

INEGI-Móvil: desarrollo de un SIG móvil en código libre para levantamiento de datos en campo

Alonso Morilla Meneses, Alejandra Albert Tejera y Silvia Franceschi

Trabajo de cuidado en las fuentes de información estadística de México

César González-González, Karina Orozco-Rocha, Mireya Patricia Arias Soto y Ma. Gregoria Carvajal Santillán

Ser entrevistadora o entrevistador. La recolección de datos como un oficio, el aprendizaje de un saber hacer

Eleonora Nuricumbo Rivera, Candi Uribe Pineda y Adriana Marcela Meza Calleja

Estandarización de las áreas geoestadísticas básicas urbanas para 100 ciudades del Sistema Urbano Nacional

Jorge A. Montejano Escamilla, F. Gerardo Ávila Jiménez y Camilo Alberto Caudillo Cos

Moda, media y mediana de la altura del relieve mexicano

Rodrigo Tovar Cabañas, José Alfredo Jáuregui Díaz y Shany Arely Vázquez Espinosa

Indicadores de actividad económica municipal 1993-2013

Jesús López-Pérez y Francisco Corona

Cubo de datos geoespaciales para el uso de las imágenes satelitales en la generación de información geográfica y estadística

Olivia Jimena Juárez Carrillo, Paloma Merodio Gómez, María del Socorro Ponce Medina, José Luis Ornelas de Anda y Abel Alejandro Coronado Iruegas

***importinegi*: un paquete de R para descargar y gestionar bases de datos del INEGI**

Cesar Rentería

Análisis de eficiencia en los servicios del agua en México con datos de los Censos Económicos

Ismael Aguilar-Benítez

Reformar la educación: ¿misión imposible? Los consejos y experiencias de los líderes educativos

Reseña

Otto Granados Roldán

Contenido

INEGI-Móvil: desarrollo de un SIG móvil en código libre para levantamiento de datos en campo	6
<i>INEGI-Mobile: Development of a Free Software Mobile GIS for Field Data Surveys</i>	
Alonso Morilla Meneses, Alejandra Albert Tejera y Silvia Franceschi	
Trabajo de cuidado en las fuentes de información estadística de México	22
<i>Care Work in the Statistical Sources of Information in Mexico</i>	
César González-González, Karina Orozco-Rocha, Mireya Patricia Arias Soto y Ma. Gregoria Carvajal Santillán	
Ser entrevistadora o entrevistador. La recolección de datos como un oficio, el aprendizaje de un saber hacer	44
<i>To Be a Female Interviewer or a Male Interviewer. Data Compilation as a Craft, the Learning of a Know-How</i>	
Eleonora Nuricumbo Rivera, Candi Uribe Pineda y Adriana Marcela Meza Calleja	
Estandarización de las áreas geoestadísticas básicas urbanas para 100 ciudades del Sistema Urbano Nacional	62
<i>Standardization of Basic Urban Geostatistical Areas for 100 Cities of the National Urban System</i>	
Jorge A. Montejano Escamilla, F. Gerardo Ávila Jiménez y Camilo Alberto Caudillo Cos	
Moda, media y mediana de la altura del relieve mexicano	90
<i>Mode, Mean and the Median of the Height of the Mexican Relief</i>	
Rodrigo Tovar Cabañas, José Alfredo Jáuregui Díaz y Shany Arely Vázquez Espinosa	
Indicadores de actividad económica municipal 1993-2013	102
<i>Municipality-Level Economic Activity Indicators 1993-2013</i>	
Jesús López-Pérez y Francisco Corona	
Cubo de datos geospaciales para el uso de las imágenes satelitales en la generación de información geográfica y estadística	124
<i>The Use of Satellite Imagery in the Generation of Geographic and Statistical Information</i>	
Olivia Jimena Juárez Carrillo, Paloma Merodio Gómez, María del Socorro Ponce Medina, José Luis Ornelas de Anda y Abel Alejandro Coronado Iruegas	
importinegi: un paquete de R para descargar y gestionar bases de datos del INEGI	140
<i>importinegi: an R Package to Download and Manage INEGI's Data Sets</i>	
César Rentería	
Análisis de eficiencia en los servicios del agua en México con datos de los Censos Económicos	150
<i>Analysis of Efficiency of Water Services in Mexico Using Data from the Economic Censuses</i>	
Ismael Aguilar-Benítez	
Reformar la educación: ¿misión imposible? Los consejos y experiencias de los líderes educativos	168
<i>Reforming Education: Impossible Mission? Sharing Advises and Experiences from Educational Leaders</i>	
Reseña	
Otto Granados Roldán	
Colaboran en este número	174

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Presidente del Instituto

Julio Alfonso Santaella Castell

Vicepresidentes

Enrique de Alba Guerra

Paloma Merodio Gómez

Enrique Jesús Ordaz López

Adrián Franco Barrios

Dirección General de Estadísticas Sociodemográficas

Edgar Vielma Orozco

Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia

Óscar Jaimes Bello

Dirección General de Estadísticas Económicas

José Arturo Blancas Espejo

Dirección General de Geografía y Medio Ambiente

María del Carmen Reyes Guerrero

Dirección General de Integración, Análisis e Investigación

Sergio Carrera Riva Palacio

Dirección General de Coordinación del Sistema Nacional de Información Estadística y Geográfica

María Isabel Monterrubio Gómez

Dirección General de Comunicación, Servicio Público de Información y Relaciones Institucionales

Eduardo Javier Gracida Campos

Dirección General de Administración

Luis María Zapata Ferrer

Contraloría Interna

Manuel Rodríguez Murillo

REALIDAD, DATOS Y ESPACIO REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Editor responsable

Sergio Carrera Riva Palacio

Editor técnico

Gerardo Leyva Parra

Coordinación editorial

Virginia Abrin Batule y Mercedes Pedrosa Islas

Corrección de estilo

José Pablo Covarrubias Ordiales y Laura Elena López Ortiz

Corrección de textos en inglés

Gerardo Piña

Diseño y formación edición impresa

Juan Carlos Martínez Méndez y Eduardo Javier Ramírez Espino

Indizada en: Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal *Latindex Catálogo*; Citas Latinoamericanas en Ciencias Sociales y Humanidades (*CLASE*) y en la Red Iberoamericana de Innovación y Conocimiento (*REDIB*).

REALIDAD, DATOS Y ESPACIO REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA, Vol. 11, Núm. 3, septiembre-diciembre, 2020, es una publicación cuatrimestral editada por el Instituto Nacional de Estadística y Geografía, Avenida Héroe de Nacozari Sur 2301, Fraccionamiento Jardines del Parque, 20276, Aguascalientes, Aguascalientes, Aguascalientes, entre la calle INEGI, Avenida del Lago y Avenida Paseo de las Garzas, México. Teléfono 55 52781069. Toda correspondencia deberá dirigirse al correo: rde@inegi.org.mx

Editor responsable: Sergio Carrera Riva Palacio. Reserva de Derechos al Uso Exclusivo del Título Núm. 04-2012-121909394300-102, ISSN Núm. 2007-2961, ambos otorgados por el Instituto Nacional del Derecho de Autor. Certificado de Licitud de Título y Contenido Núm. 15099, otorgado por la Comisión Calificadora de Publicaciones y Revistas Ilustradas de la Secretaría de Gobernación. Domicilio de la publicación, imprenta y distribución: Avenida Héroe de Nacozari Sur 2301, Fraccionamiento Jardines del Parque, 20276, Aguascalientes, Aguascalientes, Aguascalientes, entre la calle INEGI, Avenida del Lago y Avenida Paseo de las Garzas, México.

El contenido de los artículos, así como sus títulos y, en su caso, fotografías y gráficos utilizados son responsabilidad del autor, lo cual no refleja necesariamente el criterio editorial institucional. Asimismo, la Revista se reserva el derecho de modificar los títulos de los artículos, previo acuerdo con los autores. La mención de empresas o productos específicos en las páginas de la Revista no implica el respaldo por el Instituto Nacional de Estadística y Geografía.

Se permite la reproducción total o parcial del material incluido en la Revista, sujeto a citar la fuente. Esta publicación consta de 400 ejemplares y se terminó de imprimir en enero del 2021.

Versión electrónica: <http://rde.inegi.org.mx>

ISSN 2395-8537



Offline Work

This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

CONSEJO EDITORIAL

Dr. Enrique de Alba Guerra

Presidente del Consejo

Mtra. Claudia Aburto Rancaño

Instituto Tecnológico Autónomo de México
México

Dr. Clemente Ruiz Durán

Universidad Nacional Autónoma de México
México

Dr. Emilio Cunjamá López

Instituto Nacional de Ciencias Penales
México

Dr. Fernando Cortés Cázares

Profesor emérito de FLACSO PUEDE de la UNAM
México

Dra. Graciela Teruel Belismelis

Universidad Iberoamericana
Ciudad de México
México

Dra. Landy Sánchez Peña

El Colegio de México
México

Dra. María Martha Téllez Rojo Solís

Instituto Nacional de Salud Pública
México

Dr. Víctor Manuel Guerrero Guzmán

Instituto Tecnológico Autónomo de México
México

A continuación, se reseña el contenido de los artículos que forman este último número del 2020.

INEGI-Móvil: desarrollo de un SIG móvil en código libre para levantamiento de datos en campo (INEGI-Mobile: Development of a Free Software Mobile GIS for Field Data Surveys) describe el proceso del diseño de una aplicación que consiste en un sistema de información geográfica para dispositivos móviles integrado en una infraestructura de datos espaciales. Su aplicación en el trabajo de campo del Instituto Nacional de Estadística y Geografía (INEGI) pretende dar mayor certeza y agilidad en la captura de la información, aminorar tanto el tiempo para disponer de esta como el de los levantamientos de proyectos, como los censos de población y vivienda, el Directorio Estadístico Nacional de Unidades Económicas, las clasificaciones de uso del suelo y vegetación o incluso la gestión, en tiempo real, de los desastres naturales.

Trabajo de cuidado en las fuentes de información estadística de México (Care Work in the Statistical Sources of Information in Mexico) es una investigación cuyo objetivo inicial fue examinar la cobertura conceptual y la de los grupos de población dependiente, además de la suficiencia de los datos estadísticos para abordar el trabajo de atención de personas. Los resultados señalan que existe un buen cubrimiento conceptual en México, que hay un desbalance en el que corresponde a los grupos dependientes de esta actividad y que las encuestas nacionales sobre Uso del Tiempo y de Ocupación y Empleo del INEGI son insumos que permiten tratar el tema y relacionarlo con la actividad económica de mujeres y hombres.

Ser entrevistadora o entrevistador. La recolección de datos como un oficio, el aprendizaje de un saber hacer (To Be a Female Interviewer or a Male Interviewer. Data Compilation as a Craft, the Learning of a Know-How) es un artículo que pretende poner de manifiesto el trabajo realizado por las encuestadoras y los encuestadores que va más allá de la captación de datos y que, desde la práctica, van construyendo y perfeccionando las tareas para obtener la información requerida, donde ponen en marcha su capacidad de resistencia, adaptación y creatividad para construir estrategias que les permitan hacer frente a lo no previsto.

Estandarización de las áreas geoestadísticas básicas urbanas para 100 ciudades del Sistema Urbano Nacional (Standardization of Basic Urban Geostatistical Areas for 100 Cities of the National Urban System) describe el procedimiento utilizado en el ajuste espacial y la corrección topológica de esas unidades de observación a partir de diversas hipótesis sobre la relación entre forma urbana y productividad laboral (medida como valor agregado censal bruto/trabajador). La idea central planteaba que las ciudades compactas tienden a ser más eficientes que las más dispersas.

Moda, media y mediana de la altura del relieve mexicano (Mode, Mean and the Median of the Height of the Mexican Relief) muestra cómo, a través del estudio de todo el territorio nacional, se determinaron las medidas de tendencia central de las celdas del Continuo de Elevaciones Mexicano 3.0 desarrollado por el INEGI, cuya característica principal es la resolución de 15 metros.

Indicadores de actividad económica municipal 1993-2013 (Municipality-Level Economic Activity Indicators 1993-2013) es un trabajo que da continuidad a la investigación hecha por los autores en el 2019, en la cual los construyeron basados en los Censos Económicos para el periodo 2003-2013; ahora, extendieron la longitud de las series de tiempo hasta 1993 e hicieron mejoras metodológicas en términos de representatividad económica y geográfica; esto permite a los usuarios de información económica contar con series más largas para analizar la evolución de los municipios y, por ende, generar políticas en este tema.

Cubo de datos geoespaciales para el uso de las imágenes satelitales en la generación de información geográfica y estadística (The Use of Satellite Imagery in the Generation of Geographic and Statistical Information) presenta una herramienta mediante la cual se busca aprovechar el potencial de estas para la producción de ese tipo de datos; ofrece una importante posibilidad para hacer eficiente su integración y análisis para un mejor diseño y monitoreo de las políticas públicas.

importinegi: un paquete de R para descargar y gestionar bases de datos del INEGI (*importinegi: an R Package to Download and Manage INEGI's Data Sets*) es

un trabajo que muestra un instrumento que contribuye a facilitar la búsqueda, descarga y gestión de los datos abiertos del Instituto, en particular para su uso en la investigación y la enseñanza. Se describen todas las funciones y se acompañan con ejemplos. Fue desarrollado en un *software* libre y busca apoyar al creciente ecosistema de ciencia abierta para los investigadores.

En el artículo *Análisis de eficiencia en los servicios del agua en México con datos de los Censos Económicos (Analysis of Efficiency of Water Services in Mexico Using Data from the Economic Censuses)* se muestra el potencial que representa la base de datos de ese operativo censal del INEGI para el estudio del desempeño de los organismos encargados del suministro del vital líquido en las localidades del país.

Para concluir este número, se presenta la reseña del libro *Letters to a New Minister of Education. Learning to Lead Education Systems*, titulada *Reformar la educación: ¿misión imposible? Los consejos y experiencias de los líderes educativos (Reforming Education: Impossible Mission? Sharing Advises and Experiences from Educational Leaders)* que, como menciona el autor, "... es un conjunto de lecciones aprendidas pero, sobre todo, un llamado al optimismo en la medida en la que los autores comparten la convicción de que, pese a las resistencias, ofrecer una educación de alta calidad para los niños y jóvenes es, al mismo tiempo, un compromiso moral superior, un bien público...".

<http://rde.inegi.org.mx>

INEGI-Móvil:

desarrollo de un SIG móvil en código libre para levantamiento de datos en campo

INEGI-Mobile: **Development of a Free Software Mobile GIS for Field Data Surveys**

Alonso Morilla Meneses,* Alejandra Albert Tejera* y Silvia Franceschi**

* GeoAlternativa, AC, alonsomorilla@geoalternativa.com, alejandraalbert@geoalternativa.com, respectivamente.

** HydroloGIS, silvia.franceschi@hydrologis.eu



Arturo López Jiménez.

El artículo aborda el proceso de desarrollo de una *app* para el Instituto Nacional de Estadística y Geografía consistente en un sistema de información geográfica para dispositivos móviles integrado en una infraestructura de datos espaciales. Se explica cómo se llevó a cabo el diagnóstico de necesidades del equipo técnico del Instituto, la metodología elegida para desarrollar y testear la aplicación y los resultados obtenidos durante el proyecto. Su desarrollo ha sido liderado por GeoAlternativa, AC gracias al apoyo del Fondo Sectorial CONACYT-INEGI.

Palabras clave: *software* libre; SIG móvil; infraestructura de datos espaciales.

Recibido: 25 de marzo de 2019.
Aceptado: 18 de diciembre de 2019.

1. Introducción

Los entes de la administración pública federal, como el Instituto Nacional de Estadística y Geografía (INEGI), suelen realizar multitud de trabajos en campo y de diversa naturaleza. Actualmente, lo más habitual es que se utilicen métodos analógicos para el registro de información en el terreno, como: peritajes, encuestas, censos, etcétera.

El uso del papel y herramientas analógicas para este tipo de labores puede conllevar a errores humanos durante su cumplimentación, provocando problemas en la futura gestión de la información. Esta metodología obliga a una fase posterior de digitalización, lo cual puede sumar nuevos fallos, registros duplicados y valores erróneos, entre otros, todo ello ocupando recursos humanos y económicos. Además, sería necesario georreferenciar esta información para poder gestionarla de manera óptima, sumando una fase más de trabajo y ralentizando la gestión de los datos por parte de la administración pública.

INEGI-Móvil, la aplicación para dispositivos móviles desarrollada gracias al proyecto 267963 del Fondo Sectorial CONACYT-INEGI, permite obtener datos en el terreno a través de formularios personalizados, editarlos y sincronizarlos con una futura infraestructura de datos espaciales (IDE), práctica-

This article discusses the process of developing an app for the *National Institute of Statistics and Geography* consisting of a geographic information system for mobile devices integrated into a spatial data infrastructure. It explains how the needs assesment of the Institute's technical team was carried out, the methodology chosen to develop and test the application and the results obtained during the project. Its development has been led by GeoAlternativa, AC thanks to the support of the CONACYT-INEGI Sectorial Fund.

Key words: free software; mobile GIS; Spatial Data Infrastructure.

mente en tiempo real. De esta forma, conseguimos un ahorro considerable en tiempo y recursos económicos.

Por lo tanto, además del desarrollo de la aplicación móvil que se solicitaba en este proyecto, GeoAlternativa, AC¹ propuso conectar esta herramienta con un sistema de información geográfica (SIG) de escritorio y uno web, creando una *suite* completa que pueda gestionar toda la información geográfica desde un punto de vista interoperable (ver figura 1).

Las aplicaciones prácticas más notables en las que se puede utilizar esta *suite* son:

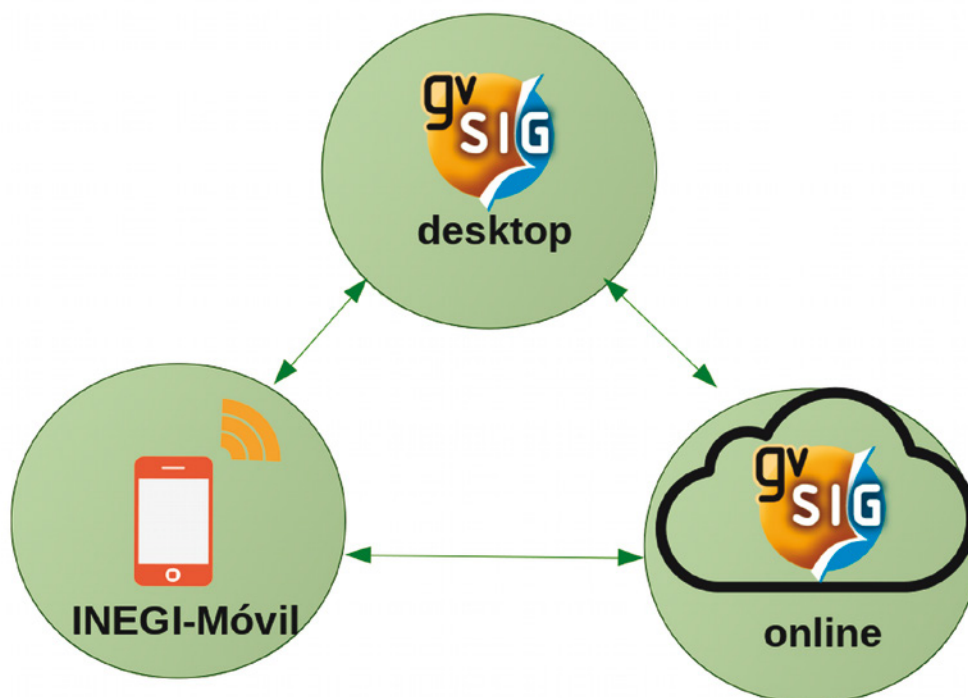
- Gestión tanto de infraestructuras urbanas y rurales como del catastro.
- Censos y peritajes.
- Encuestas y estudios sociales.
- Gestión de recursos naturales, del patrimonio histórico y arqueológico, así como de emergencias y actuaciones de protección civil.

Este proyecto ha sido liderado por GeoAlternativa. En el proceso han apoyado en el desarrollo de sof-

¹ Organización de la sociedad civil especializada en geomática, educación y ordenamiento ambiental que forma parte de la Asociación gvSIG.

Figura 1

Esquema de la propuesta de *suite* interoperable



ware otros miembros de la Asociación gvSIG, como HydroloGIS o Scolab y parte del equipo técnico del INEGI, ayudando en el diagnóstico de necesidades y en mejorar la experiencia de usuario.

2. Fase 1: diagnóstico de necesidades de los usuarios

2.1 Métodos y técnicas

Para llevar a cabo esta fase, se tuvieron en cuenta tanto los requerimientos propuestos en la convocatoria del Fondo Sectorial como los particulares del grupo que sirvió de enlace entre el INEGI y GeoAlternativa.

Además, se realizó un estudio exhaustivo de todas las aplicaciones SIG móviles libres del mercado para revisar las funcionalidades, opciones y librerías que mejor se podían ajustar a estas necesidades.

2.2 Resultados

Gracias a la fase de diagnóstico de requerimientos, se observó que el equipo técnico del INEGI encargado de recopilar datos en campo necesita superar el uso de métodos analógicos tradicionales, ya que se han convertido en un factor limitante en el desarrollo de sus proyectos. La necesidad de manipular diferentes instrumentos (como papel y lápiz, además de otros materiales, por ejemplo, brújulas, planos cartográficos y agenda de notas, entre otros) complican esta labor. Si le sumamos la incorporación de aparatos tecnológicos (dispositivos con GPS o sensores de variables biométricas), la atención y calidad de su trabajo pueden verse seriamente mermadas.

En esta convocatoria se planteó la implementación de herramientas de SIG móviles para dispositivos (como tabletas o *smartphones*) que permitieran

adquirir, editar y almacenar datos georreferenciados. Además, se aconsejaba la comunicación con un servidor que permitiera compartir, visualizar y administrar la información geográfica en un ambiente distribuido a nivel de intranet y/o internet, de cara a la posible incorporación del SIG móvil a una futura IDE. Por medio de este canal, por lo tanto, se podrá enviar información desde el dispositivo móvil hacia el servidor, en el cual un cliente usuario podrá hacer consultas a bases de datos actualizadas.

Durante esta fase de diagnóstico también se tomaron en cuenta aspectos secundarios, pero asimismo muy importantes de cara a asegurar su uso generalizado en la administración pública federal, como: la compatibilidad con estándares académicos y comerciales, la disponibilidad de cartografía digital actualizada y la recomendación de reducir costos asociados a licencias de *software* propietario. Por eso, desde un inicio, se tomó la decisión de desarrollar herramientas basadas en *software* libre y apegadas a estándares OGC (*Open Geospatial Consortium*).

En las diferentes reuniones presenciales y virtuales con el equipo del INEGI en Aguascalientes se desarrollaron y complementaron las necesidades citadas más arriba e incorporando otras de carácter práctico con el objetivo de mejorar la experiencia del usuario y el desempeño de la aplicación en los trabajos de campo.

Estas nuevas necesidades podemos agruparlas en diferentes rubros:

- Técnicas:
 - ✓ Trabajo con formularios personalizables. Uno de los mayores obstáculos a la hora de tomar datos con un dispositivo móvil es la introducción de texto a través de teclado que, además, puede conllevar a gazapos gramaticales. Esto puede producir fallos graves en los registros de las bases de datos. Si creamos formularios personalizados con respuestas preconfiguradas ganaremos tiempo y facilidad de uso, además de que evitaremos posibles errores humanos.

- ✓ Combinar fotografías con información georreferenciada. Dependiendo del tipo de levantamiento de datos que estemos llevando a cabo, puede ser necesario incorporar a nuestro registro una imagen que ilustre mejor la información.
- ✓ Posibilidad de más de una fotografía georreferenciada por registro. Especialmente necesario para censos de viviendas, inventarios de flora y otros trabajos similares.
- ✓ Importación de WMS e imágenes *raster*. La disponibilidad de servicios web del INEGI y otras imágenes *raster* en la aplicación SIG móvil permitirá mejorar el contexto y, por lo tanto, la precisión en los registros.
- ✓ Diseño intuitivo y alta usabilidad. Por lo normal, los equipos humanos que levantan datos en campo no desean una curva de aprendizaje larga, ya que es necesario primar la efectividad a corto plazo. Por eso, la interfaz de INEGI-Móvil es simple y puede utilizarse correctamente desde el primer momento.
- Logísticas:
 - ✓ Sustituir los formularios de papel por digitales. Como se mencionaba con anterioridad, el proceso de digitalización de la información analógica supone un esfuerzo y un riesgo real de errores de transcripción, por lo que introducir los datos directamente en formato digital mejora mucho el proceso.
 - ✓ Disminuir los tiempos de postprocesado de la información de campo. De la misma manera, al sustituir los formularios de papel por digitales, el tiempo de postprocesado prácticamente desaparece, optimizando el proceso completo.
- Administrativas:
 - ✓ Exportación de formularios de formato SIG a PDF o HTML para su consulta y revisión por personal no especializado en SIG. No todos en el equipo que revisa los datos obtenidos en campo son especialistas en SIG o tienen nociones de geomática, por lo que es necesario que la información obtenida pueda ser vista y comprendida por todo el mundo en la toma de decisiones.

2.3 Comparativa de soluciones SIG libres

Entre las muchas aplicaciones para dispositivos móviles revisadas (ver figura 2), finalmente se analizaron en profundidad las *app Qfield*, *gvSIG Mobile* (para Windows Mobile), *gvSIG Mini* y *Geopaparazzi*, es decir, las relativas a los dos proyectos más importantes de SIG libre: *QGIS* y *gvSIG*. La razón de elegir las, además de que son las más completas, es por ser muy estables y soportadas por una comunidad de usuarios mundial, lo cual otorga fiabilidad a sus propuestas.

La comparativa que se realizó de todas las soluciones SIG libres tuvo en cuenta, sobre todo, que cumplieran con las siguientes características y funcionalidades:

- Licencia libre.
- Sujetas a estándares OGC.
- Visualización de archivos *raster* y vectoriales.
- Visualización de servicios OGC.
- Conexión a BBDD.

- Herramientas tanto de geoposicionamiento, navegación y consulta como de edición de información vectorial.
- Toma de fotografías georreferenciadas.
- Trabajo en modalidad *offline* y *online*.
- Importación y exportación de la información desde diferentes formatos.
- Trabajo con formularios.

De todas ellas, *Geopaparazzi* y *Qfield* fueron las que cumplieron con la mayoría de los requisitos y, por lo tanto, objeto de un testeo en profundidad para observar las mejores funcionalidades y librerías de código susceptibles de desarrollarse para INEGI-Móvil.

Las que se probaron para el trabajo de campo se dividieron en tres bloques:

- Preparación del proyecto de campo:
 - ✓ Carga de información base (*raster*).
 - ✓ Importación de información vectorial.
 - ✓ Carga de un mismo proyecto en varios dispositivos.

Figura 2

Muestra de las aplicaciones de SIG móvil analizadas de cara a establecer la base del futuro desarrollo de INEGI-Móvil



- Desempeño en campo:
 - √ Tiempo de captación de satélites y de plena operatividad.
 - √ Carga del proyecto.
 - √ Grabación de *track* y *waypoint*.
 - √ Consulta de la información vectorial.
 - √ Creación de nuevas capas vectoriales.
 - √ Edición de capas vectoriales existentes.
 - √ Inclusión de fotografías al proyecto.
 - √ Integración de nueva información a la tabla de atributos.
 - √ Trabajo con formularios.
 - √ Sincronización con 3G/4G y wifi.
- Interoperabilidad:
 - √ Visualización de la información obtenida por otros usuarios.
 - √ Descarga de la información a SIG de escritorio.
 - √ Visualización de la información obtenida en la IDE.
 - √ Precisión de los datos obtenidos en campo.

2.4 Propuesta de GeoAlternativa

Con toda esta información previa y analizando las funcionalidades que mejor desempeño obtenían en el trabajo de campo, se sugirió un SIG móvil integrable en una futura IDE basada en *software* libre y estándares OGC/ISO garantizando tanto la independencia tecnológica de la solución como su interoperabilidad con el resto de la *suite* recomendada.

La aplicación se desarrolló en Java para Android, el sistema operativo más extendido en los dispositivos móviles, lo cual posibilita una mayor implantación en los entes de la administración pública federal.

Dadas las características de uso de este tipo de herramientas, que principalmente se orientan a las actividades de toma de datos en campo, la aplicación trabajaría tanto en modalidad *online* como *offline*. En este último caso se garantiza su operatividad sin acceso a datos de internet para, en el momento de conseguir conexión, alimentar una base de datos.

Al formar parte de una plataforma o *suite* integral, no solo existiría comunicación con un servidor, sino también se permitiría que los ficheros de proyecto del SIG móvil pudieran leerse automáticamente por uno web libre y visualizarse también en otro de escritorio. En ambos casos, el resto de componentes reconocerían de manera automática los distintos datos tomados en campo (elementos vectoriales de capas de puntos, líneas o polígonos, fotografías, notas, etcétera).

El sistema, para las tareas de geoposicionamiento, utilizaría el GPS interno del dispositivo, que funcionaría no solo para permitir la ubicación del usuario, sino también para poder editar e introducir nuevos datos vectoriales.

La solución móvil se pensó también para soportar, además de servicios estándar WMS o WMTS (servicios web de información *raster*), bases de datos de teselas como *Mapforge* o *MBTiles*. Esto garantiza que en zonas sin señal de datos de internet no se pierda el acceso a la cartografía necesaria para el trabajo de campo.

También se eligió, para almacenar la información geográfica vectorial, la base de datos *SpatialLite*, lo cual significa que puede soportar todo tipo de información (geometrías, datos alfanuméricos, etc.) con un peso mucho menor y con la posibilidad de guardar, en un mismo archivo, multitud de capas vectoriales.

Además, se propuso implementar características de importación y exportación a distintos formatos, como KMZ y GPX, así como la importación y exportación de datos en tiempo real a un SIG web (*gvSIG Online*) y de proyectos a uno de escritorio libre (*gvSIG Desktop*). Se garantizaría, de este modo, la posibilidad de distribuir de múltiples formas la información y en todas direcciones (hacia y desde la aplicación SIG móvil).

El equipo técnico del INEGI consideró una parte fundamental del proyecto —pensando en el trabajo de campo— el desarrollar la capacidad de la herramienta para diseñar formularios de forma ágil y sencilla. Con una interfaz clara y simple se podrían

generar aquellos que permitieran la toma de datos tanto gráficos como alfanuméricos. Durante la fase de diagnóstico se realizaron varias reuniones con el equipo técnico para anotar sus necesidades en este sentido y se realizó una salida a campo para probar dicha funcionalidad.

3. Fase 2: desarrollo de la solución

Una vez recopiladas las necesidades del Fondo y del equipo técnico que fungió como enlace con GeoAlternativa se procedió a elegir la base de desarrollo, las funcionalidades requeridas y el plan de trabajo para conseguir una herramienta SIG móvil robusta, sencilla de usar e interoperable con el resto de elementos de la *suite*.

3.1 Métodos y técnicas

La metodología que se propuso para acometer este proyecto fue la de Desarrollo Ágil, concretamente la de Programación Extrema, que otorga una importancia crucial a la interacción entre el equipo de desarrollo y el usuario final. De esta manera, siempre se obtendrá una solución informática a gusto del usuario y se salvan los posibles errores casi de inmediato.

Concretamente, la de Programación Extrema propone los siguientes principios:

- Desarrollo iterativo e incremental. Pequeñas mejoras, unas tras otras. En lugar de escribir grandes cantidades de código se desarrollan pequeñas partes del *software* para eliminar cuellos de botella de posibles errores.
- Pruebas unitarias continuas. Al mismo tiempo, cada desarrollo se testea continuamente, tanto en entorno virtual como en terreno, para localizar errores y *bugs* de una manera más rápida y tomar las medidas para solucionarlos.
- Integración usuario final/equipo de desarrollo. Para que el usuario final sepa en todo momento qué producto se está desarrollando y que la solución informática servirá a sus necesidades, es necesario que la comunica-

ción sea constante desde el inicio del proyecto. De esta forma, la retroalimentación efectiva estará asegurada.

- Corrección de errores antes de una nueva funcionalidad. El siguiente paso solo se realiza cuando los errores detectados hayan sido corregidos por el equipo de desarrollo.
- Entregas frecuentes. Se publican compilaciones de manera constante para que el equipo de usuarios finales pueda probar la solución y verificar, de la manera más real posible, el funcionamiento de la aplicación.
- La simplicidad conlleva transparencia. Explicar el trabajo que se está realizando con un lenguaje complicado solo provoca falta de entendimiento y problemas en la retroalimentación; por eso, se pretende mostrar el proyecto de manera simple para que cualquiera pueda entender en qué se están destinando los recursos.

Además de los beneficios que aporta esta metodología en un proyecto de este tipo, también se han aprovechado los que proporcionan otras herramientas basadas en *software* libre que nos han permitido mantener la comunicación y transparencia en todo momento con el cliente:

- Uso de repositorios. Se utilizó el servicio de *Git* para almacenar todas las compilaciones que se fueron desarrollando durante el proyecto. De esta forma, el código y la aplicación estuvieron disponibles en todo momento para revisión, uso y mejora. Para optimizar este trabajo de guardar versiones, se usó el motor de integración continua Jenkins.
- Virtualización. Antes del testeo en campo o en escritorio de los tres elementos de la *suite*, se probaron las aplicaciones en entornos virtualizados a través de *Android Studio* y *Proxmox*.
- Control y seguimiento. Como en el resto del proceso, para las labores de control y seguimiento también se emplearon herramientas *FLOSS* (por *Free/Libre and Open Source Software*). En este caso se eligió *GitHub*, una

plataforma de desarrollo colaborativo que utiliza el control de versiones *Git*. Además, se utilizó *Collabtive* como *software* de gestor de proyectos y asignación de tareas.

INEGI-Móvil se desarrolló completamente en lenguaje Java. Las únicas librerías nativas en C referenciadas en el proyecto son las de las bases de datos *Spatialite*. El proyecto usa *Gradle* como herramienta de compilación.

Las librerías más importantes de este proyecto son:

- *Inegimovil Core*. Es el módulo central (core), y contiene todo lo que es visible en la *app* móvil, es decir, todas las actividades, vistas y fragmentos, así como la lógica que hay detrás de la interacción con el usuario.
- *Inegimovil library*. El módulo *Biblioteca* contiene el código de funcionalidades que

puede ser replicado en otras aplicaciones desarrolladas en un futuro.

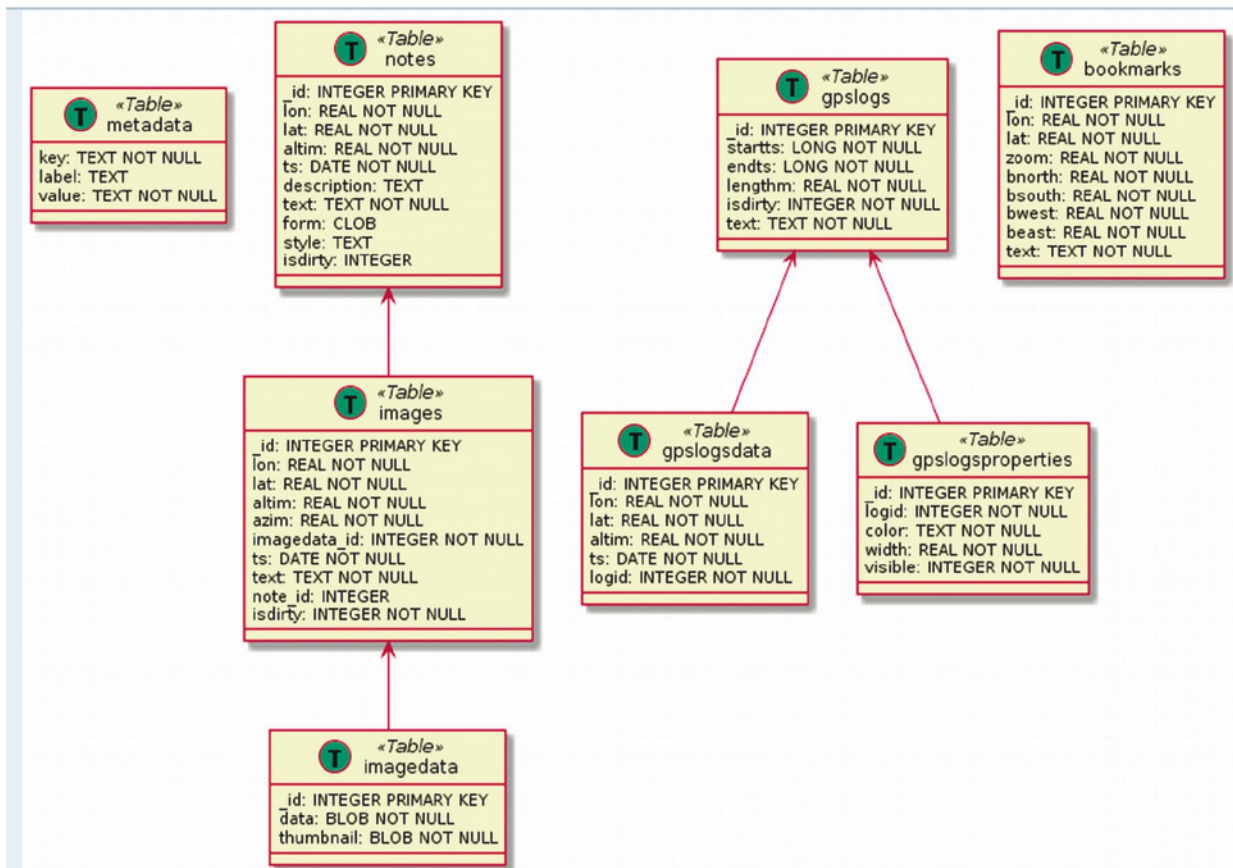
- *Inegimovil spatialite library*. El módulo *Spatialite Library* contiene elementos clave para el correcto funcionamiento de la aplicación principal. Como el propio nombre indica, contiene todas las clases que definen las tablas espaciales, sus geometrías y atributos, es decir, gestiona la parte vectorial de *Spatialite*. La base de datos del proyecto INEGI Móvil es una *sqlite* simple con la extensión *.gpap* (ver figura 3).

3.2 Resultados

Desde el punto de vista funcional y estético, INEGI-Móvil se ha desarrollado para ser una aplicación muy sencilla de utilizar y de una interfaz limpia y simple,

Figura 3

Esquema de la base de datos de los proyectos en INEGI-Móvil



diseñada en especial para el levantamiento de datos en campo con dispositivos móviles (ver figura 4).

Principalmente, consta de una pantalla inicial para acceder a todas las funcionalidades, una vista de mapa donde visualizar la capa base y los datos de trabajo, así como apartados para importar y exportar información geográfica.

La pantalla principal consiste en los accesos directos de las diferentes funcionalidades, y se diseñó predominando el espacio diáfano y sencillo para

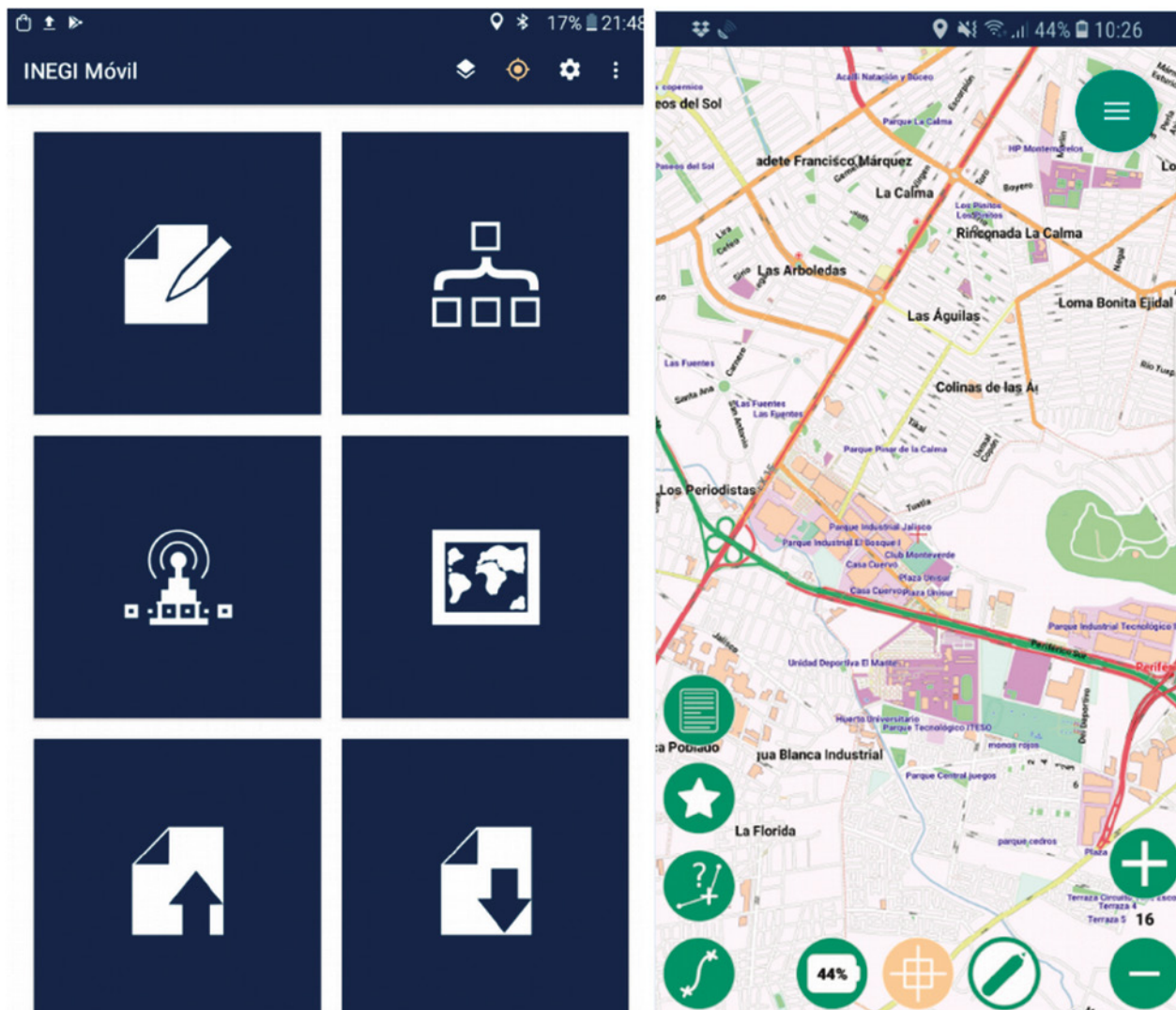
necesitar poca precisión a la hora de pulsar las opciones, algo muy valioso cuando se trabaja en el terreno. La vista de mapa también intenta prescindir de múltiples accesos para evitar errores durante el trabajo en campo.

Funcionalidades desarrolladas para INEGI-Móvil

Además del núcleo de la aplicación y sus funcionalidades más básicas, se atendieron las recomen-

Figura 4

Interfaz principal de la aplicación INEGI-Móvil



A la izquierda se observa la pantalla principal con los accesos a las diferentes funcionalidades y a la derecha, la vista de mapa.

daciones y peticiones obtenidas durante la fase de diagnóstico, desarrollando también las siguientes funcionalidades:

- Posicionamiento por GPS. INEGI-Móvil utiliza la tecnología GPS para ubicar a la persona usuaria sobre el terreno y grabar el *track* de recorrido durante el proceso de trabajo en campo. De esta manera se puede monitorear la labor de cada técnico y planear los puntos de actuación. Además, es posible importar archivos GPX (formato de intercambio GPS) a la *app*, por lo que esta planeación puede realizarse con antelación y cargar cuando sea necesario.

Esta funcionalidad también permitirá georreferenciar los datos captados en el terreno.

- Obtención de fotografías, notas y dibujos georreferenciados. INEGI-Móvil nos permite tomar información georreferenciada a través de la funcionalidad *Notas*. Por defecto, podemos tomar notas de imagen, dibujo, texto y de mapa.

- Obtención de información geográfica en campo a través de formularios personalizados. Además de estas notas simples, es posible generar formularios complejos similares a los oficiales de la administración pública federal, como el INEGI. Pueden ser pequeños formularios de incidencias o extensos de varias páginas (ver figura 5).

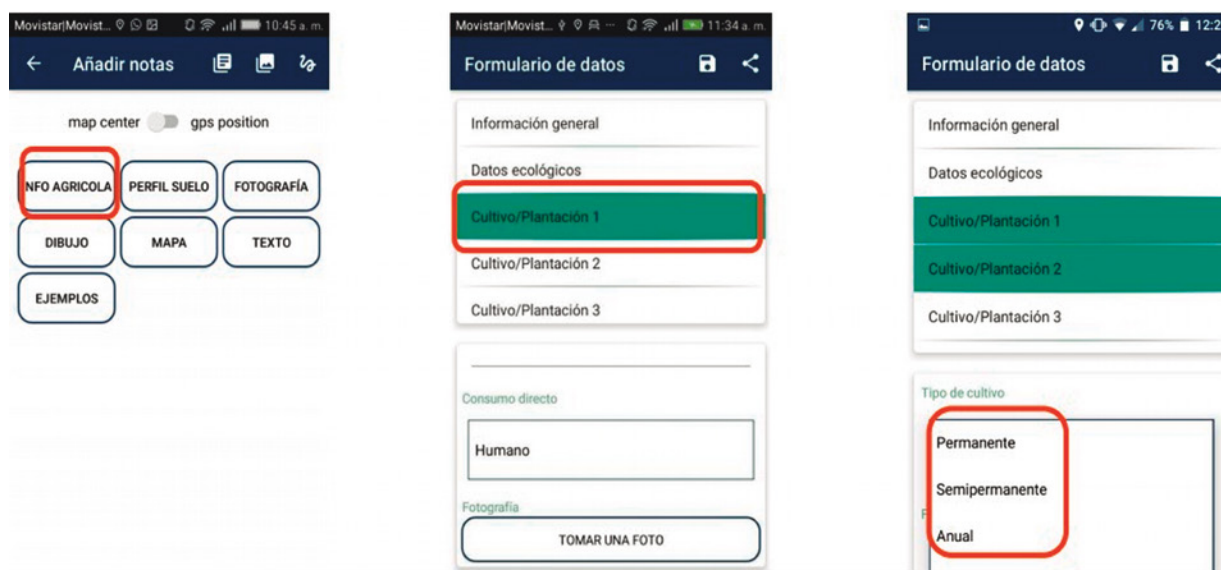
Una de las principales ventajas de los personalizados en INEGI-Móvil es que se pueden preconfigurar las respuestas para cada campo a cumplimentar, lo cual ahorra tiempo y evita errores humanos en la introducción de datos.

Además, es posible incorporar fotografías a estos formularios, algo muy valioso para la gestión de inventarios y censos.

- Utilización de mapas base (*OSM*, *WMS*, *Mapforge*, etc.). Por defecto, INEGI-Móvil tiene como mapa base la cartografía en línea del proyecto *OpenStreetMap*, muy completa y funcional. Sin embargo, en ocasiones, es necesario disponer de otras capas base propias acordes con el ámbito en el que se trabajará o que pueda utilizarse en modo *offline*.

Figura 5

Formularios personalizados en INEGI-Móvil para realizar un inventario agrícola



Permite incorporar nuevas capas base, como ortofotografías para operativos de censos, catastrales o agrarios, LIDAR para trabajos forestales, servicios WMS, etcétera.

- Funcionamiento *online* y *offline*. INEGI-Móvil nos permite trabajar utilizando wifi o redes 3G/4G para cargar mapas base (*OpenStreetMap*, *WMS*) y sincronizar los datos que se vayan obteniendo en campo con una base de datos remota o un SIG en la nube como *gvSIG Online* (ver figura 6).

Sin embargo, hay ocasiones en las que no es posible tener una conexión a internet por estar en espacios sin señal; *gvSIG* puede, en estos casos, seguir trabajando en el levantamiento de datos en campo y, al recuperar conexión, sincronizar con dichos servicios externos.

- Gestión de información vectorial. INEGI-Móvil no solo está orientado al levantamiento de datos en campo a través de formularios, sino que es un SIG completo que puede gestionar

información vectorial desde el propio dispositivo móvil.

Sus herramientas de edición permiten crear, cargar, modificar y guardar capas de puntos, líneas y polígonos, además de cambiar su información alfanumérica (ver figura 7).

- Botón de pánico. Ya que el trabajo en campo del equipo técnico del INEGI no está exento de riesgos, INEGI-Móvil cuenta con esta opción que, con una sola pulsación, enviará automáticamente un SMS personalizado con las coordenadas de ubicación al número de teléfono que se haya predefinido para que se pueda enviar ayuda inmediata.
- Opciones de importación y exportación. Como ya se ha mencionado anteriormente, uno de los puntos fuertes de esta aplicación de SIG móvil es su interoperabilidad y versatilidad.

Para preparar el trabajo en campo, se pueden importar datos de diferentes tipos. De la misma manera, podremos exportar la infor-

Figura 6

Interfaz de *gvSIG Online* representando los puntos donde se ha realizado el trabajo de campo para testear la aplicación móvil

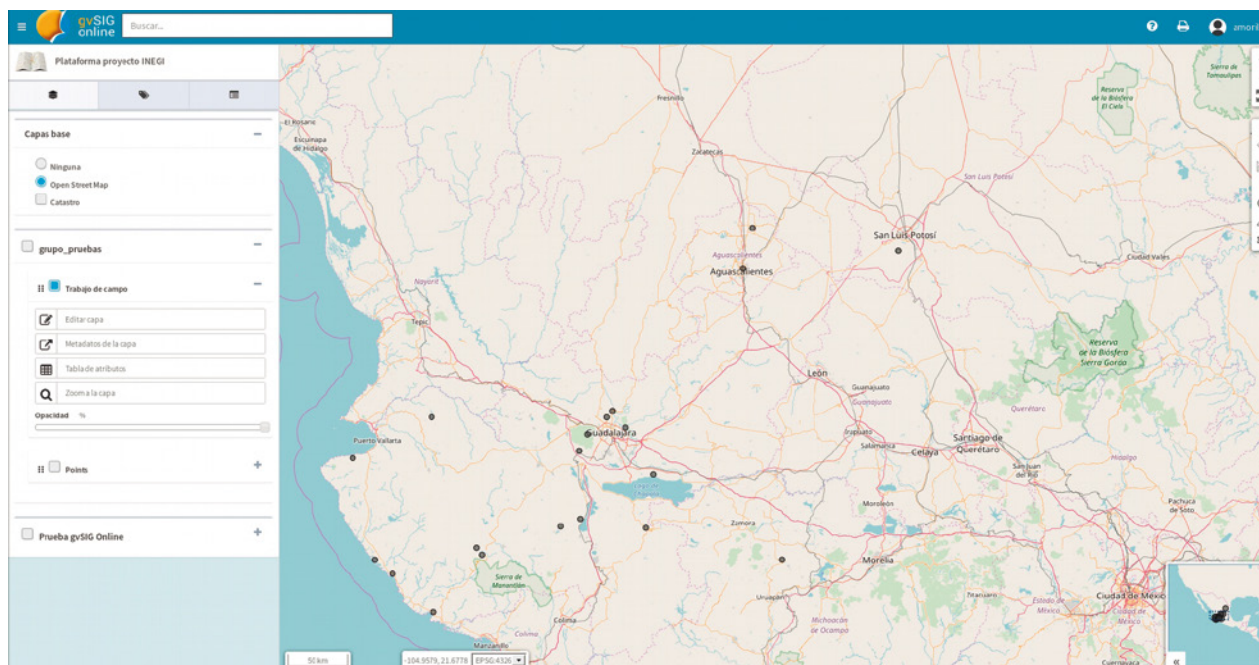
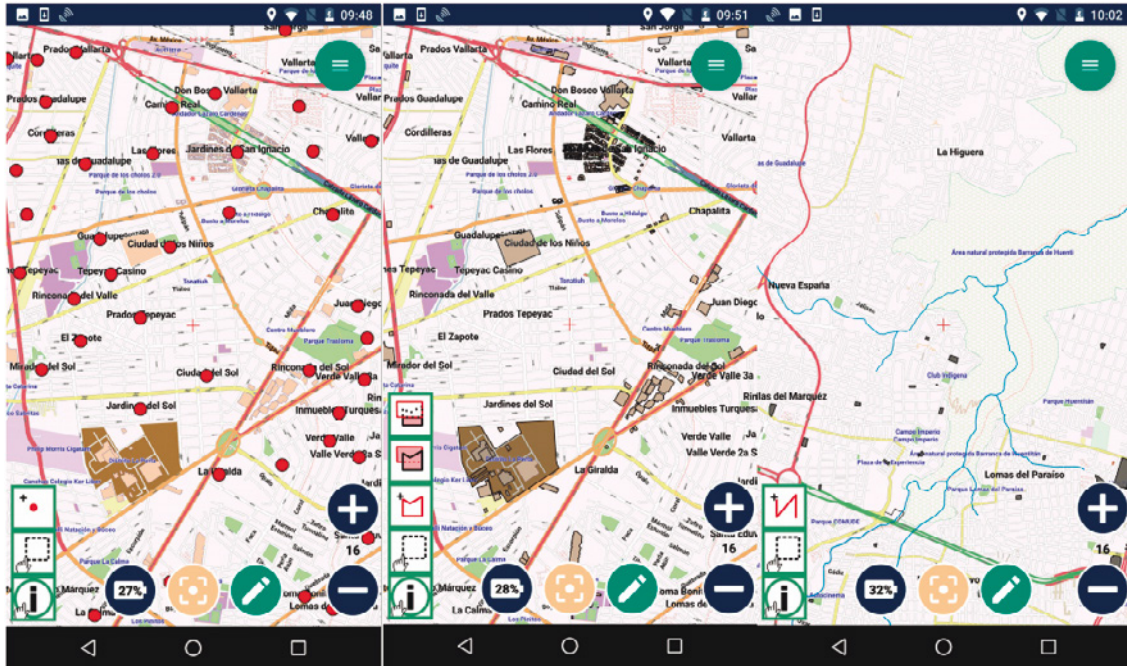


Figura 7

Modo de edición vectorial para puntos, polígonos y líneas en INEGI-Móvil



mación generada en campo a diferentes formatos. Sin duda, lo más destacado en este sentido es la opción de importar capas directamente desde *gvSIG Online*, trabajar con ellas en campo y sincronizarlas de nuevo con la base de datos remota de la IDE (ver figura 8).

Además, para mejorar la experiencia de usuario, y a petición del equipo técnico del INEGI, se desarrolló la funcionalidad de exportar los formularios de campo a formato PDF para poder visualizar la información obtenida en un entorno no SIG, de cara a que personas

Figura 8

Accesos a los diferentes tipos de importación y exportación de datos



que no son especialistas en geomática puedan revisar los datos y trabajar con ellos (ver figura 9).

4. Fase 3: testeo de la solución

4.1 Métodos y técnicas

Estrategia de pruebas de INEGI-Móvil

Estuvo alineada, al igual que todo el proceso de desarrollo, con los principios de las metodologías ágiles, específicamente de la *xTreme Programming* o Programación Extrema; esta considera el proceso de pruebas como indispensable para que la liberación de cada versión de la aplicación sea de calidad.

Con el auge de las metodologías ágiles, las estrategias para abordar estos procesos se han diversificado, y han ganado popularidad aquellas como Desarrollo Dirigido por Pruebas (TDD o *Test Driven Development*) y Desarrollo Dirigido por Pruebas de Aceptación (ATDD o *Acceptance Test Driven Development*). Para el caso de la aplicación INEGI-Móvil, se optó por seguir como línea principal para el proceso de pruebas la de ATDD por las siguientes razones:

- Toma como principal referencia a la persona usuaria final, en este caso el Instituto Nacional de Estadística y Geografía.
- Utiliza historias de usuario como requisitos a los que se asocian pruebas de aceptación (escenarios).
- Las pruebas de aceptación dirigen el diseño/ desarrollo del sistema.
- Requieren entrar tempranamente en detalles de implementación e instanciación de datos de pruebas.

En la estrategia ATDD, las necesidades de las personas usuarias quedan registradas a través de las historias de usuario, y cada una de estas tiene asociados determinados criterios de aceptación; a su vez, para cada criterio de aceptación, se di-


Figura 9

Ejemplo de formulario obtenido en INEGI-Móvil exportado a formato PDF

1. image note

Timestamp	Tue Jul 25 11:30:07 CDT 2017
Latitude	20.76423186
Longitude	-103.41610961

1.1. image note

description	Incendio 1
<p>pictures of the environment around the note</p>	 <p>IMG_20170725_163018.jpg</p>

señan las respectivas pruebas o escenarios que tienen que aprobarse. Basándose en estas pruebas, los desarrolladores programan el código fuente que permitirá cumplir con los requisitos estipulados. Por su lado, ellos definen las pruebas unitarias que evalúan el cumplimiento de los requisitos (expresados en términos de criterios de aceptación).

Antes de las pruebas de validación junto con el equipo técnico del INEGI, se llevó a cabo un proceso continuo e intensivo de pruebas de las funcionalidades de la *app* —a cargo de GeoAlternativa— en

distintos escenarios para pulir errores e identificar problemas de usabilidad. Este proceso se sistematizó en una serie de fichas denominadas *Ficha de prueba en campo*. Cuando se consideró que una versión de la *app* era funcional y se encontraba en una condición aceptable de acuerdo con los términos acordados, se procedió a realizar la prueba definitiva de campo junto con el INEGI.

Por lo tanto, la estrategia siguió este flujo de trabajo para cada versión de la aplicación liberada:

1. Diagnóstico de las necesidades de usuarios (historias de usuario).
2. Definición de los criterios de aceptación (diseño de las pruebas de aceptación).
3. Desarrollo de la aplicación de acuerdo con las historias de usuarios (programación del código fuente).
4. Pruebas continuas y paralelas al desarrollo de la aplicación (fichas de campo).
5. Liberación de la versión.
6. Pruebas de validación de la versión.
7. Retroalimentación al equipo desarrollador.

Cuando se llega al punto 7 se empieza de nuevo con el 1 para la siguiente versión, adaptando y modificando cada uno de los pasos de acuerdo con los requisitos y necesidades de la persona usuaria final identificados durante todo el proceso y siguiendo el enfoque de la metodología ágil.

Criterios de aceptación de pruebas

Estos definen si una cierta funcionalidad registrada en cada historia de usuario cumple sus objetivos o no. Concretamente, en el contexto de las metodologías ágiles, se les define como un conjunto de sentencias redactadas de tal manera que conduzca a una respuesta clara de *aceptado/rechazado*.

En los criterios se deben incluir tanto cuestiones funcionales como no funcionales, y es recomendable evitar aspectos técnicos o de implementación. Una manera de lograr esto es expresar criterios que denoten una intención y no una solución.

4.2 Resultados

La aplicación de la metodología de pruebas y validación tuvo como resultado una fase muy importante de testeo en campo, intentando reproducir, de la forma más fiel posible, los diferentes escenarios reales que se encontrarían los usuarios finales durante su labor de levantamiento de datos en el terreno.

Debido a las necesidades particulares del equipo técnico del INEGI, el testeo se realizó en las condiciones más exigentes en cuanto a duración, rendimiento del sistema, almacenamiento y localización, eligiendo ubicaciones geográficas con poca o nula señal de telefonía móvil, internet e, incluso, señal de satélites GPS. De esta manera se pudo analizar en gabinete si los datos obtenidos cumplían los requerimientos de fiabilidad y precisión necesarios para este proyecto.

Se hicieron un total de 50 salidas a diferentes escenarios de la República que tuvieron como objetivo el testeo del desempeño y las funcionalidades de la aplicación divididas en seis grupos principales:

- Testeo y comparativa *Geopaparazzi* y *Qfield*.
- Testeo y grabación de *tracks*.
- Testeo de las primeras versiones INEGI-Móvil.
- Testeo de formularios personalizados.
- Sincronización de datos con *gvSIG Online*.
- Testeo completo de la aplicación.

5. Fase 4: capacitación y transferencia de conocimiento

Aunque la transmisión comenzó desde el inicio a través de la publicación de versiones de INEGI-Móvil en *GitHub* y mediante reuniones y seguimiento continuo a lo largo de todo el proyecto, durante esta etapa también se hizo del conocimiento a parte del personal técnico del INEGI.

La capacitación consistió en la impartición de un taller de varios días para usuarios finales de las

diferentes aplicaciones de la *suite* y también para las personas encargadas de la administración del sistema.

La idea fue que el equipo del Instituto Nacional de Estadística y Geografía lograra la independencia tecnológica necesaria para poder usar y modificar, de acuerdo con sus intereses, las diferentes aplicaciones utilizadas en este proyecto.

El taller tuvo lugar los días 15, 16 y 17 de noviembre de 2017 en las instalaciones del INEGI en Aguascalientes, donde participaron, aproximadamente, 20 personas cada día (ver fotografía).

6. Conclusión

Este proyecto, financiado por el Fondo Sectorial CONACYT-INEGI, ha permitido el desarrollo de una aplicación de SIG móvil para dispositivos Android y la posibilidad de integración, en un futuro, en una posible infraestructura de datos espaciales con herramientas que permiten a esta *app* interoperar con *gvSIG Desktop* y *gvSIG Online*, además de con cualquier SIG de escritorio o de la nube que esté sujeto a estándares OGC.

Gracias a esta, el equipo técnico del INEGI, y cualquier otro ente público, podrá llevar a cabo



Proceso de capacitación al personal técnico del INEGI.

levantamientos de datos en campo de una manera más fácil y cómoda que con herramientas analógicas. Estos resultados se podrán enviar a una base de datos en tiempo real, resultando un proceso mucho más rápido y fiable, eliminando los tradicionales errores humanos ocurridos durante los procesos de toma de datos y de digitalización.

Además, el hecho de que todo este proceso se realice en un entorno SIG, permitirá aprovechar todo el potencial de visualización, edición y análisis de los sistemas de información geográfica.

Las posibilidades que se abren al incorporar esta herramienta al trabajo en campo del INEGI son muy amplias: permitirá agilizar procesos tan complejos como los censos de población y vivienda, el Directorio Estadístico Nacional de Unidades Económicas, las clasificaciones de uso del suelo y vegetación o incluso la gestión, en tiempo real, de desastres naturales.

Fuentes

De Donatis, M.; A. Antonello; L. Lanteri; S. Susini; M. Foi; y A. R. P. A. . "BeeGIS: a new open source and multiplatform mobile GIS", en: *Digital Mapping*

Techniques '09-Workshop proceedings. US Geological Survey Open-File Report, 2010-1335. Piemonte, Torino-Italy, 2010, pp. 241-246 (DE) <http://pubs.usgs.gov/of/2010/1335/>, consultado el 06/03/2017.

Bernabé-Poveda, M. A. y C. M. López-Vázquez. *Fundamentos de las infraestructuras de datos espaciales*. Serie Científica. España, UPM-Press, 2012, pp. 179-224.

Clark, J. *Location Gathering: An Evaluation of Smartphone-Based Geographic Mobile Field Data Collection Hardware and Applications*. San José State University, 2015, pp. 18-32.

Yagüe, Agustín y Juan Garbajosa. "Las pruebas en metodologías ágiles y convencionales: papeles diferentes", en: *Actas de los Talleres de las Jornadas de Ingeniería del Software y Bases de Datos*. 3(4). España, Universidad Politécnica de Madrid, 2009, pp. 67-73.

Downs, G. "Learn-agile acceptance test-driven development", en: *ACM SIGSOFT Software Engineering Notes*. 36(4), 34 (DE) https://www.researchgate.net/publication/220631270_Lean-agile_acceptance_test-driven_development_better_software_through_collaboration_by_Ken_Pugh, consultado el 13/04/2017.

Jeffries, R. *Essential XP: Card, Conversation, Confirmation*. 30 de agosto de 2001 (DE) <http://xprogramming.com/articles/expcardconversationconfirmation/>, consultado el 07/07/2017.

Auer, K. and R. Miller. *Extreme Programming Applied. Playing to win*. USA, Addison-WesleyProfessional, 2002, pp. 51-83.

Gaikwad, V. and P. Joeg. "An Empirical Study of Writing Effective User Stories", en: *International Journal of Software Engineering and Its Applications*. 10(1). South Korea, Science and Engineering Research Support Society, 2016, pp. 387-404.

Trabajo de cuidado en las fuentes de información estadística de México

Care Work

in the Statistical Sources of Information in Mexico

César González González, Karina Orozco Rocha, Mireya Patricia Arias Soto y Ma. Gregoria Carvajal Santillán*

* Universidad de Colima, cgonzalez31@uclm.mx, korozco9@uclm.mx, mireya_arias@uclm.mx y margre@uclm.mx, respectivamente.



Mexicana patio de colegio/cnibcr/Getty Images

El tema del cuidado ha adquirido importancia académica, política y social. Hay un consenso de que este es un derecho, pero no es claro de quién es la responsabilidad de otorgarlo. Se han desarrollado metodologías y generado información estadística —entre las cuales las encuestas del uso del tiempo han sido clave— para estudiarlo, así como los temas relacionados. El objetivo de esta investigación fue examinar la cobertura conceptual y la de los grupos dependientes, además de la suficiencia de los datos estadísticos para abordar el trabajo de cuidado. Los resultados señalan que existe una buena cobertura conceptual del cuidado en México, que hay un desbalance en la cobertura de los grupos dependientes de este y que tanto la ENUT y la ENOE son fuentes de información que permiten abordar el tema y relacionarlo con la actividad económica de mujeres y hombres.

Palabras clave: trabajo no remunerado; cuidados; fuentes de información estadística.

Recibido: 26 de septiembre de 2019.
Aceptado: 27 de diciembre de 2019.

Introducción

El tema del cuidado ha adquirido relevancia tanto en el plano nacional como en el internacional, esto debido a las importantes transformaciones demográficas y socioeconómicas, así como a la creciente y continua participación de la población femenina en la actividad económica. En México, el trabajo de cuidado no remunerado es provisto, principalmente, por mujeres jóvenes y adultas, quienes en su mayoría lo hacen de forma gratuita, precaria e invisible. Durante los últimos 50 años se han dado cambios importantes en la estructura por edad, producto de la disminución paulatina de la fecundidad y el descenso de la mortalidad; esto ha hecho que las relaciones de dependencia y la demanda de cuidado se modifiquen, sobre todo por el aumento de adultos mayores que lo necesitan. Además, el aumento en la escolaridad, las crisis

The topic of care has acquired academic, political and social importance. There is a consensus that care is a right but it is not clear whose responsibility it is to grant it. Methodologies and statistical information have been developed —among which Time Use surveys have been crucial—to study care work and its related issues. The objective of the research was to examine the conceptual coverage of care, the dependent group's coverage, and the sufficiency of statistical information to address the care work. Results indicate that there is a good conceptual coverage on the care topic in Mexico, that there is an imbalance in the coverage of dependent care groups, and that both ENUT and ENOE are sources of statistical information that allow to address the issue of care and relate it to the economic activity of both women and men.

Key words: unpaid work; care; statistical sources of information.

económicas recurrentes y los cambios culturales han incentivado la participación económica de las mujeres, pero no se han disminuido las cargas de trabajo no remunerado, incluido el de cuidado, lo que ha generado consecuencias negativas para este grupo poblacional.

Ante estos cambios, el cuidado pasó de ser un conjunto de actividades con elementos indispensables para garantizar la reproducción social y el bienestar de las personas a ser reconocido como un concepto amplio y con diversos atributos que permite estudiar el trabajo de cuidado no remunerado, así como la economía del cuidado, el derecho a este, además de los regímenes y la organización social del cuidado, entre otros temas (Esquivel, 2015; García Guzmán, 2019). Si bien es cierto que hay un consenso de que el cuidado es un derecho y que todo aquel que lo necesita debería recibirlo,

no está tan claro de quién debe ser el responsable de garantizárselo a la población dependiente: ¿es el Estado, el mercado, la familia o la comunidad?

Se ha documentado que una organización social del cuidado (OSC), como la que prevalece en México, genera desigualdad en las oportunidades de participación económica de mujeres y hombres. La OSC se refiere a la forma en cómo se entiende, gestiona y distribuye la necesidad de cuidado que sostiene el funcionamiento del sistema económico y de la política social de un país (Arriagada, 2010; Esquivel, 2012). En México, la OSC tiene las siguientes características, por mencionar algunas: 1) persiste la visión de ligar los derechos sociales a esquemas formales de empleo, 2) continúa la tendencia de asignar a las mujeres el cuidado de las poblaciones infantil y adulta mayor, los enfermos y las personas con discapacidad, 3) los sistemas de protección social están fragmentados y 4) prevalece el concepto minimalista del universalismo por parte del Estado con servicios de baja calidad para los más pobres y excluidos.

También, se ha examinado la demanda y oferta de cuidado a nivel institucional y dentro de la familia. En este marco, se ha documentado que en Latinoamérica y México, esta ha sido, históricamente, la principal proveedora y dentro de ella, las mujeres (Leiva, 2015; Rodríguez y Castro, 2014; Sunkel, 2007). En la actual OSC se ha informado sobre una serie de asociaciones entre el trabajo de cuidado no remunerado y las desigualdades socioeconómicas y de género, además de la contribución que la economía del cuidado realiza a la economía en general (Durán, 2018; García Guzmán, 2019; Nieves y Robles, 2016; Pedrero, 2010; Rodríguez y Marzonetto, 2015), señalando que donde el hogar provee en su mayor parte este trabajo, la actividad de cuidado se convierte en un vector de desigualdades, con desventajas para sus integrantes femeninos.

A la par de la discusión teórica y conceptual, se han desarrollado metodologías e información estadística que permiten abordar los diversos temas asociados al cuidado y a los cambios sociodemo-

gráficos y económicos que se están presentando en México. De ahí la importancia de contar con fuentes de datos estadísticos regulares para revisar continuamente la situación de la OSC; de los regímenes, la economía y el trabajo no remunerado de cuidado; además de sus implicaciones económicas tanto a nivel social como individual, entre otros.

Históricamente, las encuestas sobre uso del tiempo han sido los insumos clave para el estudio del cuidado (Aguirre y Ferarri, 2014; Marco, 2012), pues han contribuido a medir y hacer visibles las desigualdades de género en el trabajo de cuidado no remunerado, la importante aportación que este realiza en la economía de los países, así como el entendimiento de los regímenes y de la organización social de cuidado mediante el reparto de esta labor al interior de los hogares y la adquisición de servicios de este tipo en el mercado.

En México, el tema del tiempo dedicado a esta actividad no remunerada (medido en horas y minutos) ha estado presente dentro de las estadísticas oficiales desde hace varias décadas: desde finales de la de los 80 en las encuestas de empleo, como la Nacional de Empleo Urbano (ENEU), Nacional de Empleo (ENE) y, la hoy vigente, Nacional de Ocupación y Empleo (ENOE).¹ Además, desde 1996, en México se han realizado cinco encuestas especializadas sobre uso del tiempo (1996, 1998, 2002, 2009 y 2014) y ha sido reconocido como un país que dispone de valiosas fuentes de información estadística e investigaciones sobre el tema de cuidado (Aguirre y Ferarri, 2014:15).

Así, el concepto del cuidado está en constante evolución y perfeccionamiento, y las encuestas incluyen, cada vez más, reactivos para diferentes actividades relacionadas con el trabajo de cuidado. Por ello, el objetivo de esta investigación es examinar si las actuales fuentes de datos estadísticos en México tienen una cobertura conceptual de las distintas categorías de cuidado, si hay balance en la cobertura entre los grupos dependientes de este

¹ Todas estas levantadas por el Instituto Nacional de Estadística y Geografía (INEGI) de México.

y si es suficiente la información disponible para abordar los diversos temas en este ámbito; por ejemplo, para identificar tanto el régimen o la OSC que predomina en el país como el grado en que los hogares satisfacen su propia demanda de cuidado o si lo adquieren en el mercado con costo o de manera gratuita en otros hogares e instituciones, así como para profundizar en las implicaciones económicas negativas del cuidado para quienes cuidan.

Para lograr el objetivo, se tiene como punto de partida la revisión conceptual del cuidado y, a partir de ello, se examina la cobertura que presenta sobre el tema la Encuesta Nacional sobre Uso del Tiempo (ENUT) 2014 de México² y se compara con otras fuentes de información internacionales similares.

Marco conceptual de cuidado

El debate sobre su conceptualización tuvo sus orígenes en los países anglosajones en la década de los 70 y aún se está distante de lograr un consenso (Batthyány *et al.*, 2014), pues no existe un acuerdo en términos de: ¿cuáles actividades son consideradas dentro del trabajo de cuidado?, ¿quién(es) es(son) el(los) que recibe(n) o debe(n) recibirlo?, ¿quién(es) debería(n) otorgarlo? y ¿dónde y durante cuánto tiempo se realiza el cuidado? (Budig y Folbre, 2004). Justamente, sin un consenso se presentan divergencias al momento de estimar o medir la demanda y oferta de cuidado entre sociedades y países (Arriagada, 2010; Budig y Folbre, 2004; Durán, 2018).

Con la finalidad de tener un concepto y categorización del cuidado para fines del presente trabajo y en el intento por responder a la pregunta: ¿cuáles actividades son consideradas como parte del cuidado?, diversas investigaciones hacen referencia a las actividades esenciales para satisfacer las necesidades básicas de la vida y la reproducción de las personas, labores que proporcionan elementos físicos y simbólicos que representan un pilar tanto para el bienestar físico y emocional como para el

desarrollo de las capacidades y potencialidades de los individuos que les permiten vivir en sociedad y entretejiendo una compleja red para el sostenimiento de la vida humana (Aguirre *et al.*, 2014; Nieves y Robles, 2016; Rodríguez y Marzonetto, 2015). A continuación, se presenta una clasificación de las actividades asociadas al cuidado, recuperando las categorías que los mismos investigadores han empleado (Batthyány, 2008; Budig y Folbre, 2004; Folbre, 2006a, 2006b, 2011; Folbre y Yoon, 2008b; Rodríguez y Marzonetto, 2015; Suh, 2016; Suh y Folbre, 2016):

- 1) El autocuidado. Son todas las labores asociadas al cuidado, pero que puede proveerse la misma persona sin la ayuda de un tercero, entre ellas, alimentarse y asearse.
- 2) El cuidado directo o activo. Incluye un conjunto de actividades principales que supone interacción directa entre la persona que provee los cuidados y el individuo dependiente quien los recibe. En esta categoría se distinguen dos subcategorías:
 - a) Las tareas de carácter *imprescindible*, que abarcan el dar de comer, bañar, asear, vestir o ayudar a hacerlo a terceras personas; apoyar para entrar y salir de la cama, en el uso del sanitario y a moverse o caminar al interior del hogar; administrar medicamento. Estas son las más exigentes en cuanto a ritmo y tiempo para su ejecución y, en la literatura sobre cuidado en la vejez, algunas de estas son consideradas como actividades básicas de la vida diaria (ABVD).
 - b) Las labores *socialmente creadas*, que comprenden la ayuda con las tareas escolares y el apoyo en el aprendizaje a una tercera persona, el tiempo dedicado a leer y al juego, hacer alguna terapia especial o ayudar a realizar ejercicio, así como el acompañamiento.
- 3) El cuidado indirecto o pasivo. Se define como un conjunto de actividades secundarias, no necesariamente de contacto directo con el dependiente de cuidado. Sin embargo, diversos investigadores destacan su importancia porque generan un beneficio hacia este.

² También levantada por el INEGI.

Además, es la categoría donde existe mayor divergencia en cuanto a la delimitación del cuidado. En esta categoría se identifican cinco subcategorías:

- a) La *supervisión* de niños u otros dependientes de cuidado preservado en un ambiente seguro. Como el que se realiza a distancia con el monitoreo mediante cámaras y otros medios electrónicos mientras ellos pasan tiempo jugando. La supervisión puede traslaparse con actividades primarias, ya que es posible cuidar a la par de que se ve televisión o se cocina.
- b) La *gestión* del cuidado. Incluye la coordinación de horarios; realizar traslados a centros educativos o instituciones de salud y su acompañamiento; contratación y supervisión de niñeras, enfermeras, trabajadoras domésticas, así como otros cuidadores remunerados; ayudar en la compra de víveres y artículos personales; y manejar el dinero del dependiente del cuidado; por mencionar algunos.
- c) El cuidado en *beneficio de (for whom)*. Implica las actividades o el trato con terceras personas para el beneficio del dependiente, siempre que estuviese vinculado con la atención de su salud o con las necesidades educativas, como conversar con la maestra de los hijos sobre asuntos escolares o platicar con los médicos acerca de asuntos de salud del enfermo dependiente.
- d) El aspecto de *en guardia (on-call)*. Es una tarea que requiere estar disponible para responder a una alarma de cuidado durante largos periodos; por ejemplo, bebés, enfermos o personas adultas mayores, que pasan gran parte de su tiempo durmiendo en casa u hospitales, requieren a alguien *en guardia* para atender sus necesidades aun durante la siesta del cuidador, o en el caso de los niños en edad escolar que requieren de un responsable para que cuando se llegare a presentar una emergencia este acuda a su ayuda, aun cuando el responsable esté en horario laboral.

e) *Apoyo al cuidado (support care)*. Es una dimensión que suele ser olvidada porque no implica una interacción directa con los dependientes de cuidado, pero prepara el escenario del cuidado activo o directo, como la compra y preparación de alimentos y la limpieza de la casa, que le dan soporte a este. Algunos investigadores sugieren el detallar para quién (*who for*) es el tiempo de la preparación de alimentos, limpieza o aseo de la casa meramente asociada al dependiente de cuidado, más allá del detalle de la preparación de alimentos o limpieza en general de la vivienda.

Finalmente, cabe precisar que algunas de las labores asociadas a la gestión y apoyo al cuidado (como la preparación de comida, hacer compras de víveres y el manejo de dinero) son consideradas como actividades instrumentales de la vida diaria (AIVD) en la literatura sobre cuidados en adultos mayores.

La definición de las tareas que integran el cuidado y el detalle con el que se presenta facilita responder la pregunta: ¿quién(es) es(son) el(los) que lo reciben? Toda la gente lo necesita a lo largo del ciclo de vida, sin embargo, no todos pueden proveérselo por sí solos, de ahí que la acción de cuidar permite atender las necesidades de cuidado de las personas dependientes. Se entiende que un individuo está en esa condición porque no puede realizar las actividades de cuidado por sí mismo debido a la edad —los niños y ancianos—, así como por limitadas condiciones y/o capacidades temporales o permanentes —las enfermedades y/o discapacidades— (Batthyány *et al.*, 2017; Rodríguez y Marzonetto, 2015).

Acotar el cuidado a situaciones de dependencia presenta la problemática de ser entendido como un concepto dual, donde el proveedor es pensado como una persona totalmente autónoma e independiente y quien recibe el cuidado, como un individuo con total dependencia; sin embargo, se debe reconocer que darlo y recibirlo son actividades que pueden suceder de manera simultánea, y la decisión de acotar el cuidado con

la dependencia responde a la necesidad académica de especificar el alcance de los cuidados en la investigación empírica y en el análisis de las políticas públicas (Batthyány *et al.*, 2017). En este sentido, usualmente se considera proveedores de cuidado a la población en los rangos centrales de edad, por lo general entre 15 y 65 años (Durán, 2018). Además, se reconoce que el proveedor puede dedicar tiempo de forma simultánea a distintos grupos poblacionales; por ejemplo, algunos autores la han llamado la generación sándwich o *sandwich carers generation*, pues tienen que cuidar a niños y ancianos al mismo tiempo (Suh, 2016) y, en los contextos mexicano y latinoamericano, este fenómeno tomará importancia en el corto y mediano plazos dado el acelerado envejecimiento demográfico que se presenta.

Una de las características del cuidado es que puede ser desempeñado por personas que obtienen o no una remuneración, y esta condición rige en gran medida el lugar donde se realiza. La actividad no remunerada se lleva a cabo principalmente por las mujeres en la residencia de los receptores del cuidado, de los familiares o en sitios comunitarios, en tanto que el trabajo remunerado, ya sea adquirido en el mercado u otorgado por el Estado, por lo general se hace en los sectores de salud y educación, pero también mediante el servicio doméstico remunerado o asistencia médica domiciliaria, ámbitos laborales donde el sexo femenino es mayoría (Espino, 2011; Esquivel, 2011).

A pesar de los debates aún inconclusos, se aprecia un importante avance en la noción del cuidado, lo que permite examinar de manera cada vez más completa el tema. De ahí la importancia en revisar la cobertura conceptual en las fuentes de información estadísticas dados sus cambios y tendencias producto de transformaciones demográficas y socioeconómicas.

Aproximación al cuidado y fuentes de información estadística

El desarrollo conceptual del cuidado avanza de manera paralela con la generación de instrumen-

tos metodológicos para su operacionalización y medición, con énfasis desde quien lo otorga dentro de los hogares y con la finalidad de mostrar las diferencias por género, por estrato socioeconómico, las implicaciones en la actividad económica de mujeres y hombres, así como la contabilización de la contribución de la tarea de cuidado a la economía. La aproximación desde la oferta está estrechamente ligada al trabajo de cuidado que tiene sus orígenes en la medición del trabajo no remunerado en sus distintas modalidades: doméstico, de cuidado y voluntario o comunitario, mediciones que rescatan el criterio *de tercera persona* presentado en 1939 por Margaret Reid para definir como trabajo a aquellas actividades que se pueden pagar a una tercera persona para que lo realice en sustitución de uno mismo o de otro individuo (Folbre y Yoon, 2008a, 2008b).

La principal forma de medir del trabajo de cuidado no remunerado es a través del tiempo (medido en horas y minutos) que se dedique a este, ya que es un indicador que ha permitido visualizar el desigual reparto del cuidado realizado en la esfera familiar, su organización, implicaciones y valor equivalente con el trabajo de mercado (Carrasco, 2016; Marco, 2012); por ello, la agenda de investigación ha encontrado en las encuestas de uso del tiempo una alternativa para responder a las interrogantes sobre cómo se organiza socialmente el cuidado, cuáles son las implicaciones del trabajo de cuidado no remunerado para quienes lo ejercen y cuánto aportan a la economía nacional, entre otras.

En esta sección se analizan cuáles categorías de cuidado están presentes en las fuentes de información estadística de México, tomando como principal referente la ENUT 2014 y comparándola con otras similares en Latinoamérica y Norteamérica, y con las nacionales derivadas de encuestas en hogares.

En los cuadros 1 y 2 se presentan los detalles metodológicos de las encuestas o los módulos acerca del uso del tiempo para Argentina, Canadá, Chile, Costa Rica, Ecuador, Estados Unidos de América (EE. UU.), México, Paraguay y Uruguay, así como sie-

te instrumentos disponibles en México que tocan el tema de cuidado: las encuestas Intercensal (EI) 2015, Nacional de Ingresos y Gastos de los Hogares (ENIGH), Nacional de Ocupación y Empleo (ENOE), Nacional sobre Salud y Envejecimiento en México (ENASEM), Laboral y de Corresponsabilidad Social (ELCOS) —con preguntas sobre percepción del

cuidado—, Nacional de Empleo y Seguridad Social (ENESS) —sobre el uso de la oferta de cuidado en el mercado y del Estado, así como del gasto del cuidado en los hogares— y Nacional de Gastos de los Hogares (ENGASTO).³

³ De las cuales, la ENGASTO y ELCOS no capturan el tiempo.

Cuadro 1

Continúa

Encuestas que abordan el tema de cuidado en los hogares en Latinoamérica y Norteamérica

Nombre	Siglas	País	Tipo de levantamiento de la información	Periodo de levantamiento	Cobertura geográfica	Población objetivo	Tamaño de muestra	Instituto	Clasificación
Encuesta Nacional sobre Uso del Tiempo	ENUT	México	Encuesta	Del 13 de octubre al 28 de noviembre de 2014	Nacional	Población de 12 años y más de edad	16 996 viviendas	INEGI	Clasificación Mexicana de Actividades sobre Uso del Tiempo (CMAUT), 2005; es una adaptación de la Clasificación Internacional para Estadísticas de Uso del Tiempo (ICATUS)
Módulo Trabajo no Remunerado y Uso del Tiempo, 2013	Módulo EAHU		Módulo en la Encuesta Anual de Hogares Urbanos	Tercer trimestre de 2013	Localidades de 2 mil y más habitantes	Población de 18 años y más	46 mil viviendas	Instituto Nacional de Estadística y Censos (INDEC).	Ninguna, son preguntas sintéticas
Encuesta Nacional sobre Uso del Tiempo, 2015	ENUT	Chile	Encuesta	Cuarto trimestre de 2015 (21 sept.-13 de dic.)	Localidades de 2 000 y más habitantes	Población de 12 años y más	11 623 viviendas	Instituto Nacional de Estadística (INE)	Clasificación de Actividades para América Latina y El Caribe (CAUTAL)
Encuesta Nacional de Uso del Tiempo, 2017	ENUT	Costa Rica	Encuesta	Del 23 de octubre al 24 de noviembre de 2017	Nacional, urbano-rural	Población de 12 años y más	4 560 viviendas	Instituto Nacional de Estadística y Censos (INEC).	CAUTAL para el marco general y CMAUT para mayor desagregación
Encuesta Específica de Uso del Tiempo, 2012	EUT	Ecuador	Encuesta	Del 7 de noviembre al 30 de diciembre de 2012	Nacional	Población de 12 años y más	22 968 viviendas	INEC	CAUTAL

Encuestas que abordan el tema de cuidado en los hogares en Latinoamérica y Norteamérica

Nombre	Siglas	País	Tipo de levantamiento de la información	Periodo de levantamiento	Cobertura geográfica	Población objetivo	Tamaño de muestra	Instituto	Clasificación
Encuesta sobre Uso del Tiempo, 2016	EUT	Paraguay	Encuesta	Del 9 de mayo al 12 de agosto de 2016	Nacional, urbano-rural	Población de 14 años y más	4 272 viviendas	Dirección General de Estadística, Encuestas y Censos	CAUTAL
Módulo Uso del Tiempo, Encuesta Continua de Hogares, 2013	EUT-ECH	Uruguay	Módulo en la Encuesta Continua de Hogares	Mayo a julio de 2013	Nacional	Población de 14 años y más	52 199 viviendas	Instituto Nacional de Estadística (INE)	Ninguna, son preguntas sintéticas
American Time Use Survey, 2011-17	ATUS	EE. UU.	Encuesta	Mensual	Nacional, estatal.	Población de 15 años y más.	26 400 personas, divididas mensualmente	Bureau of Labor Statistics	No especificado.
General Social Survey on Time Use, 2015	GSS	Canadá	Encuesta	Abril de 2015 a abril de 2016	10 de 13 de las provincias.	Población de 15 años y más.	2 % de la población objetivo (25 mil individuos).	Statistics Canada	No especificado.

Fuente: elaboración propia con base en cuestionario de cada encuesta publicados en la página web de las instituciones ejecutoras.

Encuestas que abordan el tema de cuidado en los hogares de México

Nombre	Siglas	Tipo de encuesta	Periodicidad	Cobertura geográfica	Población objetivo	Tamaño de muestra	Oferta/demanda	Secciones	Asistencia escolar
Encuesta Nacional sobre Uso del Tiempo	ENUT	Especial	No definido (2014, 2009, 2002)	Nacional	Oferta: integrantes de 12 años o más de edad. Demanda: todos los integrantes.	16 996 viviendas (2014)	Oferta: familiar y del mercado por trabajadores domésticos, cuidadores contratados y guarderías. Demanda: integrantes con enfermedades crónicas, limitaciones físicas y mentales, y enfermos temporales.	Sección VI	5 años y más de edad, pero identifica la asistencia a guardería para la población de 0 a 4 años.

Encuestas que abordan el tema de cuidado en los hogares de México

Nombre	Siglas	Tipo de encuesta	Periodicidad	Cobertura geográfica	Población objetivo	Tamaño de muestra	Oferta/demanda	Secciones	Asistencia escolar
Encuesta Intercensal 2015		Especial	No definido (2015)	Nacional, entidad federativa, municipal, localidades con 50 mil y más habitantes	Oferta: integrantes de 12 años o más	6.1 millones de viviendas	Oferta: familiar.	Sección III, pregunta 41	3 años y más
Encuesta Nacional de Ingresos y Gastos de los Hogares	ENIGH	Regular	Bienal (año par)	Nacional y entidades federativas, área urbana (2 500 y más habitantes) y rural (menos de 2 500 habitantes).	Oferta: integrantes de 12 años o más	64 mil viviendas a nivel nacional y 2 mil por entidad federativa	Oferta: familiar y contratación en el mercado. Demanda: integrantes con discapacidad motriz, visual, habla, oír, AIVD y déficit de atención.	Demanda: cuestionario de hogares y vivienda (sección III, pregunta 7). Oferta: cuestionario para personas de 12 años o más años (sección IX).	3 años y más
Encuesta Nacional de Ocupación y Empleo	ENOE	Regular	Mensual y trimestral	Nacional, entidad federativa, ciudades autorrepresentadas y cuatro tamaños de localidades	Oferta: integrantes de 12 años o más.	122 mil viviendas	Oferta: familiar y del mercado por trabajadores domésticos.	Sección IX cuestionario básico, XI cuestionario ampliado.	5 años y más
Encuesta Nacional de Empleo y Seguridad Social.	ENESS	Regular	Cada cuatro años: 2017, 2013, 2009, 2004, 2000 y 1996	Nacional y entidad federativa	Demanda: población total de residentes habituales en las viviendas participantes.	83 585 viviendas	Oferta: familiar, del mercado y Estado. Demanda: niños de 0 a 6 años y población de 60 años y más.	Módulo de la ENOE.	Ver ENOE.
Encuesta Nacional de Gastos de los Hogares.	ENGASTO	Especial	2012 y 2013	Nacional, urbano-rural	Demanda: hogares con personas de 0 a 6 años de edad, 7 a 59, 60 y más.	71 851 viviendas	Oferta: familiar, del mercado y Estado. Demanda: se pregunta por necesidades de cuidado pero no por la intensidad de cuidado. Hogares con personas de 0 a 6 años de edad, de 7 a 59, y de 60 años y más.	Secciones VIII, IX y X del cuestionario sociodemográfico.	3 años y más.

Encuestas que abordan el tema de cuidado en los hogares de México

Nombre	Siglas	Tipo de encuesta	Periodicidad	Cobertura geográfica	Población objetivo	Tamaño de muestra	Oferta/demanda	Secciones	Asistencia escolar
Encuesta Laboral y de Corresponsabilidad Social.	ELCOS	Especial	No definido (2012)	Nacional, Ciudad de México, Guadalajara, Monterrey y agregado de 29 áreas urbanas de 100 mil y más habitantes.	Demanda: integrantes del hogar. Oferta: personas de 12 años o más de edad. Mujer de 14 a 70 años, elegida en el hogar.	15 mil viviendas	Oferta: familiar y del mercado por trabajadores domésticos, cuidadores contratados y guarderías. Demanda: integrantes con limitaciones permanentes, de 0 a 5 años de edad, de 6 a 14 años que necesitan cuidado y enfermos temporales de 6 años y más.	Secciones III y IV	5 años y más de edad, pero identifica la asistencia a guardería para la población de 0 a 4 años.
Encuesta Nacional sobre Salud y Envejecimiento en México	ENASEM	Especial	Cada tres años (2001, 2003, 2012, 2015, 2018).	Nacional, urbano-rural.	Persona de 50 o más, elegida en el hogar y su cónyuge sin importar edad.	11 mil viviendas.	Oferta. Demanda de personas de 50 años y más por ABVD y AIVD.	Sección D y H.	No pregunta asistencia escolar, solo escolaridad: 12 años y más de edad.

Fuente: elaboración propia con base en el cuestionario de cada encuesta (www.beta.inegi.org.mx).

En el cuadro 3 se sintetiza la cobertura conceptual de cada fuente de información estadística de acuerdo con las categorías y subcategorías del cuidado descritas con anterioridad: el autocuidado, el cuidado directo (en sus dos modalidades: el imprescindible y el socialmente creado), así como el indirecto (con sus subcategorías de supervisión, gestión, en beneficio de, de guardia y apoyo). Para tal propósito, se considera como categoría conceptualmente cubierta (✓) si la fuente de in-

formación contiene al menos una pregunta relacionada con las actividades descritas en las categorías de cuidado.

Un primer hallazgo, tal como se esperaba, es que las encuestas especializadas en uso del tiempo tienen mayor cobertura conceptual que las que abordan el tema del cuidado mediante módulos, por ejemplo, las de Argentina y Uruguay. Al comparar las encuestas en México, la ENUT cubre más

el tema que la EI, ENOE o ENIGH, las cuales utilizan módulos. Además, junto con Ecuador y Costa Rica, México se destaca por tener encuestas sobre uso del tiempo con una cobertura más extensa en la región, pues contemplan al menos una pregunta en todas las categorías del cuidado.

Al examinar por categorías de cuidado, son dos las que están presentes en todas las fuentes de información analizadas: *cuidado directo imprescindible* (como dar de comer o administrar medicamentos) y *gestión del cuidado* (que incluye traslados a terapias, centros escolares o citas médicas). Esta presencia puede deberse a que existe mayor consenso e importancia de las actividades que las integran, a diferencia del resto de las categorías (ver cuadro 3).

La de *autocuidado* está presente solo en las encuestas especializadas sobre uso del tiempo y captura el tiempo dedicado por las personas para dormir, alimentarse y asearse. La de *guardia* aparece de diversas formas, denotando la falta de consenso para su captación. En algunos casos se pregunta de forma clara y precisa, como en el de la Encuesta Específica de Uso de Tiempo de Ecuador;⁴ otros países incorporan de manera general el cuidado de *guardia* en combinación con otras categorías, como Chile y México.⁵ La de *apoyo al cuidado* está incorporada en todos los instrumentos de información analizados; sin embargo, es necesario precisar la importancia de separar esas tareas del resto de las labores domésticas para no incluir en el conteo de tiempo de apoyo al cuidado en actividades que no lo son, como el cuidado de animales de traspas, la reparación y mantenimiento del hogar, entre otras, como ocurre en el caso de la encuesta de Argentina y la ENIGH de México que no las separan.

Las encuestas de Ecuador y México (ENUT) representan un avance en la propuesta de detallar el

4 Por ejemplo: "¿Cuidó a algún miembro del hogar enfermo/a hospitalizado/a o lo acompañó en el hospital, clínica durante el día y/o noche?" y "Durante la noche, ¿Estuvo en vela cuidando a (Persona con discapacidad) que es miembro del hogar?"

5 Preguntas como: "¿Le(s) dio medicamentos o checó sus síntomas?", "¿(temperatura, presión, otros)?", "¿Estuvo pendiente de (nombre), si presentó algún síntoma, malestar o enfermedad durante la noche?", "¿Dar medicamentos, algún tratamiento de salud o cuidar por alguna enfermedad?"

trabajo de cuidado *para quién* (*who for*); en estas, más allá de referir a la preparación de alimentos en general, se pregunta si preparó algún remedio, comida o alimento especial para alguna persona con discapacidad. La categoría de *supervisión de cuidado* está presente en casi todas las encuestas de uso del tiempo (excepto en la de Chile); en algunas ocasiones, esta categoría puede traslaparse con otras actividades y, por ello, quizá no se captura de manera consistente en los módulos sobre el cuidado. Finalmente, la de *en beneficio de* se encuentra solo en las encuestas especializadas de uso del tiempo, aunque se orienta hacia el apoyo de las tareas escolares y la asistencia a reuniones o actividades escolares.

En cuanto a la cobertura conceptual del cuidado por grupos dependientes, se encontró que existe un desbalance con énfasis en la población infantil sobre la de adultos mayores, y esto puede deberse a la etapa de la transición demográfica en la que se encuentra cada uno de los países de la región. Como es de esperarse, en las encuestas predominan las preguntas sobre actividades del cuidado infantil y centradas en aspectos educativos; la parte de salud se aborda someramente, como ocurre en las categorías de *cuidado directo socialmente creado* y *en beneficio de*. Además, solo algunas fuentes de información incluyen reactivos relacionados con el tiempo dedicado a personas adultas mayores y con enfermedades permanentes o temporales (por ejemplo, realizar terapias y el acompañamiento diario), así como a platicar con profesionales de la salud (médicos, enfermeras, terapeutas, psicólogos, trabajadores sociales, entre otros).

Al examinar la ENUT (encuesta especializada en uso del tiempo) y la ENOE (que incorpora un módulo sobre este) de 2014 con respecto a las categorías de cuidado y por grupos dependientes, en el cuadro 4 se puede apreciar una mayor cobertura en estas y desagregación del cuidado por grupos dependientes en el primer instrumento que en el segundo; es evidente y natural que los módulos en encuestas no cuentan con el nivel de detalle que tiene una especializada. Si bien es cierto que la ENUT cubre todas las categorías de cuidado plan-

Cobertura conceptual desde la oferta de cuidado de las encuestas seleccionadas

Encuestas o módulos	Módulo de Trabajo no Remunerado y Uso del Tiempo 2013	Encuesta Nacional sobre Uso del Tiempo 2015	Encuesta Nacional de Uso del Tiempo 2017	Encuesta Específica de Uso del Tiempo 2012	Encuesta sobre Uso del Tiempo 2016	Módulo de Uso del Tiempo, Encuesta Continua de Hogares 2013	American Time Use Survey, 2011-17	General Social Survey on Time Use, 2015
Países	Argentina	Chile	Costa Rica	Ecuador	Paraguay	Uruguay	EE. UU.	Canadá
Categorías del cuidado								
1. Autocuidado		√	√	√	√		√	√
2. Cuidado directo								
a. Imprescindibles	√ Con 3b	√ Con 3d	√	√	√	√	√	√
b. Socialmente creadas	√ Escolar	√	√	√	√	√		
3. Cuidado indirecto								
a. Supervisión			√	√	√			
b. Gestión	√ Con 2a	√	√	√	√	√	√	√
c. En beneficio de		√ Educ.	√ Educ.	√ Educ.				
d. De guardia		√ Con 2a	√	√	√			
e. Apoyo al cuidado	√ Trab. Dom.	√	√	√	√	√	√	√
General							√	√

Concluye

Encuestas o módulos	Encuesta Nacional sobre Uso del Tiempo (ENUT) 2014	Encuesta Intercensal 2015	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)	Encuesta Nacional de Ocupación y Empleo (ENOE)	Encuesta Nacional de Empleo y Seguridad Social (ENESS)	Encuesta Nacional de Gastos de los Hogares (ENGASTO)	Encuesta Laboral y de Corresponsabilidad Social (ELCOS)	Encuesta Nacional sobre Salud y Envejecimiento en México (ENASEM)
Países	México	México	México	México	México	México	México	México
Categorías del cuidado								
1. Autocuidado	√							
2. Cuidado directo								
a. Imprescindibles	√ Con 3d	√ Con 3b	√ Con 3b	√		√	√ Con 3b	√ Dep.
b. Socialmente creadas	√							
3. Cuidado indirecto								
a. Supervisión	√							
b. Gestión	√	√ Con 2a	√ Con 2a	√		√	√ Con 2a	√ Dep.
c. En beneficio de	√ Educ.							
d. De guardia	√ Con 2a							
e. Apoyo al cuidado	√	√	√ Trab. Dom.	√			√ Trab. Dom.	
General					√			√ Oferta

Nota: *Con* indica que se pregunta conjuntamente con otra categoría; *Trab. Dom.* significa trabajo doméstico e indica que el apoyo de cuidado se pregunta conjuntamente con otras actividades domésticas diferentes al trabajo de cuidado; *Educ.* indica que las preguntas de cuidado giran en torno al tema educativo; *Dep.* indica que da información sobre cuidados que recibe el dependiente; y *Oferta* se refiere al cuidado que da la oferta.

Fuente: elaboración propia con base en cuestionarios de cada encuesta publicados en la página web de las instituciones ejecutoras.

teadas, al hacer un cruce entre dichos elementos se observa una menor captación de información de cuidados hacia adultos mayores no dependientes —en comparación con la población infantil no dependiente— en las categorías de *cuidado directo imprescindible y socialmente creado*, así como en la de *cuidados indirectos* (como *en beneficio de y de guardia*). Esta situación limita la equitativa aproximación a las demandas de cuidado derivadas de los grupos de edad extremos y, en particular, para los estudios sobre la *generación sándwich* que

tiene requerimientos tanto de la población infantil como de la adulta mayor. Además, la ENUT también tiene limitaciones en términos de la captura de la oferta de cuidado hacia la población con discapacidad, enfermedad crónica o temporal, pues está englobada para todas las edades y, para separarlas, es necesario implementar artificios estadísticos y tener buen manejo de bases de datos para estimar o imputar el tiempo de cuidado para las poblaciones infantil y adulta mayor con discapacidad o enfermedad crónica o temporal.

Cuadro 4

Continúa

Asignación de preguntas de la ENUT y ENOE 2014 por categorías de cuidado

Categorías del cuidado	ENUT				ENOE
	Discapacidad, enfermedad crónica o temporal	Pob. 0 a 14 años de edad (no dependientes)	Pob. 60 años y más (no dependientes)	Pob. 15 a 59 años (no dependientes)	Población general
1. Autocuidado Autocuidado	Entrevistados				
2. Cuidado directo					
a. Imprescindibles	6.11.1. ... Le(s) dio de comer o ayudó a hacerlo? 6.11.2. ... Lo(s) bañó, aseó, vistió, arregló o ayudó a hacerlo? 6.11.3. ... Lo(s) cargó, acostó o le(s) ayudó a hacerlo?	Población de 0-5 años 6.12.1. ... Le(s) dio de comer (amamantó) o dio de beber? 6.12.2. ... Lo(s) bañó, aseó (cambió pañales), vistió o arregló? 6.12.3. ... Lo(s) cargó o acostó?			9.2 ... Cuidar o atender sin pago, <i>de manera exclusiva</i> , a niños, ancianos, enfermos o discapacitados? (bañarlos, cambiarlos)
b. Socialmente creadas	6.11.7. ... Le(s) dio terapia especial o ayudó a realizar ejercicios? 6.11.9. ... Lo(s) ayudó o apoyó en las tareas de la escuela o trabajo?	6.13.2. ... Le(s) dio terapia especial o ayudó a realizar ejercicios? 6.13.3. ... Lo(s) ayudó en las tareas de la escuela?			
3. Cuidado indirecto					
a. Supervisión	6.11.11. ... Mientras hacía otra cosa , lo(s) cuidó o estuvo al pendiente?	6.13.6. ... Mientras hacía otra cosa , lo(s) cuidó o estuvo al pendiente?	6.15.4. ... Mientras hacía otra cosa , lo(s) cuidó o estuvo al pendiente?		
b. Gestión	6.11.6. ... Lo(s) llevó, recogió o esperó para que recibiera(n) atención de salud (exámenes, visitas al médico, etc.) o alguna terapia especial? 6.11.8. ... Lo(s) llevó y/o recogió de clases, trabajo u otro lugar?	6.13.1. ... Lo(s) llevó y/o recogió de la guardería, de clases, de la casa de algún familiar o amigo para ser cuidado? 6.13.5. ... Lo(s) llevó, recogió o esperó para que recibiera(n) atención de salud? (vacunas, dentista, chequeo médico, etc.)	6.15.2. ... Lo(s) llevó, recogió o esperó para que recibiera(n) atención de salud? (vacunas, dentista, chequeo médico, etc.) 6.15.3. ... Lo(s) llevó y/o recogió del trabajo, de algún trámite u otro lugar?	6.14.2. ... Lo(s) llevó, recogió o esperó para que recibiera(n) atención de salud? (vacunas, dentista, chequeo médico, etc.) 6.14.3. ... Lo(s) llevó y/o recogió de clases, trabajo, de algún trámite u otro lugar?	9.4 ... Llevar a algún miembro del hogar a la escuela, cita médica u otra actividad?

Asignación de preguntas de la ENUT y ENOE 2014 por categorías de cuidado

Categorías del cuidado	Discapacidad, enfermedad crónica o temporal	ENUT			ENOE
		Pob. 0 a 14 años de edad (no dependientes)	Pob. 60 años y más (no dependientes)	Pob. 15 a 59 años (no dependientes)	Población general
c. En beneficio de	6.11.10. ... Asistió a juntas, festivales o actividades de apoyo escolar?	6.13.4. ... Asistió a juntas, festivales o actividades de apoyo en la guardería o escuela?			
d. De guardia	6.11.5. ... Le(s) dio medicamentos o checó sus síntomas? (temperatura, presión, otros)				
e. Apoyo al cuidado	6.11.4. ... Le(s) preparó remedios caseros o algún alimento especial?	Población en general 6.4.1. ... Desgranó maíz, cocinó o molió el nixtamal o hizo tortilla de maíz o trigo para su hogar? 6.4.2. ... Encendió el fogón, horno o anafre de leña o carbón para preparar o calentar alimentos? 6.4.3. ... Cocinó, preparó o calentó alimentos o bebidas? 6.4.4. ... Sirvió la comida, recogió, lavó, secó o acomodó los trastes? 6.4.5. ... Llevó comida a algún integrante de su hogar a la escuela, trabajo u otro lugar? 6.5.1. ... Barrió la banquetta, cochera o patio de su vivienda? 6.5.2. ... Limpió o recogió el interior de su vivienda? 6.5.3. ... Recogió, separó, tiró o quemó la basura? 6.6.1. ... Lavó, tendió o puso a secar la ropa? 6.10.1. ... Llevó o recogió ropa o calzado a algún lugar para su limpieza o reparación? 6.10.7. ... Organizó o repartió los quehaceres de su hogar?			9.7 ... Realizar los quehaceres de su hogar? (lavar, planchar, preparar y servir alimentos, barrer)

Fuente: elaboración propia con base en el cuestionario de cada encuesta (www.beta.inegi.org.mx).

Además del tiempo de cuidado, en México se utilizan aproximaciones analíticas alternas relacionadas con las actividades de este; por ejemplo, las encuestas ENUT, ELCOS y ENESS identifican la cantidad de dependientes de cuidado en el hogar y por grupos de edad; la ELCOS captura la percepción de las mujeres cuidadoras; la ENASEM examina el cuidado otorgado por los adultos mayores e indaga sobre el número de veces que ellos cuidaron en el mes de referencia; la ENASEM y ENESS registran la demanda de cuidado de los adultos mayores con alguna limitación para las ABVD y las AIVD; la ENGASTO, ENESS y ENIGH captan el gasto en esta labor, aunque este se limita a algunos rubros del cuidado, como en estancias infantiles y a cuidadoras y cuidadores.

Estrategias diversas para abordar los temas del cuidado

Estos han tomado importancia al mismo tiempo que se reconocen los diversos abordajes para su estudio, entre ellos, el trabajo de cuidado no remunerado, sus implicaciones económicas, su contribución a la economía, el régimen y la organización social del cuidado, el derecho a este, entre otros. Cada uno de estos requiere seguir estrategias diversas en función de las fuentes estadísticas disponibles, pues obtener datos adecuados para desarrollar los temas de investigación requiere de recursos económicos y tiempo, por lo que surge la siguiente pregunta: ¿las fuentes de información estadística disponibles en México son suficientes

para abordar los temas de cuidado? De acuerdo con lo presentado con anterioridad, la respuesta es sí, aunque con ciertas limitaciones pues, en algunas categorías de cuidados, la información está segmentada y parcialmente cubierta, es decir, no hay una sola fuente que cubra todos los rubros expuestos.

Pensando en temas globales por cubrir, para examinar el régimen y la organización social del cuidado, se requiere que las fuentes de información capturen la *demanda y oferta de cuidado* (ver cuadro 5). La primera se capta en la ENUT, ELCOS, ENESS y ENASEM; en estas se pregunta por

el número de personas en el hogar con necesidades de cuidado de otra persona, ya sea por enfermedad crónica, limitación física o mental, si se trata de un enfermo temporal o con dificultades para realizar las ABVD o AIVD; a esta demanda le acompaña en prácticamente todas las encuestas la identificación de la demanda potencial de cuidado —población entre 0 a 15 y de 65 años y más, con algunas variantes en los límites de edad (Durán, 2018)—, pues se pregunta por la edad de los integrantes del hogar. La segunda se captura en la ENUT, EI, ENESS y ENASEM, ya que preguntan por el cuidado otorgado a grupos dependientes o se aborda de manera general en la ENIGH y ENOE.

Cuadro 5

Continúa

Características conceptuales y analíticas de las encuestas con temáticas de cuidado

Siglas	Cuidados extrahogar: recibidos y otorgados	Grupos dependientes	Aspectos del mercado laboral	Otras áreas de interés
ENUT	Cuidado otorgado a otros hogares, así como cuidados recibidos de personas ajenas al hogar.	1) Discapacitados, enfermos crónicos o temporales. 2) Niños de 0 a 5 años de edad. 3) Niños de 0 a 14. 4) 15 a 59 años, 5) Adultos de 60 o más.	Participación en el mercado laboral, jornada, tiempo de traslado al trabajo, posición en la ocupación, prestaciones e ingresos laborales.	Características de la vivienda, del hogar (como: apoyo de programas sociales, transferencias e ingresos no laborales, trabajadores contratados) y sociodemográficas de los integrantes del hogar.
El 2015	No especificado.	1) Discapacitad. 2) Enfermos. 3) Menores de 6 años. 4) De 6 a 14. 5) 60 o más.	Participación en el mercado laboral, tiempo de traslado al trabajo, posición en la ocupación y prestaciones e ingresos laborales.	Características de la vivienda y sociodemográficas de los integrantes del hogar.
ENIGH	No especificado.	General.	Participación en el mercado laboral, posición en la ocupación, prestaciones y jornada e ingresos laborales, entre otras características laborales.	Características de la vivienda, de los hogares (como asistencia a guarderías para niños de 3 años y más e ingresos por fuentes) y sociodemográficas de los integrantes del hogar.
ENOE	No especificado.	General.	Participación en el mercado laboral, posición en la ocupación, prestaciones y jornada e ingresos laborales, entre otras características laborales con mayor nivel de desagregación.	Características sociodemográficas de los integrantes del hogar.
ENESS	En adultos mayores sí especifica si no residen en el hogar; para la población infantil específica por relación de parentesco con el menor.	1) De 0 a 6 años. 2) 60 y más. 3) Con discapacidad.	Uso de guarderías en población de 0 a 6 años y razón de no uso de guarderías. Cotización actual y anterior a la seguridad social. Pensiones, riesgos de trabajo, licencias de maternidad.	Pago por el cuidado de población de 0 a 6 años, horarios de cuidados. Pago por el cuidado de adultos mayores.

Características conceptuales y analíticas de las encuestas con temáticas de cuidado

Siglas	Cuidados extrahogar: recibidos y otorgados	Grupos dependientes	Aspectos del mercado laboral	Otras áreas de interés
ENGASTO	No aplica.	1) De 0 a 6 años. 2) 7 a 59. 3) 60 y más.		Se captura gastos en cuidado por hogar y número de trabajadores domésticos, así como características sociodemográficas de los integrantes del hogar.
ELCOS	Captura apoyo de otros hogares de las mujeres elegidas por la muestra.	1) Personas con limitaciones permanentes. 2) De 0 a 5 años. 3) De 6 a 14. 4) Enfermos temporales de 6 o más años.	Participación en el mercado laboral, posición en la ocupación, ingreso por trabajo de todos los integrantes; asimismo, características laborales o de desempleo de las mujeres elegidas.	Características de la vivienda, apoyo de programas asistenciales recibidos y características sociodemográficas de los integrantes, incluida asistencia a la guardería de niños de 0 a 4 años de edad.
ENASEM	Es posible detectar si el cuidado otorgado o recibido es entre residentes o no residentes del hogar; sin embargo, estos cuidados son de y para las personas de 50 años y más de edad.	a) Cuidado otorgado a adultos enfermos o discapacitados y a niños menores de 12 años. b) Cuidado recibido por dificultades en actividades básicas e instrumentales de la vida diaria (ABVD y AIBD).	a) Participación en el mercado laboral e ingresos. Abandono del último trabajo por cuidado de familiares. b) Participación en el mercado laboral de las hijas e hijos que ofrecen directamente cuidado.	a) Características sociodemográficas, económicas y de salud de los entrevistados. b) Características de la vivienda. c) Características sociodemográficas de los residentes del hogar y de las y los hijos no residentes del hogar. d) Experiencia sobre cuidados en migrantes en EE. UU., cuidados preventivos, cuidado de los padres, percepción de contar con un cuidador en un futuro.

Fuente: elaboración propia con base en el cuestionario de cada encuesta (www.beta.inegi.org.mx).

Respecto a los cuidados externos recibidos, la ENUT se orienta en las redes informales de cuidado, pues identifica el tiempo otorgado a y recibido por otros hogares; si bien la ELCOS y ENASEM capturan ese apoyo, la primera fuente no mide el tiempo y la segunda está centrada en el cuidado de adultos mayores. En estos casos, el tema de cuidado externo proveniente del mercado o del Estado está ausente; lo más cercano a esto son las preguntas sobre asistencia a guarderías, estancias infantiles, preescolar (ENUT y ELCOS) y educación básica para la población en edad de asistir, incluidas en casi todas las encuestas, excepto en la ENASEM; el realizado por instituciones públicas y privadas

se captura en la ENGASTO y ENESS, mientras que lo referente a los cuidadores pagados se capta en la ENUT, ENGASTO, ENESS y ENASEM y la contratación de servicio doméstico, en la ENUT, ENOE, ENIGH y ELCOS. Es necesario mencionar que existe variación entre las encuestas en cuanto a la identificación de la disponibilidad del servicio, el número de personas y el tiempo contratados, así como los gastos para esta labor. En este último punto, solo la ENGASTO, ENESS y ENIGH capturan el monto gastado, pero son pocos rubros propios del cuidado, como: pago de enfermeras y personas que cuidan, el gasto en estancias infantiles y cuidado de niños; aunque la ENIGH y ENASEM registran gastos en

salud, estos reflejarían un tipo de cuidado coyuntural, por ejemplo, los que se hacen por hospitalización. La ENESS se aproxima a la demanda de cuidado infantil no satisfecha, un área importante y emergente, al preguntar si las madres llevan a sus hijos al lugar de trabajo o si los dejan solos.

Para abordajes como el trabajo de cuidado no remunerado asociado con las desigualdades socioeconómicas, su valor en el mercado laboral e implicaciones económicas, se puede decir que la mayoría de las fuentes de información estadística captura el tiempo dedicado al cuidado por los integrantes del hogar en edad de trabajar (12 años y más), así como sus características sociodemográficas (sexo, escolaridad, parentesco y estado conyugal, entre otras), lo cual permite crear perfiles en torno a su participación e intensidad con la que intervienen en esta actividad. Además, otras encuestas capturan características de la vivienda, ingresos y transferencias económicas recibidas en el hogar (EI, ENUT, ENIGH y ELCOS), lo cual permite realizar análisis más detallados a partir de estratos socioeconómicos.

Las implicaciones del cuidado en la actividad económica de quien cuida se sustentan en las características del mercado laboral; en la mayoría de las fuentes de información se pregunta sobre la participación en este, posición en la ocupación y jornada de trabajo (número de horas); otras capturan datos acerca de ingresos y prestaciones laborales (ENUT, EI, ENIGH y ENOE). En este tema, la ENOE resulta de vital importancia, pues se ha especializado en identificar el trabajo de mercado asociado a la informalidad o precariedad, así como del desempleo desalentado, aspectos importantes para el estudio de las implicaciones del cuidado en la actividad económica de mujeres y hombres. Además, cada cuatro años se le agrega el módulo de la ENESS, con el que se pueden realizar análisis más profundos en cuanto a la demanda y oferta de cuidado, así como el gasto en cuidados en que incurren los hogares.

En síntesis, existe una diversificación de los instrumentos de información que abordan el tema

de cuidado; la ENUT es el que tiene la mayor cobertura conceptual con respecto al resto de las fuentes de datos estadísticos en México; es posible analizar el tema de cuidados en diferentes niveles y relacionarlo con otras temáticas.

Estimación del tiempo de cuidado: módulos en encuestas versus encuestas especializadas

En las secciones previas se han mencionado aspectos de las encuestas especializadas de uso del tiempo y de las que incluyen módulos sobre cuidados con la finalidad de revisar qué tan semejantes o diferentes son las estadísticas obtenidas acerca del tiempo de cuidado en uno y otro tipo de fuente de información. En este apartado se realiza un ejercicio comparativo del tiempo de cuidado reportado en la ENUT 2014 y en la ENOE (tercer trimestre del 2014).

Dada la naturaleza continua de la ENOE, se puede examinar la evolución de la oferta de cuidado disponible para los integrantes del hogar. La Encuesta captura las categorías en las que hay un mayor consenso (cuidado directo imprescindible, de gestión y apoyo al cuidado) y no hace distinción del otorgado para grupos dependientes de cuidado o si este es o no para los miembros del hogar, al menos para el cuidado directo imprescindible y el apoyo al cuidado.

En el cuadro 4 se observa que la ENUT y ENOE coinciden en tres de las ocho categorías. Los cuidados directos imprescindibles que en la segunda se captura a través de una pregunta sintética, en la primera se hace con seis reactivos detallados. Los indirectos de gestión se cubren en la ENOE con una pregunta sintética y en la ENUT, con ocho detalladas. El apoyo al cuidado en la ENOE se cubre con una sintética y en la ENUT, mediante 12 detalladas. A diferencia de la ENUT, la ENOE no captura las categorías de cuidado socialmente creado, supervisión (al pendiente), en beneficio de, de guardia y de autocuidado.

En el caso de la ENUT se especifica claramente el cuidado a integrantes del hogar; captura de manera independiente el no remunerado otorgado a otros hogares y el recibido de esos. En la ENOE, esta precisión está en el cuidado de gestión y el apoyo al cuidado; sin embargo, no es claro en el caso del imprescindible, pues no se especifica a qué miembros o integrantes del hogar se les otorga este, por lo que podría estar capturando el cuidado imprescindible otorgado a otros hogares.

Dadas estas diferencias en el detalle de preguntas y en el número de categorías de cuidado abordadas, es conveniente examinar si existe una posible subestimación en este tema de la ENOE respecto a la ENUT. Por ello, en el cuadro 6 se presenta la participación que tienen hombres y mujeres en el trabajo de cuidado según la ENUT y ENOE 2014. Se considera población participante en este a aquella que acumuló/reportó al menos una hora al trabajo de cuidado durante la semana de referencia. Para ambos sexos, en la ENUT se observa un porcentaje mayor de participación en el trabajo de cuidado que en la ENOE y esta diferencia es mayor en los hombres que en las mujeres. El de ellos es de 80.7 en la ENUT, mientras que en la ENOE es de 64.9. El de ellas son de 97.2 y 95.8, respectivamente. El hecho de que la ENUT registre mayores porcentajes de participación en el trabajo de cuidado que la ENOE puede deberse al

mayor nivel de detalle en las preguntas y la mayor cobertura de las categorías de la primera respecto a la segunda fuente de información.

Además, se ha documentado que la participación en el trabajo de cuidado de hombres y mujeres debe complementarse con la intensidad con la que participan, pues ellos lo hacen con un bajo nivel de horas en relación con ellas (Orozco-Rocha, 2017; Pedrero, 2014). Por ello, en los cuadros 7.a y 7.b se presenta el tiempo dedicado a cada categoría de esta actividad con dos indicadores, el primero se refiere al promedio de horas semanales entre el total de la población de 12 años y más de edad, mientras que el segundo, al promedio de horas semanales entre la población participante en el trabajo de cuidado (la que dedicó al menos una hora a este).

En ambos casos se puede observar que el promedio de horas reportado es mayor en la ENUT que en la ENOE. En cuanto a los hombres, la primera reporta alrededor de dos horas más a la semana que la segunda y en las mujeres, alrededor de cinco. En términos de intensidad de cuidado, en la ENOE con respecto a la ENUT, la subestimación en ellas es mayor que para ellos, contrario a lo observado en la tasa de participación de trabajo de cuidado que afecta más a los hombres que a las mujeres.

Cuadro 6

Distribución porcentual de la población de 12 años y más de edad por condición de trabajo de cuidado, según sexo y encuesta, 2014

Trabajo de cuidado	ENUT ^{1/}		ENOE ^{2/}	
	Hombres	Mujeres	Hombres	Mujeres
No cuida	19.3	2.9	35.2	4.2
Cuida	80.7	97.2	64.9	95.8
Total	100.0	100.0	100.0	100.0

^{1/} Incluye seis categorías de cuidado: imprescindible, socialmente creado, gestión, en beneficio de, de guardia y apoyo al cuidado, más el cuidado no remunerado a otros hogares.

^{2/} Incluye tres categorías de cuidado: imprescindible, gestión y apoyo al cuidado.

Fuente: elaboración propia con base en datos de la ENUT 2014 y ENOE 2014 (www.beta.inegi.org.mx).

Cuadro 7.a

Promedio de horas semanales dedicadas al trabajo de cuidado de la población de 12 años y más de edad

Categoría de cuidado	ENUT		ENOE	
	Hombres	Mujeres	Hombres	Mujeres
a. Imprescindible	1.1	3.8	1.5	6.6
b. Socialmente creado	0.5	1.3		
c. Gestión	0.7	1.2	0.1	0.4
d. En beneficio de	0.0	0.2		
e. En guardia	0.0	0.1		
f. Apoyo al cuidado	5.1	23.6	4.0	20.5
g. Cuidado a otros hogares	0.7	2.0		
Total	8.1	32.0	5.6	27.5
∑ de categorías coincidentes	6.8	28.5	5.6	27.5
% (a+c+f)/Total	84.4	89.0		

Fuente: elaboración propia con base en datos de la ENUT 2014 y ENOE 2014 (www.beta.inegi.org.mx).

Cuadro 7.b

Promedio de horas semanales dedicadas al trabajo de cuidado de la población participante de 12 años y más de edad

Categoría de cuidado	ENUT		ENOE	
	Hombres	Mujeres	Hombres	Mujeres
a. Imprescindible	1.4	3.9	2.3	6.9
b. Socialmente creado	0.6	1.3		
c. Gestión	0.8	1.2	0.2	0.5
d. En beneficio de	0.1	0.2		
e. En guardia	0.0	0.1		
f. Apoyo al cuidado	6.3	24.3	6.1	21.3
g. Cuidado a otros hogares	0.8	2.0		
Total	10.0	33.0	8.6	28.7
∑ de categorías coincidentes	8.4	29.4	8.6	28.7
% (a+c+f)/Total	84.4	89.0		

Fuente: elaboración propia con base en datos de la ENUT 2014 y ENOE 2014 (www.beta.inegi.org.mx).

Dadas estas subestimaciones, se realizan las siguientes precisiones: a) las diferencias se reducen de manera sustantiva si se considera solo sumar las tres categorías coincidentes entre la ENUT y ENOE (cuidados imprescindibles, de gestión y apoyo al cuidado), lo cual lleva a suponer que hay coincidencias en el número promedio de horas reportadas en ambas encuestas para dichas categorías; b)

parece que existe una pequeña subestimación del trabajo de cuidado en la ENOE respecto a lo reportado por la ENUT, y esta deriva, principalmente, por las categorías faltantes en la ENOE. Por ejemplo, si la suma de las horas promedio de las coincidentes se divide entre la media total de las dedicadas al cuidado en la ENUT se puede observar que las tres categorías incluidas en la ENOE representan en la

ENUT 84 % para los hombres y 89 % para las mujeres. Por ello, se destaca que la ENOE es un buen instrumento para estudiar el tema en un contexto más actual al de la ENUT 2014 y considerando las limitaciones antes descritas para ambas encuestas.

Conclusiones y recomendaciones

A la par que se ha desarrollado el concepto de cuidado, las fuentes nacionales e internacionales de información han avanzado para tratar de capturar las actividades relacionadas con el trabajo de cuidado, en particular las encuestas en hogares sobre el uso del tiempo, las cuales permiten explorar la relación entre el trabajo de cuidado y diferentes aspectos de la vida de los cuidadores, entre ellos la participación en la actividad económica de mujeres y hombres.

En la región de América Latina, México se destaca por tener la mayor producción de mediciones de uso del tiempo y dentro de este, el cuidado; además, es palpable el compromiso adquirido para incorporar mediante módulos el tema en las encuestas regulares, por ejemplo, en la ENOE y ENIGH, que permite su estudio de manera continua y sintética.

Por los resultados obtenidos al comparar la ENUT y ENOE, se puede concluir que ambas encuestas capturan adecuadamente el tiempo dedicado al trabajo de cuidado. Si bien es cierto que existen diferencias en sus estimaciones al respecto, también lo es que esas desigualdades no son tan grandes y que al analizar las tres categorías principales de cuidado (imprescindible, gestión y apoyo al cuidado), estas son menores.

Con la finalidad de mejorar las fuentes de información disponibles en México, a continuación se presentan recomendaciones que, de llevarse a cabo, contribuirán a un mejor entendimiento y abordaje, así como a generar líneas de investigación sobre el tema de trabajo de cuidado no remunerado. Además de la amplia cobertura conceptual, que incluya las categorías de cuidado y

a los grupos dependientes también se requieren datos con las características asociadas al mercado laboral y a los contextos familiar y sociodemográfico de los individuos, así como de la oferta de cuidado a la que acceden los hogares para hacer frente a sus demandas a través del mercado, el Estado y la comunidad. Se sugiere que las fuentes de información que abordan el tema y que incluyen las dimensiones laborales, sociodemográficas y familiares las preserven, pues permiten analizar la interrelación que existe entre ellas y su potencial es mayor.

También, se recomienda dar continuidad a las fuentes de información como la ENUT, esto con el propósito de examinar la evolución de la distribución del trabajo de cuidado no remunerado, sus implicaciones en la actividad económica de mujeres y hombres y la posibilidad de visualizar una redistribución de este a la luz de políticas públicas implementadas hacia una OSC más equitativa e incluyente. Es necesario darle continuidad metodológica y temporal, mantenerla vigente, pues es la encuesta, en México, que proporciona la información más completa para el estudio de la OSC y sus implicaciones económicas, y es de suma importancia para el cálculo de las cuentas satélite del trabajo no remunerado, entre otros beneficios. Un punto importante es mantener el detalle y separación de las categorías del cuidado, pues con ello se posibilita profundizar en las implicaciones que estas tienen para los individuos y los hogares, y constatar para el caso de México lo que dice la literatura: que existe una segregación en las categorías del cuidado por género y nivel socioeconómico.

Otra de las recomendaciones tiene que ver con el actual proceso de envejecimiento demográfico en México, fenómeno que cambiará la oferta y demanda de cuidado, y entonces es necesario reflexionar y procurar el equilibrio en la cobertura conceptual y en la cantidad de preguntas que se incluyan en los instrumentos de obtención de la información entre los grupos dependientes, en particular entre la población infantil y la adulta mayor.

Finalmente, en México existen datos estadísticos consistentes, generados por diversas encuestas y con diferentes perspectivas del tema de cuidado, y se recomienda hacer ajustes en estas para lograr un panorama más integral, pensar en la obtención de la información como un continuo, agregar módulos que permitan mayor profundidad de análisis y realizar la homogeneización de determinados temas y conceptos entre fuentes de información acordes con el actual concepto de cuidado y tomando en cuenta sus diversos abordajes, por ejemplo, las actividades simultáneas o multitareas.

Esta revisión de la cobertura conceptual, de los grupos dependientes y de la suficiencia de información para abordar el tema de trabajo de cuidado no remunerado es pertinente, pues permite analizar su impacto sobre la contribución económica de mujeres y hombres, en virtud de que la OSC que se presenta en México tiene escasa presencia del Estado para proveer cuidados, una alta participación de la familia en el trabajo de cuidado y con la desigual distribución de esta actividad entre ellas y ellos, además de que los sistemas de protección social están asociados con el empleo formal, hay una baja intervención económica de las mujeres y una cantidad importante de ellas insertas en empleos precarios, con brechas salariales persistentes.

Se espera que el conocimiento que se genera de las investigaciones basadas en las encuestas sobre el uso del tiempo y en particular acerca del trabajo de cuidado contribuya a la creación e implementación de políticas públicas con una perspectiva de equidad de género, que procuren igualdad en la distribución de las tareas de cuidado y que garanticen acceso y condiciones de participación económica en igualdad de condiciones para mujeres y hombres.

Fuentes

- Aguirre, R.; K. Batthyány; N. Genta y V. Perrotta. "Los cuidados en la agenda de investigación y en las políticas públicas en Uruguay", en: *Íconos. Revista de Ciencias Sociales*. 50, 2014, pp. 43-60.
- Aguirre, R. y F. Ferrari. "Las encuestas sobre uso del tiempo y trabajo no remunerado en América Latina y el Caribe", en: *Caminos. Serie Asuntos de Género*. Núm. 122. Santiago de Chile, CEPAL-Naciones Unidas, 2014, pp. 1-99.
- Arriagada, I. "La crisis de cuidado en Chile", en: *Revista de Ciencias Sociales*. XXIII, 2010, pp. 58-67.
- Batthyány, K. "Género, cuidados familiares y uso del tiempo", en: *El Uruguay desde la sociología*. VII. Montevideo, FCS-Universidad de la República, 2008, pp. 177-198.
- Batthyány, K.; N. Genta y V. Perrotta. "Las representaciones sociales del cuidado infantil desde una perspectiva de género. Principales resultados de la Encuesta Nacional sobre Representaciones Sociales del Cuidado", en: *Papers. Revista de Sociología*. 99(3), 2014, pp. 335-354.
- _____. "El aporte de las familias y las mujeres a los cuidados no remunerados en salud en Uruguay", en: *Revista Estudios Feministas*. 25, 2017, pp. 187-209.
- Benach, J.; A. Vives; M. Amable; C. Vanroelen; G. Tarafa y C. Muntaner. "Precarious Employment: Understanding an Emerging Social Determinant on Health", en: *Annual Review of Public Health*. 35, 2014, pp. 229-253.
- Budig, M. J. y N. Folbre. "Activity, proximity, or responsibility? Measuring parental childcare time", en: Folbre, N. y M. Bitman (eds.). *Family time: The social organization of care*. New York, Routledge, 2004, pp. 51-68.
- Carrasco, C. "El tiempo más allá del reloj: las encuestas de uso del tiempo revisitadas", en: *Cuadernos de Relaciones Laborales*. 34(2), 2016, pp. 357-383.
- Durán, M. Á. "Las cuentas del cuidado", en: *Revista Española de Control Externo*. XX(58), 2018, pp. 57-90.
- Espino, A. "Economía feminista: enfoques y propuestas", en: Sanchis, N. (ed.). *Aportes al debate del desarrollo en América Latina. Una perspectiva feminista*. Buenos Aires, Red de Género y Comercio, 2011, pp. 8-19.
- Esquivel, V. "La economía del cuidado: un recorrido conceptual", en: N. Sanchis, (ed.). *Aportes al debate del desarrollo en América Latina. Una perspectiva feminista*. Buenos Aires: Red de Género y Comercio. 2011, pp. 20-30.
- _____. "Cuidado, economía y agendas políticas: una mirada conceptual sobre la 'organización social del cuidado' en América Latina", en: Esquivel, V. (ed.). *La economía feminista desde América Latina*. Santo Domingo, ONU Mujeres, 2012, pp. 141-189.
- _____. "El cuidado: de concepto analítico a agenda política", en: *Nueva Sociedad*. 2015, pp. 63-74
- Folbre, N. "Measuring care: gender, empowerment, and the care economy", en: *Journal of Human Development*. 7(2), 2006a, pp. 183-199.
- _____. "Rethinking the child care sector", en: *Community Development*. 37(2), 2006b, pp. 38-52.
- _____. "Medir los cuidados: género, empoderamiento y la economía del cuidado", en Carrasco, C.; C. Borderías y T. Torns (eds.). *El trabajo de*

- cuidados. Historia, teoría y políticas*. Madrid, CIP-Ecosocial-CATARATA, 2011, pp. 278-304.
- Folbre, N. y J. Yoon. "Economic development and time devoted to direct unpaid care activities", en: *NRISD Flagship Report on Poverty*. Geneva, United Nations Research Institute for Social Development, 2008a.
- _____. "El cuidado de los niños: lo aprendido mediante encuestas sobre el uso del tiempo en algunos países de habla inglesa", en: Suárez, R. y L. Jara (eds.). *La economía invisible y las desigualdades de género*. Washington, DC, OPS-CEPAL-CSIC, 2008b, pp. 205-228.
- García Guzmán, B. "El trabajo doméstico y de cuidado: su importancia y principales hallazgos en el caso mexicano", en: *Estudios Demográficos y Urbanos*. 34(2), 2019, pp. 237-267.
- Leiva, S. "Organización social del cuidado en Bolivia y Chile: Estado y ciudadanía", en: *Revista Austral de Ciencias Sociales*. 28, 2015, pp. 61-81.
- Marco, F. "La utilización de las encuestas de uso del tiempo en las políticas públicas", en: *Serie Mujer y Desarrollo*. Núm. 119. Santiago de Chile, CEPAL-Naciones Unidas, 2012, pp. 1-45.
- Nieves, M. y C. Robles. "Políticas de cuidado en América Latina. Forjando la igualdad", en: *Serie Asuntos de Género*. Núm. 140. Santiago de Chile, CEPAL-Naciones Unidas, 2016, pp. 1-79.
- Orozco-Rocha, K. "Reparto del trabajo doméstico y extradoméstico al interior de los arreglos familiares urbanos", en: Nájera, J.; B. García y E. Pacheco (eds.). *Hogares y trabajo en México en el siglo XXI*. Ciudad de México, El Colegio de México, 2017, pp. 277-326.
- Pedrero, M. "Valor económico del trabajo doméstico en México. Aportaciones de mujeres y hombres, 2009", en: *Cuaderno de Trabajo Núm. 20*. Ciudad de México, Instituto Nacional de las Mujeres, 2010, pp. 1-115.
- _____. "Importancia del trabajo no remunerado: su medición y valoración mediante las encuestas de uso del tiempo", en: García B. y E. Pacheco (eds.). *Uso del tiempo y trabajo no remunerado en México*. Ciudad de México, COLMEX-ONU MUJERES-INMUJERES, 2014.
- Rodríguez, C. y G. Marzonetto. "Organización social del cuidado y desigualdad: el déficit de políticas públicas de cuidado en Argentina", en: *Revista Perspectivas de Políticas Públicas*. 4(8), 2015, pp. 103-134.
- Rodríguez, C. y D. Castro. "Discriminación salarial de la mujer en el mercado laboral de México y sus regiones", en: *Economía, Sociedad y Territorio*. XIV(46), 2014, pp. 655-686.
- Suh, J. "Measuring the 'Sandwich': Care for Children and Adults in the American Time Use Survey 2003-2012", en: *Journal of Family and Economic Issues*. 37, 2016, pp. 197-221.
- Suh, J. y N. Folbre. "Valuing unpaid child care in the U.S.: A prototype satellite account using the American Time Use Survey", en: *Review of Income and Wealth*, 62(4), 2016, pp. 668-685.
- Sunkel, G. "Regímenes de bienestar y políticas de familia en América Latina", en: Arriagada, I. (ed.). *Familias y políticas públicas en América Latina. Una historia de desencuentros*. Santiago de Chile, CEPAL-Naciones Unidas, 2007, pp. 171-186.

Ser entrevistadora o entrevistador.

La recolección de datos como un oficio, el aprendizaje de un saber hacer

To Be a Female Interviewer or a Male Interviewer.

Data Compilation as a Craft, the Learning of a Know-How

Eleonora Nuricumbo Rivera,* Candi Uribe Pineda y Adriana Marcela Meza Calleja*****

* Universidad Autónoma de Querétaro (UAQ), ele_nr@yahoo.com.mx ** UAQ, upcandi@outlook.com *** Universidad Michoacana de San Nicolás Hidalgo, adime5@hotmail.com

Nota: durante el desarrollo de este artículo se utilizarán los términos entrevistadora y entrevistador con la intención de acentuar la participación de ambos sexos en el trabajo de recolección de datos.



Arturo López Jiménez

Este artículo tiene como objetivo visibilizar y problematizar el oficio de la recolección de datos que realizan de manera cotidiana las entrevistadoras y los entrevistadores del Instituto Nacional de Estadística y Geografía en un contexto físico, social, económico y político complejo. Se pretende redescubrir el trabajo que realizan, más allá de una simple captación de datos, es decir, pensarlo como un oficio artesanal que ellas(os) mismas(os) desde la práctica van construyendo y perfeccionando para llevar a buen término su jornada laboral, donde ponen en marcha su capacidad de resistencia, adaptación y creatividad para construir estrategias que les permitan hacer frente a lo no previsto. Para este trabajo —realizado desde una metodología cualitativa—, participaron entrevistadoras y entrevistadores del Instituto de Querétaro, San Luis Potosí y la Ciudad de México.

Palabras clave: entrevistadora; entrevistador; oficio; experiencia; estrategias.

Recibido: 23 de julio de 2019.

Aceptado: 6 de enero de 2020.

Introducción

Una de las funciones del Instituto Nacional de Estadística y Geografía (INEGI)¹ es generar información estadística sobre el país. Para llevar a cabo esta labor, requiere del despliegue de varios procesos, desde el diseño conceptual de los instrumentos de captación específicos para cada encuesta o censo, hasta la puesta en marcha de la actividad operativa que consiste en la recolección de datos provenientes de la sociedad mexicana acerca de diversas temáticas económicas y sociodemográficas.² Para ejecutar esta última actividad, el área de encuestas del Instituto necesita de la participación de cientos de entrevistadoras y entrevistadores que salen a las calles a captar datos de la ciudadanía, ya sea en hogares o en empresas, sobre diversos temas de interés que servirán como parámetros de medición y toma de decisiones para el país.³

1 Organismo autónomo del Estado mexicano desde el 2008.

2 Abarca temas de salud, cultura, violencia, tecnología, discriminación, por mencionar algunos.

3 A nivel nacional se contrataron, por ejemplo, 551 entrevistadoras(es) para la Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) 2013 y 481 para el Módulo de Condiciones Socioeconómicas (MCS) 2015.

The aim of this article is to make visible and problematize the profession of data collection carried out on a daily basis by the interviewers of the National Institute of Statistics and Geography in a complex physical, social, economic and political context. The objective is to rediscover the work they do, beyond simple data collection, that is, to think of it as a craft that they themselves build and perfect from practice in order to carry out their working day successfully, where they put into practice their capacity for resistance, adaptation and creativity to build strategies that allow them to deal with the unexpected. For this work —carried out using a qualitative methodology— interviewers from the Institute in Queretaro, San Luis Potosi and Mexico City participated.

Key words: interviewer; trade; experience; strategies.

El trabajo cotidiano de recolección de datos que realizan las entrevistadoras y los entrevistadores del INEGI se desarrolla fuera del centro laboral, es decir, en la calle. Este personal de campo contratado por el Instituto tiene como encargo salir, con cartografía en mano, a recorrer avenidas, transitar por caminos o veredas, atravesar ríos o cruzar montañas de todo el país con la finalidad de llegar a cada una de las viviendas y negocios por visitar para levantar sus entrevistas asignadas, realizando una serie de preguntas a los(las) informantes (población que proporciona información) a través de los instrumentos de captación preestablecidos por el Instituto.

En este proceso productivo de recolección de datos existen dos figuras que se introducen y participan de distintas formas durante el desarrollo de esta actividad: uno es el cliente (que puede ser una institución pública o privada que solicita los servicios del INEGI para realizar el diseño del instrumento a utilizar, levantamiento de información y procesamiento de datos), con quien la trabajadora

o el trabajador de campo no tiene contacto directo; es más, la mayoría de las veces no lo conoce; sin embargo, sus exigencias⁴ inciden en su proceso de trabajo. Por otra parte, está la figura del(la) informante, a quien el INEGI elige de forma aleatoria⁵ para participar en el proceso de captación de información.

La relación establecida con esta figura es directa y determinante para realizar y concluir el proceso de trabajo de la entrevistadora o el entrevistador. Desde la interacción, el(la) informante y las entrevistadoras o los entrevistadores conviven en diferentes circunstancias: es sana con aquellas personas que se muestran participativas a responder las preguntas y brindar la información requerida; sin embargo, el personal del Instituto también se enfrenta a malos tratos, agresiones verbales y físicas por parte de algunos(as) informantes poco dispuestos(as) o molestos(as) por participar.

Otro aspecto relevante a considerar en este panorama complejo es que se trata de un trabajo que no está anclado a un territorio específico, por ejemplo, una oficina o un establecimiento, mucho menos a una fábrica; si bien su destino es cada una de las viviendas o empresas por visitar, las entrevistadoras y los entrevistadores se mantienen en movimiento constante: la calle y sus riesgos terminan siendo su espacio de trabajo.

Dado que el hábitat laboral de la entrevistadora y del entrevistador se ubica en las áreas urbanas y rurales de todo el territorio nacional, es importante considerar que las situaciones políticas y sociales de cada región del país inciden directamente en su trabajo, por ejemplo, los bloqueos de carreteras derivados de manifestaciones de la población, la restricción de acceso a zonas ocupadas por la delincuencia o el narcotráfico, las amenazas de extorsión o la exposición a secuestros, por mencionar algunas situaciones irregulares. Mención específica requieren las inclemencias climáticas (inundacio-

nes o deslaves, entre otras) que pueden imposibilitar el acceso a las viviendas y empresas por visitar.

Estos elementos mencionados se presentan para dar cuenta del panorama en el cual se inserta el oficio de entrevistadora-entrevistador del INEGI, donde se observan condiciones laborales específicas que incluyen requerimientos y necesidades (Noriega, 1993; Noriega y Laurell, 2000) propias para ejecutar la actividad que se marca en las cargas laborales y en las jornadas de trabajo extenuantes, así como los riesgos inherentes al proceso productivo de recolección de datos, lo que nos permite caracterizar el trabajo de la entrevistadora o el entrevistador del Instituto como complejo, que incluye diversas aristas de este hecho social escasamente reconocido.⁶

Esta labor la ubicamos en el sector *Servicios*, como un trabajo propio de la heterogeneidad laboral contemporánea, desde la perspectiva de trabajo no clásico⁷ en México. De ahí que se toman conceptos⁸ construidos dentro del ámbito de los estudios laborales con la finalidad de mirar esta experiencia desde una perspectiva de concepto ampliado del trabajo como proceso de interacción y creación simbólica (De la Garza, 2011: 57). Posibilita comprender que, para desempeñar la actividad de entrevistadora-entrevistador, como ellas(os) mismas(os) lo comentan en sus relatos, se requiere de cierta pericia para la ejecución del proceso produc-

6 El estado del arte muestra estudios sobre entrevistadoras(es) de un centro de investigaciones en España y un par de estudios en México sobre trabajadoras(es) de campo en censos del INEGI, los cuales se interesan por propiciar prácticas que incentiven la calidad de la producción, así como la disminución de las deficiencias dentro del proceso de recolección de datos por muestreo. Es por lo anterior que la investigación que se plantea cobra relevancia, dado que no existen estudios que se interesen por las entrevistadoras y los entrevistadores desde un enfoque no productivo. En otras palabras, desde una perspectiva que priorice la dimensión subjetiva de las(los) trabajadoras(es); que dé cuenta de su realidad, de la forma como vivencian lo que implica el trabajo de calle, el riesgo y las estrategias creadas por ellas(os) como recolectoras(es) de datos para llevar a buen término su carga laboral. Ver la tesis de maestría *El proceso de trabajo de los entrevistadores del INEGI: riesgos psicosociales y estrategias de acción*, en introducción, apartado *Estado de la cuestión en Iberoamérica* (ver nota a pie de página 10).

7 En diferenciación al trabajo industrial clásico.

8 El concepto ordenador base es el proceso de trabajo, ya que de él se desprenden los demás; es en el proceso donde miramos las exigencias propias de este trabajo estudiado, las estrategias que las(los) trabajadoras(es) construyen para solventar lo no prescrito, los riesgos del trabajo de calle y la subjetividad como construcción del sentido del trabajo con sus particularidades. Los conceptos ordenadores no son unidades aisladas entre sí, estos se encuentran en una interacción constante, no se piensa el uno sin el otro, están relacionados tanto con el concepto ordenador base como entre ellos; es decir, cada uno tiene su dimensión particular, pero no por ello dejan de articularse entre sí.

4 Entendidas como necesidades o requerimientos propios del diseño de la encuesta.

5 En el caso de los censos, se entrevista al total del universo.

tivo. Además de tener o crear y desarrollar habilidades para *hacer bien*⁹ la recolección de datos, se necesita tener gusto y compromiso hacia el trabajo de entrevistadora-entrevistador.

Los motivos para voltear la mirada hacia las entrevistadoras y los entrevistadores son diversos: primero, se pretende mostrar el trabajo realizado por estos actores laborales que, a simple vista, no se aprecia; sin embargo, la labor cotidiana que realizan como trabajadoras(es) de campo del INEGI posibilita la generación de estadísticas económicas, demográficas y sociales y, con ello, cierta forma de conocimiento estratégico para el desarrollo del país. También, se quiere exponer que es errónea la representación social sobre el trabajo de la entrevistadora o del entrevistador como un empleo *tan sencillo* que cualquier persona pudiera desempeñarlo. Aquí se sostiene que, para ser entrevistadora-entrevistador se requiere poner en práctica todas las destrezas personales adquiridas, así como construir las habilidades específicas que se necesitan para poder llevar a buen término la carga total de trabajo.

Pretendemos que este texto sirva para darle su justo lugar al trabajo de entrevistadora-entrevistador, que no sea visto como uno menor que causa molestia a los(las) demás y que, por ser un trabajo de calle, se demerite o se menosprecie por el resto de la sociedad. La intención es que se conozca, comprenda y difunda la importancia social de la labor que realiza este grupo de trabajadoras(es) de campo del INEGI en el proceso de mejorar el desarrollo nacional.

Con base en lo anterior, esta investigación, desarrollada desde las ciencias sociales —en particular desde la sociología del trabajo—, propone que la recolección de datos llevada a cabo por las entrevistadoras y los entrevistadores es un oficio, ya que es una actividad que requiere de un aprendizaje cotidiano que suma a la experiencia con la

realización de cada una de las entrevistas, pues la entrevistadora o el entrevistador llega a tener tal dominio de sus habilidades innatas y adquiridas en la práctica para mantener interacción con su *medio de trabajo* —el(la) informante—, que se convierte en experta(o) en la recolección de datos a través de la entrevista; es decir, tiene una formación en un *saber hacer* a partir de la experiencia laboral cotidiana. Además, esto lo desarrolla con la convicción de hacer bien su principal tarea, con el compromiso por el trabajo y con satisfacción y orgullo al ejecutar su labor.

Por consiguiente, para comprender la tarea que realizan mujeres y hombres, entrevistadoras-entrevistadores, como la ejecución de un oficio artesanal, se desarrollará, primero, un recorrido histórico sobre la labor de recolección de datos en México; después, se abordará —a manera de relato construido desde la observación no participante— el proceso productivo de captación de información; luego, un apartado sobre la relación existente con el(la) informante como elemento fundamental dentro del proceso de trabajo. Más adelante, se hablará acerca de las destrezas y estrategias de acción construidas, para concluir en esta sección con la propuesta del trabajo como creador de experiencia. Por último, se presentan las reflexiones finales de lo que significa ser entrevistadora-entrevistador desde ese punto de vista.

A través de los datos presentados en este texto se pretende redescubrir el trabajo que realizan estas personas más allá de una simple captación de información, es decir, pensarlo como un oficio artesanal que ellas(os) mismas(os), desde la práctica, van construyendo y perfeccionando para llevar a buen término su jornada laboral. Por ello, se tiene como objetivo visibilizar y problematizar esta labor que realizan de manera cotidiana las entrevistadoras y los entrevistadores del INEGI en un contexto social, económico y político complejo.

El presente artículo se desprende de la investigación titulada *El proceso de trabajo de los entrevistadores del INEGI, riesgos psicosociales y estrategias*

⁹ Para Richard Sennet, hacer bien el trabajo alude al compromiso que la persona trabajadora interioriza para desempeñar su tarea laboral asignada —en este caso— por el espacio laboral, el cual funciona como impulsor para realizar lo indicado, añadiendo el gusto por la ejecución de la actividad y satisfacción al realizarla.

de acción,¹⁰ cuyo argumento central plantea la dimensión subjetiva de la recolección de datos, construida a partir de las condiciones y medio ambiente laboral, así como por los riesgos psicosociales inherentes al trabajo de entrevistadora-entrevistador, quienes ponen en marcha su capacidad de resistencia, adaptación y creatividad para construir estrategias que les permitan hacer frente a lo no previsto, darle sentido a su trabajo y, así, cumplir con las jornadas y los objetivos de la actividad prescrita por la organización que los contrata.

El interés y propósito de la investigación mencionada corre por los caminos de una metodología con enfoque cualitativo que permita comprender el hecho social desde la perspectiva de las entrevistadoras y los entrevistadores (Taylor y Bogdan, 1987). Para ello, se propuso que los datos descriptivos se construyeran en voz de las(los) propias(os) trabajadoras(es), es decir, a partir de su práctica, experiencia y discurso, mirando de manera amplia y profunda el escenario y a sus participantes desde un enfoque holístico. Haciendo un esfuerzo en no dar nada por supuesto, sin reducir el conocimiento a variables y tomando en cuenta que las entrevistadoras y los entrevistadores (sujetos de este estudio), el contexto por el que transitan, así como sus interrelaciones, se han construido en una historicidad específica (Zemelman, 2005).

En consecuencia, se recolectaron datos a través de diversas herramientas metodológicas.¹¹ Se entrevistó a personal de las coordinaciones estatales del INEGI de Querétaro, San Luis Potosí y la Ciudad de México; participaron 31 personas, de las cuales 22 tenían el puesto de entrevistadora-entrevistador, con una trayectoria laboral en el Instituto de entre 10 y 20 años por contratación eventual.

¹⁰ Realizada de enero del 2016 a diciembre del 2017 en el Posgrado en Estudios Multidisciplinarios sobre el Trabajo, de la UAQ. Está alojada para su consulta en el repositorio de la propia Universidad (ri.uaq.mx) y en el centro de información INEGI de San Luis Potosí, en versiones impresa y digital.

¹¹ El trabajo de campo —observación no participante, entrevistas a profundidad a personal del INEGI, análisis documental sobre los orígenes del oficio de entrevistadora-entrevistador y diarios de campo— se realizó de febrero a mayo del 2017.

Elementos que configuran el oficio de entrevistadora-entrevistador

Para ir reconstruyendo el trabajo de la entrevistadora y el entrevistador como un oficio, es necesario conocer algunas dimensiones y elementos que lo forman. En este apartado, primero se presentará una síntesis histórica sobre su génesis, que se liga directamente a la tradición censal de nuestro país, hasta la creación de lo que ahora se conoce como el INEGI. De igual forma, a manera de relato, se presenta el proceso productivo que llevan a cabo las entrevistadoras y los entrevistadores. También, para la comprensión de esta actividad, es importante ubicar el *medio de trabajo* de este proceso productivo de recolección de datos. Lo relevante en este caso es que el medio, a diferencia de otros oficios, por ser un sujeto con decisiones propias, determina el resultado del proceso de trabajo. De ahí que la entrevistadora o el entrevistador ponga en juego todas sus destrezas para llevar a buen término su carga de trabajo, aun cuando el(la) informante no quiera participar en proporcionar los datos solicitados.

Se parte de la idea de que, en las ciencias sociales, el fenómeno estudiado es un hecho social en construcción. No es un objeto dado, ni acabado, sino un proceso dinámico, que se reconstruye de forma continua, que está formado por las manifestaciones de sujetos sociales complejos, que interactúan y se relacionan en planos, tiempos y espacios distintos (Zemelman, 2011).

Por otra parte, consideramos necesario incluir un breve recorrido en torno a la actividad de recolección de información en México. Más allá de conocer su historia en distintos momentos, se pretende mostrar a la entrevistadora o al entrevistador y su oficio como sujeto colocado en un instante no cerrado, es decir, que deviene de un proceso inacabado y continuo, resultante de una articulación de situaciones, de múltiples relaciones con otros sujetos —jefes(as), pares e informantes— y de sus prácticas (Zemelman, 2005).

Asimismo, la historia, a partir del análisis documental, fue de gran utilidad ya que, además de

conocer la configuración originaria del trabajo de recolección de datos desde los primeros conteos hasta las encuestas en hogares y empresas como espacio laboral, ha servido para mirar la transformación que ha tenido la figura de la entrevistadora o el entrevistador, haciéndose presente el asunto del género; es decir, se pudo observar cómo, de ser un oficio que en sus orígenes era realizado por varones, en la actualidad se ha feminizado,¹² pues en su mayoría son mujeres quienes lo desempeñan.¹³

Censora-censor, encuestadora-encuestador, entrevistadora-entrevistador: protagonista de la historia

En México se tiene una tradición censal de la época precolombina a la fecha. Existe el dato de que los primeros censos poblacionales cumplían la función de conocer el número de súbditos y guerreros. Tiempo después, durante la Colonia (siglo XVII), los levantamientos permitían censar el número de vasallos por clase y casta, así como para determinar las obligaciones fiscales de cada poblador (INEGI, 2009a).

Todas las épocas han requerido de información socioeconómica que les permita organizar y controlar su territorio, por lo que se han empleado diversas formas para recolectarla. Durante el mandato de Porfirio Díaz, con el fin de impulsar el progreso nacional, el *Censo Peñafiel* se presentó como una herramienta indispensable de conocimiento y diagnóstico del país. Este evento sentó un precedente esencial para el establecimiento de una tradición que se mantendría en las siguientes décadas. Como primer censo mexicano, captó a la po-

blación presente en cada vivienda, distinguiendo a los(las) residentes y los(las) visitantes de paso, captando básicamente información sociodemográfica (Moreno, 1980).

Junto con los levantamientos censales, se presentó la figura del(la) censor(a) como la persona que registraba los datos de la población; por ejemplo, en el *Censo de Revillagigedo*: "... cuando los subdelegados emprendían el trabajo, distribuían entre los párrocos y justicias de cada pueblo los modelos para la formación de los padrones; curas, párrocos y justicias comenzaban por hacer una larga lista de los habitantes de cada lugar..." (Castro, 1977: 12).

Para el caso del *Censo Peñafiel* de 1895 "... la más importante de las operaciones preparatorias del evento era, elegir y organizar al personal responsable de vigilar y organizar los trabajos censales, de esta elección dependería todo el éxito del censo..." (INEGI, 2011a: 11). Las características que debía tener quien recolectaba los datos se enfatizaban sobre la importancia de que los funcionarios de mayor categoría fueran respetados y obedecidos por la comunidad y que "... los empadronadores tuvieran clara conciencia de su responsabilidad, por lo cual se encomendaba elegir a los empleados públicos más ilustrados, a particulares conocidos por sus virtudes, que conocieran la localidad y fueran identificados por su ilustración, patriotismo, posición social respetable y buena voluntad, con el fin de que inspiraran confianza a los habitantes..." (INEGI, 2011a: 12).

Avanzando en el tiempo, para 1970, con Luis Echeverría Álvarez en la Presidencia y en un contexto de crisis económica, desgaste político y un crecimiento demográfico superior al empleo, se tuvo la idea de realizar estudios que se acercaran a profundidad a la situación económica y social que se vivía en ese entonces. El censo de 1970 se presentó como la oportunidad de obtener información detallada y actualizada sobre la población mexicana, necesaria tanto en la investigación académica —ahora diversificada— como en los diagnósticos socioeconómicos y demográficos en los nuevos marcos de la planeación (INEGI, 2009b).

12 En las últimas décadas, el trabajo de levantamiento de encuestas ha sido ocupado preferentemente por mujeres. En nuestro estudio participaron 18 entrevistadoras contra cuatro entrevistadores. Existen diversas razones estructurales para suponer que este empleo es más favorable para las mujeres que para los hombres, entre las cuales destacan la escasez de ofertas de empleo específicas para mujeres, la flexibilidad del horario, el bajo salario y el carácter temporal del contrato.

13 Miguel Juan Cervera Flores, en ese entonces director general de Estadísticas Sociodemográficas del INEGI, en entrevista para *Notimex*, comentó que, entre el personal de planta y eventual, las entrevistadoras y los entrevistadores sumaron hasta 65 mil personas, en su mayoría jóvenes; 75 %, mujeres; con un sueldo promedio de 8 mil pesos mensuales y un nivel de escolaridad mínimo de preparatoria, aunque había hasta con doctorado (marzo del 2015).

Debido a que de los eventos censales se obtiene información básicamente sociodemográfica y, de manera general, económica u ocupacional, la Dirección General de Estadística planteó el surgimiento de un instrumento que captara datos temáticos de la población de forma más detallada, por lo que, a partir de 1972, con la creación del Sistema de Encuestas Continuas en Hogares, se establecieron oficialmente las encuestas en México (INEGI, 2011a y 2011b) (ver diagrama 1).

El INEGI, como órgano encargado de producir datos estadísticos y geográficos del territorio mexicano, fue determinando la temática de cada evento ante las distintas necesidades de medición del gobierno federal. Años después, con la obtención de su autonomía, estableció vínculos laborales con dependencias gubernamentales e instituciones privadas (nacionales e internacionales) para realizar la generación de datos estadísticos sobre temas particulares del interés de los propios contratantes.

Con las encuestas temáticas es donde se configuró el trabajo de entrevistadora-entrevistador, cuya principal tarea es la recolección de datos de la población sobre diversos temas de interés nacional. Para ello, recorren calles, caminos y veredas con mochila al hombro y cartografía en mano con el objetivo de localizar a la población seleccionada para participar en cada uno de los eventos (encuestas) que organiza el Instituto.

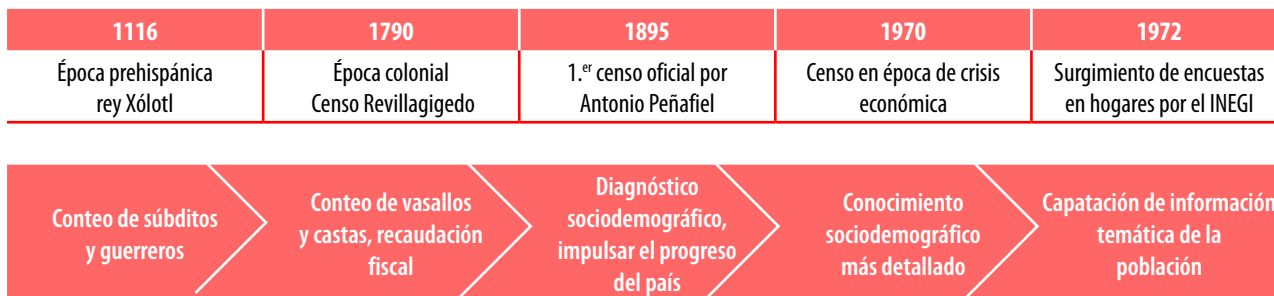
Proceso de trabajo

Cuarto día de levantamiento de un periodo de tres semanas; son las ocho y media de la mañana. De camino a su área de trabajo, Paola me comenta¹⁴ sobre las complicaciones que ha tenido en su trabajo en campo, entre ellas, está la zona que le asignaron: es de difícil acceso porque hay muchos condominios con caseta de vigilancia que no permiten el paso tan fácilmente, "... hay que hacerse amiga de los guardias para que te avisen cuándo llega tu informante..."; dice Paola; además, la mayoría de los(las) informantes que debe visitar se encuentran en un horario nocturno, pues llegan a casa después de las siete u ocho de la noche, pero lo que más le preocupa ahora es la distribución de su tiempo. Lleva ocho años trabajando como entrevistadora por contrato eventual, sin embargo, es la primera vez que lo hace desempeñando de manera simultánea la labor de crianza como madre; dice: "... antes podía andar en la calle trabajando todo el día, no me tenía que preocupar por regresar a mi casa a una hora determinada...", pero ahora, con su niño de un año cuatro meses, le está costando adaptarse para cumplir con el empleo y la maternidad, ya no puede hacerlo como antes; ahora tiene que acomodar su tiempo de otra for-

¹⁴ La narración presentada se escribe en primera persona, pues es producto del trabajo etnográfico realizado durante febrero del 2017 por parte de la estudiante-investigadora-coautora de este documento. Debido a la confidencialidad, se utiliza un pseudónimo elegido por la propia entrevistadora.

Diagrama 1

Cronología de eventos censales en México hasta el surgimiento de encuestas como espacio de trabajo de las entrevistadoras y los entrevistadores del INEGI



Fuente: elaboración propia con datos del INEGI.

ma, de manera que le alcance para atender varias cosas, "... el trabajo de entrevistadora, el niño, la comida, los quehaceres del hogar...".

Llegamos a la zona de trabajo; mientras avanzo despacio en el auto, ella saca sus controles de la mochila, el listado de viviendas y la cartografía, y me dice: "¡Detente ahí!, en lo que veo cuáles son las manzanas a las que vamos"; entretanto, continúa platicando sobre su adaptación al trabajo, ahora como madre, y saca de su mochila el chaleco y su credencial, se los pone, toma su *meebox*¹⁵ (dispositivo móvil), lo enciende y dice: "Este sistemita está mejor que el de la ENVIPE", sonrío y se muestra gustosa por la facilidad del manejo del sistema de esta encuesta; enseguida, saca algunos trípticos que contienen información de la encuesta que levanta para, al término de la entrevista, dárselos al(la) informante, pero antes los coloca en la bolsa de su chaleco.

Acto seguido, abre la puerta del auto y me pide que la espere en lo que va a la vivienda donde levantará los datos, "... es esa café de ahí enfrente...", dice; observo que, mientras va caminando, abre su *meebox* y prepara el cuestionario que debe aplicar en ese lugar. Previo a esto, platica que el sistema le muestra toda su carga de trabajo (el total de viviendas que le corresponde levantar) y ella debe seleccionar una, de acuerdo con sus listados y comenta que esto lo hace así "... para ahorrar tiempo y para que cuando llegue con el informante empiece directo con las preguntas...", para evitar que se desespere por la tardanza del aparato y el sistema.

Paola toca la puerta y pregunta por su informante; la persona que se asoma por la ventana le indica que no está, le pide que vuelva más tarde. Al parecer salió hacia un lugar cercano y no tardará en regresar.

Antes de avanzar a la siguiente vivienda por visitar, ubica de nuevo en su cartografía hacia dónde debe dirigirse. Le acompaño caminando; me

quedo a unos metros de distancia. Esta vez, veo que cuando toca a la puerta y se asoma la persona, Paola se muestra entusiasta, se presenta y pregunta por el(la) informante que va a entrevistar, le abren y la invitan a pasar. Yo espero afuera; luego de 10 minutos, sale sonriente porque después de tres visitas, en distintos momentos, en esta ocasión tiene la entrevista completa.

Como lo menciona Victoria Novelo en el prólogo de *Oficios urbanos tradicionales* (Palacios, 2014:11), el proceso de trabajo es lo fundamental de la labor, es por ello que este apartado muestra un fragmento de lo que es el proceso productivo de la recolección de datos que realizan las entrevistadoras y los entrevistadores del INEGI, así se descubren sus habilidades, conocimientos técnicos, manejo de los materiales y control de sus tiempos y movimientos. Con lo anterior, es preciso enfatizar que, para producir datos mediante el trabajo inmaterial que realizan las recolectoras y los recolectores de datos, se echa a andar un proceso de trabajo complejo en el cual intervienen destrezas, circunstancias, contextos sociales específicos y un(a) informante que imprime un rumbo sustancial al proceso de trabajo.

El propósito de este breve relato también es mostrar a la recolección de datos como un proceso complejo, dado que intervienen distintos elementos de diferente orden: los factores sociales e institucionales —mencionados en la introducción de este texto— y los que atañen en específico a quienes trabajan en esa labor. Para realizar esta actividad, se necesita, entre otras características, brindar confianza, ser observadora-observador del entorno y cuidar los detalles. Además, esta tarea se perfila como un oficio, pues de cada entrevista realizada se tiene un aprendizaje que sirve para construir su propia forma de trabajo. En relación con las habilidades que debe tener o desarrollar quien se dedique a recolectar datos, se pueden mencionar algunas, como: la persistencia —que permite regresar repetidas veces hasta encontrar al(la) informante elegido(a)—, la facilidad de palabra y disposición para no frustrarse y lograr entablar una relación cordial con los(las) vigilantes de los condominios (por ejemplo), así como con la familia de los(las) infor-

¹⁵ Parecido a una *minilaptop*; tiene programados los cuestionarios electrónicos que se aplicarán.

mantes cuando no se les encuentra en la primera visita. Hace falta persuasión y conocimiento del lenguaje, de sus significados y significantes para convencer a las personas que se niegan a participar proporcionando la información solicitada, además de la empatía y la amabilidad, todos recursos que las entrevistadoras y los entrevistadores ponen en práctica al realizar su trabajo.

El(La) informante

Dentro del proceso de trabajo de recolección de datos, las entrevistadoras y los entrevistadores interactúan en su quehacer laboral cotidiano con el(la) informante.¹⁶ Su relación es directa y determinante para realizar y concluir el proceso productivo de la entrevistadora o entrevistador. La del(la)

¹⁶ Es importante mencionar que el(la) informante, para el caso de encuestas levantadas en hogares, se elige con base en características especiales según la temática de la propia encuesta; es decir, para levantar la información específica que requiere el cuestionario es a través de un(a) informante elegido(a) o seleccionado(a), lo cual implica que, de no encontrarse en el momento que la entrevistadora o el entrevistador pasa a esa vivienda, la recolectora o el recolector de datos tendrá que volver a esa vivienda cuantas veces sea necesario para recabar la información y completar su cuestionario 100 por ciento. Esta exigencia (o requerimiento) del proceso de trabajo le representa un esfuerzo extra a la trabajadora o el trabajador de campo, ya que representa tiempo y energía invertidos en la búsqueda de ese(a) informante.

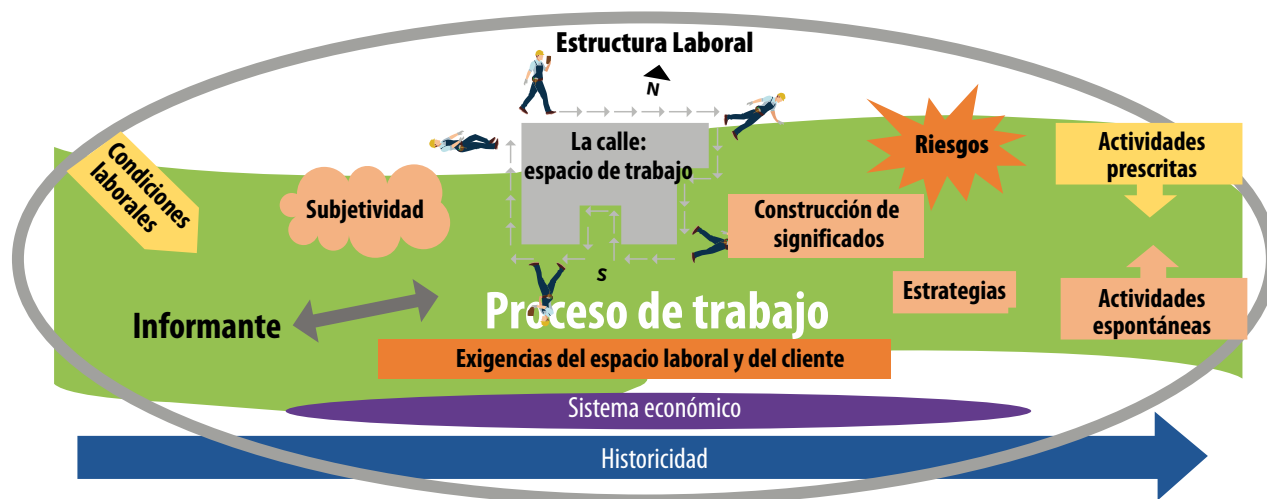
informante marca el tipo de exigencia del proceso mismo (como se ha visto en el relato de la entrevistadora): las jornadas largas y flexibles, las metas pesadas, así como los ritmos de trabajo (duración de la encuesta) están determinadas por su disponibilidad y disposición.

El contacto directo entre ambas partes es el medio para lograr el objetivo: recolectar datos de la población. En gran medida, el cumplimiento total de la carga de trabajo de la entrevistadora o el entrevistador depende de los(las) informantes. Si este(a) accede a ser entrevistado(a) hasta el término del cuestionario, significa una encuesta menos por la que se deba preocupar la trabajadora o el trabajador de campo. Por ello, mencionan las entrevistadoras y los entrevistadores, es importante el contacto que se establece desde un principio con esta persona:

“... y luego me dicen: ‘¿por qué llegas saludando, aunque no haya nadie?’, les digo: ‘es que no sabes si están y a veces con que escuches la voz de la persona te genera confianza o no, ¿verdad?’. Entonces yo siempre llevo saludando, ‘¡Buenas tardes!’” (entrevistadora B, 2017).

Diagrama 2

Elementos que forman el proceso de trabajo de las entrevistadoras o entrevistadores y configuran la construcción de un oficio



Fuente: elaboración propia.

“... hay todo tipo de personas a las que debes enfrentar, y más bien es cómo saber abordarlas, y a mí, a lo mejor, lo que me ha ayudado es eso, dirigirme con respeto y que siempre les hablas de la confiabilidad de los datos que están resguardados...” (entrevistadora T, 2017).

Existen informantes de fácil acceso, que proporcionan los datos solicitados por la entrevistadora o el entrevistador sin ninguna renuencia; algunos(as) agradecen ser visitados(as) por el INEGI, se sienten tomados(as) en cuenta al solicitarles su participación. En la relación directa, los trabajadores y las trabajadoras de campo dicen aprender de los(las) informantes, por ejemplo, sobre la situación que se vive en el país y, también, a valorar y reconocer la importancia del trabajo que desempeñan:

“... hay viviendas que te mencionan la cantidad que ganan, siendo a lo mejor familias grandes, que uno dice: ‘¿cómo sobreviven con tan poco?’ [...] te vas a la contraparte, a los niveles económicos altos, en donde dices: ‘¿en qué se gastan todo esto?’. Si este sobrevive con 2 500 pesos y son cinco integrantes y acá son tres y tienen ingreso en el hogar de 80 mil pesos por mes [...] a ver, ¿dónde está la paridad?, ¿cómo está la mala distribución de los recursos en la sociedad? Esa parte sí te deja mucho cuando escuchas a las personas; incluso, se les salen las lágrimas al mencionarte algún dato...” (entrevistador E, 2017).

“Cuando la gente te agradece, es ahí donde te das cuenta de la importancia de hacer bien tu trabajo [...] lo que ves reflejado en políticas públicas, o cuando el niño te dice: ‘¡ay!, mamá, me encargaron tarea y necesito hacer esto’, es cuando ves que ahí está y que sí se necesita, que sí es importante...” (entrevistadora J, 2017).

Sin embargo, también existen aquellos(as) informantes que implican un mayor esfuerzo para la entrevistadora o el entrevistador; que son poco accesibles debido a diversas razones: son renuentes, por la dificultad para encontrarlos(as) en casa, o bien, hay algunos(as) que se niegan de manera

rotunda a proporcionar los datos solicitados, por ignorancia o desconfianza:

“Me ha tocado gente (informantes) que hasta te mientan la madre, que relacionan todo con el gobierno, con el país, que es un país corrupto; a mí me ha tocado que me digan: ‘¡pinche barbero del gobierno!’, me han azotado la puerta en la cara, incluso, no te abren...” (entrevistador A, 2017).

Cuando este tipo de situaciones se presentan, la instrucción para las entrevistadoras y los entrevistadores es volver a solicitar la entrevista en otro momento. En algunos manuales se menciona que se le debe visitar mínimo tres veces tratando de obtener la información. Seguir la indicación prescrita por la estructura laboral coloca a la entrevistadora o el entrevistador frente a una exigencia subjetiva del proceso de trabajo, pues tiene que insistir y hacer todo lo necesario para lograr cada entrevista, independientemente de si la población seleccionada quiere participar en el proceso de captación de datos.

En otras palabras, para las entrevistadoras y los entrevistadores, una negación por parte del(la) informante es inaceptable, así lo establecen los manuales de procedimientos. También, una negativa representa una entrevista que no le permitirá cumplir con el total de su carga de trabajo o, por lo menos, obtener el mínimo establecido. Derivado de ello, un no lo consideran un reto, aunque implique un desgaste emocional, pues a partir de él se genera estrés o cierta angustia ante el imperativo institucional de regresar cuantas veces sea necesario a la vivienda y la posibilidad de no obtener la información. Como se mencionó, encontrarse con un(a) informante complicado(a) implica un esfuerzo adicional. Las entrevistadoras y los entrevistadores, ante la renuencia, comienzan a desplegar su creatividad para planear cómo abordarán al(la) informante con el objetivo de obtener lo que necesitan, la entrevista; dicho esfuerzo implica un desgaste:

“Yo lo que digo es: bueno, ahorita me dices que no, pero vamos a ver quién se sale con la suya, te voy a dejar reposar unos días y te voy a volver a insistir. O sea, no trato de hostigarlos, trato de aga-

rrarlos afuera de su casa porque sé que una vez adentro, ya me amolé. Afuera hago mi labor de convencimiento, les digo: ‘regálame 15 minutos y te juro que ya no te voy a volver a molestar, porque es desgastante para ti y también para mí estarte insistiendo, insistiendo, insistiendo! Porque tú vas a decir que no y yo voy a seguir viniendo’...” (entrevistadora N, 2017).

“... con una señora que ya estaba enchilada, le dije a mi compañera: ‘¡échamela!, ya tienes el no, ¡pus’ yo voy por el sí!’...” (entrevistador J, 2017).

La demanda prescrita por la estructura laboral, que establece recabar 100 % de las entrevistas, marca la relación con el(la) informante y abona a la configuración de las exigencias y los riesgos propios del proceso de trabajo de la entrevistadora o el entrevistador. No obstante, aun con estas condiciones descritas de exigencias y riesgos, aunque la estructura presiona con sus lineamientos prescritos, no limitan su capacidad de acción. Es justo ante esas condiciones adversas del trabajo que las entrevistadoras y los entrevistadores aprovechan los resquicios que les da lo espontáneo de la actividad para construir estrategias de acción con la finalidad de proseguir con el desempeño de su labor; si bien en un primer momento estas tácticas son para hacer frente a las condiciones desfavorables, luego pasan a ser parte de lo que les provoca satisfacción por tener la posibilidad de resolver situaciones inesperadas con su creatividad.

Estrategias de acción

A través de la observación *in situ*, así como de los relatos de las entrevistadoras y los entrevistadores, se ubica la existencia de unas estrategias de acción para la ejecución de la actividad y otras de sublimación del riesgo. En conjunto, estas acciones estratégicas refieren a la construcción de una forma de organizarse de la entrevistadora o el entrevistador que le permita hacer más práctico el trabajo, en el sentido de manejar todas las herramientas que ocupa estando en campo, a la vez que ahorra tiempo y llega preparada(o) para interactuar con el(la) informante;

estas son creadas e implementadas para diversas situaciones: operativas, trato con el(la) informante, ante riesgos y en cuidados básicos personales.

El INEGI, como institución, establece su forma de organización y los lineamientos para la ejecución del levantamiento de información, lo cual se transmite mediante las capacitaciones y los manuales operativos para aquellas personas que se contraten como entrevistadoras-entrevistadores. De igual modo, para garantizar una correcta ejecución de la actividad, el Instituto establece el mecanismo de supervisión a través de diversas figuras, como: el(la) supervisor(a) de campo, el(la) instructor(a) o el(la) responsable de la encuesta.

En el análisis de los relatos de estas(os) trabajadoras(es) de campo se encuentra que, aun con todo este despliegue de procedimientos institucionales, a las entrevistadoras y los entrevistadores se les presentan distintas situaciones inesperadas por la propia naturaleza del trabajo, las cuales deben resolver casi en el momento, para no perder la posibilidad de conseguir la entrevista.

Aunque de manera oficial está fuera de los lineamientos, las entrevistadoras y los entrevistadores, frente a lo inesperado del proceso de trabajo, cuentan con un margen de acción amplio para la resolución de situaciones espontáneas o problemáticas, por lo que estas(os) trabajadoras(es), ante la necesidad de resolverlas con prontitud, construyen tácticas para la captación oportuna de información y ponen en marcha su creatividad para construir estrategias que posibilitan el cumplimiento de sus jornadas extensas y sus cargas de trabajo pesadas.

Las estrategias de acción operativas son implementadas por cada entrevistadora-entrevistador conforme a su sentido de organización; le permiten mantener cierta sistematización de su proceso de trabajo, así como mayor dominio del manejo de los instrumentos que ocupa en campo, entre cuestionarios y otra papelería que utiliza:

“... yo lo que hago es empezar a ver qué colonia me toca, ver qué ruta de camión puedo tomar [...]

empezar a buscar accesos a los lugares que nos tocan, el poder localizar la vivienda seleccionada, el tocar, el preguntar..." (entrevistadora B, 2017).

"... en lo personal, es analizar a qué tipo de zona voy, más o menos ver cuáles son los horarios en los que realmente puedo encontrar a las personas en su vivienda y no estar dando vueltas y vueltas en un horario donde nunca lo voy a encontrar; como que esa parte sí la voy a analizar mucho [...] uno busca un horario más flexible para las personas..." (entrevistador E, 2017).

Para acercarse al(la) informante, las estrategias de acción se ponen en práctica desde el momento en que la entrevistadora o el entrevistador toca a la puerta de quien le proporcionará información. El análisis de la situación y el contexto son fundamentales para el desarrollo de la actividad:

"No es nada más ya te dieron el trabajo, ve y toca. Desde que el perro te ladra, tocas y luego, luego, '¿quién es?', ahí te das cuenta a qué persona te vas a enfrentar [...] Entonces, también se requiere estar atento al informante, estar alerta y analítico de la persona con la que vas a mantener contacto desde que te abre la puerta..." (entrevistador J, 2017).

"Cada informante es diferente, no siempre vas a encontrar que una negativa sea una negativa; tienes que valorar lo que estás viendo. Tener el contacto inicial, porque si nunca has tenido contacto no puedes valorar el decir: '¡ah, sí me van a abrir!', tienes que ver y que observes a la persona [...] nada tienes que dar por sentado..." (entrevistadora J, 2017).

Debido a la existencia de informantes con distintas características y formas de tratar a la trabajadora y el trabajador de campo, ellas(os) buscan distintas formas de abordar a las personas de quien requieren información:

"... desde que te abre la puerta, tú llegas a esa vivienda y tienes que indagar quién vive allí; si no te abrieron, vas a indagar [...] si es un señor ya grande, es jubilado, pensionado. Entonces, tienes que

agarrarlo con pincitas, tocarle despacito..." (entrevistador J, 2017).

"... te vas a enfrentar tanto a personas educadas, cultas, de buena presencia como a personas déspotas, prepotentes, groseras; y tienes que saber manejar esa situación, groserías o maldiciones. En pocas palabras, nos la han rayado, entonces, hay que tener la habilidad en ese momento de conducir a la persona para que ella suavice las palabras que está mencionando en ese rato..." (entrevistadora A, 2017).

Parte de las tácticas para enfrentar al(la) informante tienen que ver con una preparación extra, alterna a las capacitaciones que proporciona el Instituto. Las entrevistadoras y los entrevistadores se informan por su cuenta sobre los beneficios sociales que conlleva la generación de información estadística, que empieza, justo, en la parte operativa, con la recolección de datos que ellas(os) realizan:

"... hay que leer un poquito más del Instituto, así cuando nos enfrentamos a algún informante difícil, no le damos la oportunidad de que nosotros no sepamos para el lugar al que estamos trabajando..." (entrevistadora MT, 2017).

"El entrevistador debe tener un conocimiento más allá, más amplio de la encuesta [...] es para que estés segura frente a un informante [...] si tú tienes más información, eso te puede abrir la puerta [...] saber qué hace el INEGI, para qué sirve, son las preguntas que siempre te hacen..." (entrevistadora L, 2017).

Para hacer más llevadero el hecho de las agresiones, las entrevistadoras y los entrevistadores implementan ciertas tácticas que les permiten afrontar la situación y continuar el trato con el(la) informante agresor(a):

"Pues, yo lo que hago primero es escuchar, no decir nada. O sea, yo lo dejo que se desahogue, porque a veces ha resultado que está en un momento de frustración, de situaciones de la vida,

pero es una persona. A lo mejor, de seis personas es una la que no te va a dar información [...] 'ora sí que tu trabajo es estar dispuesto a soportar, a escuchar hasta cierto punto...' (entrevistadora BT, 2017).

"Yo lo que hago es no ser insistente la primera vez; si yo veo que me empieza a decir: 'no creo' o 'no quiero', empiezo a decirle que regreso mañana [...] si me dice que no, le explico que es mi trabajo y me exigen regresar cuantas veces sea necesario [...] la persona a veces sí lo entiende y lo comprende..." (entrevistadora B, 2017).

Para recuperarse de estas situaciones desagradables con ciertos informantes que expresan su negativa a participar con acciones agresivas o violentas, las entrevistadoras y los entrevistadores tienen varias formas para recuperarse del momento desagradable y continuar con su dinámica laboral:

"La primera vez que me cerraron la puerta en la cara lloré bastante y dije: '¡no vuelvo!' [...] es muy desgastante estar escuchando a la gente que te ofende [...] aprendí y ya no lo tomaba tan personal, porque hay gente que dice: '¡ahí viene el gobierno, ni hace nada!'. Yo entendí que no es a mí como persona, es por el Instituto [...] de repente sí me da coraje, pero digo: 'a ver, tranquila, no te lo tomes tan personal, hablan sobre el Instituto, no sobre mí, me tranquilizo, me relajo y el que sigue...' (entrevistador B, 2017).

"¡Con unas chelas! ¡Ja, ja, ja! No, mira [...] un día te pueden tratar muy bien; otro, muy mal. Siempre hay que tener eso en mente. Al final de cuentas, yo, como entrevistador [...] en ese momento que porto el gafete y los insumos, tengo que tener en mente que, al final de cuentas, estamos en esa representación de las instituciones. La gente está desanimada o tiene cierto recelo en contra de todas [...] hay que tener eso bien claro. Ya, si recibes una agresión, pues realmente no es una agresión contra ti, están tratando de mostrar recelo contra la institución..." (entrevistador E, 2017).

Aun con los lineamientos institucionales prescritos, las entrevistadoras y los entrevistadores, al momento de realizar su trabajo en campo, se encuentran con realidades distintas o fuera del alcance de los manuales y de las instrucciones marcadas, por lo que, para no postergar el cumplimiento de su meta, deben resolver en lo inmediato. Se puede concluir la existencia de una diferencia significativa entre los lineamientos establecidos por la institución y la realidad vivida en campo. Ante lo espontáneo del trabajo, surge la posibilidad de un margen de acción amplio por el cual la trabajadora o el trabajador pone en práctica su creatividad para construir estrategias que solucionen el imprevisto, teniendo la capacidad de decidir y ejecutar lo más conveniente para el logro de su meta laboral:

"Siempre trato de hacer lo que me dice mi lógica, como que analizo la situación, digo: 'no voy a perjudicar aquí, tampoco acá'; a veces le dices al supervisor o a alguien más arriba: 'no pues es que yo creo o siento que es esto, aunque me regañes, lo voy a hacer y asumo mi responsabilidad...' (entrevistadora N, 2017).

"En los manuales, los cursos y las capacitaciones son muy cuadrados; inclusive hasta los ejemplos visuales son como la manzana perfecta y el acomodo perfecto de las cosas. La realidad es que vas a campo y nada qué ver. Casi siempre se trata de resolver; cuando tienes la posibilidad de conseguir quién te saque de la duda, pues, ¡qué bueno!, pero cuando no, ya tendrás qué resolver rogando a Dios que haya sido lo correcto..." (entrevistadora MY, 2017).

A partir de estos elementos expuestos se puede comenzar a vislumbrar el trabajo de la entrevistadora o el entrevistador como un oficio; si bien, en términos concretos, la ejecución de su trabajo no es creadora de un objeto tangible, sí lo es del diseño de estrategias de acción que le ayudarán a conseguir cada una de las entrevistas necesarias para el cumplimiento de su meta. Haciendo alusión al texto *El artesano* de Richard Sennett (2009), lo que elabora la entrevistadora o el entrevistador es la forma en cómo ejecuta la actividad a partir del do-

minio de sus instrumentos para captar los datos requeridos; es decir, es su técnica, pensando esta no como procedimiento insensible o irreflexivo sino, más bien, como un asunto cultural, de aprendizaje, ligado a la expresión. “La comprensión de la técnica se desarrolla a través del poder de la imaginación [...] La utilización de herramientas imperfectas o incompletas estimula la imaginación a desarrollar habilidades aptas para la reparación y la improvisación...” (Sennett, 2009:13). Esta artesanía la construyen las entrevistadoras y los entrevistadores a partir de la necesidad imprevista del trabajo prescrito, sin embargo, se elabora a partir del aprendizaje lento en cada una de las viviendas visitadas y del hábito de comprender a profundidad la actividad que realiza.

Habilidades de la entrevistadora o el entrevistador: el devenir de una artesana o artesano

Existe un vínculo entre las formas de resolución de situaciones imprevistas en el trabajo de campo y la forma de afrontarlos a partir de ciertas cualidades que describe a las entrevistadoras y los entrevistadores. De ahí que ellas(os) mismas(os) comenten que: “... cada quien tiene que hacerse de sus mañas...”. Este trabajo es artesanal ya que, para realizarlo de manera exitosa, se requiere mantener una estrecha relación entre las prácticas concretas y el pensamiento, en este caso, entre el diseño de la estrategia y su ejecución y viceversa. Una buena artesana o un buen artesano mantiene una relación entre la mano y la cabeza y este diálogo lo evoluciona hasta convertirse en hábito (o en actos) que establecen un ritmo entre el descubrimiento y la solución de problemas (Sennett, 2009).

Según relatan las entrevistadoras y los entrevistadores, no todas las personas que prueban el trabajo de campo se interesan por continuar en la recolección de datos:

“... cuando contratan no conocen a la gente [...] y ellos quieren venderse, pero ya ven la realidad en el campo [...] no aguantan, pues estar pasando

calores, fríos, andar todo el día con polvo, estar soportando a gente...” (entrevistador A, 2017).

Para desarrollar esta actividad no se requiere tener un grado académico, sino contar con ciertas destrezas y gusto por el trabajo:

“Te puedo decir que yo conozco a entrevistadores muy buenos con preparatoria y a entrevistadores no tan buenos con licenciatura...” (entrevistador S, 2017).

“Que de verdad a ti te guste la chamba, yo digo que, si no te gusta andar en la calle, tocando puertas y estar viendo a diferentes personas, esa seguridad la vas perdiendo o no la transmites...” (entrevistadora R, 2017).

Así como se hacen de estrategias, también se arman de distintas habilidades, sobre todo para la interacción con el(la) informante y, en general, para la realización de su trabajo. Algunas de ellas ya las tienen y solo las potencian, y otras, las(los) propias(os) trabajadoras(es) no se explican de dónde salen:

“Sí te ocurren muchas cosas que en el manual no vienen. Sobre todo, en lo que tiene que ver con el contexto. Hasta las situaciones que te enfrentas con los informantes hay unas cosas que realmente no sabes cómo solucionarlas, pero las solucionas. Primeramente, pues accesas [sic] a tu jefe de entrevistadores, a veces no es posible porque no tienes uno personal. Pero te das tu habilidad. ¿De dónde sale? No sé —ríe—, no sé tampoco, pero la experiencia sí te va dando armas...” (entrevistadora RY, 2017).

“Aun con la experiencia que uno tiene, siempre va a ser diferente, como entrevistador tienes que buscar ese aprendizaje contigo, entonces, llega un momento en el que tienes muchas habilidades y las utilizas, por ejemplo, para andar en zonas de mucha inseguridad...” (entrevistadora A, 2017).

Entre las principales habilidades a las que refieren las entrevistadoras y los entrevistadores están

el compromiso, el buen ánimo para tratar con el(la) informante y el gusto por aprender y solucionar:

“Compromiso, tenacidad, honestidad, respeto y yo creo que mucha alegría porque el hecho de llegar, el saber enfrentar [...]; no puedo llegar triste, cabizbaja o dudosa con una persona si el otro me gritó [...] es como cerrar ese momento...” (entrevistadora JG, 2017).

“Pues para mí una es honestidad; otra, responsabilidad. Debes ser muy analítico y tener una actitud de pensar que no es fácil tener un informante, porque no lo conocemos, pero que vamos por la información. Esas serían las cuatro [...] la ética como personas, la ética en cualquier cosa, así seas entrevistador, empleado, secretario, lo que tú quieras, debes tener ética [...] y gusto, que te gusten las cosas...” (entrevistador J, 2017).

En síntesis, la actividad de recolección de datos, sencilla a simple vista, exige de la trabajadora y el trabajador un conjunto de cualidades físicas e intelectuales, una suma de hábitos, de conocimiento empírico de los instrumentos utilizados y de destrezas que van hasta los secretos más profundos del oficio (Palacios, 2014: 10). Asimismo, en palabra de las(los) propias(os) protagonistas de esta investigación, para realizar esta tarea se requiere de un gusto por desarrollarla.

Autores como Sennett (2009), Novelo (2008) y Palacios (2014) coinciden que la gente puede aprender de sí misma a través de las cosas que produce. Para hacerlo, requiere preocuparse por las cualidades para, con ello, hacer bien el trabajo, con dedicación y compromiso. Esa forma de realizar la actividad de recolección de datos es lo que ha mostrado este grupo de trabajadoras(es), aguerridas(os) para concluir con su carga laboral, cuidando que esté *bien hecho* —refiriéndonos a su particular subjetividad—; sin embargo, lo bien hecho es suma de conocimiento, habilidad y placer (Sennett, 2009), es el compromiso por llevar a cabo el trabajo hasta terminarlo y el sentimiento de satisfacción al concluirlo.

Para este grupo de mujeres y hombres, hacer bien el trabajo implica tener el conocimiento sobre la temática que abordarán en las entrevistas (desde su capacitación previa al levantamiento); asimismo, significa realizar una estrategia adecuada para recorrer y ubicar sus viviendas preestablecidas, así como levantar con el mayor porcentaje, y haciendo hasta lo imposible, cada una de las entrevistas asignadas. Hacer bien el trabajo para ellas(os) es llevarlo a cabo de una manera entusiasta, honesta y con gusto por recolectar datos.

En este sentido, no se debe dejar de lado que para estas(os) trabajadoras(es) de campo parte del motor que las(los) incentiva a hacer bien su trabajo, así como a entregar su tiempo, esfuerzo y dedicación para levantar cada entrevista, es el reconocimiento explícito por parte de la organización que las(los) contrata porque, con ello, las entrevistadoras y los entrevistadores saben que las desveladas para encontrar a los(las) informantes nocturnos, el tránsito por lugares peligrosos, los malos tratos de las personas renuentes a participar, las jornadas extenuantes y las faltas a eventos familiares han valido la pena, en tanto que el Instituto reconoce la voluntad y entrega que tienen las(los) trabajadoras(es) de campo para lograr el cumplimiento de sus metas. Por ello, se considera que al conocer los retos y dificultades que las entrevistadoras y los entrevistadores enfrentan en su cotidianidad laboral en voz propia, habrá una comprensión del significado que tiene para estas personas ser entrevistadora-entrevistador del INEGI.

Trabajo creador de experiencia

El margen de acción que surge de la propia ejecución de la actividad principal de la entrevistadora o el entrevistador, y dadas las situaciones espontáneas del trabajo de campo, les exige la construcción de estrategias para la pronta resolución de los imprevistos. Pero, de igual forma, le ofrece la posibilidad de poner en uso su creatividad para, con ello, construir las tácticas que la(lo) llevarán al buen desarrollo y conclusión de su trabajo; es decir, la labor cotidiana le exige pero, a la vez, le permite

crear experiencia.¹⁷ Al verse en la necesidad de resolver, busca las formas para hacerlo, lo cual le deja un aprendizaje sobre lo que le funciona y en qué situaciones lo hace; cada una de las viviendas que visita es una experiencia que va acumulando. Y el cúmulo de ellas, producto de la resolución y aprendizaje en su trabajo, le hace ir construyendo en lo cotidiano el *ser* entrevistadora-entrevistador. Así, pasa de ser un sujeto que trabaja en la simple ejecución de la recolección de datos, a uno que, en lo cotidiano, con el desarrollo de su trabajo, mediante el uso del ensayo y error sobre el desarrollo de la técnica, aprende a elaborar un *saber hacer* con implicaciones para la experiencia laboral y también para la personal.

Su trabajo, mediante la interacción entre los distintos elementos subjetivos, genera aprendizaje, a la vez que provoca gozo y sufrimiento en las(los) trabajadoras(es) de campo del INEGI pero, conjugado, crea experiencias significativas para ellas(os). Y aun con las condiciones específicas de precariedad, reflejadas en las exigencias laborales y en los riesgos psicosociales, configura un sentido de pertenencia —entendido como una necesidad superior, una necesidad afectiva de ser parte de, en este caso de un grupo de trabajadoras(es) pertenecientes a una institución con una tarea específica— respecto a la actividad que la entrevistadora o el entrevistador realiza.

Así es como el espacio laboral satisface en estas(os) trabajadoras(es) el sentido de pertenencia, por el reconocimiento de la institución, por la identificación entre pares, por la sensación de bienestar que le provoca cada uno de los logros, de ahí que establecen un fuerte vínculo afectivo con el Instituto, se sienten parte del grupo perteneciente a las entrevistadoras y los entrevistadores del INEGI, lo cual abona a la estructura como sujeto, su implicación a través del aprendizaje y la experiencia adquirida mediante cada una de las vivencias; eso lo hace ser entrevistadora-entrevistador:

¹⁷ Conocimiento de la vida cotidiana construida desde el aprendizaje de las vivencias, lo que les permite dar coherencia y organizar su mundo, en este caso su mundo laboral (Berger y Luckmann, 1968).

“El INEGI me ha dado parte de lo que yo soy, el estar contigo y platicar me va a cultivar más. El hecho de que yo llegue al Instituto y quiera ser entrevistador me abre no na’más las puertas de una casa, me abre las puertas de donde yo quiera, porque, por ejemplo, en la zona industrial a uno ya no lo quieren porque saben el perfil que trae, saben que con una entrevista pueden reflejar muchas cosas, entonces, yo digo que es parte de mí el estar ahí. Cada vez que yo entro a una capacitación digo: ‘¡ah, qué tema tan interesante!’, voy a aprender algo, y más aprendes de las experiencias de la gente [...] pero yo no tengo miedo, si dices: ‘tú ve con el presidente’, pues voy y hablo [...] cuando yo estaba niño no hablaba bien, tuve clase de dicción [...] pero ya no me pasa [...] el INEGI me ha dejado herramientas, ps’, yo no me atrevía a tocar una puerta...” (entrevistador J, 2017).

Reflexiones finales

De los resultados que arrojó la investigación *El proceso de trabajo de los entrevistadores del INEGI, riesgos psicosociales y estrategias de acción*, podemos dar cuenta de que las entrevistadoras y los entrevistadores del INEGI son trabajadoras(es) comprometidas(os) con su labor de recolección de datos, se sienten identificadas(os) y pertenecientes al Instituto, aun con la condición de ser trabajadoras o trabajadores eventuales, lo cual las(los) coloca en situación de incertidumbre ante la inseguridad de un trabajo continuo y permanente que brinde estabilidad laboral, económica y emocional (Castel, 2008): esto es, existe un fuerte vínculo laboral de pertenencia con el Instituto. Gozan de su quehacer laboral a pesar de ser un trabajo riesgoso, les gusta la tarea que llevan a cabo porque, para ellas(os), la recolección de datos la viven como un reto laboral, pero también personal; cuando se encuentran con informantes poco participativos(as), hacen el máximo esfuerzo para lograr que esa entrevista sea efectiva, no se rinden, buscan las formas posibles para cumplir con su carga de trabajo completa. Aunque disfrutan de su trabajo, también sufren y se decepcionan por las jornadas largas, las exigencias del proceso laboral, las cargas de trabajo cada

vez más excesivas y los riesgos, que son parte de su labor. El hecho de tener cada vivienda por visitar como un reto profesional a lograr, les da fuerza y empuje para, al final del día, conseguir cerrar una encuesta más.

Dado que la captación de datos es un trabajo que se realiza en calle, eso les permite sentirse libres —según sus comentarios— en el sentido de no estar sujetos a un solo espacio laboral y a un horario de oficina. De alguna manera, tienen la posibilidad de acomodar sus visitas para entrevista como consideran más conveniente por los horarios, el área de trabajo o los(las) mismos(as) informantes; y ese margen de acción les produce un sentimiento de control de su trabajo.¹⁸

Asimismo, el tener un margen de acción para resolver situaciones inesperadas incide en la autopercepción de ser dueñas(os) de su trabajo ya que, por esfuerzo propio, logran hacerse de estrategias que les permiten, la mayoría de las veces, llevar a buen término sus entrevistas. La apropiación del proceso de trabajo implica que las entrevistadoras y los entrevistadores han experimentado en la labor cotidiana los esfuerzos físicos e intelectuales que se requieren para lograr el cumplimiento de sus metas laborales, así que construyen su propia forma de desarrollar el proceso de recolección de datos.

Como se podrá observar en el recorrido del texto, aprender a ser o ser entrevistadora-entrevistador no es un asunto del orden académico, tampoco es algo que se aprenda con las capacitaciones que proporciona el Instituto al inicio de cada encuesta. Si bien los conocimientos académicos o las instrucciones laborales prescritas son elementos que aportan al trabajo de la entrevistadora y el entrevistador parte de los recursos cognoscitivos prácticos del sujeto, no lo son todo para aprender a ser entrevistadora-entrevistador, porque para ello se requiere gusto por el trabajo, compromiso por hacer bien la la-

bor asignada, así como *darse su habilidad*, que tiene que ver con crear y pulir sus destrezas para ejecutar su trabajo, sobre todo para interactuar con el(la) informante y saber cómo proceder en cada situación distinta que se les presente en el trabajo de campo.

Mujeres y hombres, entrevistadoras y entrevistadores, ante los desafíos que enfrentan por la propia actividad, movilizan su inteligencia práctica y su capacidad de innovar para lograr cumplir su carga de trabajo y han descubierto sus habilidades y, como producto del aprendizaje y la experiencia, han desarrollado *su técnica* que, aunque la compartan con las(los) compañeras(os), cada una(o) le imprime su sello personal para desarrollarla y ejecutarla; por lo tanto, el trabajo de las entrevistadoras y los entrevistadores no es una simple ejecución técnica (Neffa, 2015) en el sentido insensible o irreflexivo de recolección de datos, es un oficio, dado que la entrevistadora o el entrevistador, como la(el) artesana(o), muestra un compromiso a través de su práctica laboral cotidiana.

De igual forma, se puede pensar como un oficio en tanto que hay una implicación con el trabajo; cuando las entrevistadoras o los entrevistadores hablan sobre lo que hacen, utilizan una combinación de elementos concretos para describir las actividades que constituyen su proceso de producción y también echan mano de lo que la experiencia les ha dejado a partir de la ejecución cotidiana de esta labor. La mayoría, al narrar lo que hace, comienza a platicar de sus distintas experiencias del trabajo de campo; describe las situaciones que le agradan y que disfruta de la tarea que desarrolla, a la vez que refiere los hechos que le dificultan la ejecución de su actividad; es decir, la entrevistadora o el entrevistador construye la narrativa de su proceso de trabajo desde su experiencia significativa, responde desde lo que es para ella(él) ser o el hacer de la entrevistadora o el entrevistador.

Ser entrevistadora-entrevistador implica el aprendizaje por el oficio de la recolección de datos, esto

¹⁸ Para profundizar en la dimensión subjetiva de estas(os) trabajadoras(es), ver apartado 3.1, cap. 3 y reflexiones finales de la tesis de maestría mencionada.

es, aprender de un saber hacer. Para conocer de manera profunda el trabajo de las(los) otras(os)¹⁹ y para comprender la construcción de significados, de estrategias y la implicación de la trabajadora o el trabajador con su actividad, es fundamental poner al centro la experiencia subjetiva, dándole su lugar a la figura central del estudio. Escuchando su palabra y aprendiendo de sus vivencias, se podrá evidenciar que la experiencia subjetiva trasciende cualquier proceso de trabajo mecánico.

Fuentes

Berger, P. y T. Luckmann. *La construcción social de la realidad*. Buenos Aires, Argentina, Amorrortu, 1968.

Castel, R. *La inseguridad social. ¿Qué es estar protegido?* Buenos Aires, Argentina, Ediciones Manantial, 2008.

Castro, H. 1.º *Censo de Población de la Nueva España 1790. Censo de Revillagigedo, un censo condenado*. México, Dirección General de Estadística, 1977.

De la Garza, E. "Trabajo atípico, ¿identidad o fragmentación?: Alternativas de análisis", en: *Trabajos atípicos y precarización del empleo*. México, El Colegio de México, 2011, pp. 49-79.

Instituto Nacional de Estadística y Geografía (INEGI). *Cronología de la estadística en México (1521-2008)*. México, INEGI, 2009a.

_____. *125 años de la Dirección General de Estadística*. Aguascalientes, México, INEGI, 2009b.

¹⁹ Entendidas(os) como las(los) trabajadoras(es) de cualquier sector; es decir, al hacer estudios sobre el trabajo es necesario no olvidar centrar la atención en la experiencia subjetiva de los que desarrollan actividades laborales.

_____. *Precusores de las estadísticas oficiales*. Aguascalientes, México, INEGI, 2011a.

_____. *40 años de encuestas de fuerza laboral en México*. Aguascalientes, México, INEGI, 2011b.

Moreno, A. *Los censos y padrones de los siglos XVI-XIX*. México, INEGI-Comisión Nacional de Difusión Censal, 1980.

Neffa, J. C. *Los riesgos psicosociales en el trabajo: contribución a su estudio*. Buenos Aires, CEIL-CONICET, 2015.

Noriega, M. "Organización laboral, exigencias y enfermedad", en: Laurell, C. *Para la investigación sobre la salud de los trabajadores. Serie PALTEX, Salud y sociedad 2000*, 1993, pp. 167-188.

Noriega, M. y C. Laurell. *Interacción de las exigencias de trabajo en la generación de sufrimiento mental*. Río de Janeiro, Cad. Saúde Pública, 2000, pp. 1011-1019.

Novelo, V. "La fuerza del trabajo artesanal mexicana, protagonista ¿permanente? de la industria", en: *Revista Alteridades*. Vol. 18, núm. 35. México, UAM-Iztapalapa, 2008, pp. 117-126.

Palacios, L. *Oficios urbanos tradicionales*. Monterrey, México, Universidad Autónoma de Nuevo León, 2014.

Sennett, R. *El artesano*. Barcelona, España, Editorial Anagrama, 2009.

Taylor, S. J. y R. Bogdan. *Introducción a los métodos cualitativos de investigación*. Buenos Aires, Argentina, Ediciones Paidós, 1987.

Zemelman, H. *Voluntad de conocer. El sujeto y su conocimiento en el paradigma crítico*. Barcelona, España, Anthropos coedición México, Centro de Investigaciones Humanísticas de la Universidad Autónoma de Chiapas, 2005.

_____. *Conocimiento y sujetos sociales. Contribución al estudio del presente*. La Paz, Bolivia, Instituto Internacional de Integración del Convenio Andrés Bello, 2011.

Estandarización de las áreas geoestadísticas básicas urbanas para 100 ciudades del Sistema Urbano Nacional

Standardization of Basic Urban Geostatistical Areas for 100 Cities of the National Urban System

Jorge A. Montejano Escamilla,* F. Gerardo Ávila Jiménez** y Camilo Alberto Caudillo Cos***

En los últimos años se ha incrementado de forma considerable el número de estudios espaciales que utilizan las áreas geoestadísticas básicas (AGEB) como referente geográfico de los datos censales en nuestro país, los cuales produce el Instituto Nacional de Estadística y Geografía. En el proyecto *Forma Urbana y Productividad en México 1995-2015* —desarrollado por el Centro de Investigación en Ciencias de Información Espacial— se utilizaron, primero, para calcular la superficie urbana y, después, la expansión urbana en el tiempo y algunas métricas adicionales de forma urbana para las 100 ciudades más grandes del Sistema Urbano Nacional 2010 entre 1995 y el 2015. Las comparaciones entre las AGEB, realizadas en cinco fechas censales, arrojaron inconsistencias topológicas (i. e. desfases y deformaciones cartográficas), por lo que se tuvieron que corregir las bases de datos para dar certidumbre a los resultados. En este trabajo se presenta el procedimiento utilizado en el ajuste espacial y la corrección topológica de las AGEB.

Palabras clave: ajuste espacial; topología y datos espaciales.

Recibido: 11 de octubre de 2019.
Aceptado: 27 de febrero de 2020.

* Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo), jmontejano@centrogeo.edu.mx

** Universidad Autónoma Metropolitana y CentroGeo, gavila@centrogeo.edu.mx

*** CentroGeo, ccaudillo@centrogeo.edu.mx

Nota: trabajo realizado con el apoyo financiero del proyecto Fondo Sectorial INEGI-CONACYT (S0025-2016-1), solicitud núm. 278953.

In recent years there has been a considerable increase in the number of spatial studies that use basic geostatistical areas (AGEB) as a geographical reference for census data in our country, which are produced by the National Institute of Statistics and Geography. In the project *Urban Form and Productivity in Mexico 1995-2015* —developed by the Center for Research in Geography and Geomatics— were used, first, to calculate the urban area and then, the urban expansion in time and some additional metrics of urban form for the 100 largest cities of the National Urban System 2010 between 1995 and 2015. Comparisons between AGEBS, carried out on five census dates, showed topological inconsistencies (i.e. lags, and cartographic deformations), so the databases had to be corrected to give certainty to the results. In this paper we present the procedure used in the spatial adjustment and topological correction of the AGEBS.

Key words: spatial adjustment; topology and spatial data.



Aerial View Of Buildings In City/Luis Mauricio Hernández Meyer/EyeEm/Getty Images

Introducción

El Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo) logró apoyo financiero del Fondo Sectorial INEGI-CONACYT con el objetivo de probar diversas hipótesis sobre la relación entre forma urbana y productividad laboral (medida como valor agregado censal bruto/trabajador). La idea central planteaba que las ciudades compactas tienden a ser más eficientes que las más dispersas. Lo anterior es argumentado en literatura científica proveniente sobre todo de países con economías avanzadas y basadas en su mayoría en servicios. Para ahondar en la propia investigación, sus métodos y resultados, consultar Montejano *et al.* (2020) y Montejano *et al.* (2019).

La estrategia de investigación consistió en cuantificar las transformaciones de las 100 ciudades más

grandes del Sistema Urbano Nacional (SUN) 2010 en el tiempo y, posteriormente, verificar el grado y signo de correlación con la productividad laboral observada. La unidad básica para cuantificar los cambios fue el área geoestadística básica (AGEB),¹ que se utilizó para generar las métricas para 1995, 2000, 2005, 2010 y el 2015 (ver cuadro 1). Los resultados de la investigación son relevantes para la política pública territorial nacional ya que, con base en ellos, es posible matizar las políticas actuales en materia de ordenamiento del territorio, plasmadas en la mayoría de los programas de desarrollo urbano municipales que promueven la ciudad

¹ Clasificadas en: rurales y urbanas. Ambas unidades corresponden a la subdivisión de las áreas geoestadísticas municipales. Para este estudio solo se consideraron las últimas, que se definen como: "Área geográfica ocupada por un conjunto de manzanas que generalmente va de 1 a 50, perfectamente delimitadas por calles, avenidas, andadores o cualquier otro rasgo de fácil identificación en el terreno y cuyo uso del suelo sea principalmente habitacional, industrial, de servicios, comercial, etcétera, sólo se asignan al interior de las localidades urbanas". (INEGI, 2010a, p. 8).

compacta como *santo grial* sin tomar en cuenta la base económica que representa la manufactura en México y que está presente en muchas de las urbes del país analizadas, pues sigue siendo la actividad más productiva, pero su principal activo es la localización exurbana, en zonas con relativo bajo precio del suelo y alta accesibilidad a la red carretera.

Dada la gran cantidad de métricas a generarse para alimentar nuestro modelo de panel con base en la unidad AGEB,² fue indispensable revisar la consistencia espacial de las bases de datos generadas por el INEGI para reducir la incertidumbre en el modelaje.

En un ejercicio previo, el equipo ya había detectado errores en la variable *crecimiento urbano* dentro de una base de datos longitudinal 1990-2010. En ella, observamos que existían ciudades del SUN que, en lugar de crecer en tamaño, se hacían cada vez más pequeñas. Este hallazgo nos dio una pis-

2 Diversos autores, entre ellos Angel, Parent y Civco (2010) o Ewing, Pendall y Chen (2002) han hecho hincapié sobre que el *urban sprawl* o la dispersión urbana es un fenómeno multifactorial que no puede ser descrito solo con una variable, de allí el número de métricas.

ta sobre un posible error topológico en la base de datos generada a partir del Marco Geoestadístico Nacional (MGN) del INEGI en sus diferentes ediciones, disponibles en línea y con las que se generó el panel. Por lo anterior, fue necesario —antes de la generación de indicadores y armado del panel 1995-2015— realizar las tareas de ajuste espacial y corrección topológica de los datos geográficos por AGEB del propio MGN. En la segunda etapa se planteó la corrección de estas situaciones para que los índices espaciales generados a partir de la información cartográfica fueran lo más robustos posible.

En las imágenes 1 y 2 se muestran algunas de las múltiples inconsistencias detectadas en la cartografía censal utilizada, la cual fue entregada directamente por el INEGI a CentroGeo y que se asumía podría haber generado resultados erróneos en las métricas. En ellas se puede apreciar que existe rotación, desplazamiento y, en algunos casos, una reducción en el tamaño de las AGEB, errores considerables como para poner en duda cualquier resultado analítico proveniente de estos datos, máxime si el proceso analítico es espacialmente explícito.

Cuadro 1

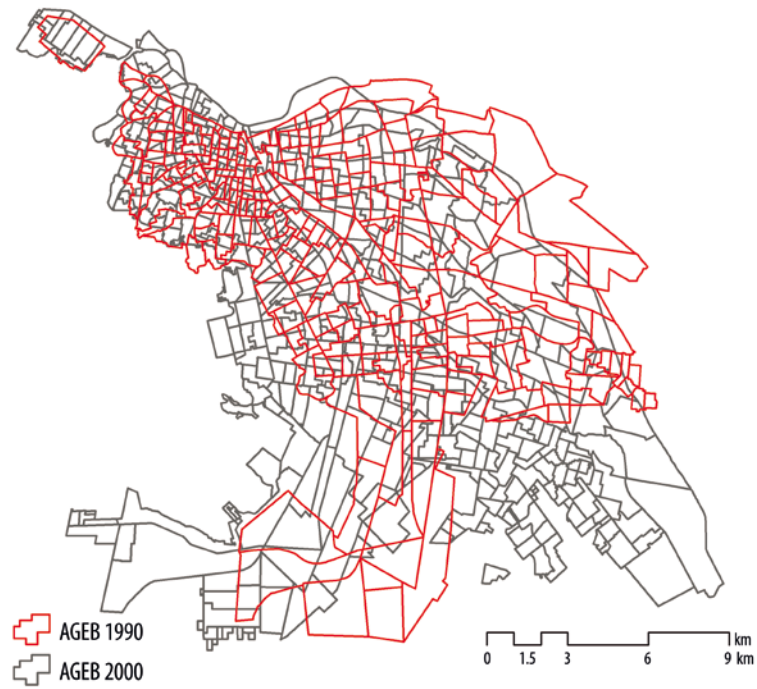
Métricas de forma y estructura urbana generadas en el proyecto *Forma Urbana y Productividad en México 1995-2015*

Categoría	Métrica	Fuentes
Densidad	Población/densidad de población	Boyko y Cooper, 2011; Galster <i>et al.</i> , 2001
Centralidad	Gradiente de densidad	Galster <i>et al.</i> , 2001
	Índice de Centralidad	Galster <i>et al.</i> , 2001
Compacidad	Índice de Proximidad	Angel, Parent y Civco, 2010a
	Índice de Forma Compacidad	VanDerWal <i>et al.</i> , 2019; Gyenizse <i>et al.</i> , 2014; Patton, 1975 Amindarbari y Sevtsuk, 2015
Fragmentación	Fragmentación	Amindarbari y Sevtsuk, 2015
Igualdad de distribución	Coefficiente de Gini	Burt, Barber y Rigby, 2009
	Índice de Agrupación Moran I	Pereira, Nadalin, Monasterio y Albuquerque, 2013 Tsai, 2005
Complejidad	Dimensión fractal	VanDerWal <i>et al.</i> , 2019; Lovejoy, 1982

Fuente: Montejano *et al.*, 2020, p. 305.

Imagen 1

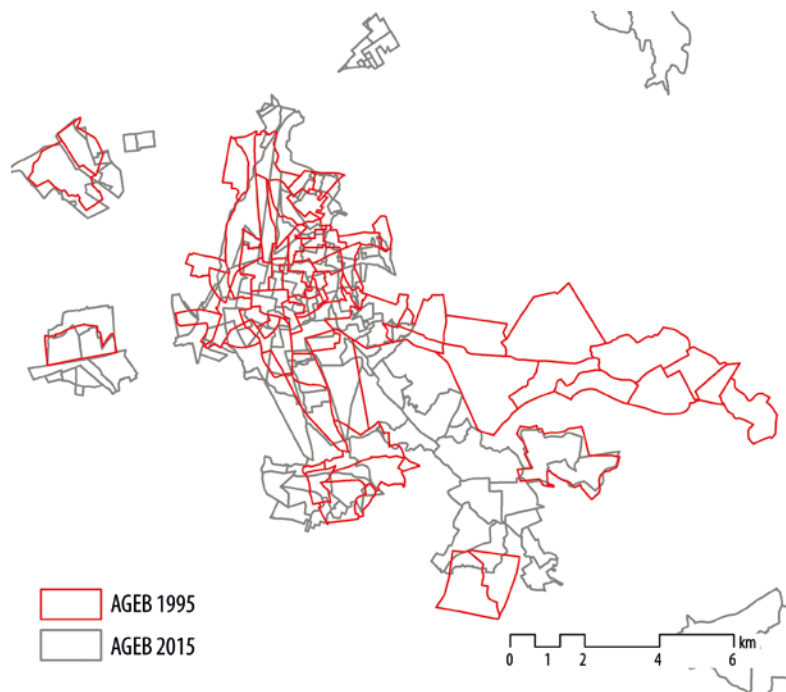
Ciudad Juárez, Chihuahua



Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional por AGEB*. 1990 y 2000.

Imagen 2

Tulancingo, Hidalgo



Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional por AGEB*. 1995 y 2015.

Sobre el SUN y la elección de las 100 ciudades

Como lo definen el Consejo Nacional de Población (CONAPO) y la extinta Secretaría de Desarrollo Social (SEDESOL), el SUN es "... el conjunto de ciudades de 15 mil y más habitantes, que se encuentran relacionadas funcionalmente, y cualquier cambio significativo en alguna de ellas propicia, en mayor o menor medida, alteraciones en las otras..." (2012, p. 11). Se clasifica en diferentes tipos de urbes, las cuales comprenden zonas metropolitanas (agrupaciones de municipios que están muy relacionados funcional y físicamente, o centros urbanos de más de un millón de habitantes), conurbaciones (cuando dos o más localidades o centros urbanos tienen contigüidad física y forman una sola unidad de más de 15 mil habitantes) y centros urbanos (ciudades de 15 mil o más pobladores que no cumplen las dos anteriores características). La versión 2010 contaba, en su momento, con 384 ciudades de más de 15 mil habitantes (CONAPO y SEDESOL, 2012); en la del 2018 hay un total de 401 (SEDATU, 2018).

El SUN, en conjunto con las distintas ediciones de la serie *Delimitación de zonas metropolitanas*, desarrolladas en conjunto por CONAPO, SEDESOL y la Secretaría de Gobernación (SEGOB), constituyen una de las principales herramientas de clasificación y delimitación espacial para la planeación del territorio en nuestro país. Ambos instrumentos dan cuenta de las transformaciones en las dinámicas exurbanas,³ de la creciente concentración de la población en núcleos urbanos y de la dispersión de gran cantidad de asentamientos humanos en el país. De allí la importancia del SUN y de la relevancia de tener una base de datos (MGN) confiable a nivel AGEB.

La razón por la cual determinamos utilizar como insumo en nuestra investigación las primeras 100 ciudades del SUN 2010 fue porque el conjunto de las primeras 93 (con población mayor a 100 mil habitantes) aportaron a la economía cer-

³ Son las que contienen características urbanas, pero en un contexto de dispersión o que no están ligadas de forma directa a entornos urbanos consolidados.

ca de 90 % de la Producción Bruta Total (PBT) del país en el 2009, concentrando también 83 % del total de personal ocupado (ONU-Hábitat y SEDESOL, 2011, p. 27). Además, cerca de 60 % de la población total del 2010 se encontraba repartida en las primeras 100 urbes. Con ellas se garantiza que la mayor parte de la PBT nacional esté cubierta.

Sobre la calidad de los datos espaciales

El problema de esta no es nuevo. Autores como Goodchild y Gopal (1989) o Gutpill y Morrison (1995) han tratado, desde hace tiempo, el tema de la precisión de los datos espaciales, incluyendo aquellos que van desde el linaje hasta el error.

Thapa y Bossler (1992) argumentan que su recolección es la fase más cara de la generación de un sistema de información geográfica (SIG), absorbiendo cerca de 80 % del costo total, por lo que es muy importante contar con métodos lo más confiables posible; en ese trabajo, clasifican los tipos de errores más comunes en la recolección de datos geoespaciales, siendo los *sistemáticos* los que más se parecen a aquellos que hemos detectado en la serie del MGN del INEGI, los cuales muchas veces se deben a imperfecciones instrumentales y a limitaciones humanas; no importando su origen, ellos señalan que, en un sentido estadístico, este tipo de errores induce desviaciones en las observaciones (p. 836).

Por fuente de información, los errores que observamos en esta serie del INEGI pueden ser clasificados en *métodos secundarios* de colección de datos, es decir, las vías en que estos son recabados desde mapas, gráficas, cartas, etc. (Thapa y Bossler, 1992, p. 836). Los más comunes en los métodos secundarios de recopilación de información espacial son los introducidos en los dibujos, debido a una generalización del mapa, la calidad de reproducción, la deformación del material, la digitalización o escaneo, etcétera. Una gran parte de ellos puede observarse en la serie del MGN, aun cuando estos productos ya son de carácter digital. El hecho de que sean en este formato no implica en lo absoluto

que sean exactos o precisos; Thapa y Bossler rematan señalando que "... la introducción de computadoras en el manejo de datos espaciales ha introducido una falsa sensación de precisión..." (p. 839).

Respecto a la precisión y la exactitud, es interesante resaltar que existe un concepto útil para reducir la carga negativa que implica la generación de errores en los datos espaciales. Es lo que Salgé define como *precisión semántica*: "... la cualidad con que los objetos geográficos son descritos en concordancia con el modelo seleccionado. Relativo al significado de 'las cosas' del universo del discurso (la realidad), la precisión semántica se refiere a la pertinencia del significado del objeto geográfico más que a su representación geométrica..." (1995, p. 10).

Es entonces que importa la precisión semántica, cuando un mismo objeto geográfico, con un nivel de precisión determinada, puede ser más o menos útil para un modelado dependiendo del objeto del propio modelado. En nuestro caso de estudio no importaría que los metadatos del objeto geográfico no estuvieran completos (por ejemplo, el método de generación), sino que es más relevante la precisión geométrica del objeto en relación con el paso del tiempo.

Supongamos que la serie de datos del año 1 tuviera un error que se repite tal cual en el año 2 hasta el n : es probable que, de mantenerse constante en todos los años, los resultados del modelo analítico serían consistentes en el tiempo, pues el error se hubiera acumulado de forma histórica. Sin embargo, cuando este pasa del año 1 al n mutando, por ejemplo, su forma geométrica (como en el caso de las imágenes 1 y 2), entonces se tiene un problema de certidumbre en el modelo analítico que no es aceptable y, por ende, es necesario enmendarlos.

Por ello, una primera tarea del grupo de investigación fue realizar el ajuste espacial y la corrección topológica de la cartografía censal del MGN para las 100 ciudades más grandes del SUN 2010 (ver Anexo A) en 1995, 2000, 2005, 2010 y el 2015. Suponemos que este trabajo —ya en manos del

INEGI al momento de redactarlo— aumenta el valor agregado de los productos del Instituto al reducir errores topológicos, aumenta la confiabilidad de los resultados de análisis posteriores y orienta a la institución el método o métodos a utilizarse para corregir todas las series del MGN. Partimos de la base de que este reporte del método utilizado abonará al objetivo que tiene trazada *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*, el cual es "... propiciar el intercambio de experiencias relevantes [...] relativas al estado del arte en la generación de información estadística y geográfica...".

Es interesante resaltar el hecho de que, al ser nuestra investigación explícitamente espacial, tuvimos especial cuidado en observar posibles errores relacionados con el territorio. Sin embargo, no sabemos cuántos estudios longitudinales que han utilizado este MGN y que no son como el nuestro —es decir, que no generan métricas donde la geometría ni la geografía importan *a priori*— circulan en la actualidad en el mundo académico, acarreando errores derivados de la utilización de cocientes que usan el territorio como denominador (i. e. densidades).

Materiales y métodos

El proceso inicial del ajuste espacial comenzó con la solicitud al INEGI de las distintas ediciones del MGN para la consulta de las AGEB con el objetivo de garantizar la calidad y consistencia de los datos. Una vez consolidado el banco de datos y seleccionadas las AGEB de cada una de las 100 ciudades más grandes del SUN 2010, se aplicaron diversos métodos, los cuales se describen en este apartado.⁴

Como principal fuente de datos cartográficos de las AGEB fue considerado el MGN para 1995, 2000, 2005, 2010 y el 2015.

⁴ Para ampliar sobre el método seleccionado, es posible consultar el *Reporte técnico de la primera etapa del FOSEC* en <http://idegeo.centrogeo.org.mx/documents/2966/download>. El apartado específico sobre el arreglo de las bases de datos se encuentra a partir de la página 93 bajo el título *Anexo 2. Proceso de estandarización de la información geográfica de 100 ciudades en nivel de AGEB*.

De manera paralela a la solicitud de información, se obtuvieron capas por medio de descarga directa en el portal (2010 y 2000), las cuales se utilizaron para realizar las primeras pruebas de similitud entre proyecciones cartográficas. El resultado de este proceso desechó el supuesto que plantea que el desfase cartográfico era originado por la definición de proyecciones cartográficas diferentes y particulares de cada capa y fecha consultada.

Se re proyectaron todas y cada una de las capas a un sistema de coordenadas geográficas, sin lograr un mejor ajuste espacial entre ellas.⁵ Durante la revisión de los datos, se detectó en la capa del 2000 la existencia de un desfase de las AGEB urbanas del estado de Colima; polígonos que se movieron utilizando puntos de control en función de su correspondencia geográfica con los polígonos de localidad; además, se identificaron 346 AGEB que fueron eliminadas por estar duplicadas en las bases geográficas.

Las pruebas realizadas en función de la proyección cartográfica no resultaron favorables: los desfases se mantuvieron, aun cuando a cada capa se le asignó un *Datum ITRF92*.

Con esta comprobación se confirmó —con base en una respuesta por escrito del propio Instituto— la existencia de desfases al sobreponer las capas de información de las AGEB obtenidas del MGN versiones 1995 y 2000 con las actuales. Dicha situación no fue posible solucionarla debido a que las capas no fueron modificadas a partir de ortofotos o versiones actualizadas del Marco Geoestadístico por escapar a la tarea central de la investigación y por no estar contemplada esta tarea en los términos de referencia.

Además, las AGEB de 1995 y el 2000 se obtuvieron de los planos incluidos en los productos del Sistema para la Consulta de Información Censal (SCINCE) 2010, los cuales no están geo-

rreferidos y fueron elaborados *a mano alzada*, esto es, no tienen una base *raster* ortorrectificada. Estos archivos fueron trabajados con los siguientes procesos:

- Conversión de formato *dxf* a formato *Shapefile*.
- Asignación de sistema global de coordenadas (UTM).
- En algunos casos se reubicó la localidad a su espacio correspondiente.
- Edición de polígonos de AGEB para eliminar huecos y traslapes.
- Obtención de polígonos de localidad a partir del proceso anterior.
- Re proyección a *CCL-ITRF 1992*.
- Edición de localidades a nivel nacional con el fin de corregir posibles traslapes parciales o totales entre ellas, siempre y cuando el proceso de edición fuera mínimo.

Derivado de lo anterior, y de que en principio los planos no fueron creados para integrarse en un SIG, los archivos originales (sin ajuste espacial) presentan las siguientes características:

- Traslapes entre localidades.
- No uniformidad en la escala de la localidad.
- Los archivos no tienen un comportamiento homogéneo en la asignación de capas; puede darse el caso de que un elemento esté en más de una de ellas.
- No cumplen con reglas topológicas (conexión, traslapes, duplicados).
- Carecen de metadatos.

Una vez que se tuvo completa la serie cartográfica, se evaluó la pertinencia de utilizar el año 1990 como una capa más en el estudio. En esta evaluación se consideró que presentaba una gran disparidad con respecto a los años posteriores y que aumentaría el tiempo de dedicación para su ajuste espacial, además de que, para ese periodo, no se cuenta con información estadística de tipo económica. Por tal motivo, la serie cartográfica utilizada fue quinquenal, iniciando en 1995 y finalizando en el 2015.

⁵ El sistema de coordenadas geográficas en el que se re proyectaron todas las capas fue *North America Lambert Conformal Conic ITRF 1992*.

Con la serie completa de las AGEB, se iniciaron las pruebas que se utilizaron para definir el proceso de estandarización de las capas del MGN. De forma inicial, se evaluó la herramienta *Modify-Stretch Multiple* de *AutoCad*, la cual permite mover, escalar y expandir o contraer polígonos.

Este proceso implicó exportar la cartografía de las 100 ciudades correspondientes a las cinco fechas seleccionadas en formato *DWG*. Como prueba inicial se trabajaron las 10 primeras del SUN 2010 para cada periodo (ver *Anexo A*). Los resultados no fueron los deseados, ya que prevalecieron los desfases ocasionados por la limitada capacidad de ajuste que brinda la herramienta. Por lo tanto, se determinó que los resultados no garantizaban una adecuada calidad de la información.

Una vez desechada la posibilidad de utilizar la herramienta antes descrita, se buscaron dentro de la caja de *ArcGis* métodos que garantizaran mejores ajustes espaciales. Para ello, se consultó la sec-

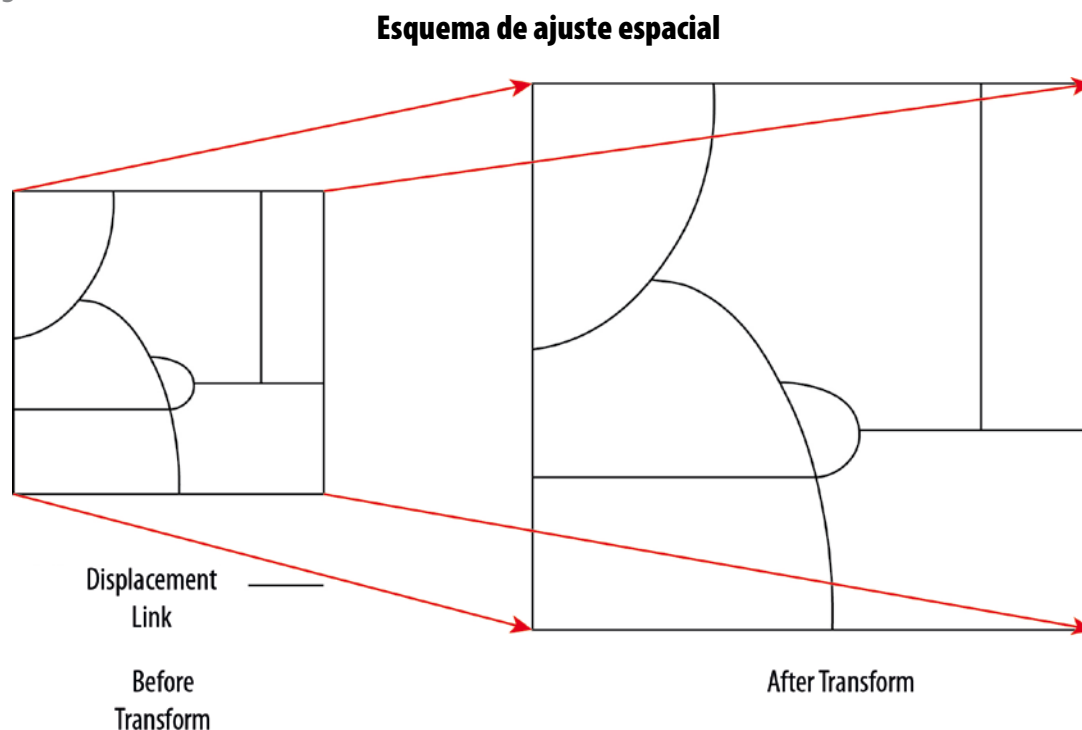
ción del escritorio de ayuda de *ArcGis 10.2* y se retomó el esquema de la imagen 3.

Las principales funciones de transformación disponibles en *ArcGis 10.2* se basan en la comparación de las coordenadas de puntos de origen y destino (también llamados puntos de control) con elementos gráficos espaciales denominados *vínculos de desplazamiento*. Se crearon vínculos de forma interactiva apuntando a ubicaciones de origen (por ejemplo, AGEB 1995) y destino conocidas (AGEB 2000).

Los métodos para el ajuste espacial que dispone *ArcGis 10.2* (*ArcMap*, *s/f*) son:

- Transformación afin (*transformation affine*).
- Transformación de similitud (*transformation similarity*).
- Transformación proyectiva (*transformation projective*).
- Media cuadrática y residual (*Rubbersheet*).
- Transformaciones ciegas (*Edge Snap*).

Imagen 3



Fuente: <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/editing-existing-features/GUID-BD818044-2DA8-40E9-B624-94B73843DE50-web.gif>

A continuación, se describe cada uno de ellos, así como los resultados del proceso de ajuste espacial para el caso de una zona de la ciudad de Aguascalientes.

- Transformación afín (*Affine*). Puede escalar los datos diferencialmente, sesgarlos, rotarlos y traducirlos. Su función es:

$$x' = Ax + By + C \quad y' = Dx + Ey + F$$

donde: x y y son las coordenadas de la capa de entrada y x' y y' son las coordenadas transformadas. A, B, C, D, E y F se determinan comparando la ubicación de los puntos de control de origen y destino. Escalan, sesgan, rotan y traducen las coordenadas de la capa. Este método requiere un mínimo de tres vínculos (ver imagen 4).

- Transformación de similitud (*Similarity*). Escala, rota y traduce los datos. No escalará los

ejes de forma independiente ni introducirá ningún sesgo. Se mantiene la relación de aspecto de las entidades transformadas, lo cual es importante si se desea mantener la forma relativa de las entidades. Su función es:

$$x' = Ax + By + C \quad y' = -Bx + Ay + F$$

donde: $A = s * \cos t$

$B = s * \sin t$

$C =$ traslación en dirección x

$F =$ traslación en dirección y

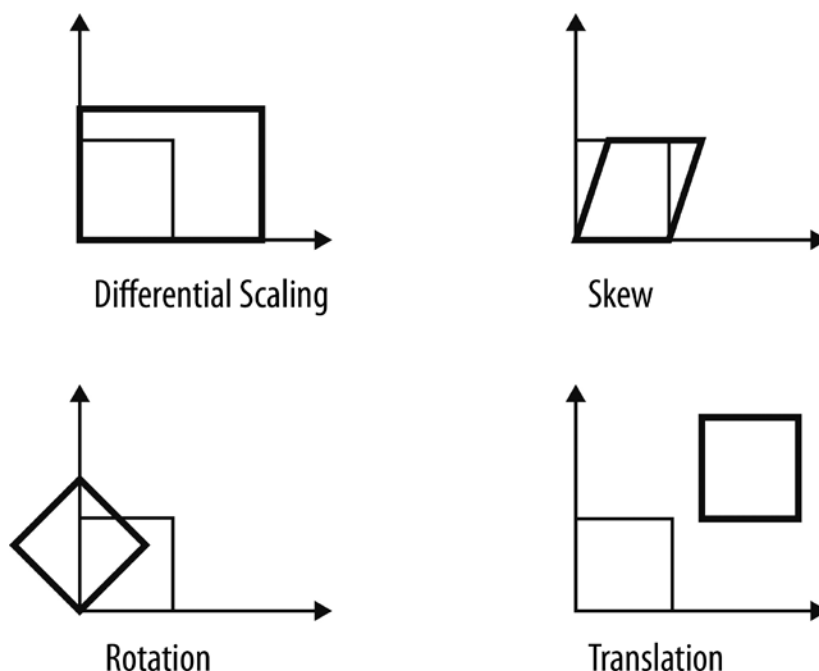
$s =$ cambio de escala (similar en las direcciones x, y)

$t =$ ángulo de rotación, medido en sentido contrario a las manecillas del reloj desde el eje x

Requiere un mínimo de dos vínculos de desplazamiento; sin embargo, se necesitan por lo menos tres para que se produzca un error cuadrático medio (ECM).

Imagen 4

Cuatro cambios posibles (*Affine*)



Fuente: desktop.arcgis.com

- Transformación proyectiva (*Projective*). Se basa en una fórmula más compleja que requiere un mínimo de cuatro vínculos de desplazamiento:

$$x' = (Ax + By + C) / (Gx + Hy + 1) \quad y' = (Dx + Ey + F) / (Gx + Hy + 1)$$

Este método se utiliza para transformar los datos capturados directamente de la fotografía aérea.

- Media cuadrática y residual (*Rubbersheet*). Los parámetros de transformación son un mejor ajuste entre los puntos de control de origen y destino; si se utilizan para transformar los puntos de control de origen actuales, las ubicaciones de salida transformadas no coincidirán con las de puntos de control de salida reales; esto se conoce como error

residual y es una medida del ajuste entre las ubicaciones reales y las transformadas de los puntos de control de salida (ver imagen 5).

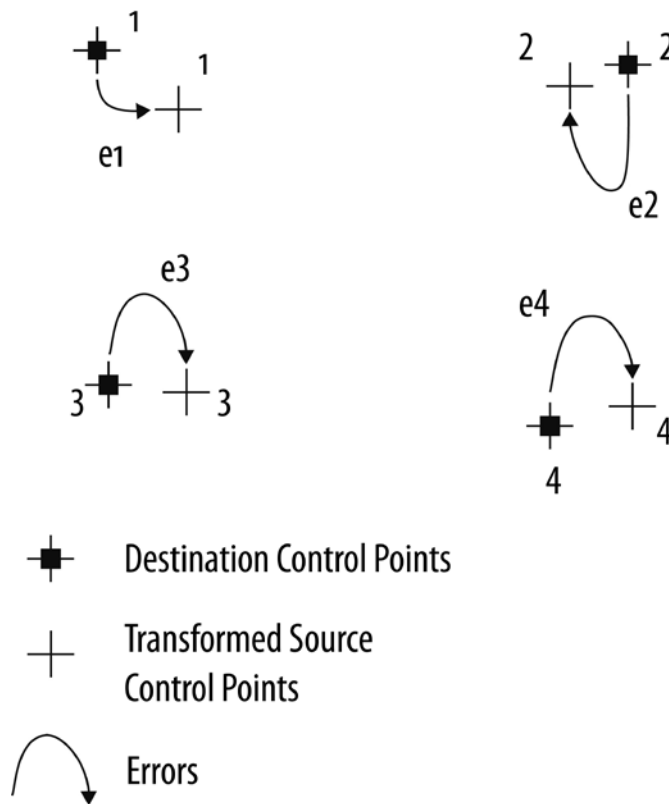
El ECM mide las fallas entre los puntos de control de destino y las ubicaciones transformadas de los puntos de control de origen:

$$ECM \text{ error} = \sqrt{\frac{e_1^2 + e_2^2 + e_3^2 \dots + e_n^2}{n}}$$

La transformación se obtiene a partir de mínimos cuadrados, por lo que se pueden proporcionar más vínculos de los necesarios; hay que especificar un mínimo de tres para obtener una transformación que dé lugar a un ECM. Por lo general, cuantos más vínculos ubique para una transformación, más precisa será.

Imagen 5

Ubicación relativa de cuatro puntos de control de destino y los puntos de control de origen transformados



Fuente: desktop.arcgis.com

- Transformaciones ciegas (*Edge Snap*). Es posible configurar la transformación creando vínculos de desplazamiento con el extremo de origen en ubicaciones conocidas y el extremo de destino en puntos en el espacio temporales. Después de esto, podrá utilizar la tabla de vínculos para editar las coordenadas de destino de estos vínculos y que pasen a ser las ubicaciones del mundo real correspondientes.

Procesamiento

Este apartado presenta el proceso desarrollado para seleccionar el método que dio el mejor ajuste espacial en las series de datos espaciales para su posterior corrección topológica. Se aplicaron cuatro de los cinco métodos citados en la sección anterior. El *Edge Snap* se omitió debido a que requiere mayor tiempo y trabajo destinado en la asignación manual de valores conocidos x, y necesarios para lograr el ajuste espacial.

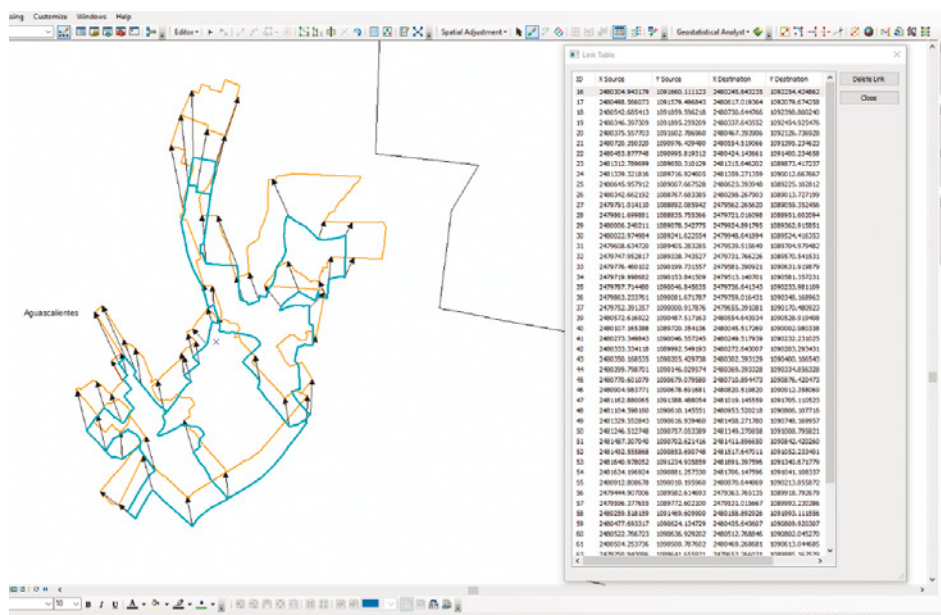
Ajuste espacial

Para seleccionar el método que garantizara el mejor ajuste espacial, se aplicaron los cuatro métodos a las primeras 10 ciudades del SUN. Se definieron dos criterios para escoger el mejor método de ajuste espacial: 1) correspondencia en su forma, que considera mayor similitud geométrica entre la capa ajustada con la de referencia (valoración visual) y 2) mayor proporcionalidad de la superficie obtenida en la capa ajustada espacialmente con respecto a la superficie de la capa de referencia (valoración proporcional), es decir, qué tanto se parecen las superficies. Para fines ilustrativos, solo se presenta una porción de la ciudad de Aguascalientes, ejemplificando la aplicación de los criterios antes mencionados.

Como primer paso para la aplicación de cualquiera de los cuatro métodos, fue necesario asignar los puntos de control en la capa por ajustar, con respecto a la capa de referencia (ajustar 1995 con referencia a la 2000), para cada una de las zonas seleccionadas (ver imagen 6).

Imagen 6

Asignación de puntos de control



Fuente: elaboración propia con base en información del INEGI de 1995 y el 2000.

Una vez que se han asignado los puntos de control, se ejecutan los métodos de ajuste espacial disponibles en *ArcGis 10.2* de la siguiente manera: abrir en el menú secundario *Spatial Adjustment* y seleccionar el método de transformación, en este caso el *Affine* y, posteriormente, *Adjust* (ver imagen 7).

Enseguida, abrir en el menú secundario *Spatial Adjustment* y seleccionar el método de transformación *Projective* y, después, *Adjust* (ver imagen 8).

Después, abrir en el menú secundario *Spatial Adjustment* y seleccionar el método de transformación *Similarity* y, enseguida, *Adjust* (ver imagen 9).

Para seguir, abrir en el menú secundario *Spatial Adjustment* y seleccionar el método de transformación *Rubbersheet* y dar clic en *Adjust* (ver imagen 10).

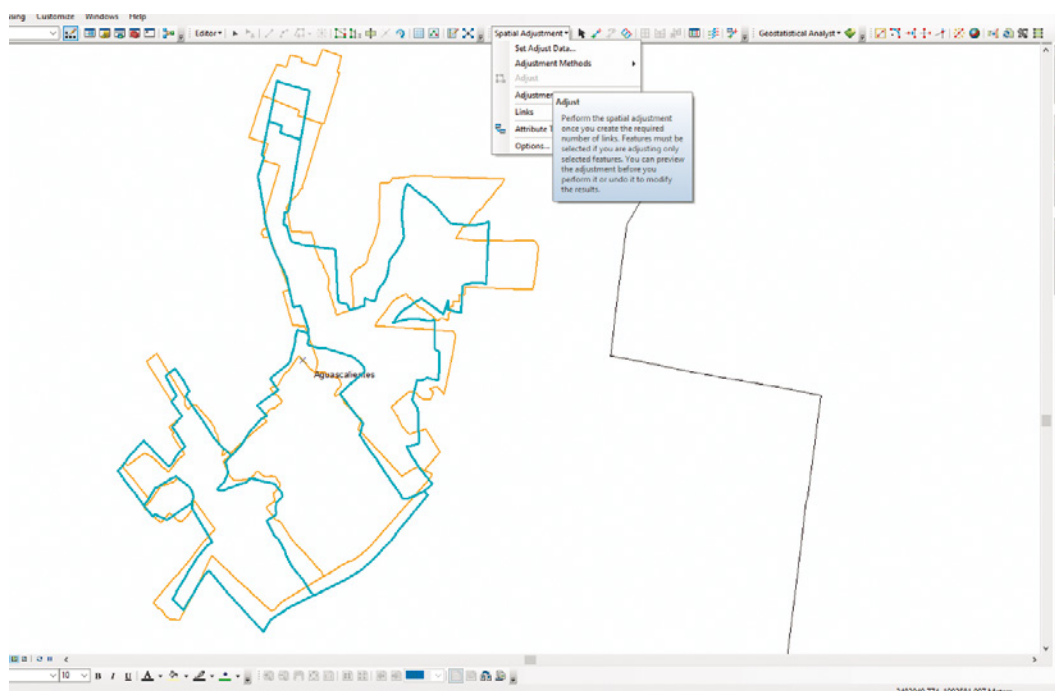
A manera de ejemplo, y con base en la aplicación de los criterios, se logra un mejor ajuste es-

pacial en función de la valoración visual con el método *Rubbersheet* (ver imagen 10) que con los otros tres. En términos de proporcionalidad, la superficie ajustada representa 96 % con respecto a la total de las AGEB de referencia. Los otros métodos obtuvieron proporcionalidades por debajo de este porcentaje.

Con base en la aplicación del método *Rubbersheet* se obtuvo un mejor ajuste espacial, por lo que fue seleccionado y aplicado para la transformación de las cuatro capas cartográficas anteriores al 2015 para cada una de las 100 ciudades del SUN. Este se realizó en dos momentos: 1) en el proceso que inició con el ajuste de la capa de AGEB 1995 con referencia a la capa del 2000, y de manera sucesiva hasta llegar al del 2010 respecto al 2015 y 2) utilizando las capas previamente ajustadas (resultado del proceso anterior). Las modificaciones se realizaron ahora en sentido inverso, tomado como referencia las AGEB 2015, para ajustar el 2010 de

Imagen 7

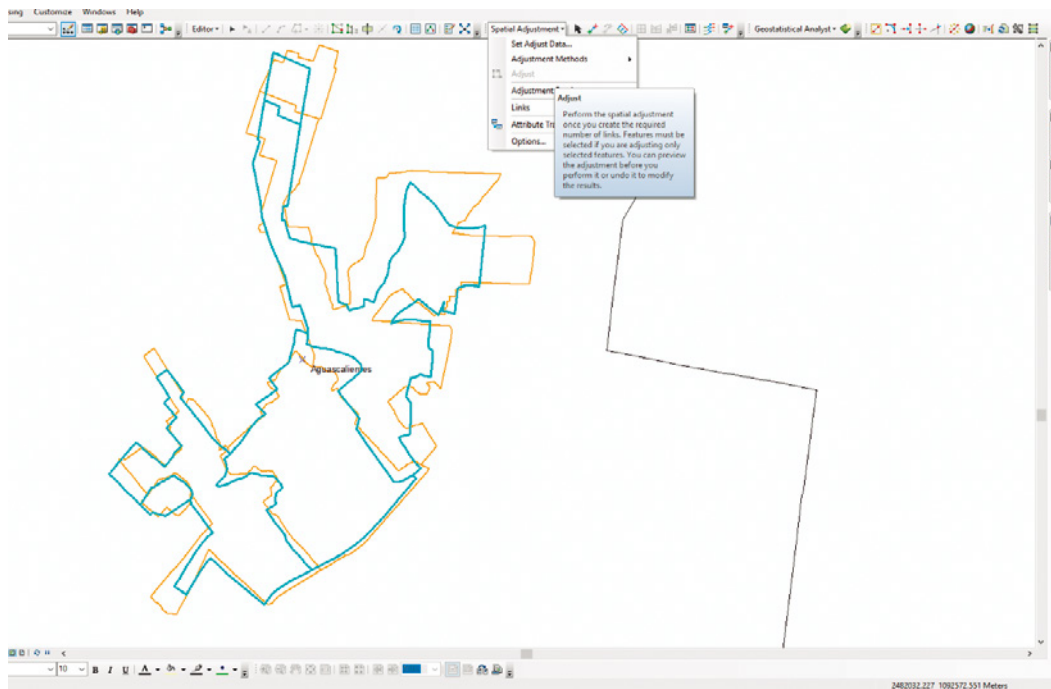
Resultado *Affine*



Fuente: elaboración propia con base en información del INEGI de 1995 y el 2000.

Imagen 8

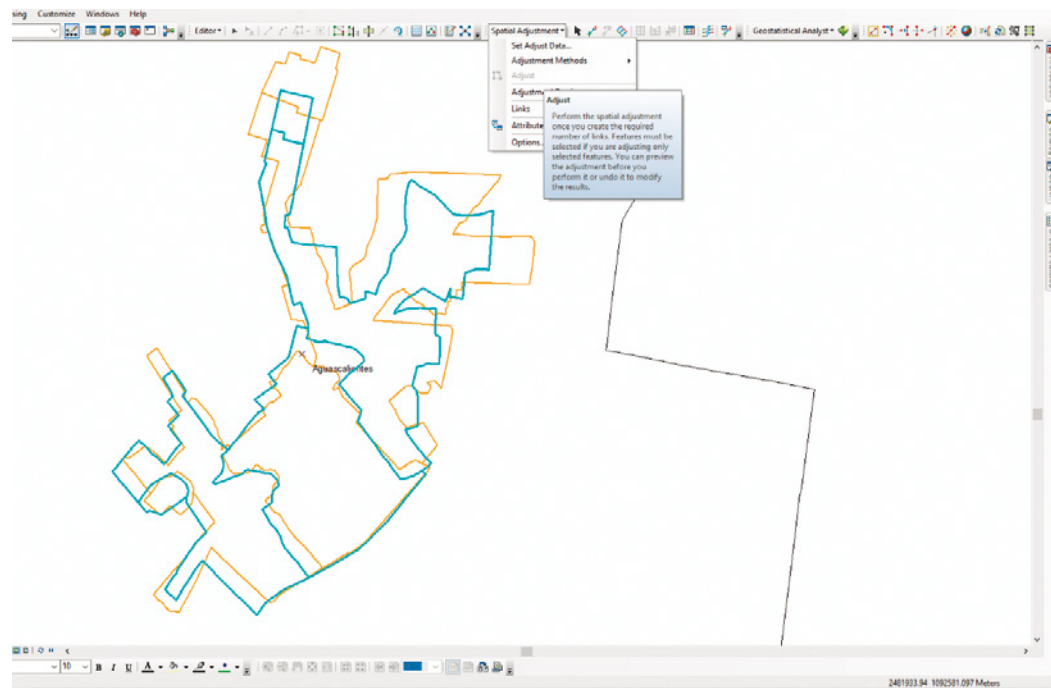
Resultado *Projective*



Fuente: elaboración propia con base en información del INEGI de 1995 y el 2000.

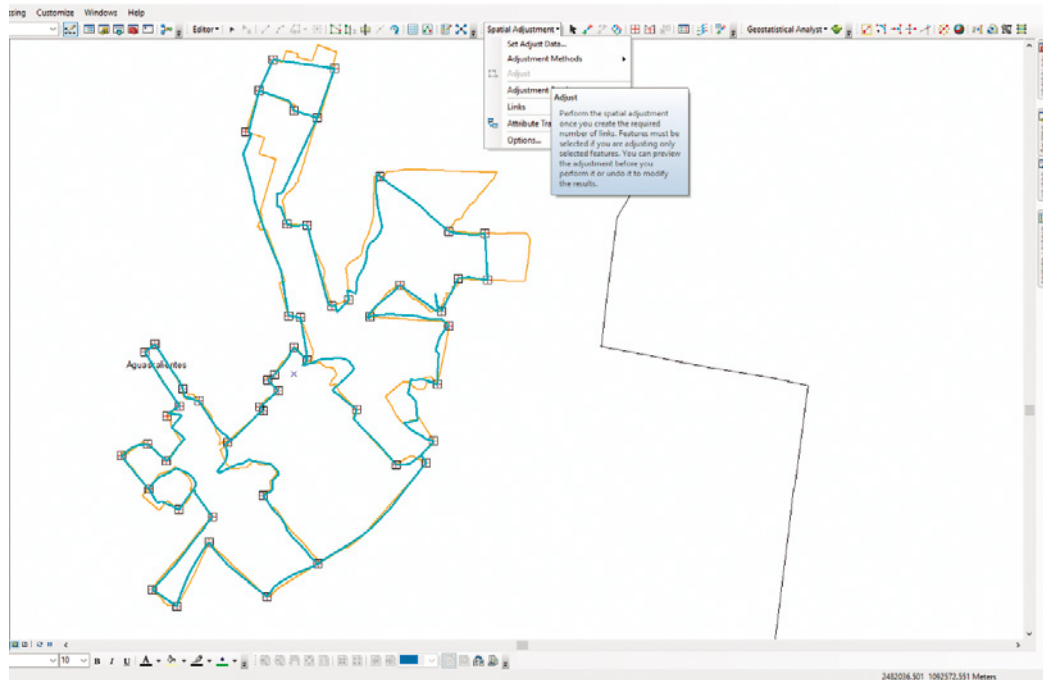
Imagen 9

Resultado *Similarity*



Fuente: elaboración propia con base en información del INEGI de 1995 y el 2000.

Resultado *Rubbersheet*



Fuente: elaboración propia con base en información del INEGI de 1995 y el 2000.

manera sucesiva hasta completar todos los de las capas posteriores. Como respaldo a los criterios considerados en este proceso, se utilizaron como referencia imágenes satelitales⁶ (procurando la correspondencia con las fechas de las capas), sobre todo para ajustar las AGEB periurbanas. Cabe señalar que las modificaciones a este tipo de polígonos requirieron de mayor dedicación para subsanar un mayor número de deformaciones geométricas, sobreescalamiento y sobredelimitaciones que incluyen territorios no urbanizados.

Corrección topológica

Una vez que se revisaron y ajustaron los polígonos de las AGEB (1995, 2000, 2005 y 2010), se obtuvieron las capas de información geográfica en formato *shapefile* para su revisión y corrección topológica.

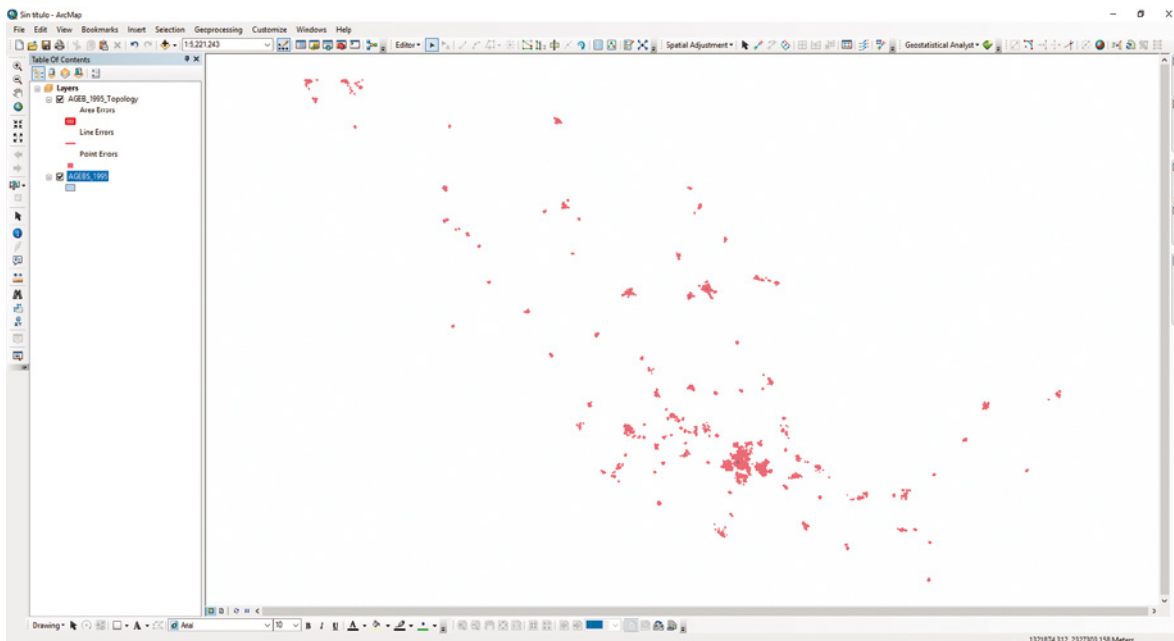
⁶ World Imagery-Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo y el GIS User Community.

La topología garantiza la calidad de la información debido a que permite obtener entidades de polígonos que comparten una geometría coincidente, es decir, deben tener límites comunes; además, incluye un modelo de datos topológicos que permite integridad, en un formato de almacenamiento abierto, reglas y coordenadas integradas entre las entidades con geometría compartida.

Para el uso de la herramienta *topology* de ArcGIS 10.2, se obtuvo un archivo *geodatabase* (*gdb*) compuesto por los archivos *dataset* (conjunto de datos) de cada una de las capas de las AGEB corregidas espacialmente. Con base en el conjunto de datos antes mencionado, el siguiente paso fue crear la topología para cada *dataset*, bajo los criterios de sobreposición y espacios entre polígonos, cuyos resultados son útiles para identificar errores y excepciones, que serán las principales referencias para consultar, editar, validar y corregir los errores identificados en las topologías obtenidas. ArcGIS 10.2 cuenta con un marco de edición y automatización de datos que se

Imagen 11

Topología de AGEB 1995 de las 100 ciudades del SUN



Fuente: elaboración propia con base en información del INEGI de 1995.

utiliza para crear, mantener y validar la integridad topológica y para editar entidades compartidas.⁷

A continuación, se presentan los reportes del proceso topológico para cada una de las capas de las AGEB, señalando el total de errores topológicos divididos en fallas de superposición de polígonos y por separación entre polígonos.

El total de los errores topológicos se muestran en la imagen 12.

Imagen 12

Errores totales

Rule Type	Class 1	Class 2	Shape	Feature 1	Feature 2	Exception
Must Not Have Ga	AGEBS_1995		Polyline	0	0	False
Must Not Have Ga	AGEBS_1995		Polyline	0	0	False

Fuente: elaboración propia con datos obtenidos de ArcGIS 10.2.

Los errores detectados para este ejemplo por sobreposición de polígonos corresponden a 51 % del total de los detectados (ver imagen 13).

⁷ Para mayor información sobre la corrección de la topología en ArcGIS, consultar <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/topologies/topology-in-arcgis.htm>

Imagen 13

Errores por superposición

Rule Type	Class 1	Class 2	Shape	Feature 1	Feature 2	Exception
Must Not Overlap	AGEBS_1995		Polygon	17768	17786	False
Must Not Overlap	AGEBS_1995		Polygon	17763	17808	False

Fuente: elaboración propia con datos obtenidos de ArcGIS 10.2.

Del total de errores identificados, 49% corresponde a separaciones entre polígonos; no obstante, hay que descontar de esta proporción aquellas separaciones naturales existentes entre ciudades (ver imagen 14).

Imagen 14

Errores por separación

Rule Type	Class 1	Class 2	Shape	Feature 1	Feature 2	Exception
Must Not Have Ga	AGEBS_1995		Polyline	0	0	False
Must Not Have Ga	AGEBS_1995		Polyline	0	0	False

Fuente: elaboración propia con datos obtenidos de ArcGIS 10.2.

Al final, para cada conjunto de datos, se revisaron y corrigieron todos y cada uno de los errores por superposición, así como por separación entre polígonos dentro de los límites de las 100 ciudades para cada una de las cinco capas geográficas.

Cuadro 2

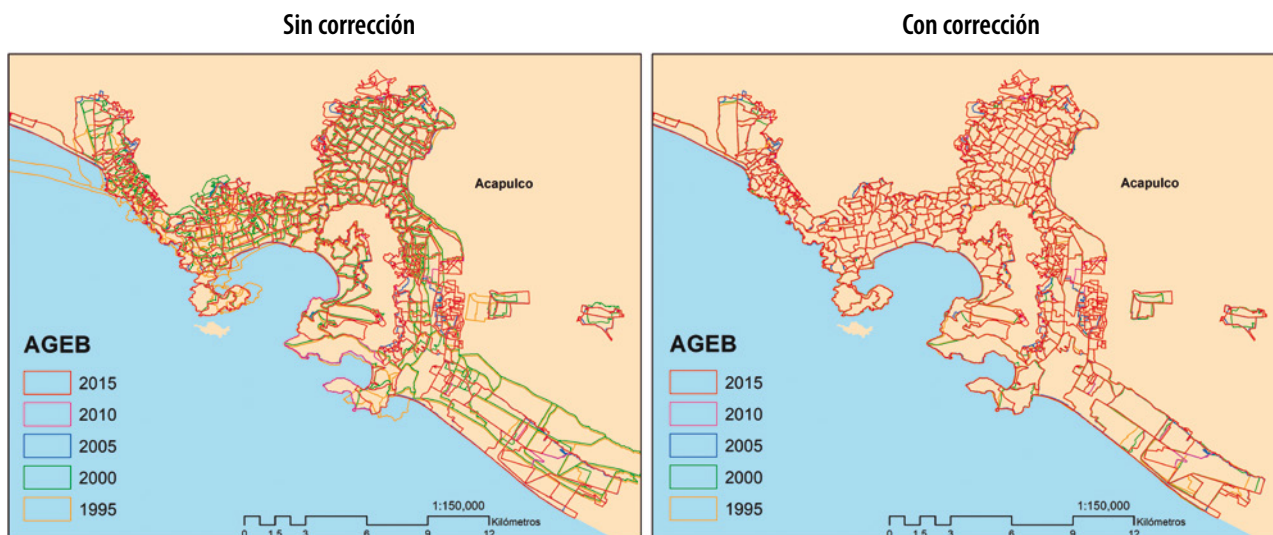
Resumen de errores en función de criterios de superposición y separación de polígonos

Dataset (AGEB)	Total	Superposición	%	Separación	%
1995	3 556	1 807	50.82	1 749	49.18
2000	1 534	530	34.55	1 004	65.45
2005	1 608	510	31.72	1 098	68.28
2010	964	43	4.46	921	95.54
2015	1 325	106	8.00	1 219	92.00
Total	8 987	2 996	33.34	5 991	66.66

Fuente: elaboración propia con base en mediciones obtenidas en ArcGIS 10.2.

Imagen 15

Resultado del ajuste espacial de la ciudad de Acapulco, Guerrero



Fuente: elaboración propia con información del INEGI de 1995, 2000, 2005, 2010 y el 2015.

Resultados

Lo obtenido en las capas ajustadas correspondientes a 1995, 2000, 2005, 2010 y el 2015 es satisfactorio de acuerdo con el criterio de valoración visual (ver imagen 15), que tal vez sea el que representa un mayor esfuerzo y reafirma que el método debe ser supervisado. De forma paralela, se realizó el cálculo de las superficies por cada una de las ciudades

ajustadas, poniendo mayor atención a las que en las capas originales reportaron una disminución de su superficie en una fecha posterior.

En términos generales, con el ajuste espacial se mejoró la forma y correspondencia espacial de los polígonos de las AGEB; además, se corrigieron las superficies de aquellas ciudades que se *comprimían*.

Comparación cartográfica con ajuste espacial

Si bien no existe una medida precisa para conocer el grado de ajuste espacial entre la capa origen con la ajustada, se optó por generar estadísticos descriptivos que dieran idea de la variabilidad de los datos de superficie y perímetro (forma). Al revisar las capas de las AGEB originales (tomadas directamente del INEGI), se encontraron errores topológicos por sobreposición, duplicidad y separación de polígonos.

Las pruebas se realizaron utilizando las variables de superficie en hectáreas y el perímetro en kilómetros en dos escalas:

- A nivel de AGEB, se analizaron en su totalidad los polígonos geoestadísticos que forman una capa geográfica en un año específico (1995, 2000, 2005, 2010 y el 2015). Para realizar comparaciones respecto a los cambios entre la capa original y la ajustada de un mismo año, se generaron estadísticos descriptivos (ver *Anexo B*).
- Para comparar los cambios tanto en la superficie como en la forma en nivel ciudad, se utilizaron los polígonos de las 100 ciudades de cada año a través de una medida de razón perímetro-área (RPA).

Para el primer caso, los estadísticos que se utilizaron para comparar los cambios fueron: la desviación estándar y la media aritmética para obtener el coeficiente de variación (CV),⁸ además de la medida de *curtosis* para ver el afilamiento de la curva de distribución del conjunto de datos y el sesgo o asimetría de los mismos.

En términos generales, las capas geográficas que registraron mayores cambios de acuerdo con sus CV (superficie y perímetro) correspondieron a 1995 y el 2000. Las principales causas que originaron cambios significativos en sus estadísticos fueron la corrección de polígonos duplicados y sobredimensionados correspondientes, sobre todo, a polígonos periféricos de las ciudades.

De forma paralela, en todos los casos analizados se mantuvo una asimetría positiva con sesgo a la derecha, con mejoría en su mayoría en las AGEB de 1995 y el 2000, considerando que el valor 0 indica que el conjunto de datos tiene una distribución simétrica. Las medidas de afilamiento o *curtosis*, obtenidas en todas las capas de AGEB ajustadas espacialmente, disminuyeron en términos absolutos. Este comportamiento demuestra que, mediante el ajuste del conjunto de las AGEB, estas continúan

⁸ Medida que se emplea para comparar la variabilidad de los datos (en este caso, la superficie) de dos capas geográficas: AGEB originales y ajustadas. A mayor valor, mayor variabilidad.

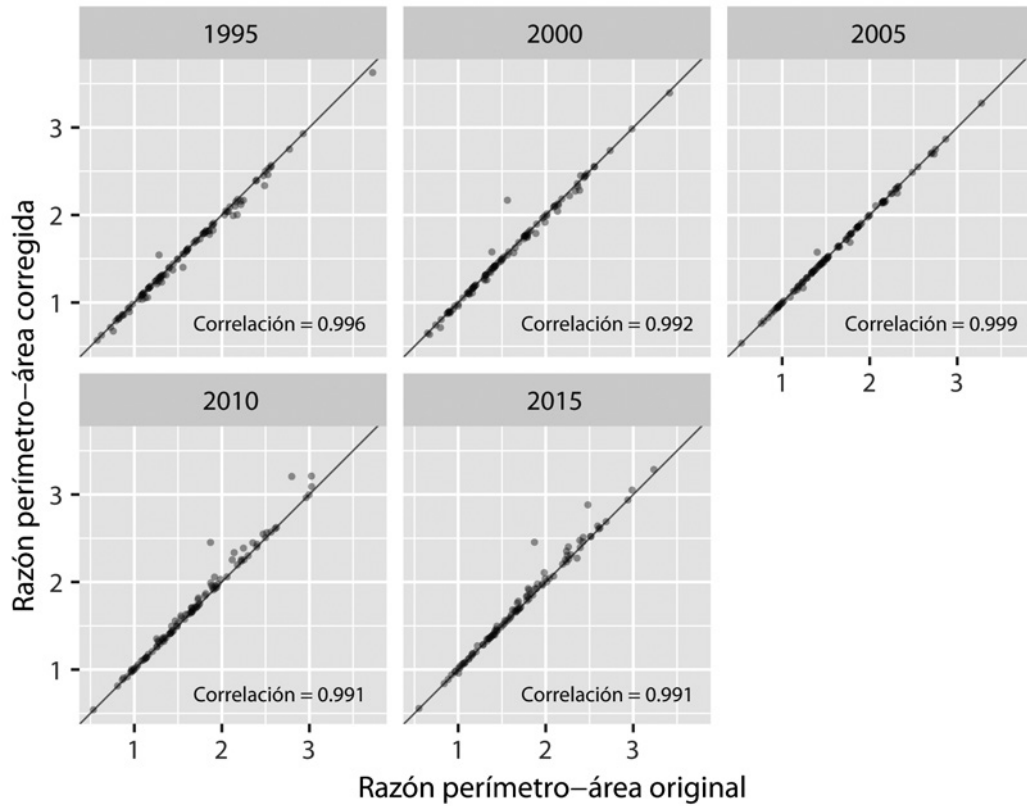
Cuadro 3

Coefficiente de variación. Variación del total de las AGEB por año

Año	Superficie (ha)		Perímetro (km)	
	Originales	Ajustados	Originales	Ajustados
1995	1.401	1.173	0.552	0.540
2000	1.461	1.220	0.586	0.573
2005	1.263	1.248	0.596	0.591
2010	1.315	1.305	0.626	0.625
2015	1.353	1.354	0.649	0.653

Fuente: elaboración propia con información del INEGI de 1995, 2000, 2005, 2010 y el 2015.

Razón perímetro-área, AGEB originales vs. ajustadas



Fuente: elaboración propia con información del INEGI de 1995, 2000, 2005, 2010 y el 2015.

con una distribución en forma de pico por arriba de una curva normal estandarizada; sin embargo, el valor absoluto disminuyó, lo cual significa que su afilamiento se aplanó (ver *Anexo B*).

Los cambios posteriores al ajuste espacial de las ciudades analizadas en relación con su superficie y forma se muestran mediante los valores obtenidos del coeficiente de correlación lineal. Dicha métrica se considera como una medida que permite calcular el grado de asociación lineal entre dos variables cuantitativas (*ciudades originales 1995* y *ciudades corregidas 1995*). Se dice que existe una asociación lineal perfecta cuando la correlación de la muestra es 1. Se trata de ver qué tan parecidos son los valores del RPA-original con el ajustado en un mismo año (ver imagen 16).

En términos generales, las ciudades que se encuentran alejadas de las líneas de ajuste son las que requirieron de mayores ajustes espaciales y su distancia refleja el grado de modificación del ajuste. Ciudades que se contraían, mediante el ajuste aumentaron tanto la superficie como el perímetro.

Conclusiones

En este trabajo se resumieron los procesos de ajuste espacial y corrección topológica mediante los cuales se corrigieron las ediciones del MGN a nivel AGEB para 1995, 2000, 2005, 2010 y el 2015 para las 100 ciudades más grandes del SUN 2010, lo que, a su vez, permitió correr un modelo longitudinal dentro del proyecto INEGI-CONACYT *Productividad y Forma Urbana en México 1995-2015* que correla-

cionara crecimiento físico, forma urbana y productividad laboral con mayor certeza geoespacial.

Con lo anterior, se pone a consideración del INEGI el método utilizado de ajuste espacial y corrección topológica con el fin de que se aplique en el resto de las ciudades del SUN para los periodos ya mencionados.

Con el método propuesto, se obtienen archivos vectoriales que integran un sistema único de carácter nacional diseñado por el INEGI para diversos años, ajustado espacialmente y corregido en su topología, con el propósito de garantizar comparabilidad espacial con productos similares de otros años.

Para el proceso de ajuste espacial de las AGEB se utilizó el *software ArcGis 10.2*. El método seleccionado fue el *Rubbersheet*, el cual permite la transformación para un mejor ajuste entre los puntos de control de origen y destino.

Por su parte, la corrección topológica se realizó una vez concluido el proceso de ajuste espacial de los polígonos (AGEB), mediante el cual se obtuvo una capa de información geográfica en formato *shapefile* para su revisión y corrección topológica para, después, generar un modelo de datos topológicos (que permite integridad) en un formato de almacenamiento abierto, reglas topológicas y coordenadas topológicamente integradas entre las entidades con geometría compartida.

Los productos obtenidos son cinco capas geográficas estandarizadas y ajustadas espacialmente, que permiten realizar comparaciones espacio-temporales y, sobre todo, obtener métricas de base espacial con mayor confiabilidad. Estos se entregaron al INEGI como parte del proceso de transferencia del conocimiento en el marco del ya mencionado Fondo Sectorial INEGI-CONACYT.

Los estadísticos obtenidos de la evaluación de las capas con ajuste espacial y corrección topológica muestran una mejoría en comparación con las bases espaciales de origen, por lo que, aun con errores derivados del nivel de precisión del traba-

jo geoespacial, se asume que esta nueva serie del MGN es, por mucho, mejor que trabajar de forma comparativa con las bases del MGN sin corregir.

Por último, debe mencionarse que se asume que existen otros métodos más precisos para corregir las desviaciones ya señaladas en las diferentes ediciones del MGN; sin embargo, también se presume que este método es mucho más económico al concentrarse más en una mayor precisión que en una exactitud superior.

Fuentes

- Amindarbari, R. y A. Sevtsuk. *Metropolitan Form Analysis Toolbox for ArcGIS v 10.2 and v 10.3.1*. City Form Lab. 2015 (DE) http://media.voog.com/0000/0036/2451/files/20150913_MFA_Help.pdf
- Angel, S., J. Parent y D. L. Civco. "Ten compactness properties of circles: Measuring shape in geography", in: *The Canadian Geographer/Le Géographe Canadien*. 54(4), 2010, pp. 441-461.
- Boyko, C. y R. Cooper. "Clarifying and re-conceptualising density", in: *Progress in Planning*. 76(1), 2011, pp. 1-61.
- Burt, J., G. Barber y D. Rigby. *Elementary statistics for geographers*. The Guilford Press, 2009.
- CONAPO y SEDESOL. *Sistema Urbano Nacional*. México, CONAPO-SEDESOL, 2012.
- ESRI. *Acerca de las transformaciones de ajuste espacial*. ArcMap. (s/f) consultado en <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/editing-existing-features/about-spatial-adjustment-transformations.htm#GUID-FA3A8070-CB17-4882-BC69-E70F23F84734>
- Ewing, R., R. Pendall y D. Chen. *Measuring sprawl and its impact*. Smart Growth America, 2002.
- Galster, G., R. Hanson, M. Ratcliffe, H. Wolman, S. Coleman y J. Freihage. "Wrestling sprawl to the ground: Defining and measuring an elusive concept", in: *Housing policy debate*. 12(4), 2001, pp. 681-717.
- Goodchild, M y S. Gopal. *The accuracy of spatial databases*. CRC Press, 1989.
- Guptill, S. y J. Morrison. *Elements of spatial data quality*. Elsevier, 1995.
- INEGI. *Marco Geoestadístico Nacional. Áreas Geoestadísticas Básicas Urbanas. Encuesta Intercensal 2015*. México, INEGI, 2015.
- _____. *Marco Geoestadístico Nacional. Áreas Geoestadísticas Básicas Urbanas. Encuesta Intercensal, versión 5.0; Censo de Población y Vivienda 2010*. México, INEGI, 2010.
- _____. *Manual de cartografía geoestadística. Levantamiento. Censo de Población y Vivienda 2010*. México, INEGI, 2010a.
- _____. *Marco Geoestadístico Nacional. Áreas Geoestadísticas Básicas Urbanas. II Censo de Población y Vivienda 2005, versión 1.0*. México, INEGI, 2005.

- _____. *Marco Geoestadístico Nacional. Áreas Geoestadísticas Básicas Urbanas. Censo General de Población y Vivienda 2000*. México, INEGI, 2000.
- _____. *Marco Geoestadístico Nacional. Áreas Geoestadísticas Básicas Urbanas. I Censo de Población y Vivienda 1995*. México, INEGI, 1995.
- Lovejoy, S. "Area-Perimeter Relation for Rain and Cloud Areas", in: *Science*. 216(4542), 1982, pp. 185-187 (DE) <https://doi.org/10.1126/science.216.4542.185>
- Montejano, J., C. Caudillo, P. Monkkonen, E. Guerra, J. M. Núñez, B. Garza, G. Ávila y S. Medina. "Urban Form and Productivity in Mexico 1995-2015", in: *European Journal of Sustainable Development*. 9(1), 2020, pp. 300-316.
- Montejano, J., P. Monkkonen, E. Guerra y C. Caudillo. *The costs and benefits of urban expansion: Evidence from Mexico, 1990-2010*. Lincoln Institute of Land Policy, 2019 (DE) <https://www.lincolninst.edu/publications/working-papers/costs-benefits-urban-expansion>
- ONU-Hábitat y SEDESOL. *Estado actual de las ciudades de México 2011*. ONU-Hábitat-SEDESOL, 2011.
- Pereira, R., V. Nadalin, L. Monasterio y P. Albuquerque. "Urban centrality: A simple index", in: *Geographical Analysis*. 45(1), 2013, pp. 77-89.
- Salgé, F. "Semantic accuracy", in: Guptill, S. y J. Morrison. *Elements of spatial data quality*. Elsevier, 1995, pp. 139-151.
- Thapa, K. y J. Bossler. "Accuracy of spatial data used information systems", in: *Photogrammetric Engineering & Remote Sensing*. 58(6), 1992, pp. 835-841.
- Tsai, Y. *Travel-efficient Urban Form: A Nationwide Study of Small Metropolitan Areas*. 2001.
- VanDerWaal, J., L. Falconi, S. Januchowski, L. Shoo y C. Storlie. *SDMTools: Tools for processing data associated with species distribution modelling exercises (Version 1.1-221.1)* [R package]. 2019.

Anexo A

Continúa

Primeras 100 ciudades del SUN 2010

Clave SUN 2010	Ciudad	Entidad
1	Aguascalientes	Aguascalientes
2	Tijuana	Baja California
3	Mexicali	Baja California
4	La Laguna	Coahuila de Zaragoza
5	Saltillo	Coahuila de Zaragoza
6	Monclova-Frontera	Coahuila de Zaragoza
7	Piedras Negras	Coahuila de Zaragoza
8	Colima-Villa de Álvarez	Colima
9	Tecomán	Colima
10	Tuxtla Gutiérrez	Chiapas
11	Juárez	Chihuahua
12	Chihuahua	Chihuahua
13	Valle de México	Ciudad de México, México e Hidalgo
14	León	Guanajuato
15	San Francisco del Rincón	Guanajuato
16	Moroleón-Uriangato	Guanajuato
17	Acapulco	Guerrero
18	Pachuca	Hidalgo
19	Tulancingo	Hidalgo
20	Tula	Hidalgo

Primeras 100 ciudades del SUN 2010

Clave SUN 2010	Ciudad	Entidad
21	Guadalajara	Jalisco
22	Puerto Vallarta	Jalisco
23	Ocotlán	Jalisco
24	Toluca	Estado de México
25	Morelia	Michoacán de Ocampo
26	Zamora-Jacona	Michoacán de Ocampo
27	La Piedad-Pénjamo	Guanajuato
28	Cuernavaca	Morelos
29	Cuatla	Morelos
30	Tepic	Nayarit
31	Monterrey	Nuevo León
32	Oaxaca	Oaxaca
33	Tehuantepec	Oaxaca
34	Puebla-Tlaxcala	Tlaxcala
35	Tehuacán	Puebla
36	Querétaro	Querétaro
37	Cancún	Quintana Roo
38	San Luis Potosí-Soledad de Graciano Sánchez	San Luis Potosí
39	Rioverde-Ciudad Fernández	San Luis Potosí
40	Guaymas	Sonora
41	Villahermosa	Tabasco
42	Tampico	Tamaulipas
43	Reynosa-Río Bravo	Tamaulipas
44	Matamoros	Tamaulipas
45	Nuevo Laredo	Tamaulipas
46	Tlaxcala-Apizaco	Tlaxcala
47	Veracruz	Veracruz de Ignacio de la Llave
48	Xalapa	Veracruz de Ignacio de la Llave
49	Poza Rica	Veracruz de Ignacio de la Llave
50	Orizaba	Veracruz de Ignacio de la Llave
51	Minatitlán	Veracruz de Ignacio de la Llave
52	Coatzacoalcos	Veracruz de Ignacio de la Llave
53	Córdoba	Veracruz de Ignacio de la Llave
54	Acayucan	Veracruz de Ignacio de la Llave
55	Mérida	Yucatán
56	Zacatecas-Guadalupe	Zacatecas
57	Celaya	Guanajuato

Primeras 100 ciudades del SUN 2010

Clave SUN 2010	Ciudad	Entidad
58	Tianguistenco	Estado de México
59	Teziutlán	Puebla
60	Ensenada	Baja California
63	Campeche	Campeche
65	Manzanillo	Colima
66	Tapachula de Córdova y Ordóñez	Chiapas
68	Guanajuato	Guanajuato
69	Irapuato	Guanajuato
73	Chilpancingo de los Bravo	Guerrero
94	Ciudad Lázaro Cárdenas	Michoacán de Ocampo
95	Uruapan	Michoacán de Ocampo
98	Zitácuaro	Michoacán de Ocampo
104	San Juan Bautista Tuxtepec	Oaxaca
119	Chetumal	Quintana Roo
121	Ciudad Obregón	Sonora
123	Cárdenas	Tabasco
134	Túxpam de Rodríguez Cano	Veracruz de Ignacio de la Llave
137	Fresnillo	Zacatecas
144	La Paz	Baja California Sur
145	Ciudad del Carmen	Campeche
148	Ciudad Acuña	Coahuila de Zaragoza
160	Comitán de Domínguez	Chiapas
172	San Cristóbal de las Casas	Chiapas
180	Cuauhtémoc	Chihuahua
181	Delicias	Chihuahua
182	Hidalgo del Parral	Chihuahua
188	Victoria de Durango	Durango
201	Salamanca	Guanajuato
211	Iguala de la Independencia	Guerrero
228	Ciudad Guzmán	Jalisco
234	Lagos de Moreno	Jalisco
242	Tepatitlán de Morelos	Jalisco
252	Apatzingán de la Constitución	Michoacán de Ocampo

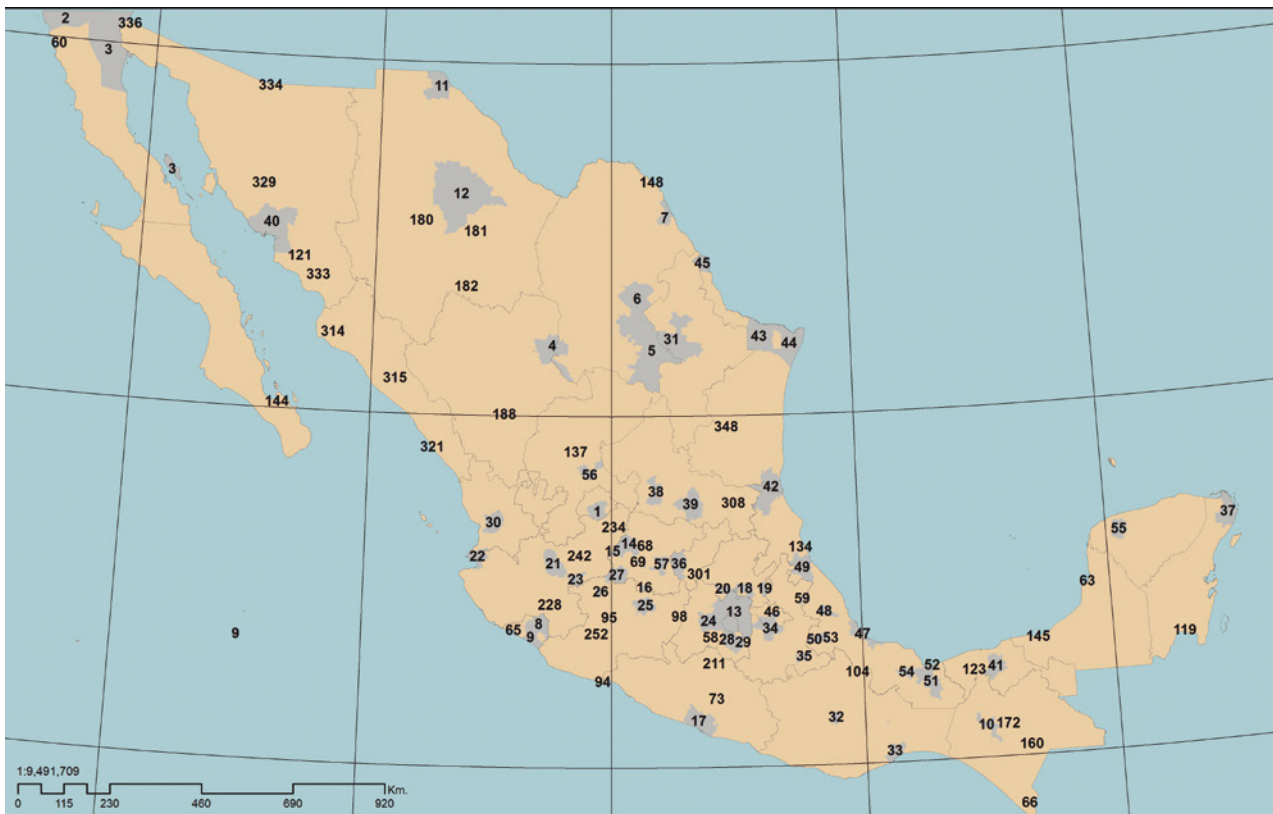
Primeras 100 ciudades del SUN 2010

Clave SUN 2010	Ciudad	Entidad
301	San Juan del Río	Querétaro
308	Ciudad Valles	San Luis Potosí
314	Los Mochis	Sinaloa
315	Culiacán Rosales	Sinaloa
321	Mazatlán	Sinaloa
329	Hermosillo	Sonora
333	Navojoa	Sonora
334	Heroica Nogales	Sonora
336	San Luis Río Colorado	Sonora
348	Victoria	Tamaulipas

Fuente: elaboración propia con base en el SUN 2010.

Imagen 17

Distribución territorial de las 100 primeras ciudades del SUN 2010



Fuente: elaboración propia con base en el SUN 2010.

Anexo B

Estadísticos de AGEB por año. Resumen estadístico de AGEB 1995 (superficie en ha)

Estadístico	Originales	Ajustados
Media	54.43	53.29
Error típico	0.53	0.44
Mediana	37.05	36.83
Moda	33.61	32.20
Desviación estándar	76.30	62.56
Varianza de la muestra	5 821.97	3 913.83
<i>Curtosis</i>	2 261.94	526.19
Coefficiente de asimetría	31.64	12.48
Rango	6 244.36	3 582.39
Mínimo	0.10	0.05
Máximo	6 244.46	3 582.44
Suma	1 127 709.28	1 094 172.92
Cuenta	20 720	20 534

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 1995.

Resumen estadístico de AGEB 2000 (superficie en ha)

Estadístico	Originales	Ajustados
Media	51.06	49.84
Error típico	0.49	0.40
Mediana	34.55	34.21
Moda	28.71	15.02
Desviación estándar	74.60	60.85
Varianza de la muestra	5 565.38	3 702.96
<i>Curtosis</i>	2 388.16	523.79
Coefficiente de asimetría	31.93	12.07
Rango	6 382.18	3 592.45
Mínimo	0.06	0.05
Máximo	6 382.24	3 592.50
Suma	1 191 858.52	1 159 554.19
Cuenta	23 341	23 266

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2000.

Resumen estadístico de AGEB 2005 (superficie en ha)

Estadístico	Originales	Ajustados
Media	45.39	45.21
Error típico	0.34	0.34
Mediana	31.37	31.38
Moda	16.76	43.52
Desviación estándar	57.36	56.43
Varianza de la muestra	3 290.14	3 183.85
<i>Curtosis</i>	569.45	593.81
Coefficiente de asimetría	12.56	12.69
Rango	3 608.51	3 583.51
Mínimo	0.03	0.03
Máximo	3 608.54	3 583.54
Suma	1 282 566.12	1 246 995.15
Cuenta	28 259	27 583

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2005.

Resumen estadístico de AGEB 2010 (superficie en ha)

Estadístico	Originales	Ajustados
Media	42.44	42.27
Error típico	0.31	0.31
Mediana	29.15	29.11
Moda	39.12	39.12
Desviación estándar	55.85	55.17
Media	42.44	42.27
Error típico	0.31	0.31
Mediana	29.15	29.11
Moda	39.12	39.12
Desviación estándar	55.85	55.17
Varianza de la muestra	3 118.67	3 043.19
<i>Curtosis</i>	587.62	607.22
Coefficiente de asimetría	12.68	12.69
Rango	3 641.25	3 641.25
Mínimo	0.03	0.03
Máximo	3 641.28	3 641.28
Suma	1 353 324.04	1 345 462.08
Cuenta	31 888	31 833

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 1995.

Resumen estadístico de AGEB 2015 (superficie en ha)

Estadístico	Originales	Ajustados
Media	40.02	40.06
Error típico	0.29	0.29
Mediana	27.29	27.27
Moda	16.38	16.38
Desviación estándar	54.15	54.26
Varianza de la muestra	2 932.40	2 944.40
<i>Curtosis</i>	610.17	602.64
Coefficiente de asimetría	12.67	12.58
Rango	3 641.25	3 641.25
Mínimo	0.03	0.03
Máximo	3 641.28	3 641.28
Suma	1 373 919.32	1 381 555.64
Cuenta	34 327	34 489

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2015.

Resumen estadístico de AGEB 1995 (perímetro en km)

Estadístico	Originales	Ajustados
Media	3.398	3.399
Error típico	0.013	0.013
Mediana	2.968	2.972
Desviación estándar	1.874	1.837
Varianza de la muestra	3.513	3.374
<i>Curtosis</i>	26.783	10.645
Coefficiente de asimetría	2.698	2.139
Rango	49.497	32.631
Mínimo	0.129	0.096
Máximo	49.625	32.728
Suma	70 402.142	69 789.444
Cuenta	20 720	20 534

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 1995.

Resumen estadístico de AGEB 2000 (perímetro en km)

Estadístico	Originales	Ajustados
Media	3.288	3.270
Error típico	0.013	0.012
Mediana	2.880	2.869
Desviación estándar	1.928	1.874
Varianza de la muestra	3.717	3.512
<i>Curtosis</i>	28.394	8.898
Coefficiente de asimetría	2.615	1.943
Rango	54.058	32.110
Mínimo	0.100	0.097
Máximo	54.158	32.207
Suma	76 745.773	76 081.241
Cuenta	23 341	23 266

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2000.

Resumen estadístico de AGEB 2005 (perímetro en km)

Estadístico	Originales	Ajustados
Media	3.093	3.100
Error típico	0.011	0.011
Mediana	2.744	2.752
Desviación estándar	1.844	1.833
Varianza de la muestra	3.401	3.361
<i>Curtosis</i>	8.984	9.176
Coefficiente de asimetría	1.901	1.916
Rango	33.832	33.825
Mínimo	0.073	0.073
Máximo	33.906	33.898
Suma	87 401.785	85 514.365
Cuenta	28 259	27 583

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2005.

Resumen estadístico de AGEB 2010 (perímetro en km)

Estadístico	Originales	Ajustados
Media	2.982	2.976
Error típico	0.010	0.010
Mediana	2.659	2.656
Desviación estándar	1.867	1.860
Varianza de la muestra	3.487	3.461
<i>Curtosis</i>	8.148	8.078
Coefficiente de asimetría	1.872	1.859
Rango	31.506	31.506
Mínimo	0.076	0.076
Máximo	31.581	31.581
Suma	95 084.781	94 744.280
Cuenta	31 888	31 833

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2010.

Resumen estadístico de AGEB 2015 (perímetro en km)

Estadístico	Originales	Ajustados
Media	2.859	2.867
Error típico	0.010	0.010
Mediana	2.567	2.569
Desviación estándar	1.856	1.872
Varianza de la muestra	3.446	3.504
<i>Curtosis</i>	7.551	7.857
Coefficiente de asimetría	1.795	1.839
Rango	31.510	31.510
Mínimo	0.071	0.071
Máximo	31.581	31.581
Suma	98 154.132	98 866.457
Cuenta	34 327	34 489

Fuente: elaboración propia con base en: INEGI. *Marco Geoestadístico Nacional*. 2015.

Moda, media y mediana de la altura del relieve mexicano

Mode, Mean and the Median of the Height of the Mexican Relief

Rodrigo Tovar Cabañas,* José Alfredo Jáuregui Díaz** y Shany Arely Vázquez Espinosa***

En este trabajo se determinaron las medidas de tendencia central de las celdas del *Continuo de Elevaciones Mexicano 3.0* desarrollado por el Instituto Nacional de Estadística y Geografía, cuya característica principal es la resolución de 15 metros. El área estudiada comprende todo el territorio nacional: 1) se trabajó sobre un universo superior a los 9 mil millones de celdas; 2) los 225 m² que cubre cada celda es la mayor resolución que se ha logrado obtener para la escala nacional; 3) de estas, 58 % cae dentro de un promedio altitudinal de los 1 040 metros sobre el nivel del mar (m. s. n. m.), mientras que el rango 0-100 m. s. n. m. corresponde a la moda, pues dicho intervalo se repite en 18 % de las celdas, en tanto que la mediana ocupa el valor de 1 045; 4) se determinó con un criterio análogo que la pequeña ciudad de Zumpango del Río, Guerrero, se sitúa sobre el promedio altitudinal, en tanto que, territorialmente hablando, la demarcación municipal de Axochiapan, Puebla,

In this work, the central tendency measurements of the cells of the *Mexican Elevation Continuum 3.0* developed by the National Institute of Statistics and Geography were determined, whose main characteristic is the 15-meter resolution. The studied area includes all the national territory: 1) we worked on a universe of more than 9 billion cells; 2) the 225 m² that covers each cell is the highest resolution that has been obtained for the national scale; 3) of these, 58 % fall within an average altitude of 1 040 meters above sea level (m.a.s.l.), while the range 0-100 m.a.s.l. corresponds to the mode, since this interval is repeated in 18 % of the cells, while the median occupies the value of 1 045 m.a.s.l.; 4) it was determined with an analogical criterion that the small city of Zumpango del Río, Guerrero, is located above the altitudinal average, while, territorially speaking, the municipal demarcation of Axochiapan, Puebla, presents the altitudinal average of the nation; and 5) re-

* Universidad de Xalapa, AC, rod_geo77@hotmail.com

** Universidad Autónoma de Nuevo León (UANL), jose.jaureguid@uanl.mx

*** Universidad Veracruzana, shanyvaz@gmail.com

presenta el promedio altitudinal de la nación; y 5) respecto a la moda, las localidades de Puerto Vallarta, Jalisco, y Acayucan, Veracruz de Ignacio de la Llave, son una buena representación de dicha tendencia e Ixtlán del Río, Nayarit, con sus 1 044 m. s. n. m., representaría la mediana.

Palabras clave: relieve mexicano; promedio altitudinal; *Continuo de Elevaciones Mexicano*.

Recibido: 31 de diciembre de 2019.
Aceptado: 6 de marzo de 2020.

garding mode, the towns of Puerto Vallarta, Jalisco, and Acayucan, Veracruz de Ignacio de la Llave, are a good representation of this trend and Ixtlan del Rio, Nayarit, with its 1 044 m.a.s.l., would represent the median.

Key words: Mexican relief; average altitude; Mexican Elevation Continuum.



Render 3D mapa topográfico frontera Neutral/FrankRamsrott/Getty Images

Planteamiento

El primer mapa hipsométrico de la República Mexicana se elaboró a partir de levantamientos topográficos, mediciones barométricas, curvas de nivel y escala altitudinal en color data de 1911 (ver figura 1A), y una segunda edición en escala 1:15 000 000 apareció en 1921; en ambas versiones se muestran las alturas y el relieve mediante *achures* coloreados, en tanto que la profundidad se representa a partir de isolíneas previamente sondeadas (Institut Kartographia Winterthur, 1921).

Ese mapa y otros hipsométricos editados en Europa en la década de los 20 (Bartholomew, 1922; Hammond, 1928) en escala 1:5 000 000 no estuvieron disponibles para el público americano; sin embargo, la mayor peculiaridad es que su confección quedó influenciada por la visión alpina de sus autores o cartógrafos, en su mayoría de origen septentrional y montañoso. Ahora bien, el lugar de enunciación (coordenadas geográficas en las cuales una altitud de 700 m se considera región alpina), desde donde se confeccionaron de forma primigenia los pisos altitudinales del territorio mexicano dio como resultado lo que se ha dado en denominar mapas hipsométricos de tradición alpina europea (ver figura 1B).

La primera serie hipsométrica de la República Mexicana elaborada a partir del reconocimiento e interpretación de fotografías aéreas, fotogramas, proyectores dobles, métodos anaglíficos y en escala militar (1:500 000) data de la década de los 30 (ver figura 1C). Dicha edición constó de 17/27 cartas producidas por el ejército de los Estados Unidos de América (EE. UU.); en estas versiones se muestran en color los pisos altitudinales cada 500 metros (m), con excepción de las llanuras costeras delimitadas por la cota de los 200 metros sobre el nivel del mar (m. s. n. m.); en tales mapas no se representa la profundidad (USAMS, 1936).

En 1942 apareció el norte de México representado en escala 1:5 385 600 con seis pisos altitudinales: de 0 a 152.5 m. s. n. m.; 152.5 a 305.0; 305-610; 610-1 525; 1 525-3 050; y 3 050 y más altura (Ohman,

1942). La base topográfica de dicho trabajo fue elaborada por el ejército de los EE. UU. mediante técnicas de aerotriangulación o triangulación topográfica y de grabación estereoscópica en escala 1:250 000; en estas últimas versiones se muestra la batimetría a 20, 50, 100 y 200 m de profundidad (USAMS, 1941).

Ahora bien, la técnica aereostereográfica permitió a los cartógrafos militares estadounidenses constatar la preponderancia del relieve llano sobre el accidentado que presenta el globo terrestre, por lo que, en esta confección hipsométrica (a la que se le ha denominado tradición hipsométrica estadounidense), los pisos altitudinales del territorio mexicano delimitan la llanura costera hasta la cota de los 300 m. s. n. m.; de 300 a 600, una zona de transición (piedemonte); las praderas (altiplano mexicano), entre los 600 y 1 500, aproximadamente; y el terreno montañoso (las Rocallosas y la Sierra Madre Occidental) lo asocian a alturas por arriba de los 1 500 (ver figura 1D).

La tradición cartográfica estadounidense, por ser secreto de Estado, no estuvo disponible al público mexicano, por lo que la visión hipsométrica europea fue la que prevaleció en los mapas altimétricos subsecuentes, como el elaborado por el Instituto Geographico di Agostini (IGA, 1952) o el publicado por Bayer (1953), donde el relieve mexicano, prácticamente, se representa como una enorme meseta con amplias regiones montañosas (ver figuras 1E y 1F).

Este hecho se debe a que, como las primeras representaciones hipsométricas del relieve mexicano fueron elaboradas por cartógrafos europeos que no recorrieron el territorio y que usaron la regla del gradiente térmico atmosférico (GTA) septentrional (fijado por esos años en 1° C por cada 100 m de altitud) como criterio para delimitar la base del relieve montañoso, supusieron que a los 1 000 m. s. n. m. en el México tropical debían existir bosques de montaña similares a los de las latitudes septentrionales. Al respecto, en la actualidad se sabe que el GTA meridional o tropical corresponde a 1° C por cada 180 m de

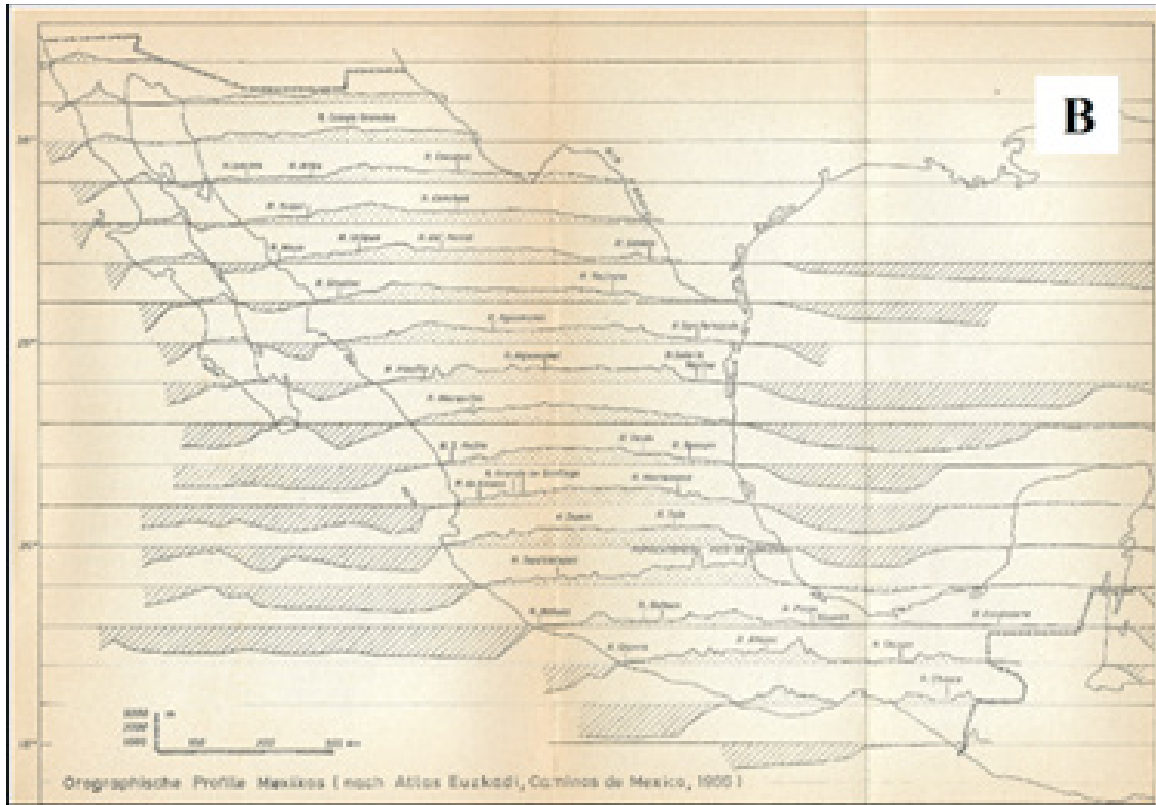
Primeros mapas hipsométricos de la República Mexicana, 1911-1953



Fuentes: Institut Kartographia Winterthur, 1921; Bartholomew, 1922; USAMS, 1936; Ohman, 1942; IGA, 1952; y Bayer, 1953.

Figuras 2

A) carta hipsométrica de la República Mexicana editada por la escuela de cartografía de Polonia en la década de los 60 y B) perfiles hipsométricos del relieve mexicano por grado de latitud



Fuentes: PATS, 1967; Alcorta, 1974.

altitud (Hallé, 1999), por lo que el piso montañoso que esos cartógrafos europeos representaron a los 1 000 m. s. n. m. hoy se conoce que en realidad se encuentra a los 1 800.

Al respecto, no todas las escuelas cartográficas cometieron o asimilaron esos riptos involuntarios, que Said (2004: 49) denomina imaginarios geográficos; por ejemplo, se debe destacar que la representación del relieve mexicano no se le escapó a la prestigiosa escuela de cartografía de Polonia, quien, en la década de los 60, enmendó en parte las exageraciones montañosas de la corriente hipsométrica de tradición alpina europea.

En efecto, el mapa hipsométrico de la República Mexicana, confeccionado a partir de fotogrametría aérea con proyección mecánica (1:500 000) elaborado por el Servicio de Topografía del Ejército Polaco, muestra en color los pisos altitudinales cada 500 m, con excepción de las llanuras costeras delimitadas por la cota de los 100 y 200 m. s. n. m., así como las cimas de 4 000 a 5 000 m. s. n. m., mientras que, en la representación batimétrica, los pisos abisales son cada 1 000 m (PATS, 1967).

Como se dijo con anterioridad, en dicha obra se aprecian los 500 km de anchura que tiene la Planicie del Noreste de México hacia su porción noroccidental, o los 110 km de ancho máximo que tiene la Planicie del Noroeste en su parte central, así como la ubicación puntual y radios pequeños de los siete picos más elevados del país; con ello, se brindó una representación cartográfica más suave con la intención de mostrar la relativa baja altitud de la topografía mexicana. Lamentablemente, este excelente mapa hipsométrico, al igual que otro de hechura soviética en escala 1:2 500 000 (USSR, 1967), así como uno alemán de perfiles hipsométricos del relieve mexicano por grado de latitud (Alcorta, 1974), solo estuvieron a disposición de los geógrafos y cartógrafos de nacionalidad cubana (ver figuras 2).

En México, el primer mapa hipsométrico del país en color estuvo disponible a partir de la década de los 60. Tal edición 1:5 000 000 estuvo a cargo del

ingeniero Tamayo (1962), quien usó los datos de 1958 de la Comisión Intersecretarial y Coordinadora del Levantamiento de la Carta Geográfica de la República Mexicana (CICLCGRM). Una segunda, corregida y aumentada en escala 1:8 000 000 apareció 10 años después (García y Falcón 1972). Es de importancia destacar que la versión de Tamayo se alinea a la corriente alpina europea, mientras que la de García y Falcón se apega a los lineamientos de la escuela polaca de cartografía, es decir, el mapa de los 60, pese a que resalta la anchura de la Planicie del Noreste, termina dando la impresión de que México es una nación alpina (ver figura 3A), en tanto que la carta altimétrica de los 70 muestra un relieve más llano, pues resalta las planicies costeras pero, además, es el primer mapa hipsométrico que le brinda un papel central a la altiplanicie mexicana (ver figura 3B).

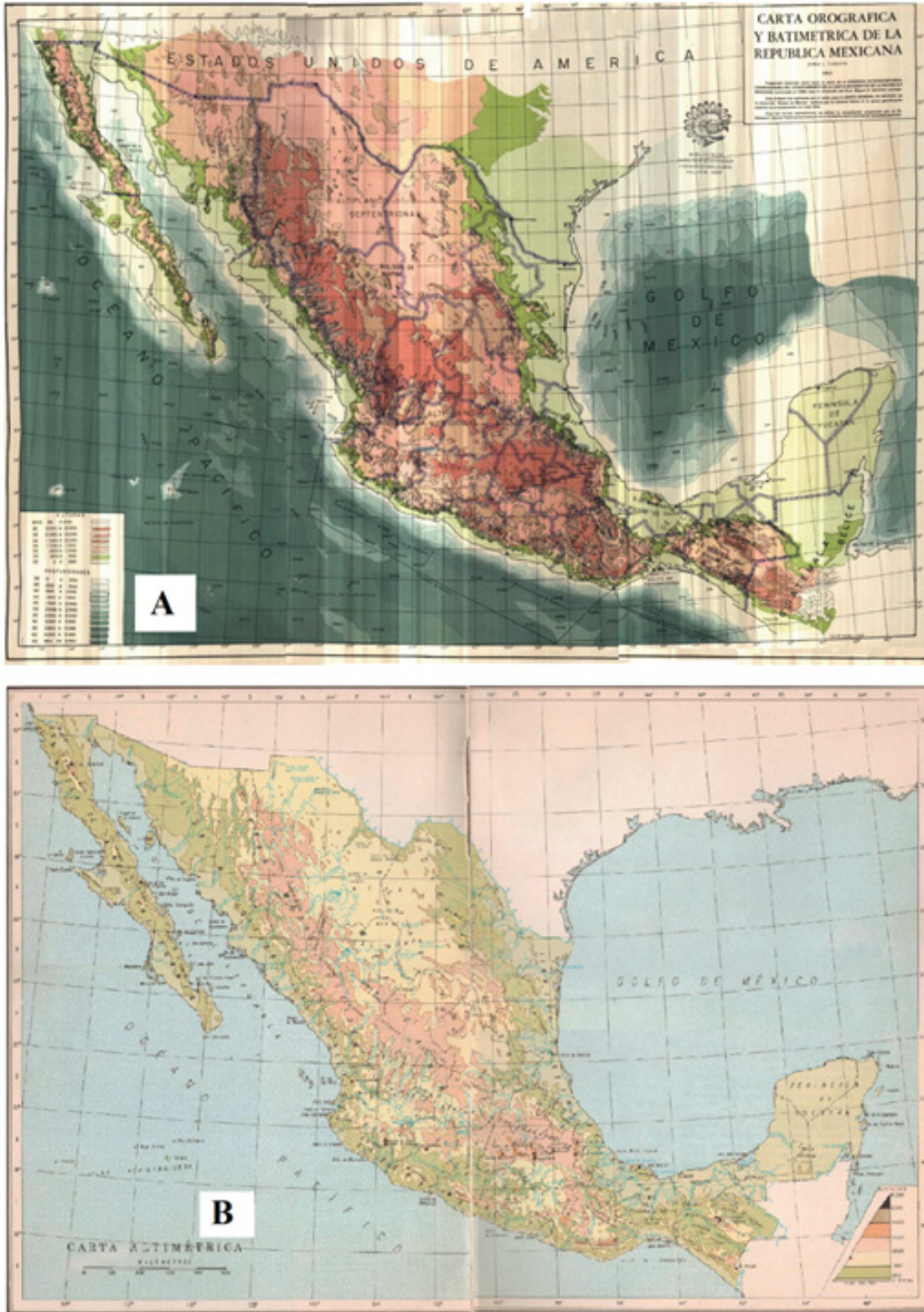
En 1958, México inició la confección moderna de la cartografía nacional en escala 1:50 000. Cabe recordar que, después de la CICLCGRM, en octubre de 1968 se creó la Comisión de Estudios del Territorio Nacional (CETENAL, 1974), que en julio de 1978 cambió de nombre a Departamento de Estudios del Territorio Nacional (DETENAL) para, finalmente, el 25 de enero de 1983, convertirse en el Instituto Nacional de Estadística, Geografía e Informática (INEGI).¹ Al caso, luego de un cuarto de siglo, se logró la elaboración de las 2 500 cartas topográficas en esa escala que se necesitaban para representar a todo el territorio nacional. Con esa cuantiosa información, los geógrafos del Instituto de Geografía de la Universidad Nacional Autónoma de México realizaron el mapa hipsométrico escala 1:5 000 000, entre otros estudios geomorfológicos.

Bajo ese marco, Lugo (1990), usando fotografías aéreas 1:75 000 y la cartografía profesional del INEGI en 1:50 000, produjo mapas del relieve mexicano con unidades mínimas de análisis espacial de hasta 500 m de resolución. Con ello, pudo elaborar la geomorfología en escala nacional de las

¹ Con el decreto de la *Ley del Sistema Nacional de Información Estadística y Geográfica*, el 14 de abril de 2008, adquirió su autonomía y cambió su nombre a Instituto Nacional de Estadística y Geografía, pero conservó las mismas siglas.

Figuras 3

Cartas hipsométricas de la República Mexicana de 1962 y 1972



Fuentes: Tamayo, 1962; García y Falcón, 1972.

provincias fisiográficas. Algunos de sus resultados señalan la altitud media de las mismas, por ejemplo, para el altiplano mexicano señaló 1 000 m. s. n. m.; al septentrional lo dividió en dos pisos: 1 400 m. s. n. m. al norte y 1 200 al sur; dio una altura media de 2 100-2 200 para la Sierra Madre Occidental; al flanco oriental de la Sierra Madre Oriental lo definió en 2 000; y a la Planicie del Noreste de México la dividió en dos pisos: 1 000 en el piedemonte y 200 m. s. n. m. en la costa.

Asimismo, sobre la Mesa Central aclaró su descenso altitudinal de sur a norte de 2 000 a 1 200, además de estimar el promedio altitudinal de la Sierra de Guanajuato en 1 100 m. s. n. m.; a la Planicie Costera del Golfo de México también la dividió en dos niveles: 400 m. s. n. m. en el piedemonte y 200 en el límite costero; al *parteaguas* de la Sierra Madre del Sur lo promedió en 2 000; la altitud mínima de la Depresión del Balsas la fijó en 400; para la Península de Yucatán fijó dos criterios: para la planicie, 25-30 m. s. n. m. y para la porción meridional, 200-400.

También, determinó ciertos rangos regionales; por ejemplo, para la Sierra de San Felipe, en Baja California, dio un valor de 1 000-1 500 m. s. n. m. y a la Sierra Pinacate (Llanura Sonorense) le calculó 1 190; cabe mencionar que detalló con ahínco al Sistema Neovolcánico Transmexicano de la siguiente forma: Fosa Tepic-Chapala, 900 m. s. n. m.; Lago de Chapala, 1 600; Fosa de Colima, 1 200; Michoacán de Ocampo (Pátzcuaro y Cuitzeo), 1 800-2 000; y Cuenca de Toluca, 2 600; asimismo, para la Sierra de Soconusco, estableció un rango que va de los 2 200 a 2 800 m. s. n. m.; para la Depresión de Chiapas, 2 000; y para la Planicie del Usumacinta, 100-200. Sin embargo, pese a todo este loable esfuerzo, se vio cauto al no determinar la altitud media del conjunto nacional.

Esta última cautela podría decirse que es una excepción, pues el riesgo subjetivo, comúnmente asumido en las interpretaciones geográficas, que los cartógrafos europeos hicieron del resto de las regiones del planeta (Santos, 1996), la mayoría de las veces ha culminado posicionándose

como una verdad objetiva; en términos de Said (2004), esos imaginarios terminan confeccionando realidades, de allí que la idea de que México es un país montañoso haya sido asumida en la década de los 80 por varios geógrafos mexicanos, por ejemplo, Coll (1985: 44) pronunció que "... en la mayor parte del territorio nacional, las condiciones naturales no son las más indicadas para el desarrollo de los cultivos. La abundancia de terrenos montañosos da lugar a suelos de pendiente mucho mayor que la adecuada agrológicamente...". Solo un experto en geomorfología mexicana de la talla de Lugo (1990) pudo advertir sobre la anomalía que presenta la cartografía hipsométrica de tradición alpina europea al señalar que "... a pesar de la poca altura de la Sierra Madre Oriental, el paisaje da una impresión de montañas altas...".

Para destrabar dicho conflicto de carácter preceptivo, hace una década se trazaron mapas del relieve mexicano usando datos de elevación digitales (o datos matriciales) con resolución de 50 m, sin embargo, como esta corresponde a 0.04 % de una ideal de 1 m², Pérez y Mas (2009) se volcaron a evaluar los errores inherentes a las interpolaciones elaboradas con distintos *softwares* o modelos digitales de elevación (MDE). Es útil señalar que la interpolación de los datos de elevación digitales mediante MDE son, actualmente, la forma más efectiva de representar la superficie.

De ese modo, la vieja interpretación del relieve mexicano, con el advenimiento de la cartografía de alta resolución, poco a poco se ha diluido, sin embargo, el mito del México montañoso aún persiste, sobre todo entre las compañías mineras, que asumen que un relieve así es sinónimo de banco de materiales o suelos aptos para la vocación minera; por ello, nos hemos dado a la tarea de determinar las medidas de tendencia central de las celdas del *Continuo de Elevaciones Mexicano (CEM) 3.0* desarrollado por el INEGI, cuya característica principal es la resolución de 15 m; con esto, se pretende cambiar la interpretación por la tendencia principal, esta última más real dada la cantidad de datos procesados.

Área de estudio

Es toda la superficie continental de la República Mexicana, es decir, 1 960 189 km² (INEGI, 2017), que podría representarse idealmente (como para trazar mapas de riesgos en escala urbano-vecinal, por ejemplo) con 2 billones de celdas de superficie de 1 m²; sin embargo, la capacidad tecnológica de los procesadores convencionales (estaciones de trabajo) aún no puede procesar semejante cantidad de información (quizá en una década, los cuánticos lo logren). Lo que actualmente existe son los datos de elevación digitales (interpolación nativa) con una resolución de pixel de 15 m proporcionados por el INEGI (2013); esta equivale a 0.44 % de la ideal a 1 m², empero ese 0.44 % es 1 100 % más preciso que la resolución de 50 m de hace una década. Esto significa que la distancia lineal mínima entre dos celdas de 15 x 15 m es de 21 m (conjunto de celdas a las que se les va a extraer sus medidas de tendencia central), mientras que la distancia mínima entre dos de 50 x 50 m es de 70 m; pues bien, con el valor altitudinal de 9 mil millones de celdas de 15 x 15 m por lado se intenta definir las especificidades hipsométricas del relieve mexicano, cuya metodología se explica a continuación.

Metodología

Antes de pasar a la descripción bayesiana del método estadístico empleado, conviene hacer una precisión sobre los MDE y demás datos procesados. Hoy en día existen diferentes fuentes de error en los MDE; el más común concierne a errores en los insumos utilizados para generar dichos MDE, por ejemplo, los posicionales (derivados de una deficiente localización geográfica de la cota o de la trayectoria de la curva de nivel) de los centroides acotados o de la trayectoria de las curvas de nivel; otro ocurre con los valores de altitud erróneos (caso de lecturas imprecisas por diversos factores).

Existen también los relacionados con el proceso de interpolación, es decir, los bilineales, bicúbicos, de krigado y de vecino más cercano, los cuales frecuentemente se obvian; sin embargo, Pérez y Mas

(2009) han descubierto que las resoluciones entre 20 y 30 m son las celdas que menos errores de interpolación suelen tener, es decir, el análisis de los fallos en los MDE a 30 x 30 m muestra un rango de error en altitud de entre 7.3 a 11.3 m, donde la mayoría de estos errores se ubican principalmente en las zonas con pendiente muy baja (planicies y planicies aluviales) y en las cimas. Además, Uribe, Cruz y Cruz (2018) señalan que, en zonas de planicie, tales errores llegan a estar, decenas de veces, por encima del valor real de la elevación, por lo que, en otros análisis en escalas menores, se debe tener cautela, pues las estimaciones del CEM podrían verse afectadas por estos.

No obstante, en México, la información más susceptible de ser utilizada para generar MDE de todo el territorio nacional es la información topográfica vectorial del INEGI (2013) en escala 1:50 000 interpolada en celdas con resolución de 15 m, donde en escala nacional tales valores altitudinales derivados de sensores satelitales y algoritmos de interpolación aún tienen la mayor cobertura que los *lidar* en una relación de 1:1.75.

Así, en esta investigación se procesaron 9 mil millones de celdas con una resolución de 225 m² proporcionadas por el INEGI (2013). Esto significa que la distancia lineal mínima entre dos celdas de 15 x 15 m es de 21 m y su error máximo de interpolación, ya sea estimado por *software* de patente o abierto, en posición XY es de 1/3 del tamaño lineal del pixel y en la coordenada Z, de 2/3 (5.1 m lineales y 11 m de altitud), lo cual, para la escala de análisis geográfico nacional, 1:1 000 000, es bastante aceptable.

Las medidas de tendencia central se refieren a una técnica estadística empleada para resumir en un solo valor a un conjunto de valores; los resultados pretenden evaluar en qué medida los datos del conjunto analizado son semejantes entre sí, por ejemplo, el valor medio de la muestra empleada en este trabajo (9 mil millones de datos) se asume como numéricamente muy cercano a la realidad matemática de la variable medida (2 billones de datos); este último valor solo es calculable si se conocieran con toda exactitud todos los casos, lo

cual excede el alcance de esta investigación. Su expresión matemática es:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n xi = \frac{x1 + x2 + \dots + xn}{n}$$

donde:

\bar{x} = media o promedio.

n = suma de todos los datos recolectados.

$\sum_{i=1}^n$ = sumatoria de las frecuencias (número de casos recolectados).

$X1$ = cualquiera de los valores de n .

de modo que su ordenamiento en esta investigación se ejemplifica:

$$n = 9.0405E+09.$$

$$\sum_{i=1}^n = 1 = 9.4166E+12.$$

$$\bar{x} = 1\ 041.60 \text{ m. s. n. m.}$$

Dado que en esta investigación n es par, la mediana es la media aritmética de los dos valores centrales, por lo que su fórmula geoestadística

empleada para calcular la mediana de la altura del relieve mexicano fue la siguiente:

$$Me = L1 + \left[\frac{\frac{n}{2} - (\sum f)1}{f_{mediana}} \right] c$$

donde:

Me = mediana aritmética.

n = número de datos (frecuencia total).

$L1$ = frontera inferior de la clase de la mediana.

$(\sum f)1$ = suma de frecuencias de las clases inferiores a la de la mediana.

$f_{mediana}$ = frecuencia de la clase de la mediana.

c = anchura del intervalo de clase de la mediana.

En este caso, la mediana del conjunto de números par ordenados en magnitud es: 1 045 metros sobre el nivel del mar.

Ahora bien, recordemos que la moda de un conjunto de números es el valor que ocurre con mayor

Cuadro 1

Intervalos modales por orden de prelación

Rango m. s. n. m.	Núm. de celdas	%	Rango m. s. n. m.	Núm. de celdas	%
-32-100	1 622 960 419	18.0	1 601-1 700	276 420 662	3.1
101-200	584 863 147	6.5	501-600	269 140 423	3.0
201-300	440 273 597	4.9	1 001-1 100	266 890 294	3.0
1 301-1 400	370 940 923	4.1	601-700	244 527 286	2.7
1 201-1 300	366 108 247	4.0	2 201-2 300	228 919 587	2.5
1 401-1 500	338 992 935	3.7	701-800	223 694 245	2.5
1 101-1 200	332 185 203	3.7	801-900	217 969 666	2.4
301-400	321 685 671	3.6	901-1 000	206 058 596	2.3
2 001-2 100	321 581 032	3.6	2 301-2 400	178 908 900	2.0
1 901-2 000	320 923 193	3.5	2 401-2 500	139 431 025	1.5
1 501-1 600	318 402 420	3.5	2 501-2 600	107 952 047	1.2
1 801-1 900	302 655 330	3.3	2 601-2 700	70 281 013	0.8
1 701-800	293 538 215	3.2	2 701-2 800	41 985 386	0.5
401-500	289 680 669	3.2	2 801-2 900	24 071 784	0.3
2 101-2 200	283 111 475	3.1	2 901-3 000	13 230 801	0.1

Nótese que 8/10 rangos están por debajo de la cota de los 1 500 metros sobre el nivel del mar.

Fuente: elaboración propia con base en INEGI, 2013.

frecuencia, el cual se puede expresar estadísticamente de la siguiente forma:

$$M = L1 + \left(\frac{D1}{D1 + D2} \right) c$$

donde:

M = moda.

$L1$ = frontera inferior de la clase modal (clase que contiene a la moda).

$D1$ = exceso de la frecuencia modal sobre la de la clase inferior inmediata.

$D2$ = exceso de la frecuencia modal sobre la clase superior inmediata.

c = anchura del intervalo de clase modal.

En cuanto a la moda, se tuvieron que trazar varios intervalos modales, los cuales se exponen en porcentaje en el cuadro 1.

Resultados y discusión

Los mapas hipsométricos de tradición alpina europea que marcan el relieve de montaña a los 1 000 m. s. n. m. (similar al de las latitudes septentrionales) fue la cartografía aceptada, más por subjetividad que por objetividad científica, en México. La cartografía derivada de la tradición hipsométrica estadounidense, que delimitó dos grandes regiones planas (la llanura costera definida hasta la cota de los 300 m. s. n. m. y el altiplano mexicano acotado entre los 600 y 1 500) no estuvo a disposición de los geógrafos mexicanos del siglo XX. Por esa circunstancia, la idiosincrasia de que México es un país montañoso fue asumida en la década de los 80 por varios profesionistas abocados al estudio del territorio mexicano, con excepción de los geomorfólogos, quienes notaron que, a pesar de la poca altura de las sierras mexicanas, el paisaje daba una impresión de montañas altas.

Con el advenimiento de los MDE, a finales del siglo XX se logró conocer la altitud verdadera de 0.04 % de la topografía nacional; en la actualidad, se ha logrado conocer la especificidad del relieve mexi-

cano hasta 0.4 %; es probable que en 10 años se pudiese computar el valor altitudinal de 4.0 % del territorio mediante celdas con resolución de 5 x 5 m; sin embargo, la cantidad que representa ese 0.4 % actual se refiere a más de 9 mil millones de píxeles, a los cuales se les determinó sus medidas de tendencia central.

El promedio que arrojó la sumatoria de esa muestra fue de 1 041 m. s. n. m. y la mediana correspondió a un valor de 1 045. Ahora bien, pese a que dichos valores nos indican que si el territorio mexicano tuviere una tendencia, esta se inclinaría hacia los 1 000 m. s. n. m.; sin embargo, dado que, para este caso, tanto la media como la mediana son datos abstractos y continuos, para comprobar dicha afirmación se tendrían que homologar los 9 mil millones de columnas de 225 m² de área, situación que no ocurre con la moda, ya que sus valores tienen la propiedad de poderse agrupar como datos discretos y concretos, por lo tanto, su interpretación es más asimilable, pues de los 9 mil millones de columnas de 225 m² de área, desde la tradición hipsométrica estadounidense, 67.4 % tiene menos de 1 500 metros de altitud.

Ahora bien, considerando las divisiones florísticas de México y las cuatro provincias serranas (sierras Madre Oriental y Occidental, así como serranías meridionales y transísmicas) situadas por arriba de los 1 800 m. s. n. m. (Rzedowski, 1978), la superficie propiamente montañosa correspondería solo a 22.5 % del total de las celdas computadas. Este hecho abre un debate en torno al uso de suelo con vocación forestal pues, actualmente, este se conceptualiza a los 1 500 m. s. n. m., cuando en realidad debe situarse a los 1 800. En otro orden de ideas, al modificar la concepción hipsométrica del país mediante otras propuestas de pisos altitudinales, por ejemplo, asumir a la planicie costera hasta los 300 m. s. n. m., también se modifican los imaginarios en torno a las actividades productivas y los nichos ecológicos.

Para asimilar los resultados, se hicieron las siguientes analogías: la pequeña ciudad de Zumpango del Río, Guerrero, que se sitúa sobre

los 1 041 (promedio altitudinal) e Ixtlán del Río, Nayarit, con sus 1 044 m. s. n. m., representarían la mediana. Respecto a la moda, las localidades de Puerto Vallarta, Jalisco, y Acayucan, Veracruz de Ignacio de la Llave, son una buena representación de dicha tendencia.

Por último, de cumplirse las previsiones observadas, en el futuro, los mapas hipsométricos de la República Mexicana cambiarán la base 100 en la que actualmente se suelen configurar los pisos altitudinales por cotas más apegadas a los datos que arrojan los satélites, cota para la planicie costera a los 300 m. s. n. m. y para el altiplano a los 1 500 o 1 800 m. s. n. m., dependiendo de la latitud del lugar.

Fuentes

- Alcorta, Ramón. *Geographie Mexikos*. Köln, Botschaft der Vereinigten Mexikanischen Staaten, 1974.
- Bartholomew, John. *The Times atlas*. London, The Times-Edinburgh Geographical Institute, 1922.
- Bayer, H. *World geo-graphic atlas: A composite of man's environment*. Chicago, Priv. Print. for Container Corporation of America, 1953.
- CETENAL. *Memoria del ciclo de conferencias: la información CETENAL, parte importante del instrumental de la planeación*. Monterrey, NL, Universidad Autónoma de Nuevo León-Centro de Investigaciones Urbanísticas, 1974.
- Coll, H. *¿Es México un país agrícola? Análisis geográfico*. México, Siglo XXI, 1985.
- García, M. y G. Falcón. *Atlas: nuevo atlas Porrúa de la República Mexicana*. México, DF, Editorial Porrúa, 1972.
- Hallé, F. *Un mundo sin invierno: los trópicos, naturaleza y sociedades*. Ciudad de México, Fondo de Cultura Económica, 1999.
- Hammond, C. S. *Our flying ambassadors and their routes*. New York, Hammond, 1928.
- Instituto Geographico di Agostini (IGA). *Grand Atlante Geographic*. Italia, IGA, 1952.
- INEGI. *Continuo de Elevaciones Mexicano 3.0*. Aguascalientes, México, INEGI, 2013.
- _____. *Anuario estadístico y geográfico de los Estados Unidos Mexicanos*. Aguascalientes, México, INEGI, 2017.
- Institut Kartographia Winterthur (IKW). *Atlas Scolaire Suisse pour l'Enseignement Secondaire*. Publ. par la Conférence des Chefs des Départements Cant. 2.ª edición. Génova, IKW Winterthur-Payot, 1921.
- Lugo, J. "El relieve de la República Mexicana", en: *Revista Mexicana de Ciencias Geológicas*. 9(1), 1990, pp. 82-111.
- Ohman. *Ohman's new relief map of the United States*. Memphis, Tenn, Ohman Co., 1942.
- PATS. *Pergamon World Atlas*. Warsaw, Poland, Polish Army Topography Service, 1967.
- Pérez, A. y J. Mas. "Evaluación de los errores de modelos digitales de elevación obtenidos por cuatro métodos de interpolación", en: *Investigaciones Geográficas*. (69), 2009, pp. 53-67.
- Rzedowski, J. *Vegetación de México*. México, Limusa, Noriega Editores, 1978.
- Said, E. W. *Orientalism*. New York, Vintage Books, 2004.
- Santos, M. *De la totalidad al lugar*. Vilassar de Mar, Oikos-Tau, 1996.
- Tamayo, J. *Atlas geográfico general de México: con cartas físicas, biológicas, demográficas, sociales, económicas y cartogramas*. México, Instituto Mexicano de Investigaciones Económicas, 1962.
- Uribe, E.; J. Cruz; y E. Cruz. "Evaluación del Modelo de Elevación Digital oficial mexicano (Continuo de Elevación Mexicano, CEM 3.0) de INEGI", en: *Tecnología y Ciencias del Agua*. 9(6), 2018, pp. 153-172.
- USSR. *The World Atlas*. Union of Soviet Socialist Republics, Moscú, 1967.
- USAMS. *Tampico Sheet NF 14 N, 1:500 000: AMS Series F 401*. Washington, DC, Army Map Service, 1936.
- _____. *Vera Cruz Sheet E-14 N-III, 1: 250 000: AMS, Series F541*. Washington, DC, Army Map Service, 1941.

Indicadores de actividad económica municipal 1993-2013

Municipality-Level Economic Activity Indicators 1993-2013

Jesús López-Pérez y Francisco Corona*

* INEGI, jesus.lopez@inegi.org.mx y francisco.corona@inegi.org.mx, respectivamente.

Nota: los autores agradecen los comentarios y sugerencias tanto del personal de la Dirección General Adjunta de Cuentas Nacionales del Instituto Nacional de Estadística y Geografía (INEGI) —en particular a Francisco Guillén, Lourdes Mosqueda, Leonel García y Jaime Rodríguez— como de Benito Arciniega, director de la Coordinación de Diseño Conceptual y Estudios Económicos, y en especial a Natalia Volkow por el acceso al Laboratorio de Microdatos del Instituto; asimismo, extienden un reconocimiento por las aportaciones de los participantes en los coloquios de investigación del INEGI, de los dictaminadores anónimos y de Gerardo Leyva Parra, director general adjunto de Investigación; sin lugar a dudas, todos contribuyeron a mejorar este documento.

Vivero de plantas en Coyoacán México/ChepeNicolli/Getty Images



Este trabajo representa una extensión de nuestra investigación del 2019, en la cual construimos estos indicadores (IAEM) basados en los Censos Económicos para el periodo 2003-2013; ahora, extendimos la longitud de las series de tiempo hasta 1993 e incluimos mejoras metodológicas en términos de representatividad económica y geográfica; también, a manera de verificar el funcionamiento en muestras finitas del uso de interpoladores para los años no censales, se estimó un experimento Monte Carlo para diferentes procesos generadores de datos evaluando el desempeño de los interpoladores usados en esta investigación y, asimismo, se verificó la preservación del movimiento entre estimaciones preliminares y finales al hacer uso de la Regla de Combinación (RC) de Guerrero y Peña (2003). Concluimos que los IAEM hasta 1993 cuentan con las características económicas y geográficas representativas que permiten a los usuarios tener una mayor cantidad de información para la toma de decisiones a nivel municipal; asimismo, hay certeza estadística de que el uso de interpoladores lineales minimiza la incertidumbre para la estimación de los años no censales e, igual, se puede comprobar de manera empírica la preservación del movimiento, lo cual garantiza la consistencia de las estimaciones al hacer uso de la RC.

Palabras clave: interpoladores; Monte Carlo; Producto Interno Bruto; Regla de Combinación; Valor Agregado Censal Bruto.

Recibido: 12 de diciembre de 2019.
Aceptado: 20 de marzo de 2020.

Antecedentes

En el 2019 presentamos la primera aproximación formal para la construcción de indicadores¹ de actividad económica municipal (IAEM) del 2003 al 2013 por Gran Actividad Económica (GA) basados en los Censos Económicos (CE). Para lo anterior, usando

¹ Nos referimos al término en plural por la naturaleza econométrica con la que se construyen, es decir, se tienen diferentes indicadores según su fase de estimación, restringiendo finalmente y de manera contemporánea con datos oficiales de índole estatal; no obstante, como resultado final, el lector puede considerar como sinónimos los términos indicadores e indicador.

This work represents an extension of our 2019 research, in which we constructed these indicators (IAEM) based on the Economic Censuses for the period 2003-2013; now, we extended the length of the time series to 1993 and included methodological improvements in terms of economic and geographic representativeness. In order to verify the functioning in finite samples of the use of interpolators for non-census years, we estimated a Monte Carlo experiment for different data generating processes evaluating the performance of the interpolators used in this research and we also verified the preservation of the movement between preliminary and final estimates by using the Combination Rule (CR) of Guerrero and Peña (2003). We concluded that the IAEMs up to 1993 have the representative economic and geographic characteristics that allow users to have a greater amount of information for decision making at the municipal level. Likewise, there is statistical certainty that the use of linear interpolators minimizes the uncertainty for the estimation of non-census years and the preservation of movement can be empirically verified, which guarantees the consistency of the estimates when using CR.

Key words: interpolators; Monte Carlo; Gross Domestic Product; Combination Rule; Gross Census Added Value.

información representativa a nivel regional, estimamos para los años censales aproximaciones del Producto Interno Bruto (PIB) de municipios² para los sectores regionalmente representativos; enseguida, hicimos uso de interpoladores lineales para tener estimaciones del PIB para los años no censales; después, utilizando la Regla de Combinación (RC) de Guerrero y Peña (2003), restringimos los resultados preliminares de tal forma que las sumas

² A lo largo de todo el documento, por simplicidad, cuando nos referimos a los municipios en general abarcamos también a las delegaciones (hoy demarcaciones territoriales) de la Ciudad de México.

de dichas estimaciones fueran equivalentes al PIB estatal; finalmente, agregamos la información de los sectores considerados para formar una aproximación por GA.

La metodología fue evaluada de manera empírica para datos estatales, donde se pudo validar un error muestral. Se concluyó que la principal fuente de error era debido a que el uso de los CE no incluía el total de las actividades económicas.

Tal investigación permite a los usuarios tener una metodología que genera información temporal de la actividad económica por municipio, lo cual puede traducirse en realizar una mayor cantidad de análisis de corto plazo y medir, por ejemplo y para el periodo considerado, el funcionamiento de las políticas económicas de los gobiernos locales; sin embargo, debido a que la longitud de las series es relativamente corta, 10 años, se limita el posible uso estadístico y econométrico de los indicadores, situación que con series de tiempo más largas se minimiza, dado que se pueden construir modelos más robustos que permitan entender, de mejor manera, los fenómenos económicos acontecidos.

Para lograr lo anterior, es claro que hay situaciones a considerar, por ejemplo, los diferentes catálogos utilizados por el Sistema de Cuentas Nacionales de México (SCNM) a través del tiempo, es decir, en los CE 1994 y 1999, la categorización de las actividades económicas se realizaba con la Clasificación Internacional Industrial Uniforme (CIIU), mientras que desde el 2003 se utiliza el Sistema de Clasificación Industrial de América del Norte (SCIAN); esto plantea algunas limitantes, ya que es necesario homologar las actividades entre los distintos catálogos. También, es importante señalar que a partir del trabajo de Guerrero y Corona (2018a, 2018b) existe información oficial del PIB estatal año base 2013 hasta 1980, lo cual permite utilizar la RC para generar IAEM, incluso antes del 2003. Asimismo, se considera apropiado evaluar tanto el desempeño estadístico por interpoladores para los años no censales como los supuestos que garantizan la optimalidad de los resultados en un sentido estadístico.

En la actualidad, el artículo 115 constitucional señala que los estados³ tienen como base de su división territorial y organización política y administrativa el municipio libre. Desde su concepción, esta instancia ha buscado contar con independencia política y económica; así, parte del discurso de la *Constitución de 1917* versa: "... El municipio independiente, que es sin disputa una de las grandes conquistas de la Revolución, como que es la base del gobierno libre, conquista que no sólo dará libertad política a la vida municipal, sino que también le dará independencia económica, supuesto que tendrá fondos y recursos propios para la atención de todas sus necesidades..." (González, 1987).

En consecuencia, es importante que los municipios cuenten con información económica desagregada y con la temporalidad suficiente para la elaboración de sus políticas públicas sustentadas. Además, las entidades federativas requieren de datos a esas escalas para la elaboración de diagnósticos hacia dentro de sus regiones territoriales para poder identificar las disparidades y desigualdades que, por lo común, aparecen en la configuración económica. Es, pues, a través de la observación detallada de la evolución económica de estos territorios que se puede identificar su progreso. Por otro lado, tener claro su tamaño y desempeño diferenciado plantea el hecho de poder diseñar mecanismos que sean capaces de internalizar desde su armado las distintas características de estos en beneficio de los habitantes.

La disponibilidad de estadísticas con este nivel de desagregación permite, también, sensibilizar a los tres niveles de gobierno sobre el rezago de unos municipios con respecto a otros, lo que puede traducirse en el diseño de programas que intenten revertir este comportamiento diferencial. Por ende, la falta de información económica de estos ocasiona limitantes en materia de diseño, análisis y ejecución de políticas públicas.

En México, la disponibilidad de datos a este nivel es, en la mayoría de los casos, para indicado-

³ Incluye la Ciudad de México con sus demarcaciones territoriales.

res demográficos, siendo escasos los de carácter económico. Por mencionar algunos ejemplos, el Instituto Nacional para el Federalismo y Desarrollo Municipal, que mantiene el Sistema Nacional de Información Municipal, contiene cierta información política y sociodemográfica y concentra algunos datos de las condiciones laborales de la población a nivel municipal para los años de los censos de población y vivienda del INEGI. Por otra parte, el Sistema Estatal y Municipal de Bases de Datos del INEGI concentra, además de la información anterior, datos de índole municipal con base en registros administrativos y de otras fuentes, por ejemplo: vehículos de motor en circulación, accidentes de tránsito y estadísticas de finanzas públicas; no obstante, esta última presenta un gran número de datos faltantes, además de que la temporalidad de estas dista de ser larga.

La experiencia de otros países muestra que, por ejemplo, la Oficina de Análisis Económico de los Estados Unidos ha publicado en fecha reciente información del PIB a nivel condado; previamente, en Guci *et al.* (2016), esta institución estableció una agenda de investigación para la medición del PIB a ese nivel de desagregación y cuyos resultados fueron liberados en diciembre del 2019 en la modalidad de prototipo de información.⁴ Otra nación que ya cuenta con datos de este tipo es Ecuador, la cual publica cifras del valor agregado cantonal, pero con la diferencia de que las genera con base en registros contables y no en estimaciones estadísticas.

En México, además de nuestros estudios (2019), existen otros esfuerzos para elaborar indicadores económicos municipales, como los de González-Estrada y Gallegos-Cedillo (2014) y, en fecha más reciente, del Instituto de Información e Investigación Geográfica, Estadística y Catastral del Estado de México (IGECEM, 2019), que los publicó para los 125 municipios de esa entidad utilizando, también, datos de los CE. Sin embargo, la principal limitación de trabajos previos es que en ellos no se hacen consideraciones respecto a la representatividad geográfica de los sectores de actividad eco-

nómica ni a restricciones de índole sectorial para limitar las cifras preliminares a los datos oficiales, situación que sí se considera en el presente trabajo; asimismo, no realiza ninguna validación estadística de la metodología utilizada.

Introducción

Así, el objetivo de esta investigación es calcular IAEM para todos los municipios para el periodo 1993-2013 usando la RC y haciendo uso de datos de los CE con representatividad geográfica, así como de las recomendaciones emitidas por el SCNM del INEGI para tener información sobre las actividades económicas a este nivel de desagregación de manera fidedigna.

Para cumplir con este propósito, nos basamos en nuestro propio trabajo del 2019 revisando de nuevo el concepto de representatividad económica y geográfica en los CE, homologando catálogos de actividades económicas entre el CIIU y el SCIAN y utilizando como restricciones contemporáneas adicionales las estimaciones provenientes de la retroproyección del PIB estatal hasta 1980 realizada por Guerrero y Corona (2018b). También, ya que existen valores atípicos en los CE que pueden limitar el uso de toda la información económica, se recurrió a una corrección de datos basada en la estimación de regresiones usando especificaciones tipo Cobb-Douglas; asimismo, para complementar el análisis, se hizo otro sobre el funcionamiento en muestras finitas del uso de interpoladores para los años no censales estimando un experimento Monte Carlo para diferentes procesos generadores de datos que pueden ocurrir en la práctica cuando se interpolan datos para los años no censales; para finalizar, y muy importante, se verificó de manera empírica la preservación del movimiento entre estimaciones preliminares y finales, lo cual garantiza de forma estadística la consistencia de los resultados al hacer uso de la RC.

Es válido dejar en claro que en este trabajo no se pretende hacer recomendaciones de política económica ni análisis estructurales de los resulta-

⁴ <https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas>

dos obtenidos sino, más bien, presentar al público en general una opción de estadística oficial validada de manera correcta para tener indicadores económicos municipales con una temporalidad relativamente larga, la cual permita a los usuarios realizar análisis para los fines que se consideren convenientes, algunos de ellos, comentados con anterioridad.

Así, el artículo se divide de las siguientes secciones: en la primera se resume la metodología usada para generar IAEM; en la segunda se presentan los resultados del experimento Monte Carlo; en la tercera se muestran las estimaciones de los indicadores para el periodo 1993-2013 y la validación de la preservación del movimiento; y, finalmente, se dan las conclusiones y consideraciones finales.

Metodología

En esta sección se sintetizan las fuentes de información utilizadas en este trabajo y el procedimiento estadístico para obtener IAEM. Para mayores detalles, se recomienda al lector revisar Corona y López-Pérez (2019), cuyo resumen se presenta en el apartado *IAEM con información de los CE*.

Fuentes de información

La que representó mayor importancia para esta investigación son los microdatos de los CE 1994, 1999, 2004, 2009 y 2014 del INEGI. No fue posible

incorporar los de la edición 1989, ya que esta no dispone de información a ese nivel.

El objetivo de los CE, de acuerdo con la definición del INEGI (2015, 4), es: "... obtener información estadística básica y actualizada, sobre los establecimientos productores de bienes, comercializadores de mercancías y prestadores de servicios, para generar indicadores económicos de México a un gran nivel de detalle geográfico, sectorial y temático...". En consecuencia, con estos datos estuvimos en condiciones de comprender la estructura y ubicación geográfica de la actividad económica, pudiendo así establecer la composición económica de los municipios que forman el país. Otras fuentes que se usaron son el PIB por entidad federativa del SCNM, la cual se complementó con la información generada al seguir la metodología de Guerrero y Corona (2018), donde obtuvieron un PIB a nivel estatal desde 1993 para 10 grupos de sectores con año base 2013; en otras palabras, homologaron cifras entre el clasificador actual de actividades económicas (SCIAN) y el anterior (CIU).

Para esta investigación, se utilizó información de los CE que incorpora solo aquellas clases de actividad económica en las que las unidades de observación registren datos referidos al lugar donde estas se realizan (ver cuadro 1), es decir, se excluyeron aquellas en las cuales las unidades censadas presentaron datos que estaban referidos al sitio donde se ubica la oficina matriz o gerencia, pero que en realidad se desarrollan en un área geográfica mayor (INEGI 2015, 12).

Cuadro 1

Continúa

Número de clases por sector de actividad económica

Código	Sector	SCIAN 2013	CE 2014	IAEM 2019
S11	Agricultura, cría y explotación de animales, aprovechamiento forestal, pesca y caza	90	12	12
S21	Minería	31	29	1
S22	Generación, transmisión y distribución de energía eléctrica, suministro de agua y de gas por ductos al consumidor final	5	4	0

Número de clases por sector de actividad económica

Código	Sector	SCIAN 2013	CE 2014	IAEM 2019
S23	Construcción	45	44	0
S31-33	Industrias manufactureras	291	288	288
S43	Comercio al por mayor	71	71	71
S46	Comercio al por menor	83	83	83
S48-49	Transportes, correos y almacenamiento	64	59	5
S51	Información en medios masivos	35	34	34
S52	Servicios financieros y de seguros	27	24	0
S53	Servicios inmobiliarios y de alquiler de bienes muebles e intangibles	27	27	27
S54	Servicios profesionales, científicos y técnicos	41	37	37
S55	Corporativos	2	2	2
S56	Servicios de apoyo a los negocios y manejo de residuos y desechos, y servicios de remediación	28	28	28
S61	Servicios educativos	38	20	20
S62	Servicios de salud y de asistencia social	64	34	34
S71	Servicios de esparcimiento culturales y deportivos, y otros servicios recreativos	41	27	27
S72	Servicios de alojamiento temporal y de preparación de alimentos y bebidas	22	20	20
S81	Otros servicios excepto actividades gubernamentales	44	40	40
S93	Actividades legislativas, gubernamentales, de impartición de justicia y de organismos internacionales y extraterritoriales	10	0	0
	Total	1 059	883	729

Fuente: elaboración propia con base en las metodologías de los CE 1994-2014.

Una vez realizado el análisis de las clases de actividad económica con representatividad regional, se consideró necesario ajustar el sesgo a la izquierda que tiene la distribución del Valor Agregado Censal Bruto (VACB) a través de ajustar la regresión tipo Cobb-Douglas:

$$\log(VACB_i) = a + b_1 \log(hrs_i) + b_2 \log(Activos_i) + e_i$$

donde el subíndice $i = 1, 2, \dots, n$ identifica cada establecimiento dentro del sector, hrs_i son las horas-hombre, $Activos_i$ son los activos totales y e_i representa el error aleatorio; finalmente,

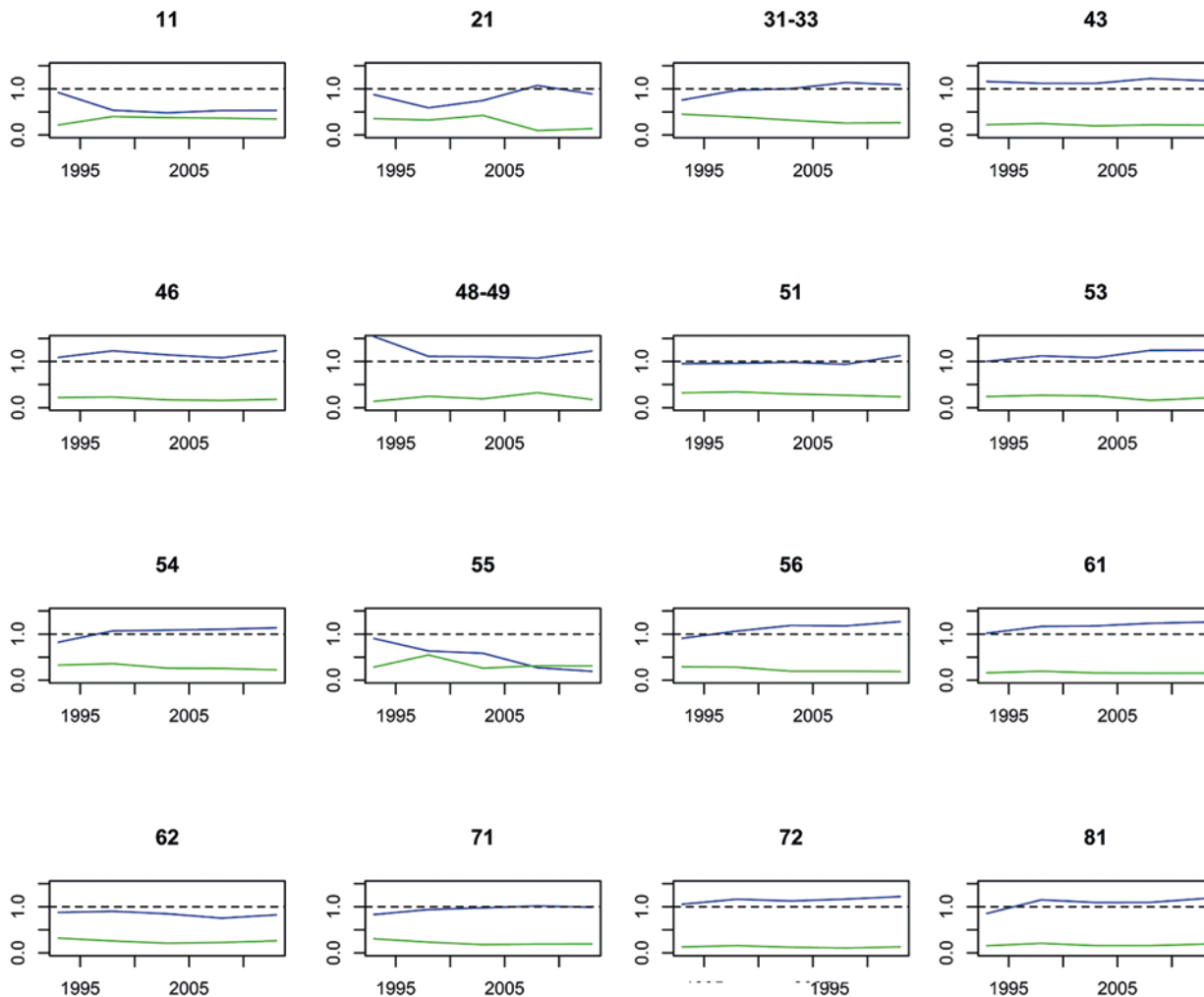
los parámetros b_1 y b_2 son las elasticidades atribuibles a los factores de la producción y a es el intercepto.

De esta manera, las cifras atípicas del VACB se ajustan con los coeficientes estimados de a , b_1 y b_2 . Es importante señalar que esto se realizó a consideración de la Coordinación de Diseño Conceptual y Estudios Económicos del INEGI.

En las gráficas 1 se muestran los coeficientes de estas estimaciones; en estas es posible observar la consistencia de los parámetros calculados a

Gráficas 1

Regresiones tipo Cobb-Douglas por sector de actividad económica



Nota: las líneas azules representan la estimación para b_1 y las verdes, para b_2 .
Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

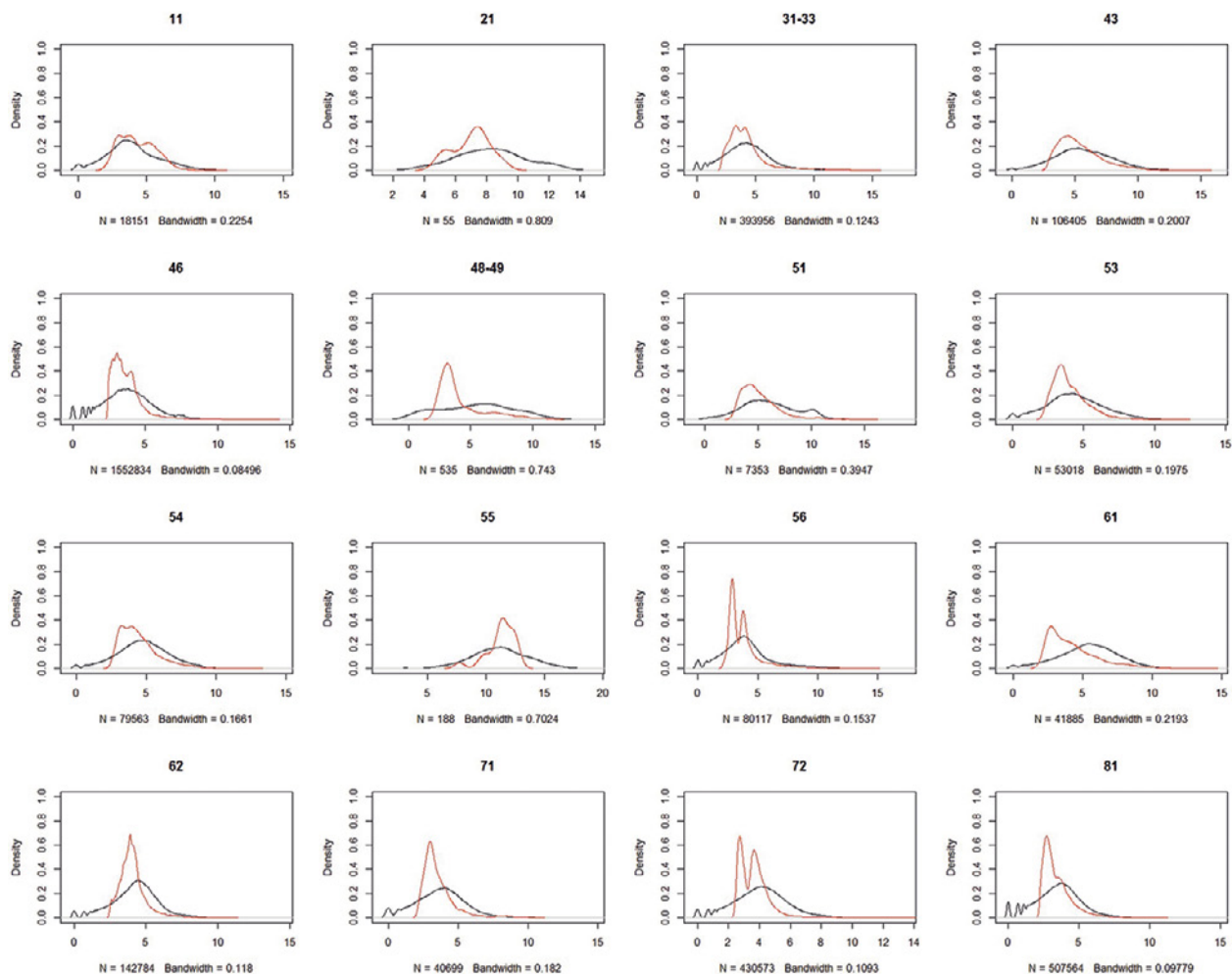
través del tiempo, salvo para el sector 55, donde la línea azul (que representa el parámetro asociado al número de horas-hombre trabajadas) decrece durante todo el periodo, la cual se puede ver que, de manera consistente, se encuentra por arriba de la línea verde (que es el parámetro asociado a los activos totales).

Se puede apreciar que, con excepción de los corporativos, las elasticidades tienen el mismo comportamiento, es decir, el VACB crece porcentualmente

más dado cambios porcentuales en las horas-hombre que cuando lo hacen los activos.

Como resultado de las estimaciones anteriores, se procedió con el ajuste de los datos por el sesgo, lo cual derivó en que la cola izquierda de la distribución fuera más corta y se mostrara asimétrica hacia ese lado. En las gráficas 2 se presenta una comparación de las densidades del VACB por sector de actividad económica: en negro antes de la corrección y en rojo una vez corregido el sesgo hacia la izquierda.

Distribución del VACB por sector antes y después de la corrección por sesgo



Nota: la línea negra representa la distribución del VACB antes de la corrección por sesgo y la roja, después de la corrección.

Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

Se puede ver cómo las distribuciones de los datos ajustados no están tan cargadas hacia la izquierda, lo cual es congruente al ajustar el VACB usando la especificación Cobb-Douglas. Vale la pena destacar que los datos considerados como atípicos son a los que se les asignó el valor ajustado o corregido; caso contrario, se considera el VACB observado.

Para visualizar el porcentaje que cubren los CE en el total de la economía y, también, las cifras —considerando solo las clases de actividad económica con representatividad regional— se computaron para

cada nueve grupos de sectores⁵ los porcentajes de cobertura promedio, es decir:

$$\frac{1}{5} \sum_{h=1}^5 Cobertura_h,$$

donde el subíndice h representa cada uno de los cinco años censales utilizados en este trabajo y la *cobertura_h* se define como $\frac{VACB_h}{PIB_h}$.

5 Concentran una serie de actividades económicas que tienen en común el origen de las mismas: 1) Agropecuario, silvicultura y pesca; 2) Minería; 3) Industria manufacturera; 4) Construcción; 5) Electricidad, gas y agua potable; 6) Comercio, restaurantes y hoteles; 7) Transporte, almacenamiento y comunicaciones; 8) Servicios financieros, seguros y bienes inmuebles; y 9) Servicios comunales, sociales y personales.

La gráfica 3 muestra los promedios, mínimos y máximos de los porcentajes de cobertura.

Se aprecia que, para S21, S22 y S52-S53, la cobertura se reduce de manera drástica respecto al total a lo que consideran los CE; no obstante, por lo comentado con anterioridad, en términos de representatividad regional se concluye que estos reflejan mejor la dinámica de las actividades económicas de los sectores en los municipios. Nótese que, en algunos casos, la cobertura de los CE es mayor que la reportada por el SCNM; sin embargo, al corregir solo con clases de actividad económica con representatividad regional, ya no ocurre esta situación, por ejemplo, en S31-S33 y S43-S46&S51. También, las cifras ligeramente mayores, al considerar las clases de actividad económica regional, vienen dadas por el ajuste de la regresión Cobb-Douglas.

Veleros *et al.* (2011) exponen con más detalle las principales razones que explican las discrepancias para los CE con datos referidos al 2003 y 2008,

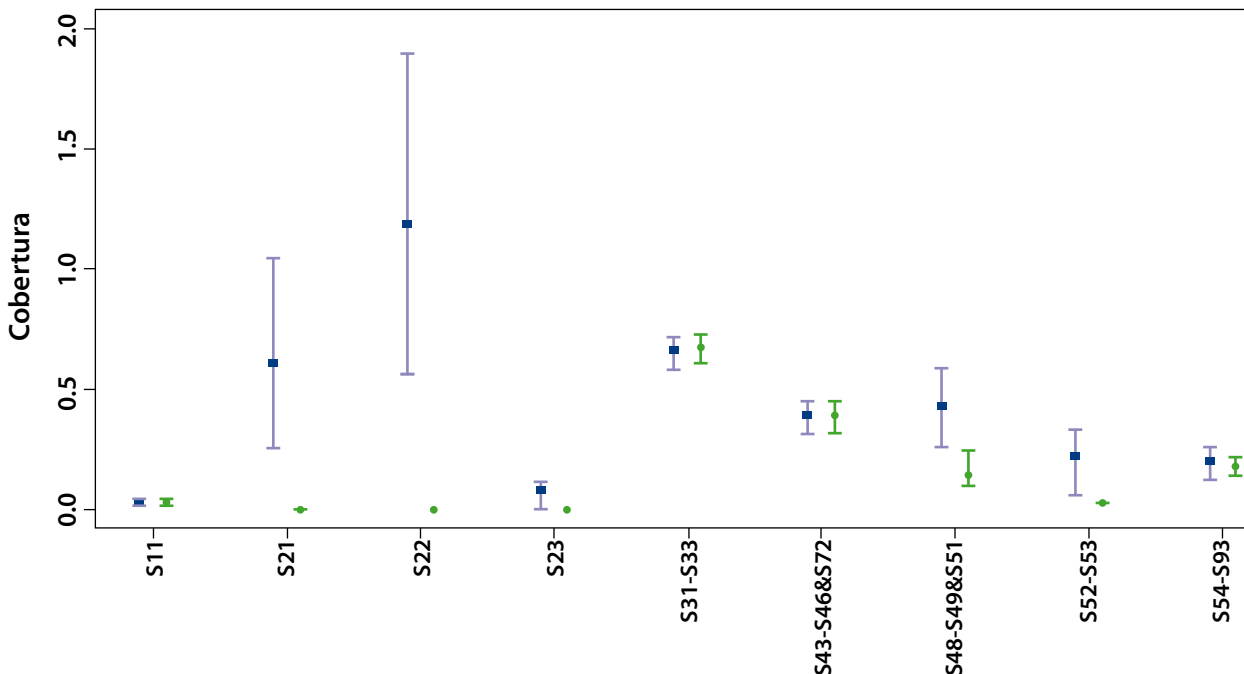
entre las que destacan: cobertura de sectores económicos (los CE no consideran al agropecuario, transporte y sector público); los diferentes criterios contables de las empresas productoras de cada entidad; que el SCNM realiza estimaciones de la actividad informal (22.7 % en 2017) y es, en especial, relevante en los sectores de agricultura (80 %), comercio al por menor (54 %) y construcción (44 %); la valoración de los CE a precios de productor, mientras que el SCNM lo hace a básicos; y la representatividad geográfica de las cifras de los CE.

El SCNM (INEGI, 2010, 8) establece: "... por lo expuesto, es posible afirmar que el VACB y el PIB calculado por el SCNM, no son comparables en sentido estricto. Ambas estadísticas pretenden divulgar datos relevantes del acontecer económico y ninguna de las dos tiene preeminencia sobre la otra, sino que deben ser dimensionadas de acuerdo con los propósitos que persiguen..."

Atendiendo al criterio de representatividad geográfica, el presente trabajo excluye entre 21 y 42 %

Gráfica 3

Porcentaje de cobertura promedio de los CE respecto al PIB considerado y no clases de actividad económica con representatividad geográfica



Nota: las líneas azules consideran el total de los CE y las verdes corresponden a las clases de actividad económica con representatividad geográfica.
Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

de la información de los CE; la principal diferencia derivada de las exclusiones radica en la cobertura de los sectores 21 (minería) y 22 (energéticos), mientras que en el 11 (actividades primarias) y 23 (construcción), ya de por sí, la cobertura de los CE es muy baja.

La gráfica 4 muestra el porcentaje promedio de cobertura respecto al PIB por entidad federativa, así como su nivel máximo y mínimo considerando todos los CE utilizados en esta investigación.

A nivel estatal, las coberturas se encuentran por debajo de 50 %, donde el más bajo es para Campeche derivado, sobre todo, de la exclusión de los sectores minero y energético. Ello da cuenta de que el VACB y el PIB calculados por el SCNM no son del todo comparables en sentido estricto; de esta manera, algunos sectores tenderán a representar mejor la dinámica que otros, sobre todo cuando su actividad económica se realice en un espacio físico determinado, como las manufacturas. De esta forma, el VACB y el PIB pretenden divulgar datos relevantes del acontecer económico y ninguno de los dos tie-

ne preeminencia sobre el otro, sino que deben ser dimensionados de acuerdo con los propósitos que persiguen (INEGI, 2010); no obstante, en un sentido econométrico, se pueden obtener estimaciones preliminares que, una vez impuestas las restricciones contables, aproximen de mejor manera lo que se desea calcular, hecho que se considera en este trabajo.

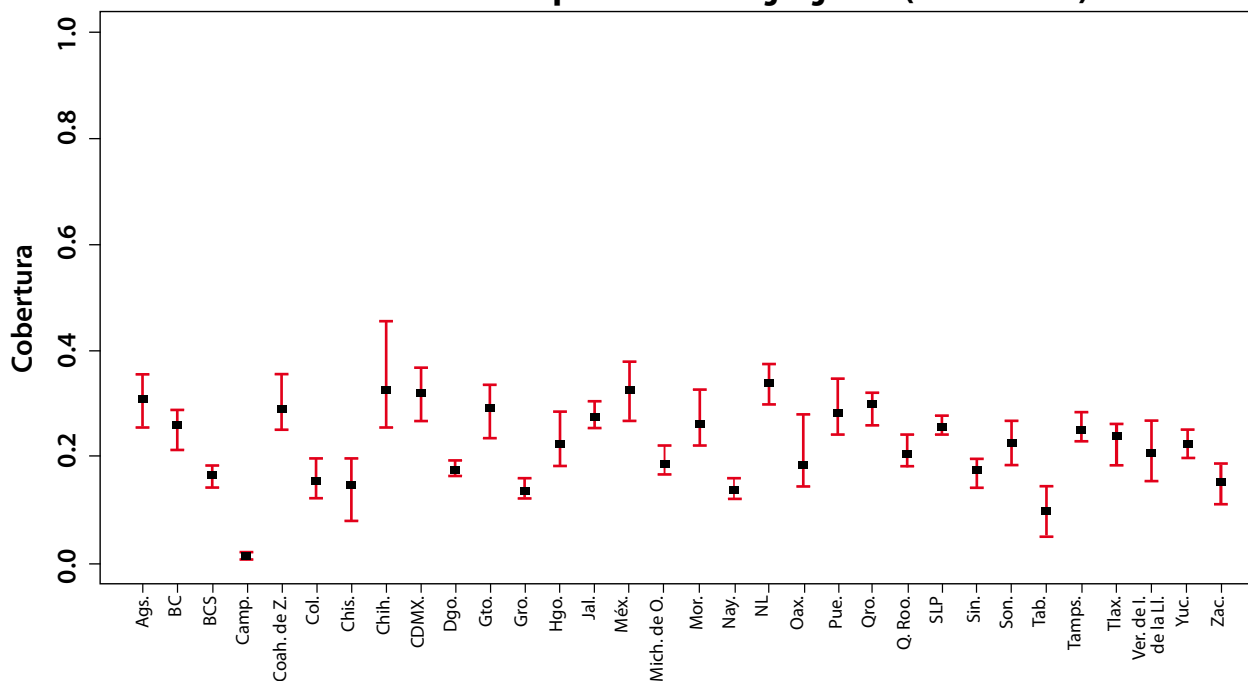
Una vez realizado este análisis se pudo concluir que, establecidas la representatividad geográfica y sectorial —dada esta por la cobertura y agrupación de la información oficial— para la construcción de los IAEM, se deberían considerar los siguientes grupos de sectores (entre paréntesis se muestra la cobertura de los IAEM respecto al SCIAN):

- S31-33: *Industrias manufactureras* (99%).
- S43-46 & S72: *Comercio y Servicios de alojamiento temporal y de preparación de alimentos y bebidas* (99 %).
- S54-S81: *Servicios no financieros* (70 %).⁶

6 No se incluye el S72.

Gráfica 4

Porcentaje de cobertura de los CE respecto al PIB por entidad federativa considerando sectores con representatividad geográfica (CE 1994-2014)



Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

IAEM con información de los CE

Para los detalles de notación estadística formal, se recomienda al lector revisar Corona y López-Pérez (2019). La idea general fue, primero, obtener estimaciones preliminares, las cuales, después, debían satisfacer restricciones oficiales, en este caso, que las sumas de los IAEM equivalieran a su respectivo agregado captado del SCNM obteniendo, así, estimaciones finales. En este artículo resumimos el procedimiento para su respectiva aplicación, que consiste en los siguientes pasos:

1. Para cada entidad y año censal, construir razones de VACB por grupo de sectores de actividad económica considerados, de tal forma que la suma de los municipios sea equivalente a 1.
2. Para cada grupo de sectores de actividad económica, estimar los años no censales usando interpoladores que pueden ser, por ejemplo: lineales, cuadráticos o cúbicos. Se considera la fecha de fundación de los municipios. Estos cálculos se consideran como preliminares.
3. Aplicar la RC de Guerrero y Peña (2003), donde la matriz de ponderación se obtiene siguiendo a Corona y López-Pérez (2019) para reconciliar las cifras preliminares, de tal forma que los IAEM del paso anterior sumen a su respectivo dato otorgado por el SCNM. Nótese que esta fase de reconciliación permite expresar los IAEM a año base 2013. Estas se consideran como las estimaciones finales.
4. Se pueden agregar los IAEM para obtener niveles superiores de agregación.
5. Se considera prudente expresar los IAEM como índices de volumen físico dado que el porcentaje de cobertura no es el total.

Nótese que el paso 2 conlleva a un error de estimación. En Corona y López-Pérez (2019) se puede ver que los interpoladores lineales minimizan el error de estimación cuando se supone que queremos estimar el PIB estatal a través de los CE y siguiendo la metodología aquí presentada; no

obstante, esto puede resultar limitativo solo a este conjunto de datos, por lo que, en este trabajo, se complementó este análisis a través de un experimento Monte Carlo, donde se simuló las condiciones bajo las cuales se interpolaron cifras.

Experimento Monte Carlo: evaluando el uso de interpoladores para años no censales

Con el objetivo de analizar el rendimiento en muestras finitas del uso de interpoladores para estimar datos faltantes en series de tiempo —explícitamente en este caso para los años no censales—, se estimó un experimento Monte Carlo simulando series con diferentes comportamientos estocásticos, es decir, estacionarias, estacionarias con tendencia determinística y caminatas aleatorias. Se pretendió que emularan características similares a las que tienen las anuales que desean ser estimadas. Lo anterior proporcionará información sobre el costo estadístico de usar como variables preliminares series de tiempo provenientes de una fase de estimación previa.

Se utilizaron siete procesos generadores de datos con una dimensión temporal $T = 21$, es decir, el horizonte temporal de los IAEM, los cuales son:

$$M1. y_t = \mu + \phi y_{t-1} + \varepsilon_t \text{ donde } \phi \sim U(0.5, 1).$$

$$M2. y_t = \mu + \gamma \varepsilon_{t-1} + \varepsilon_t \text{ donde } \gamma \sim U(0.5, 1).$$

$$M3. y_t = \mu + \phi y_{t-1} + \gamma \varepsilon_{t-1} + \varepsilon_t \text{ donde } \phi \text{ y } \gamma \sim U(0.5, 1).$$

$$M4. y_t = \mu + 0.5y_{t-1} + 0.2t + \varepsilon_t$$

$$M5. y_t = \mu + 0.5\varepsilon_{t-1} + 0.2t + \varepsilon_t$$

$$M6. y_t = \mu + 0.5y_{t-1} + 0.5\varepsilon_{t-1} + 0.2t + \varepsilon_t$$

$$M7. y_t = \mu + y_{t-1} + \varepsilon_t$$

En todos los casos se supuso que $\mu = 100$ y $\varepsilon_t \sim N(0, 1)$. Se consideraron conocidos solo $t = 1, 5, 9, 13, 17$ y 21 y se aplicaron los interpoladores que utilizamos en la investigación del 2019, los cuales son el uso del interpolador lineal (S_L), en el que se unen los puntos a través de una proporcionalidad lineal entre los datos faltantes (S_{Sp}), donde se ajusta una posición cúbica exacta a través de los cuatro puntos en cada extremo de los datos y el de Stineman

(1980) denotado como S_{ST} , donde se computan pendientes más bajas cerca de pasos o picos abruptos en la secuencia de puntos. Denotamos a w_t , como la estimación de y_t y se estimó el coeficiente de correlación $p = corr(y_t, w_t)$ en cada una de las $R = 500$ réplicas. Se presentaron los cuantiles de $\rho(0.05, 0.5, 0.95)$ para los datos en niveles como la primera diferencia del logaritmo. El cuadro 2 resume los resultados para los niveles de las series.

Se puede apreciar que, en todos los casos, el interpolador lineal supera a sus respectivos competidores y solo cuando los modelos cuentan con medias móviles sin tendencia el ajuste es débil; no obstante, cuando las series de tiempo tienen un comportamiento autorregresivo o no estacionario, ya sea determinista o estocásticamente, el S_L es muy competitivo; a este le sigue el S_{ST} y, por último, el S_{SP} .

Es importante señalar que, para todos los casos, la caminata aleatoria se ajusta relativamente bien, aunque los resultados deben ser tomados con cautela, pues la correlación solo es válida si las series de tiempo son cointegradas. El cuadro 3 muestra lo mismo que el 2, pero para las primeras diferencias del logaritmo.

En todas las situaciones, el S_L fue el que obtuvo mejores resultados. En este caso, todas las correlaciones para este interpolador fueron significativas. También, se corrobora que conforme más tendencias tengan las series de tiempo mejor será el ajuste que se obtenga al momento de utilizar los interpoladores, aun cuando las series estén en primeras diferencias.

Se realizaron diferentes experimentos Monte Carlo variando el número de puntos a estimar,

Cuadro 2

Resultados del experimento Monte Carlo: cuantiles 0.05, 0.5 y 0.95 de p para la serie en niveles

Experimento	S_L	S_{SP}	S_{ST}
M1	[0.36, 0.78, 0.96]	[0.26, 0.75, 0.95]	[0.33, 0.77, 0.96]
M2	[-0.01, 0.15, 0.35]	[-0.05, 0.10, 0.33]	[-0.03, 0.12, 0.34]
M3	[0.00, 0.34, 0.70]	[-0.09, 0.31, 0.68]	[-0.03, 0.33, 0.69]
M4	[0.63, 0.85, 0.94]	[0.50, 0.83, 0.94]	[0.58, 0.84, 0.94]
M5	[0.47, 0.72, 0.85]	[0.36, 0.68, 0.84]	[0.41, 0.70, 0.85]
M6	[0.57, 0.79, 0.90]	[0.48, 0.76, 0.89]	[0.54, 0.78, 0.90]
M7	[0.62, 0.91, 0.98]	[0.60, 0.89, 0.98]	[0.62, 0.90, 0.98]

Fuente: elaboración propia con base en simulaciones.

Cuadro 3

Resultados del experimento Monte Carlo: cuantiles 0.05, 0.5 y 0.95 de p para la serie logarítmica y en primeras diferencias

Experimento	S_L	S_{SP}	S_{ST}
M1	[0.18, 0.38, 0.61]	[0.03, 0.30, 0.58]	[0.11, 0.36, 0.59]
M2	[0.09, 0.19, 0.31]	[-0.06, 0.06, 0.24]	[-0.01, 0.09, 0.25]
M3	[0.09, 0.23, 0.41]	[-0.07, 0.10, 0.32]	[-0.01, 0.16, 0.36]
M4	[0.15, 0.30, 0.53]	[-0.01, 0.21, 0.49]	[0.07, 0.27, 0.49]
M5	[0.09, 0.19, 0.32]	[-0.09, 0.07, 0.27]	[-0.01, 0.10, 0.25]
M6	[0.12, 0.23, 0.39]	[-0.07, 0.11, 0.34]	[0.01, 0.16, 0.35]
M7	[0.20, 0.43, 0.68]	[0.10, 0.37, 0.63]	[0.15, 0.41, 0.65]

Fuente: elaboración propia con base en simulaciones.

mayores y menores a $T = 21$, y se pudo concluir que los resultados por interpolador son insensibles al tamaño de muestra, es decir, los obtenidos en los cuadros 2 y 3 no se ven modificados de manera estadística.

Por último, con el fin de verificar las condiciones del proceso de simulación, se ajustaron modelos ARIMA a las series del PIB de las 32 entidades federativas para el periodo 1993-2013⁷ por GA y el total de actividad económica (128 series), para lo cual se utilizó la función *auto.arima* del paquete *forecast* en el programa *R*. Los resultados se muestran en la gráfica 5.

Se puede ver que las condiciones en las que fue generado el experimento abarcan, de buena forma, las características de las series estatales observadas; la mayor parte de estas se ajustan con un modelo ARIMA (0, 1, 0) con deriva, es decir, caminatas aleatorias, las cuales son incluidas en el experimento Monte Carlo de este trabajo.

Nótese que este ejercicio de simulación es independiente de los resultados y/o de la aplicación de la metodología propuesta, solo informa o describe datos sobre el error de estimación que se puede obtener al hacer uso de interpoladores para los años no censales, el cual se minimiza usando el lineal.

Resultados: IAEM

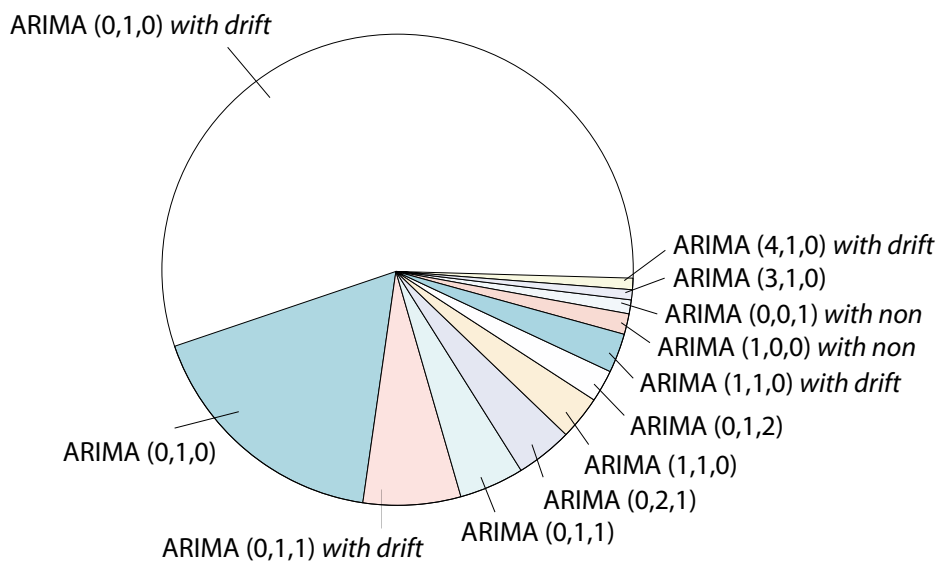
En esta sección se presentan los indicadores calculados para los 20 municipios que al 2013 registraron mayores niveles de actividad económica en el total. En primer lugar, están los resultados para todos los sectores considerados; enseguida, para las industrias manufactureras; y, por último, para los servicios no financieros que corresponden a la suma de los grupos de sectores contenidos en S43-48 y S54-S81. En todos los casos se muestran las estimaciones preliminares y las finales obtenidas al aplicar la RC.

Como se mencionó anteriormente, los IAEM que se obtuvieron no consideran el total de las actividades económicas en el sentido del SCNM, sino que

⁷ Series que consideran los cálculos del SCNM y no incluyen al periodo de retroprolación.

Gráfica 5

Ajustes de modelos ARIMA para las 128 series del PIBE por GA



Fuente: elaboración propia con base en ajustes de modelos ARIMA al PIB por entidad federativa.

son un reflejo de las que se pueden atribuir al lugar donde se desempeña la actividad económica, en este caso al municipio usando la representatividad geográfica contenida en los CE (ver gráficas 6).

Los resultados muestran que la delegación Miguel Hidalgo encabezó la lista con el mayor nivel de actividades económicas, seguida por los municipios Monterrey y Guadalajara. Se puede observar que la Miguel Hidalgo y Monterrey tuvieron un crecimiento desde 1996 hasta el 2008, pero a partir de esta última fecha presentaron una reducción; en

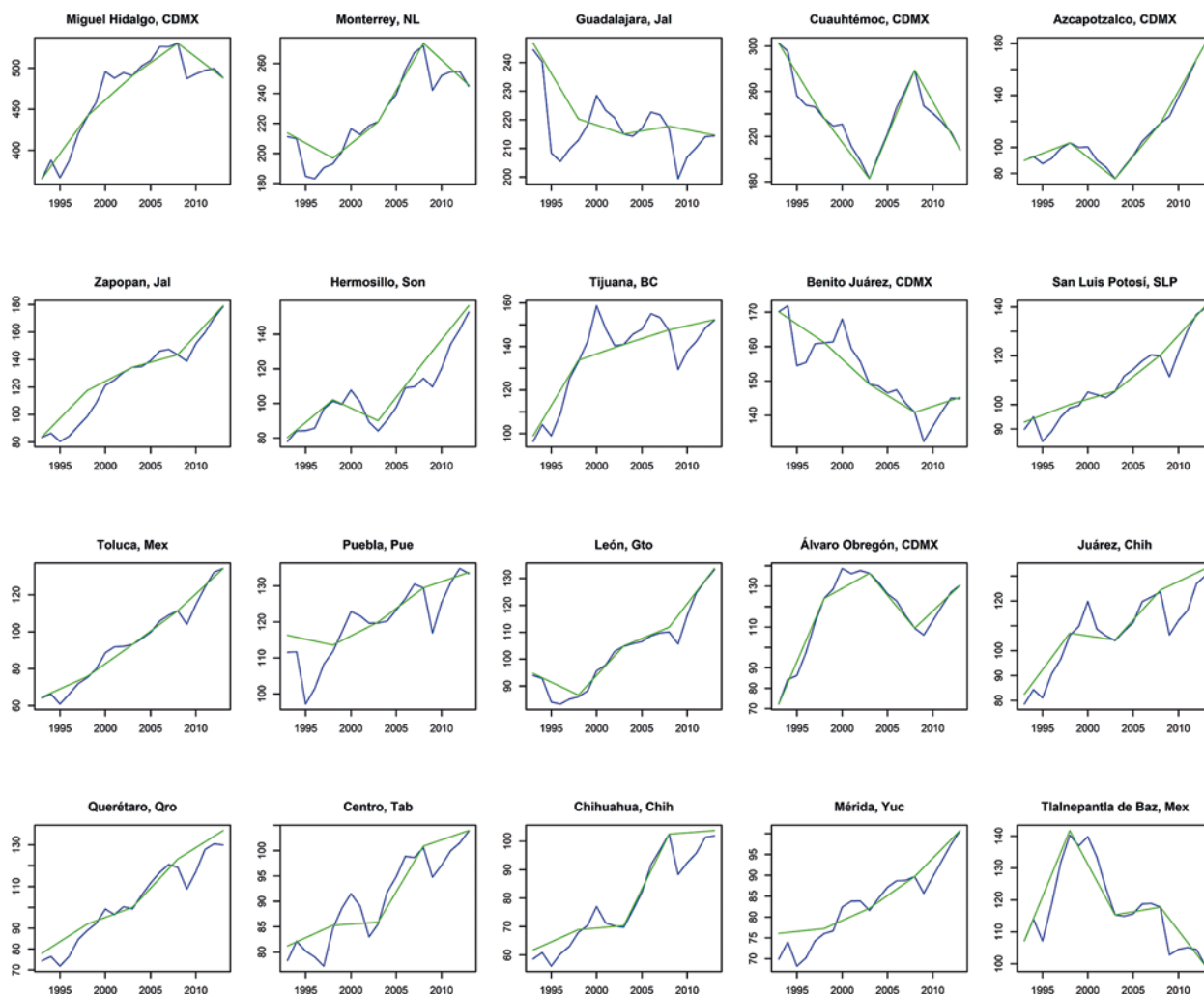
cambio, Guadalajara mostró una caída en todo el periodo de análisis. En el 2013, los tres concentraron 10.8 % del total, mientras que los 17 restantes más grandes en conjunto acumularon 37.66 por ciento.

Los resultados para las primeras diferencias del logaritmo, es decir para las variaciones porcentuales anuales, se muestran en las gráficas 7.

Las variaciones porcentuales anuales nos indica que, a pesar del uso de interpoladores lineales

Gráficas 6

IAEM en niveles para el total de la economía

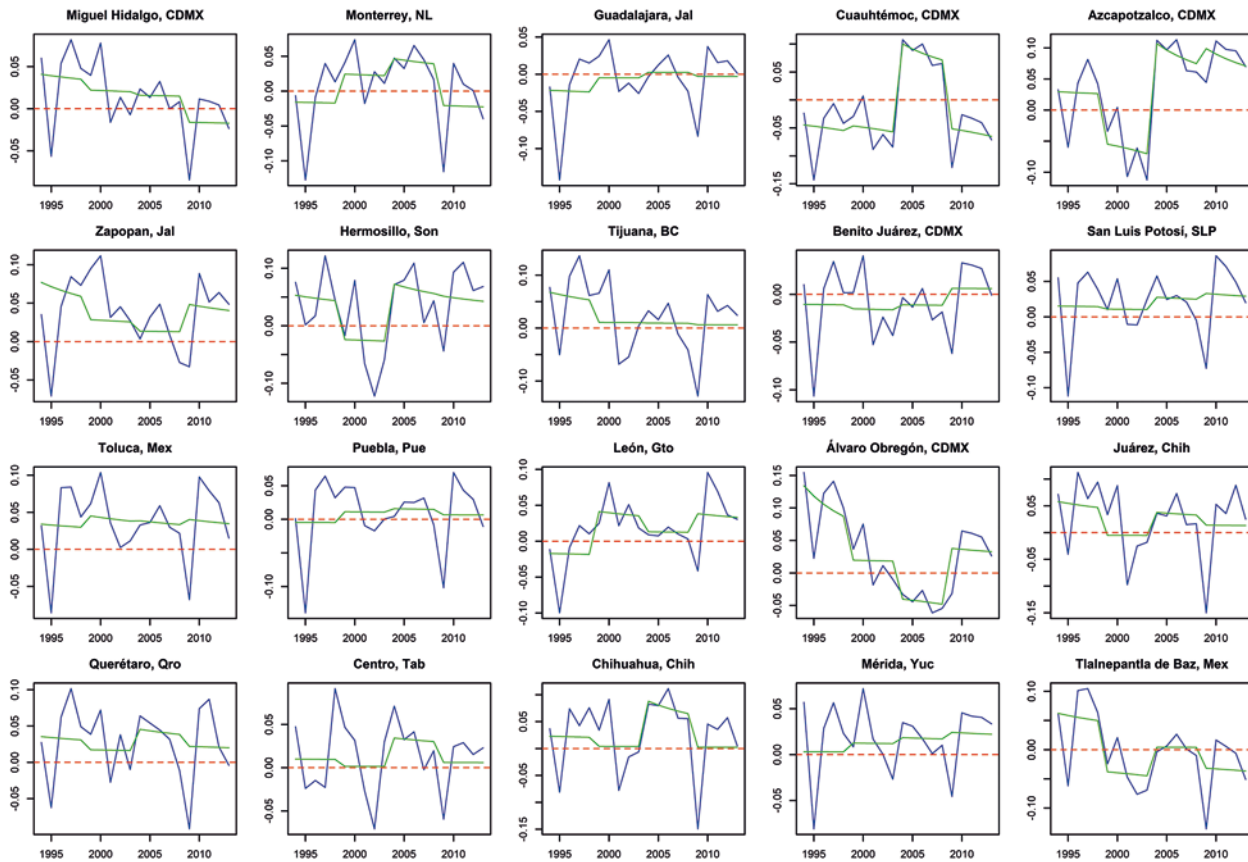


Nota: las líneas azules representan los IAEM y las verdes corresponden a las estimaciones preliminares.

Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

Gráficas 7

Primeras diferencias del logaritmo del IAEM para el total de la economía



Nota: las líneas azules representan los IAEM y las verdes corresponden a las estimaciones antes de reconciliar.

Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

les, la técnica de *abajo hacia arriba* —la cual consiste en agregar por grupos de sectores y dada la utilización de la RC— permite que las variaciones se muestren persistentes en todo el periodo, llamando la atención las caídas en 1995 y el 2009. De manera general, la Miguel Hidalgo y el municipio San Pedro Garza García son los que exhibieron las tasas de crecimiento más pronunciadas al final del periodo, mientras que la Cuauhtémoc y Monterrey mostraron ligeros decrecimientos en el 2013.

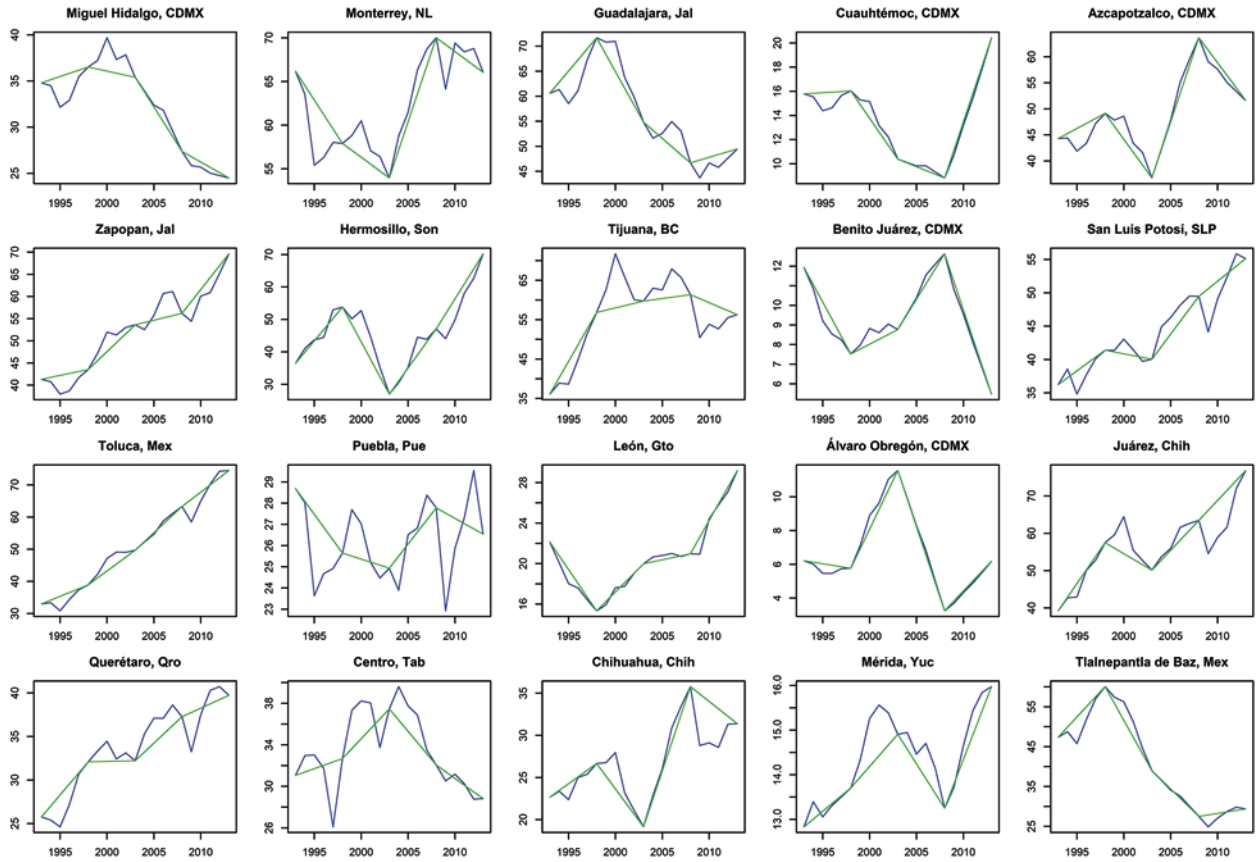
Es de interés observar que hay territorios que exhibieron caídas pronunciadas en sus tendencias (la Miguel Hidalgo, Guadalajara, la Benito Juárez y Álvaro Obregón), mientras que, por otro lado, los que crecieron de manera llamativa (Monterrey, la

Cuauhtémoc y Azcapotzalco, Zapopan, Hermosillo y San Luis Potosí). Las gráficas 8 muestran los niveles de los IAEM para las industrias manufactureras y las 9 presentan lo mismo, pero para las primeras diferencias del logaritmo.

En términos porcentuales, las industrias manufactureras exhibieron variaciones negativas para una gran cantidad de municipios durante el último periodo, donde destacaron la Miguel Hidalgo y Cuauhtémoc, así como Zapopan. Por otra parte, sobresalieron las positivas de la Cuauhtémoc, Monterrey y Guadalajara.

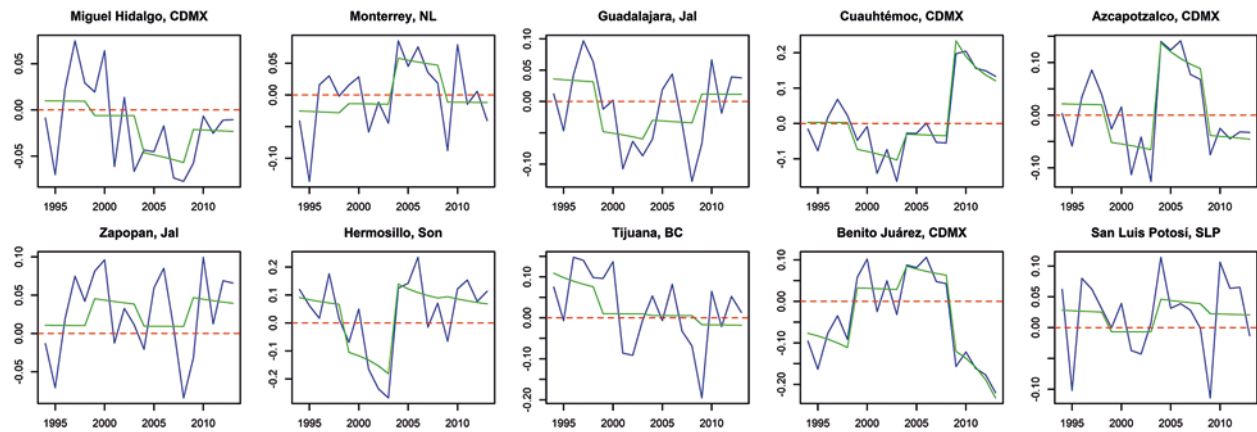
Para los servicios no financieros, las tendencias de los IAEM, en casi todos los casos, presentaron

IAEM en niveles para las industrias manufactureras

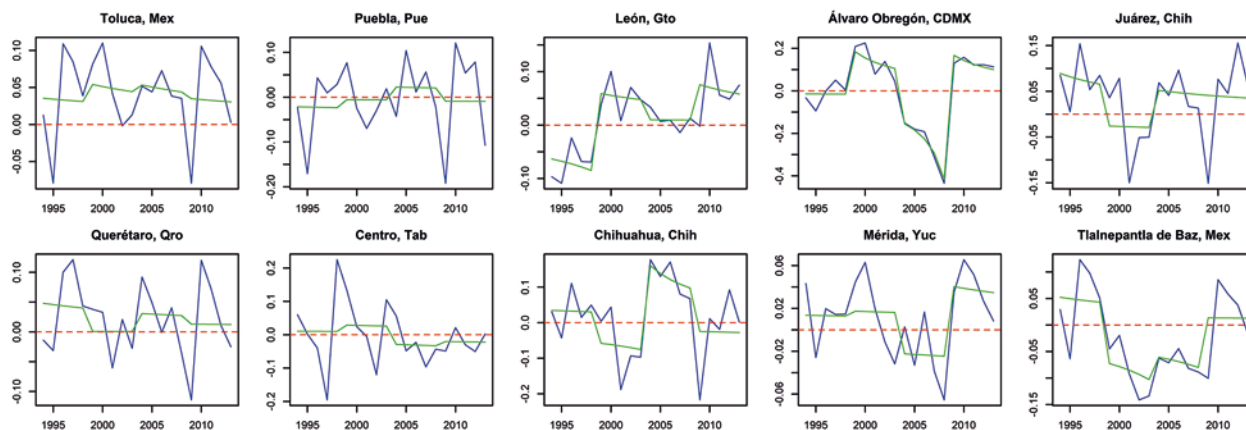


Nota: las líneas azules representan los IAEM y las verdes corresponden a las estimaciones antes de reconciliar.
Fuente: elaboración propia con base en los microdatos de los CE 1994-2014

Primeras diferencias del logaritmo del IAEM para las industrias manufactureras



Primeras diferencias del logaritmo del IAEM para las industrias manufactureras



Nota: las líneas azules representan los IAEM y las verdes corresponden a las estimaciones antes de reconciliar.
Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

una creciente clara, lo cual fue consistente con el PIB de la GA3 que presentaron todas las entidades federativas; sin embargo, de 1999 al 2003, la Cuauhtémoc mostró una caída, se recuperó en el 2008 y volvió a caer a partir de esa fecha. Los que, aunque con tendencia creciente, tuvieron una ligera caída en los últimos periodos son la Miguel Hidalgo y Monterrey.

muestra esto mismo, pero para las variaciones porcentuales anuales.

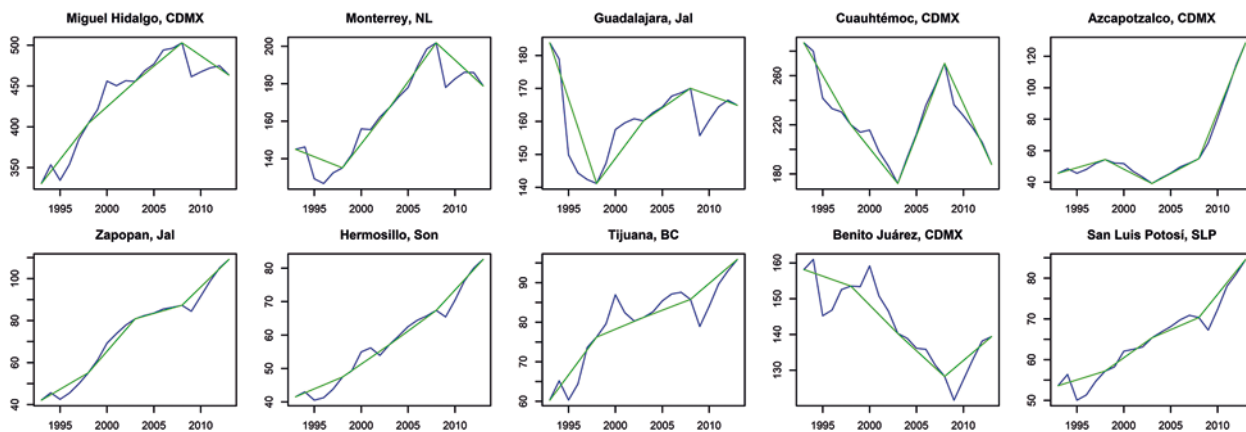
Las variaciones porcentuales anuales de los grupos de sectores que representan a los servicios no financieros mostraron características peculiares; llaman la atención las caídas por las crisis de 1995 y el 2009 para casi todos los municipios; los que, en el último dato, presentaron variaciones negativas son la Miguel Hidalgo, Monterrey, Guadalajara y la Cuauhtémoc.

Los resultados para los servicios no financieros en niveles se presentan en las gráficas10, y las 11

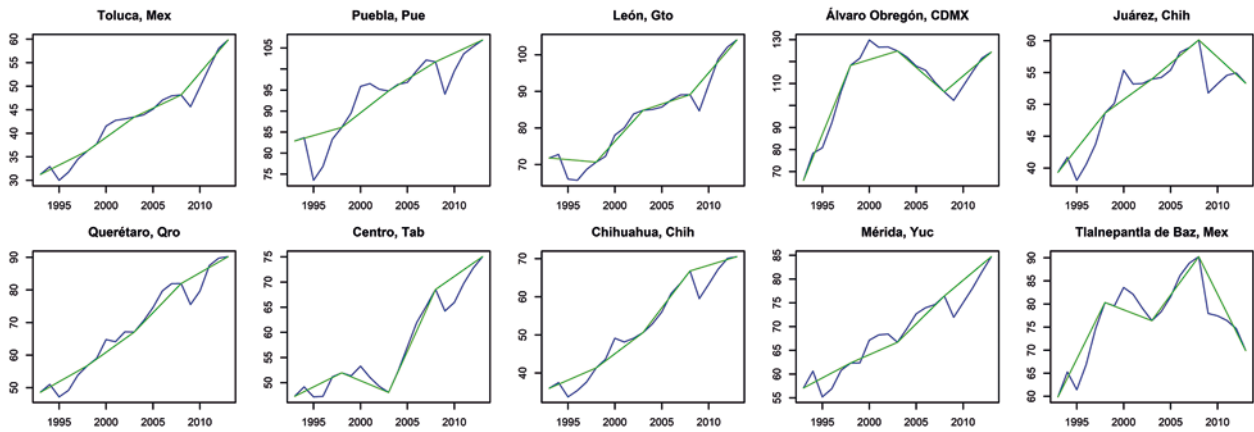
Gráficas 10

Continúa

IAEM en niveles para los servicios no financieros



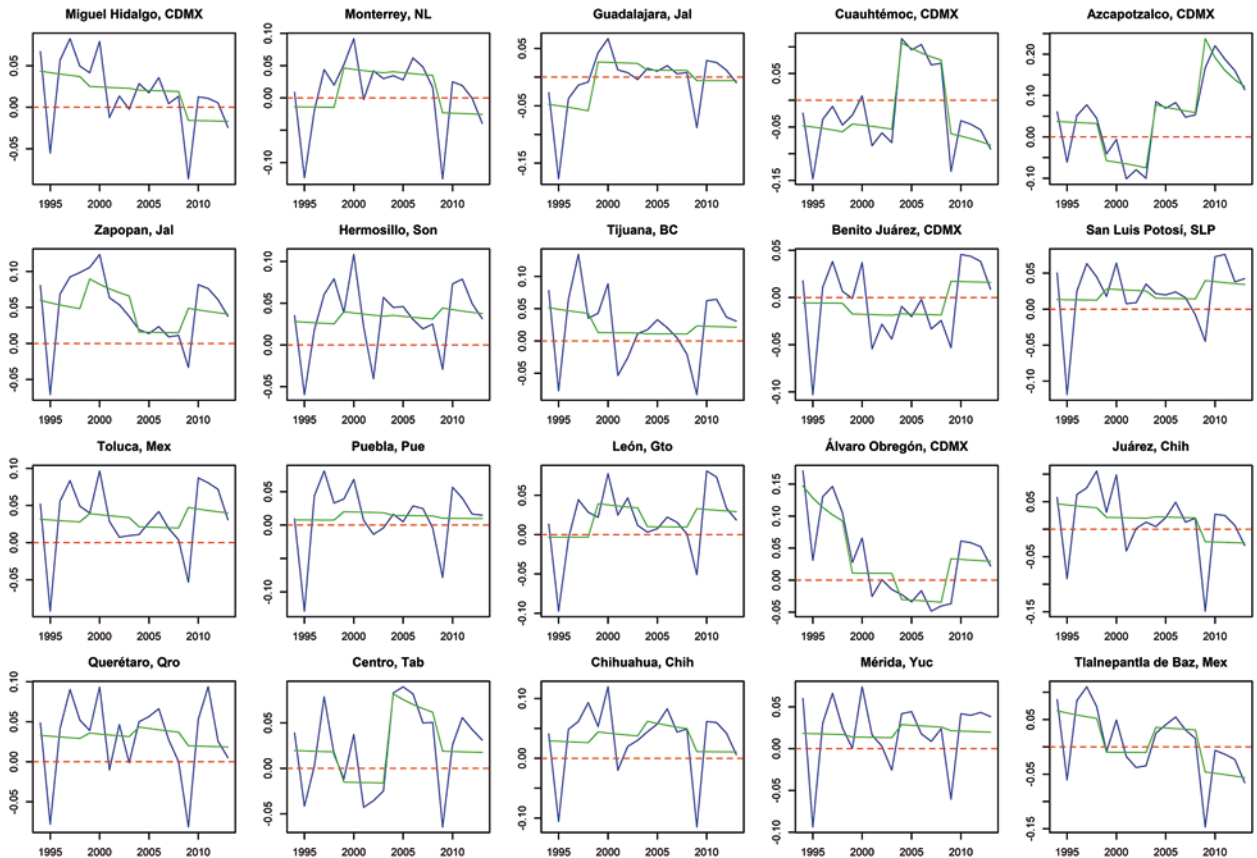
IAEM en niveles para los servicios no financieros



Nota: las líneas azules representan los IAEM y las verdes corresponden a las estimaciones antes de reconciliar.
Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

Gráficas 11

IAEM para las primeras diferencias del logaritmo de los servicios no financieros



Nota: las líneas azules representan los IAEM y las verdes corresponden a las estimaciones antes de reconciliar.
Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

Para verificar la preservación del movimiento, lo cual garantiza la consistencia de las estimaciones finales de acuerdo con Guerrero y Peña (2003), se calculó para cada municipio la diferencia entre la preliminar y la final, denotada como D_t ; después, nos centramos en verificar que dicha diferencia estuviera centrada en 0, o sea, que no hubiera sesgo y, también, que no se encontrara autocorrelacionada, es decir, que la diferencia fuera ruido blanco. Para constatar el no sesgo, se realizaron pruebas t , cuyos resultados se resumen en el cuadro 4.

Se puede observar que, con excepción de pocos casos, en la gran mayoría no existe sesgo entre las estimaciones preliminares y finales, es decir, la resta entre ambas series está centrada en 0. En los que no sucede esto, por ejemplo, son para los IAEM de los servicios no financieros de Aguascalientes, Durango y Quintana Roo.

Para verificar que D_t no estuviera autocorrelacionada, se estimaron pruebas Ljung-Box utilizando un rezago de $p = 4$, considerando que es el número de periodos entre cada censo. Los resultados se resumen en el cuadro 5.

De igual manera, se aprecian algunas excepciones donde no se puede validar que sea ruido blanco. Algunos ejemplos para las industrias manufactureras son Baja California, Baja California Sur, Campeche y Chihuahua, entre otros; no obstante, la gran mayoría de las diferencias están no autocorrelacionadas.

Cabe destacar que aun cuando no se pudo validar la preservación del movimiento y, por ende, la consistencia estadística de las estimaciones para algunos municipios, estos IAEM son la primera aproximación formal del comportamiento econó-

Cuadro 4

Medianas de p valores de pruebas de no sesgo

Estado	Industrias manufactureras	Servicios no financieros	Estado	Industrias manufactureras	Servicios no financieros
Ags.	0.79	0.02*	Mor.	0.53	0.09
BC	0.80	0.76	Nay.	0.09	0.29
BCS	0.15	0.54	NL	0.65	0.19
Camp.	0.13	0.35	Oax.	0.04*	0.74
Coah. de Z.	0.22	0.31	Pue.	0.45	0.35
Col.	0.80	0.40	Qro.	0.68	0.71
Chis.	0.10	0.85	Q. Roo	0.04*	0.00*
Chih.	0.34	0.11	SLP	0.74	0.17
CDMX	0.82	0.76	Sin.	0.30	0.27
Dgo.	0.00*	0.04*	Son.	0.27	0.42
Gto.	0.18	0.14	Tab.	0.44	0.31
Gro.	0.58	0.40	Tamps.	0.59	0.17
Hgo.	0.19	0.08	Tlax.	0.54	0.10
Jal.	0.73	0.31	Ver. de I. de la Ll.	0.33	0.85
Méx.	0.73	0.32	Yuc.	0.00*	0.43
Mich. de O.	0.68	0.71	Zac.	0.58	0.08

* Resultado de la prueba significativo a 5 por ciento.

Fuente: elaboración propia.

Medianas de *p* valores de pruebas Ljung-Box

Estado	Industrias manufactureras	Servicios no financieros	Estado	Industrias manufactureras	Servicios no financieros
Ags.	0.64	0.29	Mor.	0.08	0.02*
BC	0.01*	0.06	Nay.	0.08	0.06
BCS	0.00*	0.12	NL	0.11	0.10
Camp.	0.04*	0.10	Oax.	0.41	0.01
Coah. de Z.	0.15	0.09	Pue.	0.04*	0.08
Col.	0.21	0.05	Qro.	0.09	0.05
Chis.	0.08	0.07	Q. Roo	0.00*	0.33
Chih.	0.04*	0.10	SLP	0.17	0.16
CDMX	0.01*	0.08	Sin.	0.22	0.37
Dgo.	0.08	0.12	Son.	0.08	0.05
Gto.	0.90	0.23	Tab.	0.01*	0.04*
Gro.	0.15	0.13	Tamps.	0.02*	0.15
Hgo.	0.01*	0.05	Tlax.	0.07	0.27
Jal.	0.01*	0.04*	Ver. de I. de la LI.	0.00*	0.12
Méx.	0.09	0.08	Yuc.	0.09	0.03*
Mich. de O.	0.18	0.28	Zac.	0.05	0.82

* Resultado de la prueba significativo a 5 por ciento.

Fuente: elaboración propia.

mico a ese nivel sectorial, geográfico y temporal de la cual se tiene conocimiento, siendo importante su análisis estructural para los fines que el usuario considere conveniente.

Para terminar, y complementar el análisis, se obtuvieron las tasas medias de crecimiento anual (TMCA) en el periodo 1993-2013 para los 2 457 municipios considerados, y se encontró que 1 990 registraron una variación positiva, 449 las tuvieron negativas y 18 se mantuvieron sin cambio (ver mapa).

Por tamaño de población, para los 11 municipios con más de 1 millón de habitantes (al 2013), los que tuvieron mayores TMCA son Zapopan, Jalisco (3.89 %), Ciudad Juárez, Chihuahua (2.56 %), y Tijuana, Baja California (2.31 %), mientras que los

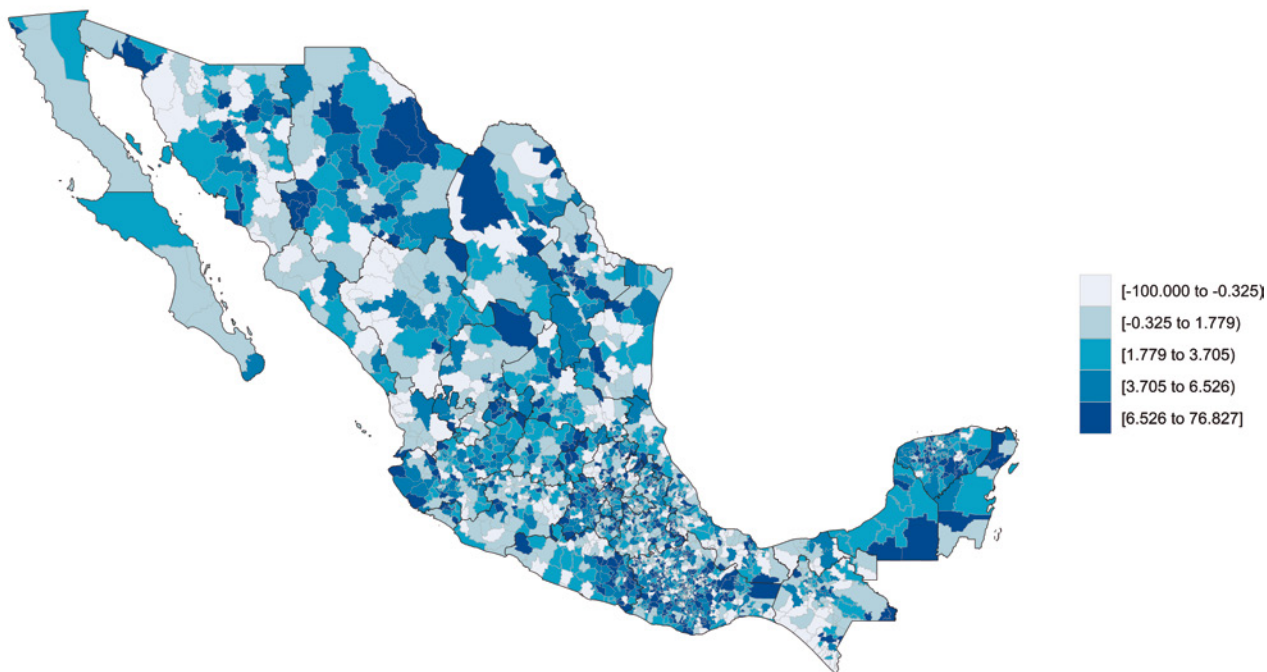
que registraron las menores, la Gustavo A. Madero, CDMX (- 0.96 %), Guadalajara, Jalisco (-0.65 %), y Nezahualcóyotl, estado de México (0.39 %).

Para los 210 con más de 100 mil habitantes y menos de 1 millón, los que mostraron TMCA más altas en el periodo son El Marqués, Querétaro (19.1 %), Solidaridad, Quintana Roo (15.83 %), y Silao de la Victoria, Guanajuato (13.42 %); los que tuvieron niveles más bajos son Chiapa de Corzo, Chiapas (- 8.66 %), Taxco de Alarcón, Guerrero (-5.6 %), y Chilón, Chiapas (-4.31 %).

Por último, para los 938 con más de 15 mil habitantes y menos de 100 mil, los que presentaron mayores TMCA son Bacalar, Quintana Roo (76.83 %), San Jacinto Amilpas, Oaxaca (28.97 %), y Tepezalá, Aguascalientes (26.88 %); los que tuvieron me-

Mapa

TMCA 1993-2013 de municipios de México



Fuente: elaboración propia con base en los microdatos de los CE 1994-2014.

nores tasas son Noria de los Ángeles, Zacatecas (-16.62 %), Atoyac, Veracruz de Ignacio de la Llave (-12.45 %), y Huixtán, Chiapas (-9.26 %).⁸

Conclusiones

En este artículo se presenta la metodología basada en Corona y López-Pérez (2019) para generar IAEM hasta 1993. Se incluyeron algunas mejoras metodológicas, como tomar solo clases de actividad económica con representatividad municipal; a manera de complementar el análisis, se realizaron experimentos Monte Carlo para tener certeza estadística del uso de interpoladores para los

años no censales; y se verificó el principio de preservación del movimiento que garantiza estadísticamente la consistencia de las estimaciones al hacer uso de la RC. Se puede observar que, dado lo anterior, lo que subyace también en utilizar un enfoque de *abajo hacia arriba* y la incorporación de restricciones contemporáneas, es decir, limitar la suma de los IAEM al total estatal por grupo de sectores, estos permiten capturar la dinámica de las actividades económicas a ese nivel de desagregación.

Se concluye que esta metodología permite a los usuarios de información económica contar con series más largas para analizar la evolución de los municipios y, por ende, generar políticas en este tema; no obstante, se recomienda tener en con-

⁸ Existen, además, 1 298 municipios con menos de 15 mil habitantes.

sideración que los IAEM, al generar un panorama global de las actividades económicas, en ningún caso pueden considerarse medidas del PIB, esto debido a que la información con la cual se construyen no abarca el total de estas de dichos ámbitos territoriales, sobre todo por la naturaleza de los CE y al uso de técnicas estadísticas para generar los IAEM; sin embargo, se considera un trabajo que combina el enfoque de micro y macrodatos para dar solución a un requerimiento de la sociedad.

Como líneas futuras de esta investigación, se recomienda adecuado usar las cifras generadas en este trabajo como estadísticas experimentales en el INEGI, a manera de incentivar su uso y discusión dentro de la sociedad. También, se pretenden actualizar los resultados una vez que se tengan los datos de los CE 2019.

Fuentes

Corona, Francisco y Jesús López-Pérez. "Obteniendo indicadores de actividad económica municipal basados en información representativa de los Censos Económicos", en: *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*. 10(2). 2019, pp. 62-81.

González, Luis. "Conformación política y jurídica del municipio: la incógnita del municipio mexicano", en: De Lameiras, Brigitte B. (ed.). *El municipio en México*. El Colegio de Michoacán, 1987, pp. 143-156.

González-Estrada, Adrián, y Gilberto Gallegos-Cedillo. "El producto interno bruto de los municipios de México: II. Estados MZ", en: *Revista Mexicana de Ciencias Agrícolas*. 5(8), 2014, pp. 1405-1421.

Guci, Ledia, Charles Ian Mead y Sharon D. Panek. *A Research Agenda for Measuring GDP at the County Level*. US Department of Commerce, Bureau of Economic Analysis. 2016.

Guerrero, Víctor M. y Daniel Peña. "Combining multiple time series predictors: a useful inferential procedure", en: *Journal of Statistical Planning and Inference*. 116(1), 2003, pp. 249-276.

Guerrero, Víctor M. y Francisco Corona. "Retropolación hasta 1980 de algunas series del Sistema de Cuentas Nacionales de México", en: *Realidad, Datos y Espacio Revista Internacional de Estadística*. 9(3), 2018, pp. 98-119.

INEGI. *Metodología de los Censos Económicos 2014*. Aguascalientes, INEGI, 2015.

_____. *Sistema de Cuentas Nacionales de México. Producto Interno Bruto por entidad federativa 2003-2007*. Aguascalientes, INEGI, 2010.

Instituto de Información e Investigación Geográfica, Estadística y Catastral del Estado de México (ICEGEM). *Producto Interno Bruto Municipal*. Gobierno del Estado de México, Secretaría de Finanzas, 2019.

Stineman, Russell W. "A Consistently Well Behaved Method of Interpolation", en: *Creative Computing*. 6(7), 1980, pp. 54-57.

Veleros, Zeus, Salvador Hernández, Gonzalo Dolores de la Merced y César Amador Ambriz. "Fundamento metodológico, discrepancias estadísticas y errores conceptuales en el uso de datos económicos", en: *Paradigma Económico*. 3(1), 2011, pp. 71-110.

Cubo de datos geoespaciales

para el uso de las imágenes satelitales
en la generación de información
geográfica y estadística

The Use
of Satellite Imagery
in the Generation of
Geographic and Statistical Information

Olivia Jimena Juárez Carrillo, Paloma Merodio Gómez, María del Socorro Ponce Medina, José Luis Ornelas de Anda y Abel Alejandro Coronado Iruegas*

Los satélites brindan mediciones a intervalos regulares, lo que presenta una oportunidad única para construir series de tiempo consistentes y derivar su análisis en tendencias de comportamiento. Gracias a esta periodicidad, sabemos que el nivel del mar realmente está subiendo, que las selvas tropicales están desapareciendo y con qué rapidez sucede todo esto. Con cientos de imágenes gratuitas liberadas de forma anual, la falta de datos ya no es una limitante; sin embargo, debido a factores como su volumen y complejidad, siguen siendo subutilizadas. En este sentido, este artículo presenta el *Cubo de Datos Geoespaciales de México*, herramienta que busca aprovechar el potencial de dichas imágenes para la generación de información.

Palabras clave: cubo de datos geoespaciales; imagen satelital; observaciones de la Tierra; *Landsat*; datos listos para el análisis.

Recibido: 3 de septiembre de 2019.
Aceptado: 1 de junio de 2020.

* Instituto Nacional de Estadística y Geografía, jimena.juarez@inegi.org.mx, paloma.merodio@inegi.org.mx, ma.ponce@inegi.org.mx, jose.ornelas@inegi.org.mx y abel.coronado@inegi.org.mx, respectivamente.

Satellites provide measurements at regular intervals, which presents a unique opportunity to build consistent time-series and derive this analysis into behavioral trends. Thanks to this periodicity, we know that the sea level is really rising, or that the rainforests are disappearing and how quickly all of this is happening. With hundreds of free images released annually, the lack of data is no longer a limitation; however, due to factors such as volume and complexity, these images remain underutilized. This article presents the *Mexican Geospatial Data Cube*, a tool that allows analyzing and storing satellite images in a more efficient way, seeking to exploit their potential in the generation of information.

Key words: Geospatial Data Cube; satellite image; Earth observations; Landsat; analysis ready data.



Planet Earth against black background/James Cawley/Getty Images

Introducción

La economía global está viviendo una revolución tecnológica. La explosión de datos es resultado de la traducción de gran parte de nuestras vidas a medios digitales. Las tecnologías que antes eran de uso exclusivo de los gobiernos y los servicios militares se han democratizado; aplicaciones para monitorear tráfico, pronósticos meteorológicos y servicios de transporte son ejemplos del uso cotidiano.

Es probable que esto tenga un efecto relevante en la generación de información geográfica y es-

taadística. Cada día se utilizan más los datos masivos generados en medios digitales para examinar de manera rigurosa la producción de indicadores, como cambios en salarios, productividad, vegetación, etcétera. Con fines de política pública, la información obtenida de dichas fuentes es crucial para la planeación y el seguimiento. Para atender fenómenos naturales, monitorear el crecimiento económico y planear ciudades inteligentes hay, hoy, un sinfín de aplicaciones tecnológicas.

Las oficinas nacionales de estadística se enfrentan a nuevos retos debido al desarrollo de tecnologías que han desatado un *boom* en la generación,

manejo y análisis de datos. Técnicas tradicionales de generación de estadística, como los censos y las encuestas, se enfrentan a los siguientes desafíos:

- Menor disposición de los informantes para proveer datos debido a desconfianza en las instituciones.
- Mayores costos de levantamiento de datos en campo. Por los altos índices de inseguridad, esta labor se ha vuelto muy costosa, comparada con el trabajo de gabinete que permite la era digital, aunada a un menor costo de almacenamiento y procesamiento.
- Los usuarios se han vuelto más exigentes requiriendo información con una mayor temporalidad y, en algunos casos, en tiempo real; por ejemplo, para algunos ya no basta tener información económica anual, buscan tener indicadores que les permitan seguir el pulso de la economía día con día.

Debido a lo anterior, existe una necesidad imperante para desarrollar metodologías que nos permitan generar información de calidad a partir de fuentes no tradicionales, que son aquellas que nos proporcionan datos provenientes de fuentes digitales. Las aplicaciones de los celulares, las facturas electrónicas, los satélites que orbitan alrededor de la Tierra, entre otros, están creando volúmenes de datos cada vez más significativos con cobertura global. La falta de estos ya no es un factor limitante.

Este artículo presenta el *Cubo de Datos Geoespaciales de México*, una herramienta que permite analizar y almacenar las imágenes satelitales de una manera más eficiente, buscando aprovechar su potencial para obtener información.

Imágenes satelitales para la generación de información geográfica y estadística

Este tipo de productos proporciona un punto de vista único para recopilar información esencial para evaluar los peligros ambientales, gestionar los recursos naturales y mejorar nuestra compren-

sión del planeta. Desde que los satélites llevaron cámaras a órbita, se tiene una perspectiva inigualable, que se ha ido enriqueciendo a medida que los sensores avanzan en términos de variedad y sofisticación.

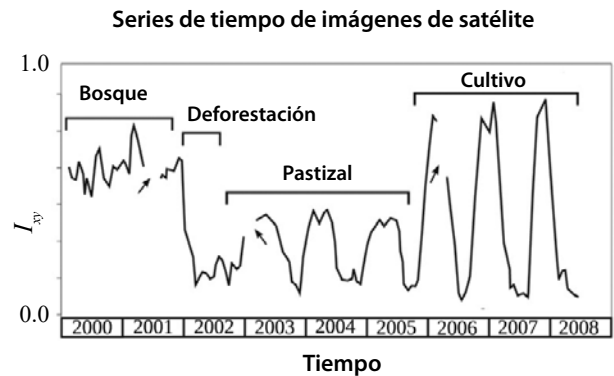
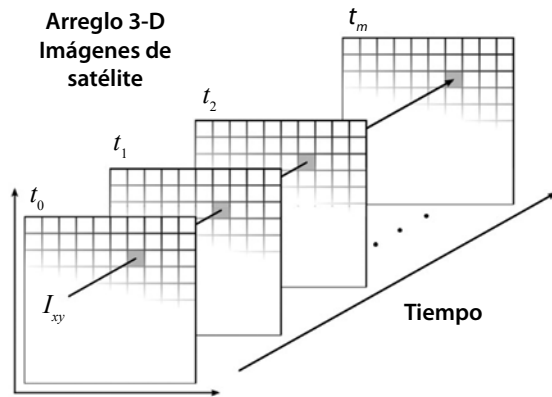
En la década de los 70 del siglo pasado, el Servicio Geológico de los Estados Unidos de América (USGS, por sus siglas en inglés) comenzó a tomar una imagen de la Tierra a través de sus satélites *Landsat* cada dos semanas. Si se reúnen los datos que captaron de México solo durante el 2015 y los almacenáramos en disquetes, se podrían apilar uno sobre otro hasta formar una torre de 1 630 metros de altura, casi la que tiene el Cerro de la Silla en Monterrey, y aprovecharlos para un análisis requeriría de cerca de ocho *laptops* de uso común, equivalentes a 7.5 *terabytes* (TB).

Conforme la tecnología evolucionó, la frecuencia de la generación de las imágenes satelitales también fue aumentando. Con un programa de más reciente lanzamiento (2015), la Agencia Espacial Europea genera una imagen del planeta cada cinco días con los satélites *Sentinel*, lo cual equivale a más de 70 imágenes de la Tierra cada año. Otro ejemplo es la empresa Planet, que cuenta con una constelación de más de 130 nanosatélites y toman una imagen de cada punto de la Tierra al día.

Información espectral, mediciones periódicas, datos abiertos

Los sensores de los satélites juegan, también, un papel en la relevancia de las imágenes satelitales como fuente de información. Así como una fotografía capta la porción del espectro de luz que es *humanamente visible* (rojo, verde y azul), el sensor de un satélite tiene la capacidad de captar, además de la luz, información infrarroja, que es imperceptible para el ojo humano. Uno de los usos comunes de las imágenes satelitales es aprovechar los datos del espectro infrarrojo para conocer información sobre la vegetación, ya que las plantas reflejan la luz en este espectro en proporción mucho mayor que en el verde (ver imagen 1).

Medición del Índice de Vegetación Mejorado (EVI) a lo largo del tiempo del pixel x, y



Fuente: Picoli *et al.*, 2018.

De manera adicional, los satélites brindan estas mediciones a intervalos regulares, por lo que representan una oportunidad única para construir series de tiempo consistentes y derivar este análisis en tendencias de comportamiento. Gracias a su periodicidad, sabemos no solo que el nivel del mar en realidad está subiendo, que las grandes plataformas de hielo se siguen rompiendo o que las selvas tropicales están desapareciendo, sino con qué rapidez están sucediendo.

El gran desarrollo en la tecnología empleada en las misiones espaciales y en los sensores ha revolucionado la industria geoespacial; ahora existen empresas con satélites del tamaño de una caja de zapatos, con sofisticados sensores. Estos cambios y avances retan a generar nuevos productos y mecanismos para el propio análisis de las imágenes.

Sin embargo, la ventaja clave que impulsa el uso a nivel global de las imágenes satelitales provenientes de los programas *Landsat* del USGS y *Sentinel-2* de la Agencia Espacial Europea dentro de las oficinas que generan información oficial es el hecho de que estas imágenes son *datos abiertos*,

lo cual significa que están disponibles en repositorios de acceso libre y gratuito para cualquier agente interesado.

Estos dos programas tienen diferencias fundamentales, como su antigüedad: *Landsat* inició su captación en la década de los 70, mientras que *Sentinel-2* comenzó en fecha mucho más reciente: 2015; otro diferenciador es el volumen de sus datos, a pesar de la posible intuición dada por la diferencia de antigüedad de los programas; gracias a la mayor frecuencia y capacidad de sensores, *Sentinel-2* ha generado un volumen que supera ya el de toda la historia de *Landsat*. Debido a estas características, ambos son fuentes de datos abiertos que se consideran complementarios entre sí.

Uso actual de las imágenes satelitales en el INEGI para generar información

El Instituto Nacional de Estadística y Geográfica (INEGI) tiene la tarea de monitorear fenómenos ambientales, socioeconómicos y demográficos

que se presentan en todo el país. En particular, el artículo 26 de la *Ley del Sistema Nacional de Información Estadística y Geográfica* indica que: "... El Subsistema Nacional de Información Geográfica y del Medio Ambiente, en su componente geográfico, generará como mínimo los siguientes grupos de datos: marco de referencia geodésico; límites costeros, internacionales, de las entidades federativas, municipales y las demarcaciones territoriales de la Ciudad de México; datos de relieve continental, insular y submarino; datos catastrales, topográficos, de recursos naturales y clima, así como nombres geográficos. A este componente también se le denominará Infraestructura de Datos Espaciales de México..."

Siendo la agencia cartográfica nacional de México, el INEGI ha producido mapas del país durante varias décadas. De inicio, en formato impreso (décadas de los 70 y 80) y ahora, en forma de conjuntos de datos geoespaciales temáticos. Con anterioridad, estos mapas se derivaban de fotografías aéreas, pero en la actualidad, las cartas y datos geoespaciales se obtienen a partir de imágenes satelitales.

No obstante, la mayoría de estos procesos involucran solo el análisis visual. Las imágenes satelitales siguen siendo subutilizadas debido a varias razones, como la complejidad del manejo de los grandes (y cada vez más crecientes) volúmenes de datos altamente estructurados, así como la falta de experiencia e infraestructura para descargar, almacenar y procesarlos eficiente y eficazmente.

Panorama internacional

Hace tiempo, las imágenes satelitales solían ser distribuidas en cintas magnéticas. En el 2011, la agencia Geoscience Australia (que custodia los datos geográficos y geológicos en esa nación) transfirió los datos de su territorio de estas cintas a un almacenamiento más moderno: discos mecánicos. Su estrategia fue clave en la iniciativa para *desbloquear el archivo Landsat* de ese país. No tardó en desarrollarse una herramienta diseñada en especí-

fico para facilitar el acceso a estos datos: el *Cubo de Datos Geoespaciales Australiano*.

Para el 2017, esta herramienta tomó un nuevo nombre: *Open Data Cube (ODC)*,¹ tras una serie de evoluciones, como la compatibilidad con distintos sistemas de referencia de coordenadas y formatos de archivos. Su objetivo es permitir enfrentar los retos que se presentan al manejar grandes volúmenes de datos como lo son las imágenes satelitales.

ODC es un proyecto de código abierto que resuelve la necesidad de un mejor manejo de los datos satelitales. La plataforma permite acceder y analizar décadas de imágenes; además, simplifica su uso para observar las condiciones espaciales que captaron, lo cual permitirá que gobiernos, empresas e individuos tomen decisiones más informadas en temas relacionados con el suelo, la vegetación, la erosión costera, la agricultura, la deforestación, los cuerpos de agua o los asentamientos humanos.

A pesar de que el propósito inicial del cubo de datos era aprovechar la riqueza temporal de las imágenes *Landsat*, la flexibilidad de la plataforma permite otra información en el análisis, como: modelos digitales de elevación o cuadrículas de datos geofísicos. Cualquier insumo disponible para el usuario puede ser utilizado en el cubo para enriquecer los resultados, sean datos comerciales, *in situ*, o bien, derivados de un análisis previo.

Varias agencias geoespaciales reconocen que es necesario distanciarse de procesos tradicionales, así como disminuir las barreras causadas por el volumen de los datos y las complejidades relacionadas con la preparación, manejo, almacenamiento y análisis de esta información, para proveer, así, a los usuarios de información para el desarrollo nacional de manera oportuna y, también, mejorar el nivel de detalle en los mapas producidos.

¹ Solución de código abierto para el acceso, manejo y análisis de grandes cantidades de datos de sistemas de información geográfica, sobre todo de observaciones de la Tierra.

Además de Australia, la comunidad de producción de información geoespacial ya se ve favorecida en varias naciones mediante el uso del cubo de datos. Suiza y Colombia son ejemplos de países que ya cuentan con un cubo operacional. Hay, también, ejemplos regionales: en febrero del 2019, la fundación benéfica Leona M. and Harry B. Helmsley destinó 10 millones de dólares a la implementación regional del *Cubo de Datos de África (Digital Earth Africa)*, mientras que el gobierno australiano asignó 10 millones de dólares australianos (alrededor de 7 millones de dólares estadounidenses), además de apoyo técnico y operacional para coordinar esta iniciativa regional.

El INEGI abordó, a su vez, los desafíos mediante la implementación del *Cubo de Datos Geoespaciales de México*. A partir del 2018 se ha invertido en planeación y desarrollo para las ca-

pacidades técnicas y de infraestructura necesaria para atender los retos de *Big Data* propios de esta tarea. A diferencia de los cubos de datos de Suiza o Colombia (ambos funcionan con infraestructura en la nube), el de nuestro país se encuentra en las instalaciones del INEGI, a pesar de que, en términos de volumen, es el mayor de estas tres versiones, derivado de la extensión del territorio mexicano.

Cubo de datos geoespaciales

Hoy en día, la mayoría de las personas elige ponerse un abrigo o no, cargar con un paraguas o no sin la necesidad de conocer conceptos meteorológicos avanzados, ya que existen varias opciones disponibles de herramientas (integradas en cualquier teléfono celular) que indican las condicio-

Imagen 2

Panorama global del cubo de datos geoespaciales



Solo 18 meses después de comenzar la iniciativa *Open Data Cube*, había 43 países con cubos de datos operativos en desarrollo o que habían expresado interés en la iniciativa. El interés global es muy diverso y está creciendo muy rápido.

Fuente: Killough, Brian. Conferencia. 2018 (DE) https://www.researchgate.net/figure/18-months-after-starting-the-Open-Data-Cube-initiative-there-are-43-countries-with_fig1_328995502

nes climáticas cada día y, así, permiten tomar decisiones oportunas y evitar contraer un resfriado. De la misma manera, un cubo de datos pretende acercar la información al usuario final en un formato listo para su uso (ver imagen 3).

Esta herramienta es, de manera significativa, más fácil de usar, eficiente y escalable que otros paradigmas de datos; está diseñada para contar con los servicios necesarios para acercar al usuario a la explotación de la información proveniente de las imágenes satelitales. Además, está alineada a los principios de *Big Data*, que ofrecen la posibilidad del procesamiento masivo, escalable y distribuido permitiendo, también, aprovechar técnicas de inteligencia artificial.

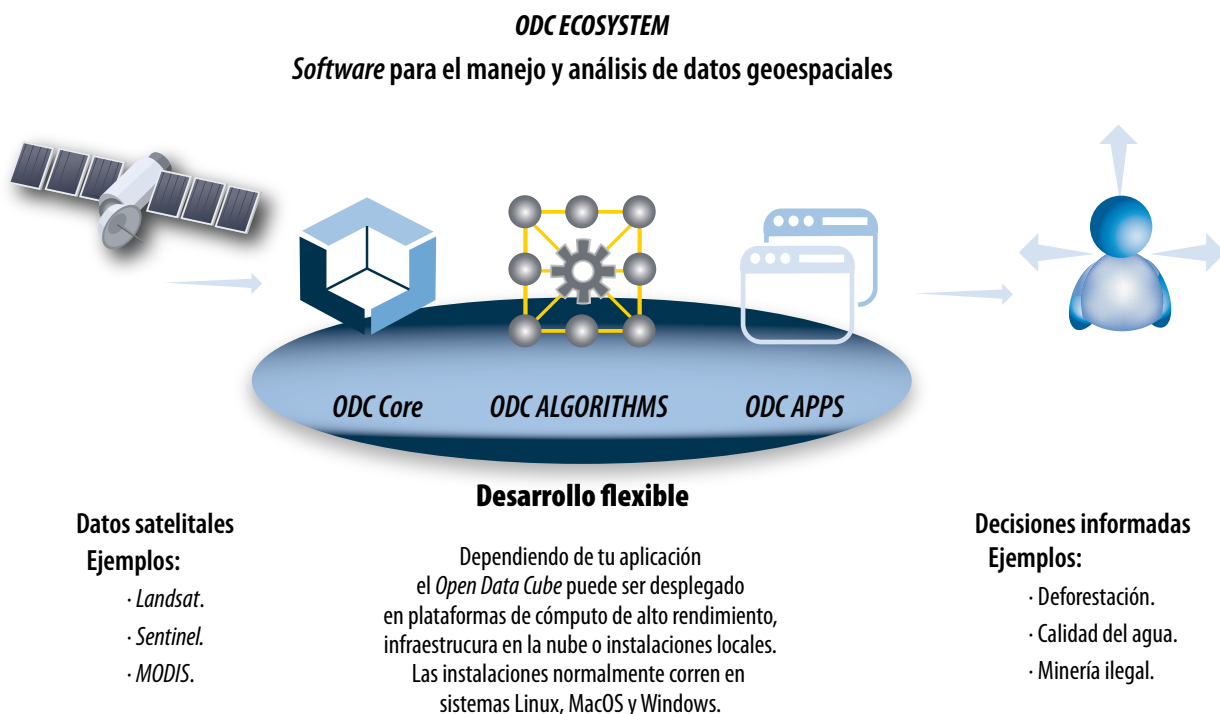
Arreglo masivo de datos *raster* listos para el análisis

Un cubo de datos es una solución para gestionar ordenada y eficientemente datos espacio-tem-

porales para su análisis; llevado al ámbito práctico, se define como un arreglo masivo de datos *raster*; estos, como las imágenes satelitales, son representaciones espaciales de información en formato matricial: cada celda de esta matriz (también llamada pixel) se puede orientar mediante coordenadas y tiene un valor asignado. Pueden existir varios canales en el mismo pixel de una imagen satelital; esto significa que, así como una foto tomada por la cámara de un teléfono celular suele generarse en el formato común —conocido como RGB, que tiene tres canales (R: rojo, G: verde y B: azul)—, las imágenes satelitales tienen, por lo regular, más canales gracias a sus complejos sensores; en el cubo de datos tienen seis.

Los datos *raster* de un cubo de datos se almacenarán en una sola estructura multidimensional, descrita por varios ejes; las coordenadas en estos son el mecanismo que permite acceder a los datos de forma inequívoca según el espacio, tiempo o canal deseado. Quizá la característica medular de un cubo es que, al integrar la información, este sis-

Imagen 3



Fuente: OpenDataCube.org

tema de coordenadas se homologa para todos los puntos y la ubicación de los pixeles se vuelve independiente de la imagen que lo generó. Así es como facilita el acceso, de entre miles de imágenes de una región o país a través del tiempo, a los datos del área y tiempo especificados por el usuario para su uso inmediato sin verse, además, obstaculizado por su volumen y complejidad.

Si bien la matriz es multidimensional, se le llama cubo de datos por la tridimensionalidad que evoca el nombre de un volumen, ya que, de forma tradicional, para acceder a datos espaciales se usan coordenadas (x, y) , y al apilar imágenes de la misma zona captadas en distintas fechas se puede observar un punto a lo largo de un periodo, lo cual añade una tercera dimensión, que es el tiempo (t) (ver imagen 4).

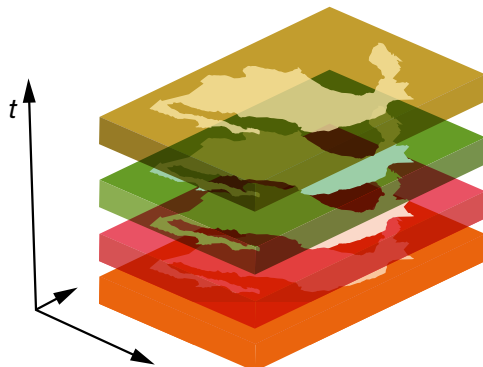
Datos *raster* listos para el análisis

La arquitectura ordenada (descrita en la sección anterior) favorece un análisis eficiente. Por otro lado, la previa generación de datos estandarizados también es crucial durante la fase de construcción de un cubo.

Las imágenes satelitales ideales para esta herramienta requieren ser preprocesadas hasta lograr un alto nivel de calidad. Estos datos se conocen como *listos para el análisis* (*Analysis Ready Data* o *ARD*) (ver imagen 5). Este término implica que el valor de cada pixel representa el mismo espacio y la misma respuesta espectral independientemente del sensor, el lugar y la época en que fueron captadas, estandarizándolas así para su uso estadístico y analítico a través de series de

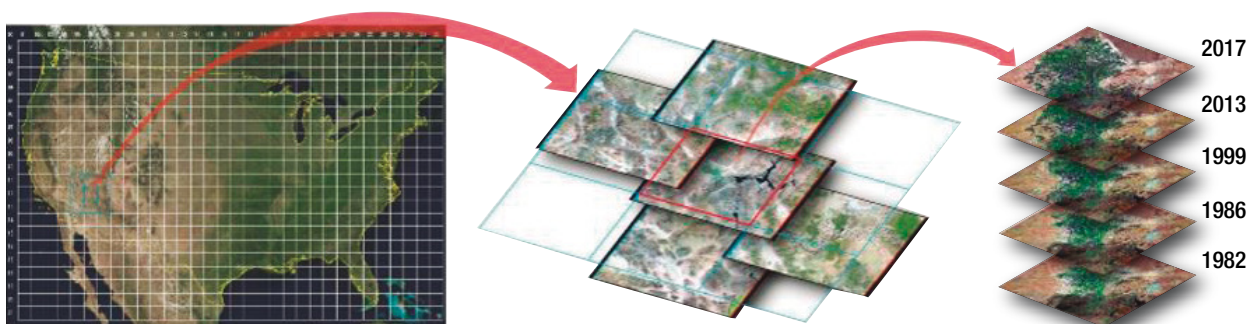
Imagen 4

Arreglo masivo de datos multidimensionales (*raster*)



Fuente: INEGI.

Imagen 5



Fuente: U.S. Landsat Analysis Ready Data (DE) <https://www.usgs.gov/>

tiempo. Además, deben cumplir con descripción de metadatos y calibraciones radiométrica, geométrica, solar y atmosférica.

Generar estas imágenes requiere de un esfuerzo computacional adicional aplicado a las comúnmente disponibles en los repositorios (datos crudos). Aunque por lo general es posible solicitar apoyo de forma directa a las agencias correspondientes, el acceso al material con estas características se irá incrementando. En la actualidad, agencias y proveedores trabajan para su pronta disponibilidad en la nube; por ejemplo, se espera que las imágenes *Landsat ARD* sean de fácil acceso de esta forma en cuestión de meses.

Podría compararse, entonces, un cubo con un video: varios cuadros (*frames*) con datos ordenados de manera temporal; para el video serían tal vez datos RGB, mientras que para una imagen satelital se agregarían las otras bandas no visibles, como el infrarrojo. Siendo aún más específicos, este sería un video en cámara rápida (*Time-Lapse*) y las correcciones de parpadeo o movimiento de la cámara (*flicker, motion control*) equivaldrían al procesamiento que genera los datos *ARD* ordenados, ya que alinean los cuadros de forma correcta y remueven efectos indeseados por variaciones en la luz.

Las imágenes *ARD* en un cubo de datos serán transformadas a una proyección geográfica definida y, así, esta organización permitirá su análisis inmediato sin que sea necesaria ninguna acción adicional por parte del usuario. De esta forma, la adopción del cubo impulsará los procesos involucrados en la explotación de imágenes satelitales en el INEGI: ahora, los datos se podrán almacenar, administrar, acceder y analizar de manera más eficiente.

Cubo de Datos Geoespaciales de México

El 4 de marzo de 2019, el INEGI recibió el archivo histórico *Landsat ARD* con más de 109 mil imágenes, que fueron generadas a partir de su versión original y luego enviadas por el personal de USGS/NASA; su volumen en formato comprimido ascien-

de a 30 TB. Este archivo ya está cargado y respaldado en la infraestructura del Instituto y su volumen es de 90 TB cuando se descomprime. Descargarlo y procesarlo —para un país del tamaño de México— desde un repositorio en línea resultaría computacionalmente demandante y poco eficiente.

Otro paso hacia la implementación del *Cubo de Datos Geoespaciales de México* fue la definición de una proyección adecuada para el archivo *Landsat ARD*. El INEGI favoreció la proyección cónica equiárea de Albers que, como lo implica el nombre, proyecta las áreas de manera proporcional. Esto facilitará las estadísticas de superficie, que es un enfoque común para aquellos indicadores de los Objetivos de Desarrollo Sostenible (ODS) —de la *Agenda 2030*— basados en información geoespacial.

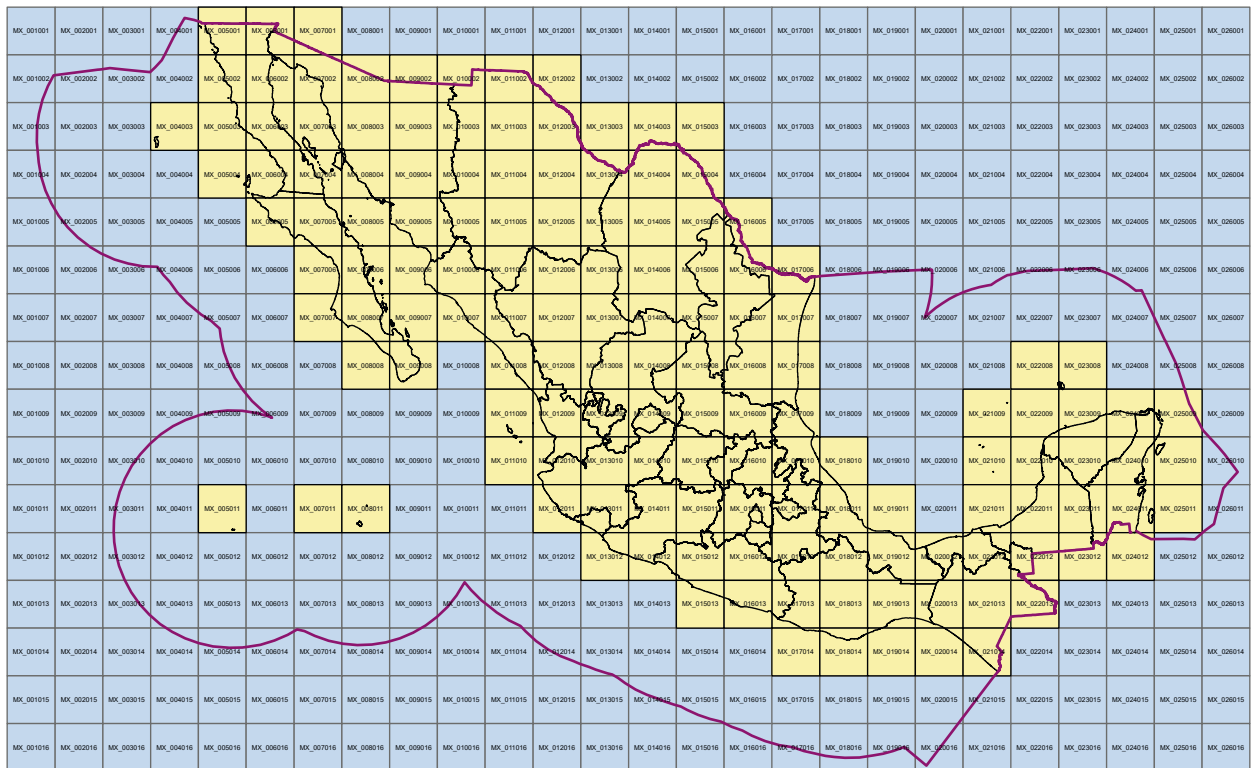
Sin embargo, una matriz de datos con toda esa información resultaría en un archivo difícil de manejar de manera tradicional, en especial si se desea realizar un análisis a nivel nacional. Considerando píxeles de 30 m por lado, provenientes de imágenes *Landsat*, sería como editar un video con resolución de alrededor de 105 mil píxeles de ancho y 70 mil de alto: el formato de ultra alta definición 4K tiene apenas 3 840 por 2 160.

Para facilitar esta y otras tareas computacionalmente complejas, la proyección de las imágenes se subdivide en mosaicos de 5 mil píxeles por lado. Este nuevo nivel de abstracción lógica en la arquitectura de los datos los distribuirá en archivos idénticos en su estructura con la finalidad de *paralelizar* los procesamientos computacionales altamente demandantes (ver imagen 6). El diseño de estos mosaicos de 5 mil píxeles por lado equivale a 150 km, lo que garantiza, además, compatibilidad para otras resoluciones como píxeles de 10 m por lado, en el caso de datos generados por el sensor de *Sentinel-2*.

Ventajas del Cubo de Datos Geoespaciales de México

Las metodologías tradicionales, en las que se incluye el uso y análisis de imágenes satelitales, conllevan

Imagen 6



Fuente: INEGI.

van un manejo tradicional y gran dependencia en la fotointerpretación (identificación de rasgos mediante análisis visual). Las dos grandes ventajas del cubo mexicano son:

- Aprovechamiento de recursos tecnológicos. Utilizando una instancia de un cubo de datos, en lugar de que cada imagen se descargue, almacene y analice en las computadoras de los especialistas, se pueden aprovechar los recursos tecnológicos y los tiempos destinados a los preparativos de los procesos. Con el análisis de datos efectuado mediante el archivo previamente generado, con imágenes alineadas, ordenadas y estandarizadas, se permitirá automatizar algunas fases de las metodologías, como la descarga y distribución de las imágenes a los especialistas. Por otro lado, la precisión espacial es mucho mayor.

Asimismo, al permitir procesamientos masivos, esta tecnología facilitará aplicar algoritmos de inteligencia artificial a grandes volúmenes de datos. Utilizando este atributo, la inspección visual también podría ser asistida por clasificaciones supervisadas a nivel nacional efectuadas de manera previa. De esta forma, se puede dirigir la atención de los pocos especialistas que deben analizar toda la extensión de un país a la inspección de regiones con características demasiado confusas para los algoritmos y, así, reducir los considerables tiempos que demandan la exhaustividad de la inspección y análisis visual.

- Aprovechamiento de los datos. Las imágenes con nubosidad no serán descartadas en su totalidad, como es común, ya que ahora se pueden eliminar solo los pixeles afectados por las nubes; por ejemplo, para construir imágenes

continuas y libres de ellas para el monitoreo del crecimiento urbano (ver imagen 7).

Además de las posibles mejoras a los procesos actuales, otra ventaja de contar con un arreglo estructurado de imágenes alineadas correctamente, como es el cubo de datos, es un mayor aprovechamiento de las imágenes satelitales y la libertad de selección de las áreas y periodos específicos de estudio que mejorarán, por ejemplo, el monitoreo de cambios en los recursos naturales a lo largo de décadas. Para ilustrar esta idea, en la imagen 8 se aprecia el cambio en la vegetación a lo largo de casi 30 años en una región (el rojo representa vegetación abundante). Estas posibilidades pueden permitir, sin duda, ampliar el catálogo de productos de información geoespacial.

Aunado a lo anterior, el *Cubo de Datos* está desarrollado de inicio a fin en el lenguaje de programación *Python*, el cual tiene una curva de aprendizaje

muy suave y, por lo mismo, es de gran popularidad en las comunidades de ciencia de datos y de procesamiento de imagen. Tras la construcción de un cubo de datos, el usuario final (que tal vez sea un especialista temático) solo debería preocuparse por dominar este lenguaje de programación de forma general y luego, las instrucciones básicas propias de la biblioteca (*set* de comandos) del cubo de datos para el manejo de la información. Esta y muchas otras bibliotecas prediseñadas se encuentran disponibles de forma pública y gratuita (OpenDataCube.org).

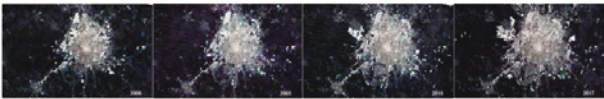
Análisis temporal a nivel pixel

El *Cubo de Datos* cuenta con una serie de algoritmos (biblioteca *Data Cube Stats*) que generan varios análisis directos, los cuales han sido aprovechados en varios temas.

Los productos se producen para una zona de interés en un periodo deseado mediante recortes o rebanadas de la matriz utilizando las coordenadas correspondientes. Se basan en el análisis de series de tiempo; las operaciones usan la serie independiente y propia de cada pixel en la zona seleccionada; en consecuencia, todos sus resultados se presentan, también, a nivel pixel.

Imagen 7

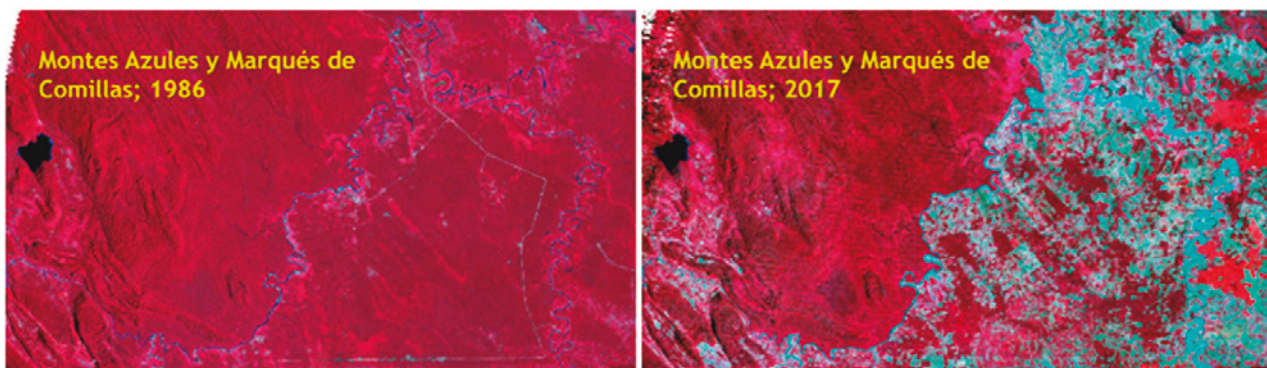
Secuencia de imágenes resumen en Mérida, Yuc., generadas con los datos *Landsat* del 2000, 2005, 2010 y 2017 mediante el *Cubo de Datos*



Fuente: INEGI.

Imagen 8

GEO-XV Plenary Session 2: EO in Support of the SDG, Kyoto, Japan, October, 2018



Fuente: INEGI.

Esta es una importante contribución que complementa e impulsa el uso del *Cubo*, ya que hace disponibles, de manera inmediata, herramientas de uso común en imágenes satelitales para generar estadística, como los índices de diferencia normalizada, por ejemplo, el Índice de Áreas Construidas de Diferencia Normalizada (NDBI, por sus siglas en inglés), que es una frecuente vía de análisis territorial en estudios urbanísticos e infraestructuras y en la comparación de la evolución del desarrollo urbano en el tiempo, ya que resalta las zonas con superficies edificadas o en desarrollo de construcción frente a las habituales áreas con vegetación o desnudas.

Quizá, una de las técnicas más comunes en el análisis de imágenes satelitales es el Índice de Vegetación de Diferencia Normalizada (NDVI, por sus siglas en inglés), que resalta la flora del resto de los elementos presentes en una imagen.

Estas versiones de índices de diferencia normalizada generadas con el *Cubo de Datos* calculan, además, los valores estadísticos (máximo, mínimo y media), lo cual permite aprovechar la información aún más, ya que estos describen el comportamiento del índice a lo largo del tiempo (en este caso, un año), lo que proporciona una mejor idea de la respuesta (ver imagen 9); por ejemplo, en el caso del NDVI, conocer el comportamiento de la vegetación durante un año permite vincular variables meteorológicas a las manifestaciones

temporales periódicas o estacionales de las plantas, lo cual favorece una mejor clasificación para cada pixel, a diferencia de lo que se podría concluir considerando solo el valor observado en una fecha determinada.

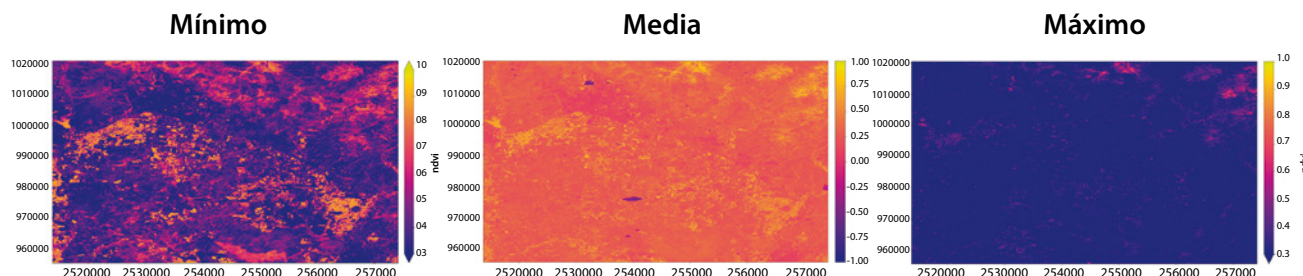
Geomediana y observaciones de agua desde el espacio

A esta colección de productos basados en el análisis a nivel pixel se agregan estos dos importantes desarrollos diseñados para el mayor aprovechamiento del *Cubo de Datos*.

La geomediana es un algoritmo utilizado para realizar un resumen estadístico multivariado para todos los valores observados de un mismo pixel en un periodo indicado. Aplicándolo a una región, se produce una imagen compuesta de *pixeles resumen*, la cual mantiene consistencia espacial, incluso en los límites entre escenas; típicamente, se obtienen imágenes continuas y libres de nubes (ya que se descartan las observaciones de pixeles donde hay nubosidad). El algoritmo trabaja con todas las bandas del pixel a la vez por lo que, además, se conserva la razón entre estos valores; debido a ello, es posible utilizar el producto de estas operaciones como el insumo de otros procesos de uso frecuente que parten de estas relaciones entre las bandas, como los presentados en la sección anterior (NDBI y NDVI).

Imagen 9

NVDI en Guanajuato, 2015



Fuente: INEGI.

La geomédiana nacional 2015 fusiona las 138 captaciones realizadas por los satélites *Landsat* durante todo ese año, que cubren el territorio continental e islas de México; cada una representa regiones de 182 km por 185 km, aproximadamente, y es captada cada ocho días. En otras palabras, esta imagen continua y libre de nubes resume los 7.5 TB de datos que son 6 074 imágenes *Landsat*, sin perder nivel de detalle espacial (ver imagen 10).

Por otra parte, una aplicación más del *Cubo* es un algoritmo para detectar y monitorear la presencia de agua mediante una clasificación supervisada (*agua o no agua*). Este se llama *Observaciones de agua desde el espacio o Water Observations from Space (WOfS)* y calcula el porcentaje de presencia de este líquido en el periodo seleccionado (ver imagen 11).

Integración de otras fuentes para realizar clasificaciones

Además de simplificar el análisis de un gran conjunto de imágenes *Landsat*, los cubos de datos geospaciales pueden complementar metodologías que implican el uso de otras fuentes de información, ya sea tipo *raster* (como un modelo digital de elevación) o vectorial (como los datos de campo).

Un ejemplo de esto es el ejercicio realizado para la identificación de cultivos y cálculo de su superficie utilizando técnicas de aprendizaje automático. En una zona del estado de Guanajuato, previamente definida, se aplicó el clasificador supervisado *Random Forest*. El cultivo de interés para este ejercicio fue el de sorgo, principal producto de la zona estudiada.

Imagen 10

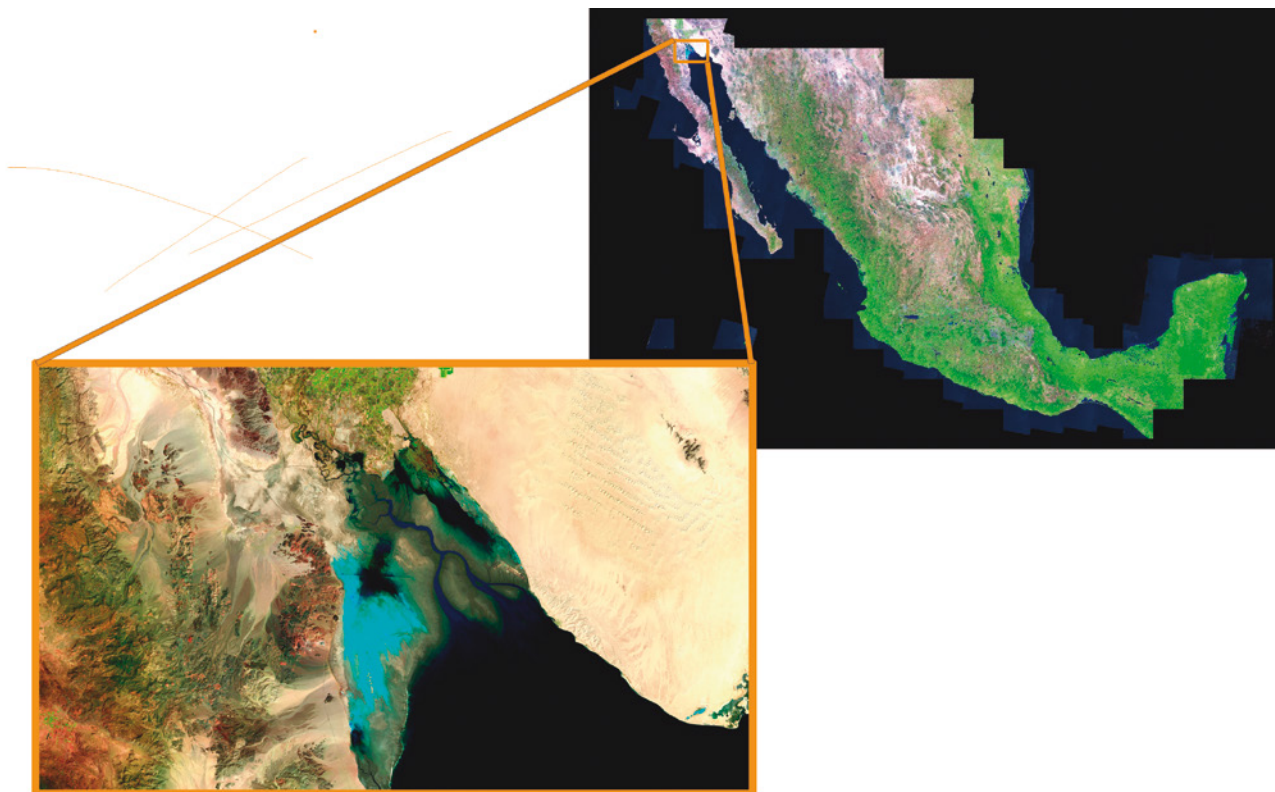


Imagen continua y libre de nubes del territorio mexicano generada con 6 074 imágenes *Landsat* del 2015 mediante el algoritmo mediana geométrica y acercamiento al delta del río Colorado, mar de Cortés.

Fuente: INEGI.

El insumo obtenido de manera directa del *Cubo de Datos Geoespaciales de México* fueron imágenes *Landsat* correspondientes a junio, julio y agosto del 2015; el periodo se eligió a partir del ciclo de cultivo del sorgo, pues en junio la planta alcanza a cubrir el suelo y en agosto se encuentra en su etapa de maduración.

De manera adicional, se utilizó información de campo correspondiente a los terrenos con cultivos en la zona de estudio durante el mismo año. Debido a la resolución espacial de *Landsat* de 30 m por

pixel, se consideraron solo los terrenos de superficie mayor (ver imagen 12).

Los resultados fueron satisfactorios en los polígonos de mayores áreas. Sin embargo, se presentó confusión para los más pequeños; por este motivo, las pruebas se están recreando en la actualidad utilizando imágenes de mayor resolución espacial y temporal (*Sentinel-2*).

También, se realizó el ejercicio usando una imagen satelital de alta resolución (6.5 m) y se com-

Imagen 11

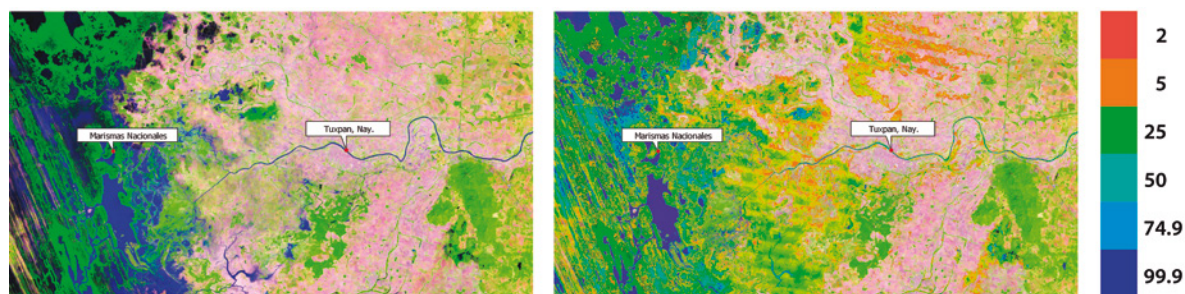
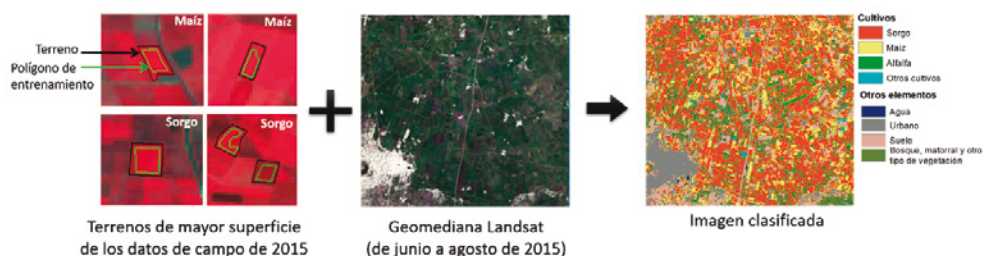


Imagen simple y sobreposición de una representación en color del porcentaje de agua presente durante el 2015 en las regiones de Marismas Nacionales y Tuxpan, Nayarit, mediante el algoritmo *WofS*.

Fuente: INEGI.

Imagen 12

Proceso para realizar clasificación supervisada de cultivos utilizando datos de campo y *Landsat* de junio a agosto del 2015 para entrenamiento del algoritmo *Random Forest*



Fuente: INEGI.

pararon las áreas resultantes de los cultivos clasificados de un municipio con cobertura completa en ambas imágenes (Valle de Santiago). Las cifras resultantes son consistentes entre sí; el cuadro de la imagen 13 muestra los resultados para superficie cultivada del cultivo principal del estudio (sorgo) y de maíz, según el insumo utilizado en el ejercicio.

Conclusiones

Con el *Cubo de Datos Geospaciales de México* hay un cambio de paradigma en cómo el INEGI entrega información a los usuarios. Históricamente, se han presentado productos finalizados, con tabulados básicos y representatividad nacional, estatal y municipal. En este caso, se pondrían a disposición del usuario las imágenes y los algoritmos con los cuales puede realizar sus propias estimaciones.

Con cada año que pasa, los nuevos satélites están creando volúmenes de datos cada vez más significativos con cobertura global. La falta de información ya no es un factor limitante, pero

almacenar y procesar esta cantidad de datos representa un nuevo reto para el desarrollo de capacidades. No obstante, hay confianza en que el enfoque basado en generar la infraestructura necesaria para el uso de los datos de observaciones de la Tierra ofrece un gran potencial para reducir las barreras técnicas para explotarlos en todo su potencial. Con una gran y creciente comunidad internacional de usuarios y desarrolladores, las aplicaciones del *Cubo de Datos* son virtualmente infinitas.

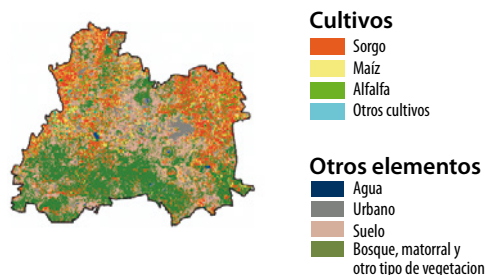
La herramienta ofrece excelentes soluciones para los problemas principales de los usuarios, como acceso, preprocesamiento y análisis eficientes para respaldar las aplicaciones de los usuarios, además de que permite minimizar el tiempo y el conocimiento especializado requerido para acceder y preparar los datos satelitales. En consecuencia, con estos cambios en el INEGI se prevé un gran éxito e impacto a escalas local, regional y mundial, apoyando agendas locales y globales prioritarias, como las que se encuentran en los ODS de la Organización de las Naciones Unidas, el Marco de Sendai y el Acuerdo de París, entre otros.

Imagen 13

Comparación de área cultivada estimada de sorgo y maíz en el municipio de Valle de Santiago, 2015

Fuente y resolución	Área de sorgo (ha)	Área de maíz (ha)
Varias imágenes de satélite <i>Landsat</i> 2015 (30 m)	17 420	7 469
Una imagen de satélite <i>RapidEye</i> 2015 (6.5 m)	18 858	5 315

Clasificación de cultivos del municipio de Valle de Santiago



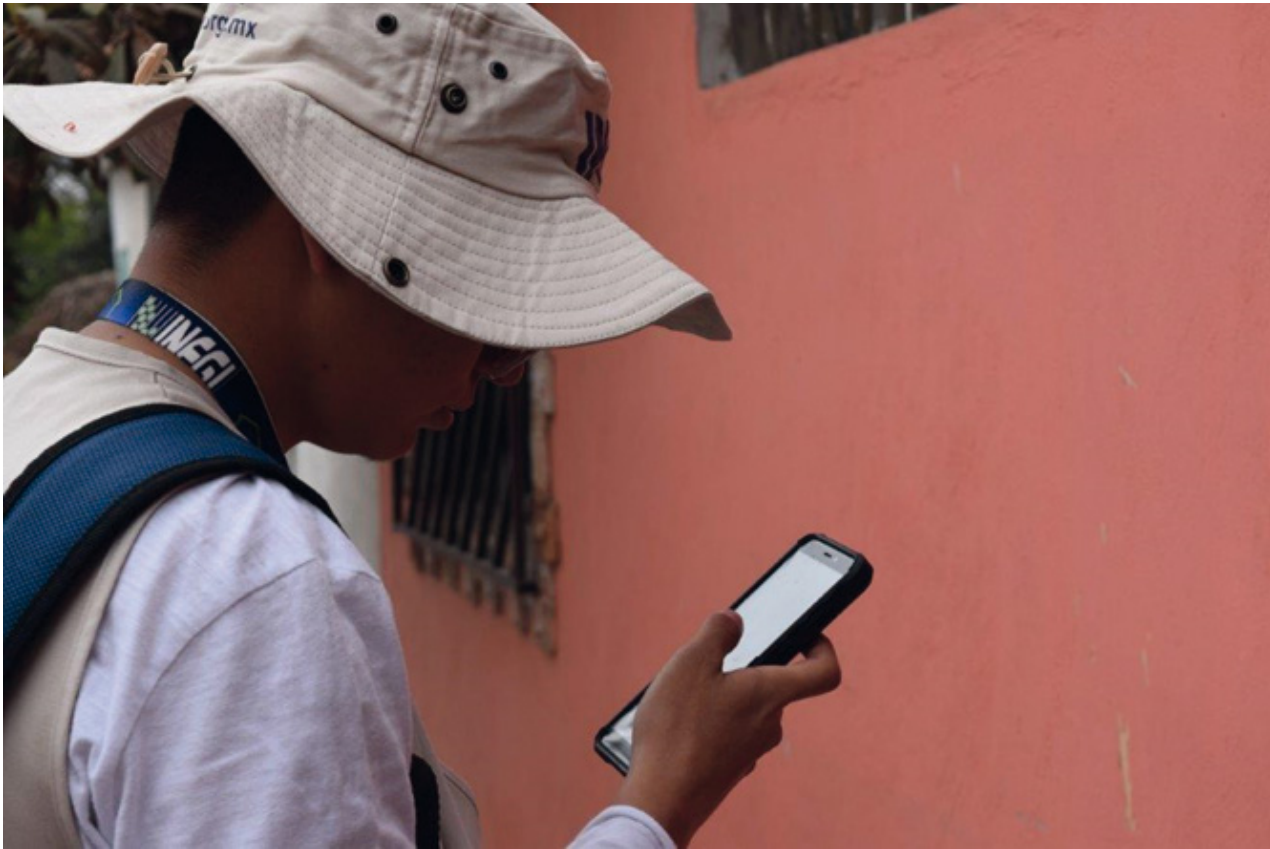
Fuente: INEGI.

Uno de los retos más relevantes al que nos enfrentamos los promotores del uso de la información geoespacial es lograr la necesidad de hacer que la dimensión geoespacial sea realmente transversal. Para ello, el INEGI cuenta con el beneficio de tener geografía y estadística en una sola institución nacional, lo que ha permitido a México una mejor integración y uso de sistemas de información complementarios. Con las herramientas asociadas de esta integración es posible georreferenciar estadísticas relevantes, así como determinar la ubicación de las desigualdades económicas y sociales, además de los riesgos y daños causados por desastres naturales. En este aspecto, el *Cubo de Datos* representa, también, una importante posibilidad para hacer eficiente la integración y el análisis de datos geográficos y estadísticos para un mejor diseño y monitoreo de las políticas públicas.

Fuentes

- Roberts, D.; N. Mueller & A. McIntyre. "High-Dimensional Pixel Composites From Earth Observation Time Series", in: *IEEE Transactions on Geoscience and Remote Sensing*. 99, 2017, 1-11.
- Killough, Brian. *Overview of the Open Data Cube Initiative*. 2018, 8629-8632. 10.1109/IGARSS.2018.8517694.

- Lewis, A. *et al.* "The Australian Geoscience Data Cube-Foundations and Lessons Learned", in: *Remote Sensing of Environment*. 2017.
- Piotrowicz, Luke. "The 40th Anniversary of Landsat, Australia's 33 year archive of Landsat data", in: *AusGeo News*. September 2012, Issue No. 107.
- Mueller, N.; A. Lewis, D. Roberts; S. Ring; R. Melrose; J. Sixsmith; L. Lymburner; A. McIntyre; P. Tan; S. Curnow & A. Ip. "Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia", in: *Remote Sensing of the Environment*. 174, 2016, 341-352.
- Picoli, Michelle *et al.* "Big earth observation time series analysis for monitoring Brazilian agriculture", in: *ISPRS Journal of Photogrammetry and Remote Sensing*. 2018. DOI: 10.1016/j.isprsjprs.2018.08.007.
- Roberts, Dale; Bex Dunn & Norman Mueller. *Open Data Cube Products Using High-Dimensional Statistics of Time Series*. 2018, 8647-8650. 10.1109/IGARSS.2018.8518312.
- Strobl, Peter; Peter Baumann; Adam Lewis; Zoltan Szantoi; Brian Killough; Matthew Purss; Max Craglia; Stefano Nativi; Alex Held & Trevor Dhu. "The Six Faces of the Data Cube", in: *Conference: Big Data from Space*. Toulouse, France, 2017.



importinegi:

un paquete de *R* para descargar y gestionar bases de datos del INEGI

importinegi:
an *R* Package to Download and
Manage INEGI's Data Sets

César Rentería*

* Centro de Investigación y Docencia Económicas, crenteria@albany.edu

Este artículo presenta *importinegi*, paquete que contribuye a facilitar la búsqueda, descarga y gestión de los datos abiertos del Instituto Nacional de Estadística y Geografía. Se describen todas las funciones y se acompañan con ejemplos. Fue desarrollado en plataforma abierta y busca apoyar al creciente ecosistema de ciencia abierta para la investigación.

Palabras clave: ciencia abierta; datos abiertos; bases de datos INEGI; *R*.

Recibido: 8 de agosto de 2019.

Aceptado: 3 de junio de 2020.

1. Introducción

El Instituto Nacional de Estadística y Geografía (INEGI) produce y provee un vasto conjunto de datos abiertos a partir de diversos proyectos. Aunque ha realizado un esfuerzo importante durante los últimos años para mejorar su disponibilidad y usabilidad, la integración de estos en ambiente de programación de *software* estadístico es todavía un área en desarrollo. Para ello, se cuenta con *R*, que es un *software* de licencia abierta que ha crecido considerablemente en popularidad y uso (especialmente en ciencias sociales).

Este artículo presenta el paquete de *R importinegi*, cuyo objetivo es facilitar la búsqueda, descarga y gestión de bases de datos abiertos del INEGI, en particular para su uso en la investigación y la enseñanza. Asimismo, al proveer funciones de código abierto —y en un *software* libre—, contribuye a fortalecer la adopción de mejores prácticas de *reproducibilidad* de la ciencia (Peng, 2011).

Facilitar el acceso y análisis de estos a través de paquetes estadísticos es un esfuerzo con muchos beneficios (Boettiger, Chamberlain *et al.*, 2015; Gandrud, 2016); por ejemplo: contribuye a expandir la *reproducibilidad* y transparencia de un artículo de investigación desde los datos en bruto hasta la publicación final; al simplificar, estandarizar y automatizar los procesos de gestión, también coadyuva en el logro eficiente del trabajo de investigación; asimismo, promueve la compartición

This article presents *importinegi*, a package that helps to facilitate the search, download and management of the *National Institute of Statistics and Geography's* open data. All functions are described and illustrated with examples. It was developed on an open platform and seeks to support the growing ecosystem of open science for research.

Key words: open science; open data; INEGI databases; *R*.

de código y la colaboración académica, minimiza errores humanos en el emparejamiento de múltiples bases de datos relacionadas y reduce problemas de acceso por cuestiones de compatibilidad o licenciamiento de formatos de archivo.

Diversos organismos a nivel internacional han desarrollado, ya sea institucionalmente o a través de sus usuarios, paquetes de *R* para facilitar el acceso a sus bases de datos y para promover los principios de *reproducibilidad* de la ciencia, por ejemplo: el Banco Mundial (Arel-Bundock, 2013), la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, por sus siglas en inglés) (Gheri & Kao, 2014), la oficina Eurostat de la Unión Europea (Lahti *et al.*, 2017) o el Buró del Censo en Estados Unidos de América (Walker, 2018).

Con *importinegi* se busca contribuir con la *reproducibilidad* de la ciencia y ampliar la base de usuarios de los datos abiertos del INEGI. Este trabajo provee una visión general de las características del paquete y describe cada una de las funciones de la versión 1.0.0.

2. El paquete *importinegi*

Se creó para facilitar la descarga y uso de las bases de datos públicas del INEGI. Entre otras ventajas, permite integrar múltiples de estas con un identificador único en común, consolidar los datos en diferentes niveles de agregación (por ejemplo, vi-

vienda, localidad, municipio o entidad federativa) y automatizar la codificación de valores perdidos en el ambiente de R.

La versión 1.0.0 de *importinegi* se publicó en el *Comprehensive R Archive Network (CRAN)* el 5 de agosto de 2019. Permite descargar y gestionar las bases de datos de los siguientes proyectos estadísticos del INEGI:

- Censo (y Conteo) de Población y Vivienda.
- Censo Nacional de Gobiernos Municipales y Delegacionales (CNGMD).¹
- Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH).
- Encuesta Nacional de Ocupación y Empleo (ENOE).

Asimismo, hace posible el acceso a las siguientes fuentes de información georreferenciada:

- Marco Geoestadístico Nacional (MGN).
- Red Nacional de Caminos (RNC).

La versión 1.0.0 *importinegi* puede descargarse desde el *CRAN* tecleando la función de instalación `install.packages()`. En el siguiente bloque de código se muestra cómo hacerlo:

```
install.packages("importinegi")
library(importinegi)
```

El paquete continúa en constante desarrollo, incorporando más datos abiertos para su descarga y refinando las funciones vigentes. La versión beta con las últimas actualizaciones se puede descargar directamente del repositorio de *Github* *crenteriam/importinegi*. En el siguiente bloque de código se muestra cómo instalarlo desde *Github*:

```
install.packages("devtools")
devtools::install_github(
  "crenteriam/importinegi")
library(importinegi)
```

¹ Ahora denominado Censo Nacional de Gobiernos Municipales y Demarcaciones Territoriales de la Ciudad de México.

La versión 1.0.0 cuenta con una extensa documentación de cada función y los valores, a la cual se puede acceder con la función de ayuda seguida del nombre de esta (por ejemplo, *importinegi*). También, es posible tener acceso a la viñeta del paquete mediante `browseVignettes("importinegi")`.

Cada función en el paquete corresponde a una fuente de datos distinta, por ejemplo, `enoe()` descarga y gestiona información de la ENOE.

En las siguientes secciones se describirá la sintaxis y uso de las funciones disponibles en la versión 1.0.0 de *importinegi*.

3. Sintaxis

El nombre de cada función representa las palabras clave o siglas del proyecto INEGI (como `censo_municipal()` o `enoe()`). En ocasiones, a esta nomenclatura le sigue una característica específica de este, por ejemplo: se puede acceder al Censo de Población y Vivienda a través de cinco formatos de microdatos, siendo ITER uno de ellos; por lo tanto, las funciones del Censo tienen una palabra clave adicional indicando el formato de microdatos: para llegar al de ITER es `censo_poblacion_iter()`.

En el caso de la información georreferenciada, el nombre de la función comienza con las siglas SIG (de sistema de información geográfica). Estas son sucedidas por una palabra clave del proyecto geoestadístico, por ejemplo, para acceder a los datos de la Red Nacional de Caminos es `sig_camino()`.

Existe una función especial, `catalogo_inegi()`, que permite explorar el catálogo de proyectos estadísticos del INEGI, así como la documentación completa de cada base de datos. La función accede a la Red Nacional de Metadatos (RNM), que es una plataforma de difusión y consulta de cada proyecto. Su sintaxis es sencilla. El único parámetro requerido es el número del proyecto (id) a consultar. En el siguiente bloque de código se ac-

cede (a través del navegador por *default* del sistema operativo) a la documentación del *Censo Nacional de Derechos Humanos Estatal 2018* mediante su id:

```
# Accede a la documentacion del
Censo Nacional de
# Derechos Humanos Estatal 2018
catalogo_inegi(id = 443)
```

Tecleando la función `catalogo_inegi()` (sin incluir nada en el paréntesis) se puede descargar y almacenar en un objeto de *R* una base de datos con el catálogo de proyectos del INEGI. Esta provee el id, su nombre, los años en que se comenzó y finalizó el levantamiento, la fecha de creación de los metadatos en la RNM y la fecha de última actualización de estos. El siguiente bloque de código ilustra la descarga de la base de datos del catálogo de proyectos y su almacenamiento en un objeto de *R*:

```
# Descarga y almacena el catálogo
de proyectos
# estadísticos del INEGI
catalogoINEGI = catalogo_inegi()
```

3.1 Censo (y Censo) de Población y Vivienda

Estos operativos proporcionan información sobre individuos, viviendas y características de las localidades en México. El INEGI ofrece varias formas de acceso a los microdatos de los censos y conteos. El paquete *importinegi* tiene una función por cada

una de estas modalidades de acceso. En el cuadro 1 se presentan los tipos de microdatos disponibles y su función correspondiente.

Todas las funciones para los cinco tipos de microdatos de los censos tienen la misma sintaxis. Aunque estas varían en los parámetros aplicables, todas comparten dos indispensables.

El primero es el año (*year*) en formato numérico. Los datos que se tienen comienzan a partir de 1990 (e incluyen los conteos), pero no todos los tipos de microdatos están presentes para todos los años censales. La documentación de cada función provee los años disponibles para cada tipo de microdatos del censo.

El segundo es la entidad federativa (*estado*) en formato alfanumérico. La selección por *default* es *Nacional*, que provee los datos para todos los estados. Estos se deben anotar con letra capital y las entidades con múltiples palabras se escriben espaciadas y sin acentos (por ejemplo, San Luis Potosí). Para el caso de la Ciudad de México (anteriormente Distrito Federal) es CDMX.

El resto de los parámetros varía entre funciones. El cuadro 2 presenta todos los disponibles en cada función. El primer grupo de ellos (`total_estado`, `total_municipio`, `total_localidad` y `total_ageb`) requiere información lógica (*TRUE* o *FALSE*), donde *FALSE* es el *default* y significa que la base de datos excluye los resultados agregados a nivel estado, municipio, localidad o AGEB, respectivamente. El segundo grupo (localidades y manzana) también necesita información

Cuadro 1

Funciones disponibles para los censos y conteos de población y vivienda

Función	Tipo de microdatos
<code>censo_poblacion_iter()</code>	Principales resultados por localidad (ITER).
<code>censo_poblacion_muestra()</code>	Muestra del cuestionario ampliado (muestra).
<code>censo_poblacion_ageb()</code>	Principales resultados por Área Geoestadística Básica (AGEB).
<code>censo_poblacion_urbano()</code>	Resultados sobre infraestructura y características del entorno urbano (urbano).
<code>censo_poblacion_rural()</code>	Resultados sobre localidades con menos de 5 mil habitantes (rural).

Cuadro 2

Parámetros disponibles por tipo de microdatos del Censo

Tipo de microdatos	Parámetros
ITER	año, estado, totalestado, totalmunicipio, localidades
Muestra	año, estado, muestra
AGEB	año, estado, totalestado, totalmunicipio, totallocalidad, totalageb, manzana
Urbano	año, estado
Rural	año, estado

lógica, donde TRUE es el *default* y significa que la base de datos incluye información a nivel manzana o localidad, en ese orden.

El objetivo de los parámetros lógicos es ofrecer flexibilidad al usuario en la descarga. Al manipularlos, este puede limitar la descarga de datos a solo uno de los niveles de gobierno o cualquier combinación de estos. En el siguiente bloque de código se presentan cuatro ejemplos de descarga entre diversos niveles de gobierno:

```
# Default: a nivel localidad, pero
sin datos a nivel estado
# ni municipio
censo.jalisco = censo_poblacion_iter
(year = 2010, estado = "Jalisco")
```

```
# Datos a nivel estado, municipio y
localidad
censo.jalisco = censo_poblacion_iter
(year = 2010,
estado =
"Jalisco", totalestado = TRUE,
totalmunicipio = TRUE)
```

```
# Datos a nivel estado y municipio,
pero sin datos a
# nivel localidad
censo.jalisco = censo_poblacion_iter
(year = 2010, estado =
```

```
"Jalisco", totalestado = TRUE,
totalmunicipio = TRUE,
localidades = FALSE)
```

```
# Datos a nivel nacional
censo.jalisco = censo_poblacion_iter
(year = 2010,
estado = "Nacional")
```

En la muestra del Censo, además del año y el estado, un tercer parámetro indispensable es *muestra*, que representa la unidad de análisis, las cuales pueden ser *Migrantes*, *Personas*, *Viviendas* u *Hogar*. En el siguiente bloque de código se ilustra la descarga del cuestionario de personas para el Censo 2010 en Jalisco:

```
# Descarga cuestionario de personas
muestra.jalisco = censo_poblacion_
muestra (year = 2010,
estado = "Jalisco", muestra =
"Personas")
```

3.2 Censo Nacional de Gobiernos Municipales y Delegacionales

Es un proyecto estadístico sobre la gestión y desempeño de las entidades gubernamentales mexicanas a nivel municipal y delegacional. El CNGMD cubre cuatro temáticas: ayuntamiento, administra-

ción pública municipal o delegacional, seguridad y justicia. Cada una representa una fuente de información independiente con múltiples bases de datos, por ejemplo, para su edición 2011, la del ayuntamiento contiene dos conjuntos de datos con variadas bases: comisiones e integrantes.

La función `censo_municipal()` descarga las bases de datos de las cuatro temáticas y tiene tres parámetros indispensables. Primero, es necesario establecer en formato numérico el año del levantamiento del CNGMD, que puede ser 2011, 2013, 2015 o 2017. En el segundo, `fuentes`, se tiene que definir en formato alfanumérico la temática o fuente de la base de datos requerida. Por último, en el parámetro `datos` se debe definir en formato alfanumérico el conjunto de bases de datos.

La relación entre los datos, conjunto de datos y fuentes puede ser complicada, en especial a partir del CNGDM 2015. Como referencia, el cuadro 3 presenta dicha relación para el CNGDM 2011.

Se recomienda consultar la disponibilidad de datos por cada fuente, conjunto de datos y año de levantamiento directamente en la documentación en línea del INEGI. Se puede acceder a esta al teclear la función `censo_municipal()` (sin incluir nada entre los paréntesis), como se muestra en el siguiente bloque de código:

```
# Acceder a la documentación en
# línea del
# CNGDM censo_municipal()
```

Al elegir una fuente y conjunto de datos específicos, la función descarga una lista de `R` con el conjunto de bases de datos correspondiente. En el siguiente bloque de código se ilustra la descarga del conjunto de datos `comisión`, de la fuente `ayuntamiento`. El resultado de la descarga es una lista de `R` con cinco bases de datos (`ACT_AYUN`, `COM_AYUN`, `COM_TASU`, `IN_TEASU`, `INI_AYUN`):

```
# Descargar conjunto de datos
# "comision" del
# censo 2011 datos.comision =
# censo_municipal(year = 2011,
# fuentes = "ayuntamiento", datos =
# "comision")
```

Esta lista, sin embargo, no es útil para el análisis estadístico; cada base de datos debe extraerse de la lista y almacenarse en un objeto de `R` distinto. Esto se logra, primero, identificando la posición de la base de datos en la lista con la función `summary()`. Posteriormente, se debe extraer la base de datos utilizando su posición dentro de la lista, como se ilustra en el siguiente bloque de código:

```
# Explorar la posición de cada
# base de datos en la
# lista summary(datos.comision)

# Extraer la base de datos
# ACT_AYUN, que esta
# en la posición número uno
datos.actas = datos.comision[[1]]
```

Cuadro 3

Relación entre fuentes y conjuntos de datos en el CNGMD del 2011

Fuentes	Datos
Ayuntamiento	Comision, Estructura
Administración	Estructura, Funciones, Marco, Participacion, Recursos, Tramites, Transparencia
Seguridad	Ejercicio, Infraestructura, Recursos
Justicia	Ejercicio, Infraestructura, Recursos

3.3 Encuesta Nacional de Ingresos y Gastos de los Hogares

La ENIGH proporciona datos estadísticos sobre las entradas y salidas económicas de las unidades de observación en cuanto a monto, procedencia y distribución. Adicionalmente, provee información acerca de las características sociodemográficas de los integrantes del hogar. En la actualidad, existen tres variaciones de la Encuesta: 1) Tradicional (1984-2014), 2) Nueva Construcción (2008-2014) y 3) Nueva Serie (2016). El paquete *importinegi* 1.0.0 provee acceso únicamente a la versión Nueva Construcción.

La función `enigh_nuevaconstruccion()` descarga las bases de datos de esta. La ENIGH cuenta con diversas en tres niveles de unidad de análisis: vivienda, hogar y personas. Como se muestra en la figura 1, la Encuesta se compone de varias cadenas de bases de datos subordinadas, por ejemplo, la de trabajos tiene dos de estas (*agro* y *no agro*) y la de *hogares*, tres (*gasto hogar*, *erogaciones* y *gasto tarjeta*). Todas contienen información que puede ser integrable en una sola matriz de datos, aunque esta labor es compleja. La principal ventaja de esta función es que reduce dicha complejidad al descargar e integrar todas estas bases de datos en una sola. La función las une a todas siguiendo el diagrama de relaciones en la figura 1.

La función requiere el parámetro del año del levantamiento de la ENIGH, que puede ser 2008, 2010, 2012 o 2014. Descarga cada una de las bases de datos que componen la Encuesta. La función tiene dos parámetros indispensables. El primero es *año del levantamiento*. El segundo es *datos*, que define la base de datos a descargar. Las opciones son *viviendas*, *hogares*, *concentrado*, *erogaciones*, *gasto hogar*, *gasto tarjetas*, *poblacion*, *ingresos*, *gasto persona*, *trabajos*, *agro* y *no agro*. El siguiente bloque de código provee varios ejemplos del uso de esta función para descargar datos de la ENIGH 2014, Nueva Construcción:

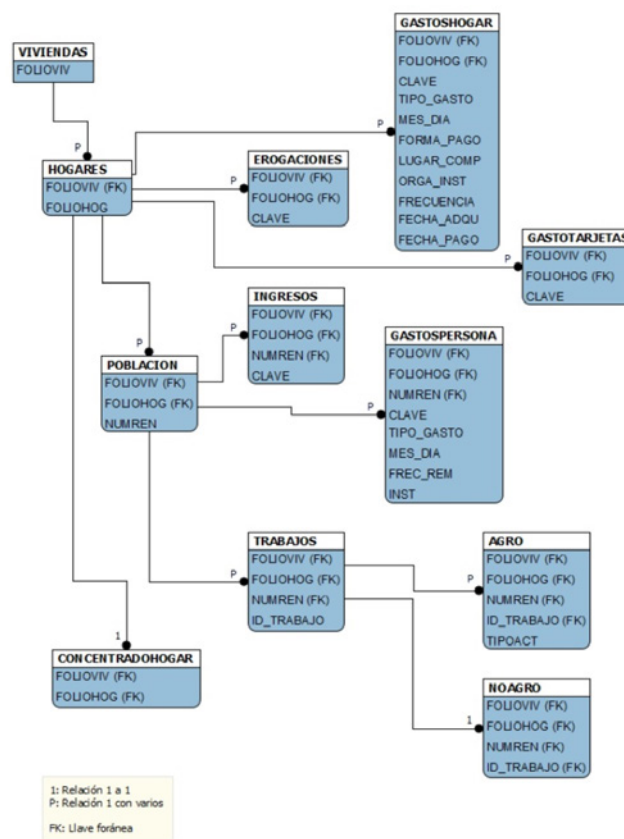
```
# Default: descargar datos de viviendas
```

```
viviendas14 = enigh_nuevaconstruccion
(year = 2014, datos = "viviendas")
```

```
# Descargar datos de hogares
hogares14 = enigh_nuevaconstruccion
(year = 2014, datos = "hogares")
```

```
# Descargar datos de agro
agro14 = enigh_nuevaconstruccion
(year = 2014, datos = "agro")
```

Figura 1
Diagrama de relación de las bases de datos de la ENIGH (INEGI, 2014, p. 5)



3.4 Encuesta Nacional de Ocupación y Empleo

Es un proyecto estadístico de encuestas en hogares especializado en información sobre el mercado laboral. La ENOE provee datos trimestrales sobre

fuerza laboral, ocupación, subocupación y desocupación de los miembros del hogar encuestado. La Encuesta se compone de dos cuestionarios: el *socio-demográfico (CS)* y el *de ocupación y empleo (COE)*, este último en sus versiones básica y ampliada. La versión ampliada del COE se levanta en el trimestre I o II. Por lo tanto, cada levantamiento de la ENOE puede estar compuesto de hasta cinco bases de datos obedeciendo al tipo de instrumento de levantamiento del COE, que son las siguientes: 1) vivienda, 2) hogar, 3) CS, 4) COE parte I y 5) COE parte II.

La función `enoe()` descarga las bases de datos trimestrales de la ENOE desde el 2005. Su sintaxis tiene tres parámetros, de los cuales los dos primeros son indispensables. Primero, es necesario definir el año de levantamiento de los datos en formato numérico. Posteriormente, se debe especificar el trimestre del año en alfanumérico, que puede ser `trim1`, `trim2`, `trim3` o `trim4`. Por defecto, la función descarga por separado las cinco bases de datos que componen la Encuesta. Sin embargo, el tercero, `integrar`, que se describe al final de esta subsección, integra las cinco bases de datos en una sola matriz.

Cuando el parámetro `integrar` está (por *default*) en `FALSE`, la función descarga una lista de *R* con cinco bases de datos separadas. La función `summary` provee un resumen de estas (nombre, tamaño y tipo de objeto) y permite identificar la posición de cada una en la lista. En el siguiente bloque de código se extrae la base de datos del cuestionario de hogares de la ENOE 2005, trimestre I:

```
# Crear un objeto con las cinco bases de
# datos de la ENOE separadas
lista.enoe05 = enoe
(year = 2005, trimestre = "trim1")
```

Esta lista, sin embargo, no es útil para el análisis estadístico; cada base de datos debe extraerse de la lista y almacenarse en un objeto de *R* distinto. Esto se logra, primero, identificando la posición de la base de datos en la lista con la función `summary()`, como se muestra en el siguiente bloque de código:

```
# Resumen de las cinco bases de
# datos
summary(lista.enoe05)
```

Posteriormente, se debe extraer la base de datos utilizando su posición dentro de la lista. En el siguiente bloque de código se ilustra la descarga de la base de datos *hogares*, que se encuentra en la tercera posición de la lista:

```
# Identificar y extraer base de
# datos del cuestionario
# hogares dt.hog105 = lista.
# enoe05[[3]]
```

Como se explicó, el parámetro por defecto `integrar = FALSE` descarga las bases de datos por separado, sin embargo, estas pueden integrarse directamente en una sola matriz utilizando la opción `integrar = TRUE`. Esta opción las conecta a todas mediante el identificador único que tienen en común. El siguiente bloque de código ilustra la descarga e integración de las bases de datos de la ENOE trimestre I:

```
# Descargar e integrar datos
# de la ENOE Trimestre I
lista.enoe.integrada = enoe (year =
2005, trimestre =
"trim1", integrar = TRUE)
```

3.5 Marco Geoestadístico Nacional

Es un proyecto que presenta información sobre la división geoestadística² del país, así como otras formas de clasificación del territorio nacional.

La función `sig_marcogeo()` descarga los mapas del MGN desde 1995 hasta el 2013. Utiliza el paquete `rdga1` para descargar y gestionar archivos de información georreferenciada. Se puede acceder

² En sucesivos niveles de desagregación. Esta división está dada por los llamados *límites geoestadísticos*, que pueden coincidir con los político-administrativos oficiales, los cuales tienen sustento legal; sin embargo, los que no cuentan con este deben entenderse como provisionales (no tienen pretensión de oficialidad), trazados solo para realizar los operativos censales.

a la documentación en línea tecleando la función `sig_marcogeo()` (sin incluir nada entre los paréntesis). Su sintaxis tiene tres parámetros, de los cuales dos son indispensables. Primero, se debe especificar el año del mapa en formato numérico; los años disponibles se encuentran en la documentación de la función. El segundo es el mapa. El Marco provee cinco mapas diferentes: `entidades`, `municipios`, `ageb`, `urbano` y `rural`, aunque la disponibilidad de algunos varía entre años. El tercer parámetro, `version`, solo es necesario en los años que cuentan con múltiples versiones, por ejemplo, para el 2010 son 4.3, 5.0 o 5.0.A. Los años que requieren versión se especifican en la documentación de la función.

El mapa descargado debe ser almacenado en un objeto de *R*. El resultado es un objeto espacial de la clase `SpatialPolygonsDataFrame`, a partir del cual se puede realizar análisis espacial y estadístico. En el siguiente bloque de código se ejemplifica la descarga del mapa de estados del MGN de 1995 y se utiliza una función para su visualización:

```
# Descargar el MGN de 1995
mapa.estados95 <- sig_marcogeo
(year = 1995, mapa =
"entidades")
```

```
# Visualización del mapa
plot(mapa.estados95)
```

3.6 Red Nacional de Caminos

La RNC provee información georreferenciada sobre las vías de comunicación interurbana e intraurbana. Adicionalmente, contiene datos acerca de la infraestructura pública urbana (por ejemplo, túneles, puentes, plazas de cobro, marcas de kilometraje, etc.) y de la de otros medios de transporte (como transbordadores, aeropuertos, puertos y estaciones de ferrocarril).

Para utilizar los mapas de la RNC, se necesitan dos funciones. La primera, `sig_camino_descarga()`, descarga los publicados entre el 2016 y 2019.

Esta función almacena en un objeto de *R* los 13 mapas que componen el conjunto de datos de la Red. En el siguiente bloque de código se ilustra la descarga del conjunto de mapas de la RNC para el 2019:

```
# Descargar mapas de la RNC de 2019
mapas.rnc <- sig_camino_descarga
(year = 2019)
```

La ventaja de la función `sig_camino_descarga()` es que descarga y almacena los 13 mapas de la RNC en una lista de objetos espaciales de *R*; sin embargo, la desventaja es que esta lista no es útil para el análisis espacial; cada mapa debe extraerse de la lista y almacenarse en un objeto espacial de *R* distinto. La segunda función, `sig_camino_extrae()`, cumple este propósito.

Para extraer cada mapa, es necesario especificar, primero, el objeto con la lista de mapas previamente creado con `sig_camino_descarga()`. En el segundo parámetro se debe especificar el mapa que se desea extraer de la lista. Los 13 mapas disponibles son: `estructura`, `localidad`, `maniobra_prohibida`, `plaza_cobro`, `poste_de_referencia`, `puente`, `red_vial`, `sitio_de_interes`, `tarifas`, `transbordador`, `red_localidad`, `red_sitio_de_interes` y `union`. También, la lista se puede consultar utilizando la función `summary()` sobre el objeto creado a partir de `sig_camino_descarga()`. Enseguida se ilustra la extracción del mapa `puentes`, que contiene información de la ubicación geográfica de puentes carreteros:

```
# Obtener un mapa de puentes
mapa.puentes <- sig_camino_extrae
(mapas.rnc, mapa = "puente")
```

4. Conclusiones y agenda de desarrollo

La provisión de herramientas para extender el acceso y procesamiento de datos para la investiga-

ción es una práctica que contribuye a mejorar la *reproducibilidad* de la ciencia y el acceso y uso de los datos abiertos (Peng, 2011; Munafò *et al.*, 2017). También, coadyuva a mejorar la transparencia, estandarización y automatización del proceso de investigación, desde la recolección de datos en bruto hasta la publicación final del producto.

El presente artículo busca apoyar tanto a la comunidad académica como a la profesional con una herramienta que facilite la obtención y el procesamiento de datos claves para la comprensión de la realidad nacional. El paquete *importinegi* provee funciones para acceder a datos abiertos del repositorio institucional del INEGI de manera simplificada, estandarizada y automatizada, con lo cual se quiere contribuir a mejorar el proceso en la investigación y enseñanza basadas en datos del Instituto. Fue desarrollado siguiendo las mejores prácticas de programación de paquetes en *R* (Wickham, 2015), *open source* y *reproducibilidad* de la ciencia (Gandrud, 2016).

La versión 1.0.0 de *importinegi*, expuesta en este documento, representa la primera fase de desarrollo del paquete, que busca seguir incorporando el acceso a más proyectos estadísticos del INEGI. La utilización de este *software* en la labor académica brinda la posibilidad de adoptar protocolos de *reproducibilidad* en el campo de las ciencias sociales, proveyendo a cada usuario el acceso a las bases de datos originales e íntegras producidas por el Instituto. Asimismo, al automatizar algunos procesos de descarga y procesamiento, reduce las capacidades de programación necesarias para utilizar las bases de datos del INEGI en el *software R*. Esto permite a las personas tanto del sector académico como profesional acceder y procesar bases de datos de amplia importancia para el análisis y la toma de decisiones.

Fuentes

- Arel-Bundock, V. *Wdi: World development indicators (world bank)*. Vol. 2, 2013.
- Boettiger, C.; S. Chamberlain; E. Hart & K. Ram. "Building software, building community: lessons from the ropensci project", en: *Journal of open research software*. 3 (1), 2015.
- Gandrud, C. *Reproducible research with r and r studio*. Chapman and Hall/ CRC, 2016.
- Gheri, F. & M. C. Kao. "Download and harmonize faostat and wdi data: the faostat package" en: *R Package Version*, 1. 2014.
- INEGI. *Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2014*. México, INEGI, 2014.
- Lahti, L.; J. Huovari; M. Kainu & M. Biecek. "Retrieval and analysis of eurostat open data with the eurostat package", en: *The R Journal*. 9 (1), 2017, pp. 385-392.
- Munafò, M. R.; B. A. Nosek; D. V. Bishop; K. S. Button; C. D. Chambers; N. P. du Sert & J. P. Ioannidis. "A manifesto for reproducible science" en: *Nature Human Behaviour*. 1 (1), 0021, 2017.
- Peng, R. D. "Reproducible research in computational science", en: *Science*. 334 (6060), 2011, pp. 1226-1227.
- Walker, K. *tidycensus: Load us census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames. r package version 0.4.1*. 2018.
- Wickham, H. *R packages: organize, test, document, and share your code*. O'Reilly Media, Inc., 2015.
- Labo. Alibus, sam et odia quam, nobis porest, te nis imus di alignim plect atenducium quo vendio omnim adictibusam iusandi oditium harioeribus eos earciis sae cuptatae eos dolore peritiones re inv ipsandae. Odis ertestio. Atiis eturistius dolendis santo quibusto exceaqua perit maio experupitat.
- Lor rehendis comnis im exped quodist, que vel ex eos dolupta ssitibustium nonsera escuscienet volor alias aruptaq uiscis suntis si ut ut laborum dem ipsam quunt.
- Occus qui consent orporporem rentem que vendi aut landit ex estrum, verio es nos ea pos molenim olorist, to magniscimi, tendige ndusapitem evellenimet renda simendam quae nos entorrum fuga. Itatem nati aspe laut estotam fugit officipsam as errum que pa ipit fuga. Henimposant quia des sapit quam vel is et quam, consequamus eat autem ilit, officae. Tiatem facerunt im ea venihil iur?

Análisis de eficiencia *en los servicios del agua en México* **con datos de los Censos Económicos**

Analysis of Efficiency *of Water Services in Mexico* ***Using Data from the Economic Censuses***

Ismael Aguilar-Benítez*

* El Colegio de la Frontera Norte, iaguilar@colef.mx

Cabadas-bridge over river Lerma./Professional geographer who loves to capture landscapes/Getty images



En este trabajo se analiza la eficiencia técnica de los organismos de captación, tratamiento y suministro de agua en México utilizando datos generados por los Censos Económicos. Para ello, se comparan cinco estimadores obtenidos mediante la aplicación de técnicas basadas en el concepto de fronteras de eficiencia. Las principales aportaciones de esta investigación son que se utiliza por primera vez una base de datos económicos censales para estimar la eficiencia técnica de los servicios del agua en nuestro país, que se realiza un análisis comparativo de la eficacia de estimadores de fronteras de eficiencia en un contexto de fuerte presencia de valores atípicos y errores en los datos y que los resultados permiten identificar mecanismos para el diseño de políticas públicas específicas para fortalecer estos servicios. En general, se recomienda una efectiva regulación y mayor profesionalización en los servicios del agua.

Palabras clave: eficiencia; organismos de captación, tratamiento y suministro de agua; censos económicos; México.

Recibido: 18 de agosto de 2020.
Aceptado: 28 de agosto de 2020.

Introducción

La evaluación de la eficiencia de las organizaciones que proveen los servicios de agua potable, drenaje y saneamiento proporciona información relevante para los administradores de esos servicios y el diseño de políticas públicas (Picazzo-Tadeo *et al.*, 2009).¹ Utilizar de manera eficiente los distintos insumos para su provisión es una condición necesaria para que la población tenga acceso a ellos, con características adecuadas de cantidad, calidad y continuidad. Su disponibilidad es esencial para proteger el derecho humano al vital líquido, al saneamiento y a la salud, como lo ha mostrado la reciente contingencia sanitaria por el COVID-19,

¹ Existen varios tipos de eficiencia que se pueden aplicar a los servicios del agua: técnica, operativa o económica. En este trabajo nos enfocamos en el análisis de la primera, que se entiende, en general, como la combinación óptima de insumos para lograr un nivel dado de producto. La segunda es la más común y se analiza por medio de indicadores de desempeño, p. e. el número de empleados por cada mil tomas o el porcentaje de dinero facturado que efectivamente se cobra. Cuando se incluye el aspecto menor costo y se propone generar el mayor producto posible, se habla de eficiencia económica.

The efficiency of water services in Mexico is analyzed using data from the Economic Censuses. We compare estimates generated with five techniques based on the concept of efficiency frontier. The three main contributions of this work are: one, a census database is used for the first time to estimate the technical efficiency of water services in Mexico; two, a comparative analysis of the effectiveness of the estimators is performed in a context of strong presence of outliers and errors in the data. Three, the results allow identifying mechanisms for the design of policy recommendations for the sector; In general, it is recommended an effective regulation and professionalization of the water services.

Key words: technical efficiency; water utilities; economic census; Mexico.

durante la cual ha sido indispensable tener agua para la higiene personal y de la vivienda.

No obstante su importancia, quienes ofrecen el servicio tienen escasos incentivos para mejorar su eficiencia. En la práctica, en México y en otros países, los servicios de agua y saneamiento operan como monopolios debido a características económicas especiales, por lo que en cada ciudad o población solo hay un proveedor (Aguilar-Benítez, 2011). Por otro lado, la gran mayoría de estas compañías operan bajo la administración del sector público y dependen financieramente del subsidio. Ambas características (monopolio y dependencia de fondos públicos) ocasionan, en parte, que esas organizaciones tengan pocos estímulos para ser eficientes en el uso de sus recursos y, aún menos, para brindar atención de calidad.

En este trabajo nos enfocamos al análisis de la eficiencia técnica de los organismos que proveen de servicios del agua a las localidades de México,

la cual puede entenderse como la mejor combinación de insumos necesarios para generar una determinada cantidad de producto.²

Aunque cada uno de estos utiliza insumos semejantes (p. e. agua, mano de obra, cloro y otros materiales para potabilización, energía eléctrica y redes para distribución, entre otros), la forma en la que se combinan difiere por múltiples razones, como la fuente principal de agua, su calidad, condicionantes topográficas en el área que cubren, así como su contexto demográfico, económico y organizacional, por mencionar algunos. La presencia de múltiples insumos y los resultados que generan (ingresos por tarifas, conexiones con servicio, volúmenes de agua potabilizada, recolección de aguas residuales, volumen de aguas residuales tratadas), así como una relación compleja entre ellos, son características que hacen necesaria la aplicación de métodos específicos para la estimación de eficiencia.

En esta investigación se analiza la eficiencia de proveedores de servicios del agua utilizando los datos generados por el cuestionario para los organismos que realizan la actividad de captación, tratamiento y suministro de agua (OCTSA),³ el cual forma parte de los Censos Económicos (CE) 2014 en México.⁴ Como único antecedente de una obra similar se registra un trabajo en el que se analizó el efecto de la descentralización en la eficiencia de una muestra de 110 organismos operadores (Anwandter y Ozuna, 2002).

Las principales aportaciones son tres: 1) desde un punto de vista empírico, se utiliza por primera

vez una base de datos censales para estimar la eficiencia técnica de los OCTSA en México y se muestra su potencial para el análisis de esta actividad; 2) desde una perspectiva metodológica, este trabajo busca contribuir al análisis comparativo de la eficacia de los estimadores analizados para el estudio de la eficiencia técnica en un contexto de fuerte presencia de valores atípicos y errores en los datos; y 3) los resultados de este ejercicio permiten identificar elementos para el diseño de políticas públicas para mejorar el desempeño de los servicios del agua en México.

Es importante mencionar que aquí se restringe el análisis de la eficiencia técnica a la información disponible sobre un grupo de variables (insumos y producto) y no incluye lo relacionado con la calidad en el servicio y el agua o la continuidad de los servicios, aspectos importantes para cuyo análisis se requiere información que no se encuentra disponible en la base de datos utilizada.

El escrito se organiza como sigue: la siguiente sección presenta las características de los OCTSA en México derivadas del cuestionario ya mencionado; en la segunda se caracteriza el enfoque de fronteras de eficiencia y se describen los estimadores que se aplicaron; en la tercera, se muestran los principales resultados de la estimación, aplicando los cinco estimadores con respecto a su eficacia para reducir el sesgo generado por una alta presencia de valores atípicos; y en la sección final se presentan las conclusiones.

Características de los OCTSA en México, CE 2014

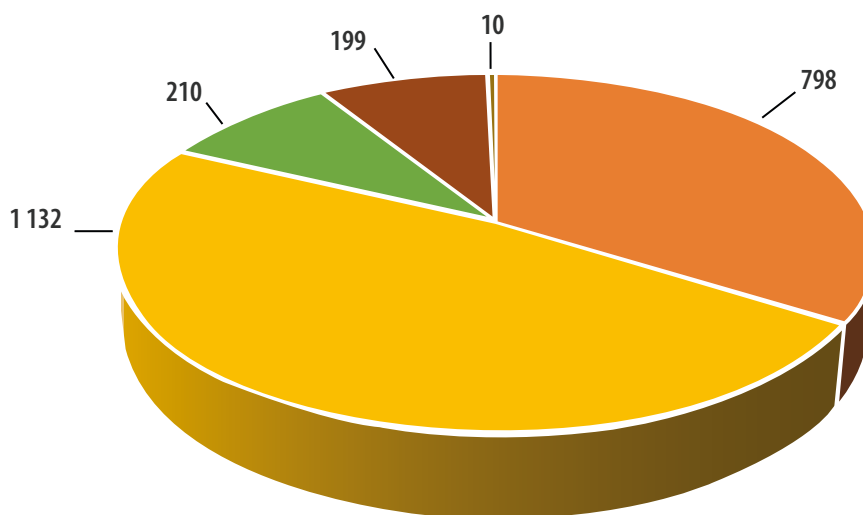
El número total de organismos censados en México en 2014 fue de 2 688. Un OCTSA puede proveer de servicios, en muchos casos, solo a la cabecera municipal, aunque también puede abastecer a varios municipios o, incluso, a toda la entidad (p. e. Servicios de Agua y Drenaje de Monterrey abastece a todo el estado de Nuevo León). Las formas de organización son diversas, principalmente: organismos operadores, sistemas de agua, juntas loca-

2 A lo largo del texto se utilizan los términos insumo y producto para referirnos a los términos en inglés *inputs* y *outputs*, aunque debe notarse que el primero se refiere a factores de producción y el segundo, a resultados de un proceso. De esta manera, aunque referirse a número de conexiones como producto no es estrictamente correcto, se hace para facilitar la exposición.

3 Este trabajo se deriva de los resultados del proyecto de investigación *Hacia una gestión integral del agua por cuenca hidrológica: un análisis de la disponibilidad y usos*, financiado por el Consejo Nacional de Ciencia y Tecnología (CONACYT) en la convocatoria Problemas Nacionales 2014, número 248 719.

4 La actividad de captación, tratamiento y suministro de agua se reporta con la clasificación 2221, de acuerdo con el Sistema de Clasificación Industrial de América del Norte (SIICAN) 2013 México. El uso de microdatos para el análisis de este trabajo se realizó en el Laboratorio de Microdatos del Instituto Nacional de Estadística y Geografía (INEGI) y fue registrado como proyecto LM588. El autor agradece el apoyo del personal del Instituto.

OCTSA por tamaño de población de la localidad atendida



■ <= 2 500 hab.
 ■ 2 501-20 000 hab.
 ■ 20 001-50 000 hab.
 ■ 50 001-1 millón hab.
 ■ de 1 millón hab.

Fuente: elaboración propia con datos de los CE 2014.

les, comités municipales y concesionarias de agua particulares. Del total, 2 401 OCTSA (90 %) se clasifican como servicios del sector público; 257 (9 %) se registran como asociaciones civiles y solo 30 (1 %) operan como sociedades mercantiles con fines de lucro o sociedad cooperativa; de la totalidad, 1 394 proveen servicios a poblaciones exclusivamente urbanas y 1 294, a centros tanto urbanos como rurales.

En México, 33 % de los OCTSA provee de servicios a localidades menores a 2 500 habitantes, mientras que el tamaño más frecuente se encuentra en el rango de 2 501 a 20 mil (48 %) y solo 10 de la base de datos final para este trabajo abastecen a asentamientos urbanos con más de 1 millón de habitantes (ver gráfica 1). Los OCTSA emplearon, en 2013, a 122 798 personas; 96 % de este personal dependía en forma directa del organismo y el resto era subcontratado. Los organismos no necesariamente funcionan en un tamaño óptimo, pueden ser muy pequeños (con algunos cuantos

empleados, para un número reducido de conexiones) o crecer más allá de su tamaño recomendable de personal.

La cantidad total de agua captada por los OCTSA fue de 12 400.6 millones de metros cúbicos anuales, considerando el volumen de extracción de fuentes superficiales (27 %), del subsuelo (61 %) y del resto que proviene de compras de agua en bloque (varios OCTSA requieren realizar esta acción para complementar el volumen del líquido que necesitan debido, sobre todo, al agotamiento o sobreexplotación de fuentes locales). De acuerdo con los volúmenes reportados, solo la mitad (51 %) del agua suministrada se cobra.

Para ofrecer servicios del agua en México, se reportaron gastos por consumo de bienes y servicios por un monto total de 24 935.3 millones de pesos. Los principales conceptos que generaron esos gastos son: energía eléctrica (39.5 %); suministro de agua en bloque (16.1 %); reparaciones

y refacciones para mantenimiento (9.8 %); compra de agentes físico-químicos, reactivos y otros materiales para la distribución de agua (7.6 %); y combustibles, lubricantes y energéticos (5.4 %); el resto (21.6 %) se destinó a otros pagos, como: publicidad, papelería, servicios de comunicación y profesionales, primas de seguros, entre otros.

Por suministro de bienes y servicios, los OCTSA obtuvieron un total de 47 561.2 millones de pesos, ingresos que se generaron, sobre todo, por: suministro de agua (69.3 %), derechos de conexión y alcantarillado (11.8 %), ingreso por ejercicios anteriores (9.2 %), venta de agua tratada (2.1 %) y otros conceptos (7.6 %). En este último rubro se incluyen la venta de lodos e ingresos tanto por publicidad como por la distribución de agua con pipas. De los proveedores de servicios del agua, 35 % reportó ingresos por suministro de bienes y servicios menores a sus gastos.

Como se ha descrito, en México se registra un gran número de OCTSA con una diversidad de formas de organización y tamaños de localidades a las que proveen servicios. Dos indicadores de las debilidades en su operación es que, en promedio, solo la mitad del agua suministrada se cobra y la tercera parte de estos operadores no recuperan sus gastos por consumo de bienes y servicios.

Noción de frontera de eficiencia y técnicas no paramétricas

Para el análisis empírico de la eficiencia técnica, se han propuesto en la literatura especializada dos enfoques: uno es el conocido como paramétrico y el otro, el no paramétrico. La diferencia principal

entre ellos es que, mientras que en los métodos que utilizan el primero se requiere especificar una función de producción con parámetros que la caracterizan, en los que emplean el segundo no se necesita determinarla. En los primeros, la especificación de una función de producción que se analiza con técnicas econométricas es muy relevante para obtener resultados confiables; en contraste, una ventaja de los no paramétricos es que no es necesario asumir o establecer una que defina la relación entre insumos y productos o resultados, esto es, no se especifica la forma funcional de la función de producción.

Debido a que, de manera usual, no existe una única combinación que genere los mejores resultados, una opción para evaluar la eficiencia de una unidad es compararla dentro de un grupo o muestra. Los métodos no paramétricos responden a esa necesidad empírica de evaluar la eficiencia como una noción relativa a la mejor práctica observada en un conjunto de unidades de referencia o grupo de comparación (Daraio y Simar, 2007).

La base de la cual parte la inferencia no paramétrica es usar un conjunto de observaciones para evaluar la eficiencia de cada unidad (empresa, organización o unidad de decisión) con respecto a la del conjunto de datos. La eficiencia técnica se obtiene al comparar el valor observado de cada unidad con el óptimo que se determina por una frontera de producción estimada. La frontera de producción puede entenderse intuitivamente como la unión de las diferentes combinaciones que optimizan los insumos necesarios para generar una cantidad de producto; por ejemplo, en el caso simple de un producto o resultado (Y) generado con dos

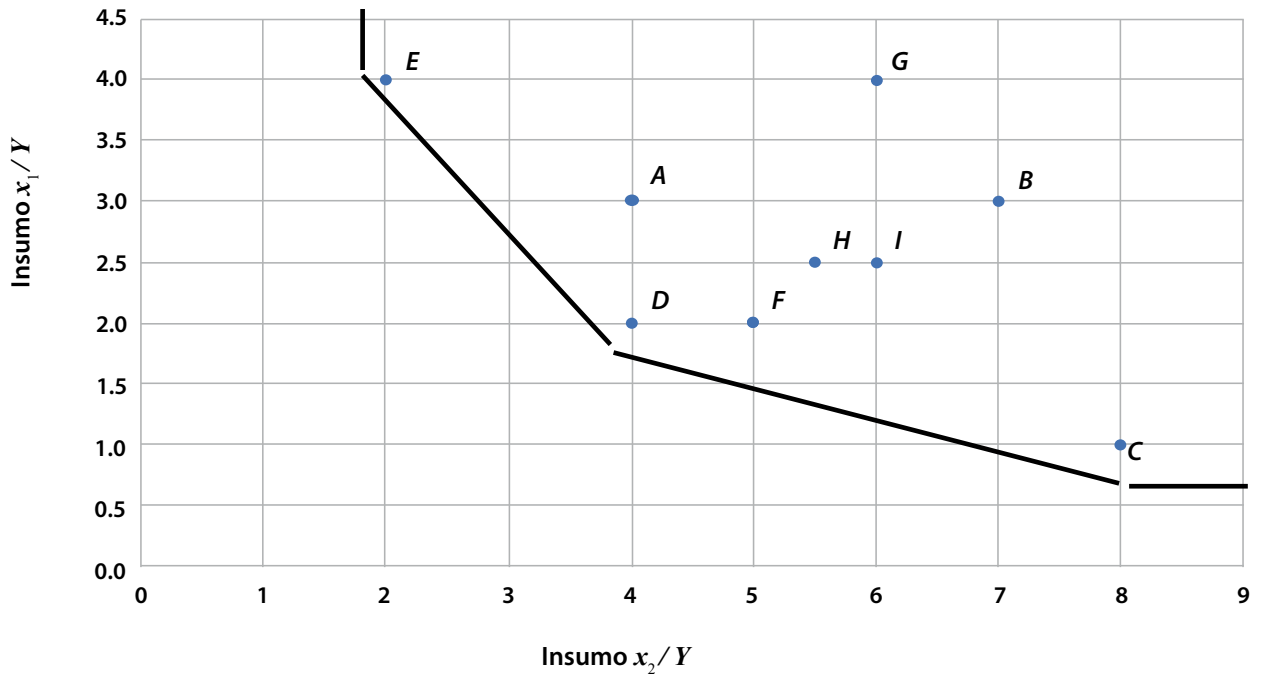
Cuadro 1

Caso simple de dos factores y un producto

Organización	A	B	C	D	E	F	G	H	I
x_1	4	7	8	4	2	5	6	5.5	6
x_2	3	3	1	2	4	2	4	2.5	2.5
Y	1	1	1	1	1	1	1	1	1

Fuente: elaboración propia, adaptado de Cooper et al., 2006.

Frontera de eficiencia para el caso de dos insumos (x_1 y x_2) y un producto



Fuente: elaboración propia, adaptado de Cooper *et al.*, 2006.

factores o insumos (x_1, x_2) en nueve diferentes organizaciones (A, B, \dots, I) se registran las distintas combinaciones que cada una utiliza para generar la unidad de producto.

Al graficar esas combinaciones, registrando como coordenadas los valores de los insumos x_1 y x_2 en los ejes horizontal y vertical de un cuadrante—esto es, las cantidades que cada organización necesita de cada insumo para generar una Y — se pueden identificar los puntos ilustrados ($A \dots I$) en la gráfica 2. Uniendo los tres puntos (E, D y C), que representan las diferentes combinaciones con los menores valores de insumos, se construye la frontera de eficiencia. En el punto E , la línea vertical proyecta la frontera de eficiencia marcando el límite para el insumo x_2 , mientras que la línea horizontal proyectada en el punto C hace lo propio para el caso del insumo x_1 .

Como puede observarse en la gráfica 2, la curva generada envuelve las distintas combinaciones observadas; de este resultado gráfico se deriva el

nombre de una de las técnicas más usuales no paramétricas: el análisis envolvente de datos (DEA, por sus siglas en inglés), que se describirá más adelante. En este ejemplo, los puntos (combinaciones de insumos) que se encuentran por encima de la frontera de eficiencia son ineficientes, pues para generar una unidad de producto requieren de una mayor cantidad de, al menos, un insumo comparado con las tres organizaciones eficientes. Como puede intuirse con esta gráfica, los puntos G y B son los más alejados de la curva y, por lo tanto, los más ineficientes.

A partir de esa noción de frontera se pueden explicar los distintos modelos utilizados para el análisis no paramétrico de la misma, el cual requiere dos especificaciones iniciales en sus modelos. La primera de ellas es la orientación de la eficiencia, que puede ser hacia los insumos, lo cual implica que las unidades tomadoras de decisiones (DMU, por sus siglas en inglés) buscan, dado un nivel fijo de producto(s), la máxima reducción proporcional en la combinación de insumos. Este es el caso del ejemplo simple ilustrado en la gráfica 2, en la que

se define la frontera de eficiencia con las combinaciones de menores cantidades de insumos para generar una unidad de producto. En el caso contrario (la eficiencia orientada hacia los productos), las DMU buscan, dado el nivel de insumos, el máximo incremento proporcional de los productos, permaneciendo dentro de la frontera de posibilidades de producción. Para el caso de los servicios del agua, es razonable asumir que, dado un contexto de limitada disposición de insumos (agua, recursos financieros para infraestructura hídrica, energía eléctrica y mano de obra, entre los principales), la orientación pertinente es hacia los insumos.

Una segunda especificación necesaria para la aplicación de modelos no paramétricos es el tipo de rendimientos de escala que se asume: constantes cuando, para generar un incremento (reducción) en la cantidad de producto, se requiere un incremento (decremento) proporcional de cada insumo necesario; o variables, cuando al incrementar (reducir) la cantidad utilizada de los factores, la cifra obtenida del producto se incrementa (reduce) en una proporción distinta a la de los insumos. En el caso de los servicios del agua, no existe un consenso en la literatura sobre este supuesto, por lo que en este trabajo se realiza una prueba en la base de datos para definir la especificación de rendimientos de escala del modelo.

Para evaluar la eficiencia técnica de los servicios del agua en México, se comparan estimaciones que resultan de la aplicación de cinco técnicas no paramétricas conocidas como: 1) análisis envolvente de datos; 2) libre disponibilidad de factores (FDH, por sus siglas en inglés), 3) de orden m , 4) de orden alfa (orden α) y 5) corregido por sesgo (BC, por sus siglas en inglés), las cuales se describen enseguida.

1) DEA

El primer estimador no paramétrico de fronteras de eficiencia instrumentado para el análisis empírico de los OCTSA es el que resulta del análisis en-

volvente de datos. Esta técnica compara las DMU, considerando todos los insumos utilizados y los productos generados (Cooper *et al.*, 2011). De manera gráfica, el conjunto de soluciones posibles genera una envolvente de todas las observaciones (como se ilustró en la gráfica 2 con un caso simple), de ahí el nombre de la técnica: DEA. La eficiencia de una DMU, definida por Farrell en 1957, se evalúa con la distancia radial del nivel en el que opera a su punto correspondiente en la frontera. DEA utiliza programación matemática y fue instrumentada para su aplicación empírica por Charnes, Cooper y Rhodes (1978). El tratamiento de los rendimientos variables a escala en el modelo DEA (conocido como BCC por haber sido desarrollado por Banker, Charnes y Cooper, 1984) introduce este supuesto en la técnica DEA. Debido a que este último es, como veremos, el supuesto más apropiado para nuestro estudio, los distintos estimadores se explican bajo ese criterio. Utilizando notación de programación lineal, el modelo BCC puede plantearse de la siguiente manera: suponga que se tiene un número n de DMU donde cada DMU_j , $j = 1, 2, \dots, n$ produce los mismos productos en diferentes cantidades y_{rj} ($r = 1, 2, \dots, s$) usando los mismos insumos, x_{ij} ($i = 1, 2, \dots, m$), también en diferentes cantidades. Formalmente, el modelo orientado a los insumos, asumiendo BCC, se formula como un problema de programación lineal de la siguiente manera (Bowlin, 1998):

$$\text{Minimizar: } \theta - \epsilon \left[\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right]$$

$$\text{Sujeto a: } 0 = \theta x_{i0} - \sum_{j=1}^n x_{ij} \lambda_j - s_i^-$$

$$y_{r0} = \sum_{j=1}^n y_{rj} \lambda_j - s_r^+$$

$$1 = \sum \lambda_j$$

$$0 \leq \lambda_j; s_i^-, s_r^+ \text{ para } i=1, \dots, m., r = 1, \dots, s; j = 1, \dots, n$$

La técnica DEA asume convexidad estricta del conjunto de posibilidades de producción; i. e., todas las combinaciones que se encuentran en el conjunto de posibilidades teóricas son factibles. Una debilidad es que requiere de datos confiables (sin error de medición) y libres de la presencia de valores atípicos, lo cual es un supuesto fuerte para el caso de los servicios del agua.

2) FDH

Parte de los mismos principios del estimador DEA, pero relaja el supuesto de convexidad; solo son factibles las combinaciones definidas por los datos observados. FDH asume la libre disponibilidad en el conjunto de posibilidades de producción, determinado por las combinaciones o puntos de frontera observados (considerando que no hay error en la medición); con ello, reduce las posibilidades al área determinada de forma empírica por los datos observados (Deprins *et al.*, 1984).

En este caso, el conjunto de producción asume una fuerte disponibilidad de *output* y libre disponibilidad de insumos, formando una envolvente poliédrica, pero no necesariamente convexa. En programación lineal, el estimador FDH con orientación a los insumos se plantea de manera formal de la siguiente manera (Marques y De Witte, 2011):

$$\theta(x, y) = \min \left\{ \theta \mid \theta x \geq \sum_{i=1}^n \lambda_i x_i; y \leq \sum_{i=1}^n \lambda_i y_i; \sum_{i=1}^n \lambda_i = 1; \lambda_i \in \{0, 1\}; 1, \dots, n \right\}$$

Estos dos primeros estimadores (DEA y FDH) son clasificados como determinísticos por sus supuestos sobre la determinación de las posibilidades de producción teóricas u observables derivadas de un conjunto de datos que, se asume, no registran errores de medición ni presencia de valores atípicos. Las estimaciones generadas por estos modelos tienen como principal limitación su alta sensibilidad a valores atípicos y a la presencia de errores en los datos (Cazals *et al.*, 2002).

3) Orden m

A diferencia de los dos anteriores, este no envuelve todos los puntos de datos observados, por lo que se considera como una comparación (*benchmarking*) entre un número m (muestra) —cuyo tamaño a

determinar depende, en parte, de la magnitud del grupo de DMU—; el procedimiento para obtenerlos se resume en cuatro etapas:

- Tener una muestra m aleatoria con reemplazo del grupo de DMU que satisface la condición de producir, al menos, tanto producto como la DMU_i .
- Calcular la eficiencia parcial usando esa muestra.
- Repetir D veces los pasos 1 y 2.
- La eficiencia orden m se calcula como el promedio de las estimaciones obtenidas en el paso 3.

De manera formal, este estimador tiene como punto de partida, desde un modelo orientado a los insumos, el concepto función de insumos mínimos esperados (Cazals *et al.*, 2002). En el caso más simple, considérese un entero fijo $m \geq 1$. El límite inferior de X de orden m se define como el valor esperado del mínimo de m variables aleatorias x^1, \dots, x^m extraído de la función de distribución de X . La definición formal es la siguiente (Daraio y Simar, 2007, p. 68):

$$\begin{aligned} \phi_m &= E[\min(x^1, \dots, x^m)] \\ &= \int_0^\infty [1 - F_x(x)]^m dx \end{aligned}$$

El valor ϕ_m es el nivel de insumo mínimo alcanzable esperado entre m DMU extraídos de la población de empresas que generan una cantidad determinada de producto. En general, y en términos económicos, se puede interpretar la frontera esperada de orden m , $\phi_m(Y)$, como el valor mínimo esperado de insumo alcanzable entre un número fijo (m) de DMU extraídas de la población que producen, al menos, un nivel de producto dado Y .

Debido al remuestreo aleatorio con reemplazo (paso 3), una DMU_i puede o no estar disponible como su propio par. Por esa razón, los estimadores orden m con orientación a los insumos pueden exceder el valor de 1 y se consideran supereficientes

por localizarse más allá de la frontera de posibilidades de producción estimada. Esa característica de los resultados del estimador permite identificar unidades que pueden ser clasificadas como valores atípicos, bien porque en realidad son tan eficientes que están fuera de la frontera de eficiencia estimada o por registrar errores en los datos.

4) Orden alfa (orden α)

Puede considerarse como una versión continua del estimador orden m . Se basa en la idea de que para cada empresa existe, en el conjunto de observaciones de comparación, una frontera de eficiencia cuantil que la toca, en la cual la compañía es eficiente (ya sea a lo largo de la dimensión de insumo o la de producto). Para evaluar la aptitud, se determina la frontera fijando primero la probabilidad $(1 - \alpha)$ de observar puntos por encima de este límite. El punto de referencia será la frontera de cuantiles orden- α definida como el nivel de insumo no superado por $(1 - \alpha) \times 100$ de las empresas que producen, al menos, el mismo nivel de resultados Y . Formalmente:

$$\phi_{\alpha}(y) = \inf\{x F_{x|y}(x|y) > 1 - \alpha\}$$

Por ejemplo, si el cuantil en el que la empresa es eficiente (en la orientación insumo) es $\alpha = 0.2$, esto significa que hay 80 % $(1 - 0.2 = 0.8)$ de las unidades en el conjunto de comparación que superan a la empresa considerada utilizando menos insumos. Cuando $\alpha < 100$, algunas DMU pueden no estar envueltas por la frontera de posibilidades de producción estimada. Esto permite, al igual que orden m , identificar a las unidades que pueden ser valores atípicos (*outliers*).

5) BC

Los cuatro anteriores miden la eficiencia relativa a partir de la estimación de una frontera de esta derivada de un conjunto de datos observados y la comparación de cada unidad, ya sea con el grupo

completo de unidades (DEA, FDH) o una muestra de ellos (orden m , orden α).

En el caso del estimador BC, se utilizan los datos observados para generar muestras de datos simulados mediante un método llamado *bootstrap*, el cual se basa en la idea de simular de manera repetida el proceso de generación de datos mediante remuestreo. Para implementar este enfoque, se proponen dos algoritmos *bootstrap* (ver planteamiento formal en Kneip *et al.*, 2008). Mediante múltiples muestras simuladas a partir de un modelo teórico que trata de explicar la forma en la que se generaron los datos (llamado proceso generador de datos), se trata de reducir el sesgo que puede representar la presencia de valores atípicos o errores de medición en el conjunto de observaciones empíricas disponibles. El algoritmo I consiste en una simulación de datos que extrae n muestras en cada réplica para construir seudomuestras y obtener estimadores de eficiencia. El II es un proceso de *suavizado* (*smoothing*) de la distribución empírica de datos y de la frontera de producción; este se enfoca a suavizar la distribución empírica de datos observados, con lo cual se trata de reducir el efecto de los valores atípicos.

Preparación de la base de datos y el modelo para el análisis no paramétrico

La base de datos original incluye información de 2 668 OCTSA captados por los CE 2014 en México. Debido a que la aplicación de métodos no paramétricos requiere de valores no negativos en insumos y productos, para el análisis se consideraron solo valores estrictamente positivos. Así, se eliminaron 188 observaciones sin registro de redes de distribución de agua potable; también, los que reportaron cero conexiones (69) o cero ingresos (45); además, se identificaron las que mostraron cifras de conexiones inconsistentes para tratarse de OCTSA que abastecen a localidades con uso doméstico (i. e. un número de conexiones domésticas menor a la suma de conexiones comerciales e industriales, que sumaron 35); por último, se detectaron valores extremos inconsistentes, como

los muy altos en longitud de redes que no corresponden con los de número de conexiones y personal (dos observaciones). La cantidad final de OCTSA con información completa para el cálculo de los estimadores fue de 2 349.

La selección de variables que se incluyeron en el modelo se basó en la revisión de la literatura especializada y la disponibilidad de datos. Se eligieron cuatro variables insumo disponibles en la base de datos; todos han sido utilizados en otros trabajos que analizan la actividad de servicios del agua: 1) la longitud de redes (en km), donde se sumaron las extensiones de tuberías de las redes primaria y secundaria reportadas; 2) los gastos en bienes y servicios (en pesos corrientes 2013), que se obtuvieron sumando los rubros: gastos por agentes físico-químicos, reactivos e insumos, por energía eléctrica y por suministro de agua en bloque; 3) el total de personal (número), que se formó con la suma del técnico y el administrativo dependiente de la razón social (contratado de forma directa, de planta, eventual y no reenumerado, sujeto a la dirección y control del OCTSA) y no dependiente de la razón social; y 4) el volumen captado de agua (en miles de m³), el cual se calculó mediante la suma de volúmenes de agua superficial y del subsuelo que obtuvo de manera directa cada OCTSA.

Como variable producto, en esta investigación se utilizó el número de conexiones domésticas, que ya ha sido usado como tal en otros trabajos.

Definir a este como la variable producto y no, por ejemplo, a los ingresos obtenidos, responde al reconocimiento que los OCTSA en México son organizaciones no lucrativas y su principal objetivo es abastecer de agua entubada, drenaje y saneamiento a todas las personas que habitan su área de influencia. En efecto, de acuerdo con los datos de los CE 2014, una proporción importante de operadores (35 %) genera ingresos menores a sus gastos, lo que claramente afecta la posibilidad de utilizar las entradas de dinero como variable producto, pues no es realista asumir que estos tienen como propósito mantener un nivel dado de ingresos u optimizarlos. Es necesario mencionar que esta elección no implica que la finalidad de los OCTSA deba ser sola o principalmente la cobertura y lograr el máximo posible de conexiones. De manera ideal, deberían incluirse otros aspectos (variables), como la calidad del agua distribuida, la continuidad de los servicios, la asequibilidad de las tarifas y la equidad en acceso, entre otros. Aunque esto es técnicamente posible, pues una de las ventajas de los métodos no paramétricos consiste, de hecho, en incluir varios *productos*, no fue factible generar estas variables con la información disponible.

En resumen, el modelo general que se utiliza en este trabajo se define por las cuatro variables insumo mencionadas y una variable producto. El cuadro 2 muestra que estas registran un amplio rango de valores; por ejemplo, se tienen OCTSA que captan desde mil hasta más de mil millones de metros

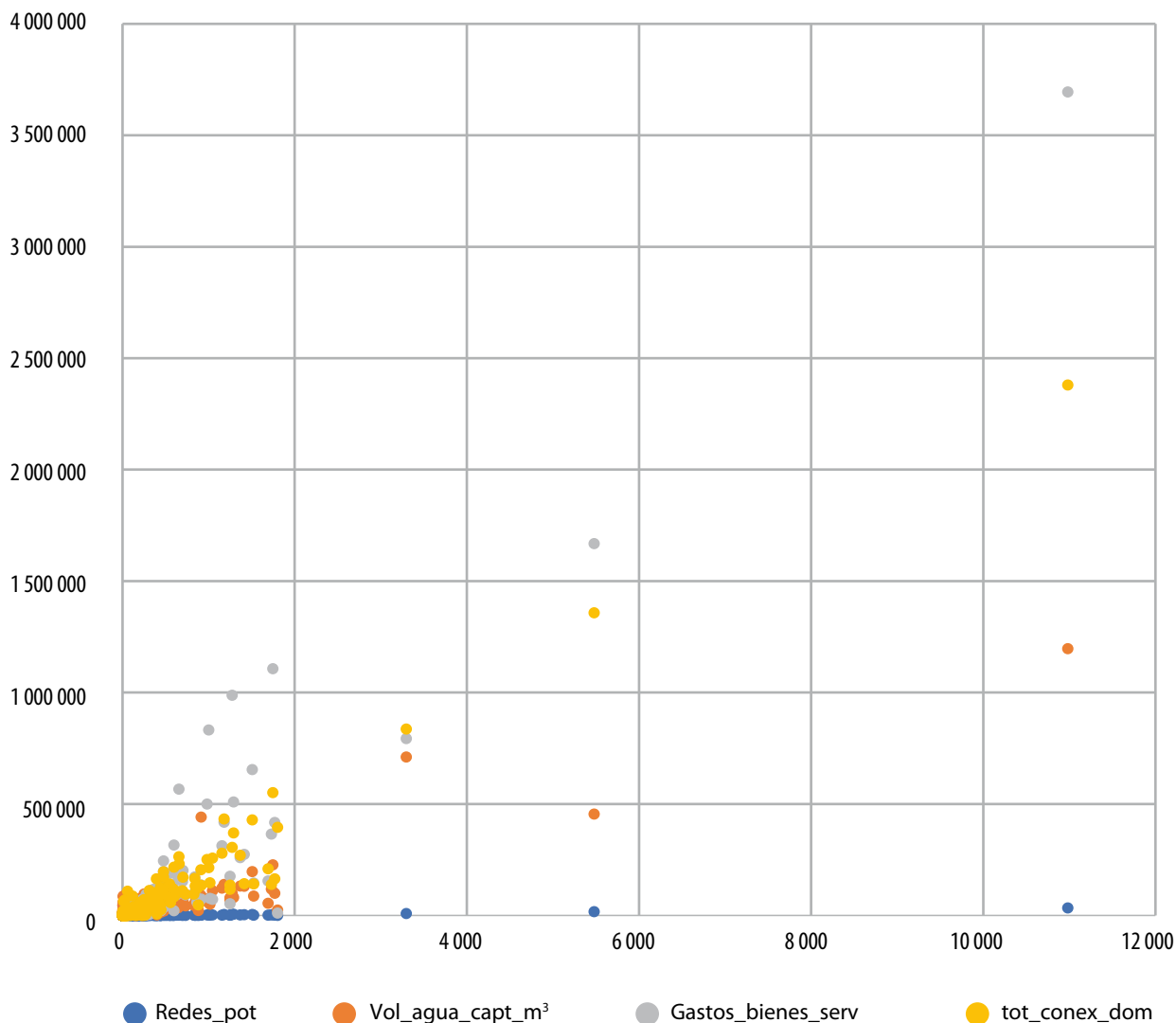
Cuadro 2

Estadísticas descriptivas de las variables seleccionadas ($n = 2\,349$)

Variable	Promedio	Std. dev.	Mín.	Máx.
Total de personal (#)	48.43	300.02	1	10 981
Longitud de redes (km)	176.79	1 207.54	1	33 180
Volumen de agua captado (miles de m ³)	5 281.43	34 274.87	1	1 197 013
Gastos en bienes y servicios (\$)	9 830.03	97 872.36	1	3 694 379
Total de conexiones domésticas (#)	9 595.85	66 452.10	1	2 379 440

Fuente: elaboración propia con datos de los CE 2014.

Gráfica 3

Principales variables del modelo OCTSA ($n = 2\ 349$)

Fuente: elaboración propia con datos de los CE 2014.

cúbicos de agua; también, se registran otros que funcionan con un solo trabajador y hasta uno con casi 11 mil empleados.

Es relevante notar que se puede esperar la presencia de errores no intencionales o, incluso, intencionales en el registro de los datos (como es conocido y por diversas razones, en México los OCTSA pueden reportar datos modificados a propósito). La presencia de valores inusualmente altos puede

observarse en la gráfica 3; mientras que el total de personal y la longitud de redes no aparecen tan dispersas, las variables volumen de agua captado, gastos en bienes y servicios y total de conexiones domésticas reportan cifras extremas.⁵

⁵ Aunque no existe una definición estándar para valores atípicos, dos aspectos que son comunes en las distintas definiciones son: que son valores extremos en la muestra y que son suficientemente extremos como para tener poca probabilidad de ocurrencia (Fox *et al.*, 2004).

Resultados del test de rendimientos a escala (200 repeticiones)

Test	Valor estadístico	<i>p-value</i>
Test # 1:		
H_0 : tecnología (<i>T</i>) es globalmente CRS	$S_1 = 0.4030$	$p = 0.0000$
H_1 : <i>T</i> es VRS		
Test # 2:		
H_0 : <i>T</i> es globalmente NIRS	$S_2 = 0.4490$	$p = 0.0000$
H_1 : <i>T</i> es VRS		

Fuente: elaboración propia con datos de los CE 2014.

El tamaño de la base de datos analizada (2 349 OCTSA) y la estimación de eficiencia con métodos no paramétricos robustos requieren de cálculos intensivos que, por fortuna, se pueden realizar mediante distintos programas estadísticos. En este trabajo se obtuvieron los cinco estimadores definidos utilizando programas de usuario que se implementan en el programa *Stata v. 15*.

Como se mencionó en la sección anterior; para este modelo se asume una orientación hacia los insumos. La segunda especificación necesaria es el supuesto de rendimientos a escala; para definir el tipo de estos utilizado en el modelo, se aplicó un test, en el cual se prueba como hipótesis nula que la tecnología es de rendimientos constantes (CRS, por sus siglas en inglés) $H_0 = \text{CRS}$. Un segundo cuestionario prueba la hipótesis nula de rendimientos de escala no crecientes (NIRS, por sus siglas en inglés) con la hipótesis alternativa (H_1) de rendimientos de escala variables (VRS, por sus siglas en inglés).⁶ Los resultados obtenidos se reportan en el cuadro 3.

El valor estadístico S_1 es el cociente entre las medidas promedio de eficiencia técnica bajo el supuesto de las tecnologías VRS y CRS. Si la hipótesis nula es cierta, entonces la distancia promedio entre las fronteras VRS y CRS es pequeña (Badunenko y Mozharovskyi, 2016); si la alternativa es verdadera, la distancia promedio entre las fronteras VRS y CRS es grande. Si en el test # 1 se rechaza la H_0 , se procede entonces al # 2; si la hipótesis nula H_0 para este test es verdadera, entonces la distancia promedio entre las fronteras VRS y NIRS es pequeña. De acuerdo con los resultados, las dos hipótesis nulas se rechazan, lo cual indica que los OCTSA operan bajo una tecnología de escala variable (VRS). Un resultado adicional de estos tests es que se clasifica a cada OCTSA como eficiente o ineficiente en su escala de operación actual. Otro dato interesante que se puede obtener es que 68.7 % de los OCTSA en México son ineficientes en escala y 31.3 % (736) operan de manera eficiente en escala.

Resultados de los cinco estimadores

La implementación del análisis no paramétrico se inició con los dos estimadores que utilizan los datos del grupo de observaciones completo de

6 Ver Simar y Wilson, 2002 y Badunenko y Mozharovskyi, 2016. Para una explicación detallada de los programas utilizados en *Stata*, ver Badunenko y Mozharovskyi, 2016 y Tauchmann, 2012.

Cuadro 4

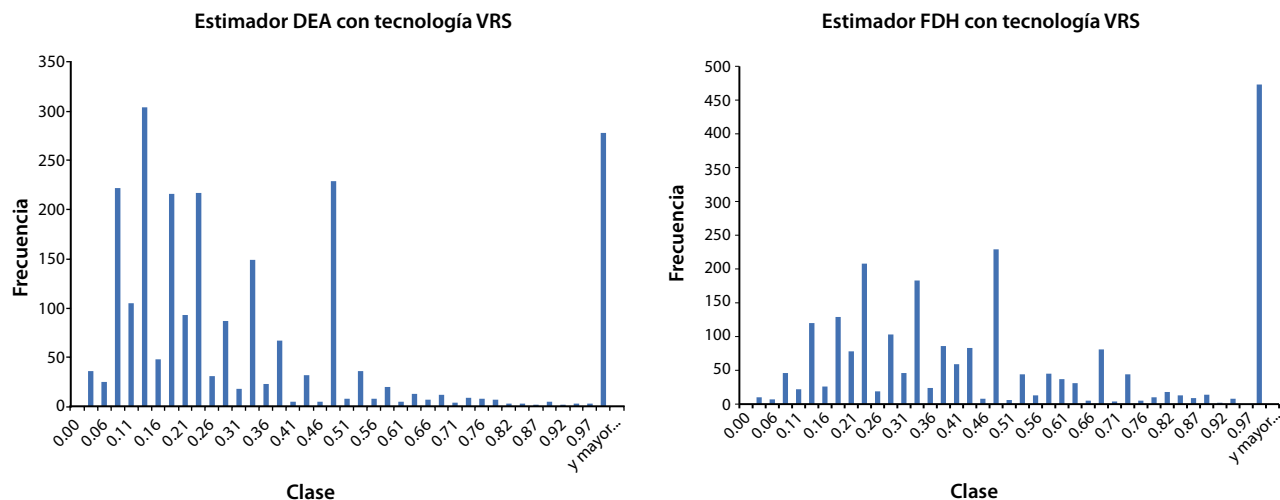
Estadísticas descriptivas de estimadores DEA y FDH ($n = 2\ 349$)

Estimador	Media	Desv. std.	Skewness	Kurtosis	OCTSA eficientes
DEA	.3492438	.2890389	1.251582	3.414062	11.7 %
FDH	.4998765	.3055849	.5508086	1.949272	19.8 %

Fuente: elaboración propia con datos de los CE 2014.

Gráfica 4

Histogramas de los estimadores de eficiencia DEA y FDH



Fuente: elaboración propia con datos de los CE 2014.

OCTSA; el DEA, que asume que todas las combinaciones de insumos teóricas son posibles y que no existen valores atípicos ni errores en los datos; y el FDH, que acepta como combinaciones de insumo factibles solo a las derivadas de las observaciones y que no existen errores de medición ni valores atípicos en los datos. Los resultados se comparan en el cuadro 4 y la gráfica 4.

Lo primero que se puede notar es que la eficiencia promedio del estimador FDH es mayor que el del DEA, lo cual es una consecuencia lógica derivada de que la frontera de eficiencia estimada con FDH se basa solo en los valores observados del grupo de OCTSA. Se encontró, también, que la proporción de unidades eficientes o en la frontera de producción (con estimadores = 1) se incrementa de 275 (11.7 %) a 465 (19.8 %). Es razonable esperar,

como lo indica la literatura, que estos dos primeros estimadores reflejen el sesgo de la presencia de valores atípicos y/o errores en los datos, lo que implica una posible sobrevaluación de la eficiencia.

El primer estimador que trata de solventar las limitaciones de los dos determinísticos (DEA y FDH) es el de orden m . Para su implementación (debido a que se basa en la evaluación de eficiencia por muestras y dado que no existe un parámetro m definido *a priori*), se prueban distintos valores de m , en un rango de 100 hasta 500; esto es, cada OCTSA fue comparado con muestras aleatorias de 100 a 500 unidades, con 200 repeticiones para obtener el estimador de eficiencia (θ). Como se explica en este trabajo, por su forma de cálculo, el estimador de orden m genera valores de eficiencia mayores a 1; a los operadores con esos valores se les deno-

mina supereficientes. Ya que tampoco hay un criterio predefinido para determinar un valor que sirva como parámetro de supereficiencia, elegimos 1.5 para que un OCTSA pueda ser considerado valor atípico. Un organismo clasificado con este valor utiliza, al menos, 50 % menos de insumos que las unidades de la submuestra m con la que se compara para producir un nivel dado de producto.

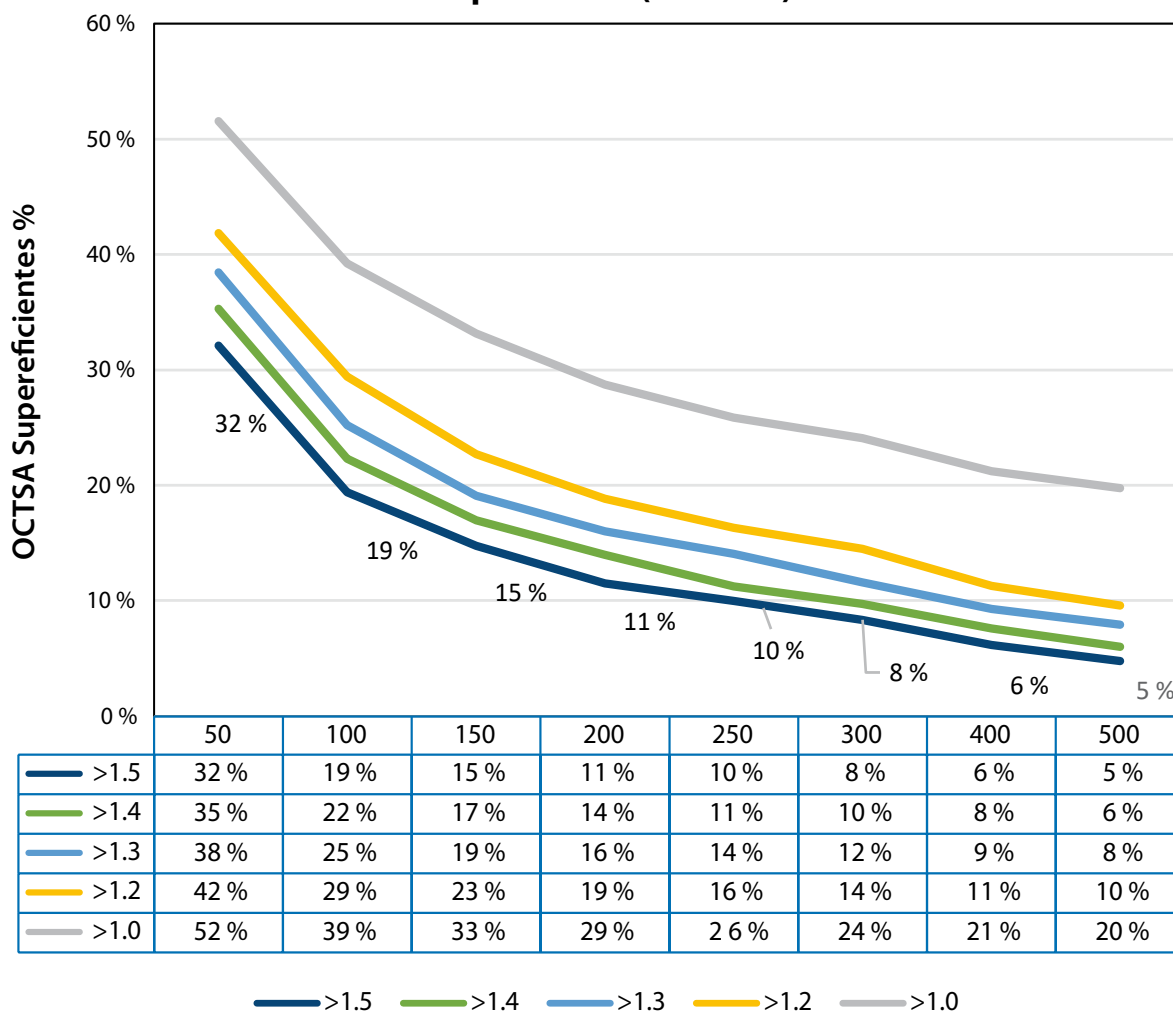
En la gráfica 5 puede observarse que con $m = 50$, la tercera parte de OCTSA resultan supereficientes (con puntuación > 1.5), mientras que con $m = 500$, la proporción se reduce a 5 por ciento. Un criterio para identificar a los organismos que son verdaderos va-

lores atípicos es su permanencia como supereficientes, aun con altos valores de m ; en este caso, al menos 5 % de operadores se mantienen como tales a altos valores de m ($m = 500$) y con un criterio extremo del estimador de eficiencia ($\theta > 1.5$).

En lugar de definir muestras con un número de OCTSA determinado (como lo hace el estimador orden m), el α utiliza cuartiles ($1 - \alpha$) como criterio para comparar la eficiencia. Con el uso de este estimador se propone evaluar el impacto de los valores atípicos y errores en la base de datos de OCTSA en la proporción de supereficiencia. Para ello, se analiza un rango de valores de α entre 0 y 1.

Gráfica 5

Proporción de OCTSA supereficientes con distintos valores de orden m y niveles de supereficiencia ($n = 2\ 349$)



Fuente: elaboración propia con datos de los CE 2014.

En la gráfica 6 se presentan los porcentajes de operadores supereficientes ($\theta > 1.5$) a diferentes valores de α . Los resultados muestran que cuando $\alpha = .80$, la proporción de estos OCTSA disminuye de manera más rápida (de 90 % a poco más de 75 %). No obstante, los valores promedio estimados de eficiencia con orden α y la proporción de OCTSA supereficientes son muy altos, esto implica que una gran proporción de organismos se ubica fuera de la frontera estimada de posibilidades de producción. En otros términos, el estimador orden α es altamente afectado por la presencia de valores atípicos o errores, los estimadores de eficiencia de orden α que resultan son sobreestimados.

Los resultados de los estimadores con fronteras parciales se pueden utilizar para identificar las unidades del grupo de operadores con valores atípicos. Para ello, definimos como criterio las puntuaciones de eficiencia > 1.5 que se generaron para cada uno de los 2 349 OCTSA de la base de datos; con los

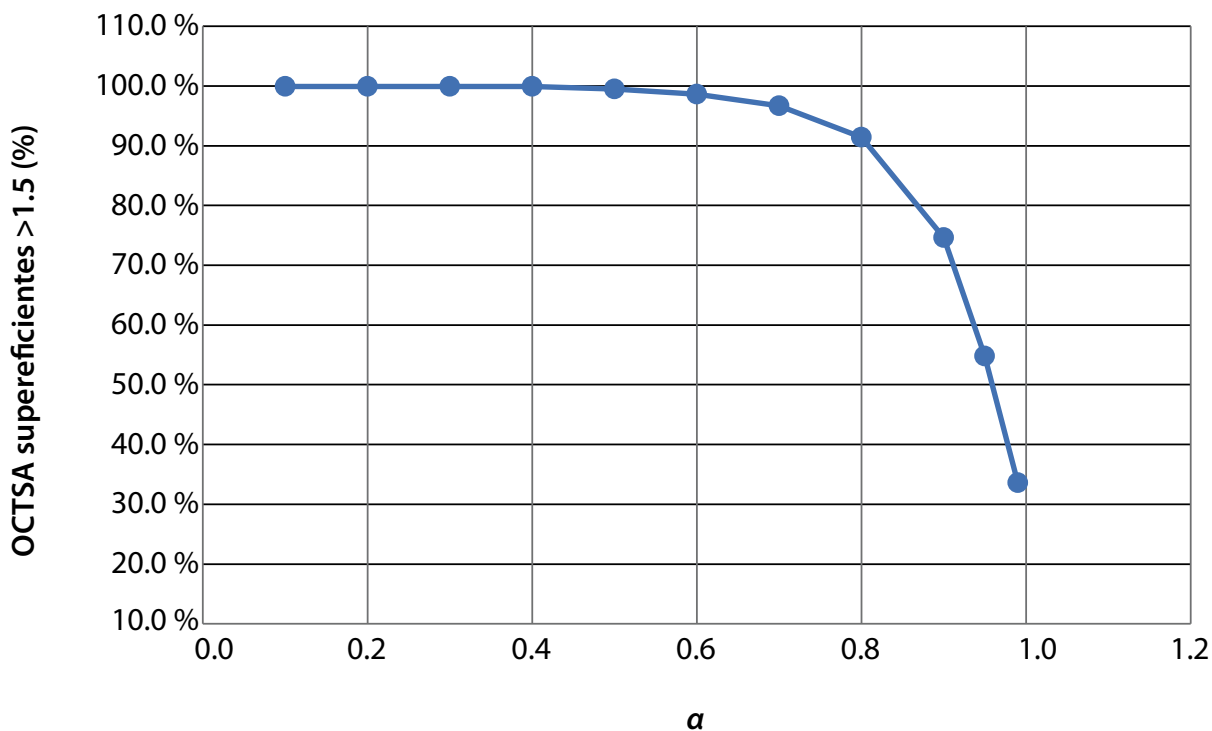
valores clave de $m = 400$ para orden m y para orden $\alpha = .80$. En este caso, se identificaron 145 OCTSA como valores atípicos.⁷ Una opción es eliminar a los atípicos y, para el resto, se podría estimar FDH, con lo que se obtendría un *ranking* de eficiencia sin el sesgo de observaciones atípicas. Esto implica, sin embargo, perder 6 % de los organismos con información completa que no podrían ser clasificados. De manera alternativa, se optó por aplicar el BC mediante la simulación de datos (*bootstrap*).

Un supuesto clave para poder definir el proceso generador de datos (necesario para la simulación de muestras que usa el estimador BC) es que las n observaciones en el grupo de datos analizado (i. e. estimaciones de eficiencia) sean variables aleato-

⁷ Una limitación de la fuente de información es que debido a que procede de los CE aplicados por el INEGI, los datos se manejan observando los principios de confidencialidad y reserva, por lo que no se cuenta con el nombre oficial de cada OCTSA, pero sí se tiene un identificador, por lo que es posible comparar el resultado individual para cada estimador.

Gráfica 6

OCTSA supereficientes ($\theta > 1.5$) con diferentes valores de α ($n = 2\ 349$)



Fuente: elaboración propia con datos de los CE 2014.

Cuadro 5

Estimaciones BC para el grupo de OCTSA con valores atípicos

Variable	Obs.	Mean	Std. dev.	Mín.	Máx.
Eficiencia BC	145	0.69	0.23	0.21	1

Fuente: elaboración propia.

Cuadro 6

Estimadores BC por grupo de tamaños de localidades atendidas

	<= 2 500 hab.	2 500-20 mil hab.	20 001-50 mil hab.	50 001-1 millón de hab.	> 1 millón	Total
Número	798	1 132	210	199	10	2 349
BC prom.	0.43	0.27	0.17	0.23	0.50	.31
Eficientes (#)	125	57	0	0	0	182
Eficientes (%)	15.6	5.03	0	0	0	7.7 %

Fuente: elaboración propia con datos de los CE 2014.

rias independientes e idénticamente distribuidas (i. i. d.) en el conjunto convexo posible. Por lo tanto, y como paso previo requerido para la estimación del BC, se realizó la prueba de independencia de las estimaciones de eficiencia (con 999 repeticiones y 95 % de nivel de confianza). El resultado de la prueba confirmó que las estimaciones de eficiencia son variables aleatorias independientes con un valor $p = 0.0000$.⁸ Para la estimación BC, se utilizó el método *bootstrap* homogéneo suavizado con 999 repeticiones (para un resumen de sus resultados, ver cuadro 5). La eficiencia promedio es 0.31, con 182 OCTSA clasificados como eficientes. Un dato notable es que la proporción global de OCTSA eficientes (7.7 %) disminuye de manera destacada cuando se compara con los estimadores determinísticos DEA (11.7 %) y FDH (19.8 %).

La corrección del sesgo que logra el estimador BC se observa en el hecho de que la eficiencia pro-

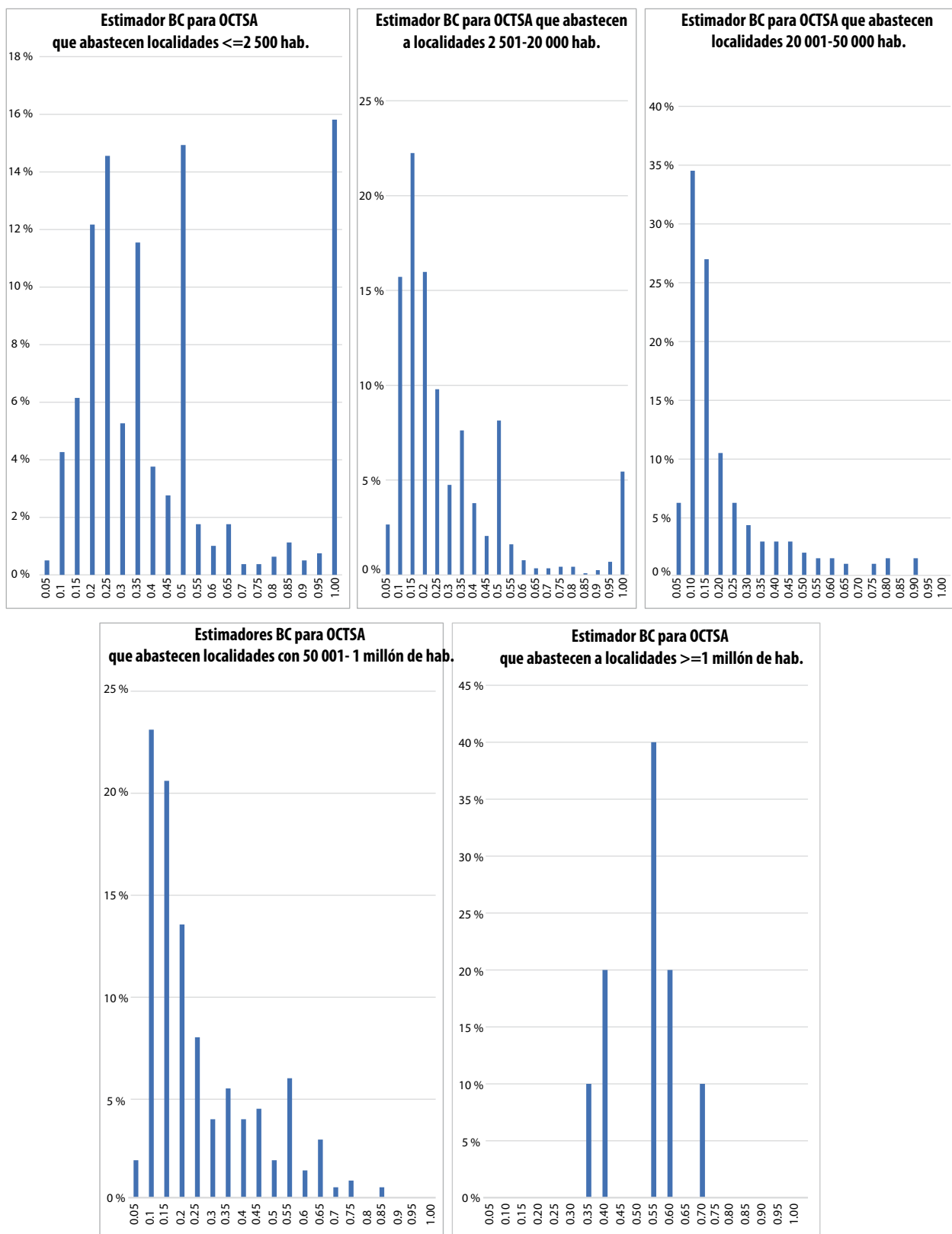
medio estimada de los 145 OCTSA identificados como valores atípicos por orden m y orden α fue de 0.69. Este es un claro resultado de la corrección por sesgo de valores atípicos que logra este estimador y la principal ventaja que aporta con respecto a los otros cuatro que, como se observó con los resultados obtenidos, son afectados por la presencia de valores atípicos, manteniendo altas proporciones de operadores supereficientes o que se encuentran más allá de la frontera de producción estimada.

Una vez que se confirma la corrección de sesgo lograda por los estimadores BC, podemos hacer diferentes análisis de los valores de eficiencia estimados. Un aspecto de interés particular en este trabajo es comparar la eficiencia técnica de los OCTSA por tamaño de población atendida. Para ello, se organiza la información por tamaño de localidad atendida en cinco secciones: menores a 2 500 (consideradas como rurales), entre 2 500 y 20 mil; entre 20 001 y 50 mil; entre 50 mil y 1 millón y mayores a 1 millón de habitantes. El cuadro 6 muestra que, en promedio, el grupo de OCTSA con mayor población atendida (más de

⁸ El comando utilizado para estimar indicadores BC (*teradialbc*) proporciona inferencia estadística para tres tipos de *bootstrap*: 1) homogéneo suavizado, 2) heterogéneo suavizado y 3) submuestreo (heterogéneo). En este caso aplicamos el tipo 1, siguiendo los resultados del test de independencia. El resultado de la prueba de independencia indicó que se puede usar *bootstrap* homogéneo.

Gráfica 7

Histogramas de estimadores BC por tamaño de OCTSA



Fuente: elaboración propia con datos de los CE 2014.

1 millón de habitantes) registra el mejor promedio de eficiencia (0.50); este resultado era de esperarse, dado que las ciudades y zonas metropolitanas con mayor población son las que pueden captar mayores ingresos y reciben, también, inversiones más elevadas, así como financiamiento público. Sin embargo, debe notarse que ninguno de los 10 operadores en localidades mayores a 1 millón de habitantes es 100 % eficiente (i. e. en uno o más de los insumos podría mejorar).

Un resultado también interesante es que, en contraste con los OCTSA de áreas urbanas grandes, 125 de localidades rurales se reportan como 100 % eficientes; esta cifra puede parecer contraintuitiva pues, como se sabe, la cobertura de servicios de agua y saneamiento en zonas rurales de México es baja y, por lo general, de mala calidad. Sin embargo, debe tomarse en cuenta que este análisis se refiere a la eficiencia técnica, entendida como las mejores combinaciones de insumos para dar servicio a un número dado de conexiones; no se analizan aspectos de calidad o cobertura. Por lo tanto, este resultado puede explicarse como consecuencia de que los organismos muy pequeños cuentan con los insumos en cantidades mínimas tanto en gastos, personal, volumen captado y redes, por lo que están obligados a optimizarlos.

En la gráfica 7 se muestran los porcentajes de OCTSA por niveles estimados de eficiencia con BC. Como puede verse, en los cuatro grupos menores a 1 millón de habitantes registran bajas eficiencias técnicas (por debajo de 0.5). Otro aspecto importante es que las distribuciones de los estimadores de eficiencia de los operadores de localidades más pequeñas (menos de 2 500 habitantes) y las más grandes (mayores de 1 millón) son los que difieren más en su distribución con respecto a los otros tres grupos.

Conclusiones

En este trabajo se muestra el potencial que la base de datos de los Censos Económicos en México tiene para el análisis de eficiencia técnica en los servicios del agua. La aplicación de cinco estimadores

no paramétricos nos permite concluir que el BC reduce de forma significativa el sesgo generado por la presencia de valores atípicos y errores de medición o reporte en los servicios del agua. Como era de esperar, los estimadores determinísticos (DEA y FDH) sobreestiman el número de unidades eficientes como consecuencia de la presencia de valores atípicos. Por otra parte, los resultados de los estimadores basados en fronteras parciales (orden m y orden a) permitieron identificar 145 OCTSA supereficientes, los cuales o bien son realmente muy eficientes o esa supereficiencia se debe a datos erróneos registrados de forma voluntaria o involuntaria. En todo caso, ambos resultados sugieren la necesidad de mejorar el proceso de validación de datos.

Una observación relevante para la política pública se deriva de la prueba de rendimientos de escala; se observa que, al menos la tercera parte de los OCTSA en México funcionan en la actualidad a una escala ineficiente; una vez identificados esos organismos pueden establecerse medidas operativas tendientes a ajustar su escala al tamaño apropiado en aspectos incluidos en el análisis, como gastos en bienes y servicios, así como la cantidad de personal. Estos aspectos específicos hacen necesaria, también, una mejor regulación y la implementación de mecanismos de profesionalización de los servicios de agua y saneamiento en México.

Los resultados de los estimadores corregidos por sesgo confirman la percepción usual de que los OCTSA más grandes (mayores a 1 millón) son, en general, el grupo más eficiente (siete de los 10 registran un estimador mayor a 0.50); sin embargo, ninguno de ellos (cuando se analizan con el estimador corregido por sesgo) es considerado 100 % eficiente. Este dato implica consecuencias importantes pues, aunque los organismos de las áreas urbanas con mayor población son usualmente los que registran mejores indicadores de desempeño tradicionales (p. e. eficiencia comercial o buena capacidad de cobrar lo facturado; o eficiencia física, un menor volumen de agua no facturada), el hecho de que no sean eficientes en su totalidad sugiere que se están desperdiciando recursos. En

específico, esto significa que en los operadores de las grandes áreas urbanas en México no se hace un uso eficiente de algunos de los insumos incluidos en este análisis (volumen de agua captado, gastos, personal o redes) con los consecuentes impactos ambientales y costos económicos que esto trae.

Los resultados también informan que 15 % de los OCTSA rurales es 100 % eficiente. Esto es importante, pues sugiere que los organismos de las localidades rurales registran buenas prácticas relacionadas con el uso de las variables incluidas en este modelo (gastos en bienes y servicios, total de personal, volumen captado de agua y longitud de redes). Sin embargo, se debe tener cuidado con esto, pues no implica que se logre buena calidad de los servicios o cobertura adecuada en localidades rurales. En general, esa es una limitación del análisis hecho en este trabajo, derivado sobre todo de la falta de indicadores apropiados y compatibles con la base de datos utilizada de calidad del agua, calidad y continuidad en los servicios y aquellos relacionados con el derecho humano al agua, como equidad y asequibilidad en las tarifas.

Por otro lado, las estimaciones de eficiencia corregidas por sesgo logradas con el estimador BC permiten la posibilidad de otro tipo de análisis; por ejemplo, el efecto de variables externas (fuentes de agua, nivel de subsidio, disponibilidad del líquido y participación privada, entre otras) en el nivel de eficiencia técnica. Estos temas se proponen para una agenda futura de investigación.

Conocer la eficiencia técnica de los OCTSA del país tiene diferentes implicaciones para los diversos niveles de administración de los servicios del agua. En el ámbito local, un operador puede identificar un referente de eficiencia (por ejemplo, dentro del mismo estado) para mejorar sus prácticas. Para instancias por entidades, como las comisiones estatales del agua, puede utilizarse como instrumento para identificar las mejores prácticas de los OCTSA eficientes o estrategias para la profesionalización de los organismos clasificados como ineficientes. A nivel federal, podría mejorar los criterios para la asignación de recursos públicos a los servicios

del vital líquido. En México, la política enfocada a mejorar la cobertura de servicios básicos de agua se aplica a través del Programa de Agua Potable, Drenaje y Tratamiento (PROAGUA), en el cual se parte, como primer criterio de priorización, de la Declaratoria de las Zonas de Atención Prioritaria, determinadas para municipios y localidades con elevados porcentajes de población en situación de pobreza y con alto grado de rezago social. Otros son el número de habitantes, acciones que corresponden a recomendaciones de derechos humanos y actividades que contribuyen al ahorro de energía. Se podría asegurar un mayor beneficio si se pondera de forma positiva a los OCTSA que muestran una mayor eficiencia en el uso de los recursos, tanto financieros (p. e. el manejo en gastos en bienes e insumos) como del agua (eficiencia en uso de volúmenes captados) y personal empleado. De otra manera, los destinados a localidades con rezago social, pero donde operan organismos ineficientes podría no mejorar la cobertura de los servicios y mucho menos otros aspectos como la calidad y asequibilidad de los mismos. Asegurar un mejor manejo del recurso e insumos mediante una mejor regulación, la profesionalización de los servicios y una mayor transparencia contribuirán a lograr el objetivo de cumplir con los derechos humanos a agua potable y saneamiento en el país.

Fuentes

- Aguilar-Benítez I. "Viabilidad financiera de los servicios del agua: una comparación de tres pares de ciudades en la frontera México-Estados Unidos", en: Aguilar-Benítez, I. (coord.). *Los servicios del agua en el norte de México: gestión, manejo financiero y aspectos socio-ambientales*. Tijuana, Colef/Colson, 2011.
- Anwandter, L., & T. Jr. Ozuna. "Can public sector reforms improve the efficiency of public water utilities?", in: *Environment and Development Economics*. 7(04), 2002. doi:10.1017/s1355770x02000414.
- Badunenko, Oleg and Pavlo Mozharovskyi. "Nonparametric Frontier Analysis Using Stata", in: *The Stata Journal*. Vol. 16, no. 3, Sept. 2016, pp. 550-589, doi:10.1177/1536867X1601600302.
- Banker, R. D., A. Charnes & W. W. Cooper. "Some models for estimating technical and scale inefficiencies in data envelopment analysis", in: *Management Science*. 30(9), 1984, pp 1078-1092.

- Bowlin, W. "Measuring Performance: An Introduction to Data Envelopment Analysis (DEA)", in: *The Journal of Cost Analysis*. 15:2, 1998, pp. 3-27, doi: 10.1080/08823871.1998.10462318.
- Cazals, C., Florens and Simar. "Nonparametric frontier estimation: a robust approach", in: *Journal of Econometrics*. 106, (1), 2002, pp. 1-25.
- Charnes, A., W. W. Cooper y E. Rhodes. "Measuring the efficiency of decision making units", in: *European Journal of Operational Research*. 1978, pp. 429-444.
- Cooper W., L. M. Seiford y Zhu. "Handbook on Data Envelopment Analysis", in: *International Series in Operations Research & Management Science*. Second edition. Volume 164, 2011.
- Cooper W., L. M. Seiford y K. Tone. *Introduction to Data Envelopment Analysis and its Uses. With DEA-Solver Software and References*. Springer, 2006.
- Daraio, Cinzia & Léopold Simar. *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. Springer, 2007. doi:10.1007/978-0-387-35231-2.
- Deprins, D. and H Tulkens. "Measuring Labour Efficiency in Post Offices", in: Marchand, M. and H. Tulkens (eds.). *The Performance of Public Enterprises: Concepts and Measurement North-Holland*. 1984, pp. 243-267.
- Daouia, A., & L. Simar. "Nonparametric efficiency analysis: A multivariate conditional quantile approach", in: *Journal of Econometrics*. 140(2), 2007, pp. 375-400. doi:10.1016/j.jeconom.2006.07.002.
- Fox, K., R. Hill & W. Diewert. "Identifying Outliers in Multi-Output Models", in: *Journal of Productivity Analysis*. 22(1/2), 2004, pp. 73-94.
- Kneip, A., L. Simar & P. Wilson. "Asymptotics and Consistent Bootstraps for DEA Estimators in Nonparametric Frontier Models", in: *Econometric Theory*. 24(6), 2008, pp. 1663-1697.
- Marques, R. C. & K. de Witte. "Is big better? On scale and scope economies in the Portuguese water sector", in: *Economic Modelling*. 28(3), 2011, pp. 1009-1016. doi:10.1016/j.econmod.2010.11.014.
- Picazo-Tadeo, A. J., F. González-Gómez & F. J. Sáez-Fernández. "Accounting for operating environments in measuring water utilities' managerial efficiency", in: *The Service Industries Journal*. 29(6), 2009, pp. 761-773. doi:10.1080/02642060802190300.
- Simar, L. and P. W. Wilson. "Non-parametric tests of returns to scale", in: *European Journal of Operational Research*. 139(1), 2002, pp. 115-132.
- Tauchmann, Harald. "Partial Frontier Efficiency Analysis", in: *The Stata Journal*. Vol. 12, no. 3, Sept. 2012, pp. 461-478, doi:10.1177/1536867X1201200309.

Reformar la educación: ¿misión imposible?

Los consejos y experiencias de los líderes educativos

*Reforming Education: Impossible Mission?
Sharing Advises and Experiences from Educational Leaders*

Reimers, Fernando M.
*Letters to a New Minister of Education. Learning
to Lead Education Systems.*

Cambridge, MA, Independently Published, 2019, 203 pp.

Reseña

Otto Granados Roldán*

Probablemente, desde la década de los 90 del siglo pasado, en buena parte del mundo (en especial en los países emergentes) se han emprendido reformas educativas con diversos diseños, alcances y perfiles, pero la mayor parte ha sido con el objetivo común de querer mejorar la calidad de la educación que se provee a niños y jóvenes, prepararlos para competir mejor y elevar los niveles de equidad e inclusión. Desde el punto de vista teórico, la formulación e instrumentación de las políticas adecuadas que produzcan ese resultado parece no tener mayor controversia —en la medida de que se trata de un bien público compartido—, se cuen-

ta con mayor información y evidencia de lo que sí funciona y de lo que no, la investigación especializada es creciente y los padres de familia, aun con reacciones ambiguas, asumen que la educación es buena para la vida de sus hijos.

Sin embargo, sobre todo en América Latina y el Caribe, los avances para lograr una educación de calidad —medidos a través de las distintas evaluaciones internacionales, del impacto real sobre las tasas de crecimiento económico o del incremento de la innovación y la productividad— han sido escasos, lentos y, en ciertos países, lo suficientemente de-

* Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura y Universidad de Harvard, otto.granados@oei.es

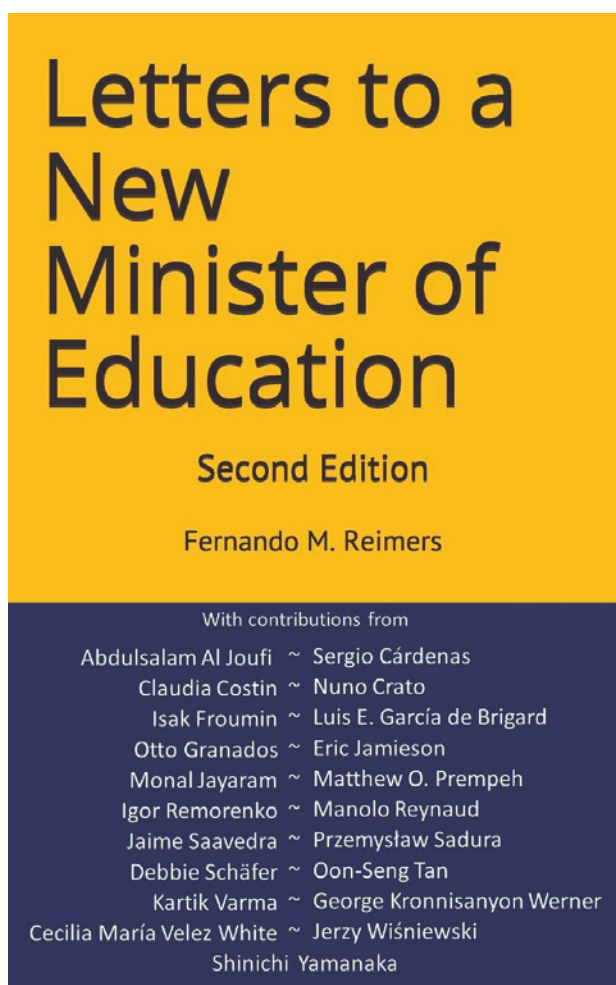
cepcionantes como para llegar a la hipótesis de que la educación es irreformable. Pero, ¿lo es en realidad? Veamos.

Con el cambio de siglo empezaron a publicarse distintos estudios que pretendían explicar las condiciones que hacen posible o dificultan hacer reformas educativas de gran calado. Destacan, por ejemplo, los trabajos de Javier Corrales (*Aspectos políticos en la implementación de las reformas educativas*, Santiago de Chile, PREAL, 1999), Robert R. Kaufman y Joan M. Nelson (*Crucial Needs, Weak Incentives. Social Sector Reform, Democratization and Globalization in Latin America*, Baltimore/Washington, D. C., Johns Hopkins University Press/Woodrow Wilson Center Press, 2004), Merilee S. Grindle (*Despite the Odds: The Contentious Politics of Education Reform*, Princeton University Press, 2004), Michael Barber y Mona Mourshed (*How the*

World's Best-Performing School Systems Come Out On Top, McKinsey & Co., 2007) y M. Mourshed, Ch. Chijioke y M. Barber (*How the world's most improved school systems keep getting better*, McKinsey and Co., 2010), quienes, desde una perspectiva internacional comparada, identificaron una serie de variables que influyen en la instrumentación de las políticas y las reformas educativas.

En términos generales, ese esfuerzo analítico mostró que, normalmente, existen dos tipos de reformas. Uno es el llamado *de acceso o de primera generación*, las cuales son las que emprenden los países de desarrollo muy incipiente y van, en esencia, dirigidas a responder a demandas de cobertura: satisfacen necesidades básicas de escolarización; su ejecución depende de una buena distribución de insumos y bastan presupuestos adecuados y una organización eficiente; tienen una visibilidad positiva, concreta y rápida; se reproducen a sí mismas; crean anticuerpos eficientes; y, con el tiempo, generan incentivos inerciales. El segundo, más propias en naciones de desarrollo medio, son las de carácter *sistémico y estructural* y van orientadas, sobre todo, a la calidad: introducen cambios relevantes en ecosistemas complejos habituados a la estabilidad y con roles asignados para los diversos agentes; por lo normal, alteran —y de allí deriva su complejidad— los arreglos políticos, económicos e institucionales ensamblados por largo tiempo; afectan intereses creados de todo tipo; y generan desafíos de sinergia y coordinación entre agencias públicas y lógicas encontradas, así como tensiones y conflictos a corto plazo y resultados a largo plazo. En el balance, estas últimas son muy difíciles de instrumentar desde el punto de vista técnico, político, institucional y administrativo; pero, no obstante, se estima que entre 1977 y el 2000 se hicieron, al menos, 50 modificaciones de ese tipo en unos 30 países, sin contar algunas a nivel subnacional.

¿Cuáles fueron sus resultados? Por un lado, las reformas de primera generación (aunque con variaciones de país a país) arrojaron algunos logros importantes en cobertura. Por ejemplo, en América Latina, hasta finales del siglo pasado, se redujeron los niveles de analfabetismo de manera significa-



tiva, comparados con los de otras regiones, con excepción de Asia Oriental; las tasas netas de matrícula en educación básica fueron relativamente altas (aun cuando Brasil, República Dominicana, El Salvador y Nicaragua todavía permanecían muy a la zaga de otras naciones como Chile, Colombia, Costa Rica, Cuba y México) y los sistemas funcionaron de forma razonable para hacer más accesible el servicio educativo. Sin embargo, el historial del subcontinente en cuanto a educación de calidad fue menos impresionante. Las tasas de repetición eran —y en algunos casos siguen siendo— muy altas y más todavía en los sectores rurales de la región; a comienzos de la década de los 90, un niño latinoamericano permanecía en la escuela por espacio de siete años (hoy son nueve), aunque el verdadero nivel de escolarización alcanzado llegaba apenas al cuarto grado de educación básica; las pruebas nacionales e internacionales mostraban un bajo nivel de logro académico, así como una mala calidad de la educación entregada, y los pobres tenían menos probabilidad que los ricos de concluir la educación primaria y secundaria.

Al analizar las razones que explican este recorrido, la mayor parte de la literatura académica encontró que el destino de las reformas ejecutadas en esas dos décadas dependió, en muy buena medida, de siete factores: 1) la naturaleza del régimen político de cada país; 2) el hecho de que toda reforma estructural arroja costos y perdedores en sectores muy concentrados (sindicatos, por ejemplo) y, en cambio, la percepción de los beneficios es difusa, dispersa y gradual entre la población abierta; 3) los presidentes muestran un compromiso escaso y volátil con las reformas, sobre todo si les dan demasiados problemas o surgen otras prioridades; 4) el empuje ciudadano en un sentido real (no el de las élites académicas) es bajo; 5) los conflictos con los grupos afectados por las reformas pueden derivar en un poder de veto que las nulifica; 6) los logros de las reformas, a nivel nacional, solo se concretan y miden a mediano y largo plazos; y 7) los cambios de gobierno y la alta inestabilidad en el liderazgo de los ministerios o secretarías de Educación interrumpen la necesaria continuidad de las reformas.

Con esa perspectiva, Fernando M. Reimers, profesor de la Práctica en Educación Internacional en la *Ford Foundation* y director del Programa de Política Educativa Internacional en la Escuela de Graduados en Educación de Harvard, reunió a un grupo de 21 especialistas que han sido (o lo son, a la fecha) secretarios de Educación, viceministros o funcionarios de alto nivel en 15 países de todos los continentes para que contribuyeran con una reflexión sobre su paso por los ministerios liderando reformas educativas ambiciosas y compartieran su experiencia viva pensando en quienes podrían ocupar esa misma responsabilidad ahora o en el futuro.

Su libro, que va en su segunda edición y fue originalmente publicado en inglés (*Letters to a New Minister of Education. Learning to Lead Education Systems*), se ha traducido al español, portugués y árabe, y es un ejercicio que tiene como propósito (no planteado de manera explícita, pero subyacente a lo largo de todas sus páginas) abordar la frecuente tensión que existe entre *policy* y *politics* o, en otras palabras, discutir cómo el diseño, la formulación y la ejecución de las políticas y las reformas educativas —es decir, el *ser*— pueden chocar con frecuencia con el mundo ideal del *deber ser*, y cómo evitar quedarse atrapado, desde el liderazgo de un ministerio de Educación, en ese aparente callejón sin salida para hacer viable un conjunto ambicioso de cambios que permitan conducir a una educación de calidad.

Como dice Reimers en la introducción: "... lo primero que un ministro entrante ha de hacer para «dar sentido» al desorden existente será «dar sentido» a su trabajo y a los recursos de que dispone para hacerlo. Para «dar sentido al desorden», primero hay que convertir los distintos desafíos relacionados entre sí en una serie de problemas organizados y priorizados y, posteriormente, en una estrategia sobre la que se pueda actuar [... y] cómo influir en esos factores políticos y organizativos. Sin duda alguna, para «dar sentido al desorden», es útil hablar con el mayor número posible de personas bien informadas...". Esa fue, justo, la idea de este libro.

Una primera *lección* que arrojan las reflexiones es que no hay lecciones preconcebidas ni recetas únicas o válidas para todos, sino que las características de las reformas están muy asociadas, en cada país, con el tipo de régimen político, la cultura cívica, el nivel de desarrollo, el grado de institucionalidad y la madurez de la sociedad civil, entre otros aspectos. Por ejemplo, las cartas referidas a los casos de Colombia, México o Perú son reveladoras de algunas de las limitaciones observadas en las reformas que la región hizo en la década de los 90, es decir, la poderosa influencia de la política partidista o, dicho de otra forma, la dificultad de lograr avances en contextos democráticos de mediano desarrollo, sistemas institucionales frágiles y ciudadanías de baja intensidad en el sentido en el que las define Guillermo O'Donnell, lo que, en buena medida, impide que la educación sea vista como una política de Estado, esto es, una que supone asegurar una planeación y ejecución de largo plazo de forma razonable, independiente de los reemplazos abruptos de partidos y gobiernos. "Comencé a ver con preocupación —comparte Luis García de Brigard, que fue viceministro en Colombia— que parte de nuestras reformas más emblemáticas estaban perdiendo su rumbo o los nuevos gobiernos las suspendían por completo [...] El vacío no tardó en llenarlo el cinismo: el juego político sucio, la burocracia, la incompetencia y la corrupción se tornaron en excusas fáciles para racionalizar el fracaso del sistema y mi decisión de abandonarlo...". En cambio, como se advierte en los textos de Australia o Japón, las reflexiones van mucho más en la dirección de cómo hacer los cambios de sustancia, es decir, el modelo curricular, la evaluación de los docentes, la preparación para el futuro o el dinero requerido, todas ellas cuestiones pertinentes en un ambiente público estable y predecible como el de ambos países.

Un segundo rasgo del libro es que sugiere que, en todas partes, ser secretario o ministro de Educación es un oficio de equilibrista: se avanza poco, despacio, en la cuerda floja, sin red de protección y nunca sabe si se camina, decía Alfred de Musset, sobre cenizas o semillas. Por ejemplo, entre 1988 y 1993 hubo nueve ministros en Perú; seis en Brasil; cinco en Colombia y Venezuela; cuatro

en Chile; y tres en Argentina, España y México, en cada uno; del 2000 al 2014, en Francia hubo ocho y en Japón, 12. En tiempos más recientes, la estabilidad no ha cambiado mucho en nuestro continente: entre el 2000 y 2015, el periodo promedio de duración de un responsable educativo en América Latina y el Caribe fue de dos años y un mes;¹ y entre el 2019 y septiembre del 2020, Brasil ha tenido cinco ministros de Educación; en la región aparecen solo tres excepciones: una ministra en Colombia que duró ocho años; la actual de Nicaragua, que ha permanecido a la fecha 10 años en la posición; y el de Bolivia, que estuvo 11 años en el cargo hasta la caída de Evo Morales. Por lo tanto, en las reflexiones que podría compartir cualquiera que ha estado por años en la función pública (ya sea desde la política o desde la Academia asesorando a *policy makers*) con quienes le suceden en un cargo tan relevante como ser ministro, siempre hay una mezcla —confusa de manera inevitable— de logros que defender, experiencias vividas, aprendizajes valiosos, lecciones aprendidas y, desde luego, frustraciones y desencantos.

Una tercera consideración de los trabajos, clave para comprender el aspecto político de la ejecución de las reformas, es que la evolución de los sistemas educativos nacionales y, por consecuencia, sus avances y retrocesos, se encuentra determinada y condicionada por sus sistemas de gobernanza, la economía política, la forma en la que está estructurada la gestión pública y por esa percepción, un poco binaria por las circunstancias propias de cada país, de que el terreno es muy propicio, hay grandes oportunidades y está todo por hacerse, o bien de que el sistema educativo está tan mal, tan corrompido o tan esclerótico que es casi imposible reformarlo y progresar, y entonces hay que moverse en medio de esa lógica contradictoria y, a veces, extrema. Aquí destaca una de las interrogantes que el libro sugiere, en virtud de que las reformas estructurales normalmente conllevan cierto grado de conflicto: ¿son menos complejas de forma operativa en regímenes menos democráticos? Hay al-

¹ Comisión para la Educación de Calidad para Todos. *Construyendo una educación de calidad: un pacto con el futuro de América Latina*. Buenos Aires, Inter American Dialogue/Fundación Santillana, 2016.

gunos países incluidos en las *Cartas*... cuyos casos califican, en el *Democracy Index 2019* del semanario *The Economist*, como plenamente democráticos (Australia y Portugal); otros tienen democracias *defectuosas* (Singapur, Polonia, India, México, Perú, Colombia, Sudáfrica y Ghana), regímenes *híbridos* (Liberia) o, de plano, *autoritarios* (Rusia y Yemen), y todos ellos, en mayor o menor medida, parecen haber alcanzado algunos logros educativos. Visto de otra forma, ¿estos ejemplos hacen suponer que los progresos son viables con independencia del régimen político? La respuesta más clara es que no hay un denominador común y se tiene que estudiar cada caso por separado.

Por otro lado, a pesar de las diferencias en las meditaciones de cada líder, muchos de los problemas y retos que describen —más de los que uno puede imaginarse— son muy similares (comunicación insuficiente, resistencias a las evaluaciones de los docentes, demandas salariales exageradas y sistemáticas, limitaciones presupuestales, diferencias políticas dentro de los gobiernos, etc.) y, por lo tanto, el ejercicio de diseñar y formular políticas eficaces a partir de otras experiencias y prácticas en sistemas en apariencia tan distintos arroja pistas e ideas para mejorar las propias.

Una cuarta conclusión, muy evidente sobre todo en los testimonios de los líderes de cuatro nuevos países incluidos en la segunda edición del libro (Ghana, Liberia, Sudáfrica y Yemen, que vienen de una situación de muy bajo desarrollo o de guerras civiles) es que ejemplifican muy bien las menores complicaciones cuando se trata de *reformas de acceso* y, en buena medida, ofrecen lecciones para algunos países de Centroamérica y el Caribe. En su carta, el actual ministro de Educación de Ghana parte de una lógica impecable: “Un hospital no continuaría existiendo si el 70 % de sus pacientes muriesen. Se convertiría rápidamente en un edificio vacío si los pacientes decidieran buscar tratamiento en otro lugar, porque los organismos reguladores harían todo lo posible para cerrarlo. ¿Por qué entonces una escuela, un distrito, una región y un país continúan tolerando malos resultados de aprendizaje sin hacer las preguntas necesarias

cuando nuestro sistema escolar falla constantemente a aproximadamente el 70 % de nuestros estudiantes?”. A partir de allí, este ministro (que es un *outsider* en el ámbito educativo de su país porque venía de la medicina y la salud pública) relata que se enfocó en una agenda que es de sentido común: desarrollo curricular, evaluaciones, materiales didácticos modernos, una política integral para la carrera docente, buen liderazgo escolar y “... una gestión adecuada que inspire esperanza entre los estudiantes...”. En suma, nada nuevo pero que, en su simplicidad, constituye un mapa de navegación viable y eficaz.

Un quinto aspecto subyacente en varios textos (Colombia, México y Ghana) es estrictamente político. Por un lado, un ministro de Educación no es un funcionario poderoso como lo son, por ejemplo, quienes manejan el presupuesto o la seguridad, o bien los consejeros áulicos del presidente en turno. Su poder es vicario; su utilidad, desechable y, en el mejor de los casos, su fuerza deriva, si acaso, de que por momentos gane cierta autoridad moral y, eventualmente, política. Y nada más. Así que, a diferencia de los mercados económicos o financieros, donde es normal que la inversión se use de manera cuidadosa para que crezca y las personas promedio tienden a mostrar una prudente aversión al riesgo, en política es al contrario: se debe usar el capital de que se dispone al principio de la gestión, que tal vez sea alto, para tomar decisiones difíciles e impopulares, pero que son las más eficaces con frecuencia; desde luego, hay otro camino: tomar decisiones populares, pero que son, por lo general, malas para todo efecto práctico.

La razón es simple: el poder, normalmente muy relativo en el caso de un ministro de Educación, desgasta rápido y las circunstancias son siempre inciertas y cambiantes. De las docenas de reformas educativas de todo tipo que se han instrumentado en los últimos 50 años en numerosos países, son muy raras aquellas que no han provocado tensiones y conflictos porque, sencillamente, casi todas tocan intereses creados (en ocasiones, muy corruptos) de sindicatos, burocracias, partidos, legisladores y organizaciones civiles. Es más, reformas

sin conflicto, no son reformas de fondo. Y por otro lado, como lo identifican bien varios ministros o exministros, hay ciclos políticos inevitables en los gobiernos y de ello depende la mayor o menor fortaleza para impulsar reformas. Esto quiere decir, aunque haya una íntima renuencia a aceptarlo, que para ser un buen ministro de Educación es mejor no tener muchas aspiraciones políticas personales. Descartarlas puede dar una ventaja para que los compañeros de gabinete no lo vean como competidor o enemigo, sino como su colega y, por lo tanto, apoyen su labor. Casi todos los testimonios coinciden en que hay dos personas que deben fungir como aliados clave: uno es el presidente o primer ministro y el otro, el ministro de Finanzas o de Hacienda. Para ambos, el de Educación es la cuarta o quinta prioridad; para este, en cambio, son sus aliados más importantes.

Una sexta experiencia (y todos los textos coinciden) es que lo ideal es una reforma que concite el más alto nivel de apoyo ciudadano que sea posible, porque no solo permite amortiguar los conflictos eventuales, sino que también ayuda a abrir puertas políticas, mediáticas y presupuestales si los interlocutores respectivos ven que el ministro cuenta con una base de apoyo relevante. Alcanzar ese objetivo, sin embargo, es muy difícil en la realidad de las sociedades abiertas y complejas. Me explico: los sistemas educativos de países emergentes —o, mejor dicho, las estructuras políticas, burocráticas o sindicales que los dominan— son, por definición, muy conservadoras. Están acostumbradas al *statu quo*, a vivir dentro de un ecosistema muy predecible de donde medran —en ocasiones de manera escandalosa— con la carrera de los docentes, la administración de las escuelas y los presupuestos públicos, entre otras situaciones. No es comprobable que los maestros estén indefensos y dependan de la buena voluntad de los gobiernos para poder progresar en su desarrollo profesional. Al inicio de cada gobierno, suelen elaborar un discurso y una narrativa que ilustre los presuntos sacrificios de la profesión docente para fortalecer su posición a la hora de negociar mil y un aspectos con los ministerios, buena parte de ellos imposibles de satisfacer. Por lo tanto, el desafío del ministro es arti-

cular coaliciones políticas, mediáticas y civiles que asuman que, si se trata de hacer cambios de fondo, lo que menos importa es la estabilidad a corto plazo y que vale la pena apoyar la transformación, aunque cause problemas durante el proceso de instrumentación. Al contrario, si las reformas sacuden el *avispero* significa que van en la dirección correcta: se trata de crear incentivos correctos para que la ciudadanía apoye, a pesar de lo complejo de una reforma.

Para finalizar, todas las cartas tienen un toque —a ratos psicoanalítico— tanto de entusiasmo como de modestia en cuanto a los progresos registrados en las muy variadas experiencias de cada autor que, en el fondo, se convierte en un consejo práctico de la mayor importancia: se debe ser honesto con el público desde un principio; si lo que se quiere es que, al final de la gestión, la mejora educativa sea visible en los logros de aprendizaje y las evaluaciones, así como en la movilidad social y económica o, sencillamente, generar la percepción, más o menos medible en los sondeos, de que las cosas algo avanzan, hay que poner en su justa dimensión lo que en verdad se puede alcanzar en unos años. Es cierto que las victorias tempranas son muy útiles, como asegura la exministra Cecilia Vélez de Colombia, pero es necesario estar conscientes de que construir un sistema educativo de excelencia y, eventualmente, uno de los mejores del mundo, toma muchos años y varios gobiernos. Prometer lo contrario por razones de corto plazo es un engaño a la sociedad y, peor aún, a uno mismo.

Letters to a New Minister of Education... es un conjunto de lecciones aprendidas pero, sobre todo, un llamado al optimismo en la medida en la que los autores comparten la convicción de que, pese a las resistencias, ofrecer una educación de alta calidad para los niños y jóvenes es, al mismo tiempo, un compromiso moral superior, un bien público que bajo ninguna circunstancia debe ser negociable en el mercadeo político y una condición *sine qua non* para que los países crezcan, sean más competitivos y alcancen niveles altos de equidad y calidad de vida.

Colaboran en este número

Alonso Morilla Meneses

De nacionalidad española. Es maestro en Análisis Geoespacial por la Universidad de Girona y en Arqueología y Territorio por la Universidad de Granada. Actualmente se desempeña como coordinador del Departamento de Geomática Libre de GeoAlternativa, AC y es responsable de operaciones de la Asociación gvSIG.

Contacto: alonsomorilla@geoalternativa.com

Alejandra Albert Tejera

Nació en España. Es licenciada en Ciencias Ambientales por la Universidad Pablo de Olavide, con los posgrados Experto en Sistemas de Información Geográfica (SIG) por la Universidad de Huelva y en Geoinformática y Programación SIG por la Universidad de Girona. En la actualidad es miembro de GeoAlternativa, AC.

Contacto: alejandraalbert@geoalternativa.com

Silvia Franceschi

De nacionalidad italiana. Es ingeniera ambiental, con especialización en Hidrología, Geomorfología e Hidráulica por la Universidad de Bolzano (Italia). Se especializa en análisis ambiental a través de sistemas de información geográfica y evaluación de recursos de energías renovables. Es cofundadora de HydroloGIS y codesarrolladora de *Geopaparazzi* y *Geopaparazzi Survey Server*.

Contacto: silvia.franceschi@hydrologis.eu

César González-González

De nacionalidad mexicana. Es economista por la Universidad de Colima (UC), maestro en Demografía por el Colegio de la Frontera Norte (El COLEF) y doctor en Estudios de Población por El Colegio de México (El COLMEX). Es profesor de tiempo completo en la Facultad de Trabajo Social de la UC y su tema de investigación está relacionado con el envejecimiento de la población, principalmente con el estado de salud, cuidados y dependencia, arreglos familiares, la situación económica y lo que pasa en el último año de vida del adulto mayor.

Contacto: cgonzalez31@uclm.mx

Karina Orozco-Rocha

Nació en México. Es economista por la UC, doctora en Estudios de Población por El COLMEX y con posdoctorado en la UC. Actualmente, es profesora en dicha universidad. Pertenece al Sistema Nacional de Investigadores (SNI), nivel 1 del Consejo Nacional de Ciencia y Tecnología (CONACYT). Sus líneas de investigación y pu-

blicaciones examinan las desigualdades de género en el trabajo remunerado y no remunerado, así como el bienestar económico y redes de protección en las edades avanzadas en México.

Contacto: korozco9@ucol.mx

Mireya Patricia Arias Soto

De nacionalidad mexicana. Es licenciada en Trabajo Social y maestra en Trabajo Social con Orientación en Desarrollo Humano y Familia por la UC, así como doctora en Educación *Ph D* por la Universidad de Baja California (UBC). En la actualidad, se desempeña como profesora e investigadora en la UC. Es miembro del SNI, nivel candidata. Su línea de investigación es acerca de grupos sociales y familia. Ha publicado artículos, capítulos y libros sobre familia, adultos mayores, grupos vulnerables, migración, entre otros.

Contacto: mireya_arias@ucol.mx

Ma. Gregoria Carvajal Santillán

Nació en México. Es licenciada en Trabajo Social y maestra en Ciencias: área Investigación Educativa por la UC, así como doctora en Gerencia y Política Educativa por la UBC. Con 33 años de servicio, es profesora-investigadora de tiempo completo en la UC. Es miembro del SNI, nivel candidata. Su línea de investigación es sobre grupos sociales y familia. Ha publicado documentos acerca de jóvenes, género y adultos mayores.

Contacto: margre@ucol.mx

Eleonora Nuricumbo Rivera

De nacionalidad mexicana. Es estudiante-investigadora del Doctorado en Estudios Multidisciplinarios sobre el Trabajo (DEMST), perteneciente al Programa Nacional de Posgrados de Calidad (PNPC), Facultad de Psicología, de la Universidad Autónoma de Querétaro (UAQ), campus Aeropuerto, así como becaria del CONACYT. Actualmente, desarrolla investigación sobre el trabajo de entrevistador(a), entretejiendo las categorías de riesgos psicosociales y género.

Contacto: ele_nr@yahoo.com.mx

Candi Uribe Pineda

De nacionalidad mexicana. Es doctora en Ciencias Sociales con Especialidad en Estudios Rurales por el Colegio de Michoacán y tiene un posdoctorado CONACYT en la Unidad Multidisciplinaria de Estudios sobre el Trabajo de la UAQ. En la actualidad, es profesora-investigadora de la Maestría y Doctorado en Estudios Multidiscipli-

narios sobre el Trabajo (MEMST-DEMST), Facultad de Psicología de la UAQ, campus Aeropuerto. Sus líneas de investigación son: jóvenes, trabajo, riesgos psicosociales y precariedad laboral.

Contacto: upcandi@outlook.com

Adriana Marcela Meza Callejas

De nacionalidad mexicana. Es doctora en Ciencias Sociales con terminal en Sociología por la Universidad de Guadalajara y tiene una especialización en Estudios de Género por la Universidad Pedagógica Nacional. Es docente-investigadora de tiempo completo en la Facultad de Psicología de la Universidad Michoacana de San Nicolás de Hidalgo. Sus líneas de investigación son: estudios psicosociales sobre la vejez y el envejecimiento y sociología del trabajo.

Contacto: adime5@hotmail.com

Jorge A. Montejano Escamilla

De nacionalidad mexicana. Es doctor en Urbanismo, con especialidad en temas metropolitanos: metropolización del territorio y lógicas de localización, por la *Universitat Politècnica de Catalunya*. Es profesor-investigador de tiempo completo en el Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo).

Contacto: jmontejano@centrogeo.edu.mx

F. Gerardo Ávila Jiménez

De nacionalidad mexicana. Es maestro en Gestión de la Información, con especialidad en Análisis Espacial y Sistemas de Información Geográfica por la Universidad de la Habana. Es profesor en la Universidad Autónoma Metropolitana e investigador asociado en CentroGeo.

Contacto: gavila@centrogeo.edu.mx

Camilo Alberto Caudillo Cos

De nacionalidad mexicana. Es maestro en Geomática por el CentroGeo y maestro en Estudios de Población por la Facultad Latinoamericana de Ciencias Sociales (FLACSO). En la actualidad, estudia el doctorado en Ciencias de Información Geoespacial en el CentroGeo. Es profesor-investigador de tiempo completo en esa misma institución.

Contacto: ccaudillo@centrogeo.edu.mx

Rodrigo Tovar Cabañas

De nacionalidad mexicana. Es doctor en Geografía por la Universidad Nacional Autónoma de México (UNAM). Se desempeña como profesor-investigador en el Instituto Interdisciplinario de Investigación de

la Universidad de Xalapa, AC. Su trabajo más citado es *Tipificación de los acuíferos del campo Mina, Nuevo León, utilizando sistemas de información geográfica*.

Contacto: rod_geo77@hotmail.com

José Alfredo Jáuregui Díaz

Nació en México. Es doctor en Demografía por la Universidad Autónoma de Barcelona. Es investigador adscrito a la Universidad Autónoma de Nuevo León. Sus principales publicaciones son *Cambios en la mortalidad por eventos climáticos extremos en México entre el 2000 y 2015* y *Factores que influyen en las intenciones de retorno de latinoamericanos residentes en España a sus países de origen*.

Contacto: jose.jaureguid@uanl.mx

Shany Arely Vázquez Espinosa

De nacionalidad mexicana. Es candidata a doctora en Historia Regional por la Universidad Veracruzana (UV). Actualmente, es profesora-investigadora en el Instituto de Investigaciones Histórico-Sociales de la UV. Es autora de *Línea de costa teórica de la República Mexicana bajo el supuesto del descongelamiento total de los casquetes polares*.

Contacto: shanyvaz@gmail.com

Jesús López-Pérez

De nacionalidad mexicana. Es licenciado en Economía y maestro en Estadística Aplicada por el Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM). Ocupó diversos cargos en áreas de administración y análisis de riesgo crediticio en instituciones financieras de los sectores público y privado. En la actualidad, es investigador invitado en el Instituto Nacional de Estadística y Geografía (INEGI) en temas relacionados con el análisis econométrico de series de tiempo.

Contacto: jesus.lopezp@inegi.org.mx

Francisco Corona

De nacionalidad mexicana. Es licenciado en Economía por la Universidad Autónoma de Baja California, maestro en Estadística Aplicada por el ITESM y doctor en Economía y Métodos Cuantitativos por la Universidad Carlos III de Madrid (UC3M). En la actualidad, es investigador en el INEGI y sus líneas de estudio están relacionadas con el análisis econométrico y pronóstico de series de tiempo, así como en *Sport Analytics*; de ambas ha publicado trabajos en revistas arbi-

tradas de circulación internacional. Pertenece al SNI, nivel candidato.

Contacto: franciscoj.corona@inegi.org.mx

Olivia Jimena Juárez Carrillo

De nacionalidad mexicana. Es ingeniera en Mecatrónica por la Universidad Panamericana y tiene el *Master of Science in Computer Graphics, Vision and Imaging* por la *University College London*. En la actualidad, es asesora en la Junta de Gobierno del INEGI.

Contacto: jimena.juarez@inegi.org.mx

Paloma Merodio Gómez

De nacionalidad mexicana. Es licenciada en Economía por el Instituto Tecnológico Autónomo de México y maestra en Administración Pública en Desarrollo Internacional por la *Kennedy School of Government, Harvard University*. En abril del 2017, fue designada como vicepresidenta en la Junta de Gobierno del INEGI coordinando el Subsistema Nacional de Información Geográfica, Medio Ambiente, Ordenamiento Territorial y Urbano; a su vez, asumió la presidencia de UN-GGIM: Américas.

Contacto: paloma.merodio@inegi.org.mx

María del Socorro Ponce Medina

De nacionalidad Mexicana. Es ingeniera en Sistemas Computacionales por la Universidad Autónoma de Aguascalientes (UAA). Ha participado en el Proyecto de Identificación de Cultivos con Imágenes Satelitales en varios municipios del país. Se desempeña como jefa del Departamento de Cobertura y Análisis Espacial de la Dirección General Adjunta de Censos Económicos y Agropecuarios del INEGI

Contacto: ma.ponce@inegi.org.mx

José Luis Ornelas de Anda

De nacionalidad mexicana. Es biólogo por el Instituto Politécnico Nacional y tiene estudios de postgrado en Ecología por la misma institución. En el ámbito laboral, cuenta con una carrera en la producción de información desde 1981, cuando se unió al Departamento de Geografía y Medio Ambiente del INEGI (entonces Comisión de Estudios del Territorio Nacional); ha participado en varios programas, como la *Carta de uso de suelo y vegetación*; en la actualidad, es director de Recursos Naturales en la Dirección General de Geografía y Medio Ambiente.

Contacto: jose.ornelas@inegi.org.mx

**Abel Alejandro Coronado
Iruegas**

De nacionalidad mexicana. Es ingeniero en Sistemas Computacionales por la UAA y maestro en Estadística Oficial por el Centro de Investigación en Matemáticas, AC. Tiene una carrera de 17 años de trabajo en el INEGI, donde ha participado en varios proyectos; actualmente, es subdirector de investigación en Ciencia de Datos.

Contacto: abel.coronado@inegi.org.mx

César Rentería

Mexicano. Es licenciado en Mercadotecnia por la Universidad de Guadalajara, maestro en Administración y Políticas Públicas por el Centro de Investigación y Docencia Económicas (CIDE) y candidato a doctor en Administración y Políticas Públicas por la Universidad Estatal de Nueva York. Desde el 2020, es profesor-investigador invitado en la División de Administración Pública del CIDE. Se especializa en el estudio de tecnologías de la información en organizaciones.

Contacto: cesar.renteria@cide.edu

Ismael Aguilar-Benítez

De nacionalidad mexicana. Obtuvo el grado de doctor en Planeación Urbana y Regional en la Universidad de California, campus Irvine, California, Estados Unidos de América. Coordinó los libros *Los servicios del agua en el norte de México* (2011); *Gestión del agua en México: sustentabilidad y gobernanza* (2020) y *La gestión de los usos del agua en tres subregiones hidrológicas: San Juan, Valle de México y Bajo Grijalva* (2020), los tres publicados por El COLEF. Es miembro del SNI del CONACYT desde el 2007 y responsable técnico de la Red Temática en Gestión e Investigación del Agua. En la actualidad es profesor-investigador en El COLEF.

Contacto: iaguilar@colef.mx

Otto Granados Roldán

De nacionalidad mexicana. Es licenciado en Derecho por la UNAM y cuenta con estudios de maestría en Ciencia Política por El COLMEX. Fue gobernador del estado de Aguascalientes, embajador de México en Chile y secretario de Educación Pública, entre otros encargos. Actualmente, es presidente del Consejo Asesor de la Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI) y miembro del *Chen Yidan Visiting Global Fellows Program* en la Escuela de Graduados en Educación de la Universidad de Harvard.

Contacto: otto.granados@oei.es

Política y lineamientos editoriales

REALIDAD, DATOS Y ESPACIO REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA es una publicación cuatrimestral que sirve de enlace entre la generación de la información estadística y geográfica oficial y la investigación académica para compartir el conocimiento entre especialistas e instituciones con propósitos similares.

Se publicarán sólo artículos inéditos y originales relacionados con la situación actual del uso y aplicación de la información estadística y geográfica a nivel nacional e internacional.

Es una revista técnico-científica, bilingüe, cuyos trabajos son arbitrados por pares (especialistas), bajo la metodología doble ciego, con los siguientes criterios de evaluación: trabajos inéditos, originalidad, actualidad y oportunidad de la información, claridad en la definición de propósitos e ideas planteadas, cobertura de los objetivos definidos, estructura metodológica adecuada y congruencia entre la información contenida en el trabajo y las conclusiones.

El resultado del proceso de dictaminación se comunica por correo electrónico y contempla tres variantes: recomendado ampliamente (con modificaciones menores), recomendado (pero condicionado a modificaciones sugeridas) y no recomendado (rechazado). Dos dictámenes aprobados, se notifica al autor que se publica y se envía a corrección de estilo; un aprobado y uno rechazado, se le solicita realizar cambios; y dos rechazados, se notifica la no publicación.

Indizaciones y registros

- LATINDEX Catálogo (Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal).
- CLASE (Citas Latinoamericanas en Ciencias Sociales y Humanidades).
- REDIB (Red Iberoamericana de Innovación y Conocimiento Científico).

Lineamientos para publicar

Se publicarán trabajos en español e inglés: artículos de investigación, revisión y divulgación; ensayos; metodologías; informes técnicos; comunicaciones cortas; reseñas de libros; revisiones bibliográficas y estadísticas, entre otros.

1. El artículo —o cualquier otro tipo de escrito de los mencionados— deberá entregarse con una carta dirigida al editor responsable de REALIDAD, DATOS Y ESPACIO. REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA en la que se proponga el texto para su publicación, que se declare que es inédito y que no ha sido postulado de manera paralela en otro medio. Asimismo, deben incluirse los datos completos del(os) autor(es), nacionalidad(es), institución(es) de adscripción y cargo(s) que ocupa(n), domicilio(s) completo(s), correo(s) electrónico(s) y teléfono(s). Esto debe dirigirse a la atención de la M. en C. Virginia Abrin Batule, virginia.abrin@inegi.org.mx (tel. 5278 10 00, ext. 1161).
2. El trabajo se debe presentar en versión electrónica (formato *Word* o compatible) con: a) extensión no mayor de 20 cuartillas; b) letra Helvética, Arial o Times de 12 puntos y c) interlineado de 1.5 líneas. El material adicional al texto se requiere por separado: a) las imágenes, con resolución de 300 ppp y un tamaño no menor a 17 centímetros de base (ancho) en formato JPG o TIF —no remuestrear (ampliar) imágenes de menor resolución—; si son líneas o mapas, deben entregarse en formato vectorial (EPS o Ai), en caso de incluirse imágenes en mapa de bits, incrustarlas o enviarlas con el nombre con el cual se creó el vínculo (conservando los requerimientos de resolución y tamaño estipulados); y para fotografías, éstas no deben ser menores a 5 megapíxeles; b) las fórmulas o expresiones matemáticas tienen que elaborarse con el editor de ecuaciones propio de *Microsoft*[™], pero en caso de usar *software* de terceros, incluir en la entrega PDF testigo en el cual figuren exactamente cómo deben representarse; c) las gráficas, que incluyan el archivo en *Excel* con el cual se desarrollaron o, en su defecto, la imagen JPG legible, de origen, en alta resolución; y d) los cuadros, que sean editables, no se deben insertar como imagen.
3. La colaboración debe incluir: título del trabajo (en español e inglés o viceversa); resúmenes del trabajo en español e inglés (que no excedan de un párrafo de 10 renglones); palabras clave en español e inglés (mínimo tres, máximo cinco); bibliografía u otras fuentes; así como breve(s) semblanza(s) del(os) autor(es) que no exceda(n) de un párrafo de cinco renglones y que incluya(n) nacionalidad(es), grado(s) académico(s), principal(es) experiencia(s) profesional(es), adscripción(es) laboral(es) actual(es) y dirección(es) electrónica(s) de contacto.
4. Las referencias bibliográficas u otras fuentes deberán presentarse al final del artículo de la siguiente manera: nombre(s) del(os) autor(es) comenzando por el(los) apellido(s); título de la publicación con cursivas (si se trata de un artículo, debe estar entrecomillado, seguido de coma y la preposición en con dos puntos y, enseguida, el título de la revista o libro donde apareció publicado, con cursivas); país de origen; editorial; lugar y año de edición; página(s) consultada(s). En el caso de las fuentes electrónicas (páginas web) se debe seguir el mismo orden que en las bibliográficas, pero al final se pondrá entre paréntesis DE (dirección electrónica), la fecha de consulta y la liga completa. Se tienen que omitir aquellas que se mencionen como notas a pie de página. Si se aplica la opción de incluir en cuerpo de texto la referencia de nombre de autor y año de la fuente consultada entre paréntesis, sí deben aparecer todas las referencias mencionadas.

Página electrónica: <http://rde.inegi.org.mx>

Editorial Guidelines and Policy

REALITY, DATA AND SPACE INTERNATIONAL JOURNAL OF STATISTICS AND GEOGRAPHY is a four-monthly publication that connects statistics and geographic official information with academic research in order to share knowledge among specialists and institutions with similar aims.

We will publish only original and unpublished articles related to the current use and appliance of statistical and geographical information at both national and international levels.

It is a technical-scientific and bilingual magazine, with articles previously peer-reviewed by specialists under a double-blind methodology with the following evaluation criteria: unpublished works, originality, information related to opportunity and current affairs, we expect clarity in the definition of aims and ideas stated, defined objectives coverage, accurate methodological structure and coherence between the information of the paper as well as its conclusions.

The result of the paper-assessment process is delivered by email, and it involves three possibilities: fully recommended (with slight modifications), recommended (on condition of suggested modifications) and not recommended (i.e. rejected). When there are two reports of approval, the author gets notified that his/her paper will be published and it is sent to a style editing process. When one report approves the paper for publication and another one rejects it, the author is requested to make some changes for the text to be published. If the text submitted receives two non-favourable reports, the author is notified that the text will not be published.

Index and Registers

- LATINDEX Catalogue (Online Regional Information System for Scientific Journals from Latin America, the Caribbean, Spain and Portugal).
- CLASE (Latin American Quotations in Humanities and Social Sciences)
- REDIB (Latin American Net of Innovation and Scientific Knowledge)

Publishing Guidelines

Articles will be published in Spanish or English: research, revision and scientific-spreading articles; methodological; technical reports; short texts; book reviews; and bibliographical and statistical revisions, among others.

1. The article —or any other kind of text from those aforementioned— must be delivered with an attached letter addressed to the chief editor of Reality, Data and Space. International Statistics and Geography Magazine in which the text intended for publication will be submitted. There it must be stated that the text has not been published, and that it has not been submitted for publication in any other media. The names in full of the authors must be included, as well as their nationalities, adscription institutions, position in those institutions, postal address, e-mail address, and telephone numbers. This must be addressed to MSc Virginia Abrin Batule, virginia.abrin@inegi.org.mx (tel (+52) (55) 52.78.10.00, extension 1161).
2. The article must be submitted in an electronic version (a Microsoft Word file or a compatible one) with the following format: a) the text should not exceed the 20 pages of length; b) typography must be Helvetic, Arial or Times (12 points); and c) there should be a 1.5 line spacing in each paragraph. Additional material to the text will be delivered separately: a) images with a resolution of 300 ppp and no smaller than 17 cm width will be delivered in format JPG or TIF —please do not amplify images with lower resolution—. If the added materials are lines or maps, these must be delivered in vectorial format (EPS or Ai). If there are images in bits map, these must be embedded or attached with the name of the original file with which the link was created (keeping the resolution and size requirements above stated). As regards to photographs, these should not be inferior as 5 megapixels; b) mathematical expressions or formulae have to be created with the equations editor by *Microsoft*[™], but in case of using third-parties software, please attach a witness PDF in which the exact representation of mathematical formulae or expressions is contained; c) graphics must include the Excel file in which they were created or a legible image in the original JPG format in high resolution; and d) charts must be editable, and must not be inserted as images.
3. The text must include the following: the article's title (both in English and Spanish); the abstract of the article—both in English and Spanish (not longer than a 10-line paragraph); key words—both in English and Spanish (three as minimum and five as maximum); bibliography and other sources; as well as brief biographical sketches of the authors not exceeding a five-line paragraph each including nationalities, academic titles, main professional experiences, current work-related affiliations, and electronic addresses for the authors to be contacted.
4. Bibliographical references and other sources must be included at the end of the article in the following way: author's name (Surname first), and publication's title (in italics). If it is an article, the title must be in quotation marks followed by a comma and the preposition "en" with semicolon (in Spanish), then it should appear the title of the book or magazine in which the article was published (in italics); country of origin; publishing house, edition year, and consulted pages. As regards to electronic sources (web pages) the same order of the bibliographical references must be followed, but at the end the word "EA" (as for Electronic Address) ("DE" in Spanish) must be added within parenthesis followed by consultation date and the complete reference link. Those web links referred previously as footnotes, must be omitted in this section. However, if the name of the author and the year of the consulted source were included in the main body of the text within parenthesis, all these must be included as part of the bibliographical references.

Webpage: <http://rde.inegi.org.mx>



¿Cuántos hay?

En 2019 había en México 6 373 169 establecimientos, 719 155 más que en 2014 por lo que su número creció 2.4 % anual y dan empleo a 36 038 272 personas.

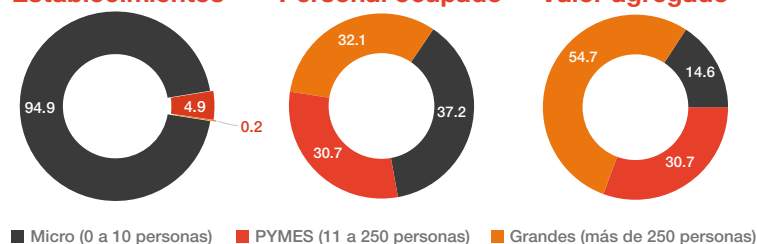
Por tamaño, los micronegocios son mayoría y emplean a 4 de cada 10 trabajadores.

(porcentajes)

Establecimientos

Personal ocupado

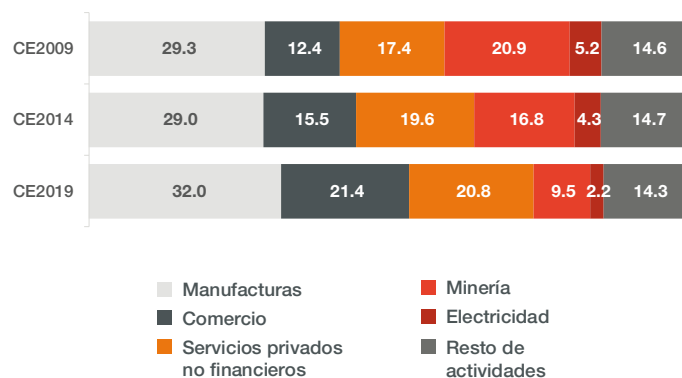
Valor agregado



¿A qué se dedican?

El sector **manufacturero genera mayor valor agregado**, seguido del comercio.

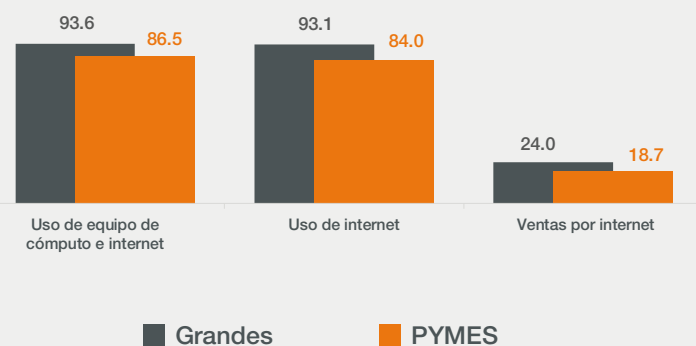
Participación de las actividades económicas en el valor agregado (porcentajes)



¿Cuántos usan el comercio electrónico?

Uno de cada cinco establecimientos PYMES declaró tener **ventas por internet**, mientras que en los grandes, uno de cada cuatro.

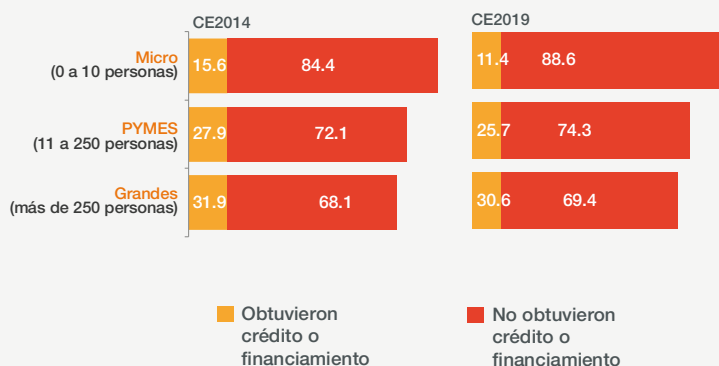
(porcentajes)



Disminuyó el número de empresas con financiamiento de 2014 a 2019

El acceso a **crédito o financiamiento** por parte de los establecimientos se ha reducido.

Acceso a financiamiento de los establecimientos según tamaño (porcentajes)



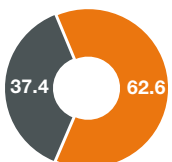


Formalidad e informalidad

En nuestro país, **6 de cada 10 establecimientos son informales** y su personal representa **19 % del total de trabajadores.**

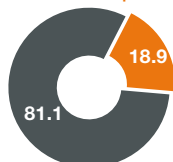
Establecimientos **formales e informales,*** su personal ocupado y valor agregado (porcentajes)

Establecimientos



Absoluto: 4 800 157

Personal ocupado

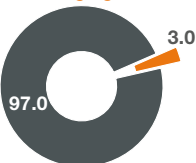


Absoluto: 27 132 927

Personal no remunerado = 78.3 %
Personal remunerado = 20.2 %
Remuneración promedio anual = \$44 089

Personal no remunerado = 9.9 %
Personal remunerado = 69.1 %
Remuneración promedio anual = \$133 968

Valor agregado



Absoluto: 9 983 800 millones de pesos

■ Formales
■ Informales

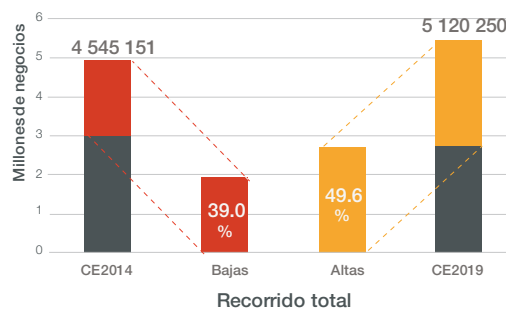
* Según la definición censal sobre informalidad basada en variables de los Censos Económicos.



Esperanza de vida de los establecimientos

De 2014 a 2019 **cerraron casi 4 de cada 10** establecimientos.

Altas y bajas de establecimientos en los últimos cinco años



Esperanza de vida de los negocios al nacer:
7.8 años

Por sector:

- ✓ Manufacturas **9.7 años**
- ✓ Comercio **6.9 años**
- ✓ Servicios **8.0 años**

Participación de las mujeres como dueñas

Se ha incrementado la proporción de **mujeres dueñas de negocios** pequeños en los últimos 10 años.



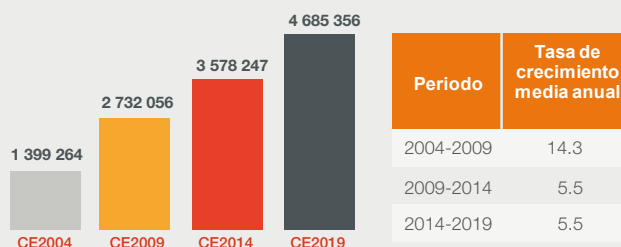
Establecimientos con hasta 100 personas ocupadas de manufacturas, comercio y servicios

	(porcentajes)	
	CE2009	CE2019
Establecimientos cuya dueña es una mujer	35.8	36.6
Establecimientos cuyo dueño es un hombre	42.2	35.7
Establecimientos de propiedad mixta	8.6	9.3
Establecimientos que no reportaron propietarios(as) en personal ocupado	13.4	18.4



Personal por outsourcing

Desde 2004 que se empezó a medir este concepto, **ha ido en aumento el personal ocupado por outsourcing.**



Fuente: INEGI.
Censos Económicos 2019.

Encuentra más información en el Sistema Automatizado de Información Censal (SAIC) y en Panorama económico de México. Disponibles en www.inegi.org.mx

