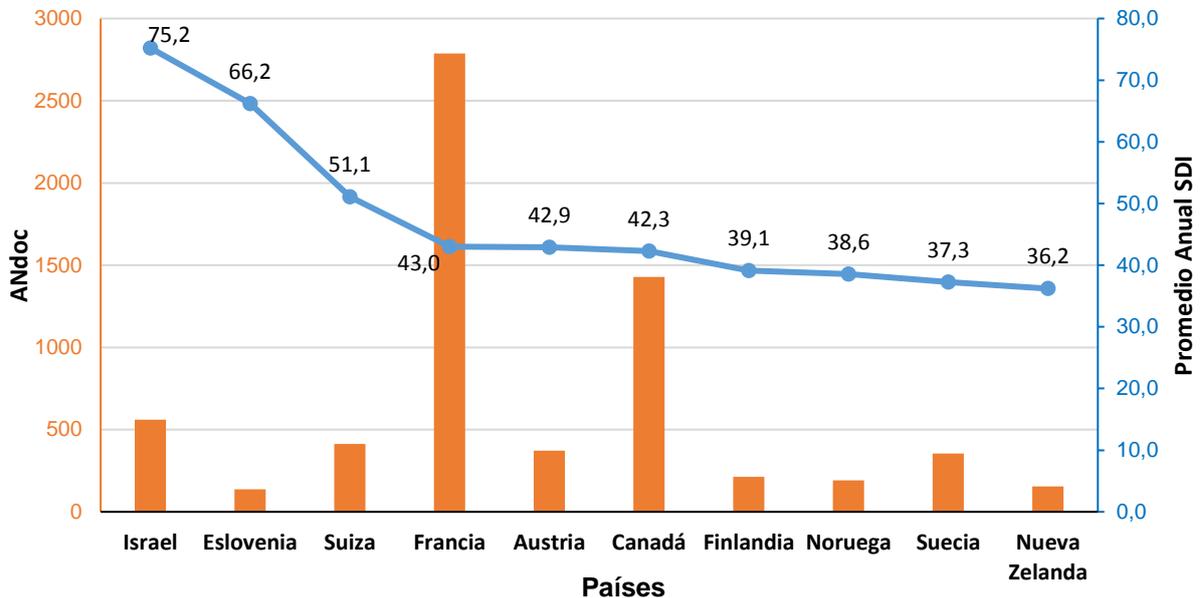


Figura 8. Países con mayor índice de desarrollo científico y su producción anual en el campo de la matemática.

Los países con mayor SDI resultaron ser Israel (75.2), Eslovenia (66.2) y Suiza (51.1), que no destacan en cuanto al volumen de producción científica. Los países con mayor volumen de producción científica no poseen los mejores valores SDI y no se muestran en la gráfica: Estados Unidos (SDI=26), China (SDI=4), Alemania (SDI=28), Italia (SDI=29) y Reino Unido (SDI=29). Los únicos países que se incluyen en la gráfica de los que poseen un alto volumen de producción científica y además un alto valor SDI son, Francia (SDI=49) y Canadá (SDI=42.3), estándares a nivel internacional. En el caso particular de México su valor en el índice de desarrollo científico (SDI=2.38) es bajo en comparación con los países ilustrados en la figura 8.

Al realizar el análisis de la producción científica por cada millón de habitantes se identificó un panorama totalmente diferente al observado con el volumen de producción científica. Esta situación ha sido descrita por otros autores (Cáceres & et al., 2014) donde se evidencia que la producción científica puede ser alta debido a la cantidad poblacional de cada país, lo cual provoca una sobrevaloración de la misma. De ahí la importancia de tener en cuenta el

número de habitantes de cada país, con el objetivo de obtener datos más precisos que permitan una mejor interpretación entre los mismos.

### 3.1.2 Producción en las mejores revistas

Lograr publicar los resultados de investigación en revistas del cuartil uno de acuerdo con el indicador de impacto del WoS condiciona que generalmente estos artículos sean citados. Ahora bien, publicar en revistas de alto factor de impacto es difícil, las tasas de rechazo de publicaciones de estas revistas son muy elevadas; el rigor, la exigencia y el alcance de la investigación no puede ser de carácter local. Por lo tanto, se hace necesario conocer cómo ha evolucionado en el tiempo el porcentaje de documentos en el cuartil de más alto factor de impacto.

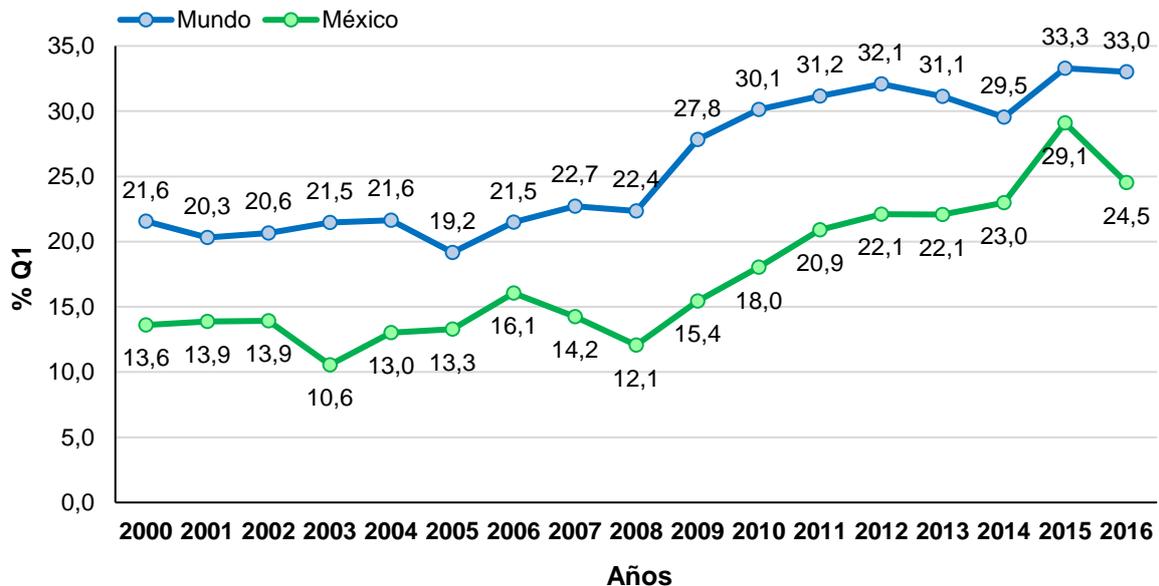


Figura 9. Evolución del porcentaje de documentos en Q1.

En la figura 9, se evidencia que tanto las publicaciones del Mundo como de México en el campo poseen una tendencia creciente en los últimos ocho años. México a inicios del período en el año 2000 tenía un porcentaje de documentos en el cuartil uno de 13.6 y para el año 2015 ya contaba con un 29.1. Esto quiere decir que la producción científica mexicana en la actualidad está muy cerca de alcanzar la media mundial en cuanto al porcentaje de artículos publicados en las revistas de más alto factor de impacto.

### 3.1.3 Evolución del volumen de citas

En la figura 10 se muestra la evolución del volumen de citas recibidas en el campo de la matemática a nivel internacional en correspondencia con México. A nivel internacional el campo de la matemática tiene acumulado más de cuatro millones de citas y tiene un promedio de siete citas por documento publicado. Los matemáticos mexicanos poseen un volumen de 23,189 citas y un promedio de cuatro citas por documento publicado en el período de análisis. El impacto de México tiene dos años de irregularidad (2006 y 2009); los cuales coinciden con la disminución en la producción científica.



Figura 10. Evolución del impacto del Mundo y México.

Se debe destacar que a partir del año 2012 México y en el 2009 el Mundo, tienden a disminuir sus citas. Este evento es normal y se corresponde con estudios de ventanas de citación (Dorta & Dorta, 2016; Adler & et al., 2008), donde se plantea que la ventana de citas varía en cada campo de estudio. Estas investigaciones demostraron que el campo de la Matemática tiene una ventana de citas entre 5 y 10 años, en otras disciplinas como las Ciencias Sociales y las Humanidades el impacto de las citas es más lento.

### 3.1.4 Países con mayor volumen de citas a nivel internacional

Durante el período de estudio, el volumen de citas en el WoS a nivel internacional en el campo de la matemática se ve reflejado un total de 448,747 artículos citados. En la tabla 3 se muestran los países que poseen más citas recibidas por encima de México en el período.

Tabla 3. Países con mayor número de citas recibidas en el campo de la matemática en WoS.

No.	Países	Cr	% DC	Cr/Ndoc	No.	Países	Cr	% DC	Cr/Ndoc
1	Estados Unidos	1,472,244	77.06	10.04	19	Bélgica	62,141	78.69	8.78
2	China	577,008	66.73	5.77	20	Taiwán	59,527	72.75	6.32
3	Francia	418,908	78.44	8.35	21	Irán	58,045	63.37	4.87
4	Alemania	336,441	77.78	8.01	22	Turquía	57,833	66.05	5.23
5	Reino Unido	305,215	77.91	9.10	23	Austria	54,671	76.71	8.14
6	Italia	244,052	77.78	7.73	24	Suecia	53,130	76.52	8.31
7	Canadá	205,330	76.75	7.98	25	Rumania	45,064	67.23	5.21
8	España	182,811	77.22	7.09	26	República Checa	39,096	74.15	5.99
9	Japón	151,144	72.08	5.52	27	Arabia Saudita	36,588	65.72	5.81
10	Australia	110,590	76.59	8.80	28	Grecia	34,497	74.52	6.98
11	Rusia	96,397	61.08	3.51	29	Portugal	34,436	74.33	5.81
12	Corea del Sur	80,062	69.70	5.38	30	Noruega	31,805	77.28	9.18
13	Israel	79,976	77.26	7.94	31	Finlandia	31,262	77.33	8.15

### Capítulo 3. Análisis y Discusión de Resultados

14	India	72,821	65.80	4.80	32	Hungría	30,129	71.27	5.23
15	Polonia	72,125	72.81	4.95	33	Singapur	29,340	77.09	9.38
16	Brasil	70,955	71.26	5.80	34	Dinamarca	28,356	77.24	9.17
17	Suiza	70,440	78.37	9.47	35	Chile	26,358	74.95	6.93
18	Holanda	64,718	78.67	9.03	36	México	23,189	69.02	4.53

En el análisis del impacto se puede constatar que el listado de los países más destacados coincide en sus primeros seis lugares con los países de mayor volumen de producción científica en el período de estudio. En el caso particular de México desciende de la posición 30 que tenía en cuanto al volumen de producción científica a la posición 36 en el impacto a nivel internacional. Algo similar le ocurrió a Rusia que contaba con un séptimo lugar a nivel internacional (27,472 artículos publicados) y no logra estar entre los países que tienen más de 100,000 citas. A diferencia de Canadá que ocupaba un décimo lugar en el mundo (25,723 artículos publicados) y logra subir al séptimo lugar del ámbito internacional con el 77% de sus artículos citados.

Por otra parte, tenemos otra situación muy diferente con China que es el segundo país en productividad (100,023 artículos publicados) y también es el segundo en la lista de los más citados (577,008 citas) ya que los autores chinos se han dedicado a publicar un volumen apreciable de artículos en inglés. Toda esta situación se debe al manejo del idioma; ya que los primeros países junto a Canadá tienen la gran parte de su producción científica en inglés a diferencia de Rusia, que publica en su idioma, de hecho, el volumen de su producción científica es tan grande que dentro del propio *Web of Science* hay una base de datos sólo para recolectarla (*Russian Science Citation Index*). Este fenómeno ha sido discutido ampliamente en la literatura (PLoS Medicine Editors, 2006; Besemer & Parr, 2013) donde hay autores que plantean que los artículos que no se publican en inglés no reciben casi citas, lo que disminuye su visibilidad.

### 3.1.5 Colaboración internacional

En la figura 11 se muestra la evolución del porcentaje de artículos en colaboración internacional de México y el mundo. Este último, posee una tendencia lenta de crecimiento de la colaboración internacional, a diferencia de México que se mantiene relativamente constante en sus valores, aunque comienzan a disminuir a partir del año 2011.

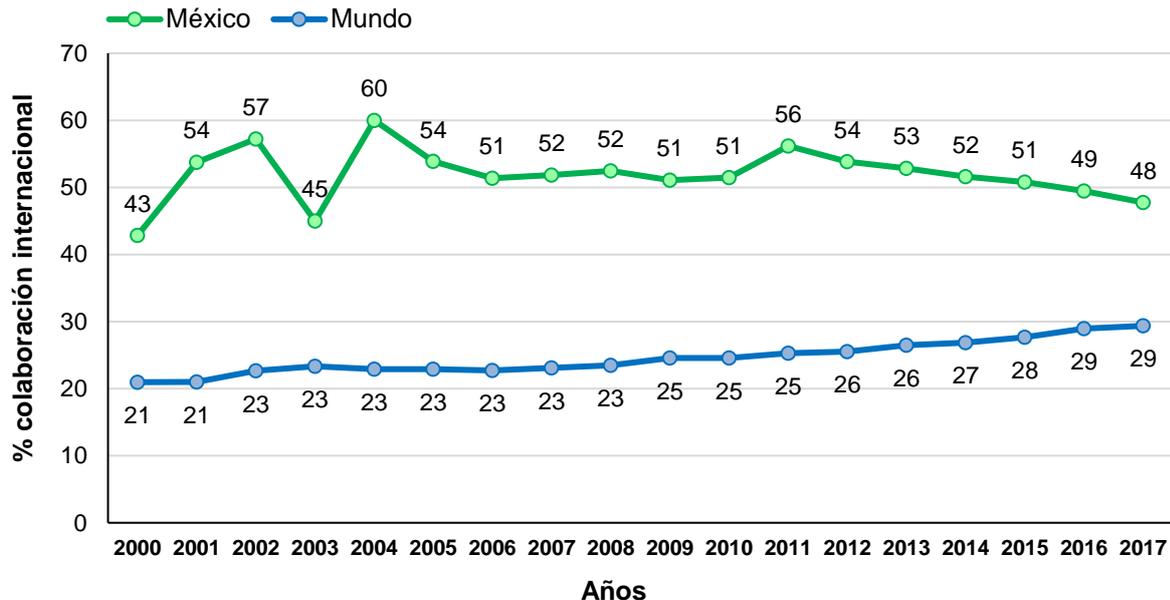


Figura 11. Evolución del porcentaje de documentos en colaboración internacional a nivel mundial y México.

El porcentaje promedio de colaboración de México en el período es (52%) y duplica al presentado a nivel internacional (25%). En este sentido, México tiende a comportarse de manera irregular con picos de 60% a 43%, a diferencia del mundo que aumenta de manera paulatina y sus valores oscilan entre 21 y 29%. Por otra parte, a nivel internacional se publicaron 152,638 artículos en colaboración internacional, de ellos 2,654 (1.7%) corresponden a México. Este análisis fue abordado en otra investigación (Dang & Zhang, 2003) durante un período precedente (1985-2000) donde se determinó que el aporte a la internacionalización a nivel de mundial por parte de México fue de 0.57%, lo que evidencia un aumento paulatino de la representación de México a nivel internacional.

### 3.2 Desarrollo de América Latina en el campo de las matemáticas

#### 3.2.1 Evolución de la producción científica

La región de América Latina está conformada por 29 países que han publicado en el período por lo menos un artículo en el WoS (anexo 2). En la figura 12 se observa que la evolución de América Latina y México; tuvieron un crecimiento promedio anual muy parecido (7%).

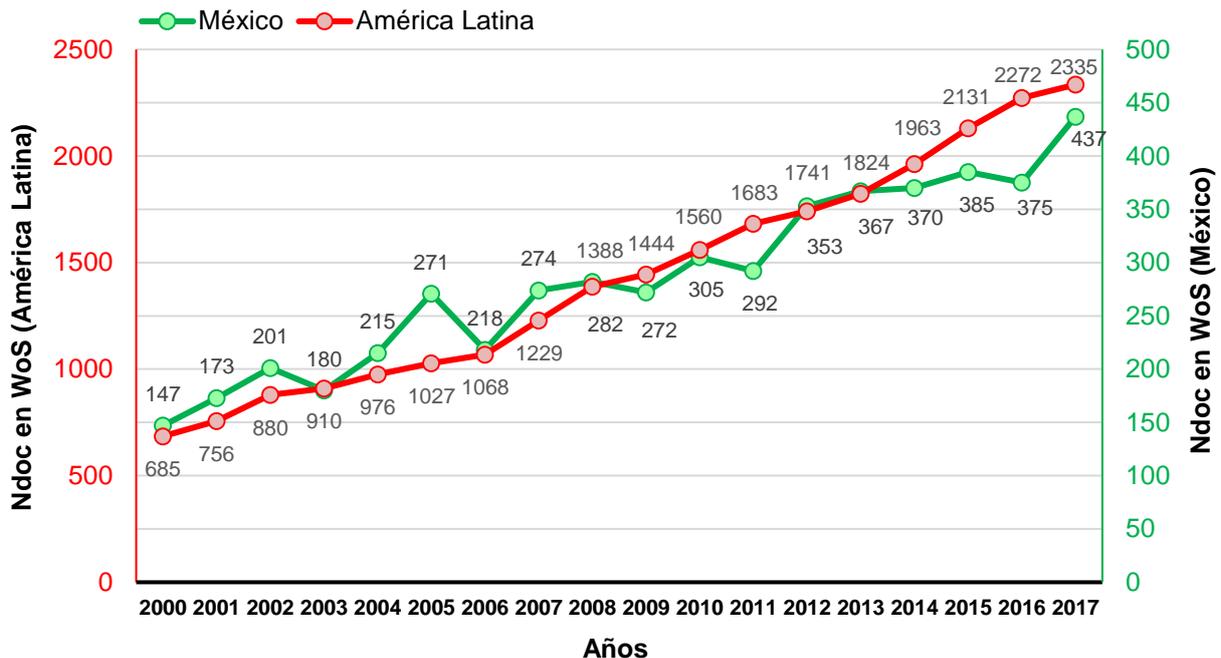


Figura 12. Evolución del número de documentos publicados en el WoS para América Latina y México.

Estos resultados coinciden con un estudio realizado entre 1972-2012, en el cual se determinó que América Latina comienza a aumentar su producción científica en Matemática a partir de los años 90 y los países que más se destacan son precisamente aquellos que tienen iniciativas que apoyan a la docencia y las investigaciones en este campo (Cáseres & et al., 2014).

Este crecimiento lineal de la producción de México es causado en cierta medida por la creación de la Facultad de Ciencias y el Instituto de Investigaciones de Matemática de la UNAM, así como otras Facultades y Centros de Investigación a lo largo del país, junto con la implementación de políticas científicas. Tanto en México desde la década de los 40, con su programa de becas y apoyos a estudiantes; como en América Latina con la

creación de las Olimpiadas Regionales de Matemática, los Programas de Doctorado y de Maestría en este campo de conocimiento.

### 3.2.2 Países con mayor volumen de publicaciones en la región

Durante el período de estudio, el volumen de producción científica en el WoS de los países de la región fue de 25,872 documentos publicados en un total de 556 revistas. En la figura 13 se muestran los países con un volumen de producción científica superior a 200 publicaciones.

En base al planteamiento anterior se determinó el comportamiento de nueve países destacados en la región. México aporta un 19% a la producción científica de la región y Brasil, que lleva el liderazgo en toda América Latina contribuye un 42%.

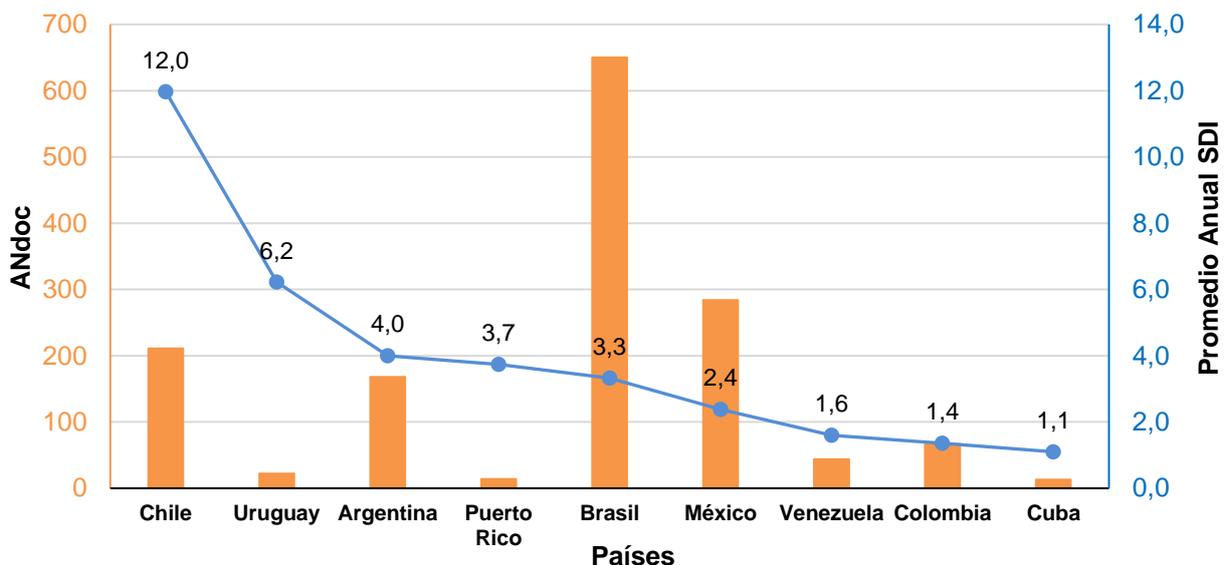


Figura 13. Países de la región con mayor número de publicaciones anuales y su índice de desarrollo científico (SDI).

En la figura se observa el promedio anual del número de publicaciones por cada uno de los países más destacados de la región y el promedio anual del Índice de Desarrollo Científico (SDI). En cuanto al volumen de producción científica, Brasil contó con 12,239 documentos publicados, México con 5,117, Chile y Argentina con 3,804 y 3,029, respectivamente. Se observa que no hay una correlación entre el Índice de Desarrollo Científico y el volumen de producción: Chile pasó de la tercera posición (ANdoc=211) a la primera (SDI=12) y Brasil que se consideraba el líder de la región con un (ANdoc=650),

llegó a situarse en quinto lugar (SDI=3.3). En el caso particular de México ocurrió algo similar, contaba con un (ANdoc=284) que situaba al país en la segunda posición en la región y terminó con un (SDI=2.4) posicionado en un sexto lugar.

### 3.2.3 Evolución del indicador de impacto

En la figura 14 se observa el comportamiento de las citas por documento publicado a nivel internacional, regional y se representan los países con mayor visibilidad en la región (México, Chile y Brasil). A nivel mundial el campo de la matemática tiene más de cuatro millones de citas acumuladas en el período. La región acumula un total de 144,157 citas, representando un tres por ciento a nivel internacional. En la figura se identificó que, de los tres países con mayor visibilidad de la región, sólo Chile tiene una evolución en las citas por documento superior al mundo en el período.

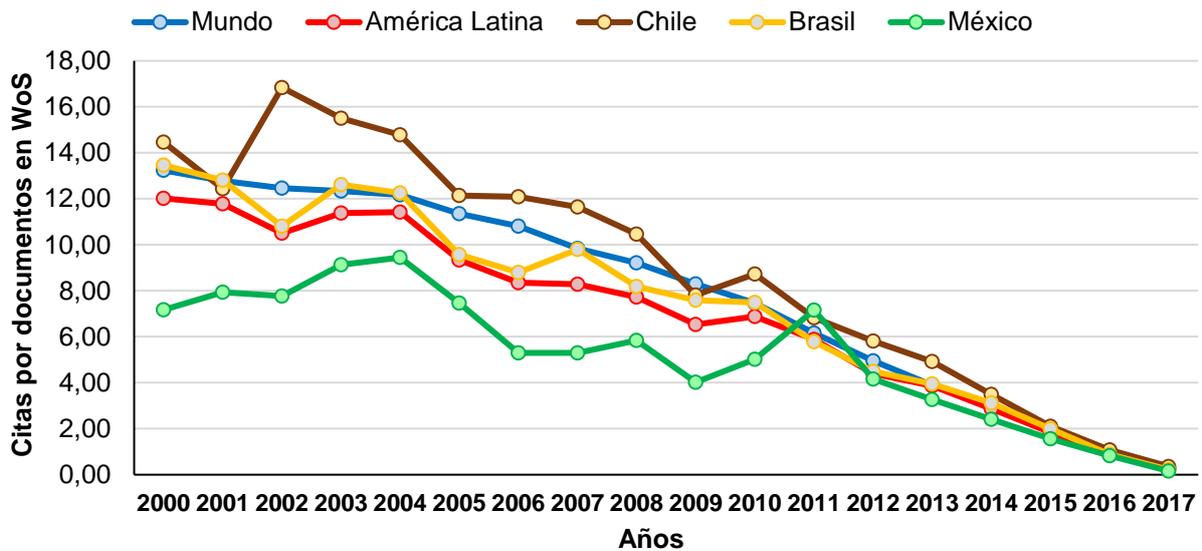


Figura 14. Evolución de las citas por documento del mundo, América Latina, Chile, Brasil y México.

Como se observa, existe una tendencia a la baja en todos los casos, esto sucede por las características propias de las citas que mientras más reciente sean las publicaciones menor probabilidad para obtener citas poseen.

### 3.2.4 Países con mayor impacto científico en la región

En la tabla 4 se muestra el impacto identificado por cada uno de los países de la región que recibieron al menos una cita en el período. Se utilizó el listado de países miembros de la ONU (ONU, 2018).

Tabla 4. Impacto de los países de la región.

No.	Países	Cr	% DC	Cr/Ndoc	No.	Países	Cr	% DC	Cr/Ndoc
1	Brasil	70,955	71.26	5.80	13	República Dominicana	150	50.00	6.82
2	Chile	26,358	74.95	6.93	14	Barbados	87	66.67	3.62
3	México	23,189	69.02	4.53	15	Paraguay	36	62.50	4.50
4	Argentina	17,760	73.03	5.86	16	Bahamas	18	75.00	2.25
5	Venezuela	4,588	69.58	5.81	17	Bolivia	9	33.33	1.50
6	Colombia	3,674	62.23	3.04	18	Trinidad y Tobago	8	25.00	0.67
7	Uruguay	2,269	71.90	5.74	19	Panamá	5	50.00	1.25
8	Cuba	1,018	64.56	4.30	20	Surinam	3	100	1.00
9	Perú	516	62.39	4.41	21	Guatemala	2	100	2.00
10	Jamaica	435	78.18	7.91	22	Dominica	2	100	2.00
11	Costa Rica	410	71.21	6.21	23	Haití	1	100	1.00
12	Ecuador	233	52.24	3.48					

Al analizar el impacto se puede constatar que el listado de los países más destacados coincide con los más productivos, pero en algunos casos se intercambian las posiciones entre ellos. Este es el caso particular de México y Chile que se intercambian, demostrando que Chile (3,804 artículos publicados) con una productividad inferior a México (5,117 documentos), logró tener 3,169 citas más. Chile, posee una mayor cantidad de citas por artículo publicado (7). Este intercambio ocurrió también entre Colombia y Venezuela, donde Colombia posee 1,210 artículos publicados y Venezuela sólo publica 789 documentos. Sin embargo, esta última obtuvo 914 citas por encima, logrando duplicar (6) la cantidad de citas por artículo publicado de Colombia. Este análisis nos permite concluir que Chile publica menos investigaciones que México, pero al tener un volumen de citas mayor logra un número superior de citas por documento.

### 3.2.5 Colaboración internacional

En la figura 15 se muestra el porcentaje de artículos en colaboración internacional de América Latina y la posición que ocupa México en esta. En la región existe una tendencia a mantener la colaboración internacional entre el 30% y 40%, a diferencia de México que se mantiene relativamente constante con valores cercanos al 50%.

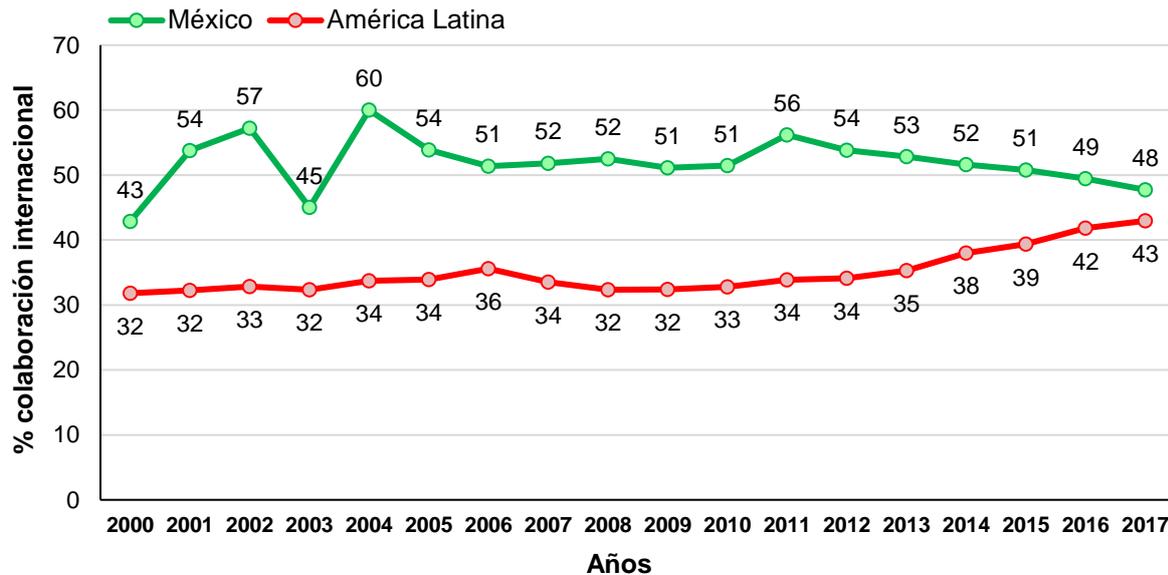


Figura 15. Evolución del porcentaje de documentos en colaboración internacional.

En el caso de la región se identificó una producción científica de 25,872; de los cuales el 35% son en colaboración internacional. Sin embargo, México publicó apenas la quinta parte de esa cantidad (5,117 artículos); pero publicó un 52% de sus artículos en colaboración internacional. No obstante, a partir del 2011 hay una tendencia en México a disminuir sus publicaciones en colaboración internacional, a diferencia de la región que va en aumento. Una de las razones por la que América Latina ha logrado elevar sus investigaciones en colaboración internacional es debido a la creación de la Unión Matemática de América Latina y el Caribe (UMALCA), la cual organiza y coordina eventos regionales de intercambio científico como el Congreso Latinoamericano de Matemáticos y Programas de Estudios a nivel de Pregrado y Posgrado en toda Latinoamérica; mediante la Escuela de Matemática de América Latina y del

Caribe (EMALCA) y la Escuela Latinoamericana de Matemática (ELAM), (UMALCA, 1995).

### 3.3 Evolución de la producción matemática de México

#### 3.3.1 Volumen de producción científica por entidad federativa

En la figura 16 se muestra la distribución de la producción matemática de los Estados de la República Mexicana. En este análisis se excluyó a la UNAM, IPN, IMSS y el IMP, ya que tienen representación en casi todo el país.

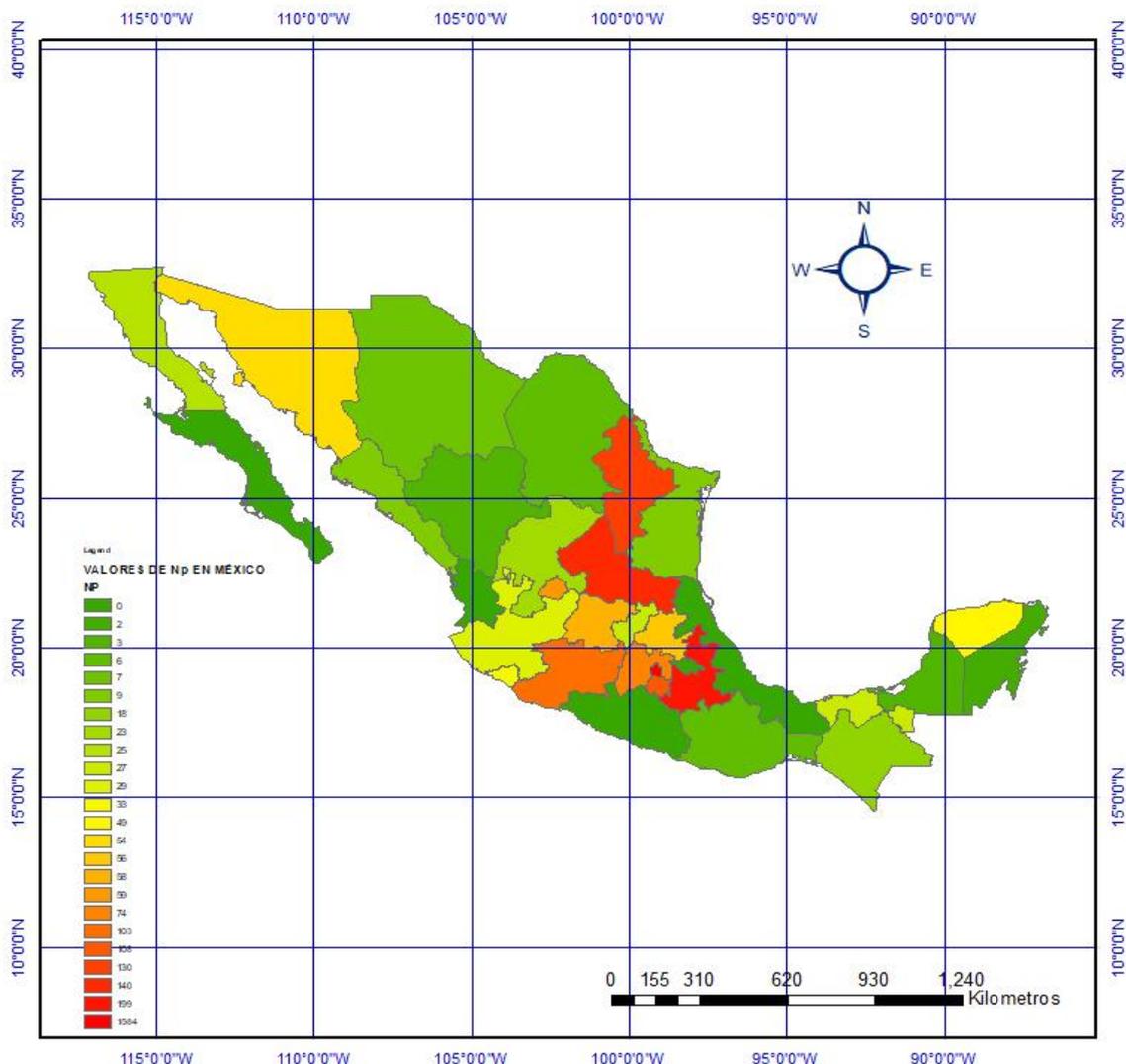


Figura 16. Distribución de la producción matemática de los Estados de la República Mexicana.

Teniendo en cuenta la leyenda del mapa, se determinó que la Ciudad de México es el único estado que supera los 1,500 artículos publicados en el

período. El peso de su productividad recae en diez instituciones (Universidad Autónoma Metropolitana, CINVESTAV, Centro de Investigación en Matemáticas, Instituto Tecnológico Autónomo de México, Universidad Panamericana y otras cinco instituciones que se muestran en el (anexo 3), las cuales tienen 1,584 artículos (31% respecto al total de la producción científica de México). Le siguen: Nuevo León, San Luis Potosí, Michoacán, Puebla y Morelos con más de 100 documentos cada uno. Son diez las instituciones (Benemérita Universidad Autónoma de Puebla, Universidad Autónoma de San Luis Potosí, Universidad Autónoma de Nuevo León, Universidad Autónoma del Estado de Morelos, Universidad Michoacana de San Nicolás de Hidalgo y otras cinco que se muestran en el (anexo 3). Estas instituciones encabezan la producción científica en estos estados y la cantidad de artículos publicados por ellas es de un total de 680 publicaciones, lo que representa un 13.3% del total de los artículos publicados por toda la nación mexicana en el campo. Sin embargo, hay estados que no tienen representación en el campo de la Matemática y estos son: Baja California del Sur, Nayarit, Veracruz y Guerrero.

### **3.3.2 Evolución del volumen de *highly cited papers***

Los artículos altamente citados (*highly cited papers*) permiten identificar la excelencia de la investigación científica y cuánta influencia tienen las investigaciones matemáticas mexicanas para la comunidad científica internacional. Este indicador se calcula para cada una de las 22 áreas temáticas en la herramienta cuantitativa, *Essential Science Indicators* (ESI). Los artículos altamente citados son aquellos que están en el uno por ciento de los más citados a nivel mundial.

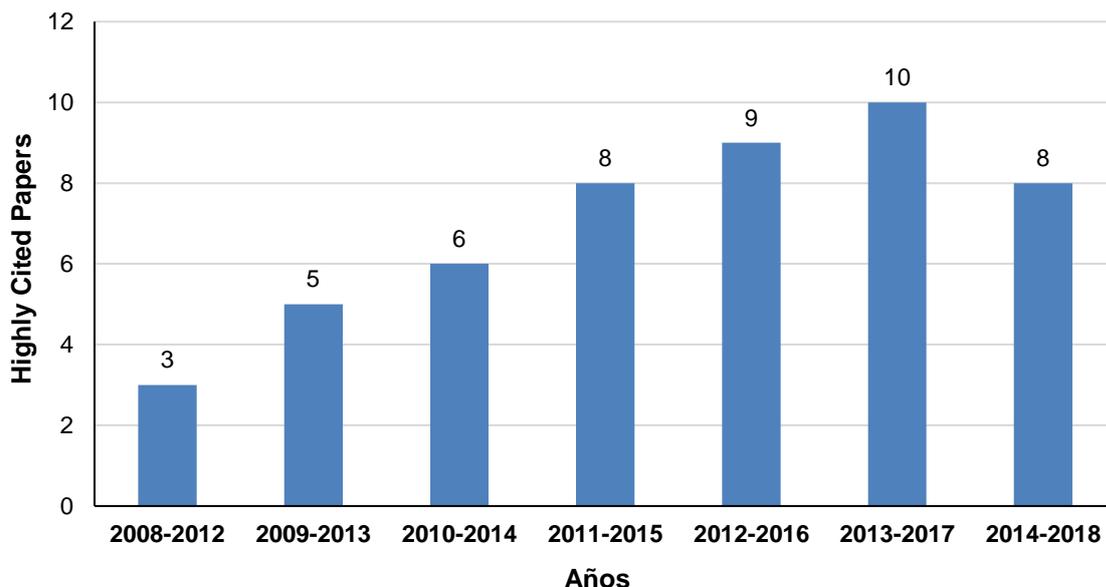


Figura 17. Evolución del número de *highly cited papers* en el período.

Este indicador no debe confundirse con el porcentaje de documentos en el uno por ciento superiores que se calcula en la herramienta del *InCites*; estos no son idénticos debido a que el ESI normaliza el cálculo teniendo en cuenta el campo, el año y el tipo de documento.

En la siguiente tabla se muestra el umbral de citas que debe superar un artículo en el campo de la matemática para que sea considerado un artículo altamente citado.

Tabla 5. Umbral de citas para ser considerado un artículo altamente citado.

2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
73	64	61	48	40	30	24	17	10	4	2
			A1-577	A3-52	A4-35	A6-29	A7-20	A9-13	A12-11	
			A2-92		A5-31		A8-19	A10-12	A13-9	
								A11-12		

En la tabla 5 se muestran los 13 artículos *highlys* identificados durante el período de estudio (en el anexo 4 se muestran las referencias bibliográficas de estos artículos). Como se pudo observar en la tabla anterior, el umbral de citas para ser considerado un artículo altamente citado puede variar con el tiempo; de manera tal que es posible que algunos artículos considerados altamente citados en un período luego dejen de serlo. Todo va a depender de que estos artículos altamente citados reciban un cúmulo de citas superior al

umbral calculado para el siguiente período. Considerando lo anterior a partir del año 2011 los matemáticos mexicanos alcanzan a tener dos artículos altamente citados como promedio anual, al acumularlos por año se da ese incremento por período que se evidencia en la figura 17.

### 3.3.3 Evolución de la excelencia científica y alto desempeño

Una tendencia actual en las comunidades científicas es promover que sus investigadores incrementen su producción científica por año. Muchas veces las publicaciones son necesarias para mejorar las calificaciones de las instancias evaluadoras por las que se rigen el sistema de ciencia de cada país, en detrimento de la calidad de las investigaciones.

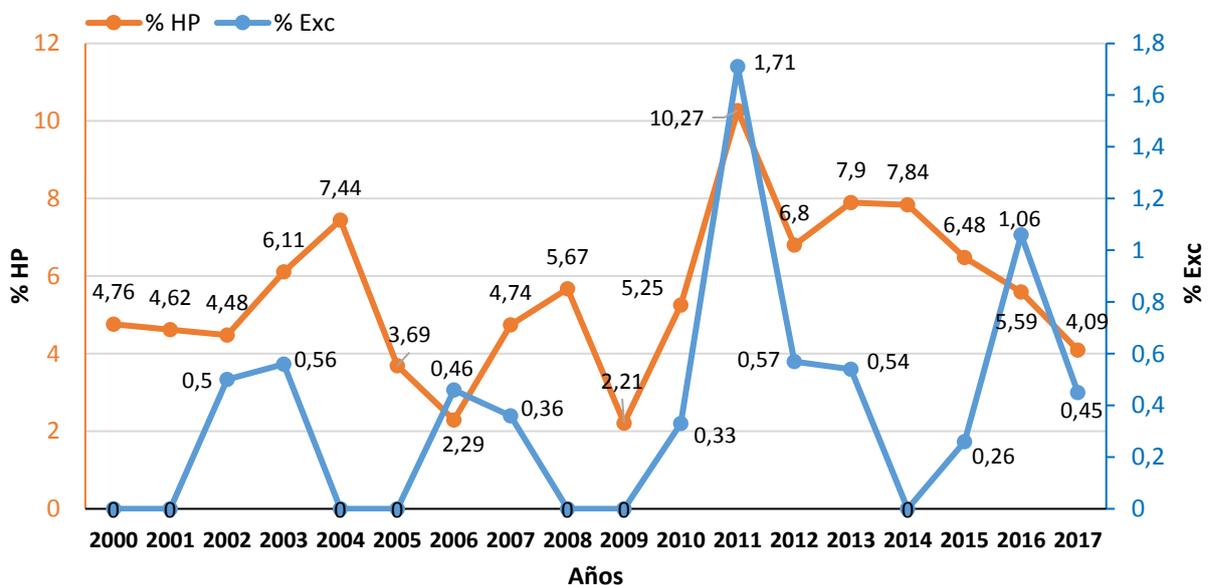


Figura 18. Evolución de los indicadores de excelencia (%Exc) y de alto desempeño (%HP).

Para que una publicación sea considerada como de alto desempeño, está debe haber obtenido un cúmulo de citas superior al 10% de las citas a nivel mundial. Como se observa en la figura 18 el indicador de alto desempeño (%HP) tiene un promedio de 5.57 en el período de estudio. Se identificaron dos años de muy bajo desempeño de los matemáticos del país, los años 2006 y 2009. Destacar el año 2011 que es el pico de mayor alcance para este indicador, además de los años 2004, 2013 y 2014 que en promedio casi alcanzan el ocho por ciento de alto desempeño.

En el caso de la excelencia científica (%Exc) es un poco más difícil de alcanzar, considerando que para calificar en este indicador es necesario que el artículo alcance un volumen de citas superior al uno por ciento a nivel mundial. La excelencia científica de los matemáticos mexicanos tiene un promedio de 0.38% en todo el período, este dato se considera bajo. El valor de este indicador se ve muy afectado porque existen siete años en los que México no posee ningún artículo de excelencia científica. Se debe destacar el año 2016 donde se alcanza un uno por ciento de documentos de excelencia científica, además en el año 2011 se logró 1.7 por ciento de excelencia científica.

### 3.3.4 Análisis de coautoría

La coautoría múltiple es una característica de los campos de ciencias básicas y aplicadas, no obstante, es una tendencia actual realizar las investigaciones en equipos de trabajo. La figura 19 muestra la evolución de los artículos teniendo en cuenta su coautoría.

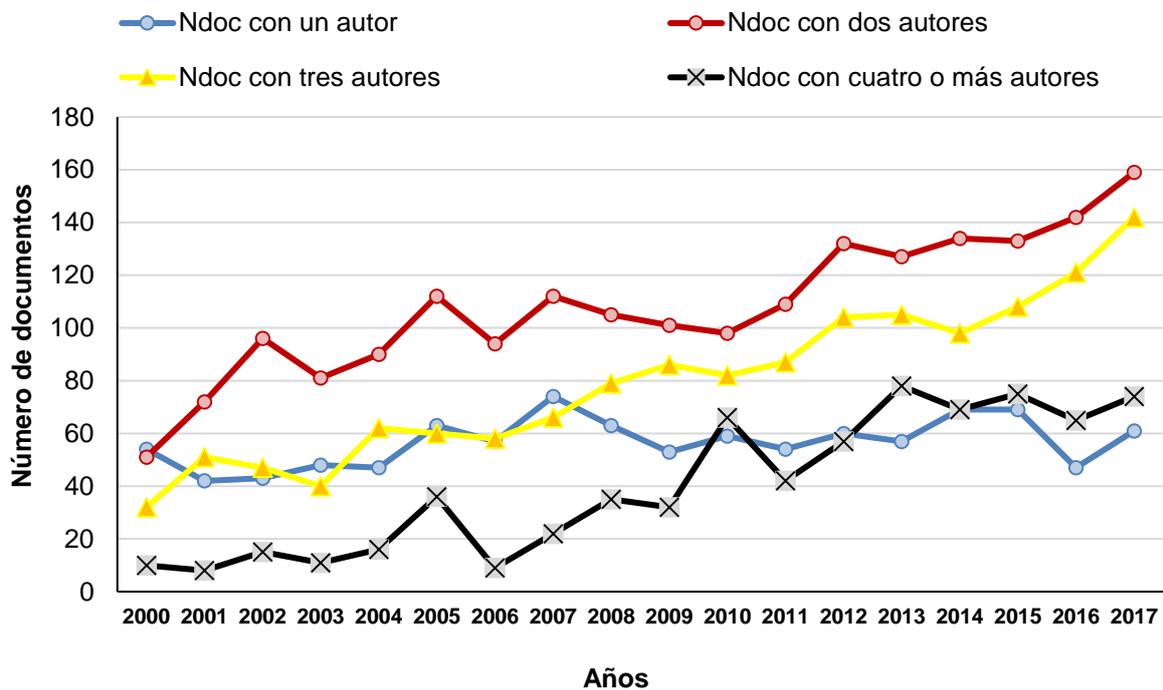


Figura 19. Número de autores por documento publicado.

Estos resultados nos muestran que existe una tendencia del aumento de la coautoría múltiple de los matemáticos mexicanos; en otros estudios a nivel internacional se corrobora esta tendencia (Dang & Zhang, 2003; Behrens & Luksch, 2011). En el período de estudio el promedio de artículos en coautoría simple es de 57 artículos, mientras que, para los trabajos publicados en coautoría de dos y tres autores, es de 108 y 79 artículos en promedio, respectivamente. En este sentido, se observa que a partir del año 2011 hay un aumento de las publicaciones con dos y tres autores, siendo este año el que alcanza el mayor impacto en el período de estudio.

Por otra parte, al analizar específicamente la colaboración internacional vemos que comienza a aumentar a partir de ese año; aunque los resultados sobre el impacto no son observados debido al período de ventana de las citas en este campo. No obstante, la relación impacto-colaboración ha sido estudiada desde varios ángulos donde se ha demostrado que a mayor colaboración internacional se genera mayor cantidad de citas por documento (Russell & et al., 2007; De Filippo & et al., 2008; Zhou & Tian, 2014).

### **3.4 Instituciones mexicanas que producen matemática**

#### **3.4.1 Análisis del impacto de las instituciones**

El volumen de la producción científica de México en el WoS se debe a las investigaciones de 55 instituciones del país que publicaron por lo menos un artículo durante el período de estudio. En la tabla 6 se representan las instituciones con más de 100 documentos publicados en el período (el resto de las instituciones identificadas en este estudio se muestran en el anexo 5). El 40% de las instituciones que conforman el estudio pertenecen a la capital, el otro 40% a entidades federativas diferentes y el 20% restante son de entidades que tienen representación a nivel nacional como: la Universidad Nacional Autónoma de México (UNAM) y el Instituto Politécnico Nacional (IPN).

Tabla 6. Instituciones mexicanas con mayor volumen de producción e impacto.

No	Instituciones	Ndoc	%5117*	Cr	Cr/Ndoc	CD	Cr/CD
1	Universidad Nacional Autónoma de México	2,215	43.29	10,111	5	1,552	6
2	Instituto Politécnico Nacional	544	10.63	3011	6	411	7
3	Universidad Autónoma Metropolitana	490	9.58	1920	4	320	6
4	Centro de Investigación y de Estudios Avanzados	487	9.52	3335	7	371	9
5	Centro de Investigación en Matemáticas	439	8.58	2638	6	295	9
6	Instituto Tecnológico Autónomo de México	132	2.58	691	5	88	8
7	Benemérita Universidad Autónoma de Puebla	125	2.44	357	3	81	4
8	Universidad Autónoma de San Luis Potosí	111	2.17	604	5	83	7
9	Universidad Autónoma del Estado de Morelos	104	2.03	331	3	67	5
10	Universidad Michoacana San Nicolás de Hidalgo	103	2.01	302	3	67	4

La UNAM es la institución más destacada; ya que representa el 43.3% de la producción matemática de México y el 42.6% del impacto que genera la producción del país. No obstante, al realizar el análisis comparativo entre la UNAM y el CIMAT; se pudo constatar que este último tiene un tercio de las citas recibidas de la UNAM, pero la cantidad de citas por publicación y por artículo citado, es superior. Un caso similar ocurre con la UAM y el ITAM, donde la primera tiene casi el triple de las citas de la segunda, en cambio la cantidad de citas por artículo publicado y por artículo citado es inferior. Tener en cuenta estos dos indicadores nos da un valor agregado sobre el impacto en las instituciones. De hecho, el cociente entre las citas y los artículos publicados nos dan una idea del impacto que puede llegar a alcanzar, pero el cociente entre las citas y los artículos citados nos muestra el impacto real que se alcanza.

### 3.4.2 Porcentaje de artículos publicados en cuartil uno versus volumen de citas por documento

Las publicaciones científicas son el instrumento por excelencia de la comunicación científica (Vessuri, 1995), razón por la cual los investigadores buscan publicar en revistas que les proporcionen visibilidad e impacto, considerando una mayor posibilidad de recibir citas en revistas de alto factor de impacto (Van Raan, 2001). En la figura 20 se muestra la relación entre el porcentaje de artículos publicados en el cuartil uno de impacto (%Q1) y la cantidad de citas por documento que reciben las instituciones (Cr/Ndoc), el diámetro de la burbuja denota el porcentaje de documentos en colaboración internacional. En el análisis de las instituciones que publicaron al menos un artículo de matemática en el período, se identificó que existen algunas instituciones que no poseen programas de estudios en matemática, como es el caso de las instituciones de salud (INCMNSZ, HGM e INSP). Sin embargo, poseen investigaciones en el campo.

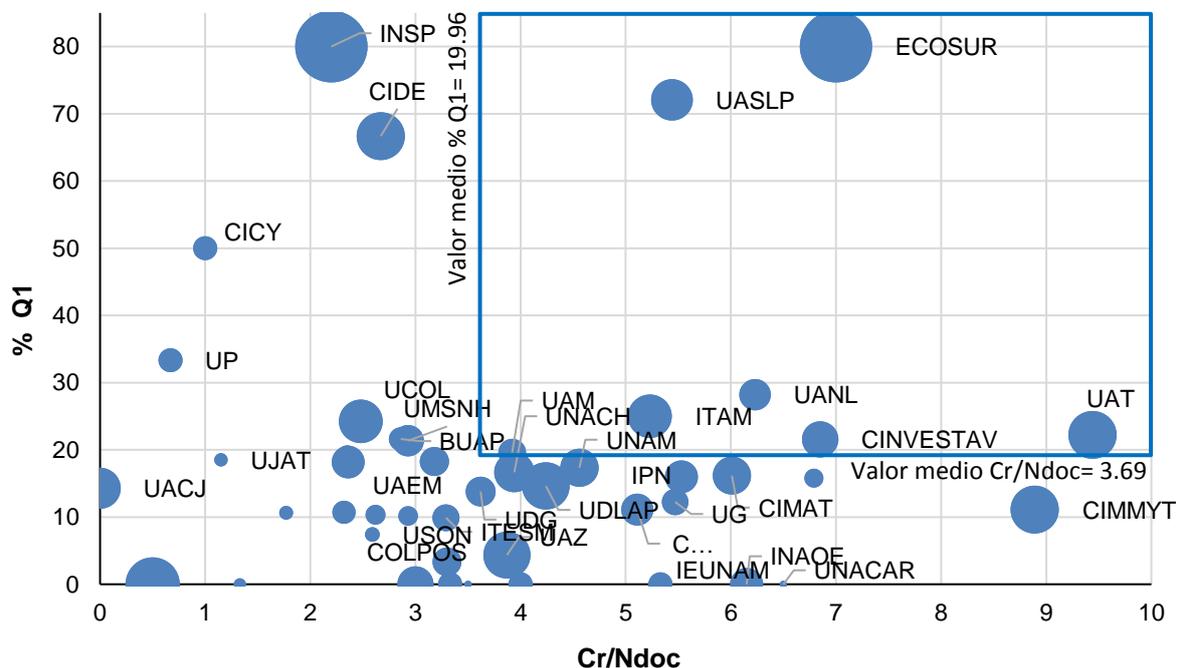


Figura 20. Análisis del porcentaje de documentos en cuartil uno (%Q1) y el volumen de citas por documento (Cr/Ndoc), en la figura no se muestran las instituciones (HGM y INCMNSZ) por tener valores extremos en el %Q1, tampoco se muestra la institución CICESE por tener el valor más extremo en el indicador de Cr/Ndoc.

En la figura se observa en el recuadro superior derecho las seis instituciones que rebasan el promedio de citas/documentos publicados y el porcentaje de artículos publicados en el cuartil uno. Entre ellas tenemos a la Universidad Autónoma de San Luis Potosí (UASLP) que publica el 72% de sus artículos en Q1 sin embargo presenta 5.4 citas/documento; a diferencia de la Universidad Autónoma de Tamaulipas (UAT) que tiene un (22%) de los artículos en Q1, pero logra 8.9 citas/documento. Esta situación se repite en la zona que está por debajo del promedio de las variables que se analizan; tal es el caso del Instituto Nacional de Salud Pública (INSP) que publica cinco artículos en el período, pero cuatro son en revistas del Q1 (80%) y logra 2.2 citas/documento. La línea de investigación de los artículos que esta institución publicó en el período es de la aplicación de técnicas estadísticas a grandes muestras de datos médicos, de hecho, los cinco artículos se publicaron en revistas de estadística. En contraste con la institución anterior, el Centro Internacional del Mejoramiento de Maíz y Trigo (CIMMYT) tiene sólo 11% de los artículos publicados en Q1, pero logra un mayor número de citas/documento (8.6).

Este es un fenómeno que ha sido poco observado en las investigaciones bibliométricas, donde se ha demostrado que el hecho de publicar en revistas de alto factor de impacto no siempre favorece recibir un gran número de citas por documento (Herrera-Vallejera & et al., 2017). No obstante, se debe destacar al Centro de Educación Científica y de Educación Superior de Ensenada (CICESE), el cual presenta el 14.29% de sus artículos en Q1 y la mayor cantidad de citas/documento (13.4).

#### **3.4.3 *Category Normalized Citation Impact (CNCI) y porcentaje en Q1***

En la figura 21 se representa la relación entre el porcentaje de artículos publicados en el cuartil uno (%Q1) y el *category normalized citation impact* (CNCI).

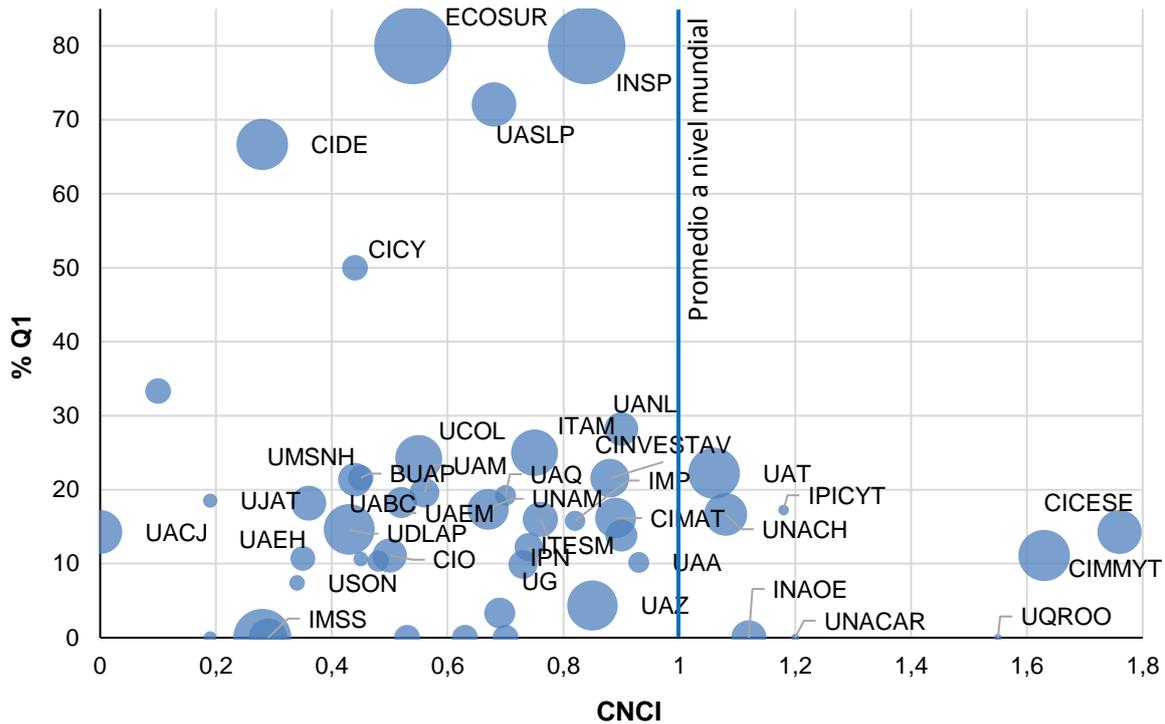


Figura 21. Porcentaje de artículos en cuartil uno (%Q1) y CNCI en el campo, teniendo en cuenta que el diámetro de la burbuja es el porcentaje de colaboración internacional. En la figura no se muestran las instituciones (HGM y INCMNSZ) por tener valores extremos en el %Q1.

En la figura anterior se observa la institución del Colegio de la Frontera Sur (ECOSUR), las investigaciones de esta institución están basados en la aplicación de modelos matemáticos al proceso de la modelación y rotación de cultivos. Continuando con el análisis de la figura anterior, se destacan ocho instituciones que se colocan por encima del promedio mundial de citas por documento en el campo de la Matemática. Entre ellas se destaca la UAT y el CIMMYT, pero este último presenta un mejor valor en el indicador del CNCI con respecto a la primera. En este resultado se evidencia que el CIMMYT a pesar de que publica el 11% de sus artículos en revistas de alto factor de impacto, posee un valor más alto en el indicador CNCI (1.6); por lo que supera a la UAT que publica el doble de sus artículos en revistas de alto factor de impacto (22%), pero sólo logra un 1.1 en el valor del indicador CNCI. Esto nos da una medida del valor que tiene el indicador CNCI para identificar instituciones destacadas en el campo. El CIMMYT publica en el período un total de nueve

artículos, estos se enfocaron en la aplicación de técnicas estadísticas biológicas y agrícolas medioambientales.

Una vez más se destaca el CICESE, por ser la institución que presenta mayor número de citas por documento y la que obtiene el mayor valor en el indicador CNCI (1.76), publicando artículos que reciben citas por encima del promedio mundial en el campo de la matemática y casi duplicando este. Se evidencia que los 14 artículos publicados por el CICESE en el período alcanzaron a tener un impacto destacado para la comunidad matemática, los que acumulan un mayor número de citas se publicaron en la revista *International Journal of Bifurcation and Chaos*. Esta revista se especializa en temas relacionados a la teoría del caos y la ciencia no lineal en la ingeniería y la ciencia aplicada.

### 3.4.4 Porcentaje de documentos citados y CNCI

En la figura 22 se visualiza el porcentaje de documentos citados (%CD) y los valores en el indicador CNCI; teniendo en cuenta en el diámetro de la burbuja, el número de documentos publicados.

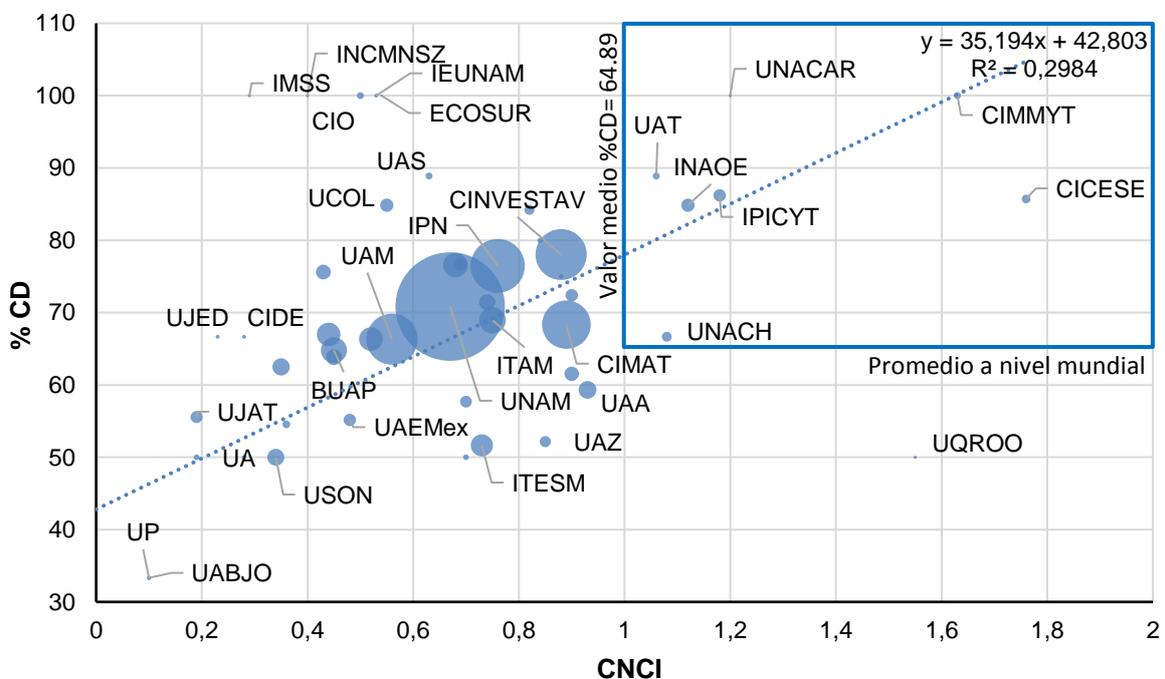


Figura 22. Análisis del porcentaje de documentos citados (%CD) y el CNCI, donde el tamaño de la burbuja representa el volumen de producción científica de la institución.

En este resultado se muestra la tendencia que existe entre las instituciones estudiadas con respecto al porcentaje de artículos citados y el indicador CNCI. A manera general las instituciones con mayor productividad como la UNAM, el CIMAT, la UAM, el IPN y el CINVESTAV; tienen entre el 60 y el 80% de sus artículos citados, pero se encuentran por debajo del promedio mundial en el campo de la Matemática, según el CNCI. Por otra parte, la Universidad de Quintana Roo (UQROO) con dos artículos publicados, logra un valor en el indicador CNCI de 1.6, ubicándola muy cerca del CICESE, el cual sí logra un 86% de sus documentos citados y además en el indicador CNCI logra un valor de 1.76.

### 3.4.5 Análisis de la colaboración internacional y el alto desempeño

En la figura 23 se presenta la relación entre el porcentaje del desempeño (%HP) y el de la colaboración internacional (%IC), teniendo en cuenta que el porcentaje del cuartil uno es el diámetro de la burbuja.

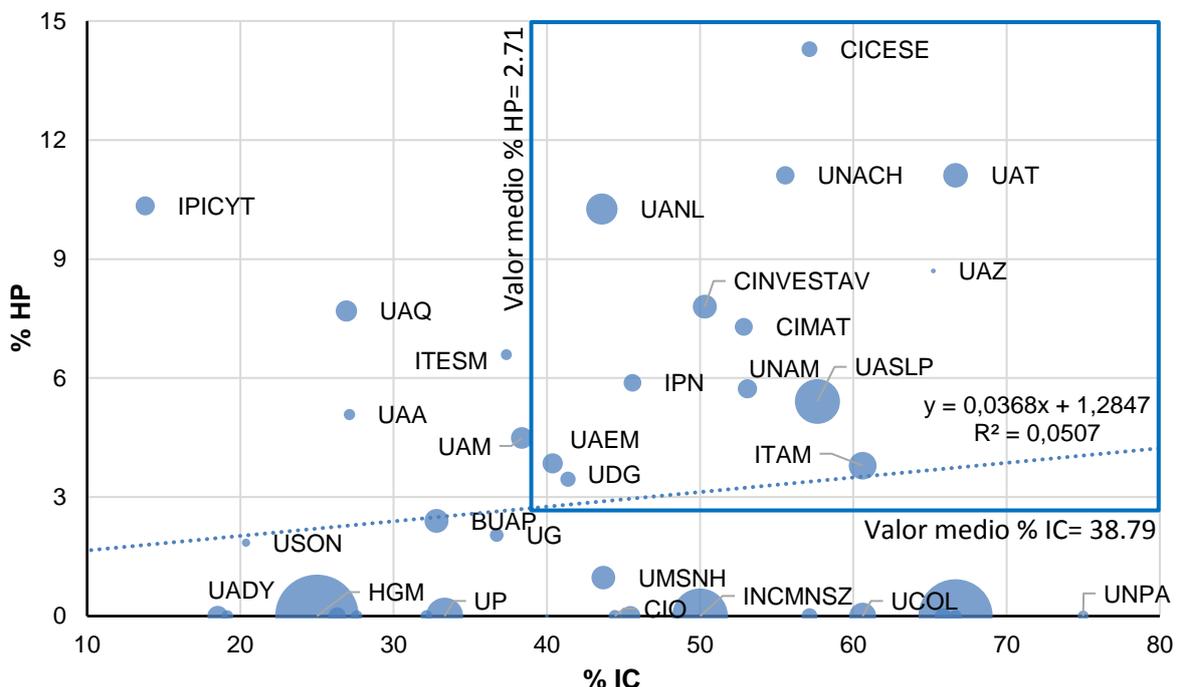


Figura 23. Análisis del desempeño científico (%HP) y la colaboración internacional (%IC), en la figura no se muestran las instituciones (INSP y ECOSUR) por tener valores extremos en el %IC.

En la figura se observan las 14 instituciones que se encuentran por encima del promedio del porcentaje de alto desempeño científico y el porcentaje de la colaboración internacional. Entre ellas se analizan a la UASLP y al CICESE, ambas publicaron un 57% de sus investigaciones en colaboración internacional, sin embargo, el CICESE tiene un 14.29% de sus artículos de alto desempeño científico, a diferencia de la UASLP que sólo tiene un 5.41%. En la zona de afuera de los valores promedio se observa como el Instituto Potosino de Investigación Científica y Tecnológica (IPICYT) tiene un bajo porcentaje de artículos en colaboración internacional 13.79% y sus artículos de alto desempeño superan el 10%; a diferencia de la Universidad Panamericana (UNPA) que tiene un elevado porcentaje de sus artículos en colaboración internacional (75%), pero su porcentaje de artículos de alto desempeño científico es nulo. En este caso se demuestra que el elevado desempeño no está directamente relacionado con la marcada presencia de la colaboración internacional. Este resultado difiere a la literatura donde se ha planteado que la colaboración internacional aumenta elevadas cantidades de citas (Zhou & Tian, 2014; De Filippo & et al., 2008; Russell & et al., 2007).

#### **3.4.6 Análisis multiparamétrico del perfil de las instituciones mexicanas**

Aprovechando las fortalezas que ofrece la inteligencia artificial se utilizó una red neuronal para realizar un análisis simultáneo de cinco indicadores cuantitativos (%Q1, %HP, %CD, CNCI y %IC), este método se denomina: análisis multiparamétrico y ha sido utilizado en otras investigaciones (Villaseñor & et al., 2016; Lozano, 2016).

La red neuronal permite identificar automáticamente los perfiles cuantitativos, observar cuáles instituciones tienen perfiles diferentes a todas las demás y también identificar las instituciones que comparten perfiles cualitativamente similares.

Con la técnica de visualización se construyeron seis mapas (cartografías), un mapa de *clustering* y otros cinco de componentes. Estos últimos son fundamentales porque nos permiten leer el significado de la distribución de los datos en la red neuronal, utilizando una escala cromática se

muestra la distribución del valor de los indicadores. En este caso, se utilizó un mapa de 20 X 20 neuronas que se entrenaron por 4,000 iteraciones.

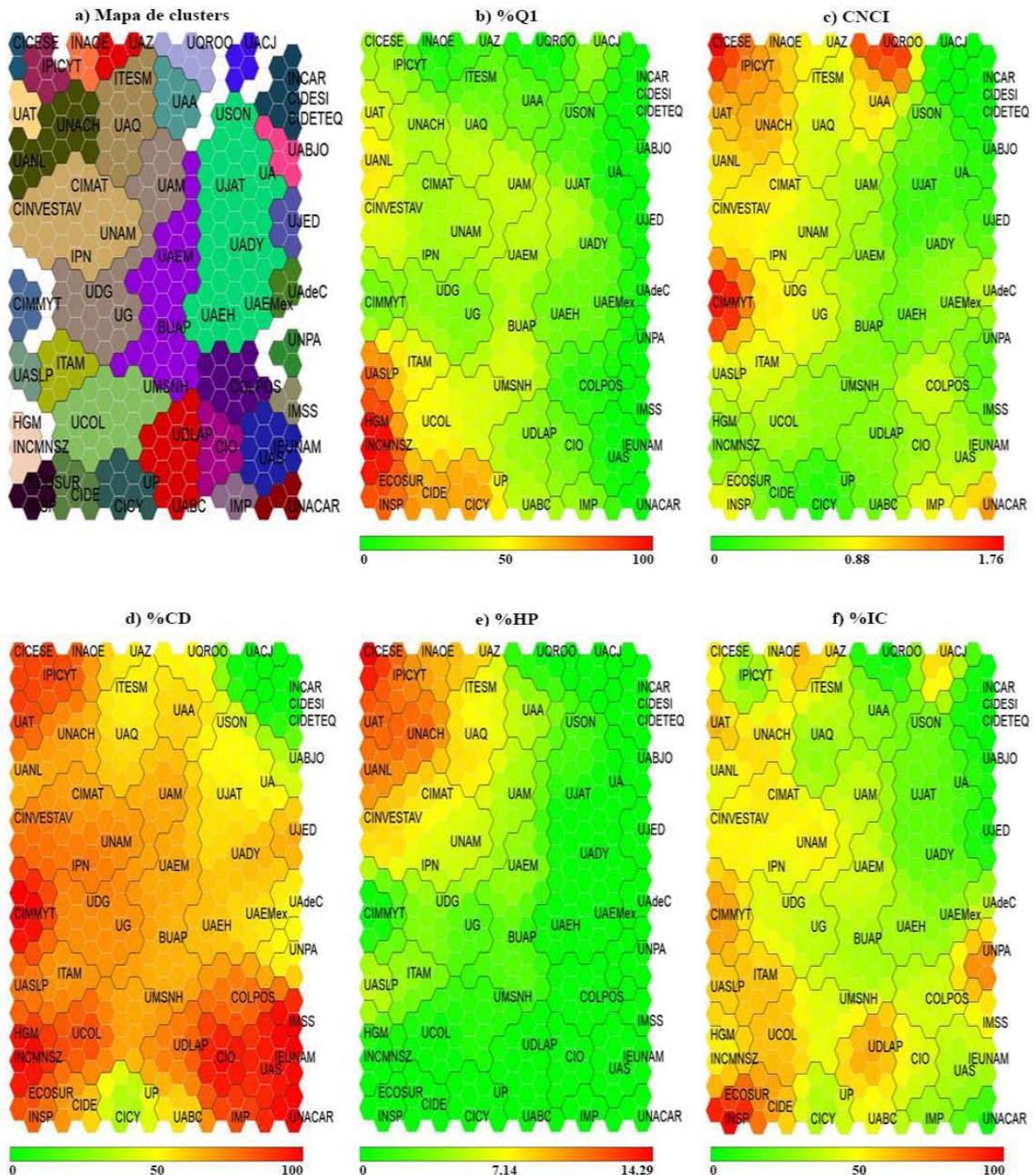


Figura 24. Análisis multiparamétrico de las instituciones que producen matemática de acuerdo con: a) Mapa de *clusters*, b) Porcentaje de artículos publicados en revistas del cuartil uno de impacto (%Q1), c) *Category Normalized Citation Impact* (CNCI), d) Porcentaje de Documentos Citados (%CD), e) Porcentaje de *High Performance* (%HP), f) Porcentaje de artículos

en Colaboración Internacional (%IC). En los anexos 6, 7 y 8 se muestra por separado cada mapa.

En la figura 24a se muestra el mapa de *clusters*; se identificaron 34 perfiles de desempeño cuantitativo diferentes, de los cuales 22 perfiles están constituidos por una sola institución, las cuales poseen un perfil peculiar de investigación y se diferencian del resto del grupo, en inglés se denominan *outliers*.

Las cinco figuras restantes representan los mapas (de componentes) correspondientes a los cinco indicadores considerados en este estudio: %Q1, CNCI, %CD, %HP y %IC. De acuerdo con la escala cromática utilizada, en estos mapas se colorean de verde aquellas regiones del mapa en las que el indicador que representa el mapa asume los valores más bajos y se colorea de rojo aquellas regiones donde se tienen los valores más altos. Por ejemplo, en el mapa 24b se colorean de rojo aquellas regiones del mapa en las que se encuentran instituciones que tienen un porcentaje de artículos en cuartil uno cercano al 100%. En este caso tenemos dos instituciones (HGM y INCMNSZ) que están en una región roja porque tienen %Q1 = 100%, lo cual se debe a que estas son instituciones médicas que sólo publicaron menos de cinco artículos de matemáticas en el período y todos ellos en revistas de cuartil uno. También se observan cinco instituciones (INSP, CIDE, ECOSUR, UASLP y CICY) que están en una región anaranjada, significando que estas tienen porcentajes de Q1 mayores al 50%. En este mismo mapa se identifica una amplia región coloreada de verde; todas las instituciones que están contenidas en esta región tienen entonces un porcentaje de artículos en cuartil uno menor al 50%.

El análisis del mapa 24d revela que la gran mayoría de las instituciones se encuentran en regiones con un color superior al amarillo, de acuerdo con la escala cromática. Esto significa que casi todas las instituciones que publican en matemáticas tienen más de un 50% de sus investigaciones citadas. Son excepciones siete instituciones: (CICY, INCAR, CIDESI, CIDETEQ, UACJ, UABJO, UP).

En el mapa 24e se determina en la región superior izquierda cinco instituciones que tienen valores del indicador %HP > 10 (más de un 10% de sus

artículos están en el 10% de los más citados internacionalmente). Estas son: (UANL, IPICYT, UNACH, CICESE, UAT). También se identifican cinco instituciones que tienen valores de %HP superiores a 7 y menores que 10: (CINVESTAV, CIMAT, INAOE, UAQ, UAZ). Sin embargo, la mayoría de las instituciones tienen valores de este indicador inferiores al 7%.

En el mapa 24c se identifican con el color rojo tres instituciones que tienen los más altos valores del indicador de impacto CNCI: CICESE, CIMMYT y UQROO. Todas ellas con un impacto de citación superior a la media mundial (ver tabla de valores de los indicadores en el anexo 9).

Se identifican también tres instituciones (UAA, UANL, UDG) en regiones con color superior al amarillo y por lo tanto tienen valores de CNCI en el intervalo [0.9, 1], lo cual significa que tienen un impacto cercano a la media mundial en el campo de las matemáticas.

En el caso de los perfiles identificados en el mapa de *clusters* 24a sobresalen instituciones con perfiles muy peculiares como es el caso del Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE). Este centro alcanza a tener el valor más alto en el impacto normalizado (CNCI=1.76) y en el porcentaje de artículos de alto desempeño científico (%HP=14.29), el 86% de sus investigaciones reciben citas y el 57% de estas son publicadas en colaboración internacional; sin embargo, tiene publicado dos artículos en revistas de alto factor de impacto, representando esto un bajo porcentaje (%Q1=14). Esta institución publicó 14 artículos en todo el período, por lo tanto, eso indica que publica menos de un artículo al año. En resumen, es una institución con una baja producción científica, pero las pocas publicaciones que realiza obtienen un alto reconocimiento en la comunidad científica internacional.

Otra institución que destaca por su perfil muy singular es la Universidad de Quintana Roo, esta institución no tiene artículos publicados en las revistas de mayor factor de impacto, no publicó ningún artículo en colaboración internacional y tampoco tiene artículos de alto desempeño científico; sin embargo, el 50% de sus documentos publicados recibieron citas y su impacto normalizado es alto (CNCI=1.55). Esta institución no destaca en casi ningún indicador, sólo destaca en el impacto normalizado lo cual hace que la misma aparezca en un solo *cluster*, con

un perfil independiente. Un último ejemplo de las instituciones agrupadas en un solo *cluster* es el Centro Internacional de Mejoramiento de Maíz y Trigo. Este centro publicó nueve artículos en el período, es decir, en promedio un artículo por cada dos años. De estos artículos seis fueron publicados en colaboración internacional y recibieron ocho citas en promedio; sin embargo, su impacto normalizado es alto (CNCI=1.63) a pesar de tener un artículo publicado en una revista de alto factor de impacto. Por lo tanto, esta institución es de las que menos publicaciones tiene, pero destaca con un 100% de sus documentos citados y un alto valor en el indicador CNCI.

Además de los perfiles individuales, también se identificaron un grupo de *cluster* que agrupan a varias instituciones por sus características. Destacar el caso de un *cluster* que agrupa a cuatro grandes instituciones del país como son: el Centro de Investigación en Matemáticas (CIMAT), el Centro de Investigación y de Estudios Avanzados (CINVESTAV), la Universidad Nacional Autónoma de México (UNAM) y el Instituto Politécnico Nacional (IPN). El perfil cuantitativo de estas cuatro instituciones se caracteriza por tener un impacto normalizado superior al (CNCI=0.7), este valor es el que registra México en el campo de la matemática. Más del 70% de los artículos que publican estas instituciones son citados y casi la mitad de sus publicaciones son realizadas en colaboración internacional. Tienen un bajo porcentaje de artículos publicados en revistas de alto factor de impacto y su desempeño científico a pesar de ser bajo, tiene más del cinco por ciento de sus artículos en el 10% de los más citados en el campo de la matemática. Por lo tanto, es para destacar que estas instituciones poseen valores por encima de la media en los indicadores de impacto y colaboración internacional. Se ha de señalar que estas cuatro instituciones son las de mayor volumen de producción científica y a pesar de eso tienen un bajo porcentaje de artículos publicados en revistas de alto factor de impacto. Para poder observar los perfiles del resto de las instituciones que no se mencionan en el texto, en la tabla del anexo 10 se muestran todos.

### 3.5 Categorías temáticas

#### 3.5.1 Número de documentos y citas recibidas

En la figura 25 se muestra la relación entre las citas recibidas y la producción científica de cada una de las categorías temáticas que componen el campo de la Matemática, de acuerdo con el esquema de clasificación de la Colección Principal del WoS.

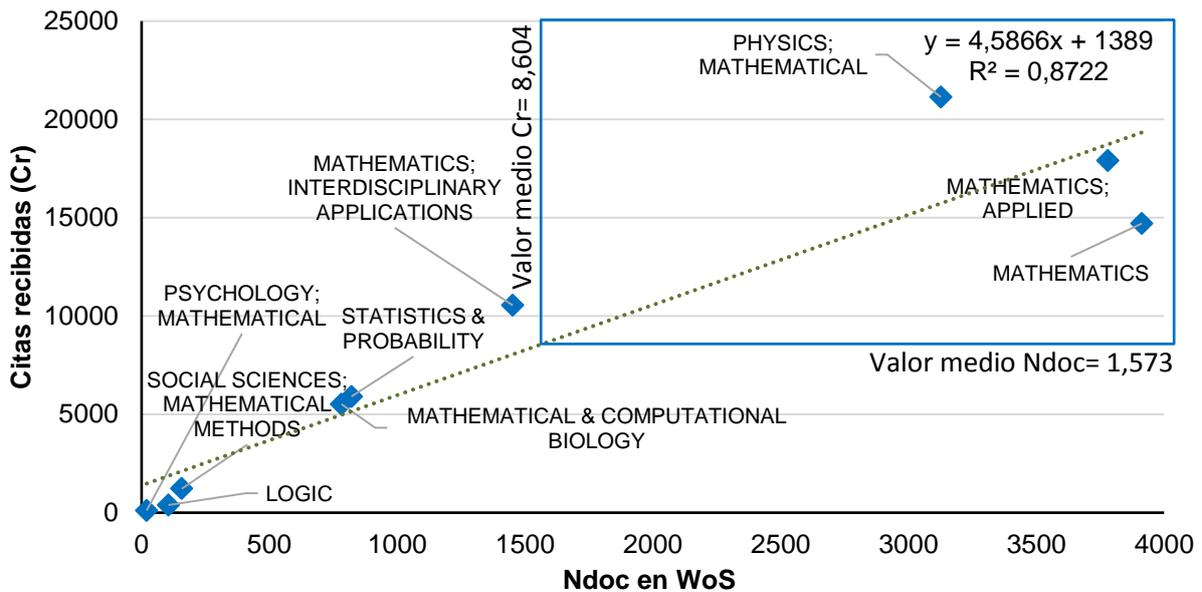


Figura 25. Número de documentos publicados y citas recibidas.

De las nueve categorías temáticas que componen el campo de la Matemática tres (Física-Matemática, Matemática Aplicada y Matemática) están por encima de la media en cuanto al volumen de producción científica y las citas que estas publicaciones acumulan en el período. Entre ellas la tercera posición en cuanto al volumen de producción científica la ocupa Física Matemática con 3,127 artículos publicados, pero se sitúa en el primer lugar en cuanto al impacto científico con 21,136 citas recibidas. Por otro lado, Matemática es la categoría temática con mayor producción científica (3,913 artículos), pero se encuentra en el tercer lugar en cuanto al impacto (14,705 citas).

### 3.5.2 Visibilidad esperada y real de las categorías temáticas

En la figura 26 se representan las categorías temáticas y su posición teniendo en cuenta el porcentaje de los documentos citados entre el 10% de los más citados (%HP) y el porcentaje de documentos publicados en el cuartil uno de impacto (%Q1), considerando el diámetro de la burbuja como el porcentaje de colaboración internacional.

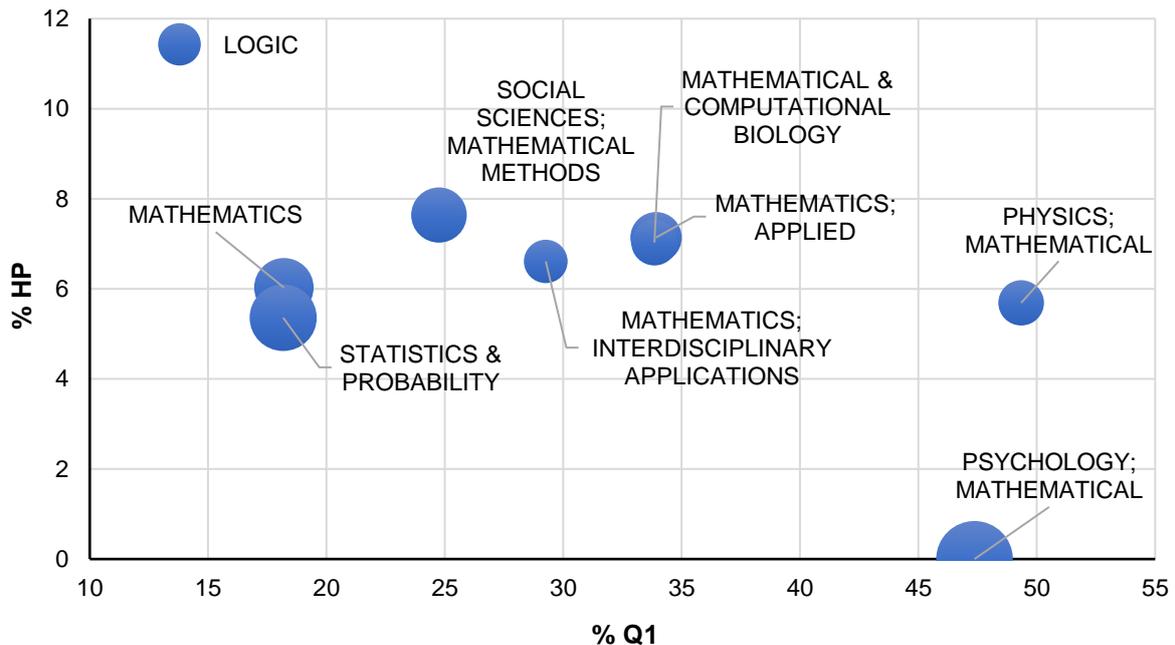


Figura 26. Porcentaje de documentos en cuartil uno (%Q1) y el alto desempeño científico (%HP) de las categorías temáticas, el tamaño de la burbuja representa el porcentaje de colaboración internacional.

Se han analizado dos casos Física-Matemática y Lógica. En el caso de la primera es una de las categorías temáticas con mayor volumen de producción científica y genera un alto impacto (3,127 artículos publicados y 21,136 citas recibidas), el 49% de sus artículos fueron publicados en revistas de alto factor de impacto, pero su porcentaje de documentos de alto desempeño científico (5.7%) es menor que Lógica; la cual publica 105 documentos, de ellos el 14% en revistas de alto factor de impacto y recibe 389 citas, pero alcanza la mayor excelencia científica con 11.4%. Este resultado nos demuestra que el número de citas y de artículos publicados en el cuartil de más alto factor de impacto de la base de datos no implica tener un elevado desempeño científico.

### 3.5.3 Impacto en las categorías temáticas

En la figura 27 se representan las categorías temáticas y su posición teniendo en cuenta el indicador de impacto CNCI y las citas por artículos publicados, considerando el diámetro de la burbuja como el número de artículos publicados. La línea azul que corta la gráfica en el valor de 0.7 del indicador CNCI representa el valor que posee México de manera general en el campo de la matemática. Por lo tanto, se identificó que la mitad de las categorías superan este valor y la otra mitad está por debajo.

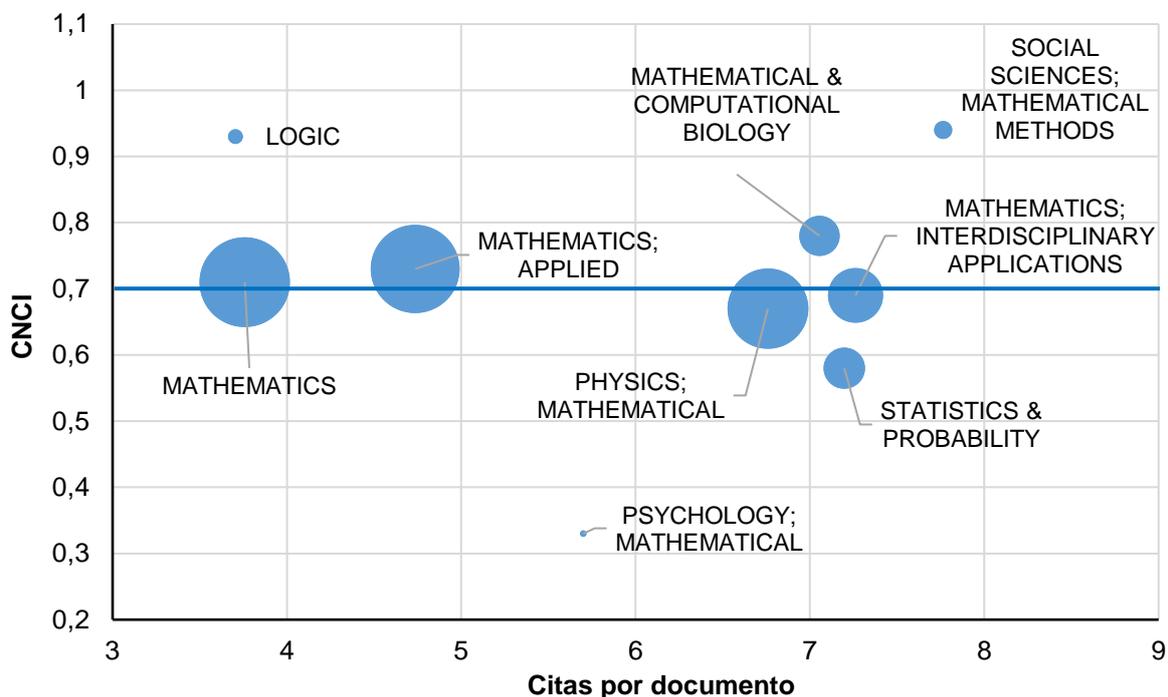


Figura 27. Análisis del indicador CNCI y las citas por documento, el tamaño de la burbuja denota el volumen de producción científica.

Se determinó que ninguna de las categorías temáticas alcanza el promedio a nivel mundial en el campo de la Matemática. Inclusive seis de las nueve temáticas se encuentran entre 0.6 y 0.8 para el valor de este indicador, sólo Lógica y Métodos Matemáticos en Ciencias Sociales logran acercarse con valores de 0.9 cada una.

Por otra parte, se observa que las temáticas de mayor componente Matemático como son: la propia Matemática y Matemática Aplicada son las de mayor producción científica, pero a su vez son las que generan el menor

impacto científico del campo con 3.8 y 4.7 citas/documento, respectivamente. Sin embargo, las temáticas donde se relaciona la Matemática con otros campos como son: Matemática y Biología Computacional, Matemática y Aplicaciones Interdisciplinarias, Estadística y Probabilidades, Física-Matemática, Métodos Matemáticos en Ciencias Sociales son los que generan un mayor número de citas por documento (5.7 - 7.8 citas/documento). Esto demuestra que el mayor impacto de la Matemática no proviene de sus líneas de investigación más fuertes, si no de comunidades científicas de fronteras como: la Biología, la Física, las Ciencias Sociales y la Estadística.

### 3.5.4 Producción de alto desempeño

En la figura 28 se muestra la posición de las categorías temáticas en relación con el indicador CNCI y el alto desempeño científico (%HP), teniendo en cuenta que el diámetro de la burbuja es el porcentaje de documentos en colaboración internacional.

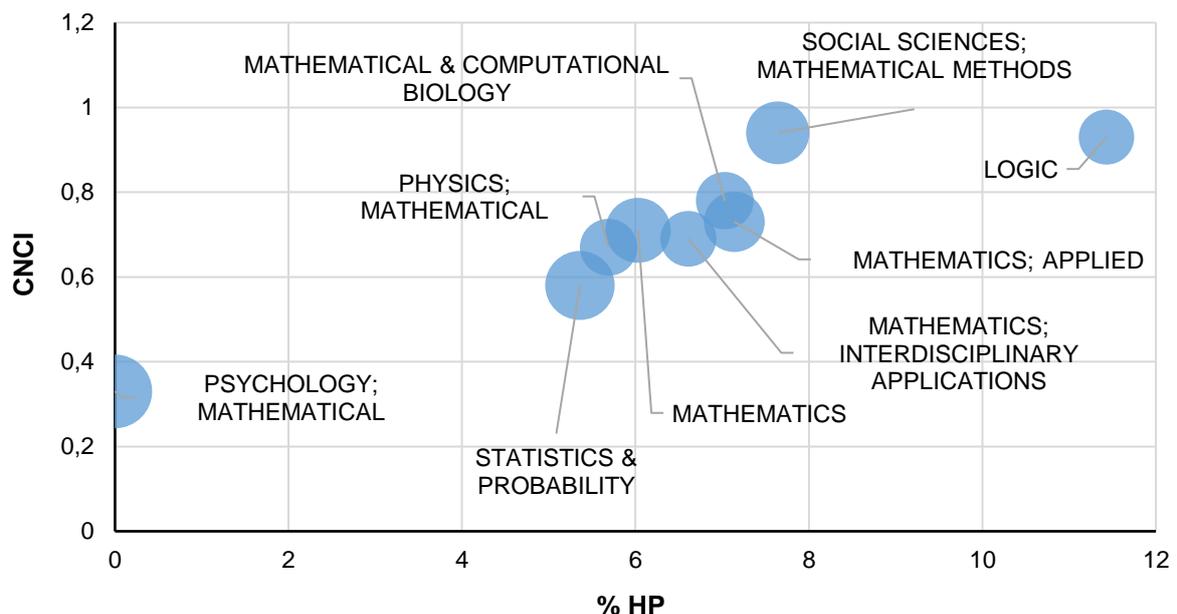


Figura 28. Desempeño científico (%HP) y CNCI de las categorías temáticas, el tamaño de la burbuja denota el porcentaje de colaboración internacional.

La temática de Matemática-Psicología es un caso particular porque a pesar de que publica el 47% de sus artículos en revistas de alto factor de impacto y recibe 5.7 citas/documentos; no tiene ningún artículo de alto desempeño científico.

Por otra parte, se comprobó las amplias colaboraciones internacionales en las nueve temáticas que conforman el estudio (36% - 65%). En cuanto a la relación entre las variables que se analizan se observa un solapamiento entre seis de ellas y nuevamente son las temáticas de Lógica y Métodos Matemáticos en Ciencias Sociales las que logran destacarse del resto de las temáticas; ya que además de tener un valor en el indicador CNCI cercano al promedio a nivel mundial (0.9), cuentan con el mayor porcentaje de artículos de alto desempeño científico, con 11.4% y 7.6%, respectivamente. En este resultado se corrobora el alto desempeño de estas dos temáticas en particular dentro del campo de la Matemática.

### 3.5.5 Excelencia científica e impacto normalizado

En la figura 29 se muestra la posición de las categorías temáticas en relación con el indicador CNCI y el porcentaje de documentos en el *Top 1%*, considerado como un indicador de excelencia científica (%Exc) en la investigación, ya que sólo los artículos más citados obtendrían el uno por ciento más alto en su campo, año y tipo de documento respectivos. Este indicador se puede aplicar a cualquier nivel de agregación (países, campos, instituciones e investigadores). En la siguiente figura el diámetro de la burbuja denota el porcentaje de desempeño científico.

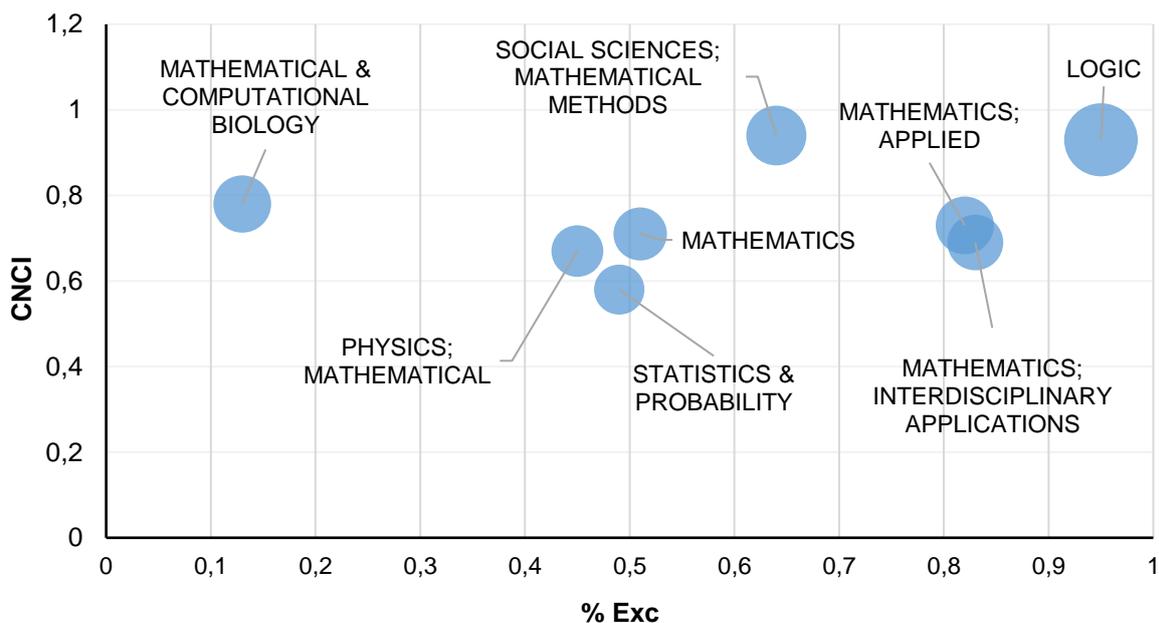


Figura 29. Excelencia científica (%Exc) y CNCI de las categorías temáticas.

En este resultado se evalúan las temáticas que ejercen el mayor impacto en la comunidad científica. Su objetivo es lograr diferenciar la posición que ocupan las temáticas que se visualizaron en la figura anterior solapas entre sí. Por supuesto, la temática de Matemática-Psicología desaparece en este resultado debido a que no posee ningún artículo de alto desempeño científico y mucho menos tiene artículos de excelencia científica.

El resto de las temáticas se agrupan en dos grupos bien definidos donde están Matemática, Física Matemática, Estadística y Probabilidades con un rango de 0.4%-0.6% de artículos citados entre el 1% de los más citados y por otra parte está Matemática Aplicada aún solapada con Matemática Aplicaciones Interdisciplinarias con 0.8% de sus artículos entre el 1% de los más citados. Con respecto a estas dos últimas temáticas se analizó que logran aumentar su excelencia por encima de Métodos Matemáticos en Ciencias Sociales. Esto quiere decir que, aunque el desempeño de las dos primeras sea menor, hay un mayor porcentaje (0.8%) que logra una mayor excelencia por encima de Métodos Matemáticos en Ciencias Sociales. Por último, se destaca Lógica ratificando su posición no sólo con un alto desempeño (11.4%); sino también con una elevada excelencia (0.95%) dentro del campo de la Matemática.

### **3.5.6 Análisis multiparamétrico de las categorías temáticas**

Aprovechando las fortalezas que ofrece la inteligencia artificial se utilizó una red neuronal para realizar un análisis simultáneo de seis indicadores cientiométricos (CNCI, %HP, %Exc, %CD, %Q1 y %IC).

La red neuronal permite identificar automáticamente los perfiles cientiométricos de las nueve categorías temáticas que componen el campo de la matemática, según el esquema de clasificación de la Colección Principal del WoS.

Con la técnica de visualización se construyeron siete mapas, un mapa de *clustering* y otros seis de componentes. En este caso, se utilizó un mapa de 10 X 10 neuronas que se entrenaron por 100 iteraciones.

En la figura 30a se muestra el mapa de *cluster*, se identificaron ocho perfiles de desempeño cientiométrico diferentes, de los cuales un sólo perfil comparten dos categorías temáticas, el resto posee un perfil peculiar de investigación y se diferencian del resto del grupo, en inglés se denominan *outliers*.

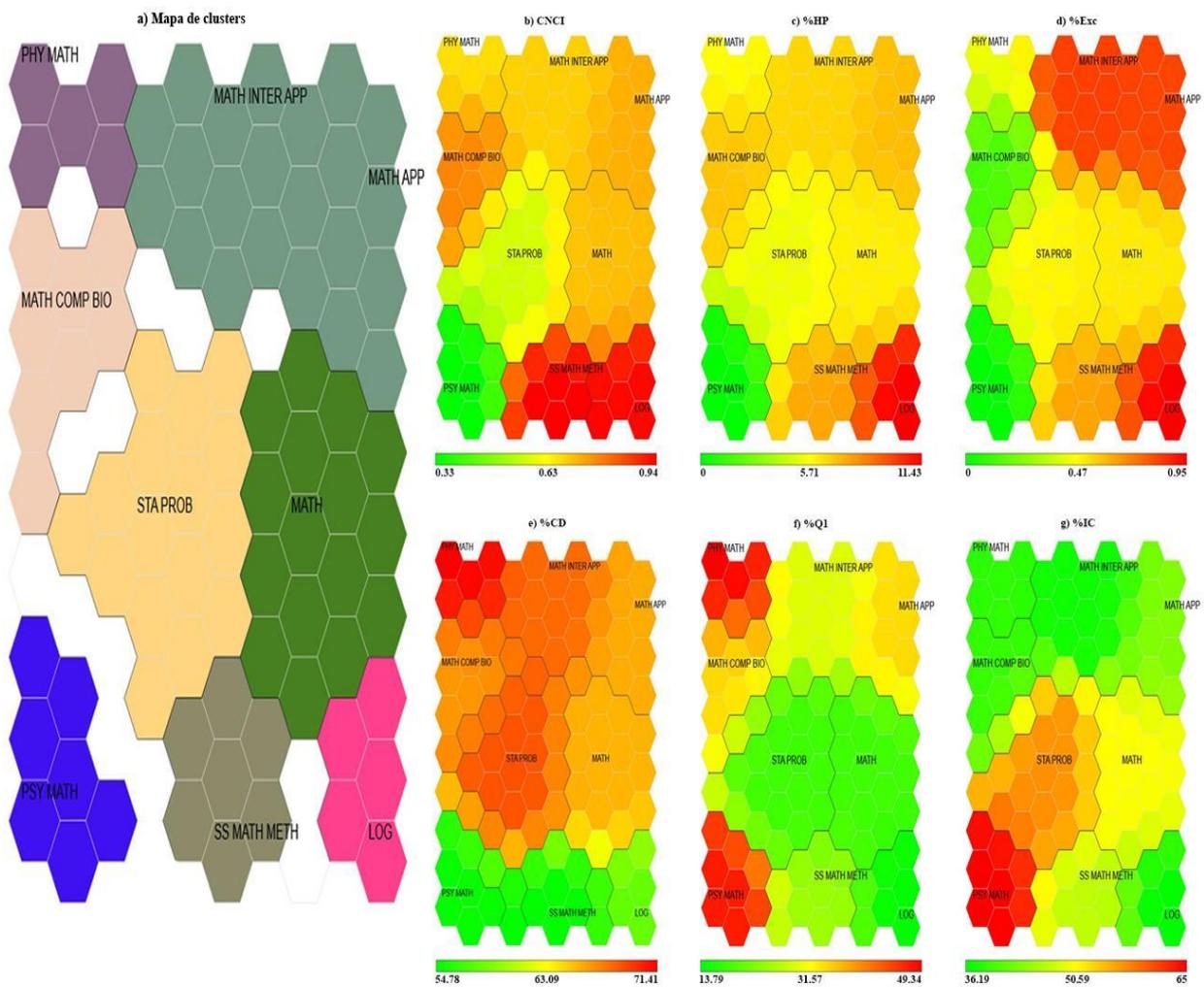


Figura 30. Análisis multiparamétrico de las categorías temáticas que componen el campo de la matemática: a) Mapa de *clusters*, b) *Category Normalized Citation Impact* (CNCI), c) porcentaje de *High Performance* (%HP), d) porcentaje de Excelencia Científica (%Exc), e) porcentaje de Documentos Citados (%CD), f) porcentaje de artículos publicados en revistas del Cuartil Uno (%Q1), g) porcentaje de artículos en Colaboración Internacional (%IC). En los anexos 11, 12, 13 y 14 se muestran los mapas por separado.

Las figuras 30b-30g representan los mapas (de componentes) correspondientes a los seis indicadores considerados en este estudio: CNCI, %HP, %Exc, %CD, %Q1 y %IC. De acuerdo con la escala cromática utilizada, en estos mapas se colorean de verde aquellas regiones del mapa en las que el indicador asociado asume los valores más bajos y se colorea de rojo aquellas regiones donde se ubican los valores más altos. Por ejemplo, en el mapa 30b se colorean de rojo aquellas regiones del mapa en las que se encuentran categorías temáticas con un valor en el indicador CNCI cercano a la media mundial en el campo. En este caso tenemos dos categorías temáticas (SS MATH METH y LOG) que están en una región roja porque tienen un valor del indicador CNCI superior al 0.9. También se observa un área anaranjada bastante grande que cubre los valores de cinco categorías temáticas (MATH COMP BIO, MATH APP, MATH INTER APP, MATH y PHY MATH), esto significa que tienen valores de CNCI en el intervalo [0.6, 0.8]. En este mismo mapa se identifica en la parte inferior izquierda una región coloreada de verde que representa las categorías temáticas con valores inferiores al 0.6.

El análisis del mapa 30c (asociado al indicador %HP) se observa una alta correlación con lo obtenido en el mapa 30b (asociado al indicador de impacto normalizado), se identifica coloreada en color rojo una sola categoría temática (LÓGICA) con el valor más alto del indicador de alto desempeño (11%). Esta figura también revela que la gran mayoría de las categorías temáticas se encuentra en regiones con un color amarillo o superior, de acuerdo con la escala cromática. Esto significa que casi todas las categorías temáticas tienen un valor de este indicador superior al 5%. Es excepción la categoría temática (PSY MATH) que aparece en la región del mapa coloreado de verde, porque posee un valor nulo en este indicador.

En el mapa 30d (asociado al indicador %Exc) se identifica con color rojo tres categorías temáticas (MATH INTER APP, MATH APP y LOG) con un porcentaje de excelencia superior al 0.8%. Se observa también dos categorías temáticas (PSY MATH y MATH COMP BIO) con valores a 0.2%, por lo tanto, su excelencia científica es baja.

En el caso del mapa 30e (asociado al indicador %CD) se identificó un grupo de seis categorías temáticas que más del 65% de sus artículos fueron citados. Tres categorías temáticas (PSY MATH, SS MATH METH y LOG) aparecen coloreadas de verde con valores inferiores al 60%.

En el mapa 30f (asociado al indicador %Q1) se observan sólo dos categorías temáticas (PHY MATH y PSY MATH) coloreadas en rojo, las cuales tienen más del 40% de sus artículos publicados en revistas de cuartil uno. En las regiones amarillas se observan tres categorías temáticas (MATH INTER APP, MATH APP y MATH COMP BIO) que tienen aproximadamente un 30% de sus artículos en revistas de cuartil uno. Se identificaron cuatro categorías (en región verde) que tienen menos de un 30% de sus artículos en revistas de cuartil uno.

En el mapa 30g (asociado al indicador %IC) se observa una categoría temática (PSY MATH) que destaca con el valor máximo para este indicador (65%). Seis categorías temáticas tienen menos del 50% de sus artículos en colaboración internacional.

En el caso de los perfiles identificados en el mapa de *clusters* (30a) se distinguen dos categorías temáticas (Matemática Aplicada y Matemática Aplicaciones Multidisciplinarias), por poseer un perfil muy similar de investigación y en esto se diferencian del resto del grupo. Su perfil se caracteriza por tener un porcentaje considerable de artículos de excelencia científica (%Exc=0.82), y más del 38% de sus publicaciones en colaboración internacional. También, el 30% de sus investigaciones están siendo publicadas en revistas de alto factor de impacto, lo cual es congruente con el hecho de que más del 65% de sus investigaciones son citadas. En cuanto al impacto normalizado, su desempeño es medio (CNCI=0.70), no alcanzan a estar cerca del impacto promedio a nivel mundial en el campo y su alto desempeño científico oscila alrededor del seis por ciento. Puede entenderse que tengan perfiles similares ya que la comunidad de investigadores de ambas trabaja en la aplicación de la matemática en otros campos de conocimiento. En resumen, el perfil de estas dos categorías temáticas se destaca por tener una pequeña comunidad de investigadores consolidados que publican artículos de alta

excelencia científica y de alto desempeño, sin embargo, no tienen una gran cantidad de documentos en revistas de alto factor de impacto y la colaboración internacional de sus investigaciones es relativamente baja.

Del resto de los *cluster* identificados individualmente, existe una categoría temática con un perfil muy peculiar (*outlier*). Se trata de la categoría temática de Lógica, se evidencia que existe un grupo consolidado de autores publicando excelentes trabajos de investigación, es cierto que publican pocos trabajos, pero estos tienen un gran impacto en la comunidad científica internacional. Se ve reflejado en los altos valores que obtiene esta comunidad de autores en el impacto normalizado (CNCI=0.93), tanto es así, que están muy cercanos al promedio mundial del impacto normalizado en su campo. Un 11% de las publicaciones de esta comunidad son de alto desempeño científico y tienen un porcentaje considerable de artículos de excelencia científica (%Exc=0.95). Lo interesante de esta comunidad de autores es que logran cumplir con estos estándares internacionales a pesar de tener un bajo porcentaje de artículos publicados en revistas de alto factor de impacto y menos del 40% de sus artículos son realizados en colaboración internacional.

Totalmente opuesto al perfil de la categoría temática Lógica se encuentra la categoría temática de Psicología Matemática, la cual publica casi la mitad de sus investigaciones en revistas de alto factor de impacto y el 65% de sus publicaciones son en colaboración internacional, sin embargo, en los indicadores de impacto normalizado, excelencia científica y alto desempeño sus valores son nulos. En este sentido, los trabajos de investigación de esta comunidad de autores tienden a ser publicados en las revistas de más alto factor de impacto y tienen mucha colaboración internacional, pero sus investigaciones no tienen un reconocido impacto en el gremio internacional del campo de la matemática. Para poder observar el resto de los perfiles identificados por las categorías temáticas que no se mencionan en el texto, en la tabla del anexo 15 se muestran todos.

En el análisis realizado a las categorías temáticas se identificó que existe un núcleo consolidado de investigadores en el área de Lógica, los cuales producen un conjunto de artículos que conforman una élite (excelencia científica, desempeño científico y CNCI), a diferencia de áreas como Matemática Aplicada y Matemática Aplicaciones Multidisciplinarias que tienen un notable porcentaje de investigaciones de excelencia científica, pero esta comunidad de autores al parecer no poseen aún la solidez necesaria para estar al nivel de los artículos de Lógica. Al igual que sucedió con las instituciones, en el caso de las categorías temáticas, las de mayor porcentaje de artículos publicados en revistas de alto factor de impacto no son las de mayor excelencia, impacto y alto desempeño científico. Por lo tanto, para el gremio de las matemáticas, publicar en revistas de alto factor de impacto y en colaboración internacional no garantiza que las investigaciones tengan un impacto en la comunidad científica, sino que el impacto está dado por la calidad de las investigaciones. En cuanto al análisis de la colaboración internacional todas las categorías temáticas exceptuando Psicología Matemática, tienen un porcentaje de colaboración internacional por encima de la media a nivel mundial.

## Consideraciones finales

Fue fundamental para esta investigación utilizar los indicadores de la Bibliometría Evaluativa, las herramientas bibliométricas del WoS, la metodología del *KDD* y la metodología ViBlioSOM. Las especialidades métricas de la información en la actualidad constituyen uno de los frentes de investigación más utilizados para describir la evolución de la ciencia.

El sistema de indicadores bibliométricos utilizados en la investigación fue muy útil para identificar la evolución del campo de la matemática de México. Los indicadores independientes de tamaño se utilizaron con el fin de obtener resultados y comparaciones más justas entre las entidades de análisis. Por ejemplo, se identificaron centros de investigación que poseen muy buenos resultados en el impacto y excelencia científica, a pesar de poseer una plantilla de investigadores inferior a otras grandes universidades del país.

El período de análisis de 18 años (2000-2017) permitió observar el desempeño de la comunidad científica mexicana de la Matemática, en comparación a la evolución internacional. A inicios del período en el año 2000 el campo de la matemática a nivel internacional registraba un total de 21,525 documentos publicados y para finales del período en el año 2017 se habían publicado un total de 43,473 documentos; con una tasa promedio de crecimiento anual del 4%; logrando duplicar el volumen de producción científica en 13 años. Sin embargo, el volumen de la producción científica de México aumenta a mayor ritmo (tasa de crecimiento promedio anual del 7%) y se duplica con mayor velocidad (10 años).

En cuanto al índice de desarrollo científico, México era el segundo país de mayor volumen de producción científica de la región, después de Brasil, pero cuando se realizó el análisis por el volumen de habitantes de cada país, México es desplazado a un sexto lugar, superado por Chile, Uruguay, Argentina, Puerto Rico y Brasil. Por lo tanto, la productividad a nivel de países se debería de medir teniendo en cuenta la cantidad de artículos científicos por habitantes, ya que

este indicador nos muestra una panorámica más real del desarrollo científico (independiente de tamaño).

En el período se identificó que existe una tendencia a nivel internacional, regional e incluso de México, de aumentar los artículos publicados en colaboración internacional.

En la actualidad, existe una tendencia internacional de publicar en revistas de alto factor de impacto, motivado de alguna manera por las presiones que ejercen los comités científicos de evaluación o por la alta expectativa que existe de recibir citas. En esta investigación durante el período de estudio se identificó que existe un aumento paulatino de las publicaciones en revistas de alto factor de impacto a nivel nacional e internacional.

Los indicadores de citas utilizados en esta investigación nos muestran un nivel de análisis más complejo, entre ellos están los artículos altamente citados (HCP), el porcentaje de alto desempeño (%HP) y el porcentaje de artículos de excelencia científica (%Exc). En esta investigación, los mejores valores alcanzados por México se evidencian en el año 2011, relacionado estos altos valores a un artículo publicado en una revista de cuartil tres de impacto que alcanza a recibir un total de 540 citas, lo que confirma que el impacto científico depende de la calidad de la investigación.

La metodología ViblioSOM permitió de manera automática realizar un análisis simultáneo de varios indicadores a través de una red neuronal. Se identificaron los perfiles de desempeño cientiométrico de las instituciones que producen Matemática en el país y se identificaron los perfiles de desarrollo de las categorías temáticas que conforman el campo de la matemática en el WoS. En el caso de las instituciones se obtuvo resultados muy interesantes, se identificaron un total de 34 perfiles de desempeño. Se observó que las instituciones que poseen un mayor porcentaje de documentos publicados en revistas de alto factor de impacto son las de mayor porcentaje de artículos publicados en colaboración internacional; sin embargo, en lo relacionado al impacto estas instituciones no son las de mejor desempeño. En el impacto normalizado se destacan dos centros de investigación y una universidad (CICESE, CIMMYT y UQROO), las investigaciones de estas instituciones

tienen un alto reconocimiento en la comunidad matemática internacional a pesar de tener muy pocos artículos publicados en el período. Por otra parte, se deben destacar otras instituciones insignias del desarrollo científico del país (CINVESTAV, CIMAT, IPN, UNAM), siendo las más productivas en el país, su perfil cuantitativo se caracteriza por tener valores por encima de la media nacional en todos los indicadores utilizados.

En el análisis realizado a las comunidades científicas por área de investigación se identificaron ocho perfiles de desempeño cuantitativo, uno de los resultados más interesantes fue, la agrupación de dos categorías temáticas (MATH INTER APP, MATH APP) en un mismo perfil, ambas categorías agrupan a la comunidad de autores que realizan investigaciones aplicadas fuera y dentro del campo de la matemática. Un perfil muy peculiar (outliers) identificado en el estudio fue la categoría temática de Lógica, esta comunidad de autores publica excelentes trabajos de investigación y su impacto está muy cerca de la media a nivel mundial en el campo. Sin embargo, no tienen un elevado porcentaje de sus documentos citados, publican muy pocos trabajos en colaboración internacional y en revistas del cuartil uno de impacto.

Se espera que esta investigación contribuya como instrumento estratégico para los decisores y gestores de políticas científicas, debido a que revela la estructura del campo de la matemática mexicana, a través del análisis de su producción científica durante 18 años. Además, la metodología utilizada constituye una herramienta muy útil para identificar la evolución en otros campos de conocimiento de la ciencia mexicana.

## Referencias

- Abrizah, A., Zainab, A. N., Kiran, K. & Raj, R. G. (2013). LIS journals scientific impact and subject categorization: A comparison between Web of Science and Scopus. *Scientometrics*, 94(2), 721-740. <https://doi.org/10.1007/s11192-012-0813-7>
- Acosta-Ramírez, D., Ordaz-Cortes, A. & Baquero-Parra, R. (2015). Análisis del desarrollo científico en la península de Yucatán: el punto de vista de los investigadores. *International Journal of Educational Research and Innovation*, (4), 111-122. Recuperado de <https://www.upo.es/revistas/index.php/IJERI/article/view/1464>
- Adem-Chahín, J. (1991). Antecedentes históricos de las matemáticas en México. *Avance y perspectiva*, (10), 155-158.
- Adler, R., Ewing, J. & Taylor, P. (2008). Joint Committee on Quantitative Assessment of Research: Citation Statistics. *A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)*. Recuperado de: <https://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf>
- Ahmed, T., Johnson, B., Oppenheim, C. & Peck, C. (2004). Highly cited old papers and the reasons why they continue to be cited. Part II. The 1953 Watson and Crick article on the structure of DNA. *Scientometrics*, 61(2), 147-156. <https://doi-org.pbidi.unam.mx:2443/10.1023/B:SCIE.0000041645.60907.57>
- Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159-170. <https://doi.org/10.3152/147154403781776645>
- Aksnes, D. W. & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. *Research Evaluation*, 13(1), 33-41. <https://doi.org/10.3152/147154404781776563>
- Alfaraz, P. H. & Calvino, A. M. (2004). Bibliometric study on food science and technology: Scientific production in Iberian-American countries (1991-2000). *Scientometrics*, 61(1), 89-102. <https://doi.org/10.1023/B:SCIE.0000037365.53469.91>
- Alonso-Gamboa, J. O. & Russell, J. M. (2012). Latin American scholarly journal databases: a look back to the way forward. *Aslib Proceedings*, 64(1), 32-45. <https://doi.org/10.1108/00012531211196693>
- Archambault, É., Campbell, D., Gingras, Y. & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the Association for Information Science and Technology*, 60(7), 1320-1326. <https://doi.org/10.1002/asi.21062>
- Arencibia, R. J. (2010). *Visibilidad Internacional de la Ciencia y Educación Superior Cubanas: desafíos del estudio de la producción científica*. Tesis Doctoral. Universidad de Granada, Granada.
- Arencibia-Jorge, R., Atenogenes-Villaseñor, E., Lozano-Díaz, I. A. & Carrillo-Calvet, H. (2016). Elsevier's Journal Metrics for the Identification of a Mainstream Journals Core: A Case Study on Mexico. *Library and Information Science Research Electronic Journal*, 26(1), 1-13.

- Recuperado de: <https://pdfs.semanticscholar.org/9530/ebe3da91cb2198abd032700a8c2e6cda142f.pdf>
- Arunachalam, S. (2001). Mathematics research in India today: What does the literature reveal? *Scientometrics*, 52(2), 235-259. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:1017915823434>
- Arvanitis, R., Russell, J. M. & Rosas, A. M. (1996). Experiences with the national citation reports database for measuring national performance: The case of Mexico. *Scientometrics*, 35(2), 247-255. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02018482>
- Bailón-Moreno, R., Jurado-Alameda, E., Ruiz-Baños, R. & Courtial, J. P. (2005). Analysis of the field of physical chemistry of surfactants with the Unified Scientometric Model. Fit of relational and activity indicators. *Scientometrics*, 63(2), 259-276. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-005-0212-4>
- Baldi, S. & Hargens, L. L. (1995). Reassessing the N-Rays Reference Network - the Role of Self Citations and Negative Citations. *Scientometrics*, 34(2), 239-253. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02020422>
- Banco Mundial (2018). *Search and Share Development Data*. Recuperado de: <https://datacatalog.worldbank.org/>
- Behrens, H. & Luksch, P. (2011). Mathematics 1868-2008: a bibliometric analysis. *Scientometrics*, 86(1), 179-194. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-010-0249-x>
- Bence, V. & Oppenheim, C. (2004). The influence of peer review on the research assessment exercise. *Journal of information Science*, 30(4), 347-368. <https://doi.org/10.1177/0165551504045854>
- Berry, M. & Linoff, G. (1997). *Data Mining Techniques for Marketing, Sales, and Customer Support*. New York, Estados Unidos: John Wiley.
- Berthold, M. & Hand, D. J. (2000). *Intelligent data analysis: An Introduction*. Berlín, Alemania: Springer.
- Bertin, J. (1999). Graphics and graphic information processing. En Stuart K. Card, Jock D. Mackinlay & Ben Shneiderman. (Ed.), *Readings in information visualization* (pp. 62-65). San Francisco, Estados Unidos: Morgan Kaufmann Publishers.
- Besemer, H. & Parr, M. (2013). *Scholarly use of R4D documents: a bibliometric exploration*. United Kingdom: CIMRC.
- Bonilla, C. A., Merigo, J. M. & Torres-Abad, C. (2015). Economics in Latin America: a bibliometric analysis. *Scientometrics*, 105(2), 1239-1252. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-015-1747-7>
- Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation*, 23(2), 166-173. <https://doi.org/10.1093/reseval/rvu002>
- Bornmann, L. & Leydesdorff, L. (2012). 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. <https://doi.org/10.1016/j.joi.2011.11.002>
- 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. <https://doi.org/10.1108/OIR-12-2012-0214>

- Börner, K. (2005). *Extracting and visualizing semantic structures in retrieval results for browsing*. Conference. San Antonio, Estados Unidos: ACM Digital Libraries.
- Börner, K., Chen, C. & Boyack, K. (2003). Visualizing Knowledge Domains. In Blaise Cronin (Ed.). *Annual Review of Information Science & Technology*, 37, 179-255. <https://doi.org/10.1002/aris.1440370106>
- Brachman, R. J. & Anand, T. (1996). *The process of knowledge Discovery in databases*. *Advances in Knowledge Discovery and Data Mining*. California, Estados Unidos: American Association for Artificial Intelligence.
- Bravo-Vinaja, A. & Sanz-Casado, E. (2008). Análisis bibliométrico de la producción científica de México en ciencias agrícolas durante el período 1983-2002. *Revista Fitotecnia Mexicana*, 31(3), 187-194. Recuperado de: <https://www.revistafitotecniamexicana.org/documentos/31-3/1a.pdf>
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J. & Zanasi, A. (1998). *Discovering Data Mining From concept to implementation*. New Jersey, Estados Unidos: Prentice Hall.
- Cáceres, L., de la Peña, J. A., Pineda, A. R., Di Prisco, C. & Solotar, A. (2014). Mathematics in Latin America and the Caribbean: So Much Happening, So Much to Do. *Notices of the American Mathematical Society*, 61(9), 1052-1055. Recuperado de: <https://www.ams.org/notices/201409/rnotip1052.pdf>
- Cañedo-Andalia, R. (1999). Los análisis de citas en la evaluación de los trabajos científicos y las publicaciones seriadas. *ACIMED*, 7(1), 30-39. Recuperado de: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94351999000100004&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94351999000100004&lng=es&tlng=es)
- Card, S. K., MacKinlay, J. & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. San Francisco, Estados Unidos: Morgan Kaufmann.
- Case, D. O. & Higgins, G. M. (2000). How can we investigate citation behavior? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science*, 51(7), 635-645. [https://doi.org/10.1002/\(SICI\)1097-4571\(2000\)51:7<635::AID-ASI6>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1097-4571(2000)51:7<635::AID-ASI6>3.0.CO;2-H)
- Castanha, R. C. G. & Gracio, M. C. C. (2013). Contribution of brazilian scientific production to mainstream science in the field of mathematics: a scientometrics analysis (2002-2011). En 14th International Society of Scientometrics and Informetrics Conference (ISSI), Vienna: Austria.
- Castillo-Pérez, J. J., Muñoz-Valera, L., García-Gómez, F. & Mejía-Aranguréc, J. M. (2015). Análisis bibliométrico de la producción científica sobre la influenza en México, 2000-2012. *Revista Médica del Instituto Mexicano del Seguro Social*, 53(3), 294-301. Recuperado de: [http://revistamedica.imss.gob.mx/editorial/index.php/revista\\_medica/article/view/33](http://revistamedica.imss.gob.mx/editorial/index.php/revista_medica/article/view/33)
- Catarci, T. & Cruz, I. F. (1996). Information visualization. *SIGMOD*, 25(4): 14-15. <https://doi.org/10.1145/245882.245888>
- Chen, C. (1999). Visualising semantic spaces and author cocitation networks in digital libraries. *Information Processing & Management*, 35(3), 401-420. [https://doi.org/10.1016/S0306-4573\(98\)00068-5](https://doi.org/10.1016/S0306-4573(98)00068-5)
- Chen, C. (2002). Information visualization. *Information Visualization*, 1(1), 1-4. <https://doi.org/10.1057/palgrave/ivs/9500009>

- Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4), 12-16. Recuperado de: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1463074>
- Chen, C. (2017). Science Mapping: A Systematic Review of the Literature. *Journal of Data and Information Science*, 2(2), 1-40. <https://doi.org/10.1515/jdis-2017-0006>
- Chinchilla-Rodriguez, Z., Zacca-Gonzalez, G., Vargas-Quesada, B. & Moya-Aregon, F. (2015). Latin American scientific output in Public Health: combined analysis using bibliometric, socioeconomic and health indicators. *Scientometrics*, 102(1), 609-628. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-014-1349-9>
- Clarivate Analytics (2018). *Web of Science: Core Collection*. Recuperado de: <https://clarivate.com/products/web-of-science/>
- Cole, S. (2000). The role of journals in the growth of scientific knowledge. En Cronin, B. & Atkins, H. B. (Eds.), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Medford: ASIS.
- Collazo-Reyes, F. (2014). Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. *Scientometrics*, 98(1), 197-209. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-013-1036-2>
- Collazo-Reyes, F., Flores-Vargas, X., Munoz-Garcia, M. L. & Perez-Angon, M. A. (2014). Las prácticas de citación como interpretantes semióticos de acreditación de saberes locales en astronomía: México 1952-1972. *Transinformação*, 26(3), 269-279. <http://dx.doi.org/10.1590/0103-3786201400030004>
- Collazo-Reyes, F., Luna-Morales, M. E. & Luna-Morales, E. (2017). Change in the publishing regime in Latin America: from a local to universal journal, *Archivos de investigacion Medica/Archives of Medical Research (1970-2014)*. *Scientometrics*, 110(2), 695-709. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-016-2207-8>
- Collazo-Reyes, F., Luna-Morales, M. E. & Russell, J. M. (2004). Publication and citation patterns of the Mexican contribution to a "Big Science" discipline: Elementary particle physics. *Scientometrics*, 60(2), 131-143. <https://doi-org.pbidi.unam.mx:2443/10.1023/B:SCIE.0000027676.41604>
- Collazo-Reyes, F., Luna-Morales, M. E., Russell, J. M. & Perez-Angon, M. A. (2008). Publication and citation patterns of Latin American & Caribbean journals in the SCI and SSCI from 1995 to 2004. *Scientometrics*, 75(1), 145-161. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-007-1841>
- Collazo-Reyes, F., Luna-Morales, M. E., Russell, J. M. & Perez-Angon, M. A. (2010). Enriching knowledge production patterns of Mexican physics in particles and fields. *Scientometrics*, 85(3), 791-802. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-010-0229-1>
- Collazo-Reyes, F., Luna-Morales, M. E., Russell, J. M. & Perez-Angon, M. A. (2011). Emergence and convergence of scientific communication in a developing country: Mexico 1900-1979. En 13th Conference of the International Society for Scientometrics and Informetrics (ISSI), Durban: South Africa.
- CINVESTAV. (2017). *Atlas de la Ciencia Mexicana*. Ciudad de México. Recuperado de: <http://bibliometria.bfm.cinvestav.mx/Serviciosinf.html>

- Cosijn, E. & Ingwersen, P. (2000). Dimensions of relevance. *Information Processing & Management*, 36(4), 533-550. [https://doi.org/10.1016/S0306-4573\(99\)00072-2](https://doi.org/10.1016/S0306-4573(99)00072-2)
- Dang, Y. & Zhang, W. L. (2003). Internationalization of mathematical research. *Scientometrics*, 58(3), 559-570. <https://doi-org.pbidi.unam.mx:2443/10.1023/B:SCIE.0000006880.38873.dd>
- Debackere, K. & Glanzel, W. (2004). Using a bibliometric approach to support research policy making: The case of the Flemish BOF-key. *Scientometrics*, 59(2), 253-276. <https://doi-org.pbidi.unam.mx:2443/10.1023/B:SCIE.0000018532.70146.02>
- de Filippo, D., Morillo, F. & Fernández M. F. (2008). Indicadores de colaboración científica del CSIC con Latinoamérica en bases de datos internacionales. *Revista Española de Documentación Científica*. 31(1). 66-84. Recuperado de: <http://redc.revistas.csic.es/index.php/redc/article/view/413/425>
- Delgado, H. & Russell, J. M. (1992). Impact of studies published in the international literature by scientists at the National University of Mexico. *Scientometrics*, 23(1), 75-90. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02020915>
- Dorta González, M. I. & Dorta González, P. (2016). ¿Se ajustan las ventanas fijas de citación a las velocidades de maduración del impacto de las revistas científicas? *Investigación Bibliotecológica*, 30(68), 73-89. <https://doi-org/10.1016/j.ibbai.2016.02.004>
- Dravec Braun, J. (2012). Effects of war on scientific production: mathematics in Croatia from 1968 to 2008. *Scientometrics*, 93(3), 931-936. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-012-0735-4>
- Dunkel, B., Soparkar, N., Szaro, J. & Uthurusamy, R. (1997). System for KDD: From Concepts to Practice. *Future Generation Computer Systems*, 13(2-3): 231-242. [https://doi.org/10.1016/S0167-739X\(97\)00023-X](https://doi.org/10.1016/S0167-739X(97)00023-X)
- Fayyad, U. (1997). Data mining and knowledge discovery in databases: Implications for scientific databases. En 9th International Conference on Scientific and Statistical Database Management, Olympia: Estados Unidos.
- Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3): 37-54. Recuperado de: <https://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1230/1131>
- Ferrer-Sapena, A., Sánchez-Pérez, E. A., Peset, F., González, L. M. & Aleixandre-Benavent, R. (2016). The Impact Factor as a measuring tool of the prestige of the journals in research assessment in mathematics. *Research Evaluation*, 25(3), 306-314. <https://doi.org/10.1093/reseval/rvv041>
- Freeman, J. A. & Skapura, D. M. (1993). *Redes Neuronales. Algoritmos, aplicaciones y técnicas de propagación*. México: Addison-Wesley.
- García-Silberman, S., Arana, D., Martínez, R., Infante, R. & Jiménez, A. (2004). Research of epidemiological and psychosocial aspects of mental health: A bibliometric analysis. *Salud Mental*, 27(5), 8-22. Recuperado de: <http://www.redalyc.org/articulo.oa?id=58252702>
- Garfield, E. (1979). *Citation indexing: Its theory and applications in science, technology and humanities*. New York, Estados Unidos: Wiley.

- Garreton, M., Murmis, M., de Sierra, G. & Trindade, H. (2005). Social sciences in Latin America: A comparative perspective - Argentina, Brazil, Chile, Mexico and Uruguay. *Social Science Information Sur Les Sciences Sociales*, 44(2-3), 557–593. <https://doi.org/10.1177/0539018405053297>
- Geisler, E. (2005). The measurement of scientific activity: Research directions in linking philosophy of science and metrics of science and technology outputs. *Scientometrics*, 62(2), 269-284. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-005-0020-x>
- Gershon, N., Eick, S. G. & Card. S. (1998). Information Visualization. *ACM Interactions*, 5(2), 9-15. <https://doi-org/10.1145/274430.274432>
- Glänzel, W. & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics*, 56(3): 357–367. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:1022378804087>
- González, R. R. (1996). Neural Networks - A Systematic Introduction. Springer-Verlag, New York: Estados Unidos. Recuperado de: <https://page.mi.fu-berlin.de/rojas/neural/>
- González, E., Arenas, M. & Licea de Arenas, J. (2003). Estudio bibliométrico de la actividad científica de los matemáticos mexicanos graduados en Estados Unidos en el período 1980-1998. *Anales de Documentación*, 6, 89-108. Recuperado de: <http://revistas.um.es/analesdoc/article/view/2021>
- Gorbea-Portal, S. (1998). Modelación matemática de la actividad bibliotecaria: una revisión. *Investigación Bibliotecológica*, 12(24), 6-23. <http://dx.doi.org/10.22201/iibi.0187358xp.1998.24.3869>
- 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. Recuperado en: <http://www.periodicos.ufpb.br/ojs/index.php/pgc/article/view/14175/9321>
- Gorostiza, L. G. (1991). Las perspectivas de las matemáticas. *Avance y perspectiva*, 10, 31-40.
- Guevara-Villanueva, A. (2007). La comunidad de matemáticos en México: su comportamiento en la búsqueda de información. *Anales de Documentación*, 10, 163-184. Recuperado de: <http://revistas.um.es/analesdoc/article/view/1131/1181>
- Guzmán, M. V. (2009). *ViBlioSOM: Metodología para la Visualización de Información métrica con Mapas Auto-organizados*. Tesis Doctoral. Universidad de La Habana, Cuba.
- Guzmán, M. V., Carrillo, H., Villaseñor, E., Valencia, E., Calero, R., Morán, L. E. & Acosta, A. (2004). *Minería de Datos con Redes Neuronales Artificiales: Aplicación en Vacunas - Tuberculosis*. Congreso Internacional de la Información, INFO'2004, La Habana, Cuba.
- Guzmán-Sánchez, M. V. & Sotolongo-Aguilar, G. (2002). Mapas tecnológicos para la estrategia empresarial. Situación tecnológica de la neisseria meningitidis. *ACIMED*, 10(4): 1-2. Recuperado en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94352002000400001](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352002000400001)
- Guzmán-Sánchez, M. V., Carrillo-Calvet, H., Jiménez-Andrade, J. L. & Villaseñor-García, E. A. (2010). *Bioinformetric studies on TB vaccines*

- research*. En: The art and science of Tuberculosis vaccine development. Penerbit Universiti Sains Malaysia, Malaysia: Oxford University Press.
- Han, J. & Kamber, M. (2006). Data mining: concepts and techniques. Massachusetts, Estados Unidos: Morgan Kaufmann Publishers. Recuperado de: [http://ccs1.hnue.edu.vn/hungtd/DM2012/DataMining\\_BOOK.pdf](http://ccs1.hnue.edu.vn/hungtd/DM2012/DataMining_BOOK.pdf)
- Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science*, 43(9), 602-615. [https://doi.org/10.1002/\(SICI\)1097-4571\(199210\)43:9<602::AID-ASI3>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-4571(199210)43:9<602::AID-ASI3>3.0.CO;2-Q)
- Hartigan, J. (1975). *Clustering algorithms*. New York, Estados Unidos: John Wiley & Sons. Recuperado de: [http://people.inf.elte.hu/fekete/algorithmusok\\_msc/klaszterezes/John%20A.%20Hartigan-Clusterig%20Algorithms-John%20Wiley%20&%20Sons%20\(1975\).pdf](http://people.inf.elte.hu/fekete/algorithmusok_msc/klaszterezes/John%20A.%20Hartigan-Clusterig%20Algorithms-John%20Wiley%20&%20Sons%20(1975).pdf)
- Hätönen, K., Klemettinen, M., Mannila, H., Ronkainen, P. & Toivonen, H. (1996). *Knowledge discovery from telecommunication network alarm databases*. En 12th International Conference on Data Engineering (ICDE '96). New Orleans, Estados Unidos. Recuperado de: [http://arbor.ee.ntu.edu.tw/~chyun/dmpa\\_per/hatokd96.pdf](http://arbor.ee.ntu.edu.tw/~chyun/dmpa_per/hatokd96.pdf)
- Hernández, Y. G., Dray, M. K. & Russell, J. M. (2013). Enfoques metodológicos para identificar y caracterizar la investigación mexicana en química en bases de datos bibliográficas. *Investigación Bibliotecológica*, 27(59), 35-66. [http://dx.doi.org/10.1016/S0187-358X\(13\)72530-9](http://dx.doi.org/10.1016/S0187-358X(13)72530-9)
- Hernandez-Garcia, Y., Chamizo, J. A., Kleiche-Dray, M. & Russell, J. M. (2016). The scientific impact of mexican steroid research 1935-1965: A bibliometric and historiographic analysis. *Journal of the Association for Information Science and Technology*, 67(5), 1245-1256. <https://doi.org/10.1002/asi.23493>
- Herrera-Vallejera, D., Sánchez-Perdomo, R., Rosario-Sierra, M., & Rodríguez-Sánchez, Y. (2017). Estudio cuantitativo de la actividad científica de Cuba en las Ciencias Naturales e Ingeniería y Matemática-Ciencias de la Computación. *Investigación Bibliotecológica*, 31(72), 113-137. doi:<http://dx.doi.org/10.22201/iibi.0187358xp.2017.72.57826>
- Honkela, T., Kaski, S., Lagus, K., & Kohonen, T. (1996). *Newsgroup exploration with WEBSOM method and browsing interface*. Espoo. Helsinki, University of Technology, Laboratories of Computer and Information Science. Recuperado de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.47.37&rep=rep1&type=pdf>
- Huber, J. C. & Wagner-Döbler, R. (2001). Scientific production: A statistical analysis of authors in mathematical logic. *Scientometrics*, 50(2), 323-337. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:1010581925357>
- Jain, A. K., Mao, J. & Mohiuddin, K. (1996). Artificial Neural Networks: A Tutorial. *IEEE Computer Special Issue on Neural Computing*. Recuperado de: <https://pdfs.semanticscholar.org/f6d5/92af572ebf890a06aa228049055923fb.pdf>
- Kamada, T. & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1), 7-15. [https://doi.org/10.1016/0020-0190\(89\)90102-6](https://doi.org/10.1016/0020-0190(89)90102-6)

- Kaplan, N. (1965). The norms of citation behavior: Prolegomena to the footnote. *American Documentation*, 16(3), 179-184. <https://doi.org/10.1002/asi.5090160305>
- Kaski, S. (1997). *Data Exploration Using Self-Organizing Maps*. Tesis de Doctorado. Helsinki University of Technology: Finlandia.
- Keim, D. A. & Kriegel, H. P. (1994). *Using visualization to support data mining of large existing databases*. En Database Issues for Data Visualization. Serie Lecture Notes in Computer Science, 871: 210-229.
- Kapseon, K. (2004). The motivation for citing specific references by social scientists in Korea: The phenomenon of co-existing references. *Scientometrics*, 59(1), 79-93. <https://doi-org.pbidi.unam.mx:2443/10.1023/B:SCIE.0000013300.68224.c6>
- Kohonen, T. (1993). Physiological interpretation of the self-organizing map algorithm. *Neural Networks*, 6(7), 895-905. [https://doi.org/10.1016/S0893-6080\(09\)80001-4](https://doi.org/10.1016/S0893-6080(09)80001-4)
- Kohonen, T. (1997). *Exploration of very large databases by self-organizing maps*. Proceedings of International Conference on Neural Networks. IEEE Service Center, Houston: Estados Unidos.
- Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, 37(1), 52-65. <https://doi.org/10.1016/j.neunet.2012.09.018>
- Koljatic, M. & Silva, M. (2001). The international publication productivity of Latin American countries in the economics and business administration fields. *Scientometrics*, 51(2), 381-394. <https://doi.org/10.1023/A:101275360179>
- Korevaar, J. C. & Moed, H. F. (1996). Validation of bibliometric indicators in the field of mathematics. *Scientometrics*, 37(1), 117-130. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02093488>
- Kornilov, A. R. (1997). Intelligent technologies new opportunities for modern industry. *Information Technology*, 3(6): 1-14.
- Kostoff, R. N. (1995). Federal research impact assessment: axioms, approaches, applications. *Scientometrics*, 34(2), 163-206. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02020420>
- Kostoff, R. N. (2001). The metrics of science and technology. *Scientometrics*, 50(2), 353-361. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:101059011245>
- Kostoff, R. N., del Río, J. A., Cortés, H. D., Smith, C., Smith, A., Wagner, C., Leydesdorff, L., Karypis, G., Malpohl, G. & Tshiteya, R. (2005). The structure and infrastructure of Mexico's science and technology. *Technological Forecasting and Social Change*, 72(7), 798-814. <https://doi.org/10.1016/j.techfore.2005.02.001>
- Kostoff, R. N., del Río, J. A., Cortés, H. D., Smith, C., Smith, A., Wagner, C., Leydesdorff, L., Karypis, G., Malpohl, G. & Tshiteya, R. (2007). Clustering methodologies for identifying country core competencies. *Journal of Information Science*, 33(1), 21-40. <https://doi.org/10.1177/0165551506067124>
- Kruskal, J. B. & Wish, M. (1978). *Multidimensional scaling* (quantitative applications in the Social Sciences). California, Estados Unidos: Sage Publications. <http://dx.doi.org/10.4135/9781412985130>
- Laclette, J. P. & Zúñiga-Bello, P. (2011). *Ranking de la producción científica Mexicana*. Foro Consultivo Científico y Tecnológico. Ciudad de México,

- México. Recuperado de: [http://www.foroconsultivo.org.mx/libros\\_editados/ranking\\_por\\_institucion\\_2011.pdf](http://www.foroconsultivo.org.mx/libros_editados/ranking_por_institucion_2011.pdf)
- Latour, B. & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. New Jersey, Estados Unidos: Princeton University Press.
- Lena, M. L. (1997). Scientific productivity in environmental psychology in Mexico- A bibliometric analysis. *Environment and Behavior*, 29(2), 169-197. <https://doi.org/10.1177/001391659702900202>
- Leydesdorff, L. & Bornmann, L. (2016). The Operationalization of “Fields” as WoS Subject Categories (WCs) in Evaluative Bibliometrics: The Cases of “Library and Information Science” and “Science & Technology Studies” *Journal of the Association for Information Science and Technology*, 67(3): 707-714. <https://doi.org/10.1002/asi.23408>
- Licea de Arenas, J., Castaños-Lomnitz, H. & Arenas Licea, J. (2002). Significant Mexican research in the health sciences: A bibliometric analysis. *Scientometrics*, 53(1), 39-48. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:1014879803333>
- Licea de Arenas, J., Sandoval, M. & Arenas, M. (2003). La investigación agrícola en México. Un estudio bibliométrico con enfoque de género. *Anales de Documentación*, 6, 145-154. <http://www.redalyc.org/articulo.oa?id=63500610>
- Licea de Arenas, J., Valles, J. & Arenas, M. (2000). Educational research in Mexico: socio-demographic and visibility issues. *Educational Research*, 42(1), 85-90.
- Lima, M., Liberman, S. & Russell, J. M. (2005). Scientific group cohesiveness at the National University of Mexico. *Scientometrics*, 64(1), 55-66. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-005-0237-8>
- Lomnitz, L. A. & Cházaro, L. (1999). Basic, Applied and Technological Research: Computer Science and Applied Mathematics at the National Autonomous university of Mexico. *Social Studies of Science*, 29(1), 113-134.
- Longar, M. P. & Ríos E. (2014). Análisis Cienciométrico sobre el estado del arte de la Hidroponía. Caso México. *Mundo Siglo XXI*, 10(34), 63-74. Recuperado de: <http://www.mundosigloxxi.ipn.mx/pdf/v10/34/05.pdf>
- López-Olmedo, R., Marmolejo-Leyva, R., Pérez-Angón, M. A., Villa-Vázquez, L. L. & Zayago-Lau, E. (2017). The role of public policies in the decentralization process of Mexican science and the formation of new researchers in institutions outside the Mexico City area. *Scientometrics*, 112(3), 1343-1366. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-017-2423-x>
- Lozano, I. A. (2016). *Análisis de la producción científica de México en el Web of Science, durante el período 2005-2015, utilizando inteligencia computacional*. Tesis de Maestría. Universidad Nacional Autónoma de México, México.
- Luna-Morales, E., Russell, J. M. & Mireles-Cárdenas, C. (2012). Evolución e impacto de la investigación en la Universidad Autónoma de San Luis Potosí, México. Patrones de publicación y Sistema Nacional de Investigadores. *Investigación Bibliotecológica*, 26(58), 175-213. <http://dx.doi.org/10.22201/iibi.0187358xp.2012.58.35985>
- Luna-Morales, M. E. (2012). La colaboración científica y la internacionalización de la ciencia mexicana de 1980 a 2004. *Investigación Bibliotecológica*,

- 26(57), 103-129. <http://dx.doi.org/10.22201/iibi.0187358xp.2012.57.33841>
- Luna-Morales, M. E., Collazo-Reyes, F., Russell, J. M. & Pérez-Angón, M. A. (2009). Early Patterns of Scientific Production by Mexican Researchers in Mainstream Journals, 1900-1950. *Journal of the American Society for Information Science and Technology*, 60(7), 1337-1348. <https://doi.org/10.1002/asi.21065>
- Macías-Chapula, C. A. (2002). Bibliometric and webometric analysis of health system reforms in Latin America and the Caribbean. *Scientometrics*, 53(3), 407-427. <https://doi.org/10.1023/A:1014829214237>
- Macías-Chapula, C. A. (2013). Comparative analysis of health public policy research results among Mexico, Chile and Argentina. *Scientometrics*, 95(2), 615-628. <https://doi.org/10.1007/s11192-012-0855-x>
- Macías-Chapula, C. A., Mendoza-Guerrero, J. A., Rodea-Castro, I. P. & Gutierrez-Carrasco, A. (2007). *Institutional health research collaboration in Mexico: A bibliometric study*. International Society of Scientometrics & Informetrics. Leuven, Bélgica.
- Maimon, O. & Rokach, L. (2010). *Data Mining and Knowledge Discovery Handbook*. Massachusetts, Estados Unidos: Springer. <https://doi.org/10.1007/b107408>
- Maltrás-Barba, B. (2003). *Los indicadores bibliométricos: fundamentos y aplicación al análisis de la ciencia*. Asturias, España: TREA.
- Mannila, H. (1996). *Data Mining: Machine Learning, Statistics and Databases*. SSDBM '96, En Eighth International Conference on Scientific and Statistical Database Management, 2-9.
- Mariscal, O. (2002). *La física mexicana 1990-1999: Indicadores bibliométricos de producción científica documental y recursos humanos*. Tesis de Licenciatura en Biblioteconomía. Escuela Nacional de Biblioteconomía y Archivonomía, México.
- Martí Lahera, Y. (2011). *Ciencia y Género en Cuba (Web of Science, 2001-2007)*. Universidad de Granada, España.
- Martínez Arellano, F. F. & Ruíz Vaca, J. O. (2011). Visibilidad de la producción científica de la UNAM en Humanidades y Ciencias Sociales a través de Humanindex. Coloquio de Acceso Abierto a La Información en: *Las Bibliotecas Académicas de América Latina y el Caribe*. Universidad Nacional Autónoma de México, 137-147. Recuperado de: <http://www2.uadec.mx/pub/pdf/informacion.pdf>
- Matheus, C. J., Chan, P. K. & Piatetsky-Shapiro, G. (1993). Systems for Knowledge Discovery in Databases. IEEE Learning & Discovery in Knowledge Based Databases. Recuperado de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.49.169&rep=rep1&type=pdf>
- May, K. O. (1966). Quantitative growth of the mathematical literature. *Science*, 154(3757), 1672-1673. <https://doi.org/10.1126/science.154.3757.1672>
- Michalski, R. S., Brakto, I. & Kubat, M. (1998). *Machine Learning and Data Mining. Methods and Applications*. New York, Estados Unidos: John Wiley & Sons.
- Michán, L., Russell, J. M., Pereyra, A. S., Cruset, A. L. & Beltrán, C. L. (2008). Análisis de la sistemática actual en latinoamérica. *Interciencia*, 33(10), 754-761. Recuperado de: <http://www.redalyc.org/html/339/33931008/index.html>

- Michán, L. (2010). *Análisis bibliométrico de la Botánica Mexicana*. En XVIII Congreso Mexicano de Botánica. Guadalajara, México.
- Michán, L. & Muñoz-Velasco, I. (2013). Cienciometría para ciencias médicas: definiciones, aplicaciones y perspectivas. *Investigación en Educación Médica*, 2(6), 100-106. [https://doi.org/10.1016/S2007-5057\(13\)72694-2](https://doi.org/10.1016/S2007-5057(13)72694-2)
- Mingers, J. & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1), 1–19. <https://doi.org/10.1016/j.ejor.2015.04.002>
- Mitchell, T. (1997). *Machine Learning*. New York, Estados Unidos: McGraw Hill.
- Moravcsik, M. J. & Murugesan, P. (1975). Some results on the function and quality of citation. *Social Studies of Science*, 5(1), 86-92. <https://doi.org/10.1177/030631277500500106>
- Moya-Anegón, F., Herrero-Solana, V. & Jiménez-Contreras, E. (2006). A connectionist and multivariate approach to science maps: The SOM, clustering and MDS applied to library and information science research. *Journal of Information Science*, 32(1), 63–77. <https://doi.org/10.1177/0165551506059226>
- Murphy, J. (2013). International trends in health science librarianship. Part 5 Latin America and the Caribbean. *Health Information and Libraries Journal*, 30(1), 76-82. <https://doi.org/10.1111/hir.12016>
- Nalimov, V. V. & Mulchenko, Z. M. (1969). *Naukometrija: izuchenie razvitija nauki kak nformatsionnogo processa*. Moscow, Russia: Nauka.
- Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81-100. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-006-0007>
- Nicolaisen, J. (2007). Citation analysis. *Annual Review of Information Science and Technology*, 41(1), 609-641. <https://doi.org/10.1002/aris.2007.1440410120>
- OECD (2016). *Main Science and Technology Indicators*, 2016(1): OECD Publishing. Recuperado de: [http://www.oecd.org/sti/inno/MSTI\\_documentation\\_e.pdf](http://www.oecd.org/sti/inno/MSTI_documentation_e.pdf)
- ONU (2018). Estados Miembros de la Organización de las Naciones Unidas. Recuperado de: <http://www.un.org/es/member-states/>
- Otlet, P. (1934). *Traité de documentation*. Le livre sur le livre. Théorie et pratique. Bruxelles, Bélgica: Editions Mundaneum.
- Pinto, M., Isabel Escalona, M., Pulgarin, A. & Uribe-Tirado, A. (2015). The scientific production of Ibero-American authors on information literacy (1985-2013). *Scientometrics*, 102(2), 1555-1576. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-014-1498-x>
- PLoS medicine Editors (2006). Ich Weiss nicht was sol les bedeuten. Language matters in medicine. *PLoS Med*, 3(2), e122. <https://doi.org/10.1371/journal.pmed.0030122>
- Polanco, X., François 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. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:1010537316758>
- Price, D. J. S. (1963). *Little Science, Big Science*. New York, Estados Unidos: Yale University Press. Recuperado de: [http://www.andreasaltelli.eu/file/repository/Little\\_science\\_big\\_science\\_and\\_beyond.pdf](http://www.andreasaltelli.eu/file/repository/Little_science_big_science_and_beyond.pdf)

- Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25(4), 344-349. Recuperado de: <http://dx.doi.org.pbi.unam.mx:8080/10.1108/eb026482>
- Restrepo-Arango, C. & Urbizagastegui-Alvarado, R. (2016). Acercamiento a los Estudios Bibliométricos, Cienciométricos e Informétricos en México. *Informação & Sociedade: Estudos*, 26(1), 51-71. Recuperado de: <http://www.periodicos.ufpb.br/ojs2/index.php/ies/article/view/22799/15528>
- Restrepo, L. C. (2015). *Modelo de evaluación del desempeño académico de los investigadores en las ciencias sociales en México*. Tesis Doctoral. Universidad Nacional Autónoma de México, México.
- RICYT (2018). Red Iberoamericana de Indicadores de Ciencia y Tecnología: Publicaciones en SCI por habitante. Recuperado de: <http://dev.riicyt.org/ui/v3/comparative.html?indicator=SCIxH>
- Rinia, E. J., Van Leeuwen, T. N., Van Vuren, H. G. & Van Raan, A. F. J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria: Evaluation of condensed matter physics in the Netherlands. *Research Policy*, 27(1), 95-107. [https://doi.org/10.1016/S0048-7333\(98\)00026-2](https://doi.org/10.1016/S0048-7333(98)00026-2)
- Robertson, G. G., Card, S. K. & MacKinlay, J. D. (1993). Information visualization using 3D interactive animation. *Communications of the ACM*, 36(4): 57-71. <https://doi.org/10.1145/255950.153577>
- Rodríguez-Navarro, A. (2011). Measuring research excellence: Number of Nobel Prize achievements versus conventional bibliometric indicators. *Journal of Documentation*, 67(4), 582-600. <https://doi.org/10.1108/002204111111145007>
- Rojas-Sola, J. I. & Jorda-Albiñana, B. (2011). Análisis bibliométrico de la producción científica mexicana sobre ingeniería hidráulica en revistas de la base de datos Science Citation Index-Expanded (1997-2008). *Tecnología y Ciencias del Agua*, 2(4), 195-213. <https://doi.org/10.24850/j-tyca-imta>
- 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. *Annalen Der Physik*, 18(4), 198-205. <https://doi.org/10.1002/andp.200910346>
- Russell, J. M. (2004). Obtención de indicadores bibliométricos a partir de la utilización de las herramientas tradicionales de información. *Congreso internacional de la información INFO 2004*. La Habana, Cuba.
- Russell, J. M., Hernández-García, Y. & Kleiche-Dray, M. (2016). Collaboration dynamics of Mexican research in Chemistry and its relationship with communication patterns. *Scientometrics*, 109(1), 283-316. <https://doi-org.pbi.unam.mx:2443/10.1007/s11192-016-2069-0>
- Russell, J. M., Ainsworth, S. & Díaz-Aguilar, J. (2012). Web visibility or wasted opportunity? Case studies from Mexican research institutes. *Aslib Proceedings*, 64(1), 67-82. <https://doi-org/10.1108/00012531211196710>.
- Russell, J. M., Ainsworth, S., Del Río, J. A., Narváez-Berthelemot, N. & Cortés H. D. (2007). Colaboración científica entre países de la región Latinoamericana. *Revista Española De Documentación Científica*, 30(2), 180-198. <https://doi.org/10.3989/redc.2007.v30.i2>

- Sancho, R. (1992). Manual de Frascati para la medición de las actividades científicas y técnicas. *Política Científica*, 45(3), 21-26.
- Sancho, R. (2001). Medición de las actividades de Ciencia y Tecnología, estadísticas e indicadores empleados. *Revista Española de Documentación Científica*, 24(4), 382-404. <https://doi.org/10.3989/redc.2001.v24.i4>
- Sang, J. L. & Siau, K. (2001). *A Review of data mining techniques*. MCB University Press, 41-46.
- Schadl, S. M. & Todeschini, M. (2015). Cite Globally, Analyze Locally: Citation Analysis from a Local Latin American Studies Perspective. *College & Research Libraries*, 76(2), 136–149. <https://doi.org/10.5860/crl.76.2.136>
- Schvaneveldt, R. (1990). *Pathfinder associative networks: studies in knowledge organization*. New Jersey, Estados Unidos: Ablex Publishing Corp.
- Sierra-Flores, M. M. & Russell, J. M. (2009). Los grupos de investigación más productivos de la Universidad Nacional Autónoma de México (UNAM) en el área de Física: 1990-1999. *Investigación Bibliotecológica*, 23(48), 127-155. <http://dx.doi.org/10.22201/iibi.0187358xp.2009.48>
- Sierra-Flores, M. M., Guzmán, M. V., Raga, A. C. & Pérez, I. (2009). The productivity of Mexican astronomers in the field of outflows from young stars. *Scientometrics*, 81(3), 765-777. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-008-2264-8>
- Skupin, A., Biberstine, J. R. & Börner, K. (2013). Visualizing the Topical Structure of the Medical Sciences: A Self-Organizing Map Approach. *PLoS ONE*, 8(3), e58779. <https://doi.org/10.1371/journal.pone.0058779>
- Small, H. G. (1978). Cited documents as concept symbols. *Social Studies of Science*, 8(3), 327-340. <https://doi.org/10.1177/030631277800800305>
- Smolinsky, L. & Lercher, A. (2012). Citation rates in mathematics: a study of variation by subdiscipline. *Scientometrics*, 91(3), 911-924. <https://doi.org/10.1007/s11192-012-0647-3>
- Snizek, W. E. (1995). Some Observations on the Use of Bibliometric Indicators in the Assignment of University Chairs. *Scientometrics*, 32(2), 117-120. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02016888>
- So, C. Y. K. (1998). Citation ranking versus expert judgment in evaluating communication scholars: Effects of research specialty size and individual prominence. *Scientometrics*, 41(3), 325-333. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02459049>
- Solano, J. & Orihuela, A. (2010). Análisis bibliométrico de las publicaciones de investigadores mexicanos sobre ovino de pelo en revistas científicas nacionales y extranjeras. *Universidad y Ciencia*, 26(1), 93-105. Recuperado de: <http://www.scielo.org.mx/pdf/uc/v26n1/v26n1a7.pdf>
- Sotolongo-Aguilar, G. & Guzmán-Sánchez, M. V. (2001). Aplicaciones de las redes neuronales. El caso de la bibliometría. *Ciencias de la Información*, 32(1), 27-34. Recuperado de: <http://cinfo.idict.cu/index.php/cinfo/article/view/252>
- Sotolongo-Aguilar, G., Guzmán-Sánchez, M. V. & Carrillo-Calvet, H. (2002). ViBlioSOM: Visualización de información bibliométrica mediante el mapeo auto-organizado. *Revista Española de Documentación Científica*, 25(4), 477-484. <https://doi.org/10.3989/redc.2002.v25.i4>

- Sotolongo-Aguilar, G., Guzmán-Sánchez, M. V., Saavedra-Fernández, O. & Carrillo-Calvet, H. (2001). *Mining informetric data with self-organizing maps*. En 8th International Society for Scientometrics and Informetrics: Sydney, Australia: BIRG, 665-673.
- Spence, R. (2000). *Information Visualization*. New York, Estados Unidos: Addison-Wesley.
- Spinak, E. (2001). Indicadores cientimétricos. *ACIMED*, 9(4), 16-18. Recuperado de: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94352001000400007](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352001000400007)
- Suárez-Balseiro C. & Maura-Sardó M. (2005). *Bibliometría evaluativa. Seguimiento de la actividad científica y tecnológica* (SACT). En Foro sobre Estudios Métricos de la Información. Escuela de Graduados en Ciencias y Tecnologías de la Información.
- Sánchez, A., Carrillo, O. & Garrido, P. (2015). Análisis bibliométrico de la Revista Mexicana de Sociología basado en indicadores de citación. *Revista Mexicana de Sociología*, 77(1), 45-70 <http://dx.doi.org/10.22201/iis.01882503p.2015.0.51757>
- Sánchez-Perdomo, R., Rosario-Sierra, M., Herrera-Vallejera, D., Rodríguez-Sánchez, Y. & Carrillo-Calvet, H. (2017). Revisión bibliométrica de las Ciencias de la Información en América Latina y el Caribe. *Investigación Bibliotecológica*, 31(nesp), 79-100. <http://dx.doi.org/10.22201/iibi.24488321xe.2017.nesp1>
- Tarango, J., Hernandez-Gutierrez, P. Z. & Vazquez-Guzman, D. (2015). Evaluation of scientific production in mexican state public universities (2007-2011) using principal component analysis. *El Profesional de la Información*, 24(5), 567-576. <https://doi.org/10.3145/epi.2015.sep.06>
- Tarango, J. & Machin-Mastromatteo, J. D. (2016). Scientific production in Mexican universities: Rates and expectations toward competitiveness. *Information Development*, 32(1), 107-111. <https://doi.org/10.1177/026666691561373>
- Thuraisingham, B. (1999). *Data Mining: Technologies, Techniques, Tools and Trends*. CRC Press.
- Torres, J. A. (2009). *Desarrollo científico de las Ciencias Sociales en México; análisis bibliométrico del período 1997-2006: Social Science Citation Index (SSCI-ISI) y CiteSpace*. Tesis Doctoral, Universidad de Granada, España.
- Torres Ponjuan, D. (2009). Aproximaciones a la visualización como disciplina científica. *ACIMED*, 20 (6), 161-174. Recuperado de: <http://scielo.sld.cu/pdf/aci/v20n6/aci051209.pdf>
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison Wesley.
- Uddin, A., Singh, V. K., Pinto, D. & Olmos, I. (2015). Scientometric mapping of computer science research in Mexico. *Scientometrics*, 105(1), 97-114. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-015-1654-y>
- UMALCA (1995). *Unión Matemática de América Latina y el Caribe*. Congresos y escuelas. Recuperado de: <http://www.umalca.org/>
- UNAM (2017). *Estudio Comparativo de Universidades Mexicanas*. Recuperado de: <http://www.dgei.unam.mx/hwp/ecum/>
- UNESCO (2015). Informe de la UNESCO sobre la ciencia hacia 2030. Recuperado de: <http://unesdoc.unesco.org/images/0023/002354/235407s.pdf>

- Valcárcel Asencios, V. (2004). Data mining y el descubrimiento del conocimiento. *Revista de la Facultad de Ingeniería Industrial*, 7(2), 83-86. <http://dx.doi.org/10.15381/idata.v7i2.6140>
- Van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics*, 36(3), 397-420. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02129602>
- Van Raan, A. F. J. (2001). Competition amongst scientists for publication status: toward a model of scientific publication and citation distributions. *Scientometrics*, 51(1), 347-357. <https://doi-org.pbidi.unam.mx:2443/10.1023/A:1010501820393>
- Van Raan, A. F. J. (2004). Measuring science. Capita selecta of current main issues. En Moed, H. F., Glänzel, W. & 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.
- Vélez-Cuartas, G., Lucio-Arias, D. & Leydesdorff, L. (2016). Regional and global science: Publications from Latin America and the Caribbean in the SciELO Citation Index and the Web of Science. *El profesional de la información*, 25(1), 35-46. <https://doi.org/10.3145/epi.2016.ene.05>
- Vessuri, H. (1995). Recent strategies for adding value to scientific journal in latin America. *Scientometrics*, 34(1), 139-161. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02019178>
- Villaseñor, E. A. (2004). *Análisis inteligente de datos con redes neuronales*. Tesis de licenciatura. Universidad Nacional Autónoma de México, México.
- Villaseñor, E. A. (2016). *Aportaciones al análisis y visualización informétrica con la red neuronal SOM*. Universidad Nacional Autónoma de México, México.
- Villaseñor, 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*, 110(1), 77-104. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-016-2166-0>
- Vinkler, P. (1987). A quasi-quantitative citation model. *Scientometrics*, 12(1-2), 47-72. <https://doi-org.pbidi.unam.mx:2443/10.1007/BF02016689>
- Vinkler, P. (2006). Composite scientometric indicators for evaluating publications of research institutes. *Scientometrics*, 68, 629-642.
- Wang, Q. & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus. *Journal of Informetrics*, 10(2), 347-364. <https://doi-org/10.1016/j.joi.2016.02.003>
- Warner, J. (2000). A critical review of the application of citation studies to the Research Assessment Exercises. *Journal of Information Science*, 26(6), 453-459. <https://doi.org/10.1177/016555150002600607>
- Waltman, L., Calero- Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., 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. <https://doi.org/10.1002/asi.22708>

- White, H. D. (2004). Reward, persuasion, and the Sokal Hoax: A study in citation identities. *Scientometrics*, 60(1), 93-120. <https://doi-org.pbidi.unam.mx:2443/10.1023/B:SCIE.0000027313.91401.9b>
- Zhou, P. & Leydesdorff, L. (2007). The citation impacts and citation environments of Chinese journals in mathematics. *Scientometrics*, 72(2), 185-200. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-007-1713-0>
- Zhou, P. & Tian, H. B. (2014). Funded collaboration research in mathematics in China. *Scientometrics*, 99(3), 695-715. <https://doi-org.pbidi.unam.mx:2443/10.1007/s11192-013-1212-4>
- Zhu, B. & Chen, H. (2005). Information Visualization. *Annual Review of Information Science and Technology*, 39(1), 139-177. <https://doi.org/10.1002/aris.1440390111>

## Anexos

Anexo 1. Ranking de 127 países con un volumen de producción científica en matemática superior a 18 artículos, durante el período 2000-2017, según el *InCites*.

No.	Países	Ndoc	Cr	% CD	Cr/Ndoc	SDI
1	ESTADOS UNIDOS	146,677	1,472,244	77.06	10.04	26.49
2	CHINA	100,023	577,008	66.73	5.77	3.97
3	FRANCIA	50,176	418,908	78.44	8.35	43.05
4	ALEMANIA	41,985	336,441	77.78	8.01	28.1
5	REINO UNIDO	33,537	305,215	77.91	9.10	29.39
6	ITALIA	31,553	244,052	77.78	7.73	29.32
7	RUSIA	27,472	96,397	61.08	3.51	10.47
8	JAPÓN	27,397	151,144	72.08	5.52	11.81
9	ESPAÑA	25,777	182,811	77.22	7.09	31.46
10	CANADÁ	25,723	205,330	76.75	7.98	42.27
11	INDIA	15,180	72,821	65.80	4.80	0.66
12	COREA DEL SUR	14,884	80,062	69.70	5.38	16.34
13	POLONIA	14,576	72,125	72.81	4.95	20.67
14	AUSTRALIA	12,571	110,590	76.59	8.80	31.85
15	BRASIL	12,239	70,955	71.26	5.80	3.32
16	IRÁN	11,924	58,045	63.37	4.87	8.09
17	TURQUÍA	11,062	57,833	66.05	5.23	8.05
18	ISRAEL	10,078	79,976	77.26	7.94	75.25
19	TAIWÁN	9,415	59,527	72.75	6.32	8.27
20	RUMANIA	8,651	45,064	67.23	5.21	23.07
21	SUIZA	7,438	70,440	78.37	9.47	51.13
22	HOLANDA	7,168	64,718	78.67	9.03	24.08
23	BÉLGICA	7,078	62,141	78.69	8.78	36.09
24	AUSTRIA	6,719	54,671	76.71	8.14	42.89
25	REPÚBLICA CHECA	6,531	39,096	74.15	5.99	33.98
26	SUECIA	6,396	53,130	76.52	8.31	37.28
27	ARABIA SAUDITA	6,295	36,588	65.72	5.81	11.17
28	PORTUGAL	5,925	34,436	74.33	5.81	30.66
29	HUNGRÍA	5,765	30,129	71.27	5.23	31.73
30	MÉXICO	5,117	23,189	69.02	4.53	2.38
31	GRECIA	4,942	34,497	74.52	6.98	24.84
32	UCRANIA	4,714	18,018	61.33	3.82	5.61
33	SUDÁFRICA	3,945	21,107	68.49	5.35	4.14
34	FINLANDIA	3,838	31,262	77.33	8.15	39.10
35	CHILE	3,804	26,358	74.95	6.93	11.97
36	NORUEGA	3,464	31,805	77.28	9.18	38.57
37	SINGAPUR	3,129	29,340	77.09	9.38	35.88

38	DINAMARCA	3,093	28,356	77.24	9.17	30.20
39	SERBIA	3,057	22,448	70.66	7.34	29.55
40	EGYPTO	3,030	20,567	69.60	6.79	2.00
41	ARGENTINA	3,029	17,760	73.03	5.86	4.00
42	NUEVA ZELANDIA	2,795	21,254	75.46	7.60	36.20
43	PAKISTÁN	2,761	15,234	66.24	5.52	0.81
44	TÚNEZ	2,677	10,650	61.71	3.98	13.23
45	VIETNAM	2,547	9,137	58.23	3.59	1.46
46	IRLANDA	2,481	20,162	75.70	8.13	31.20
47	ESLOVENIA	2,477	14,223	76.46	5.74	66.20
48	ARGELIA	2,129	9,039	56.69	4.25	3.01
49	MARRUECOS	2,105	8,771	62.23	4.17	3.59
50	BULGARIA	1,919	9,666	70.30	5.04	14.04
51	CROACIA	1,888	8,969	73.20	4.75	24.03
52	ESLOVAQUIA	1,800	10,385	73.00	5.77	18.35
53	TAILANDIA	1,728	7,866	59.95	4.55	1.36
54	MALASIA	1,669	7,384	68.66	4.42	3.12
55	BIELORRUSIA	1,292	3,361	50.15	2.60	7.49
56	COLOMBIA	1,210	3,674	62.23	3.04	1.36
57	LITUANIA	1,192	4,781	70.89	4.01	20.71
58	GEORGIA	1,014	3,681	61.83	3.63	14.42
59	JORDANIA	948	7,919	67.09	8.35	7.10
60	EMIRATOS ÁRABES UNIDOS	793	5,057	67.72	6.38	6.52
61	VENEZUELA	789	4,588	69.58	5.81	1.60
62	AZERBAIYÁN	752	2,306	57.98	3.07	4.31
63	KAZAJISTÁN	637	2,055	59.34	3.23	1.87
64	KUWAIT	571	2,553	65.85	4.47	10.92
65	CHIPRE	569	3,875	78.03	6.81	28.64
66	UZBEKISTÁN	564	1,797	58.33	3.19	1.07
67	ARMENIA	554	1,058	49.64	1.91	10.34
68	ESTONIA	446	2,424	75.11	5.43	17.94
69	URUGUAY	395	2,269	71.90	5.74	6.22
70	NIGERIA	377	1,681	63.13	4.46	0.13
71	YUGOSLAVIA	372	3,868	83.60	10.40	1.15
72	LÍBANO	370	1,131	58.92	3.06	4.28
73	OMÁN	359	1,772	66.30	4.94	6.62
74	IRLANDA DEL NORTE	325	1,927	67.08	5.93	4.09
75	CAMERÚN	269	1,050	64.31	3.90	0.69
76	CATAR	264	1,200	64.77	4.55	8.68
77	MACAO	258	680	49.22	2.64	22.96
78	INDONESIA	254	910	67.72	3.58	0.54
79	CUBA	237	1,018	64.56	4.30	1.13
80	FILIPINAS	229	858	66.38	3.75	0.13
81	LETONIA	192	322	55.73	1.68	5.14
82	SENEGAL	189	470	56.61	2.49	0.78

83	BOSNIA-HERZEGOVINA	188	499	61.17	2.65	2.87
84	ISLANDIA	179	1,291	74.30	7.21	32.66
85	BANGLADÉS	176	534	56.25	3.03	0.06
86	MONTENEGRO	172	612	63.95	3.56	20.14
87	MACEDONIA	170	620	58.24	3.65	4.23
88	IRAQ	163	597	52.15	3.66	0.26
89	BOTSUANA	161	558	62.11	3.47	4.33
90	MOLDAVIA	145	542	58.62	3.74	2.25
91	SERBIA Y MONTENEGRO	136	2,130	89.71	15.66	0.97
92	TAYIKISTÁN	120	114	45.83	0.95	0.92
93	PERÚ	117	516	62.39	4.41	0.20
94	MALTA	91	385	70.33	4.23	12,93
95	COSTA DE MARFIL	85	325	48.24	3.82	0.80
96	ETIOPÍA	81	625	66.67	7.72	0.72
97	SRI LANKA	78	592	61.54	7.59	0.57
98	TURKMENISTÁN	75	751	85.33	10.01	0.53
99	BRUNÉI	71	382	71.83	5.38	7.09
100	BURKINA FASO	69	163	52.17	2.36	0.48
101	ECUADOR	67	233	52.24	3.48	0.23
102	COSTA RICA	66	410	71.21	6.21	2.06
103	MONGOLIA	64	161	57.81	2.52	1.32
104	YEMEN	62	342	62.90	5.52	0.33
105	COREA DEL NORTE	61	177	59.02	2.90	0.39
106	ALBANIA	57	134	54.39	2.35	1.39
107	BARÉIN	56	224	64.29	4.00	0.70
108	JAMAICA	55	435	78.18	7.91	0.69
109	ZIMBABUE	55	513	63.64	9.33	0.37
110	SIRIA	52	39	21.15	0.75	0.16
111	BENÍN	50	122	58.00	2.44	0.37
112	MAURICIO	50	347	84.00	6.94	3.96
113	LIBIA	49	452	57.14	9.22	0.48
114	GHANA	44	91	52.27	2.07	0.28
115	KENIA	40	146	67.50	3.65	0.10
116	UGANDA	38	146	60.53	3.84	0.10
117	MAURITANIA	35	146	68.57	4.17	0.23
118	KIRGUISTÁN	32	21	40.62	0.66	1.15
119	GABÓN	31	45	54.84	1.45	2.02
120	TANZANIA	30	125	66.67	4.17	0.09
121	FIYI	28	153	67.86	5.46	2.23
122	SUAZILANDIA	26	165	76.92	6.35	2.16
123	BARBADOS	24	87	66.67	3.62	10.53
124	SUDÁN	24	59	79.17	2.46	0.08
125	REPÚBLICA DOMINICANA	22	150	50.00	6.82	0.56
126	NÍGER	22	91	59.09	4.14	0.14
127	MADAGASCAR	18	78	66.67	4.33	0.04

Anexo 2. Ranking de 25 países de la región con producción científica matemática durante el período 2000-2017, según el *InCites*.

No.	Países	Ndoc	Cr	% CD	Cr/Ndoc	SDI
1	BRASIL	12239	70955	71.26	5.8	3.32
2	MÉXICO	5117	23189	69.02	4.53	2.38
3	CHILE	3804	26358	74.95	6.93	11.97
4	ARGENTINA	3029	17760	73.03	5.86	4
5	COLOMBIA	1210	3674	62.23	3.04	1.36
6	VENEZUELA	789	4588	69.58	5.81	1.6
7	URUGUAY	395	2269	71.9	5.74	6.22
8	CUBA	237	1018	64.56	4.3	1.13
9	PERÚ	117	516	62.39	4.41	0.2
10	ECUADOR	67	233	52.24	3.48	0.23
11	COSTA RICA	66	410	71.21	6.21	2.06
12	JAMAICA	55	435	78.18	7.91	0.69
13	BARBADOS	24	87	66.67	3.62	10.53
14	REPÚBLICA DOMINICANA	22	150	50	6.82	0.56
15	TRINIDAD Y TOBAGO	12	8	25	0.67	1.47
16	BAHAMAS	8	18	75	2.25	2.55
17	PARAGUAY	8	36	62.5	4.5	0.30
18	BOLIVIA	6	9	33.33	1.5	0.18
19	PANAMÁ	4	5	50	1.25	0.25
20	SURINAM	3	3	100	1	1.79
21	DOMINICA	1	2	100	2	13.6
22	EL SALVADOR	1	0	0	0	0.16
23	GUATEMALA	1	2	100	2	0.06
24	HAITI	1	1	100	1	0.09
25	HONDURAS	1	0	0	0	0.11

Anexo 3. Distribución de la producción científica matemática de los Estados de la República Mexicana, se excluyeron del análisis la UNAM, IPN, IMSS y el IMP por la distribución de sus dependencias a lo largo del territorio nacional.

Instituciones	Ndoc	Estados
Universidad Autónoma Metropolitana	490	Ciudad de México
Centro de Investigación y de Estudios Avanzados	485	
Centro de Investigación en Matemáticas	438	
Instituto Tecnológico Autónomo de México	132	
Hospital General de México	4	
Universidad Panamericana	4	
Instituto de Ecología	3	
Centro de Investigación y Docencia Económicas	3	
Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán	2	
Instituto Nacional de Cardiología	2	
Benemérita Universidad Autónoma de Puebla	125	Puebla
Universidad de las Américas Puebla	41	
Instituto Nacional de Astrofísica; Óptica y Electrónica	33	
Universidad Autónoma de San Luis Potosí	111	San Luis Potosí
Instituto Potosino Investigación Científica y Tecnológica	29	
Tecnológico de Monterrey	91	Nuevo León
Universidad Autónoma de Nuevo León	39	
Universidad Autónoma del Estado de Morelos	103	Morelos
Instituto Nacional de Salud Pública	5	
Universidad Michoacana de San Nicolás de Hidalgo	103	Michoacán
Colegio de Postgraduados	30	Estado de México
Universidad Autónoma del Estado de México	29	
Centro Internacional de Mejoramiento de Maíz y Trigo	9	
Universidad Anáhuac	6	
Universidad Autónoma de Aguascalientes	59	
Universidad de Guanajuato	49	Guanajuato
Centro de Investigaciones en Óptica	9	
Universidad Autónoma del Estado de Hidalgo	56	Hidalgo
Universidad de Sonora	54	Sonora
Universidad Autónoma de Yucatán	47	Yucatán
Centro de Investigación Científica de Yucatán	2	
Universidad de Colima	33	Colima
Universidad de Guadalajara	29	Jalisco
Universidad Autónoma de Querétaro	25	Querétaro
Centro de Ingeniería y Desarrollo Industrial	1	
Centro de Investigación y Desarrollo Tecnológico en Electroquímica	1	
Universidad Juárez Autónoma de Tabasco	27	Tabasco
Centro de Investigación Científica y de Educación Superior de Ensenada	14	Baja California
Universidad Autónoma de Baja California	11	
Universidad Autónoma de Zacatecas	23	Zacatecas
Universidad Autónoma de Chiapas	18	Chiapas
Universidad Autónoma de Sinaloa	9	Sinaloa
Universidad Autónoma de Tamaulipas	9	Tamaulipas
Universidad Autónoma de Ciudad Juárez	7	Chihuahua
Universidad Autónoma de Coahuila	6	Coahuila
Universidad del Papaloapan	3	Oaxaca
Universidad Autónoma Benito Juárez de Oaxaca	3	
Universidad Autónoma del Carmen	2	Campeche

El Colegio de la Frontera Sur	1	
Universidad Juárez del Estado de Durango	3	Durango
Universidad de Quintana Roo	2	Quintana Roo

Anexo 4. Listado de los artículos considerados *highly cited papers*, según el *ESI*.

Artículos <i>highly cited papers</i>	Citas
<b>A1-</b> Blaauw, M & Christen, J. A.* (2011). Flexible Paleoclimate Age-Depth Models Using an Autoregressive Gamma Process. <i>Bayesian Analysis</i> , 6(3), 457-474. DOI: 10.1214/11-BA618 *Centro de Investigaciones Matemáticas – CIMAT.	577
<b>A2-</b> Capella, A.,* Davila, J., Dupaigne, L. & Sire, Y. (2011). Regularity of Radial Extremal Solutions for Some Non-Local Semilinear Equations. <i>Communications in Partial Differential Equations</i> , 36(8), 1353-1384. DOI: 10.1080/03605302.2011.562954 *Universidad Nacional Autónoma de México, Instituto de Matemática.	92
<b>A3-</b> Cingolani, S., Clapp, M.* & Secchi, S. (2012). Multiple solutions to a magnetic nonlinear Choquard equation. <i>Zeitschrift fur Angewandte Mathematik und Physik</i> , 63(2), 233-248. DOI: 10.1007/s00033-011-0166-8 *Universidad Nacional Autónoma de México, Instituto de Matemática.	52
<b>A4-</b> Trejo-Guerra, R.,* Tlelo-Cuautle, E.,** Carbajal-Gómez, V. H.** & Rodríguez-Gómez, G.** (2013). A survey on the integrated design of chaotic oscillators. <i>Applied Mathematics and Computation</i> , 219(10), 5113-5122. DOI: 10.1016/j.amc.2012.11.021 *SEMTECH Centro de Investigaciones en Diseño Industrial, México **INAOE, Departamento de Electrónica, México	35
<b>A5-</b> Clapp, M.* & Salazar, D.* (2013). Positive and sign changing solutions to a nonlinear Choquard equation. <i>Journal of Mathematical Analysis and Applications</i> , 407(1), 1-15. DOI: 10.1016/j.jmaa.2013.04.081 *Universidad Nacional Autónoma de México, Instituto de Matemática.	31
<b>A6-</b> Gómez Aguilar, J. F.* & Baleanu, D. (2014). Solutions of the telegraph equations using a fractional calculus approach. <i>Proceedings of the Romanian Academy Series A-Mathematics Physics Technical Sciences Information Science</i> , 15(1), 27-34. *Universidad Nacional Autónoma de México, Instituto de Energías Renovables.	29
<b>A7-</b> Shah, N. H. & Cardenas-Barron, L. E.* (2015). Retailer's decision for ordering and credit policies for deteriorating items when a supplier offers order-linked credit period or cash discount. <i>Applied Mathematics and Computation</i> , 259, 569-578. DOI: 10.1016/j.amc.2015.03.010 *Tecnológico de Monterrey, México	20
<b>A8-</b> Gomes, D. A., Pimentel, E. A. & Sánchez-Morgado, H.* (2015). Time-Dependent Mean-Field Games in the Subquadratic Case. <i>Communications in Partial Differential Equations</i> , 40(1), 40-76. DOI: 10.1080/03605302.2014.903574 *Universidad Nacional Autónoma de México, Instituto de Matemática.	19
<b>A9-</b> Rodríguez, J. M. & Sigarreta, J. M.* (2016). Spectral properties of geometric-arithmetic index. <i>Applied Mathematics and Computation</i> , 277, 142-153. DOI: 10.1016/j.amc.2015.12.046 *Universidad Autónoma de Guerrero, Facultad de Matemática.	13
<b>A10-</b> Gómez-Aguilar, J. F.*, Torres, L.**, Yopez-Martínez, H.***, Baleanu, D., Reyes, J. M.*** & Sosa, I. O.*** (2016). Fractional Lienard type model of a pipeline within the fractional derivative without singular kernel. <i>Advances in Difference Equations</i> , 173. DOI: 10.1186/s13662-016-0908-1 *Centro Nacional de Investigación & Desarrollo Tecnológico, Morelos. **Universidad Nacional Autónoma de México, Instituto de Ingeniería. ***Universidad Autónoma de la Ciudad México.	12

<p><b>A11-</b> Davini, A., Fathi, A., Iturriaga, R.* &amp; Zavidovique, M. (2016). Convergence of the solutions of the discounted Hamilton-Jacobi equation. <i>Inventiones Mathematicae</i>, 206(1), 29-55. DOI: 10.1007/s00222-016-0648-6</p> <p>*Centro de Investigaciones Matemáticas.</p>	12
<p><b>A12-</b> Morales-Delgado, V. F.* , Gómez-Aguilar, J. F.** , Yopez-Martínez, H.***, Baleanu, D., Escobar-Jiménez, R. F.** &amp; Olivares-Peregrino, V. H.** (2017). Laplace homotopy analysis method for solving linear partial differential equations using a fractional derivative with and without kernel singular. <i>Advances in Difference Equations</i>, 164. DOI: 10.1186/s13662-016-0891-6</p> <p>*Universidad Autónoma de Guerrero, México.  **Centro Nacional de Investigación &amp; Desarrollo Tecnológico, Morelos.  ***Universidad Autónoma de la Ciudad México.</p>	11
<p><b>A13-</b> Lanery, S. &amp; Thiemann, T. (2017). Projective limits of state spaces I. Classical formalism. <i>Journal of Geometry and Physics</i>, 111, 6-39. DOI: 10.1016/j.geomphys.2016.10.010</p> <p>*Universidad Nacional Autónoma de México</p>	9

Anexo 5. Ranking de 55 instituciones mexicanas con producción científica matemática durante el período 2000-2017, según el *InCites*.

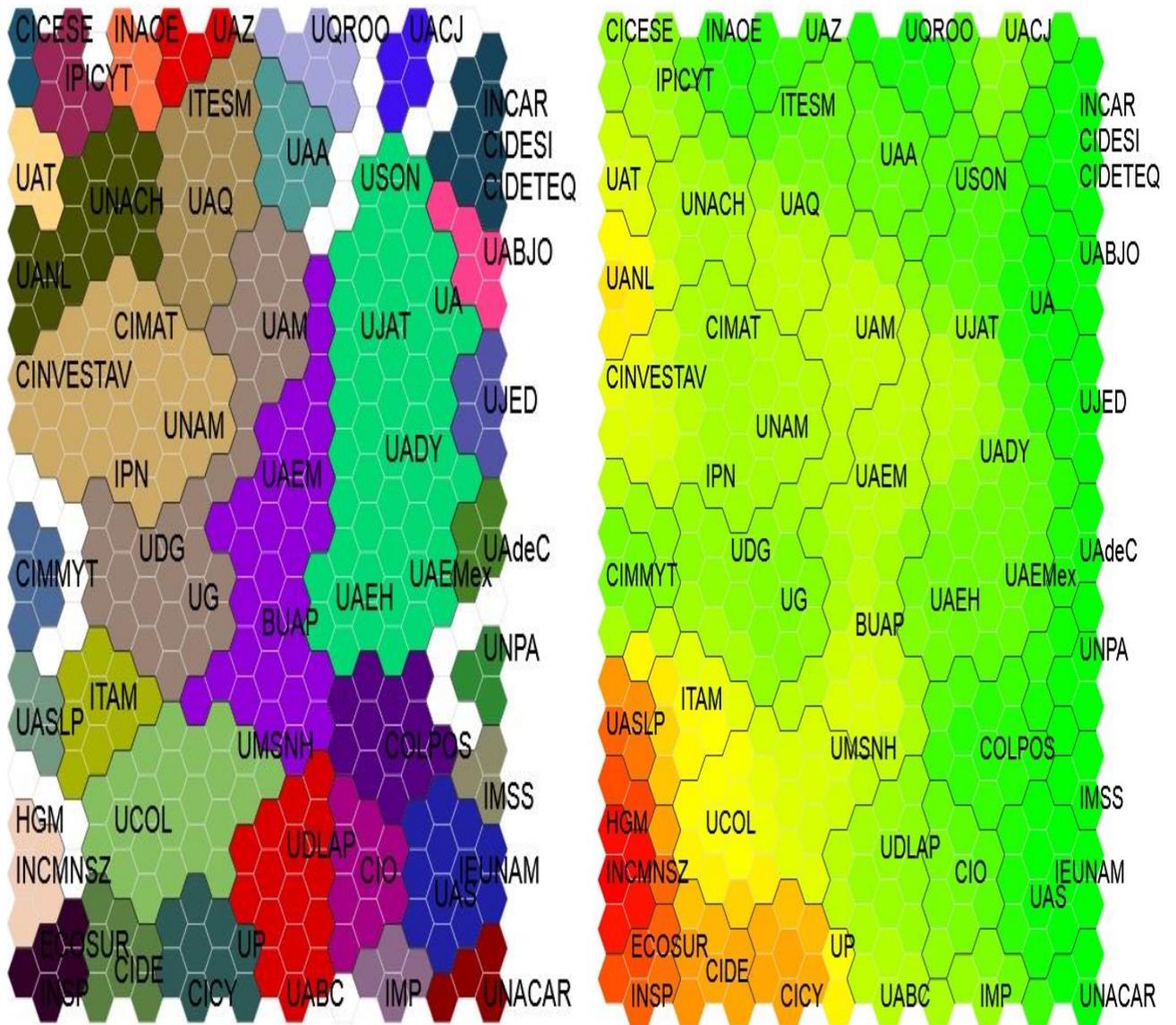
No.	Instituciones	Acrónimo	Ndoc	Cr	Cr/Ndoc	DC	Cr/Dc
1	Universidad Nacional Autónoma de México	UNAM	2,215	10,111	5	1,552	6
2	Instituto Politécnico Nacional	IPN	544	3,011	6	411	7
3	Universidad Autónoma Metropolitana	UAM	490	1,920	4	320	6
4	Centro de Investigación y de Estudios Avanzados	CINVESTAV	487	3,335	7	371	9
5	Centro de Investigación en Matemáticas	CIMAT	439	2,638	6	295	9
6	Instituto Tecnológico Autónomo de México	ITAM	132	691	5	88	8
7	Benemérita Universidad Autónoma de Puebla	BUAP	125	357	3	81	4
8	Universidad Autónoma de San Luis Potosí	UASLP	111	604	5	83	7
9	Universidad Autónoma del Estado de Morelos	UAEM	104	331	3	67	5
10	Universidad Michoacana de San Nicolás de Hidalgo	UMSNH	103	302	3	67	4
11	Tecnológico de Monterrey	ITESM	91	299	3	47	6
12	Universidad Autónoma de Aguascalientes	UAA	59	173	3	35	5
13	Universidad Autónoma del Estado de Hidalgo	UAEH	56	130	2	35	4
14	Universidad de Sonora	USON	54	140	3	27	5
15	Universidad de Guanajuato	UG	49	268	5	35	8
16	Universidad Autónoma de Yucatán	UADY	47	83	2	30	3
17	Universidad Américas Puebla	UDLAP	41	174	4	31	6
18	Universidad Autónoma de Nuevo León	UANL	39	243	6	24	10
19	Instituto Nacional de Astrofísica, Óptica y Electrónica	INAOE	33	203	6	28	7
20	Universidad de Colima	UCOL	33	82	2	28	3
21	Colegio de Postgraduados	COLPOS	30	99	3	23	4
22	Instituto Potosino Investigación Científica y Tecnológica	IPICYT	29	129	4	25	5
23	Universidad de Guadalajara	UDG	29	105	4	21	5
24	Universidad Autónoma del Estado de México	UAEMex	29	76	3	16	5
25	Universidad Juárez Autónoma de Tabasco	UJAT	27	31	1	15	2
26	Universidad Autónoma de Querétaro	UAQ	26	61	2	15	4
27	Universidad Autónoma de Zacatecas	UAZ	23	89	4	12	7
28	Instituto Mexicano del Petróleo	IMP	19	129	7	16	8
29	Universidad Autónoma de Chiapas	UNACH	18	71	4	12	6
30	Centro de Investigación Científica y de Educación Superior de Ensenada	CICESE	14	192	14	12	16
31	Universidad Autónoma de Baja California	UABC	11	26	2	6	4
32	Universidad Autónoma de Tamaulipas	UAT	9	85	9	8	11
33	Centro Internacional de Mejoramiento de Maíz y Trigo	CIMMYT	9	80	9	9	9
34	Centro de Investigaciones en Óptica	CIO	9	46	5	9	5
35	Universidad Autónoma de Sinaloa	UAS	9	36	4	8	4
36	Universidad Autónoma de Ciudad Juárez	UACJ	7	0	0	0	0

37	Universidad Autónoma de Coahuila	UAdeC	6	20	3	3	7
38	Universidad Anáhuac	UA	6	8	1	3	3
39	Instituto Nacional de Salud Pública	INSP	5	11	2	4	3
40	Hospital General de México	HGM	4	10	3	3	3
41	Universidad Panamericana	UNPA	4	2	1	2	1
42	Instituto de Ecología	IEUNAM	3	16	5	3	5
43	Centro de Investigación y Docencia Económicas	CIDE	3	8	3	2	4
44	El Colegio de la Frontera Sur	ECOSUR	3	7	2	3	2
45	Universidad Juárez del Estado de Durango	UJED	3	6	2	2	3
46	Universidad del Papaloapan	UP	3	2	1	1	2
47	Universidad Autónoma Benito Juárez de Oaxaca	UABJO	3	1	0	1	1
48	Universidad Autónoma del Carmen	UNACAR	2	13	7	2	7
49	Universidad de Quintana Roo	UQROO	2	7	4	1	7
50	Instituto Mexicano del Seguro Social	IMSS	2	6	3	2	3
51	Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán	INCMNSZ	2	3	2	2	2
52	Centro de Investigación Científica de Yucatán	CICY	2	2	1	1	2
53	Instituto Nacional de Cardiología	INCAR	2	0	0	0	0
54	Centro de Ingeniería y Desarrollo Industrial	CIDESI	1	0	0	0	0
55	Centro de Investigación y Desarrollo Tecnológico en Electroquímica	CIDETEQ	1	0	0	0	0

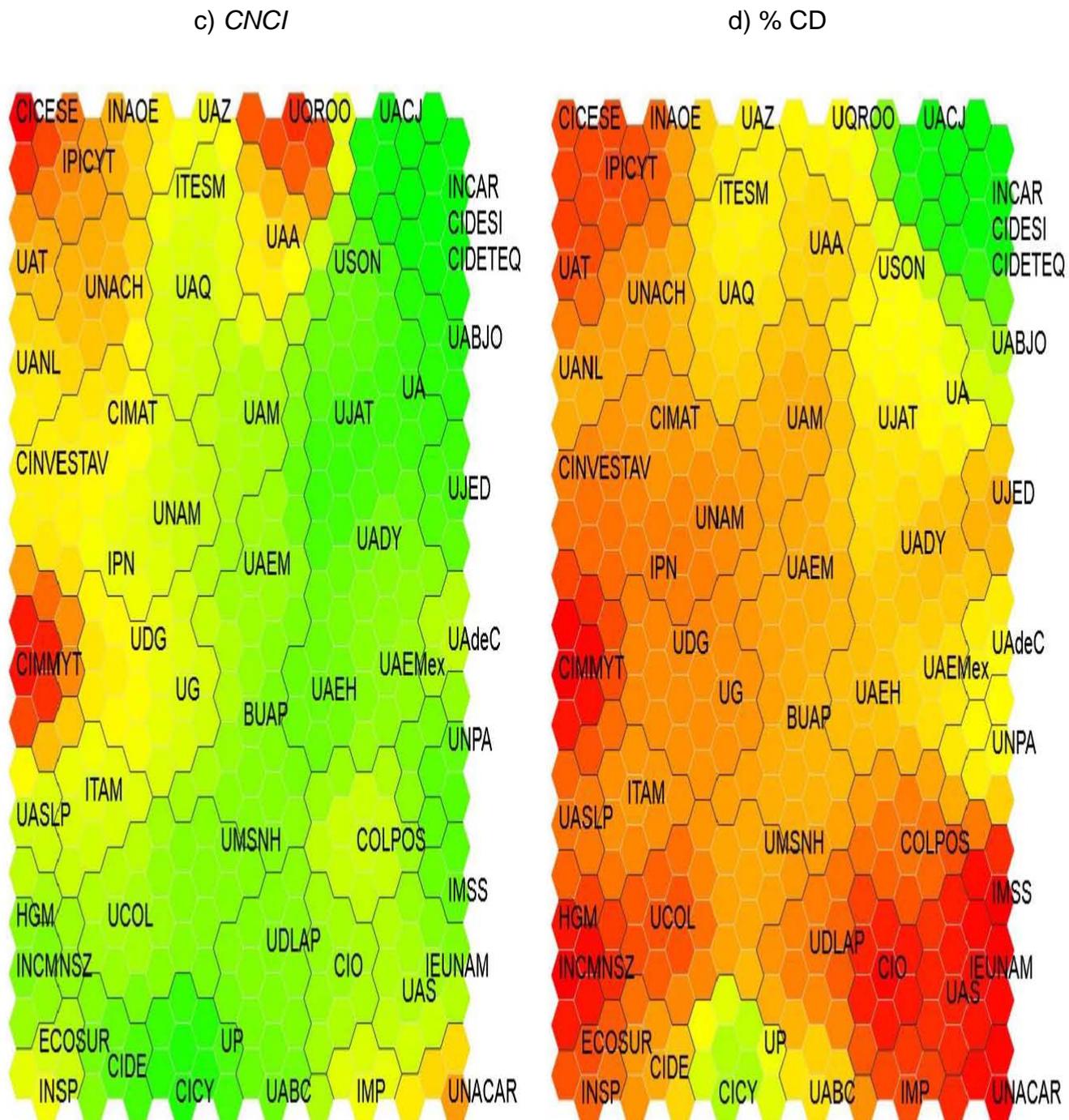
Anexo 6. Mapa de *cluster* de las instituciones mexicanas que producen matemáticas, donde se analizan multiparamétricamente los indicadores (%Q1, CNCI, %CD, %HP y %IC). Mapa de componente del indicador del porcentaje de documentos en el cuartil (%Q1).

a) Mapa de *Clusters*

b) % Q1



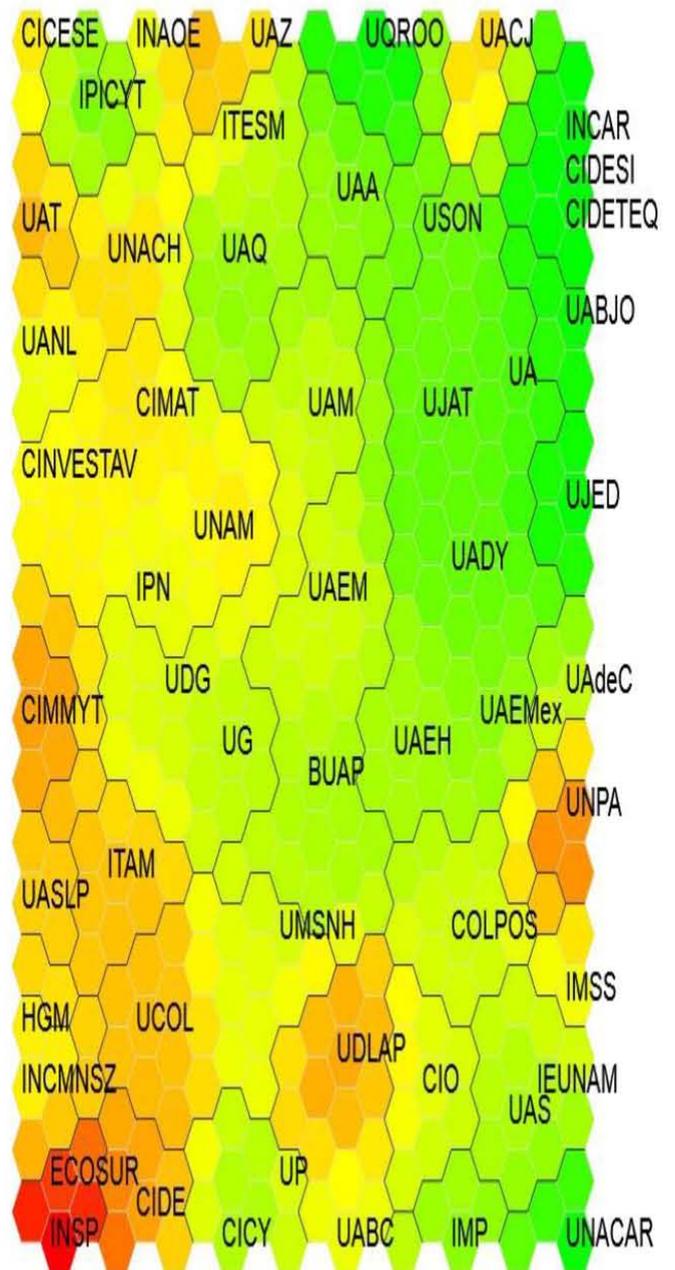
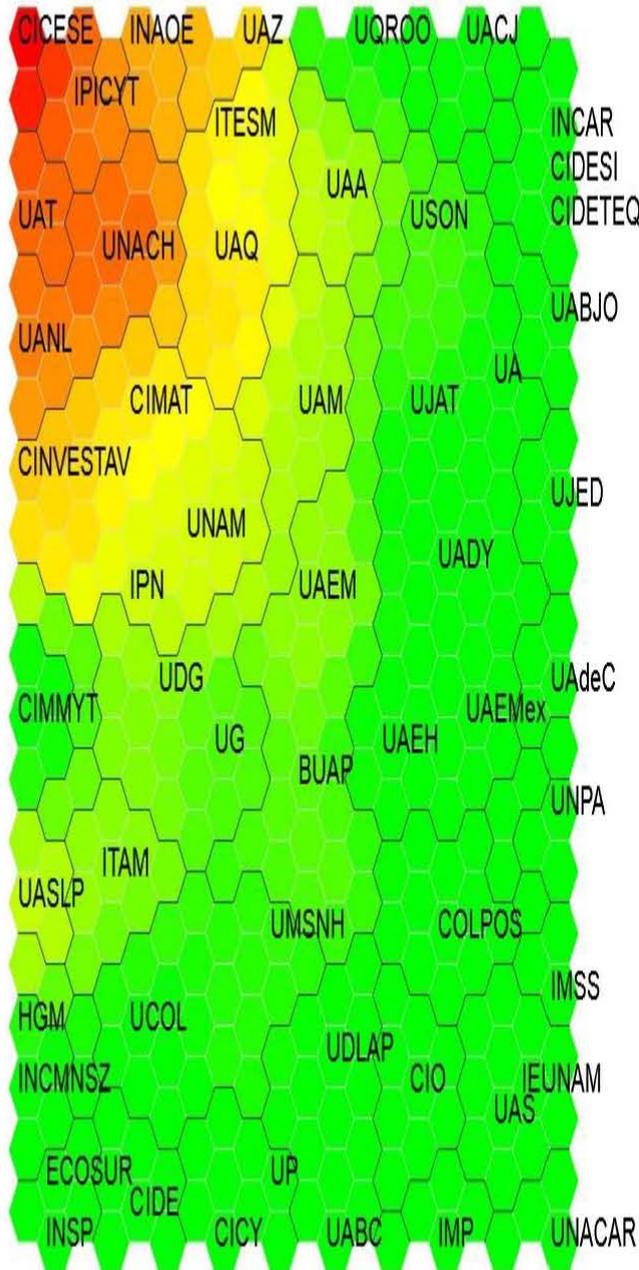
Anexo 7. Mapas de componentes de las instituciones para los indicadores del *Category Normalized Citation Impact* (CNCI) y el porcentaje de Documentos Citados (%CD).



Anexo 8. Mapas de componentes de las instituciones para los indicadores del porcentaje de *High Performance* (%HP) y el porcentaje de artículos en Colaboración Internacional (%IC).

f) % HP

g) % IC



Anexo 9. Tabla con los valores de los indicadores (%Q1, %HP, %CD, CNCI y %IC) utilizados en el análisis multiparamétrico realizado a las instituciones mexicanas que producen matemática.

Nombre de la institución	Siglas	CNCI	% Q1	% DC	% HP	% IC
Centro de Investigación Científica y de Educación Superior de Ensenada	CICESE	1.76	14.29	85.71	14.29	57.14
Centro Internacional de Mejoramiento de Maíz y Trigo	CIMMYT	1.63	11.11	100	0	66.67
Universidad de Quintana Roo	UQROO	1.55	0	50	0	0
Universidad Autónoma del Carmen	UNACAR	1.20	0	100	0	0
Instituto Potosino Investigación Científica y Tecnológica	IPICYT	1.18	17.24	86.21	10.34	13.79
Instituto Nacional de Astrofísica; Óptica y Electrónica	INAOE	1.12	0	84.85	9.09	45.45
Universidad Autónoma de Chiapas	UNACH	1.08	16.67	66.67	11.11	55.56
Universidad Autónoma de Tamaulipas	UAT	1.06	22.22	88.89	11.11	66.67
Universidad Autónoma de Aguascalientes	UAA	0.93	10.17	59.32	5.08	27.12
Universidad Autónoma de Nuevo León	UANL	0.90	28.21	61.54	10.26	43.59
Universidad de Guadalajara	UDG	0.90	13.79	72.41	3.45	41.38
Centro de Investigación en Matemáticas	CIMAT	0.89	16.17	68.34	7.29	52.85
Centro de Investigación y de Estudios Avanzados	CINVESTAV	0.88	21.56	78.03	7.80	50.31
Hospital General de México	HGM	0.88	75	75	0	25
Universidad Autónoma de Zacatecas	UAZ	0.85	4.35	52.17	8.70	65.22
Instituto Nacional de Salud Pública	INSP	0.84	40	80	0	100
Instituto Mexicano del Petróleo	IMP	0.82	15.79	84.21	0	26.32
Instituto Politécnico Nacional	IPN	0.76	15.99	76.47	5.88	45.59
Instituto Tecnológico Autónomo de México	ITAM	0.75	25	68.94	3.79	60.61
Universidad de Guanajuato	UG	0.74	12.24	71.43	2.04	36.73
Tecnológico de Monterrey	ITESM	0.73	9.89	51.65	6.59	37.36
Universidad Autónoma de Querétaro	UAQ	0.70	19.23	57.69	7.69	26.92
Universidad Autónoma de Coahuila	UAdeC	0.70	0	50	0	33.33
Colegio de Postgraduados	COLPOS	0.69	3.33	76.67	0	40
Universidad Autónoma de San Luis Potosí	UASLP	0.68	40.54	76.58	5.41	57.66
Universidad Nacional Autónoma de México	UNAM	0.67	17.34	70.84	5.73	53.09
Universidad Autónoma de Sinaloa	UAS	0.63	0	88.89	0	33.33
Universidad Autónoma Metropolitana	UAM	0.56	19.59	66.33	4.49	38.37
Universidad de Colima	UCOL	0.55	24.24	84.85	0	60.61
El Colegio de la Frontera Sur	ECOSUR	0.54	100	100	0	0
Instituto de Ecología	IEUNAM	0.53	0	100	0	33.33
Universidad Autónoma del Estado de Morelos	UAEM	0.52	18.27	66.35	3.85	40.38
Centro de Investigaciones en Óptica	CIO	0.50	11.11	100	0	44.44
Universidad Autónoma del Estado de México	UAEMex	0.48	10.34	55.17	0	27.59
Benemérita Universidad Autónoma de Puebla	BUAP	0.45	21.60	64.80	2.40	32.80

Universidad Autónoma de Yucatán	UADY	0.45	10.64	63.83	0	19.15
Universidad Michoacana de San Nicolás de Hidalgo	UMSNH	0.44	21.36	66.99	0.97	43.69
Centro de Investigación Científica de Yucatán	CICY	0.44	100	50	0	0
Universidad de las Américas Puebla	UDLAP	0.43	14.63	75.61	0	65.85
Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán	INCMNSZ	0.40	50	100	0	50
Universidad Autónoma de Baja California	UABC	0.36	18.18	54.55	0	45.45
Universidad Autónoma del Estado de Hidalgo	UAEH	0.35	10.71	62.50	0	32.14
Universidad de Sonora	USON	0.34	7.41	50	1.85	20.37
Instituto Mexicano del Seguro Social	IMSS	0.29	0	100	0	50
Centro de Investigación y Docencia Económicas	CIDE	0.28	33.33	66.67	0	66.67
Universidad Panamericana	UNPA	0.28	0	50	0	75
Universidad Juárez del Estado de Durango	UJED	0.23	0	66.67	0	0
Universidad Juárez Autónoma de Tabasco	UJAT	0.19	18.52	55.56	0	18.52
Universidad Anáhuac	UA	0.19	0	50	0	16.67
Universidad del Papaloapan	UP	0.10	33.33	33.33	0	33.33
Universidad Autónoma Benito Juárez de Oaxaca	UABJO	0.10	0	33.33	0	0
Universidad Autónoma de Ciudad Juárez	UACJ	0	14.29	0	0	57.14
Instituto Nacional de Cardiología	INCAR	0	0	0	0	0
Centro de Ingeniería y Desarrollo Industrial	CIDESI	0	0	0	0	0
Centro de Investigación y Desarrollo Tecnológico en Electroquímica	CIDETEQ	0	0	0	0	0

Anexo 10. Perfiles de los 34 *clusters* identificados en el análisis multiparamétrico realizado a las instituciones.

Indicadores	A*	M**	B***	Tabla de los valores de los intervalos para caracterizar los perfiles de desempeño Dónde: A*: valor alto M**: valor medio B***: valor bajo
% Q1	>80%	50-80%	<50%	
CNCI	>1.0	0.7-1.0	<0.7	
% DC	>80%	50-80%	<50%	
% HP	>10.0	6.0-10.0	<6.14	
% IC	>80%	50-80%	<50%	

No.	Instituciones	% Q1	CNCI	% DC	% HP	% IC
1	CICESE	B	A	A	A	M
2	UAT	B	A	A	A	M
3	ECOSUR-INSP	A	M	A	B	A
4	UNACH-UANL	B	A	M	A	M
5	CIMMYT	B	A	A	B	M
6	HGM-INCMNSZ	A	B	A	B	M
7	UNACAR	B	A	A	B	B
8	IPICYT	B	A	M	M	B
9	UQROO	B	A	M	B	B
10	IMP	B	M	A	B	B
11	IMSS	B	B	A	B	M
12	UASLP	B	B	A	B	M
13	CIO	B	B	A	B	B
14	UAS-IEUNAM	B	B	A	B	B
15	UAZ	B	M	M	M	M
16	IPN-UNAM- CINVESTAV-CIMAT	B	M	M	M	M
17	UAQ-ITESM	B	M	M	M	B
18	ITAM	B	M	M	B	M
19	UMSNH-UCOL	M	B	M	B	M
20	UNPA	B	B	M	B	M
21	UAA	B	M	M	B	B
22	INAOE	B	M	M	B	B
23	UAdeC	B	M	M	B	B
24	UDG-UG-UAM	B	M	M	B	B
25	CIDE	B	B	M	B	M
26	UDLAP-UABC	B	B	M	B	M
27	UACJ	B	B	B	B	M
28	UJED	B	B	M	B	B
29	UADY-UAEH- UAEMex-USON-UA- UJAT	B	B	M	B	B
30	COLPOS	B	B	M	B	B
31	UAEM-BUAP	B	B	M	B	B
32	CICY-UP	M	B	B	B	B
33	UABJO	B	B	B	B	B
34	INCAR-CIDESI- CIDETEQ	B	B	B	B	B

Anexo 11. Mapa de *clusters* de las categorías temáticas, donde se analiza multiparamétricamente los indicadores (%Q1, CNCI, %CD, %HP, %Exc y %IC).

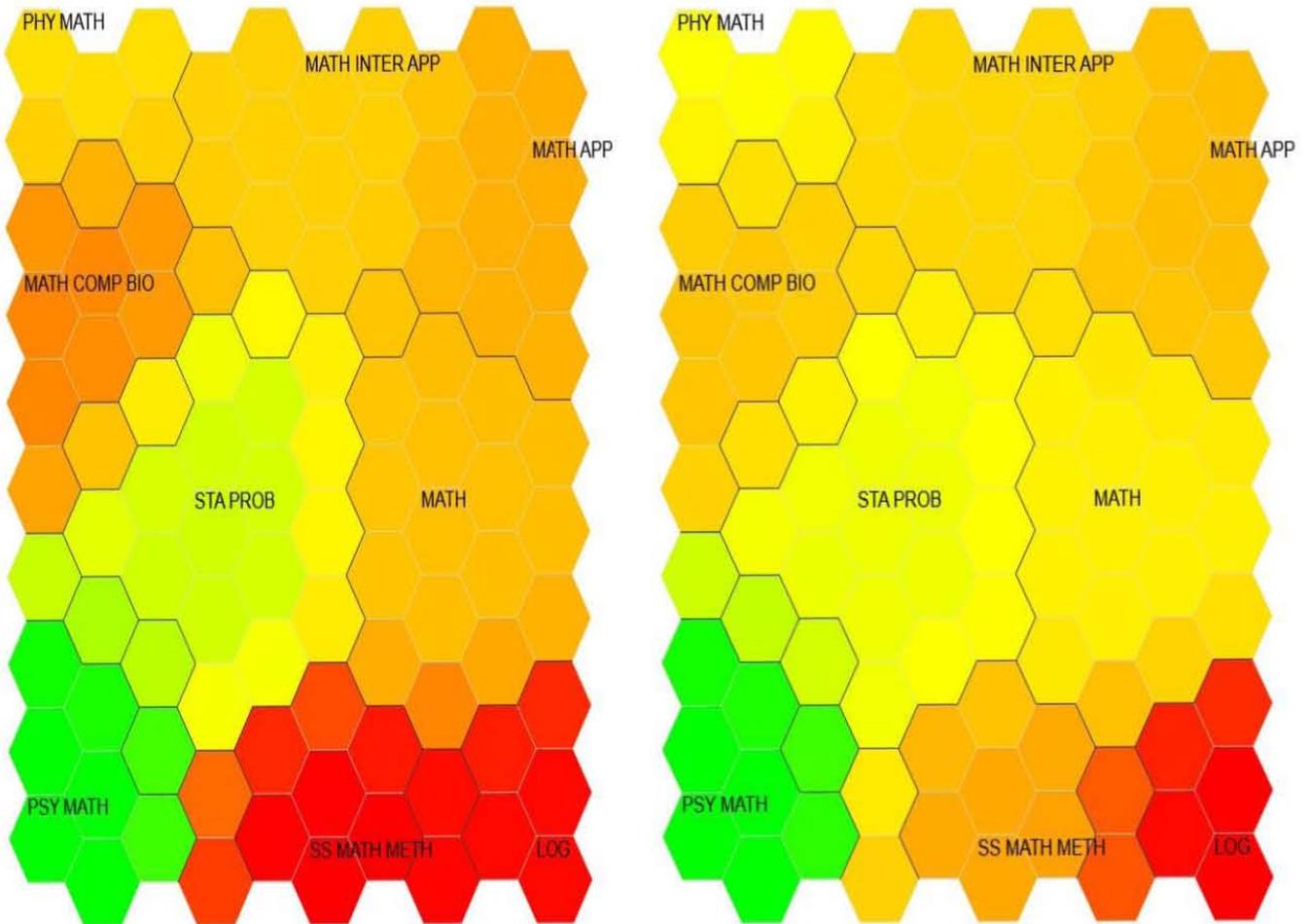
a) Mapa de *Clusters*



Anexo 12. Mapas de componentes de las categorías temáticas para los indicadores *Category Normalized Citation Impact* (CNCI) y el porcentaje de *High Performance* (%HP).

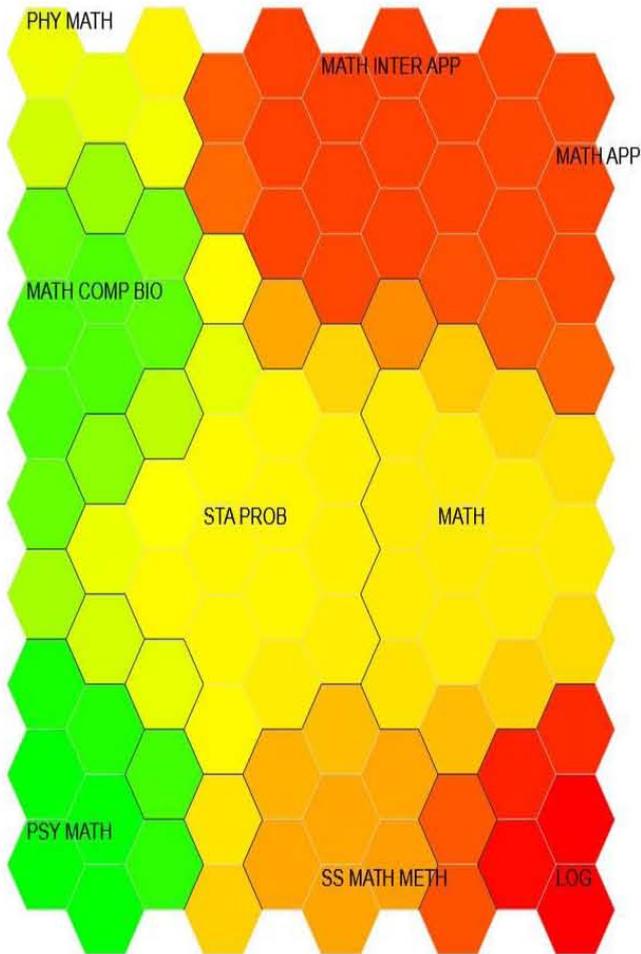
b) CNCI

c) % HP

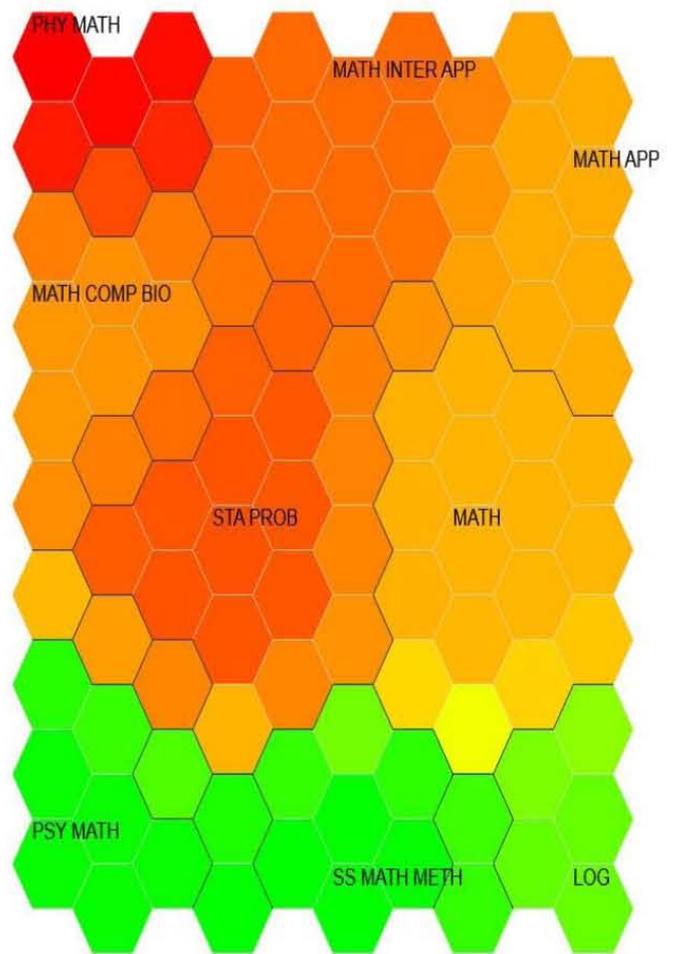


Anexo 13. Mapas de componentes de las categorías temáticas para los indicadores del porcentaje de Excelencia Científica (%Exc) y el porcentaje de Documentos Citados (%CD).

d) % Exc

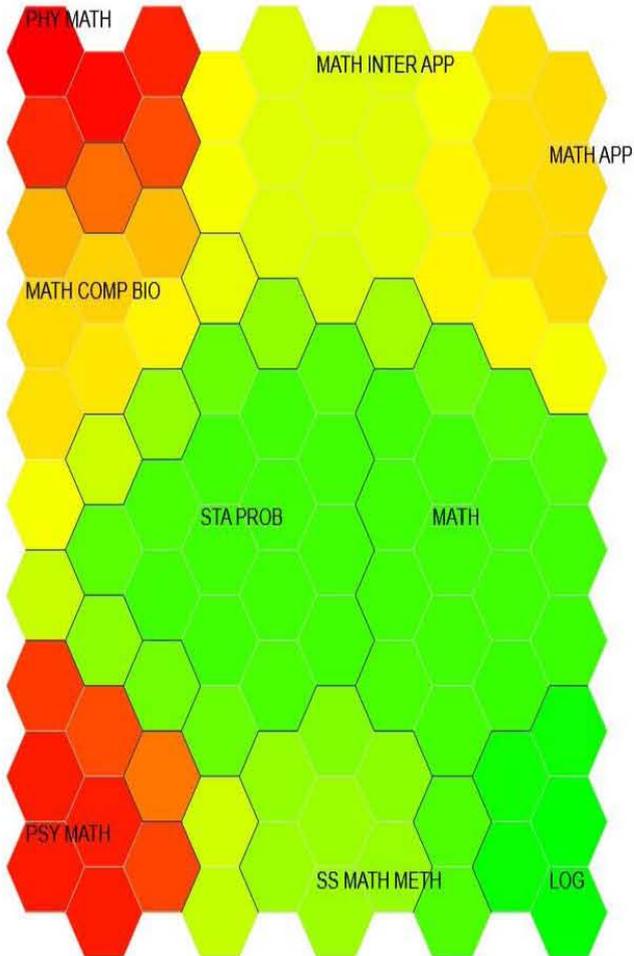


e) % CD

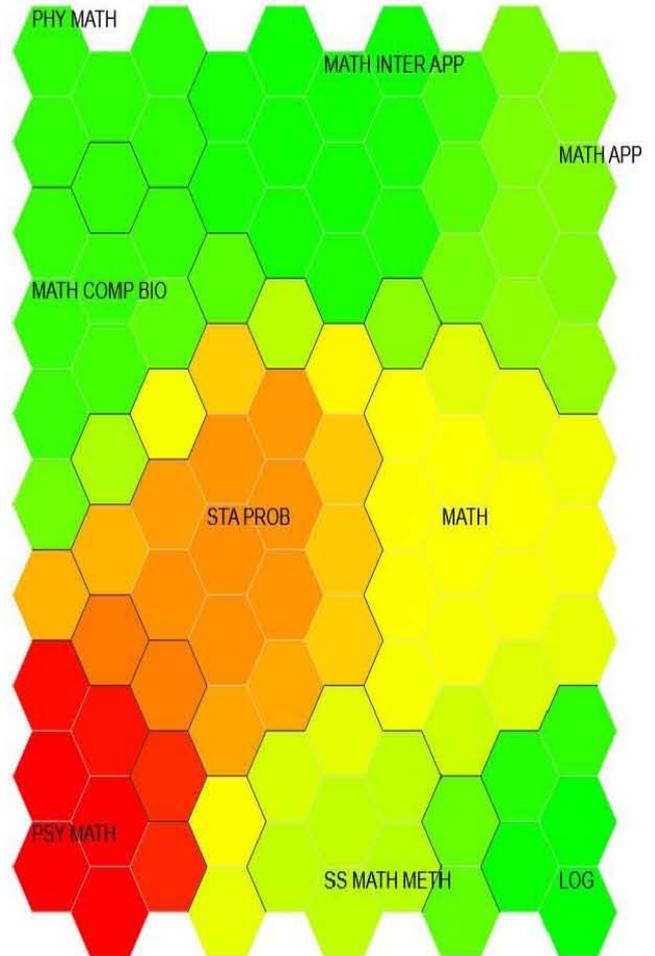


Anexo 14. Mapas de componentes de las categorías temáticas para los indicadores del porcentaje de artículos publicados en revistas del Cuartil Uno (%Q1) y el porcentaje de artículos en Colaboración Internacional (%IC).

f) % Q1



g) % IC



Anexo 15. Perfiles de los ocho *clusters* identificados en el análisis multiparamétrico realizado a las categorías temáticas.

Indicadores	A*	M**	B***	Tabla de los valores de los intervalos para caracterizar los perfiles de desempeño Dónde: A*: valor alto M**: valor medio B***: valor bajo
% Q1	>80%	50-80%	<50%	
CNCI	>1.0	0.7-1.0	<0.7	
% DC	>80%	50-80%	<50%	
% HP	>10.0	6.0-10.0	<6.14	
% IC	>80%	50-80%	<50%	
% Exc	>0.90	0.50-0.90	<0.50	

Categorías Temáticas	CNCI	%HP	%IC	%Q1	% Exc	%CD
MATH	M	M	M	B	M	M
MATH APP	M	M	B	B	M	M
PHY MATH	B	B	B	B	B	M
MATH INTER APP	M	B	B	B	M	M
STA PROB	B	B	M	B	B	M
MATH COMP BIO	M	M	B	B	B	M
SS MATH METH	M	M	B	B	M	M
LOG	M	A	B	B	M	M
PSY MATH	B	B	B	B	B	M