

Número

# REVISTA cartográfica

Número 100 / enero-junio 2020



INSTITUTO PANAMERICANO DE  
GEOGRAFÍA E HISTORIA

**AUTORIDADES DEL  
INSTITUTO PANAMERICANO DE GEOGRAFÍA E HISTORIA  
2017-2021**

<b>PRESIDENTE</b>	Lic. Israel Sánchez	Panamá
<b>VICEPRESIDENTE</b>	Mag. Alejandra Coll	Chile

**SECRETARIO GENERAL**  
Mag. César Rodríguez Tomeo  
Uruguay

**COMISIÓN DE CARTOGRAFÍA**  
(Costa Rica)  
Presidente:  
Mag. Max Lobo

Vicepresidente:  
Mag. Álvaro Antonio Álvarez

**COMISIÓN DE HISTORIA**  
(México)  
Presidenta:  
Dra. Patricia Galeana Herrera

Vicepresidente:  
Mtro. Rubén Ruiz

**COMISIÓN DE GEOGRAFÍA**  
(Estados Unidos de América)  
Presidenta:  
Dra. Patricia Solís

Vicepresidenta:  
Geóg. Jean W. Parcher

**COMISIÓN DE GEOFÍSICA**  
(Ecuador)  
Presidente:  
Dr. Mario Ruiz

Vicepresidente:  
Dra. Alexandra Alvarado

**MIEMBROS NACIONALES DE LA COMISIÓN DE CARTOGRAFÍA**

Argentina  
Belice  
Bolivia  
Brasil  
Chile  
Colombia  
Costa Rica  
Ecuador  
El Salvador  
Estados Unidos de América  
Guatemala  
Guatemala  
Haití  
Honduras  
México  
Nicaragua  
Panamá  
Paraguay  
Perú  
República Dominicana  
Uruguay  
Venezuela

Agrimensor Sergio Rubén Cimbaro Sr.  
Wilbert Vallejos  
Cnl. DAEN José Oviedo Bustillos  
João Bosco de Azevedo  
Dr. Carlos Mena Frau  
Fernando León Rivera  
  
Sr. Xavier Molina  
Sr. Mario Antonio Zeledón Flores  
Eric Van Praag  
Ing. Marcos Osmundo Sutuc Carillo Ing.  
Arch. Pierre Alexilien Versaille Arq.  
Marilyn Villatoro  
Ing. Raúl Ángel Gómez Moreno  
  
Lic. Israel Sánchez  
Cnel. SG Antonio Gavilán Estigarribia  
Mayor EP Reynaldo Flores Rivero  
Geóg. Susana Hernández  
Cnel. Daniel Píriz  
Ing. Sergio Rodríguez

**COMITÉS DE LA COMISIÓN DE CARTOGRAFÍA**

Normas y estándares de Información Geográfica	(Colombia)
Cartografía Temática	
Infraestructura de Datos Geospaciales	(Ecuador)

# INSTITUTO PANAMERICANO DE GEOGRAFÍA E HISTORIA

## REVISTA **cartográfica**



NÚMERO 100

ENERO-JUNIO 2020

## REVISTA CARTOGRÁFICA

Publicación anual fundada en 1952  
Desde 2016 publicación semestral  
Indizada en PERIÓDICA y Latindex  
Disponible en: Cengage Learning, Ebsco y ProQuest

La preparación de la REVISTA CARTOGRÁFICA está a cargo de la editora  
María Ester González  
Departamento de Geografía, Facultad de Arquitectura, Urbanismo y Geografía  
Universidad de Concepción  
Victoria 486-490, CP 4030000, Concepción, Chile  
Correo electrónico: editor\_revista\_cartografica@ipgh.org  
www.revistasipgh.org/index.php/rcar

Canje, venta y distribución de publicaciones, escribir a:  
Instituto Panamericano de Geografía e Historia  
Secretaría General  
Apartado Postal 18879, C.P. 11870 Ciudad de México, México  
Teléfonos: (52-55)5277-5888, 5277-5791 y 5515-1910  
Correo electrónico: publicaciones@ipgh.org      <http://www.ipgh.org>

Las opiniones expresadas en notas, informaciones, reseñas y trabajos publicados en la REVISTA CARTOGRÁFICA, son de la exclusiva responsabilidad de sus respectivos autores. Los originales que aparecen sin firma ni indicación de procedencia son de la Dirección de la Revista.

En cumplimiento con la resolución IX de la XIV Reunión del Consejo Directivo del IPGH, celebrada en julio de 1972, en Buenos Aires, Argentina, se advierte que: "Los límites que aparecen en los mapas de esta publicación no están, en algunos casos, finalmente determinados y su reproducción no significa aprobación oficial o aceptación por el Instituto Panamericano de Geografía e Historia (IPGH)".

---

**Diseño de portada / Cover design:** Hanss Erick Rivera Gutiérrez.

**Fuente de la imagen / Source:** Composición en base a distintas portadas de la *Revista Cartográfica*, 1952-2019.

© 2020 Instituto Panamericano de Geografía e Historia.

*Revista Cartográfica*, núm. 100, enero-junio 2020, es una publicación semestral editada por el Instituto Panamericano de Geografía e Historia, Ex-arzobispado núm. 29, Col. Observatorio, Alcaldía Miguel Hidalgo, C.P. 11860, Ciudad de México, México. Tels. (52-55)5277-5888, 5277-5791, 5515-1910, [www.ipgh.org](http://www.ipgh.org) / [publicaciones@ipgh.org](mailto:publicaciones@ipgh.org). Editora: María Ester González, correo electrónico: [editor\\_revista\\_cartografica@ipgh.org](mailto:editor_revista_cartografica@ipgh.org). Reserva de Derechos al Uso Exclusivo núm. 04-2015-090212390600-203, otorgado por el Instituto Nacional del Derecho de Autor. ISSN (en línea) 2663-3981. Licitud de título y contenido: en trámite. Responsable de la última actualización de este número: Departamento de Publicaciones del IPGH, Ex arzobispado núm. 29, Col. Observatorio, Alcaldía Miguel Hidalgo, C.P. 11860, Ciudad de México, México. Fecha de última modificación: 23 de abril de 2020.

Se autoriza cualquier reproducción parcial o total de los contenidos o imágenes de la publicación, incluido el almacenamiento electrónico, siempre y cuando sea para usos estrictamente académicos y sin fines de lucro, citando la fuente sin alteración del contenido y otorgando los créditos autorales.

## Contenido

Nota Editorial	7
<b>Artículos científicos</b>	
Modelización y predicción espacio-tiempo de la irradiancia solar global a corto plazo mediante redes neuronales artificiales y geoestadística <i>Space-time short-term solar radiation modeling and forecasting through artificial neural networks and geostatistics</i>	
Federico Vladimir Gutiérrez-Corea Miguel Ángel Manso-Callejo Francisco Serradilla-García	13
Patrones espaciales asociados a la infestación comunitaria por vectores de la enfermedad de Chagas <i>Spatial patterns of human settlement infestation by Chagas disease vectors</i>	
Gabriela Ramírez-Hernández Jean François-Mas Janine M. Ramsey	41
Propuesta de una guía para la evaluación de la exactitud posicional de datos espaciales <i>Proposal for a guide for the positional accuracy assessment of spatial data</i>	
Francisco Javier Ariza-López José Luis García-Balboa Joselyn Robledo Ceballos José Rodríguez-Avi Rosario Casanova	61
<i>Admissible relative errors of the parcel area while creating a land parcel map base using an orthoimage as a reference.</i>	
Determinación de errores admisibles en el área de las parcelas al ajustar un parcelario usando un ortomosaico de referencia	
Irene Balado Carlos López-Vázquez	81

---

Atlas escolares para la educación geográfica de niños y jóvenes <i>School atlases for the geographical education of children and young people</i>	
José Jesús Reyes Nunez	95
Columbus: aportes de la teledetección y análisis de coste anisotrópico a un debate histórico <i>Columbus: contributions of remote sensing and anisotropic cost analysis to an historical debat</i>	
Ángel M. Felicísimo	117
Geospatial data quality (ISO 19157-1): evolve or perish <i>Calidad de datos geoespaciales (ISSO 19157-1): evoluciona o muere</i>	
Francisco Javier Ariza-López Pablo Barreira González Joan Masó Pau Alaitz Zabala Torres Antonio Federico Rodríguez Pascual Gonzalo Moreno Vergara José Luis García Balboa	129
Reproducibilidad en AGILE: Experiencias, logros y recomendaciones <i>Reproducibility in AGILE: experiences, achievements, and recommendations</i>	
Carlos Granell Barbara Hofer Daniel Nüst Frank O. Ostermann Rusne Sileryte	155
La situación de la producción científica latinoamericana en el área de la Ciencia de Información Geográfica <i>The state of affairs of the latin-american GIScience scientific production</i>	
Carlos López-Vázquez Miguel Ángel Bernabé-Poveda	173
Sinfonía del Nuevo Mundo: el modelaje espacio-temporal de HGIS de las Indias, una infraestructura SIG para la América hispana borbónica <i>The New World Symphony: the spatio-temporal modeling of HGIS of the Indies; a GIS infrastructure for Bourbon Hispanic</i>	
Werner Stangl Paul Stangl	195

---

Megaciudad Buenos Aires: Cartografía de su última expansión y conurbación mediante el procesamiento digital de imágenes satelitales nocturnas

*Megacity Buenos Aires: Cartography of its latest expansion and conurbation with digital processing of nocturnal satellite images*

Gustavo D. Buzai

Eloy Montes Galbán

215

### **Reseñas**

Manual of Digital Earth

Luis M. Vilches-Blázquez

241

Análisis espacial con R: Usa R como un Sistema de Información Geográfica

Elena Chicaiza Mora

245

Instructivo para Autores

247

## Comité Editorial

---

**Carlos Granell Canut**

Universidad Jaume I de Castellón  
(España)

**Carlos López Vázquez**

Universidad ORT Uruguay  
(Uruguay)

**Daniela Ballari**

Universidad de Azuay  
(Ecuador)

**Diego Seco Naveiras**

Universidad de Concepción  
(Chile)

**Jean Francois Mas**

Universidad Nacional Autónoma de  
México  
(México)

**Joep Crompvoets**

KU Leuven Public Governance Institute  
(Bélgica)

**Miguel Ángel Bernabé Poveda**

Universidad Nacional de Catamarca  
(Argentina)

**Montserrat Gómez Delgado**

Universidad de Alcalá  
(España)

**Sandra Staller Vázquez**

Universidad Politécnica de Madrid  
(España)

**Stefan Steiniger**

Pontificia Universidad Católica de Chile  
(Chile)

---



## Nota Editorial

Es una satisfacción y un privilegio presentar el número 100 de la *Revista Cartográfica*. Como editora (desde noviembre de 2017), miro el pasado, el presente y el futuro de una publicación que nació en el año 1952, y que cumple 68 años de vida.

Miro el pasado de la *Revista Cartográfica* y pienso en su origen, en el momento en que empezó a dar sus primeros pasos. En este punto, el reconocimiento y tributo al presidente de la Comisión de Cartografía Gral. Carlos Alberto Levene (Argentina), quién ante el Comité Ejecutivo reunido en Washington en el año 1952 propuso la idea de creación de la *Revista Cartográfica*, con el fin de dar respuesta a la necesidad de disponer de un medio de intercambio de información científico-técnica.

El Gral. Levene fue el editor de los primeros cuatro números, desde 1952 hasta junio de 1955, asistido por Heliodoro Negri (Argentina).

Durante muchos años, hasta diciembre de 1996, los editores de la revista eran los propios presidentes de la Comisión de Cartografía, contando con el apoyo de asistentes en algunos números. El reconocimiento y agradecimiento a cada uno de los editores que hicieron posible su publicación durante tantos años: Félix M. F. Renault (Argentina) asistido por Heliodoro Negri (Argentina), Robert Randall (EUA), Juan J. Nano (Argentina) asistido por Heliodoro Negri y Pablo Dragan (Argentina), Oscar J. Colombo (Argentina) asistido por Leopoldo F. Rodríguez (Argentina), Porfirio García Gallont (Guatemala), José A. Villasana (México), Álvaro González Fletcher (Colombia) y Luz M. Valencia Carvajal (Colombia), Paul Peeler (EUA) y coeditores Scott W. Mooney (EUA) y Carlos S. Galindo Contreras (México), Antonio Hernández Navarro (México), Hermann Manríquez Tirado (Chile) y Valéria Oliveira Henrique de Araújo (Brasil)<sup>1</sup>.

Desde 1952 la *Revista Cartográfica* se ha publicado de forma ininterrumpida. Prevista con una frecuencia anual, alternó con otras periodicidades (semestrales, anuales y bianuales) según distintas circunstancias. En el año 2016 comenzó su publicación con una frecuencia semestral, publicando un número regular y otro temático. A partir de 2020 comienzan a publicarse dos números regulares al año, para responder a distintos requisitos de indexación. También durante sus 68 años de vida la revista cambió su “cara” o imagen de presentación, cambios que se han querido reflejar en la portada de este número 100.

Miro el presente de la *Revista Cartográfica*, orientando sus esfuerzos en el cumplimiento de requisitos que harán posible su indexación en distintos índices, bases de datos y catálogos. Para lo cual, uno de los primeros esfuerzos ha sido disponer de la

<sup>1</sup> Fuente: Pena, H. (2018) *El IPGH. Una historia de 90 años*, Ed. IPGH, Publ. 552, pp. 186-197, <https://publicaciones.ipgh.org/publicaciones-ocasionales/LIBRO-90-IPGH-Digital.pdf>

infraestructura técnica que permite la presentación en línea de los artículos y gestión de todo el flujo editorial a través del sistema de administración de revistas *Open Journal Systems* (OJS), accesible desde julio de 2019 en: <https://revistasipgh.org/index.php/rcar>. También se ha registrado el DOI (*Digital Object Identifier*) en la agencia de registro [CrossRef](#), garantizando el acceso directo, visibilidad, propiedad intelectual e interoperabilidad de cada artículo que se publica. Estas acciones han marcado el presente como una nueva época de la revista, cumpliendo requisitos tecnológicos para seguir avanzando de acuerdo a los escenarios y exigencias que rigen actualmente el complejo mundo de las revistas científicas.

Miro al futuro de la *Revista Cartográfica*, con el deseo de que logre mayor visibilidad en la comunidad científica a la que se dirige; que resulte atractiva para publicar a partir de su indexación en índices, bases de datos y catálogos reconocidos por los organismos relacionados con la investigación y continúe siendo accesible con una clara apuesta por el acceso libre y gratuito, sumándose al movimiento Open Access.

El pasado y presente, que ha sido posible gracias al esfuerzo colectivo que permitieron publicar cada número: editor, autor, revisor, diseñador, equipo de producción, etc. En este punto quiero agradecer a los autores que confiaron en esta revista como medio científico para publicar sus investigaciones, transmitir sus ideas y conocimiento; y al trabajo silencioso y eficaz de todos los revisores que han participado en la evaluación de los artículos. El futuro también será posible gracias al esfuerzo colectivo y por encima de los nuevos indicadores y exigencias que surjan para la indexación, resulta fundamental no olvidar el norte que ha regido la vida de la *Revista Cartográfica* desde su fundación en 1952: ser un medio de comunicación o en palabras del Gral. Levene un medio de intercambio de información científico-técnica. No olvidar ese norte que determina que al final, las revistas existen por y para sus autores y lectores.

La *Revista Cartográfica* inició su camino o recorrido en 1952, y hoy 68 años después en tiempos de grandes cambios en el mundo de las publicaciones científicas, el objetivo es seguir avanzando con el apoyo de todos los que hacen posible cada número, esperando que en el futuro un/a editor/a escriba: *Es una satisfacción y un privilegio presentar el número 200 de la Revista Cartográfica*.

Sin más preámbulos, los invito a disfrutar de la lectura del número 100 y a ser partícipes de la nueva época de la *Revista Cartográfica*.

María Ester Gonzalez

Editora

## Editorial

It is a satisfaction and a privilege to present issue 100 of the *Revista Cartográfica*. As editor (since November 2017), I look at the past, present and future of a publication that was born in 1952 and that celebrates 68 years of life.

I look to the past of the *Revista Cartográfica* and think about its origin, at the moment when this magazine began to take its first steps. At this point, the acknowledgment and gratitude to the President of the General Cartography Commission, who before the Executive Committee meeting in Washington DC in 1952 proposed the idea of creating the *Revista Cartográfica*, in order to respond to the need of having a means of exchanging scientific-technical information.

General Levene was the editor of the first four issues, from 1952 to June, 1955, assisted by Heliodoro Negri (Argentina).

During many years, until December 1996 the editors of the magazine were the presidents of the Cartography Commission, with the support of assistants in some issues. Acknowledgment and thanks to each of the editors that made its publication possible during so many years: Félix M. F. Renault (Argentina) assisted by Heliodoro Negri (Argentina), Robert Randall (EUA), Juan J. Nano (Argentina) assisted by Heliodoro Negri y Pablo Dragan (Argentina), Oscar J. Colombo (Argentina) assisted by Leopoldo F. Rodríguez (Argentina), Porfirio García Gallont (Guatemala), José A. Villasana (Mexico), Álvaro González Fletcher (Colombia) y Luz M. Valencia Carvajal (Colombia), Paul Peeler (USA) and coeditors: Scott W. Mooney (USA), Carlos S. Galindo Contreras (Mexico), Antonio Hernández Navarro (Mexico), Hermann Manríquez Tirado (Chile) and Valeria Oliveira Henrique de Araújo (Brazil)\*

Since 1952, the *Revista Cartográfica* has been published continuously. Planned with an annual frequency, it was alternated with other publications (semi-annual, annual and biannual), according to different circumstances. In 2016 its publication began with a biannual frequency, publishing a regular and a thematic issue. Starting in 2020, two regular issues began to be published to respond to indexing requirements. Also during its 68 years of life, the magazine has changed its “face” or presentation image. It was intended to reflect these changes on the cover of this issue 100.

I look at the present of the *Revista Cartográfica*, guiding its efforts in the fulfillment of requirements that will make possible its indexation in different indices, databases and catalogs. In order to achieve this, one of the first efforts has corresponded to have the technical infrastructure that allows the online presentation

\* Source: Pena, H. (2018), *El IPGH. Una historia de 90 años*, Ed. IPGH, Publ. 552. pp. 186-197 <https://publicaciones.ipgh.org/publicaciones-ocasionales/LIBRO-90-IPGH-Digital.pdf>

of articles and management of the entire editorial flow, through the *Open Magazine Systems* administration system of magazines accessible since July 2019 at: <https://revistasipgh.org/index.php/rcar>. The DOI (*Digital Object Identifier*) has also been registered at the CrossRef registration agency, guaranteeing direct access, visibility, intellectual property and interoperability of each article that is published. These actions have marked the present as a new era for the magazine, meeting technological requirements to continue advancing according to the scenarios and demands that currently govern the complex world of scientific magazines.

I look to the future of *Revista Cartográfica*, with the hope that will achieve greater visibility in the scientific community to which it is directed; that becomes attractive to publish, based on its indexation of indices, databases and catalogs recognized by research-related organization and continues to be accessible with a clear commitment for the free and open access, being joined to the Open Access movement.

The past and present have been possible thanks to the collective effort that allowed publishing each issue, editor, author, designer, production team, etc. In this point, I want to thank the authors who trusted this magazine as a scientific means to publish their research, transmit their ideas and knowledge, as well as the silent and efficient work of all reviewers who have participated in the evaluation of the articles. The future will also be possible thanks to the collective effort of the new indicators and requirements that arise for indexing. Therefore, it is essential not to forget the course that has governed the life of the *Revista Cartográfica* since its foundation in 1952: to be a means of communication, or in the words by General Levene, a means of exchanging scientific-technical information. It should not be forgotten that north that determines that in the end, magazines exist by and for their authors and readers.

The *Revista Cartográfica* began its journey in 1952 and today, 68 years later in times of great changes in the world of scientific publications, the objective is to continue advancing with the support of all who make each issue possible, hoping that in the future and editor will write “It is a satisfaction and a privilege to present the 200th issue of the *Revista Cartográfica*”.

Without further ado, I invite you to enjoy reading the issue 100 and to be part of the new era of *Revista Cartográfica*.

María Ester Gonzalez  
Guest Editor

ARTÍCULOS  
**científicos**





# Modelización y predicción espacio-tiempo de la irradiancia solar global a corto plazo mediante redes neuronales artificiales y geoestadística

## *Space-time short-term solar radiation modeling and forecasting through artificial neural networks and geostatistics*

Federico Vladimir Gutiérrez-Corea\*

Miguel Ángel Manso-Callejo\*\*

Francisco Serradilla-García\*\*\*

*Recibido 20 de noviembre de 2018; aceptado 03 de enero de 2019*

### Resumen

Mejorar el conocimiento de la Irradiancia Solar (IS) sobre la superficie terrestre, así como su predicción (pronóstico), cobra especial interés por su importancia para las energías renovables, a como lo son los sistemas basados en Energía Solar (ES), y para distintas aplicaciones industriales o ecológicas. En la presente investigación se ha experimentado con cinco técnicas de estimación espacial de la IS a intervalos de 15 minutos, en el territorio peninsular español, con distintas configuraciones espaciales. Encontrándose que la geoestadística mediante el Kriging con Regresión, usando variables auxiliares —una de ellas la IS estimada a partir de imágenes satelitales— permite estimar espacialmente la IS más allá de los 25 km, identificados en las investigación científicas previas, como límite de distancia máxima al punto de estimación.

\* Universidad Politécnica de Madrid, España-Instituto Nicaragüense de Estudios Territoriales, Nicaragua, correos electrónicos: [fv.gutierrez@upm.es](mailto:fv.gutierrez@upm.es); [vladimir.gutierrez@ineter.gob.ni](mailto:vladimir.gutierrez@ineter.gob.ni).  
ORCID: <https://orcid.org/0000-0002-5081-2934>.

\*\* Universidad Politécnica de Madrid, España, correo electrónico: [m.manso@upm.es](mailto:m.manso@upm.es).  
ORCID: <https://orcid.org/0000-0003-2307-8639>.

\*\*\* Universidad Politécnica de Madrid, España, correo electrónico: [francisco.serradilla@upm.es](mailto:francisco.serradilla@upm.es).  
ORCID: <https://orcid.org/0000-0001-7621-0627>.

Se ha experimentado con el modelado de Redes Neuronales Artificiales (RNA) para la predicción en tiempo futuro (temporal) —a corto plazo— de la IS utilizando observaciones próximas (componentes espaciales) en sus entradas y los resultados son prometedores. Así los niveles de errores disminuyen, en relación a investigaciones relacionadas, bajo las siguientes condiciones: cuando el horizonte temporal de predicción es inferior o igual a tres horas, las estaciones vecinas que se incluyen en los modelos deben encentrarse a una distancia máxima aproximada de 55 km.

Palabras clave: *redes neuronales artificiales, geoestadística, pronóstico.*

## Abstract

The enrichment of knowledge about solar irradiance (SI) on the Earth's surface and its prediction (forecast) has a great interest in renewable energy (RE), such as systems based on solar energy (SE) and for different applications industrial and environmental. At the present research it has been investigated five techniques of spatial estimation of the SI in 15 minutes of temporal resolutions for the Spanish mainland, with several spatial configurations. It's been found that the Geostatistics through Regression Kriging, using auxiliary variables —one of this: the SI estimated from Satellite Images— allows spatially estimates the SI beyond the 25 km identified by the related researches as the maximum distance limit to the estimation point. It has been experimented with the Artificial Neural Networks (ANN) modelling for the short-term forecasting of the SI, using close observations (spatial component) as part of its inputs, and the results are promising. In this way the error levels diminish, regarding to the related researches, under the following conditions: when the temporal horizons of the forecast is lower or equal to 3 hours, the neighbors stations to be included as input to the models should be at a 55 km of maximum distance.

Key words: *artificial neural networks, geostatistics, forecast.*

## Introducción

En las sociedades modernas la utilización de los Recursos y Energías Renovables (RER) como la energía solar, hidroeléctrica, eólica o geotérmica entre otras, han tomado fuerza como una de las soluciones a largo plazo con mayor potencial, sostenible y más adecuado para ayudar a minimizar el impacto ambiental y ayudar a solucionar los problemas climáticos (Kaya, 2006). La importancia de las fuentes de energía alternativas en el consumo energético mundial, se evidencia en el incremento superior al 100% de la utilización de estas fuentes en las últimas décadas (IEA, 2014). Este valor aún se puede considerar bajo, si se tiene en cuenta la capacidad de las fuentes de energía renovables existente, como es el caso del Sol. Según Moebius (2006), para cubrir la demanda energética actual, bastaría con un área total de celdas solares inferior al 1% de la superficie total del planeta, ya considerando celdas solares con el



10% de efectividad, la rotación de la tierra (día-noche), la interacción de la cobertura nubosa y otras componentes meteorológicas. Conscientes de la necesidad de aprovechar mejor los RER y promover el desarrollo sostenible, las grandes economías mundiales como en Estados Unidos, Unión Eueopea, China, entre otras; han definido sus objetivos de uso de RER a medio-largo plazo (EC, 2013; Toole *et al.*, 2010; EEN, 2010: 11; MITYC, 2009; IDAE, 2010).

El conocimiento y la predicción de los valores de la Irradiancia Solar (IS) desempeñan un papel importante en el desarrollo y la explotación de los sistemas basados en ES como son las plantas termo-solares y las fotovoltaicas (Mellit y Pavan, 2010), para las cuales la información de la irradiancia, y en particular la Irradiancia Solar Global (ISG), es especialmente relevante para el cálculo de numerosos procesos que van desde, la selección de sitios idóneos para la construcción de nuevas plantas de ES, pasando por su diseño, hasta finalmente llegar a la etapa de explotación; requiriéndose datos de la IS para la predicción de la producción energética, para la planificación y administración de la energía misma (Voyant *et al.*, 2014; Hernández *et al.*, 2012; Perpiñan, 2008: 21).

A pesar de la importancia de la información de la IS para investigaciones relacionadas con la ES (así como para otros tipos de estudios y aplicaciones), estos datos no siempre están disponibles en los sitios de interés debido a la ausencia de estaciones meteorológicas y/o sensores de IS (Yadav y Chandel, 2014).

El objetivo principal es investigar la predicción espacial y temporal de la ISG en el territorio peninsular español mediante RNA y geoestadística, con el propósito de generar mejores modelos predictivos que ayuden en la estimación de la producción de energía renovable procedente del Sol, tanto en sitios desde donde se disponen observaciones de esta variables como en emplazamientos donde no existen.

## **Estado de la cuestión**

En esta sección se describen los trabajos relacionados que permiten identificar los aportes científicos a realizar en materias tanto, de la estimación espacial en intervalos de 15 minutos de la IS, como de su predicción temporal a corto plazo.

### ***Disponibilidad de observaciones meteorológicas e IS***

Se sabe que los mejores valores de la IS son los observados por los sensores de IS en las Estaciones Meteorológicas (EM) emplazadas en la superficie terrestre (Polo *et al.*, 2008). Los datos de IS son necesarios en la generación y verificación de modelos, inferencia de datos o, asimilación de datos en modelos usados en contextos tales como la meteorología, climatología o hidrología, entre otros campos de estudio. Sin embargo, existe una baja densificación espacial de estos. De acuerdo con Hernández

*et al.* (2012), Antonanzas-Torres *et al.* (2013) y Badescu (2008), la escasa densificación de EM en superficie o sensores que observen la IS se debe a la inversión necesaria para su instalación y mantenimiento.

### ***La Radiación Solar***

La Radiación Solar (RS) es el término utilizado para describir la energía que por fusión nuclear se libera desde el centro del Sol cuando este convierte átomos de hidrógeno en helio, generándose así una pérdida de masa y una liberación de energía (Mayo 2013). La IS que alcanza a la Tierra puede verse en dos etapas. La primera es la IS Extraterrestre (ISE), que puede ser calculada de forma determinista, y la segunda es la ISG que llega a la superficie de la Tierra como resultado de la ISE que se transmite desde la periferia de la atmósfera hasta la superficie y, que como resultado de este viaje, se ve afectada por las condiciones estocásticas atmosféricas. A la vez, la ISG en un plano horizontal está compuesta por la suma de las irradiancias Directa, Difusa y Albedo (Gueymard y Myers, 2008).

### ***Estimación espacial de la ISG intra-horaria***

Los valores de la IS en la superficie terrestre pueden obtenerse por observación directa *insitu* o estimada de un modo indirecto. A continuación se presentan brevemente los métodos utilizados en la presente investigación para estimar superficies de ISG en intervalos inferiores a una hora, así como su relación con la literatura científica asociada.

#### ***Método Inverse Distance Weighting (IDW)***

En este método de interpolación la influencia entre observaciones está definida de modo determinista/análítico. Los pesos que se le asignan a las observaciones de IS de cada estación son inversamente proporcionales a la distancia entre el punto a estimar (interpolar) y las estaciones de referencia. Distintos estudios han utilizado IDW para interpolar la IS (Pons y Ninyerola, 2008; Zelenka *et al.*, 1999; Perez *et al.*, 1997).

#### ***Método Ordinary Kriging (OK)***

Tanto el método Kriging Ordinario (OK) como el Kriging con Regresión (RK), son métodos geoestadísticos. Para definir objetivamente los pesos a utilizar en la interpolación, la geoestadística modela un semivariograma que, considera tanto la distancia entre los puntos observados como la variación de sus valores. El método OK ha sido utilizado en distintas ocasiones para la estimación de la IS (Bojanowski *et al.*, 2013; Alsamamra *et al.*, 2009).

### *Método Regression Kriging (RK)*

Kriging con Regresión (RK) incluye información exógena a la variable objeto de la interpolación (variables auxiliares explicativas) compensándose así la relativa escasa densidad de distribución espacial de la variable objetivo, implicando que el RK está compuesto por la suma de una parte estocástica y una determinista de la variación espacial (Alsamamra *et al.*, 2009; Bojanowski *et al.*, 2013).

El método RK también ha sido utilizado en distintas ocasiones para estimar la IS (Antonanzas-Torres *et al.*, 2013; Alsamamra *et al.*, 2009; Evrendilek y Ertekin, 2007), así como para interpolar otras variables meteorológicas y climáticas (Hengl *et al.*, 2012; Pebesma, 2006; Hudson, 1993).

### *Estimación mediante Sensores Remotos*

La base para la estimación de la IS mediante sensores remotos es la relación inversa que existe entre, la reflectividad en el exterior de la atmósfera y la IS asociada a la nubosidad; esta relación fue derivada, en la década de los sesenta, a partir de la correlación entre las medidas de IS registrada por EM y la reflectividad obtenida por el satélite TIROS III (Wielicki *et al.*, 1996). A partir de ese momento, varios estudios han usado de imágenes de satélite para estimar la IS (Posselt *et al.*, 2011; Laszlo *et al.*, 2008; Pinker y Laszlo, 1992; Gautier *et al.*, 1980; Tarpley, 1979).

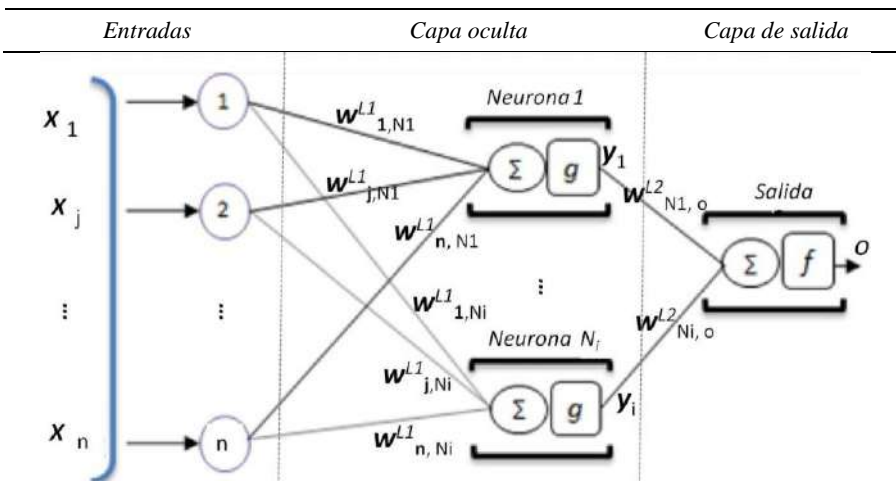
En Europa, a partir de la familia de satélites geoestacionarios Meteosat, se han generado distintas bases de datos de superficies de IS, entre ellas destaca la base de datos HelioClim3 (SoDA-is, 2011), la cual se genera a partir del método HelioSat-2 (Rigollier *et al.*, 2004).

### *Predicción de la ISG mediante Redes Neuronales Artificiales*

Las Redes Neuronales Artificiales (RNA) forman parte del área de conocimiento de la Inteligencia Artificial (IA) y del Aprendizaje Automático (*Machine Learning*), emulando mediante programas informáticos, la capacidad humana de aprender, memorizar y encontrar relaciones. Las RNA en particular intentan reproducir de manera extremadamente simple el comportamiento de las redes neuronales biológicas (Hagan *et al.*, 1996). La habilidad para aprender relaciones no lineales y su capacidad de modelar sistemas complejos, las han convertido en una herramienta útil en distintos ámbitos científicos (Yadav y Chandel, 2013; Paoli *et al.*, 2010; Reed y Marks, 1998). La unidad básica de las RNA es la Neurona Artificial (NA), que es una abstracción matemática simplificada del comportamiento de una Neurona Biológica (NB). Las RNA se componen de un gran número de NA agrupadas en capas, y altamente conectadas entre sí, para trabajar conjuntamente en la solución de un problema.

### Perceptrón Multicapa

El Perceptrón Multicapa (Multi Layer Perceptron: MLP) es la arquitectura de RNA más popular empleada en la solución de problemas científicos (Tymvios *et al.*, 2008), por su capacidad demostrada para aproximar relaciones no lineales (Reed y Marks, 1998: 37). En la Figura 1 se ilustra una RNA de tipo MLP con “n” valores de entrada, una capa oculta con “Ni” neuronas artificiales y una sola neurona “o” en la capa de salida. En ella también se aprecian los pesos “w” asociados a cada conexión entre neuronas. Los superíndices “L1” o “L2” indican la capa a la cual pertenece el peso; el primer subíndice indica la entrada o neurona desde donde parte la conexión y el segundo subíndice indica la neurona hacia la que se dirige la conexión.



**Figura 1.** Ilustración de una RNA de tipo Perceptrón Multicapa (MLP).

### Modelización de la IS a corto plazo con RNA

Las RNA han sido ampliamente utilizadas en la modelización de la predicción de IS en distintos horizontes de tiempo y en distintos lugares del mundo como, España (Linares-Rodríguez *et al.*, 2011; Bosch *et al.*, 2008; Hontoria *et al.*, 2002), Italia (Mellit y Pavan, 2010), Isla de Córcega en Francia (Voyant *et al.*, 2014; Voyant *et al.*, 2011; Paoli *et al.*, 2010), Turquía (Koca *et al.*, 2011), Arabia Saudita (Mohandes *et al.*, 1998), China (Wang *et al.*, 2012), India (Yadav y Chandel, 2012), Argelia (Mellit *et al.*, 2006), Uganda (Mubiru y Banda, 2008; Mubiru, 2008), entre otros (Mellit *et al.*, 2008; Mellit y Kalogirou, 2008).

Las investigaciones de Voyant *et al.* (2014), Wang *et al.* (2012), y Mellit y Pavan (2010), se relacionan a la presente investigación al focalizarse directamente en la predicción a corto plazo de la IS, utilizando RNA del tipo MLP y utilizando distintos estadísticos: Error Medio Cuadrático Normalizado (nRMSE), Error Medio Absoluto Porcentual (%MAE), Error Medio Cuadrático (RMSE) y Error Medio Cuadrático Porcentual (%RMSE). En la Tabla 1 se resumen las principales características de estas investigaciones.

**Tabla 1**  
**Niveles de errores en trabajos relacionados y otras características**

<i>Autores</i>	<i>Estadísticos</i>	<i>Periodo de cálculos</i>	<i>Nivel de agregación</i>	<i>Niveles de errores</i>
Voyant <i>et al.</i> (2014)	nRMSE	Estimaciones en las horas de luz	Promedios verano invierno	(27.3%) 21.7% 42.4%
Wang <i>et al.</i> (2012)	RMSE	Estimaciones para las 24 horas	Promedios claros nubosos	63.47 42.29 W/m <sup>2</sup> 84.65 W/m <sup>2</sup>
Mellit y Pavan (2010)	nRMSE	Estimaciones en las horas de luz	Promedios claros nubosos	(62.5%) 58% (promedio) 67%

### Estimación espacial de la ISG

El principal objetivo en esta sección es investigar la bondad de las estimaciones espaciales de la IS en intervalos de 15 minutos, de cinco métodos, en el territorio peninsular español y el archipiélago Balear, aplicado en seis configuraciones espaciales de los datos.

### Fuentes de datos

Se utilizaron cuatro fuentes de datos distintas, tres son Redes de Estaciones Meteorológicas (REM) (AEMet, Castilla y León y Meteoclimatic) que registraron observaciones de ISG en la superficie terrestre con periodicidades subhoras, y la última fuente es la base de datos de superficies de ISG HelioClim3v2, estimadas a partir de imágenes del satélite Meteosat.

Se han utilizado datos de dos años de observaciones de la red estaciones de la Agencia Estatal de Meteorología (AEMet), desde noviembre de 2010 hasta octubre de 2012. Durante este periodo, se identificaron 19 estaciones que recolectan información de ISG a intervalos de 10 minutos en el área de estudio. Se utilizaron las observaciones de 50 estaciones que registraron valores de ISG en el año 2011 del Instituto

Tecnológico Agrario de Castilla y León (Estévez y Gavilán, 2008). La tercera fuente de datos REM fueron las observaciones de 206 estaciones la de la Red Meteoclimatic (red de aficionados voluntarios), las cuales solo se muestran como referencia de sus ubicaciones, para posibles futuras consideraciones. Las estimaciones de ISG derivadas a partir de imágenes de satélite utilizadas provienen de la base de datos HelioClim3 versión 2 (HC3v2) (SoDA-is, 2011), creada con el método HelioSat2 (Rigollier *et al.*, 2004). La incertidumbre de la ISG HelioClim3v2 la estiman siguiendo un proceso de benchmarking recomendado por la tarea 36 del Programa de Calentamiento y Enfriamiento Solar (Solar Heating and Cooling Programme: SHC) de la Agencia Internacional de Energía (International Energy Agency: IEA) (IEA-SHC, 2010). Este *benchmark* define el uso de umbrales (0.1; 10; 50 y 200) W/m<sup>2</sup>, resultando en errores cuadráticos medios de 22.9%, 22%, 20.4% y 16.3% respectivamente (SoDa-is, 2013b).

## Metodología

### *Colecta de datos y adaptación temporal*

La información de ISG (observaciones) fue descargada de forma periódica mediante robots web programados para el propósito de esta investigación. Todas las observaciones se transformaron a un marco temporal común: tiempo UTC y valores de la ISG en los minutos múltiplos de 15 (minutos: 0, 15, 30 y 45).

### *Preparación de variables explicativas*

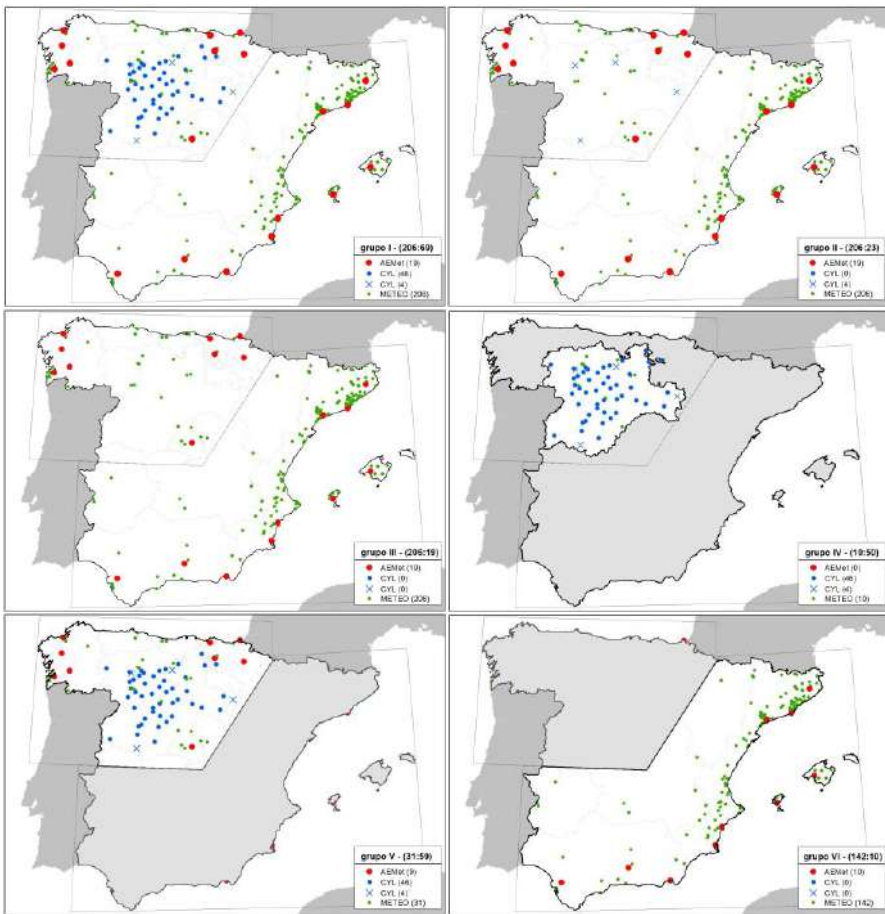
Son varios los factores que pueden influir en la cantidad de ISG percibida en una posición de la superficie terrestre. Así diferentes autores utilizan distintas variables explicativas (Bojanowski *et al.*, 2013; Moreno *et al.*, 2011; Alsamamra *et al.*, 2009; Antonanzas-Torres *et al.*, 2009; Kumar *et al.*, 2009; Evrendilek y Ertekin, 2007), entre otros.

Sin embargo ninguna estas investigaciones se focalizó la estimación de valores de ISG a 15 minutos. En este trabajo se seleccionaron tres variables auxiliares. La primera es una abstracción de las horas de luz durante el día y los cambios continuos en los ángulos del Sol, estructurada como una distancia estandarizada calculada para cada observación de la ISG desde su Tiempo Solar Verdadero (True Solar Time: TST) hasta su Medio Día Solar (MDS) con respecto al amanecer y el atardecer, para cada estación meteorológica y para cada día en cuestión. La variable se ha denominado Distancia al Medio Día Solar (DMDS). La segunda variable es la Estimación de la Irradiancia Solar Global (EISG) desde sensores remotos, siguiendo la tendencia de otras investigaciones como es el caso de Hengl *et al.* (2012). La tercera variable utilizada es la latitud, esta se utiliza en los dos métodos de RK y su selección se basó

en la identificación de tendencias con esta variable en Evrendilek y Ertekin (2007), y Kumar *et al.* (2009).

### Agrupación de los datos fuente

Se agruparon los datos en seis configuraciones geográficas, a como se ilustra en la Figura 2, bajo distintos criterios: solo estaciones oficiales, romper aglomeraciones, solamente inclusión de estaciones de AEMet, consideración de toda el área de estudio o solo una subárea.



**Figura 2.** Seis agrupaciones los datos para las estimaciones espaciales. (a) Grupo I (AEMet + CYL); (b) Grupo II (AEMet + 4 CYL); (c) Grupo III (AEMet); (d) Grupo IV (CYL); (e) Grupo V (CYL + 9 AEMet); (f) Grupo VI (AEMet).

### Estimación espacial de la ISG para intervalos de 15 minutos

La estimación espacial de la ISG, se ha realizado mediante los métodos descritos en la sección anterior. Se ejecutaron de forma independiente un conjunto de procedimientos para cada una de las seis agrupaciones antes mencionadas. Los procedimientos y cálculos de la metodología se realizaron con el paquete informático R (R Development Core Team, 2012). Para los procesos de interpolación utilizó el paquete geoestadístico para R gstat (Pebesma, 2004), con excepción de la generación de los semivariogramas, en el que se utilizó el paquete para R automap (Automap, 2011).

### Detección de valores atípicos

La detección de los valores atípicos (*outliers*) se realizó mediante el método de Box-Plot (Croarkin, 2010), en base a las diferencias entre las observaciones y las estimaciones, denotadas por ( $Y$ ) e ( $\hat{Y}$ ) respectivamente. Sus diferencias se agruparon por día y por estación, y se agregaron usando dos estadísticos RMSE (ecuación 1) y MAE (Ecuación 2), a los que se les aplicó el BoxPlot, considerando *outliers* aquellas estaciones identificadas por este método tanto por RMSE como por el MAE.

$$RMSE = \sum_{i=1}^n \sqrt{\frac{1}{n} (Y_i - \hat{Y}_i)^2} \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2)$$

A continuación se procedió a implementar los métodos IDW, OK y las dos formas de RK. Ambas formas de RK (denominadas RK1 y RK2 respectivamente) se diferencian en una de sus dos variables auxiliares (DMDS en el caso de RK1 y EISG en el caso de RK2), sin embargo ambas comparten como segunda variable auxiliar la “Latitud” geográfica. No se juntaron ambas variables (EISG y DMDS) en un mismo método de RK, por el hecho de que ambas presentan valores de multicolinealidad casi extrema, a pesar de ser ambas variables conceptualmente distintas (Allison, 1998: 141, 150), con valores de correlación de a.m. es 0.7825 y un coeficiente de determinación R2 de 0.6124, y en p.m. el valor de correlación es -0.8102 con R2 de 0.6566.

### Validación de los métodos

Para determinar la precisión de los métodos se realizaron validaciones cruzadas. La precisión del modelo HelioSat2 fue obtenida desde la base de datos HC3v2, calculándose las diferencias entre las observaciones en tierra y estos valores.



En todos los casos se calcularon los estadísticos RMSE relativo (%RMSE) y el MAE relativo (%MAE) en base al ratio entre el error estadístico y la media de la variable observada en tanto por 100, como se puede observar en las Ecuaciones (3) y (4) respetivamente:

$$\%RMSE = RMSE / \left( \frac{1}{n} \sum_{i=1}^n (Y_i) \right) \cdot 100 \tag{3}$$

$$\%MAE = \frac{MAE}{\frac{1}{n} \sum_{i=1}^n (Y_i)} \cdot 100 \tag{4}$$

Con el fin último de poder comparar los resultados de los distintos modelos con el benchmarking de referencia en (SoDa-is, 2013b), se calculan los errores para cinco umbrales de la ISG observada en tierra (todo, 0.1, 10, 50 y 200) W/m2, los que también son utilizados por SoDa-is (2013b) para el cálculo de la incertidumbre del modelo HelioSat2.

### Resultados y discusión

#### Puntuación de los métodos de estimación

La precisión y puntuación de los métodos de estimación en cada una de las agrupaciones geográficas, se presentan en la Tabla 2. Se obtuvieron calculando agregaciones diarias de los estadísticos %RMSE y %MAE en base a las diferencias 15-minutales (ISG – EISG) y descartando los días identificados como atípicos por el método BoxPlot.

**Tabla 2**  
**Clasificación de los métodos sin umbralizar los valores de irradiancia (agregación diaria) de acuerdo a los 6 grupos de datos: I (AEMet + CYL); II (AEMet + 4 CYL); III (AEMet); IV (CYL); V (CYL + 9 AEMet); VI (10 AEMet)**

Puntuación	I		II		III	Estadístico	
1°	RK2	23.75	RK2	27.09	SAT	26.93	
2°	SAT	29.09	SAT	27.49	RK2	28.02	
3°	RK1	30.59	RK1	39.28	IDW	39.86	%RMSE
4°	IDW	30.79	IDW	39.50	RK1	41.17	
5°	OK	31.61	OK	40.56	OK	43.22	

Continuación Tabla 2

Puntuación	IV		V		VI	Estadístico
1°	RK2	21.77	RK2	23.27	SAT	23.12
2°	RK1	25.25	RK1	28.59	RK2	29.82
3°	OK	25.41	IDW	28.86	IDW	34.16
4°	IDW	25.86	OK	28.92	OK	36.75
5°	SAT	28.95	SAT	29.94	RK1	37.91

Desde los valores anteriores se deduce que los métodos RK2 y SAT son siempre los ganadores (por simplicidad solo se muestra %RMSE). Siendo el método RK2 el ganador en las formas de agrupamiento (I, II, IV y V) y SAT en las formas de agrupamiento (III y VI) en estas últimas, solo participan las estaciones de AEMet con el menor número de estaciones por área de trabajo (poca densificación).

Se seleccionó como mejor método el que presenta el valor agregado de la media más bajo para cada tipo de agrupamiento. En la Tabla 3 se presentan los métodos ganadores junto con otros valores de relevancia (por simplicidad solo se muestra %RMSE). Se aprecia que aplicando el método RK2 al grupo IV (CYL) se obtiene el menor valor medio de error (el mejor resultado). También se observa que el grupo VI (10 AEMet) con el método SAT tiene el segundo valor más bajo de media, sin embargo, la unión de las áreas de los dos mejores grupos/métodos (IV y V) no cubren toda el área de estudio. Por lo anterior, también se toma en consideración el Grupo (V) con el método RK2, por presentar el siguiente mejor valor de media.

**Tabla 3**  
El mejor método por grupo y por estadístico, sin discriminar valores de irradiancia (agregación para el número total de días)

Grupo	Mejor método	% RMSE	Stdv (absoluta)	Estadístico
I (CYL + AEMet)	RK2	23.75	11.38	
II (AEMet + 4 CYL)	RK2	29.09	10.02	
III (AEMet)	SAT	26.93	10.03	
IV (CYL)	RK2	21.77	11.57	%RMSE
V (CYL + 9 AEMet)	RK2	23.27	11.63	
VI (10 AEMet)	SAT	23.12	10.53	

*Análisis detallado de los mejores métodos*

Los análisis presentados a continuación incluyen todos los días que forman parte de los experimentos (es decir, sin descartar los días identificados como atípicos), esto permite comparar los resultados del benchmarking en (SoDa-is, 2013b) y los trabajos relacionados.

- RK2 en la forma de agrupación IV (CYL)

La Tabla 4 presenta los estadísticos para el método RK2 en el agrupamiento (IV-CYL) con seis umbrales de ISG (todo, 0.1, 10, 50, 200 W/m<sup>2</sup>). Se puede apreciar que los estadísticos porcentuales (%RMSE y %Stdv) van disminuyendo a medida que el umbral aumenta (ISG), además estos valores son más bajos que los respectivos de (69, 72, 77, 83) W/m<sup>2</sup> y (22.9, 20.2, 20.04, 16.3) % y para umbrales (0.1, 10, 50 y 200) W/m<sup>2</sup> del benchmarking en (SoDa-is, 2013b).

**Tabla 4**  
**Estadísticos para el método RK2 en el grupo IV (CYL)**

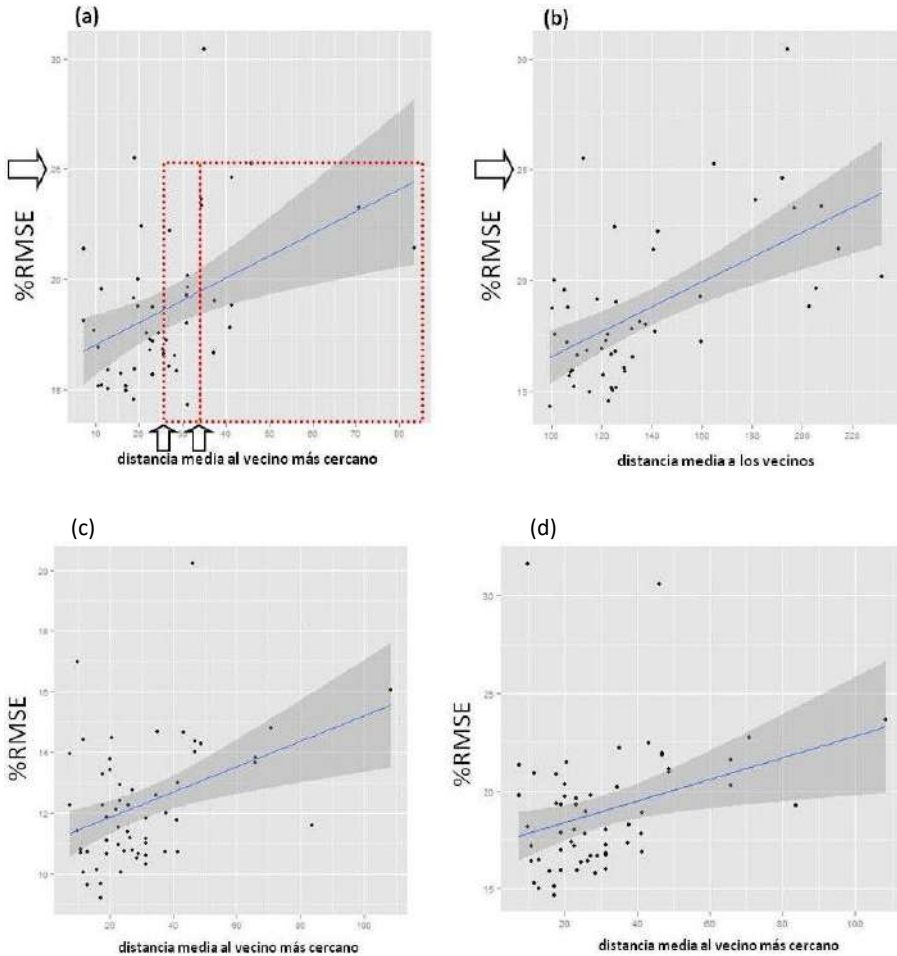
<i>Umbral</i>	<i>%RMSE (%)</i>	<i>Stdv %RMSE (%)</i>	<i>RMSE (W/m<sup>2</sup>)</i>	<i>Stdv RMSE (W/m<sup>2</sup>)</i>	<i>Rango RMSE (W/m<sup>2</sup>)</i>	<i>%RMSE (All)</i>	<i>Stdv RMSEp (All)</i>
NA	22.46	2.87	68.95	10.36	59-79	247.31	508.06
0.1	22.46	2.87	68.95	10.36	59-79	148.74	362.46
10	21.03	2.75	68.97	10.59	59-81	51.90	25.22
50	17.51	2.45	72.86	11.51	61-84	32.73	9.15
200	11.26	1.35	75.21	12.97	62-88	17.88	3.23

El análisis de la relación entre el %RMSE y la distancia que separa el punto a interpolar de los vecinos, en esta forma de agrupamiento se encontró que, el %RMSE es inferior o igual al 25.5%, independientemente de la distancia que separa el punto a interpolar del vecino más cercano, o independientemente de la distancia media a todos los vecinos, a como se muestra en la Figura 3, esto implica que el método desarrollado no se superan el límite de 25% de %RMSE para observaciones horarias indicado por Zelenka et al. (1999) y Pérez et al. (1997).

- RK2 en la forma de agrupación V (CYL + 9 AEMet)

En la Tabla 5 se presentan los estadísticos resultantes de aplicar el método RK2 en la forma de agrupamiento V (CYL + 9 AEMet) para 6 umbrales de la ISG

(todo, 0.1, 10, 50, 200). De aquí cabe destacar que los valores %RMSE para los umbrales 50 y 200 (18.95 y 12.29) %, son mejores que los respectivos valores de (20.04 y 16.3) % del benchmarking en (SoDa-is, 2013b); mientras que los valores RMSE a partir de los umbrales 10 (70.49, 68.95 y 63.39) W/m<sup>2</sup>, son mejores que los equivalentes del citado benchmarking.



**Figura 3.** %RMSE de las interpolaciones en relación a los vecinos para el método RK2 en CYL; (a) y (b) muestra el análisis sin umbrales aplicados; (c) y (d) muestra los valores con un umbral aplicado de mayor igual a 50 w/m<sup>2</sup>.

**Tabla 5**  
**Estadísticos para el método RK2 en el grupo V (CYL + 9 AEMet)**

<i>Umbral</i>	<i>%RMSE (%)</i>	<i>Stdv %RM SE (%)</i>	<i>RMSE: Wm2</i>	<i>Stdv RMSE (Wm2)</i>	<i>Rango RMSE (Wm2)</i>	<i>%RMSE (All)</i>	<i>Stdv %RMSE (All)</i>
NA	24.70	4.26	71.23	23.38	48-95	533.20	1'114.30
0.1	24.69	4.26	71.23	23.38	48-95	357.53	933.02
10	22.90	3.83	70.49	22.96	48-93	63.97	53.70
50	18.95	3.22	68.95	19.11	50-88	37.98	17.61
200	12.29	2.01	63.39	12.25	51-76	22.42	13.76

- SAT en la forma de agrupación VI (10 AEMet):  
En la Tabla 6 se presentan los estadísticos para el método HelioSat2 en la forma de agrupar (VI- 10 AEMet) para los seis umbrales de la ISG (todo, 0.1, 10, 50, 200) W/m2. Estos se comparan con al benchmarking en (SoDa-is 2013b), destacando que aunque los % RMSE para los umbrales 0.1 y 10 W/m2 son muy parecidos, los valores de los umbrales 50 y 200 W/m2 son más bajos que los del citado benchmarking y parecidos a los de los casos RK2 de las agrupaciones IV y V.

**Tabla 6**  
**Estadísticos para las estimaciones de la ISG derivadas de imágenes de satélite en el grupo VI (10 AEMet)**

<i>Umbral</i>	<i>% RMSE</i>	<i>Stdv % RMSE</i>	<i>RMSE (Wm2)</i>	<i>Stdv RMSE (Wm2)</i>	<i>Rango RMSE (Wm2)</i>	<i>% RMSE (All)</i>	<i>Stdv % RMSE (All)</i>
NA	22.85	0.87	105.13	15.67	89-121	713.09	823.02
0.1	22.84	0.87	105.13	15.68	89-121	500.27	476.25
10	21.55	1.03	104.64	15.78	89-120	84.05	33.26
50	18.04	1.27	103.59	15.12	88-119	46.06	8.14
200	11.46	1.25	101.21	15.91	85-117	23.13	2.64

**Predicción espacio-tiempo de la ISG**

Cuando se trabaja con contextos espaciales reducidos y en el corto plazo de tiempo (pequeña escala espacio-tiempo), la ISG se ve afectada por las condiciones dinámicas

de la atmósfera y a un alto ruido debido a factores como a la nubosidad. Estas condiciones presentan relaciones no lineales que vuelven compleja la predicción a corto plazo de la ISG. Las RNA forman parte de la IA y el ML, las que han demostrado su utilidad para investigar modelos asociados a procesos naturales que aproximan funciones no lineales, como es el caso de la predicción de la ISG a corto plazo.

El objetivo principal de esta sección ha sido investigar científicamente la predicción de la ISG en distintos horizontes temporales a corto plazo, donde este tipo de modelación es compleja, mediante el modelado de RNA, incluyendo la componente espacial como entrada a los mismos.

### ***Ámbito de estudio y datos experimentales***

Al igual que los trabajos relacionados (Voyant *et al.*, 2014; Wang *et al.*, 2012; Mellit y Pavan, 2010), se realizó la predicción de la ISG a corto plazo sobre un punto fijo (objetivo de predicción), sin embargo una innovación de esta investigación reside en el incluir la componente espacial en los datos de entrada, mediante el uso de datos en paralelo procedentes de estaciones vecinas al punto objetivo.

#### ***Sitio de estudio***

La selección de las EM a utilizar ha estado condicionada al uso de REM oficiales, a la distribución geográfica y a la densidad de estaciones con sensores de la ISG en ellas contenidas, quedando los experimentos acotados, al uso de 10 estaciones cercanas al centro de gravedad de Castilla y León.

#### ***Datos experimentales***

Los datos experimentales utilizados son las observaciones del año 2011 obtenidos desde esta REM de Castilla y León, utilizándose cinco variables meteorológicas: (1) ISG, (2) Temperatura Ambiente (TA), (3) Humedad Relativa del Aire (HRA), (4) Dirección del Viento (DV) y (5) Velocidad del Viento (VV). Adicionalmente se utilizaron 4 variables, relacionadas, calculadas de un modo determinista y referidas al instante de tiempo exacto de la toma de las observaciones meteorológicas, siendo éstas: (6) la Irradiancia Solar Extraterrestre Horizontal (ISEH), (7) el Índice de Claridad Instantáneo (KTi), (8) Distancia al Medio Día Solar (DMDS) y (9) el Ángulo Cenit (AC).

### ***Metodología***

Las fases de la metodología utilizada para investigar las predicciones de la ISG a corto plazo mediante RNA se vio influenciadas por el objetivo de experimentar in-

cluyendo un gran número de entradas en estas RNA. Todos los procesos se desarrollaron con el lenguaje de programación y *software* estadístico R.

#### *Normalización y selección de datos (entrenamiento y test)*

Con el objeto de lograr convergencia en las redes, los datos se normalizaron entre [0, 1] según la ecuación 5. Donde,  $x \in [x_{min}, x_{max}]$  y,  $y \in [y_{min}, y_{max}]$ ; “x” es el dato original y “y” es el su valor normalizado; Asumiendo los valores de  $y_{min} = 0$ ;  $y_{max} = 1$ .

$$y = \frac{(y_{max} - y_{min})(x - x_{min})}{x - x_{min}} + y_{min} \quad (5)$$

#### *Parámetros de entrenamiento de las RNA*

Con el propósito de determinar las mejores configuraciones de RNA aplicadas a una arquitectura de tipo Perceptrón Multicapas (MLP), manteniendo la mayor cantidad de entradas posibles, se estudiaron algunos parámetros que influyen en el aprendizaje y convergencia de la red como lo son: (i) el número de capas ocultas, (ii) la cantidad de neuronas en cada capa, (iii) el intervalo de normalización de los datos, (iv) el ratio de aprendizaje, (v) la cantidad de veces que se entregan todos los ejemplos a la RNA para su entrenamiento (Epochs). Otros parámetros se dejaron fijos, basados en los buenos resultados que estos ya han demostrado en estudios previos: (vi) el algoritmo de entrenamiento utilizado, ha sido BackPropagation, (vi) la función de comparación seleccionada ha sido la “suma de los errores al cuadrado”.

#### *Modelos de RNA y estadísticos para medir la calidad de las predicciones*

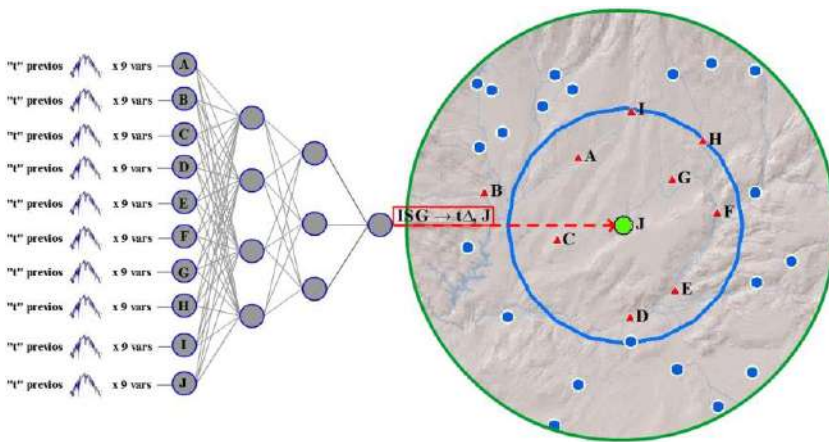
Los modelos de RNA creados se concentraron en la predicción futura de la ISG en horizontes temporales desde una hasta seis horas, experimentándose de forma independiente con un conjunto de RNA por cada horizonte de tiempo. En todos los casos, los datos de entrada utilizaron información de las distintas variables en instantes de tiempo previos, lo que se conoce como el método de ventana deslizante. Este método consiste en predecir el valor futuro de la ISG ( $\Delta t$ ) —el cual define el horizonte temporal de la predicción— utilizando como entradas un conjunto de variables correspondientes a una ventana de tiempo previa.

Los modelos de RNA aquí creados se pueden agrupar en dos tipos que, se diferencian entre sí por la utilización, o no, de datos procedentes de estaciones vecinas. En la figura 4 se representa la arquitectura general simplificada de los métodos basados en RNA desarrolladas. Aquí se puede apreciar cómo llegan a las entradas de la RNA datos de estaciones vecinas, y valores de sus variables para distintos intervalos

de tiempo mediante (método de ventana deslizante), también se aprecia que estas RNA poseen dos capas ocultas de neuronas, y como salida se obtiene el valor de la ISG predicho para un horizonte  $\Delta t$ .

La calidad de las predicciones de las distintas RNA creadas, se llevó a cabo utilizando cuatro estadísticos: RMSE, %RMSE, %MAE, nRMSE. El nRMSE se define en la siguiente ecuación (los otros se definieron previamente). Donde  $Y_i$  es el valor observado de la ISG, mientras que  $\hat{Y}_i$  es el valor predicho por la RNA. Se seleccionaron estos estadísticos para poder hacer comparaciones con investigaciones relacionadas.

$$nRMSE = \sum_{i=1}^n \sqrt{\frac{(Y_i - \hat{Y}_i)^2}{Y_i^2}} \quad (6)$$



**Figura 4.** Representación simplificada de una RNA para la predicción la ISG en un horizonte “ $\Delta t$ ” futuro, en base a datos de distintas estaciones vecinas y de sus variables en diferentes instantes de tiempo.

### Resultados y discusión

#### Parámetros para la selección de las arquitecturas de RNA

El proceso de identificación de los parámetros idóneos para el correcto entrenamiento de las redes fue iterativo, programándose en R experimentos con varias arquitecturas y configuraciones de red. A continuación se presentan los resultados del %RMSE de las validaciones cruzadas (utilizando los datos en test) de una muestra representativa



de algunos de los experimentos, distribuidos en distintas tablas. Los experimentos han permitido identificar los parámetros prometedores para, posteriormente utilizarlos en el entrenamiento de nuevas RNA que optimizaran el resultado de las predicciones.

- **Parámetros y RNA para horizonte temporal de 1 hora.**  
En esta tabla 7, las primeras RNA (1-3) realizan la predicción utilizando datos de estaciones vecinas, mientras que las ultimas (4-6) lo hacen con sólo observaciones de instantes previos de la estación objetivo. En ambos casos, se puede apreciar que los mejores niveles de error se encuentran con un MLP de una capa oculta (HL1). También se observa cómo el uso de dos capas ocultas (HL1 y HL2) asemeja los niveles de errores al del Perceptron Simple, sin embargo cabe destacar que tanto el Perceptron Simple como las otras configuraciones de RNA mantienen niveles de errores no muy alejados entre sí, lo cual es un indicio de que para este horizonte temporal el problema tiene una fuerte relación lineal.

**Tabla 7**  
**Identificación de parámetros para RNA que predicen la ISG:**  
**Pruebas con horizonte de una hora**

<i>Red #</i>	<i>Otros parámetros</i>	<i>Arquitectura de Red</i>	<i>% RMSE en TEST</i>
1	Delta-Deriv = 0.03	PERCEPTRON	24.62%
2	Delta-Deriv = 0.1	10HL1	23.99%
3	Delta-Deriv = 0.05	10HL1-5HL2	25.10%
4	Delta-Deriv = 0.03	PERCEPTRON	26.25%
5	Delta-Deriv = 0.025	30HL1	25.77%
6	Delta-Deriv = 0.025	10HL1-5HL2	25.82%

- **Parámetros y RNA para horizontes de más de una hora.**  
A continuación se presenta una muestra de las RNA con las que se experimentó la predicción futura de la ISG para un horizonte temporal de dos horas, mostrándose los resultados en la tabla 8. Se aprecia que las mejores predicciones se alcanzan con MLP de dos capas ocultas, también que los Perceptrones Simples no logran converger ni alcanzar niveles de errores aceptables, lo que implica que la predicción de la ISG para un horizonte de 2 o más horas, deja de ser un problema con relaciones lineales. En esta Tabla, no se muestran resultados de los experimentos para horizontes temporales de 3-6 horas puesto que el patrón encontrado fue similar a lo antes mencionado.

**Tabla 8**  
**Identificación de parámetros para RNA que predicen la ISG:**  
**Pruebas con horizonte de dos horas**

<i>Red #</i>	<i>Otros parámetros</i>	<i>Arquitectura de Red</i>	<i>%RMSE en TEST</i>
7	Delta-Deriv = 0.03	PERCEPTRON	mayor al 100%
8	Delta-Deriv = 0.05	10HL1	26.97%
9	Delta-Deriv = 0.05	100HL1-10HL2	24.23%
10	Delta-Deriv = 0.035	300HL1	mayor al 100%
11	Delta-Deriv = 0.035	300HL1-25HL2	27.67%
12	Delta-Deriv = 0.055	50HL1-25HL2	No convergió

#### *Predicciones de la ISG a corto plazo*

A continuación se presentan los resultados definitivos de modelar la predicción a corto plazo de la ISG mediante RNA. Una vez identificada la arquitectura RNA y sus parámetros de entrada, se generaron más experimentos para la optimización de los modelos de RNA definitivos. Se identificó que para la predicción a corto plazo de la ISG con estaciones vecinas, los mejores resultados se alcanzan con 100 neuronas en la primera capa oculta y 10 neuronas en la segunda. En el caso de no utilizar datos de estaciones vecinas, se encontró que los mejores resultados los genera un MLP con una capa oculta con 30 neuronas, siendo estos los parámetros definitivos utilizados.

#### *Análisis global de los resultados*

En la Tabla 9 se presentan los resultados globales (sin desagregar) de las predicciones de la ISG calculados a partir de las validaciones cruzadas (utilizando los datos en test) y agrupadas según los distintos horizontes de predicción (1-6 horas) y por la inclusión de estaciones vecinas en los datos de entrada.

Los resultados mostrados en la Tabla 9 muestran cómo en todos los casos los modelos de RNA creados, logran predecir la ISG a corto plazo con niveles de errores inferiores al 20% de nRMSE. Además desde la tabla se destaca cómo las RNA que predicen la ISG para horizontes temporales de una a tres horas y que incluyen en sus entradas los datos de estaciones vecinas, presentan mejores valores de nRMSE (12.88%, 13.80%, 15.98%), contrario a lo que ocurre con las RNA que sólo usan datos de instantes previos del mismo punto objetivo de predicción (13.38%, 14.84%, 16.04%). También es importante destacar que para los horizontes temporales de entre cuatro a seis horas los resultados se invierten, siendo mejores el segundo tipo de RNA

con valores de nRMSE de (17.36%, 17.27% y 18.27%) contra los valores de (17.94%, 20.07% y 19.62%) del primer tipo de RNA.

**Tabla 9**  
**Resumen de resultados de las validaciones cruzadas de las RNA que estiman la ISG a corto plazo para horizontes desde una a**

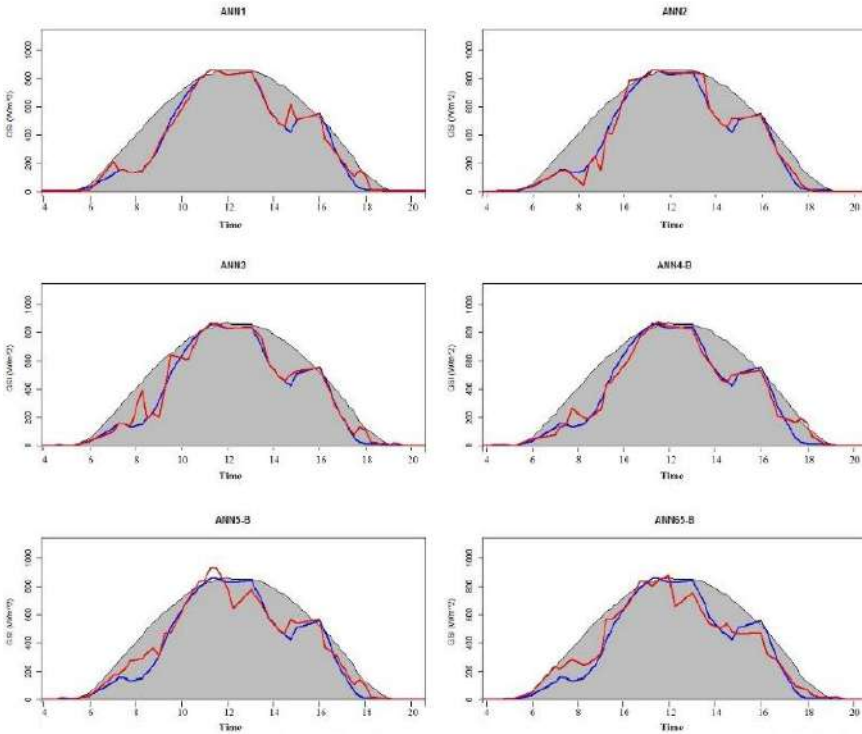
Nombre de Red	Arquitectura de Red	% RMSE	% MAE	RMSE W/m2	nRMSE
<i>RNA que utilizan datos de estaciones vecinas así como valores de instantes de tiempo previos de las distintas variables</i>					
ANN1	100-10 (D-0.05)	22.63	35.55	43.33	12.88
ANN2	100-10 (D-0.035)	24.24	38.15	47.49	13.80
ANN3	100-10 (D-0.05)	28.07	37.52	54.99	15.98
ANN4	100-10 (D-0.05)	31.52	46.96	61.75	17.94
ANN5	100-10 (D-0.05)	35.25	45.32	69.04	20.07
ANN6	100-10 (D-0.05)	34.47	43.95	67.52	19.62
Promedios:		29.36	41.24	57.36	16.72
<i>RNA que sólo utilizan datos de instantes de tiempo previos de la propia estación objetivo de predicción</i>					
ANN1-B	30 (D-0.05)	23.05	26.21	46.04	13.38
ANN2-B	30 (D-0.05)	26.06	44.52	51.05	14.84
ANN3-B	30 (D-0.05)	28.81	33.78	56.43	16.04
ANN4-B	30 (D-0.055)	30.49	44.02	59.74	17.36
ANN5-B	30 (D-0.055)	30.34	46.74	59.43	17.27
ANN6-B	30 (D-0.055)	32.09	47.99	62.86	18.27
Promedios		28.95	40.92	56.69	16.47

Los resultados alcanzados en esta investigación presentan un valor promedio en nRMSE de 16.45%. Este valor es mejor, y a la vez consistente, con los valores alcanzados en los trabajos de Voyant *et al.* (2014) y Paoli *et al.* (2010), en los que obtuvieron valores de nRMSE de 27.8% y ~20% respectivamente.

El valor medio del RMSE resultante es de 57.02 W/m2, este valor es mejor promedio, entre días claros y nublado, que el valor de 63.47 W/m2 obtenido por Wang *et al.* (2012). La comparación de los resultados con el otro estadístico común (%MAE) y el citado trabajo previo, no se ha podido establecer. Es posible que las diferencias se deban a las distintas localizaciones de las áreas de estudio, estando Wang *et al.*

(2012) ubicado en China, mientras que, tanto la investigación aquí desarrollada como las de Paoli *et al.* (2010) y Voyant *et al.* (2014) están ubicadas en el Sur de Europa. Otra diferencia es el hecho de que, en Wang *et al.* (2012), el análisis de los errores se hace mezclando las predicciones en periodos de día como en periodos nocturnos.

La Figura 5 muestra un ejemplo de pronóstico de serie temporal para un día (20 de agosto de 2011), utilizando los modelos de RNA desarrollados.



**Figura 5.** Ejemplo de series temporales basadas en valores pronosticados por las RNA.

Las Figuras 5 (“a”, “b” y “c”) presenta las predicciones para RNA que usan datos de estaciones vecinas (ANN1, ANN2 y ANN3) para horizontes temporales de una, dos y tres horas respectivamente; las figuras 5 (“d”, “e” y “f”) muestra las predicciones de RNA que solo usan observaciones de la propia estación objetivo (ANN4-B, ANN5-B y ANN6-B) para horizontes temporales de cuatro, cinco y seis horas respectivamente. Las líneas azules representan los valores observados por los sensores. Las líneas rojas representan los valores de las predicciones realizados por los modelos de RNA. El área gris representa el pronóstico basado persistencia de tipo *clear-sky*

*expectations*. Aquí se puede apreciar como la curva de predicción se acerca a la curva de las observaciones, a pesar de que la primera se proyecta horas antes de la segunda.

## Conclusiones

El mejor método de estimación espacial de la IS en intervalos de 15 minutos para el territorio español es el Kriging con Regresión (RK), soportado por los valores de la IS estimados desde imágenes de satélite y la latitud geográfica de la localización de la estación como variables auxiliares (KR2).

Se encontró que el método KR2 permite predecir espacialmente (interpolando) la IS en superficie más allá de los 25 km, límite que indica la bibliografía de referencia. El método KR2 ha sido probado hasta los 108 km, distancia para la que el %RMSE nunca alcanza el 25% indicado como límite.

Por lo tanto, el método KR2 desarrollado en esta investigación aprovecha las fortalezas de ambas fuentes de información (observaciones en superficie y valores derivados de imágenes de satélite); esto permite cubrir distancias de interpolación más allá del límite que implica el hecho de sólo utilizar las observaciones de las estaciones en tierra pero con un nivel de error menor al de sólo utilizar valores derivados desde imágenes de satélite.

Lo anterior habilita algunas aplicaciones prácticas. Por ejemplo, la futura generación de superficies de IS en intervalos de 15 minutos con el propósito de estimar el potencial de producción de energía generada por el Sol en lugares donde no hay observaciones directas de la IS y con un error inferior al 25% de RMSE.

Para cualquiera de los métodos (Interpolaciones o Estimación por Satélite), los outliers encontrados son mayores cuando se trata de estimar la IS en franjas de tiempo donde los valores de la IS son bajos (amanecer/anochecer).

La consideración en los modelos de RNA de la mayor cantidad de entradas, mediante la inclusión de la componente espacial, ha implicado experimentar con el modelado de RNA de hasta 900 entradas (altas dimensiones).

Los resultados de esta investigación científica han permitido generar modelos de RNA que predicen la ISG a corto plazo con niveles de exactitud inferiores al 20% del nRMSE, en todos los horizontes temporales, lo cual en valores promedios, es una mejora en cuanto a las investigaciones relacionadas.

Los resultados también han permitido identificar el aporte y los límites de la inclusión de estaciones vecinas en los datos de entrada a los modelos de RNA; así se encontró que existe una relación entre la distancia de los vecinos y el aporte que hacen al horizonte temporal de predicción. Para horizontes de predicción de hasta 3 horas, lo mejor es utilizar datos de las estaciones vecinas a distancias inferiores a 55 km respecto a la estación objetivo de predicción, mientras que para horizontes de entre 4 y 6 horas es mejor no utilizar dichas estaciones en ese rango de distancias.

La relación entre las predicciones en distintos horizontes de tiempo y las distancias en que se encuentran las estaciones vecinas, se puede explicar por el traslado del efecto meteorológico local de éstas hacia la estación objetivo de predicción, de tal forma que estaciones a 55 km o menos de lejanía y las condiciones metrológicas temporalmente cercanas (una a tres horas) influirán en la estación que es objetivo de predicción, mientras que los efectos de las observaciones de estas estaciones ya habrán pasado para horizontes de (cuatro a seis) horas, por lo cual, en trabajo futuros se debería investigar otras configuraciones especiales para la inclusión de la componente especial en los modelos RNA para mayores horizontes de tiempo que aporten mejoras en los niveles de errores.

## Bibliografía

- Allison, P. (1998). Federico Valdimir *Multiple regression: A primer*, Thousand Oaks, California, USA, Pine Forge Press.
- Alsamamra, H.; Ruiz-Arias, J.A.; Pozo-Vázquez, D.; Tovar-Pescador, J. (2009). “A Comparative Study of Ordinary and Residual Kriging Techniques for Mapping Global Solar Radiation Over Southern Spain”, *Agric. For. Meteorol*, 149: 1343-1357. <http://dx.doi.org/10.1016/j.agrformet.2009.03.005>.
- Antonanzas-Torres, F.; Cañizares, F. and Perpiñán, O. (2013). “Comparative Assessment of Global Irradiation from a Satellite Estimate Model (CM SAF) and on-Ground Measurements (SIAR): A Spanish Case Study”, *Renewable and Sustainable Energy Reviews*, 21: 248-261. <http://dx.doi.org/10.1016/j.rser.2012.12.033>.
- Automap (2011). Automatic Interpolation Package. Recuperado de <http://cran.r-project.org/web/packages/automap>.
- Badescu, V. (ed.) (2008). *Modeling Solar Radiation at the Earth's Surface*, Springer-Verlag Berlin Heidelberg, pp. vi-xvii.
- Bojanowski, J.S.; Vrieling, A.; Skidmore, A.K. (2013). “Calibration of Solar Radiation Models for Europe using Meteosat Second Generation and Weather Station Data”, *Agric. For. Meteorol*, 176: 1-9. <http://dx.doi.org/10.1016/j.agrformet.2013.03.005>.
- Bosch, J.L., Lopez, G. & Batlles, F.J. (2008). “Daily solar irradiation estimation over a mountainous area using artificial neural networks”, *Renewable Energy*, 33(7): 1622-1628.
- Croarkin, C. & Tobias, P. (2010). *NIST/SEMATECH e-Handbook of Statistical Methods. The National Institute of Standards and Technology 2010*, Gaithersburg, MD, USA.

- Estévez, J. and Gavilán, P. (2008). “Procedimientos de validación de datos de estaciones meteorológicas automáticas. Aplicación a la Red de Información Agroclimática de Andalucía”, II Jornadas de Agrometeorología 2008.
- EC (European Commission) (2013). *Joint Research Center Scientific and Policy Report. PV Status Report 2013*, Publication office of the European Union 2013, Ispra, Italy. Recuperado de [https://helapco.gr/pdf/pv\\_status\\_report\\_2013.pdf](https://helapco.gr/pdf/pv_status_report_2013.pdf).
- EEN (European Economy News) (2010). SMART, Sustainable and Inclusive Growth. EUR-LEX 2010. Recuperado de <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC2020&from=EN>.
- Evrendilek, F. and Ertekin, C. (2007). “Statistical Modeling of Spatio-Temporal Variability in Monthly Average Daily Solar Radiation Over Turkey”, *Sensors*, 7 (11): 2763-2778. <http://dx.doi.org/10.3390/s7112763>.
- Finn, P.G.; Udy, N.S.; Baltais, S.J.; Price, K. and Coles, L. (2010). “Assessing the Quality of Seagrass Data Collected by Community Volunteers in Moreton Bay Marine Park, Australia”, *Environmental Conservation*, 37(1): 83-89.
- Gautier, C.; Diak, G. and Masse, S. (1980). “A Simple Physical Model to Estimate Incident Solar Radiation at the Surface from GOES Satellite Data”, *J. Appl. Meteor*, 19: 1005-1012.
- Gueymard, C.A. and Myers D.R. (2008). Solar radiation measurement: Progress in radiometry for improved modeling. *Modeling Solar Radiation at the Earth Surface*, Springer-Verlag Berlin Heidelberg, pp. 1-27. [http://dx.doi.org/10.1007/978-3-540-77455-6\\_1](http://dx.doi.org/10.1007/978-3-540-77455-6_1).
- Hagan, M.T.; Demuth, H. B. and Beale, M.H. (1996). *Neural network design*, vol. 1, Boston: Pws.
- Hartley, B. (2012). “The Internet of Things - Weather Monitoring Too”, *WMO technical conference on meteorological and environmental instruments and methods of observation*, Belgium, Brussels.
- Hengl, T.; Heuvelink, G.B.; Tadić, M.P. and Pebesma, E.J. (2012). “Spatio-Temporal Prediction of Daily Temperatures using Time-Series of MODIS LST Images”, *Theoretical and applied climatology*, 107(1-2): 265-277.
- Hernández, L.; Baladrón, C.; Aguiar, J.M.; Calavia, L.; Carro, B.; Sánchez-Esguevillas, A.; Cook, D.J.; Chinarro, D. and Gómez, J. (2012). “A Study of the Relationship between Weather Variables and Electric Power Demand Inside a Smart grid/smart World Framework”, *Sensors*, 12(9): 11571-11591.
- Hontoria, L.; Aguilera, J. and Zufiria, P. (2002). “Generation of hourly irradiation synthetic series using the neural network multilayer perceptron”, *Solar Energy*, 72(5): 441-446.
- Hudson, G. (1993). “Kriging temperature in Scotland using the external drift method 2”, *Geostatistics Troia '92*, 1: 577-588, Dordrecht, The Netherlands Springer Netherlands. [http://dx.doi.org/10.1007/978-94-011-1739-5\\_45](http://dx.doi.org/10.1007/978-94-011-1739-5_45).

- IDAE (2010). *Informe de previsión. Plan de Energías Renovables 2011-2020*. Recuperado de <https://www.idae.es/tecnologias/energias-renovables/plan-de-energias-renovables-2011-2020>.
- IEA (2014). *2014 key world energy statistics*. Recuperado de <http://www.iea.org/publications/freepublications/publication/KeyWorld2014.pdf>.
- IEA-SHC (2010). *Task 36 Objectives*. Recuperado de <http://task36.iea-shc.org/publications>.
- Kaya, D. (2006). "Renewable energy policies in Turkey", *Renewable and Sustainable Energy Reviews*, 10 (2): 152-163.
- Knights, M. (2007). "Web 2.0", *Communications Engineer*, 5(1): 30-35.
- Koca, A., Oztop, H.F., Varol, Y. & Koca, G.O. (2011). "Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey", *Expert Systems with Applications*, 38(7): 8756-8762.
- Kumar, L.; Skidmore, A.K.; Knowles, E. (2009). "Modelling Topographic Variation in Solar Radiation in a GIS Environment", *Int. J. Geogr. Inf. Sci.*, 11(5):475-497.
- Laszloa, I.; Cirenba, P.; Liuba, H.; Kondraguntaa, S.; Tarpleya, J.D. & Goldberga, M.D. (2008). "Remote sensing of aerosol and radiation from geostationary satellites", *Advances in Space Research*, 41(11): pp. 1882-1893.
- Linares-Rodríguez, A., Ruiz-Arias, J. A., Pozo-Vázquez, D. & Tovar-Pescador, J. (2011). "Generation of synthetic daily global solar radiation data based on ERA-Interim reanalysis and artificial neural networks", *Energy*, 36(8): 5356-5365.
- Mayo, L. (2013). "The Sun: 101!. Sun-Earth Days", NASA. Recuperado de [http://sunearthday.nasa.gov/2013/articles/thesun\\_101.php](http://sunearthday.nasa.gov/2013/articles/thesun_101.php)
- Mellit, A., Benghanem, M. & Kalogirou, S.A. (2006). "An adaptive wavelet-network model for forecasting daily total solar-radiation", *Applied Energy*, 83(7): 705-722.
- Mellit, A. & Kalogirou, S.A. (2008). "Artificial intelligence techniques for photovoltaic applications: A review", *Progress in Energy and Combustion Science*, 34(5): 574-632. <http://dx.doi.org/10.1016/j.pecs.2008.01.001>.
- Mellit, A.; Kalogirou, S.A.; Hontoria, L. & Shaari, S. (2008). "Artificial intelligence techniques for sizing photovoltaic systems: A review", *Renewable and Sustainable Energy Reviews*, 13(2): 406-419. <http://dx.doi.org/10.1016/j.rser.2008.01.006>.
- Mellit A. & Pavan A.M. (2010). "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy", *Solar Energy*, 5(84): 802-821.
- MITYC (2009). Informe en relación con la cumplimentación del artículo 4.3 de la directiva 2009/28/ce, de 23 de abril de 2009, relativa al fomento del uso de energía procedente de fuentes renovables.



- Moebius, E. (2006). Solar Wind Electricity? NASA's Cosmicopia. [http://helios.gsfc.nasa.gov/qa\\_sun.html#swelectric](http://helios.gsfc.nasa.gov/qa_sun.html#swelectric)
- Mohandes, M.; Rehman, S. and Halawani, T. O. (1998). "Estimation of global solar radiation using artificial neural networks", *Renewable Energy*, 14(1): 179-184.
- Miorandi, D.; Sicari, S.; De Pellegrini, F. & Chlamtac, I. (2012). "Internet of things: vision, applications and research challenges", *Ad Hoc Netw*, 10(7): 1497-1516.
- Moreno, A.; Gilabert, M. & Martínez, B. (2011). "Mapping Daily Global Solar Irradiation Over Spain: A Comparative Study of Selected Approaches", *Sol. Energy*, 85(9): 2072-2084.
- Mubiru, J. (2008). "Predicting total solar irradiation values using artificial neural networks", *Renewable Energy*, 33(10): 2329-2332.
- Mubiru, J. & Banda, E.J.K.B. (2008). "Estimation of monthly average daily global solar irradiation using artificial neural networks", *Solar Energy*, 82(2): 181-187.
- Paoli, C., Voyant, C., Muselli, M. & Nivet, M.L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks, *Solar Energy*, 84(12): 2146-2160.
- Pebesma, E.J. (2004). "Multivariable Geostatistics in S: The Gstat Package", *Comput. Geosci.*, 30(7): 683-691.
- (2006). "The Role of External Variables and GIS Databases in Geostatistical Analysis", *Trans. GIS*, 10(4): 615-632.
- Perpiñan, O. (2008). "Grandes centrales fotovoltaicas: Producción, seguimiento y ciclo de vida", tesis doctoral, ETSII, UNED, Madrid, España.
- Perez, R.; Seals, R. & Zelenka, A. (1997). "Comparing Satellite Remote Sensing and Ground Network Measurements for the Production of site/time Specific Irradiance Data", *Sol. Energy* 60(3): 89-96.
- Pinker, R.T. and Laszlo, I. (1992). "Modeling Surface Solar Irradiance for Satellite Applications on a Global Scale", *J. Appl. Meteor.*, 31(2):194-211.
- Polo, J.; Zarzalejo, L.F. and Ramírez, L. (2008). "Solar radiation derived from satellite images", *Modeling Solar Radiation at the Earth's Surface*, Springer-Verlag Berlin Heidelberg, pp. 449-462.
- Pons, X. & Ninyerola, M. (2008). "Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data", *Int. J. Climatol.*, 28 (13): 1821-1834. <http://dx.doi.org/10.1002/joc.1676>.
- Posselt, R.; Mueller, R.; Stöckli, R. & Trentmann, J. (2011). "Spatial and Temporal Homogeneity of Solar Surface Irradiance across Satellite Generations", *Remote Sens.*, 3(5): 1029-1046. <http://dx.doi.org/10.3390/rs3051029>.
- Rigollier, C.; Lefèvre, M. & Wald, L. (2004). "The Method Heliosat-2 for Deriving Shortwave Solar Radiation from Satellite Images", *Solar Energy*, 77(2): 159-169.

- Reed, R. D. & Marks, R.J. (1998). *Neural smithing: supervised learning in feedforward artificial neural networks*, Mit Press.
- SoDA-is (2011). “Main Properties of the HelioClim Databases”, France. Recuperado de <http://www.soda-pro.com/help/helioclim/introduction>.
- (2013b). “HelioClim-3 Uncertainty”, France. [https://web.archive.org/web/20130511070410/http://www.soda-is.com/eng/helioclim/helioclim3\\_uncertainty\\_eng.html](https://web.archive.org/web/20130511070410/http://www.soda-is.com/eng/helioclim/helioclim3_uncertainty_eng.html).
- Tarpley, J.D. (1979). “Estimating Incident Solar Radiation at the Surface from Geostationary Satellite Data”, *J. Appl. Meteor.*, 18: 1172-1181.
- Toole, G.L.; Fair, M.; Berscheid, A. & Bent, R. (2010). “Electric power transmission network design for wind generation in the western United States: Algorithms, methodology, and analysis”. Transmission and distribution conference and exposition. *IEEE PES*.19-22. <http://dx.doi.org/10.1109/TDC.2010.5484549>.
- Tymvios, F.S., Michaelides, S.C. & Skouteli, C.S. (2008). “Estimation of surface solar radiation with artificial neural networks”, *Modeling Solar Radiation at the Earth's Surface*, Springer Berlin Heidelberg, pp. 221-256.
- Voyant, C.; Muselli, M.; Paoli, C. & Nivet, M.L. (2011). “Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation”, *Energy*, 36(1): 348-359.
- Voyant, C.; Randimivololona, P.; Nivet, M.L.; Paoli, C. and Muselli, M. (2014). “Twenty four hours ahead global irradiation forecasting using multi-layer perceptron”, *Meteorological Applications*, 21(3). <http://dx.doi.org/10.1002/met.1387>.
- Wang, F.; Mi, Z.; Su, S. & Zhao, H. (2012). “Short-Term Solar Irradiance Forecasting Model Based on Artificial Neural Network Using Statistical Feature Parameters”, *Energies*, 5(5): 1355-1370. <http://dx.doi.org/10.3390/en5051355>.
- Wielicki, B.A.; Barkstrom, B.R.; Harrison, E.F.; Lee, R. B.; Smith, G.L. & Cooper, J.E. (1996). “Clouds and the Earth’s Radiant Energy System (CERES): An Earth Observing System Experiment”, *Bull. Amer. Meteor. Soc.*, 77(5): 853-868.
- Yadav, A.K. and Chandel, S.S. (2012). “Artificial Neural Network based Prediction of Solar Radiation for Indian Stations”, *International Journal of Computer Applications*, 50(9): 1-4.
- Yadav, A.K. and Chandel, S.S. (2014). “Solar radiation prediction using Artificial Neural Networks Techniques: A review”, *Renewable and Sustainable Energy Reviews*, 33: 772-781.
- Zelenka, A.; Perez, R.; Seals, R. & Renné, D. (1999). “Effective Accuracy of Satellite-Derived Hourly Irradiances”, *Theoretical and Applied Climatology*, 62(3): 199-207.

# Patrones espaciales asociados a la infestación comunitaria por vectores de la enfermedad de Chagas

*Spatial patterns of human settlement infestation by Chagas disease vectors*

Gabriela Ramírez-Hernández\*

Jean François Mas\*\*

Janine M. Ramsey\*\*\*

*Recibido el 18 de agosto de 2019; aceptado 4 de noviembre de 2019*

## Resumen

Factores socioculturales y ambientales afectan el contacto en el paisaje entre triatominos infectados con *Trypanosoma cruzi* y la población humana, etiología de la enfermedad de Chagas. En este trabajo se analizan los efectos de la estructura del paisaje sobre la presencia de triatominos (vectores) en asentamientos humanos de una zona rural de Oaxaca, México. Mediante la búsqueda y colecta participativa de vectores, se calcularon índices de infestación en 35 localidades, en 2005 y en 2015. Se elaboró cartografía de cubierta y uso de suelo para ambas fechas y se estimó la cobertura de cada cubierta dentro de diez áreas de influencia. Se encontraron dos especies principales en el ambiente doméstico: *Triatoma dimidiata* Haplogrupo 2 (Hg2) y *Triatoma mazzottii*. La altitud, la densidad de vivienda y el área de bosque tropical son factores que permitieron explicar la presencia de cualquier de triatominos en ambiente doméstico. Se identificaron zonas prioritarias para implementar acciones a

\* Universidad Nacional Autónoma de México (UNAM), México, correo electrónico: [gabriela.ramrezhdz@gmail.com](mailto:gabriela.ramrezhdz@gmail.com). ORCID: <https://orcid.org/0000-0001-7082-8311>

\*\* Universidad Nacional Autónoma de México (UNAM), México, correo electrónico: [jfmas@ciga.unam.mx](mailto:jfmas@ciga.unam.mx). ORCID: <https://orcid.org/0000-0002-6138-9879>

\*\*\* Instituto Nacional de Salud Pública, México, correo electrónico: [jramsey@insp.mx](mailto:jramsey@insp.mx). ORCID: <https://orcid.org/0000-0002-5950-0830>

corto y mediano plazo y así atender casos actuales y detener la incidencia de esta zoonosis.

Palabras clave: *vectores, infestación, paisaje, Trypanosoma cruzi, enfermedad de Chagas.*

## Abstract

Environmental and sociocultural factors affect landscape contact between the human population and triatomines infected with *Trypanosoma cruzi*, etiology of Chagas disease. This paper analyzes the effects of landscape structure on the presence of triatomines (vectors) in human settlements in a rural area of Oaxaca, Mexico. Through participatory search and collection of vectors, infestation rates were calculated in 35 communities, in 2005 and 2015. Cover and land use mapping were developed for both periods and the coverage of each type within ten areas of influence was estimated. Two primary species were found in the domestic environment: *Triatoma dimidiata* Haplogroup 2 (Hg2) and *Triatoma mazzottii*. The altitude, density of housing and surface area of tropical forest are factors that explained presence of either triatomine in houses. Priority surveillance areas were identified for short and medium-term prevention and healthcare attention to eliminate incidence of this disease.

Key words: *vectors, infestation, landscape, Trypanosoma cruzi, Chagas disease.*

## Introducción

Presente de forma endémica en 21 países de Latinoamérica, la enfermedad de Chagas ocasiona la muerte de 10 600 personas anualmente. Tan solo en México se estima que el número de personas afectadas por la enfermedad es entre uno y dos millones (Sánchez-González *et al.*, 2016). En años recientes la enfermedad ha adquirido gran importancia de carácter urbano y global debido a los movimientos migratorios de las personas que la padecen (Schmunis y Yadon, 2010; WHO, 2015, 2017).

La enfermedad de Chagas es causada por la infección del parásito *Trypanosoma cruzi* en órganos internos de mamíferos, entre ellos los humanos. Se le conoce como una enfermedad silenciosa ya que desde que se contamina una persona hasta que se desarrollan síntomas o afectación patológica puede pasar entre 5 y 20 años. Durante la fase crónica, una proporción de las personas desarrollarán alteraciones en el tracto intestinal, trastornos neurológicos y cardíacos (Rassi *et al.*, 2010). El principal mecanismo de transmisión del parásito hacia los humanos es el vectorial, es decir, mediante el contacto con las heces de insectos de la subfamilia Triatominae mejor conocidos como chinches o “pic” en México (Ramsey *et al.*, 2015).

El método más efectivo para impedir la incidencia de la enfermedad es la prevención de la transmisión vectorial (WHO, 2017). Las poblaciones del vector están presentes en diferentes tipos de ambientes, sin embargo, el grado en que están en

contacto con los humanos difiere dependiendo de múltiples variables socioeconómicas, ambientales y culturales (Valdez-Tah *et al.*, 2015a). Se ha documentado que la calidad de la vivienda, las características del área peri-doméstica, la disponibilidad de hospederos domésticos (Ibarra-Cerdeña *et al.*, 2017) y la ubicación de la vivienda son factores que favorecen significativamente la domiciliación de los vectores, es decir, la infestación doméstica o presencia de vectores dentro y en los alrededores de las viviendas (Cohen *et al.*, 2006; Dumonteil *et al.*, 2013; Bustamante *et al.*, 2014; Rossi *et al.*, 2015; López-Cancino *et al.*, 2015). Estas condiciones permiten a los vectores encontrar refugios, sitios idóneos para la reproducción y acceso a recursos alimenticios constantes. Sin embargo, ha sido poco estudiada la influencia de la variación espacio-temporal del paisaje en el proceso de infestación. De acuerdo con Valdez-Tah *et al.* (2015a), el grado de modificación del paisaje derivado de las diferentes modalidades de su aprovechamiento da como resultado variaciones en las tasas de infestación doméstica. Los procesos de cambio de uso de suelo como la deforestación para la cría extensiva de ganado y la perturbación de la cubierta forestal para la producción de carbón vegetal o extracción de leña están asociados a elevadas tasas de infestación (Vázquez-Prokopec *et al.*, 2012).

La exposición de la población a los vectores constituye el principal factor de peligro para la transmisión del parásito, lo cual sumado a la vulnerabilidad humana, da como resultado un incremento en el riesgo de contacto vector-humano e incidencia de la enfermedad. Considerando que la identificación de los factores que incrementan la probabilidad de infestación doméstica permite el diseño de estrategias de control de las poblaciones de vectores a diferentes escalas espaciales (Vázquez-Prokopec *et al.*, 2012), se planteó la siguiente investigación con la finalidad de comprender cuáles elementos del paisaje, así como sus cambios, están relacionados con la presencia del vector en viviendas de comunidades rurales de Oaxaca. Bajo el marco conceptual de la ecología del paisaje (Castro *et al.*, 2019) se caracterizó espacio-temporalmente la estructura del paisaje para identificar variables de mayor relevancia para la infestación de asentamientos humanos.

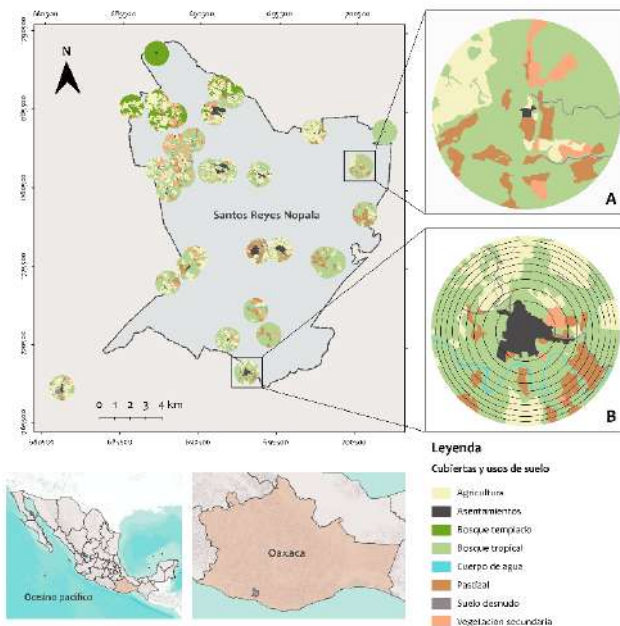
## Métodos

### *Área de estudio*

El municipio de Santos Reyes Nopala, es uno de los 570 que conforman el estado de Oaxaca, México, como se ilustra en la Figura 1. El territorio municipal abarca una extensión de 226 km<sup>2</sup> y comprende dos subprovincias fisiográficas —costa del sur y cordillera costera del sur—, en las cuales se distinguen cuatro tipos de vegetación: selva baja caducifolia, selva mediana subcaducifolia, bosque mixto de pino-encino y bosque de encino, en donde las selvas o bosques tropicales tienen mayor representa-

tividad (INEGI, 2009). El sistema agrícola de tumba-roza y quema, así como las actividades intensivas de recolección de leña han favorecido la predominancia del uso del suelo con fines ganaderos, agrícolas y para el desarrollo de infraestructura. Prevalece la agricultura de temporal como actividad económica en el municipio, seguida del comercio. Los principales cultivos son: maíz, caña, frijol y café (Pérez, 2010).

El elevado grado de marginación que presenta el municipio deriva de importantes rezagos sociales, lo que se traduce en carencia de recursos básicos como agua entubada o red de drenaje, rezago educativo y escasez en los servicios de salud (CONAPO, 2010). De acuerdo con cifras oficiales para los periodos de evaluación, dos tercios de la población total —10 441 habitantes— vive en 35 localidades rurales reportadas para el municipio (INEGI, 2010). Las viviendas en dichas localidades se caracterizan por estar construidas con tablones de madera como paredes, con techo de lámina y piso de tierra. Usualmente, la vivienda es compartida con animales domésticos como gallinas y mascotas (perros y gatos), y dentro se suele almacenar la cosecha (Ramsey *et al.*, en prensa).



**Figura 1.** Ubicación de la zona de estudio. Dentro de los círculos se representan las cubiertas y usos de suelo de cada localidad para el año 2004, interpretados en un área de 200 ha (A). Las áreas de influencia para describir la estructura del paisaje se establecieron a partir del centroide de la localidad (B).

Fuente: imágenes satelitales SPOT V, elaboración cartográfica propia.

### ***Colecta de vectores***

En colaboración con la población, auxiliares y comités de salud comunitarios, el grupo de trabajo implementó un programa de diálogo y concientización sobre la enfermedad de Chagas. Dicho programa incluyó la vigilancia permanente de presencia de vectores en las viviendas y donde mayormente pernoctan los campesinos (parcelas, campo), basado en la capacidad de reconocer (con alto valor predictivo > 95%) y coleccionar las dos especies principales conocidas en el área de estudio (Ramsey *et al.*, 2000). Se solicitó a los pobladores y los comités técnicos comunitarios la colecta de los vectores encontrados dentro de su vivienda o en el hábitat peridoméstico. Los vectores coleccionados fueron mantenidos en un contenedor de plástico debidamente etiquetado y entregados al personal del grupo de trabajo mediante actores voluntarios de los comités comunitarios.

La infestación doméstica fue estimada utilizando los registros de colecta participativa durante dos periodos de estudio: diciembre de 2003-mayo de 2006 y junio de 2015-febrero de 2016. El índice de infestación que corresponde al porcentaje de viviendas con presencia de al menos un triatomino de cualquier especie entre el total de viviendas revisadas fue calculado para cada una de las localidades (WHO, 2002).

### ***Datos para la caracterización y descripción de la estructura del paisaje***

Se elaboró cartografía de cubiertas y usos del suelo a escala 1:10 000, utilizando como fuente de información imágenes satelitales SPOT V de dos fechas: 25 de abril de 2004 y 24 de febrero de 2015. La temporalidad de las imágenes satelitales corresponde a los periodos de colecta de vectores, por lo que representan la configuración espacial de los elementos del paisaje para cada periodo de estudio.

La cartografía de cubiertas y usos de suelo de la primera fecha, se generó mediante interpretación visual de una superficie circular de 200 ha trazada a partir del centroide de cada localidad como se ilustra en la Figura 1A. La superficie se determinó con base en la revisión de literatura especializada sobre la historia natural de las especies de vectores (Barbu *et al.*, 2010; Ramsey *et al.*, 2000; Ibarra-Cerdeña *et al.*, 2009; Pech-May *et al.*, 2019). La cartografía de la segunda fecha se elaboró siguiendo el método de interpretación interdependiente (Organización de las Naciones Unidas para la Alimentación y la Agricultura —FAO, 2002). Las clases utilizadas para la interpretación fueron las siguientes: asentamientos humanos, pastizales, agricultura, bosque templado, bosque tropical, suelo desnudo y cuerpo de agua. Los asentamientos humanos corresponden al área ocupada con una o más edificaciones utilizadas como viviendas, y en caso de tratarse de un conglomerado, se integraron los elementos naturales y obras materiales como canchas deportivas, escuelas o mercados. Las áreas de pastizales son aquellas donde la vegetación dominante son los pastos, las cuales en su mayoría son de uso ganadero. El bosque de encino y el bosque mixto

fueron incluidos en la clase de bosque templado, mientras que la selva baja caducifolia y mediana subcaducifolia se catalogaron como bosque tropical. La visualización y manipulación de la información se hizo en el programa ArcGIS 10.1 (ESRI, 2011).

Para la descripción de la estructura del paisaje se calcularon métricas en diez áreas de influencia concéntricas a cada una de las localidades. En la Figura 1B se ejemplifica la delimitación de las áreas de influencia. Se calcularon las siguientes métricas utilizando el programa FRAGSTATS v3 (McGarigal *et al.*, 2012): proporción del paisaje ocupada por cada clase de cubierta o uso de suelo —PLAND— y el grado de fragmentación del bosque tropical —Splitting index, SPLIT—. Se estimó el “efecto del área de influencia” para cada variable estructural mediante una regresión logística, cuya variable dependiente fue la presencia/ausencia del vector en el ambiente doméstico. Dicho efecto alude al área en la cual la estructura del paisaje colindante predice mejor la presencia del vector en el ambiente doméstico (Miguet *et al.*, 2016).

### *Análisis estadístico*

Se realizaron análisis estadísticos paramétricos y no paramétricos en el programa R (R Core Team, 2019). Para contrastar la igualdad de proporciones de localidades donde se reportó la colecta de vectores, entre ambos periodos de estudio, se utilizó el estadístico de McNemar. En el caso de los valores del índice de infestación por especie o sin importar la especie de vector, fueron comparados entre ambos periodos de estudio usando la prueba U de Mann-Whitney.

El análisis de la influencia de la estructura del paisaje sobre la presencia de vectores en ambientes domésticos se basó en nueve hipótesis a priori, en las cuales se consideró como variables de análisis la elevación altitudinal de la localidad en msnm, el PLAND de bosque tropical, el PLAND de pastizales, el PLAND de agricultura, la fragmentación del bosque tropical y la densidad de viviendas por ha.

Las hipótesis fueron evaluadas con los siguientes modelos: 1) efecto de la altitud; 2) efecto del PLAND de bosque tropical; 3) efecto del PLAND de pastizales; 4) efecto del PLAND de pastizales y agricultura; 5) efecto de la fragmentación del bosque tropical; 6) efecto de la altitud, la densidad de viviendas y el PLAND de bosque tropical; 7) efecto de la altitud, el PLAND de pastizales y la fragmentación del bosque tropical; 8) modelo global el cual incluye todas las variables sin interacciones; 9) modelo nulo, sin la inclusión de variables explicativas. Para evaluar el soporte de los modelos se usaron modelos lineales generalizados, al igual que el criterio de información de Akaike corregido para tamaños de muestra pequeños, AICc, y los pesos de Akaike,  $w_i$ . Aquellos modelos con un  $\Delta AICc \leq 6$  entre el mejor modelo y un modelo dado, fueron considerados como el subconjunto de modelos con mayores niveles de soporte empírico. Los parámetros ponderados y sus errores estándar fueron utilizados para calcular la razón de momios —odds ratio—, y sus intervalos de con-



fianza asociados —95% IC— de cada variable de análisis. Esta métrica permite determinar la importancia relativa de cada variable para predecir los datos (Symonds y Moussalli, 2011).

## Resultados

Durante el periodo 2003-2006 el índice de infestación promedio fue de 2,7% con desviación estándar de 5,6%, encontrándose vectores en 13 localidades. En el segundo periodo 2015-2016 fueron 15 localidades en las cuales se reportó un índice de infestación promedio de 5,5% con desviación estándar de 7,7%. La infestación se mantuvo en diez localidades, mientras que para el periodo 2015-2016 se identificó nueva infestación en cuatro localidades. La variación en el índice de infestación global encontrado entre ambos periodos de estudio no difiere significativamente —Mann-Whitney  $W=524$ ,  $p > 0,05$ .

Las dos especies de triatomino colectadas en el ambiente doméstico y sus alrededores en ambos tiempos corresponden a las especies *T. dimidiata* Hg2 y *T. mazzottii*. En el primer periodo también se colectaron ejemplares aislados de *T. bolivari*, *Panstrongylus rufotuberculatus* y *Eratyrus cuspidatus*. El índice de infestación promedio y la proporción de localidades con presencia del vector *T. dimidiata* no difiere significativamente entre un periodo y otro —McNemar's chi-squared= 0,1666,  $df= 1$ ,  $p > 0,05$ . Se observó un incremento en el número total de localidades infestadas por esta especie: de 12 localidades con infestación promedio de 2,1% con desviación estándar de 5,1%, a 14 localidades con una infestación promedio de 2,1% con desviación estándar de 3,1% —Mann-Whitney,  $W= 560$ ,  $p > 0,05$ . En el caso de *T. mazzottii*, se reportó un aumento del 100 % en el número de localidades con su presencia en el segundo periodo de estudio, pasando de siete a 14 —McNemar's chi-squared=3,125,  $df= 1$ ,  $p > 0,05$ . Durante el periodo 2003-2006 el vector se encontró en siete localidades con un índice de infestación promedio de 0,8% con desviación estándar de 1,9 %, mientras que para el segundo periodo el índice de infestación promedio resultó en un incremento significativo de 4,1% con desviación estándar de 6,5%, —Mann-Whitney,  $W= 471$ ,  $p < 0,05$ .

Existe un traslape en el rango altitudinal al cual se encuentran ambas especies de vectores dentro del territorio municipal. El límite inferior altitudinal de *T. dimidiata* es de 120 msnm con presencia hasta los 1 017 msnm y 1 287 msnm, en el primero y segundo periodo de estudio, respectivamente. La presencia de la especie *T. mazzottii* en el primer periodo tuvo un rango altitudinal de 216-1017 msnm y en el segundo periodo desde los 120 hasta los 1287 msnm. El índice de infestación promedio de la especie *T. mazzottii* muestra una correlación inversa con la altitud, es decir, los índices de infestación por dicha especie decrecen conforme aumenta la altitud, mientras que para la especie *T. dimidiata* no existe una tendencia clara (Tabla 1).

**Tabla 1**  
**Índice de infestación promedio (SD) de dos especies de vectores de la enfermedad de Chagas de acuerdo con la altitud de la localidad de colecta**

<i>Rango altitudinal (msnm)</i>	<i>Periodo 2003-2006 (SD, N)</i>	<i>Periodo 2015-2016 (SD, N)</i>
<i>T. dimidiata</i>		
120-556	0,0385 (0,0372, 9)	0,0580 (0,0285, 9)
557-1017	0,1240 (0,1004, 3)	0,0361 (0,0161, 3)
1018-1420	0	0,0525 (0,0063, 2)
<i>T. mazzottii</i>		
120-556	0,0400 (0,0243, 6)	0,1167 (0,0662, 8)
557-1017	0,0294 (0, 1)	0,1052 (0,0479, 3)
1018-1420	0	0,0910 (0,0322, 2)

**Fuente:** Elaboración propia.

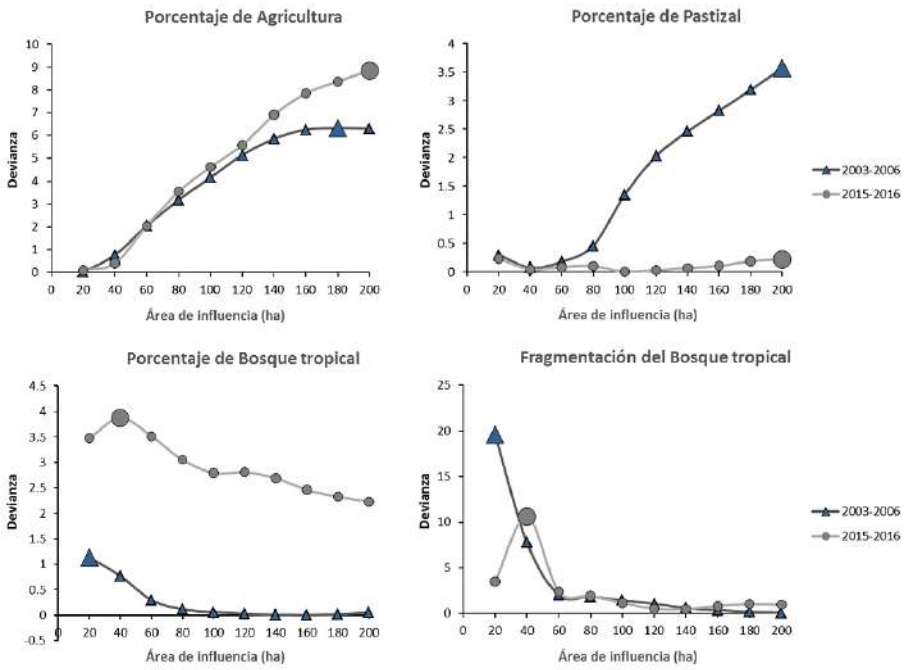
### *Efecto del área de influencia*

En una primera aproximación el PLAND de áreas agrícolas, pastizales y bosque tropical presentaron bajos valores de AIC y alta devianza explicada en las áreas correspondientes a 180 ha, 200 ha y 20 ha respectivamente (Tabla 2). Es notorio que las variaciones relativas al porcentaje ocupado por bosque tropical son importantes para la infestación en un entorno inmediato al ambiente doméstico. No se encontraron cambios marcados en cuanto a la identificación del “efecto del área de influencia” para el segundo periodo como se compara en la Figura 2. Se detectó un efecto a las 200 ha para el PLAND de áreas agrícolas y pastizales, mientras que el efecto para el PLAND de bosque tropical y fragmentación del mismo fue más evidente en un área de 40 ha.

**Tabla 2**  
**Valores de Devianza y AIC para determinar “efecto de área de influencia” de las variables de la estructura del paisaje**

<i>Variable de la estructura del paisaje</i>	<i>Periodo 2003-2006</i>	<i>Periodo 2015-2016</i>
Porcentaje de agricultura	6,33 (47,25)	8,85 (47,56)
Porcentaje de pastizal	3,57 (48,52)	0,22 (51,69)
Porcentaje de bosque tropical	1,13 (49,65)	3,87 (49,95)
Fragmentación del bosque	19,59 (41,12)	10,61 (45,71)

**Fuente:** Elaboración propia.



**Figura 2.** Efecto del área de influencia en la cual la estructura del paisaje predice mejor la presencia de vectores en el ambiente doméstico. El símbolo de mayor tamaño indica el área de influencia con el valor de devianza más grande y el valor de AIC más bajo. Fuente: elaboración propia.

### Factores que influyen en la infestación de las localidades

La identificación de los modelos *a priori* que mejor predicen la presencia del vector en ambientes domésticos se realizó en dos etapas, de acuerdo con el período de estudio. En ninguna etapa hubo multicolinealidad entre las variables,  $VIF < 4$ . Los modelos 8, 6 y 7 fueron los que presentaron mayor soporte para el primer periodo,  $\Delta AICc \leq 6$ . Estos modelos explican el 77,8%, 52,7% y el 48% de la variación en la probabilidad de presencia de un vector en una localidad, respectivamente (Tabla 3). En esta primera etapa de selección de modelos, la densidad de viviendas y la altitud a la cual se encuentra la localidad son variables con un efecto importante en la presencia del vector en las localidades (Tabla 4). Se identificó una asociación de protección estadísticamente significativa, lo que sugiere que la presencia del vector en ambientes domésticos de las localidades de estudio disminuye cuando la altitud y la densidad de viviendas aumentan. En las localidades donde la elevación altitudinal es mayor a los 1 122 msnm y existían más de 21 viviendas/ha no hubo infestación doméstica.

**Tabla 3**  
**Modelos *a priori* que explican la presencia del vector en el ambiente doméstico durante el periodo de estudio 2003-2006**

<i>Modelo</i>	<i>Variable</i>	<i>K</i>	<i>Log<sub>e</sub>(L)</i>	<i>AIC<sub>c</sub></i>	<i>ΔAIC<sub>c</sub></i>	<i>w<sub>i</sub></i>
Global *	Todas las variables	7	- 6,0501	26,2718	0,0000	0,8077
6 *	Efecto de la altitud, la densidad de viviendas y el porcentaje de bosque tropical en 20 ha	4	-10,9200	29,9009	3,6292	0,1316
7 *	Efecto de la altitud, el porcentaje de pastizal en 200 ha y la fragmentación del bosque tropical en 20 ha	4	-11,9203	31,9015	5,6298	0,0484
1	Efecto de la altitud	2	-15,4093	34,8368	8,5650	0,0112
5	Fragmentación del bosque tropical en 20 ha	2	-18,5646	41,1474	14,8757	0,0005
4	Efecto del porcentaje de pastizal en 200 ha y el porcentaje de agricultura en 180 ha	3	-18,5044	43,0454	16,7736	0,0006
Nulo		1	-23,0899	48,1859	21,9141	0,0001
3	Efecto del porcentaje de pastizal en 200 ha	2	-22,2645	48,5473	22,2755	0,0000
2	Efecto del porcentaje de bosque tropical en 20 ha	2	-22,8284	49,6750	23,4032	0,0000

*K*, número de parámetros estimado para cada modelo; *Log<sub>e</sub>(L)*, función logística de máxima verosimilitud; *AIC<sub>c</sub>*, criterio de información de Akaike; *ΔAIC<sub>c</sub>*, diferencia de *AIC<sub>c</sub>* y *w<sub>i</sub>* pesos de Akaike. Con el símbolo \* se indican los modelos de mayor soporte.

**Fuente:** Elaboración propia.

Fueron cinco los modelos equivalentes al mejor modelo con  $\Delta AIC_c \leq 2$  de la selección de modelos de la segunda etapa de evaluación (Tabla 5). El modelo con mayor soporte fue el que incluye el efecto de la fragmentación del bosque tropical a 40 ha, el modelo 5. Sin embargo, todos los modelos, incluyendo el modelo global y nulo, tienen un  $\Delta AIC_c \leq 6$ . La variabilidad explicada por el mejor modelo es del 11,17%, aunque el modelo global explica el 23,33% de la variación en la probabilidad de presencia del vector en el ambiente doméstico de las localidades (Tabla 5). Se sugieren como variables importantes la altitud y el porcentaje ocupado por el bosque tropical en un entorno inmediato al ambiente doméstico, debido a que los intervalos de confianza de la razón de momios no tienen un traslape con uno o el traslape es ligero (Tabla 6). En este caso, la asociación de protección sugiere que la presencia del vector en el ambiente doméstico disminuye conforme el porcentaje ocupado por el bosque tropical aumenta.

**Tabla 4**  
**Coefficientes de regresión de los modelos promediados (COEF), error estándar (SE), razón de momios (OR) e intervalos de confianza (IC) del 95 %, a partir de los modelos a priori durante el periodo 2003-2006**

Variable	$\Sigma w_i$	COEF	EE	OR (IC 95 %)
Altitud	0,9988	-0,0055	0,0034	0,9945 (0,9879-1,0011)*
Porcentaje de Bosque tropical en 20 ha	0,9393	0,0572	0,1014	1,0588 (0,8680-1,2917)
Porcentaje de Pastizal en 200 ha	0,8567	0,2359	0,1693	1,2660 (0,9086-1,7641)
Porcentaje de Agricultura en 180 ha	0,8083	0,3230	0,1937	1,3813 (0,9451-2,0190)
Fragmentación del bosque tropical en 20 ha	0,8566	0,0201	0,0326	1,0203 (0,9571-1,0876)
Densidad de viviendas	0,9393	-0,3386	0,1982	0,7128 (0,4833-1,0513)*

Con el símbolo \* se indican las variables consideradas como importantes.

**Fuente:** Elaboración propia.

**Tabla 5**  
**Modelos a priori que explican la presencia del vector en el ambiente doméstico durante el periodo de estudio 2015-2016**

Modelo	Variable	K	$Log_e(L)$	$AIC_c$	$\Delta AIC_c$	$w_i$
5 *	Fragmentación del bosque tropical en 40 ha	2	-21,2317	46,4816	0,0000	0,1598
1 *	Efecto de la altitud	2	-21,3266	46,6714	0,1898	0,1453
6 *	Efecto de la altitud, la densidad de viviendas y el porcentaje de bosque tropical en 40 ha	4	-19,3715	46,8040	0,3225	0,1360
7 *	Efecto de la altitud, el porcentaje de pastizal en 200 ha y la fragmentación del bosque tropical en 40 ha	4	-19,4704	47,0018	0,5202	0,1232
4 *	Efecto del porcentaje de pastizal en 200 ha y el porcentaje de agricultura en 200 ha	3	-20,5870	47,2105	0,7289	0,1110
Nulo		1	-23,9018	49,8096	3,3281	0,0908
2	Efecto del porcentaje de bosque tropical en 40 ha	2	-22,9762	49,9707	3,4891	0,1675
Global	Todas las variables	7	-18,3257	50,8229	4,3413	0,0547
3	Efecto del porcentaje de pastizal en 200 ha	2	-23,8486	51,7152	5,2337	0,0117

K, número de parámetros estimado para cada modelo;  $Log_e(L)$ , función logística de máxima verosimilitud;  $AIC_c$ , criterio de información de Akaike;  $\Delta AIC_c$ , diferencia de  $AIC_c$  y  $w_i$  pesos de Akaike. Con el símbolo \* se indican los modelos de mayor soporte.

**Fuente:** Elaboración propia.

Tabla 6

**Coefficientes de regresión de los modelos promediados (COEF), errores estándar (SE), razón de momios (OR) e intervalos de confianza (IC) del 95%, a partir de los modelos *a priori* durante el periodo 2015-2016**

<i>Variable</i>	$\Sigma w_i$	<i>COEF</i>	<i>EE</i>	<i>OR (IC 95 %)</i>
Altitud	0,4592	-0,0024	0,0012	0,9976 (0,9952-1,0000)*
Porcentaje de Bosque tropical en 40 ha	0,3582	-0,0244	0,0193	0,9759 (0,9397-1,0136)*
Porcentaje de Pastizal en 200 ha	0,3006	0,0294	0,0538	1,0299 (0,9268-1,1444)
Porcentaje de Agricultura en 200 ha	0,1657	0,0597	0,0375	1,0615 (0,9863-1,1425)
Fragmentación del bosque tropical en 40 ha	0,3377	0,0062	0,0039	1,0062 (0,9984-1,0140)
Densidad de viviendas por ha	0,1907	0,0238	0,0805	1,0241 (0,8746-1,1990)

Con el símbolo \* se indican las variables consideradas como importantes.

**Fuente:** Elaboración propia.

## Discusión y conclusiones

El riesgo de contacto vector-humano en ambientes domésticos no es enteramente explicado por la calidad y carencias de la vivienda (Enger *et al.*, 2004; Ramsey *et al.*, 2005; Cohen *et al.*, 2006; Bustamante *et al.*, 2009; Dumonteil *et al.*, 2013; Grijalva *et al.*, 2017), como la literatura temprana sobre la epidemiología de la enfermedad de Chagas en Sudamérica lo señalaba. En la actualidad varios estudios señalan que una proporción de la transmisión de *T. cruzi* al humano no ocurre en el hábitat doméstico, sino en áreas de ecotono o silvestres (Ramsey *et al.*, 2012; López-Cancino *et al.*, 2015; Valdez-Tah *et al.*, 2015a, b). Los resultados del presente estudio muestran evidencia de que la pérdida de bosque tropical y sus variaciones espaciales en un entorno inmediato al ambiente doméstico, en combinación con una baja densidad de vivienda, y la elevación altitudinal, están asociados con un incremento en la presencia doméstica de las dos especies más abundantes de triatomos.

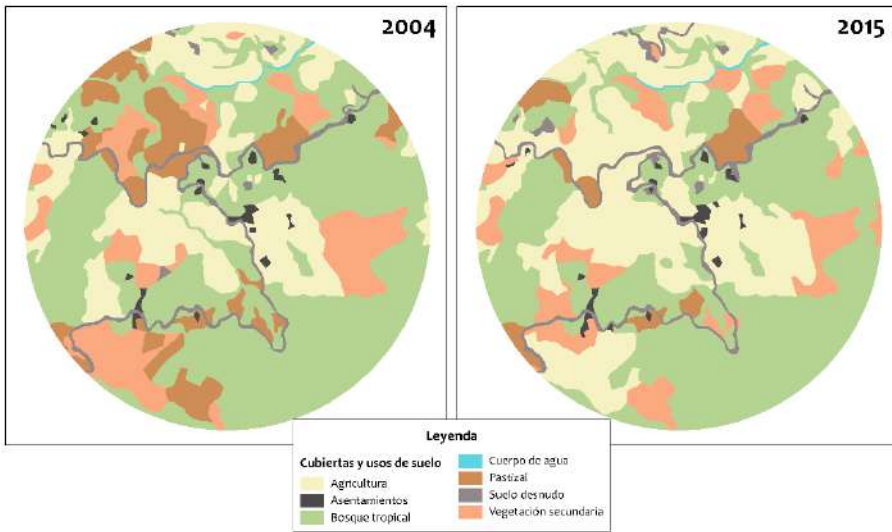
Históricamente en el municipio de Santos Reyes Nopala se ha reportado la presencia de tres especies de vectores (Ramsey *et al.*, 2000; Ibarra-Cerdeña *et al.*, 2009). Una era invasora, *Rhodnius prolixus* (Hashimoto y Schofield, 2012) y no se ha colectado en México desde 2005, mientras *T. dimidiata* y *T. mazzottii* han sido las dos especies con una dispersión y capacidad sinantrópica más importantes. En los 10 años que hubo entre las evaluaciones del presente estudio se ha presentado una ampliación en el gradiente altitudinal de la distribución de ambas especies de vectores, llegando hasta los 1 785 msnm para *T. dimidiata* y 1 395 msnm para *T. mazzottii*. La evidencia

encontrada en este estudio indica que la infestación a nivel municipal se ha duplicado con el tiempo y que dicho cambio se debe principalmente a un incremento en la infestación por el vector *T. mazzottii* en altitudes menores. Se trata de una especie distribuida tanto en bosques tropicales como en templados y con alta capacidad para alimentarse de sangre de aves con un desarrollo ventajoso (Ibarra-Cerdeña *et al.*, 2009). Con un incremento del 100% en el número de localidades infestadas por *T. mazzottii*, al igual que un aumento significativo en el porcentaje de viviendas infestadas —de 0,8% a 4,1%—, esta especie es el vector predominante no sólo en el municipio, sino también en el estado de Oaxaca (Ramsey *et al.*, 2015).

La deforestación de los bosques tropicales para su conversión en zonas agropecuarias favorece la presencia de vectores en el ambiente doméstico (Ibarra-Cerdeña *et al.*, 2017) como se ha reportado en paisajes rurales de Panamá (Gottenker *et al.*, 2011) y Argentina (Vázquez-Prokopec *et al.*, 2012). Los bosques tropicales albergan una gran variedad de hospederos reservorios del parásito *T. cruzi* y hospederos de los vectores, mamíferos silvestres (Izeta-Alberdi *et al.*, 2016). Los diferentes procesos de cambio de esta cubierta modifican las comunidades de hospederos en términos de abundancia y diversidad (Ibarra-Cerdeña *et al.*, 2017). Los animales pecuarios y mascotas (ganado, gallinas, perros, gatos) y la misma población humana son alimento abundante y continuo para los vectores, es por ello que las actividades de pastoreo y la explotación de los recursos forestales, como la recolección de leña y cacería, han sido documentadas como factores asociados a elevados índices de infestación doméstica (Vázquez-Prokopec *et al.*, 2012). Aunque en el presente estudio las zonas agrícolas o los pastizales *per se* no tuvieron efecto significativo en la infestación de las localidades, se ha hipotetizado que estas actividades podrían desempeñar un papel importante en la incidencia de la enfermedad en algunos estados del sur de México (Ramsey *et al.*, 2012; López-Cancino *et al.*, 2015). La agricultura es una de las principales actividades ocupacionales de la población y es en estas áreas de trabajo donde se sugiere que puede tener lugar la transmisión vectorial (Ibañez-Cervantes *et al.*, 2018).

Los factores socioeconómicos no fueron incluidos en este trabajo debido a que la mayoría de las viviendas en el municipio cuentan con las características de áreas, documentadas en estudios previos, que presentan altos porcentajes de infestación (Ramsey *et al.*, 2002; Bustamante *et al.*, 2014). Los esfuerzos gubernamentales para abatir la pobreza en el municipio se reflejan en el mejoramiento de las condiciones de las viviendas y el suministro de servicios básicos, debido a la implementación de diversos programas para dignificar dicho espacio (Ayuntamiento de Nopala, 2014). Sin embargo, estos programas han promovido la construcción de viviendas —frecuentemente sin la destrucción de viviendas viejas— y el subsecuente desarrollo de las actividades productivas, lo que puede tener consecuencias directas o indirectas (positivas o negativas) sobre la prevalencia de infestación. Un ejemplo de ello es el

caso particular de la localidad El Aguacatal, ubicada al noroeste de la cabecera municipal. En esta localidad se registró la colecta del vector *T. dimidiata* únicamente para el periodo 2015-2016, con una infestación doméstica de 5,8%. La deforestación fue un proceso predominante alrededor de la localidad desde 2011 debido a la expansión de la frontera agrícola como se ilustra en la Figura 3 (SAGARPA y FAO, 2015).



**Figura 3.** Comparativa del paisaje en El Aguacatal, Santos Reyes Nopala, donde se muestra la expansión de las áreas agrícolas (cultivo de caña de azúcar).

**Fuente:** imágenes satelitales SPOT V, elaboración cartográfica propia.

En México, desde el 2007 se ha documentado un aumento considerable en la detección de nuevos casos de la enfermedad de Chagas, principalmente debido al incremento en investigación y presencia de grupos académicos en las comunidades, y en menor medida a programas estatales encaminados a la detección de la presencia del vector en vivienda y a actividades para eliminar la transmisión transfusional (Ibañez-Cervantes *et al.*, 2018). Actualmente, las estrategias y líneas de acción para el control de la transmisión intradomiciliar se basan en la estratificación de riesgo (a pesar de no tener definición clara de esta), mejoramiento de la vivienda y control químico del vector en áreas prioritarias, definidas así por ser endémicas y por tener la necesidad de acciones de respuesta (CENAPRECE, 2018). Los resultados que aquí se presentan pueden ayudar a enriquecer la definición de riesgo de exposición al vector en México, de manera que se sugiere priorizar las estrategias de información y de vigilancia co-



munitaria de presencia y contacto en las localidades de menor altitud de este municipio. Es en estos lugares donde prevalece el movimiento continuo entre el ambiente doméstico y los cultivos o el bosque, y una carencia por calidad y servicios básicos de la vivienda, más aún en los asentamientos pequeños (Ramsey, en prensa). Acciones como el almacenaje de alimentos y confinamiento de animales en lugares lejanos del ambiente doméstico, al igual que el uso de pabellones durante la noche en el caso de dormir en las áreas de pastoreo, ayudarán a prevenir el contacto de los humanos con el vector. Por lo tanto es importante tomar medidas preventivas en el paisaje completo debido a las diferentes relaciones que los pobladores establecen con su entorno natural (Valdez-Tah *et al.*, 2015b).

El control de poblaciones de vectores requiere una apropiada estrategia de implementación, cobertura, aceptación y uso de medidas de control para generar un impacto adecuado. En el presente trabajo se enfatiza la comprensión de las características del paisaje que propician la presencia en el hábitat doméstico y el peligro de contacto entre el vector y el humano en un contexto local. El enfoque de paisaje con la finalidad de determinar patrones espaciales relacionados con la ampliación de riesgo de enfermedades transmitidas por vectores ha sido poco abordado e incluido en estudios epidemiológicos o eco-epidemiológicos. El uso de dicho enfoque brinda una oportunidad para discriminar los diferentes componentes del paisaje que influyen en la infestación doméstica y provee evidencia para dar a conocer a la población y con ello cambiar prácticas peligrosas. Se debe fomentar el diálogo entre la población y los diferentes sectores institucionales para evitar prácticas y programas que promuevan o amplifiquen la infestación (agropecuario, de salud y académico). Es a través de los esfuerzos intersectoriales e interinstitucionales que se fortalecen los programas y actividades de gobierno que mitigan la exposición, aunque las prácticas y cambios estructurales o de vigilancia permitirán reducir la vulnerabilidad para la transmisión vectorial en procesos locales.

## Bibliografía

- Ayuntamiento de Nopala (2014). *Plan municipal de desarrollo 2014-2016*, Santos Reyes Nopala, Oaxaca.
- Barbu C., Dumonteil E. & Gourbière S. (2010). "Characterization of the dispersal of non-domiciliated *Triatoma dimidiata* through the selection of spatially explicit models", *PLoS Neglected Tropical Diseases*, 4(8), e777. <https://doi.org/10.1371/journal.pntd.0000777>
- Brito R.N., Gorla D.E., Diotaiuti L., Gomes A.C.F., Souza R.C.M. & Abad-Franch F. (2017). "Drivers of house invasion by sylvatic Chagas disease vectors in the Amazon-Cerrado transition: A multi-year, state-wide assessment of municipality-aggregated surveillance data", *PLoS Neglected Tropical Diseases*, 11(11), e0006035. <https://doi.org/10.1371/journal.pntd.0006035>

- Bustamante D.M., De Urioste-Stone S.M., Juárez J.G. & Pennington P.M. (2014). "Ecological, social and biological risk factors for continued *Trypanosoma cruzi* transmission by *Triatoma dimidiata* in Guatemala", *PLoS ONE*, 9(8), e104599. <https://doi.org/10.1371/journal.pone.0104599>
- Castro F., Bunting S.C., Strand E. & Godinho-Ferreira P. (eds.) (2019). *Applied Landscape Ecology* (1a), UK. Editorial Wiley, pp. 461.
- CENAPRECE (2018). "Centro Nacional de Programas Preventivos y Control de Enfermedades. Avances, logros y retos 2013-2018", Secretaría de Salud, México. Recuperado de [https://www.gob.mx/cms/uploads/attachment/file/398612/ALR\\_CENAPRECE\\_opt\\_1.pdf](https://www.gob.mx/cms/uploads/attachment/file/398612/ALR_CENAPRECE_opt_1.pdf)
- Cohen J.M., Wilson M.L., Cruz-Celis A., Ordoñez R. & Ramsey J. M. (2006). "Infestation by *Triatoma pallidipennis* (Hemiptera: Reduviidae: Triatominae) is associated with housing characteristics in rural Mexico", *Journal of Medical Entomology*, 43(6): 1252-1260. <https://doi.org/10.1093/jmedent/43.6.1252>
- CONAPO (2010). "Índice de marginación por entidad federativa y municipio 2010". Recuperado de [http://www.cnapo.gob.mx/en/CONAPO/Indices\\_de\\_Marginacion\\_2010\\_por\\_entidad\\_federativa\\_y\\_municipio](http://www.cnapo.gob.mx/en/CONAPO/Indices_de_Marginacion_2010_por_entidad_federativa_y_municipio)
- Dumonteil E., Nouvellet P., Rosecrans K., Ramirez-Sierra M.J., Gamboa-León R., Cruz-Chan V., Rosado-Vallado M. & Gourbière, S. (2013). "Eco-Bio-Social determinants for house infestation by non-domiciliated *Triatoma dimidiata* in the Yucatan Peninsula, Mexico", *PLoS Neglected Tropical Diseases*, 7(9), e2466. <https://doi.org/10.1371/journal.pntd.0002466>
- Dumonteil E., Ramírez-Sierra M.J., Ferral J., Euan-García M. & Chavez-Núñez L., (2009). Usefulness of community participation for the fine temporal monitoring of house infestation by non-domiciliated triatomines. *The Journal of Parasitology*, 95(2), 469–471. <https://doi.org/10.1645/GE-1712.1>
- Enger K.S., Ordoñez R., Wilson M.L. & Ramsey J.M. (2004). "Evaluation of risk factors for rural infestation by *Triatoma pallidipennis* (Hemiptera: Triatominae), a Mexican vector of Chagas disease", *Journal of Medical Entomology*, 41: 760-767. <https://doi.org/10.1603/0022-2585-41.4.760>
- ESRI (2011). ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
- FAO (2002). "FRA2000, pan-tropical survey of forest cover changes 1980-2000 – methods and implementation, FAO FRA working paper 49a", Rome. Recuperado de <http://www.fao.org/3/a-ak592e.pdf>
- Gottdenker N.L., Calzada J.E., Saldaña A. & Carroll C.R. (2011). "Association of anthropogenic land use change and increased abundance of the Chagas disease vector *Rhodnius pallescens* in a rural landscape of Panama", *American Journal of Tropical Medicine and Hygiene*, 84(1): 70-77. <https://doi.org/10.4269/ajtmh.2011.10-0041>.

- Grijalva M.J., Villacis A.G., Moncayo A.L., Ocaña-Mayorga S., Yumiseva C.A., Baus E.G. (2017). “Distribution of triatomine species in domestic and peridomestic environments in central coastal Ecuador”, *PLoS Negl Trop Dis*, 11(10): e0005970.
- Hashimoto K. & Schofield C.J. (2012). “Elimination of *Rhodnius prolixus* in Central America”, *Parasites & Vectors*, 5: 45. <https://doi.org/10.1186/1756-3305-5-45>
- Ibarra-Cerdeña C.N., Sánchez-Cordero V., Peterson A. T. & Ramsey, J.M., (2009). “Ecology of North American Triatominae”, *Acta Tropica*, 110: 178-186. <https://doi.org/10.1016/j.actatropica.2008.11.012>
- Ibarra-Cerdeña C.N., Valiente-Banuet L., Sánchez-Cordero V., Stephens C.R. & Ramsey J.M. (2017). “*Trypanosoma cruzi* reservoir-triatomine vector co-occurrence networks reveal meta-community effects by synanthropic mammals on geographic dispersal”, *PeerJ*, 5, e3152. <https://doi.org/10.7717/peerj.3152>
- INEGI (2009). “Prontuario de información geográfica municipal de los Estados Unidos Mexicanos. Santos Reyes Nopala, Oaxaca”. Recuperado de [http://www3.inegi.org.mx/contenidos/app/mexicocifras/datos\\_geograficos/20/20526.pdf](http://www3.inegi.org.mx/contenidos/app/mexicocifras/datos_geograficos/20/20526.pdf)
- (2010). Censo de Población y Vivienda 2010. Principales resultados por localidad (ITER). Recuperado de [http://www.inegi.org.mx/sistemas/consulta\\_resultados/iter2010.aspx](http://www.inegi.org.mx/sistemas/consulta_resultados/iter2010.aspx)
- Izeta-Alberdi I., Ibarra-Cerdeña C.N., Moo-Llanes D.M. & Ramsey J.M. (2016). “Geographical, landscape and host associations of *Trypanosoma cruzi* DTUs and lineages”, *Parasites & Vectors*, 9: 631. <https://doi.org/10.1186/s13071-016-1918-2>
- López-Cancino S.A., Tun-Ku E., De la Cruz-Felix H.K., Ibarra-Cerdeña C.N., Izeta-Alberdi A., Pech-May A., Mazariegos-Hidalgo C.J., Valdez-Tah A. & Ramsey J. M. (2015). “Landscape ecology of *Trypanosoma cruzi* in the southern Yucatan Peninsula”, *Acta Tropica*, 151: 58-72. <https://doi.org/10.1016/j.actatropica.2015.07.021>
- McGarigal K., Cushman S.A. & Ene E., (2012). *FRAGSTATS v4: Spatial pattern analysis program for categorical maps and continuous maps*, Amherst, University of Massachusetts.
- Miguet P., Jackson H.B., Jackson N.D., Martin A.E. & Fahrig L. (2016). “What determines the spatial extent of landscape effects on species?”, *Landscape Ecology*, 31: 1177-1194. <https://doi.org/10.1007/s10980-015-0314-1>
- Pech-May A., Mazariegos-Hidalgo C.J., Izeta-Alberdi A., López-Cancino S.A., Tun-Ku E., De la Cruz-Felix K., Ibarra-Cerdeña C.N., González R.E. & Ramsey J.M., (2019). “Genetic variation and phylogeography of the *Triatoma dimidiata* complex evidence a potential center of origin and recent divergence of haplogroups

- having differential *Trypanosoma cruzi* and DTU infections”, *PLoS Neglected Tropical Diseases*, 13(1). <https://doi.org/10.1371/journal.pntd.0007044>
- Pérez J.G. (ed.) (2010). *Buscando el origen de los Chatinos de Nopala* (3a), Oaxaca, México, Editorial Carteles, pp. 160.
- Ramírez-Sierra M.J., Herrera-Aguilar M., Gourbière S. & Dumonteil E. (2010). “Patterns of house infestation dynamics by non-domiciliated *Triatoma dimidiata* reveal a spatial gradient of infestation in rural villages and potential insect manipulation by *Trypanosoma cruzi*”, *Tropical Medicine and International Health*, 15(1): 77-86. <https://doi.org/10.1111/j.1365-3156.2009.02422.x>
- R Core Team, (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ramsey J.M., Ordoñez R., Cruz-Celis A., Alvear A. L., Chavez V., Lopez R., Pintor J. R., Gama F. & Carrillo S. (2000). “Distribution of domestic Triatominae and stratification of Chagas Disease transmission in Oaxaca, Mexico”, *Medical and Veterinary Entomology*, 14: 19-30. <https://doi.org/10.1046/j.1365-2915.2000.00214.x>
- Ramsey J.M., Alvear A.L., Ordoñez R., Muñoz G., Garcia A., Lopez R. & Leyva R., (2005). “Risk factors associated with house infestation by the Chagas disease vector *Triatoma pallidipennis* in Cuernavaca metropolitan area, Mexico”, *Medical and Veterinary Entomology*, 19: 219-228. <https://doi.org/10.1111/j.0269-283X.2005.00563.x>
- Ramsey J.M., Gutierrez-Cabrera A.E., Salgado-Ramírez L., Peterson A.T., Sanchez-Cordero V. & Ibarra-Cerdeña C.N. (2012). “Ecological connectivity of *Trypanosoma cruzi* reservoirs and *Triatoma pallidipennis* hosts in an anthropogenic landscape with endemic Chagas Disease”, *PLoS ONE*, 7(9): e46013. <https://doi.org/10.1371/0046013>
- Ramsey J.M., Peterson A.T., Carmona-Castro O., Moo-Llanes D.A., Nakazawa Y., Butrick M., Tun-Ku E., De la Cruz-Félix K. & Ibarra-Cerdeña C.N. (2015). “Atlas of Mexican Triatominae (Reduviidae: Hemiptera) and vector transmission of Chagas disease”, *Memórias Do Instituto Oswaldo Cruz*, 110(3): 339-352. <https://doi.org/10.1590/0074-02760140404>
- Rossi J.C., Duarte E.C. & Gurgel-Gonçalves R. (2015). “Factors associated with the occurrence of *Triatoma sordida* (Hemiptera: Reduviidae) in rural localities of Central-West Brazil”, *Memórias Do Instituto Oswaldo Cruz*, 110(2): 192-200. <https://doi.org/10.1590/0074-02760140395>
- SAGARPA & FAO (2015). PESA en México. Recuperado de [www.pesamexico.org/PESAenMéxico/EIPESAenMéxicoysumetodología.aspx](http://www.pesamexico.org/PESAenMéxico/EIPESAenMéxicoysumetodología.aspx)
- Sánchez-González G., Figueroa-Lara A., Elizondo-Cano M., Wilson L., Novelo-Garza B., Valiente-Banuet L. & Ramsey J.M., (2016). “Cost-Effectiveness of Blood Donation Screening for *Trypanosoma cruzi* in Mexico”, *PLoS Neglected*

- Tropical Diseases*, 10(3): e0004528.  
<https://doi.org/10.1371/journal.pntd.0004528>
- Schmunis G. & Yadon Z.E. (2010). “Chagas disease: a Latin American health problem becoming a world health problem”, *Acta Tropica*, 115(1-2): 14-21.  
<https://doi.org/10.1016/j.actatropica.2009.11.003>
- Symonds M.R.E. & Moussalli A. (2011). “A brief guide to model selection, multi-model inference and model averaging in behavioural ecology using Akaike’s information criterion”, *Behavioral Ecology and Sociobiology*, 65: 13-21.  
<https://doi.org/10.1007/s00265-010-1037-6>
- Valdez-Tah A., Huicochea-Gómez L., Nazar-Beutelspacher A., Ortega-Canto J. & Ramsey J.M. (2015a). “La vulnerabilidad humana a la transmisión vectorial de *Trypanosoma cruzi* a través de los procesos de salud-enfermedad y la apropiación social del territorio”, *Salud Colectiva*, 11(2): 191-210.  
<https://doi.org/10.18294/sc.2015.683>
- Valdez-Tah A., Huicochea-Gómez L., Ortega-Canto J., Nazar-Beutelspacher A. & Ramsey J.M. (2015b). “Social representations and practices towards Triatomines and Chagas Disease in Calakmul, México”, *PLoS ONE*, 10(7): e0132830.  
<https://doi.org/10.1371/journal.pone.0132830>
- Vázquez-Prokopec G.M., Spillmann C., Zaidenberg M., Gürtler R.E. & Kitron U., (2012). “Spatial heterogeneity and risk maps of community infestation by *Triatoma infestans* in rural northwestern Argentina”, *PLoS Neglected Tropical Diseases*, 6(8): e1788. <https://doi.org/10.1371/journal.pntd.0001788>
- WHO (2002). “Control of Chagas Disease. Technical Report Series 905”. Second report of the WHO Expert Committee. Ginebra. World Health Organization, pp. 120. Recuperado de [https://apps.who.int/iris/bitstream/handle/10665/42443/WHO\\_TRS\\_905.pdf?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/42443/WHO_TRS_905.pdf?sequence=1)
- WHO (2015). “Chagas disease (American trypanosomiasis)”. Recuperado de <http://www.who.int/mediacentre/factsheets/fs340/en/>
- WHO (2017). “Respuesta mundial para el control de vectores 2017-2030”. Recuperado de [https://www.who.int/malaria/areas/vector\\_control/Draft-WHO-GVCR-2017-2030-esp.pdf](https://www.who.int/malaria/areas/vector_control/Draft-WHO-GVCR-2017-2030-esp.pdf)



# Propuesta de una guía para la evaluación de la exactitud posicional de datos espaciales

## *Proposal for a guide for the positional accuracy assessment of spatial data*

Francisco Javier Ariza-López\*

José Luis García-Balboa▲

Joselyn Robledo Ceballos♦

José Rodríguez-Avi•

Rosario Casanova■

*Recibido 20 de marzo de 2019; aceptado 13 de junio de 2019*

### Resumen

Los participantes en el proyecto de asistencia técnica del IPGH (PAT2018) (Agenda del IPGH 2010-2020) “Propuesta de adopción de metodologías y procedimientos para la evaluación de la calidad de la información geográfica para los Estados Miembros del IPGH”, decidieron abordar la realización de una propuesta de guía genérica para la evaluación de la exactitud posicional de datos espaciales. En la elaboración de la propuesta de guía se han aplicado el ciclo de Deming y un método específico de desarrollo de guías basadas en evidencias. Se ha asumido su alineación con las

\* Universidad de Jaén, España, correo electrónico: [fjariza@ujaen.es](mailto:fjariza@ujaen.es).

ORCID: <http://orcid.org/0000-0002-5204-3630>.

▲ Universidad de Jaén, España, correo electrónico: [jibalboa@ujaen.es](mailto:jibalboa@ujaen.es).

ORCID: <http://orcid.org/0000-0002-3109-5888>.

♦ Servicio Aerofotogramétrico de la Fuerza Aérea, Chile,

correo electrónico: [joselyn.robledo@gmail.com](mailto:joselyn.robledo@gmail.com). ORCID: <https://orcid.org/0000-0001-6869-3665>

• Universidad de Jaén, España, correo electrónico: [jravi@ujaen.es](mailto:jravi@ujaen.es).

ORCID: <https://orcid.org/0000-0002-1673-9876>.

■ Universidad de la República, Uruguay, correo electrónico: [casanova@fing.edu.uy](mailto:casanova@fing.edu.uy).

ORCID: <https://orcid.org/0000-0002-1873-207X>.

normas internacionales de la familia ISO 19100 y con los métodos de evaluación de la exactitud posicional más extendidos en Hispanoamérica. Igualmente, se presentan los distintos alcances asumidos para su elaboración, tanto en lo que se refiere a contenidos como a la perspectiva genérica desarrollada.

Palabras clave: *exactitud posicional, evaluación de la calidad, guía de evaluación.*

## Abstract

The participants in the Technical Assistance Project (TAP2018) of the Pan American Institute for Geography and History entitled “Proposal for the adoption of methodologies and procedures for the evaluation of the quality of the geographic information for the Member States of the PAIGH”, have decided to address the realization of a generic guide proposal for the evaluation of spatial data positional accuracy. In the elaboration of the proposal, the Deming cycle and a specific method have been applied to develop evidence-based guides. The assumed lineaments are, the alignment with the international standards of the ISO 19100 family and with the most widespread positional accuracy assessment methods in Latin America stand out. Likewise, the different assumed scopes are presented, both in terms of content and the generic perspective developed. In relation to the contents, the propositive part of the guide, which includes: a general method for the evaluation of positional accuracy, aspects related to execution, aspects related to statistics and meta-quality and a proposal for evaluation report; is the one that receives more dedication in this presentation.

Key words: *positional accuracy, quality evaluation, evaluation guide.*

## Introducción

Desde siempre, la exactitud posicional ha sido considerada como un aspecto definitorio y primordial de la calidad de todo producto cartográfico (Ariza-López, 2002), debido a que afecta a factores como la geometría, la topología y la calidad temática, y que está directamente relacionada con la interoperabilidad del dato espacial.

Dada la masificación del uso de la información geoespacial y las necesidades de interoperabilidad que exigen las diferentes aplicaciones geomáticas e Infraestructuras de Datos Espaciales (IDE), es fundamental asegurar la calidad de la información, pues es la única forma de garantizar soluciones fiables al momento de tomar decisiones.

La calidad de los productos de datos espaciales se ha de definir en sus especificaciones (p.ej. utilizando ISO 19131 —ISO, 2007). Los procesos de producción deben estar diseñados y gestionados adecuadamente para que la calidad diseñada sea asegurada en la producción y, así, alcanzada en el producto. Para verificar que lo anterior se consigue, se aplican evaluaciones de la calidad, que deben basarse en métodos



estandarizados y bien definidos. Los resultados de las evaluaciones deben ser utilizados por los productores para conocer, manejar y mejorar sus procesos de producción. Estos resultados también deben ser publicados en metadatos para que los usuarios tengan conocimiento de la calidad real de los productos que desean utilizar.

Las fuentes de error en las diferentes etapas del proceso productivo son múltiples, variadas y dependientes del propio proceso (p.ej. captura del vuelo, proceso geodésico, proceso de orientación, restitución, captura y estandarización de bases de datos, etc.). Dado este complejo escenario es imposible garantizar la perfección del producto, en nuestro caso: la ausencia de errores posicionales. Lo importante es controlar que el nivel de exactitud existente (sesgos e incertidumbres) esté acotado al final del proceso productivo, y que se cumpla con los estándares internacionales, sus parámetros y tolerancias para cada escala (resolución).

Dentro de este marco, los métodos de evaluación de la exactitud posicional (MEEP) son procesos estandarizados que permiten, o bien estimar la calidad, o controlarla. La estimación consiste en determinar un valor fiable de la propiedad de interés en el producto de datos, mientras que el control de calidad consiste en tomar una decisión de si la propiedad de interés en ese producto de datos alcanza, o no, un nivel determinado.

Existe un trabajo preliminar que permitió obtener un diagnóstico sobre los MEEP en los Países Miembros del Instituto Panamericano de Geografía e Historia (IPGH) a través del proyecto “Diagnóstico de la situación actual sobre las metodologías y procedimientos empleados para la evaluación de la calidad de la Información Geográfica”, financiado en la convocatoria de proyectos de asistencia técnica del año 2015. Este trabajo está publicado en la *Revista Cartográfica* número 94 correspondiente al año 2017 (Ariza-López *et al.*, 2017). En este trabajo se identificaron los MEEP que se están aplicando (Tabla 1), se analizó la situación actual y se valoró el grado de madurez en este aspecto. Con respecto a la situación y a los MEEP utilizados se indica: “La mayoría de los casos analizados tienen como base de la evaluación posicional métodos procedentes de los Estados Unidos. Estos métodos son el NMAS y el NSSDA. La norma del IPGH (1978) también se menciona en algunos países. De modo general, se aprecia un claro interés por mantenerse actualizados y mejorar los procesos de evaluación de la exactitud posicional, lo que lleva a adoptar los cambios que van surgiendo fuera de sus fronteras”. Y con respecto a la aplicación de estos MEEP se concluye: “...el principal problema que se detecta es el que se ha denominado “informalidad”, es decir, la situación en la que no se dispone de una norma nacional publicada o una referencia explícita a la adopción de un método concreto (p.ej. IPGH, NMAS, NSSDA, etc.)”. Además, esta informalidad, no sólo significa que no exista una adopción explícita de un MEEP ya existente o una norma o estándar propio, sino que, por lo general, también faltan documentos detallados que restrinjan la variabilidad de las múltiples interpretaciones y opciones que se pueden desarrollar

a la hora de aplicar los MEEP existentes. Por otra parte, los autores no conocemos la existencia de ninguna guía que tenga un contenido similar a la desarrollada. El manual sobre la aplicación del NSSDA (MPLMIC, 1999) presenta un carácter aplicado pero muy limitado, y otros documentos centrados en esta materia (p.ej. Poggioli, 2010, o el capítulo 3 de Congalton y Green, 2009) no realizan un tratamiento con suficiente detalle y con una orientación de guía aplicable.

**Tabla 1**  
**Uso de estándares identificados en Hispanoamérica**

<i>País</i>	<i>Norma</i>	<i>Referencia</i>
Argentina (Provincia de Santa Fe)	Sí	UNL&SCIT (2004). Norma cartográfica de la Provincia de Santa Fe
Brasil	Sí	ET-CQDG (2016): Norma da especificação técnica para controle de qualidade de dados geoespaciais (ET-CQDG). Decreto Nº 89.817, DE 20 DE JUNHO DE 1984. Estabelece as Instruções Reguladoras das Normas Técnicas da Cartografia Nacional
Chile	No	Informantes. Se aplican: NSSDA, ASPRS, EMAS
Colombia	Sí	NTC 5205 - Precisión de datos espaciales ASPRS (2015). ASPRS Positional Accuracy Standards for Digital Geospatial Data IPGH (1978). Especificaciones para mapas topográficos. Panamá: Instituto Panamericano de Geografía e Historia —IPGH, Panamá, 1978
Ecuador	No	Informantes. Se aplica NSSDA
España	Sí	UNE 148002:2016. Metodología de evaluación de la exactitud posicional de la información geográfica
México	Sí	INEGI. Norma técnica sobre estándares de exactitud posicional INEGI. Compendio de criterios y especificaciones técnicas para la generación de datos e informaciones de carácter fundamental. 07 geodesia- 07.2 Estándares de Exactitud Posicional
Perú	No	Instituto Geográfico Nacional normas técnicas “ especificaciones técnicas para la producción de mapas topográficos a escala de 1:100 000”
Puerto Rico	Sí	Estándar TIG-005-004: Exactitud posicional de los datos. Oficina de Gerencia y Presupuesto
Uruguay	No	Informantes
Estados Unidos	Sí	NMAS (USBB, 1947), NSSDA (FGDC, 1998) , ASPRS (2015)
Venezuela	No	IGVSB (2012). Especificaciones técnicas para la revisión de cartografía básica a escala 1:1 000. Instituto Geográfico de Venezuela Simón Bolívar

**Fuente:** Ariza-López *et al.* (2017).

Finalmente, con base en el análisis realizado en el mencionado trabajo se cita como idea principal a desarrollar la “Visión de Aplicación del Marco Normativo de las Américas” (NN.UU., 2013), el cual indica: “Es preciso definir y adoptar un marco normativo compatible en la región, que establezca las convenciones comunes y acuerdos técnicos necesarios para alcanzar mayor eficiencia en la respuesta a las demandas de información geoespacial, donde los datos se generen y mantengan el común denominador de: compatibilidad, comparabilidad, confiabilidad, consistencia y completitud, siendo éste la base para el establecimiento de un esquema interoperable de colaboración, que contribuya al desarrollo de la Infraestructura de Datos Geoespaciales de las Américas (IDEA)”.

Dados los antecedentes expuestos, dentro del proyecto de asistencia técnica del IPGH (PAT2018) (Agenda del IPGH 2010-2020) “Propuesta de adopción de metodologías y procedimientos para la evaluación de la calidad de la información geográfica para los Estados Miembros del IPGH” ([https://coello.ujaen.es/investigacion/web\\_giic/SubWeb\\_IPGH2018/index.htm](https://coello.ujaen.es/investigacion/web_giic/SubWeb_IPGH2018/index.htm)), en el que participaban técnicos de Argentina, Brasil, Chile, Ecuador, España, México y Uruguay, se consideró pertinente la creación de una guía genérica teórico-práctica para la evaluación de la exactitud posicional que pudiera ayudar a los técnicos de los organismos públicos cartográficos, y demás agentes del sector (profesionales, docentes, alumnos, etc.), a mejorar la calidad de los procesos de evaluación de la exactitud posicional.

El objetivo de este trabajo es presentar el método seguido para desarrollar esta guía genérica, así como sus contenidos más sustanciales. El documento se organiza de la siguiente manera: la segunda sección presenta el método aplicado, posteriormente en la tercera se indican los lineamientos y los alcances adoptados; el apartado cuarto muestra una visión general de los contenidos entrando en detalle en aquellos de carácter propositivo que son el propio método general para la evaluación de la exactitud posicional y el informe de evaluación; se finaliza con unas conclusiones.

### **Método para el desarrollo de la guía**

Una guía ha de ser un documento que sirva de ayuda real al caso propuesto, es decir, que ofrezca recomendaciones de valor para los usuarios, en nuestro caso para la evaluación de la exactitud posicional. Entendemos por recomendación lo que se induce a los usuarios de la guía para realizar una actividad o método de una manera determinada. Por ello, en la creación de una guía se ha de adoptar un método que sea eficaz. Existen numerosas guías para muy diversos temas (p.ej. muestreos estadísticos (EPA, 2002), atlas urbano europeo (EU, 2012), levantamientos topográficos (MSHA, 2001), etc.), pero no es así en lo relativo a cómo desarrollar el propio proceso de elaborar una guía. Basado en el ámbito sanitario (MinSalud, 2014), Ariza-López (2017) propone una metaguía, es decir, una guía de guías donde se presenta un método genérico que, parcialmente, es el que se ha aplicado en este trabajo para

realizar la propuesta que se presentará más adelante. Un aspecto relevante de este método es que toda propuesta que se realice en la guía (p.ej. tamaño de muestra, método de observación, etc.), se ha de basar en evidencias, es decir, en cualquier parte de una referencia (p.ej. artículo científico, manual, norma, etc.) que, basada en ensayos, demostraciones o cualquier otro tipo de prueba científica o experiencia, puede sustentar de manera objetiva esa recomendación.

El proceso base de desarrollo de guías es el ciclo PHCA (Planificar, Hacer, Chequear y Actuar) de Deming (Ariza-López, 2002), se aplica de manera global y también de manera iterada dentro de cada una de las siguientes etapas:

- **Planificar.** En esta fase se incluyen todas las actividades orientadas a definir el alcance del documento a desarrollar, establecer el equipo de trabajo, etc., con vistas a asegurar el éxito del trabajo. Incluye los siguientes pasos:
  1. **Priorización y focalización en una guía basada en la evidencia.** Esta decisión se toma en el seno de una organización competente tomando en consideración criterios de oportunidad, necesidad, etc.
  2. **Constitución del grupo de trabajo de la guía.** Se deben buscar y proponer técnicos competentes en todas las materias implicadas en el desarrollo de la guía, repartiendo responsabilidades según el tamaño del grupo y el funcionamiento que se haya determinado.
  3. **Identificación y análisis de conflictos de intereses.** Este paso es importante cuando se desea realizar una recomendación neutra, sin vínculo con casas comerciales o tecnologías concretas.
  4. **Definición del alcance y los objetivos de la guía.** Se debe establecer la trascendencia y los objetivos iniciales de la guía, así como algún lineamiento básico.
  5. **Formulación de los casos de uso y requisitos funcionales.**
  6. **Socialización de alcances y objetivos, casos de uso y requisitos.**

En el caso de la guía para la evaluación de la exactitud posicional de datos espaciales, la selección y priorización de esta temática se tomó por consenso por parte de los miembros Proyecto. En relación al grupo de desarrollo, dentro del equipo del Proyecto se constituyó un equipo corto formado por la ingeniera Joselyn Robledo Ceballos y el doctor Francisco Javier Ariza López, cuya misión era avanzar en los contenidos de la guía y someterlos al resto del equipo del Proyecto. No se consideró la existencia de conflicto de intereses. Como objetivo se adoptó redactar una guía aplicada que permita desarrollar evaluaciones de la exactitud posicional absoluta de una manera correcta y fiable y que, además, ofreciera suficientes apoyos teóricos y prácticos. Como lineamiento básico se adoptó el marco establecido por las normas de la familia ISO 19100 del Comité Técnico 211 aclarando, en la medida de lo posible, aquellos aspectos de este proceso que

suelen causar mayor duda. Todas estas decisiones fueron tomadas y compartidas dentro de los miembros del proyecto.

- Hacer (desarrollo). Son todas las actividades orientadas a la elaboración del material de la guía, como documento. De manera genérica se incluyen revisiones de documentación, realización de experimentos o pilotos, redacción del documento, etc., todas ellas dentro del alcance establecido. La Figura 1 presenta un esquema del proceso lógico orientado a generar recomendaciones y que abarca desde el paso 7 al 10, los cuales se describen a continuación. Finalmente, con el conjunto de evidencias se redacta la guía (paso 11). En mayor detalle los pasos son los siguientes:
    7. Búsqueda de antecedentes. Consiste en localizar normas, estándares, documentos técnicos, documentos científicos, etc., que traten los distintos contenidos a desarrollar en la guía (p.ej. tamaño de la muestra de evaluación, distribución espacial de los elementos de evaluación, test de normalidad de los errores, etc.).
    8. Evaluación general de la calidad de las referencias. Todas las referencias encontradas no tienen el mismo valor. Las referencias deben ser analizadas con una perspectiva crítica, aceptando aquellas que se consideren fundadas y de buena calidad y rechazando las demás. En esto es importante trabajar con la metacalidad de esas referencias. Algunas veces, dada la limitación de la información disponible, este análisis se limita a la reputación de la fuente.
    9. Construcción del conjunto de evidencias. Las evidencias que pasan el filtro de calidad del paso 8 se conforman como un conjunto de evidencias que podrá ser utilizado.
    10. Formulación de las recomendaciones. Consiste en la redacción sencilla de los ítems de recomendación que se incluirán en la guía.
    11. Redacción de la guía. Es la fase final de compilación ordenada y sistemática de las recomendaciones que se ofrecerán.
- Se acaban de describir los pasos de este proceso genérico, pero el desarrollo dentro del Proyecto no ha sido exactamente el arriba indicado pues se ha contado con la participación del equipo de expertos en evaluación de la exactitud posicional de la Universidad de Jaén. Este equipo posee una amplia experiencia investigadora sobre la evaluación de la exactitud posicional y en el desarrollo de estándares y guías, por lo que los pasos 7 al 10 se han podido realizar de una manera más ágil. Un elemento clave de esta propuesta es el manual *Fundamentos de evaluación de la calidad de la Información Geográfica* (Ariza-López, 2013).
- Chequear (validación). El alcance de esta fase son las actividades de validación externa del documento elaborado en la fase anterior y su aplicación por parte de los usuarios.
    12. Revisión por pares (evaluación externa de primer nivel). Antes de su publicación, todo documento se ha de revisar por pares externos e independientes, es

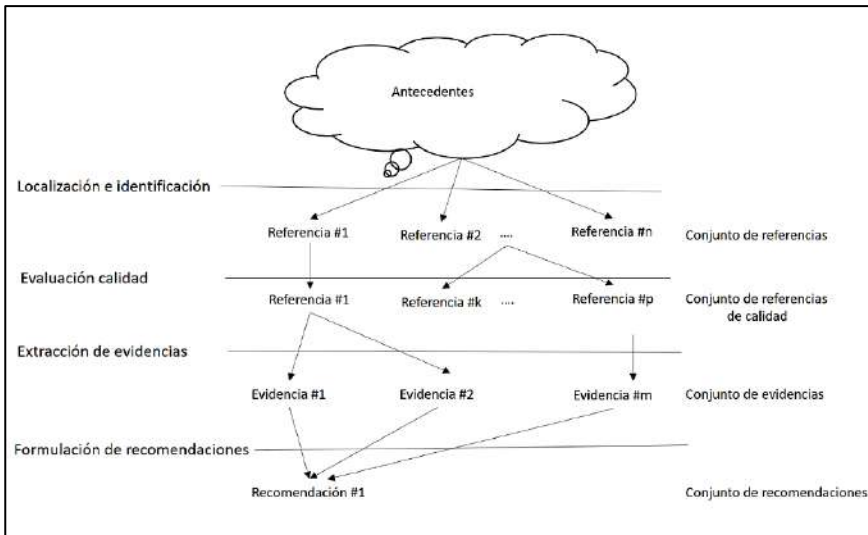
decir, por terceras personas no pertenecientes al grupo de trabajo, que posean competencia en la materia y que, igualmente, no tengan conflicto de intereses.

- Difusión y revisión pública (evaluación externa de segundo nivel). La difusión es la publicación. Con ella los usuarios tienen constancia de su existencia y conocen sus contenidos. A lo largo de todo el periodo de vigencia la guía se ha de favorecer que los usuarios puedan realimentar el proceso con sus aportaciones (p.ej. comentarios, dudas, experiencias, etc.).

13. En nuestro caso, la revisión por pares de la Guía la realizó un panel reducido de expertos internacionales. En cuanto a la publicación, se le ha asignado el número 557 de la colección de Publicaciones Ocasionales del IPGH y está disponible en [http://publicaciones.ipgh.org/publicaciones-ocasionales/Guia\\_Evaluacion\\_Exactitud\\_Posicional\\_Datos\\_Espaciales.pdf](http://publicaciones.ipgh.org/publicaciones-ocasionales/Guia_Evaluacion_Exactitud_Posicional_Datos_Espaciales.pdf). En la guía se incorpora una dirección de correo electrónico para enviar la retroalimentación por parte de los usuarios.

- Actuar (mejora). El alcance de esta fase son las actividades de revisión del documento en base a las aportaciones recibidas por parte de los usuarios o de los nuevos conocimientos y experiencias disponibles en el campo técnico y científico.

14. Mejora. Son cada uno de los ciclos de revisión que ha de sufrir el documento. En nuestro caso, esta fase todavía no se ha iniciado. Se realizará con mayor o menor rapidez en función del número y tipología de las aportaciones recibidas en el proceso de realimentación.



**Figura 1.** Proceso lógico para derivar recomendaciones basadas en evidencias.  
Fuente: Ariza-López (2017).

## Lineamientos y alcance de la propuesta

Es importante que una guía como la que se propone esté anclada en bases sólidas y aceptadas, tanto científicas como sociales o de aplicación. Por ello, a la hora de considerar su desarrollo se consideraron fundamentales los siguientes lineamientos:

- Lineamiento con las normas internacionales de la familia ISO 19100. Esto supone adoptar el modelo de la Norma Internacional ISO 19157 (ISO 2013) para describir, evaluar e informar sobre la calidad de los datos. Junto a ISO 19157 también se deben considerar otras normas como son:
  - ISO 19131 (ISO 2007). Tiene como propósito el establecer directrices para la elaboración de especificaciones de productos de datos. La evaluación de la exactitud posicional deberá ser consistente con lo indicado en las especificaciones del producto con el que se esté trabajando.
  - ISO 19115-1 (ISO 2014). Ofrece un modelo de metadatos para datos espaciales, dentro del cual cabe la información relativa a los resultados de la evaluación de la calidad. Estos metadatos pueden ser complementados con el informe de la calidad independiente que se propone en la Norma Internacional ISO 19157.

Junto a lo anterior, en la redacción y presentación de la guía también se ha adoptado el modo y estructura que se utiliza en las normas internacionales.

- Lineamiento con los métodos de evaluación de la exactitud posicional más utilizados en Hispanoamérica. Como resultado del Proyecto “Diagnóstico de la situación actual sobre las metodologías y procedimientos empleados para la evaluación de la calidad de la Información Geográfica” ([https://coello.ujaen.es/investigacion/web\\_giic/SubWeb\\_IPGH2016/](https://coello.ujaen.es/investigacion/web_giic/SubWeb_IPGH2016/)), financiado en la convocatoria de Proyectos de Asistencia Técnica del año 2015 se disponía de una publicación (Ariza-López y col., 2017) que indicaba los métodos base de evaluación de la exactitud posicional más extendidos en los Países Miembros del Instituto Panamericano de Geografía e Historia. De esta forma, se consideró adecuado abarcar los métodos NMAS (*National Map Accuracy Standard*) (USBB, 1947), EMAS (*Engineering Map Accuracy Standard*) (ASCE, 1983) y NSSDA (*National Standard for Spatial Data Accuracy*) (FGDC, 1998). Con ellos, además, se ofrecen tres perspectivas complementarias desde el punto de vista estadístico, lo cual ofrece opciones múltiples a los usuarios y enriquece la guía.

En relación al alcance de sus contenidos, se consideraron las siguientes premisas generales:

- Parte conceptual: ha de ser abarcadora. La guía debía presentar un marco conceptual esclarecedor sobre los múltiples aspectos que suelen ser problemáticos a la hora de realizar evaluaciones de la exactitud posicional.
- Parte aplicada: ha de ser de ámbito general. La guía debía ofrecer directrices generales para los procesos de diseño de la evaluación, para la determinación de los datos de referencia (p.ej. por trabajos de campo), y para los procesos de cálculo e informe de los resultados. Cada organización puede elaborar su propia guía a partir de esta propuesta general.

En relación al alcance de aplicabilidad, se consideró lo siguiente:

- Elementos de evaluación: han de ser de geometría puntual. La guía se debía limitar a la tradicional evaluación de la exactitud posicional por medio de elementos puntuales por ser la opción operativa actualmente más extendida.

Finalmente, dentro de su alcance se consideró oportuno no ofrecer deliberadamente exigencias de conformidad dado el mero carácter de orientación que se ha pretendido.

### **Contenidos de la guía**

Los contenidos que se han desarrollado dentro de la guía para la evaluación de la exactitud posicional de datos espaciales son los que se presenta en la Tabla 2 donde, además, se muestra el carácter de ese contenido. Existen contenidos de introducción al documento, de especificación del propio documento y de apoyo al documento. Las partes más relevantes son, por un lado, los contenidos que establecen el marco conceptual, donde se explica el modelo ISO 19100, algunos aspectos concretos de la exactitud posicional, las diferencias entre estimación versus control y entre métodos de base paramétrica y no paramétrica y, finalmente, los tres métodos base de evaluación de la exactitud posicional más utilizados en Hispanoamérica (NMAS, EMAS y NSSDA) y, por otro, los contenidos de carácter propositivo (general o aplicado), pues estos últimos son los que tienen una vocación de guía. También existen dos contenidos de apoyo a la aplicación de la guía: un conjunto de preguntas frecuentes y un ejemplo de aplicación detallado centrado en el cálculo y en el informe.

En los dos siguientes subapartados se presentarán los contenidos de tipo propositivo general: primeramente, el método general para la evaluación de la exactitud posicional y, posteriormente, el informe de evaluación propuesto.

### **Método general para la evaluación de exactitud posicional**

Se va a considerar que la evaluación de la exactitud posicional de un producto de datos espaciales se realiza comparando las posiciones de una selección (muestra)

**Tabla 2**



### Guía para la evaluación de la exactitud posicional de datos espaciales: contenidos y carácter

<i>Contenido</i>	<i>Carácter</i>
Presentación	Introductorio
Introducción	
Objetivo	Especificativo
Alcance	
Conformidad	
Referencias normativas	Apoyo al documento
Términos y definiciones	
Abreviaturas	
Calidad de datos espaciales basada en normas ISO del TC/211	Marco conceptual
La norma ISO 19157 aplicada a la evaluación de exactitud posicional	
Exactitud y error posicional	
Estándares de exactitud posicional	
NMAS	
EMAS	
NSSDA	
Método general para la evaluación de exactitud posicional	Propositivo general
Aspectos relativos a la ejecución	Propositivo aplicado
Datos de referencia para la evaluación	
Exactitud de los trabajos de evaluación	
Cantidad y distribución de los elementos de evaluación	
Identificación y observación de los elementos de evaluación en campo	
Identificación y observación de alturas en campo	
Identificación y observación de elementos en imágenes	
Aspectos estadísticos	
Aspectos de metacalidad de la evaluación de la exactitud posicional	
Informe de la evaluación	Propositivo general
Referencias	Apoyo al documento
Preguntas y respuestas frecuentes	Apoyo a la aplicación
Anexo I. Ejemplos de cálculo	

de objetos de tipo puntual del producto, o conjunto de datos a evaluar —CDE—, frente a las posiciones de los objetos puntuales homólogos en una referencia más

exacta que el producto, o conjunto de datos de referencia —CDR. De esta forma el proceso es el siguiente:

- Definición del proceso de evaluación. Si existen unas especificaciones de producto adecuadas, los aspectos clave de la evaluación se extraerán de ellas. En otro caso, deberán ser definidos de manera previa a la evaluación. Estos aspectos son:
  - El elemento de calidad. Podrá ser cualquiera de los establecidos por ISO 19157 (p.ej. exactitud posicional absoluta o externa).
  - El ámbito de la evaluación de la calidad. Se establecerán los aspectos temáticos, espaciales, temporales, etc., que delimiten de manera adecuada el conjunto de objetos que forman parte de la evaluación (p.ej. capa temática de construcciones).
  - Unidad de calidad de datos. Es la conjunción de los dos anteriores.
  - Nivel de conformidad. En el caso de un control de calidad, deberá existir un nivel o niveles de conformidad que se utilice para tomar la decisión de aceptación/rechazo del producto.
  - Método de evaluación. Debe ser la especificación del procedimiento de evaluación de forma integral. No basta con indicar la aplicación de un estándar de evaluación de la exactitud posicional (p.ej. NMAS, EMAS, NSSDA). El método de evaluación al que nos estamos refiriendo debe abarcar todos los aspectos relevantes de la evaluación incluyendo: exactitud necesaria en los trabajos de evaluación, tamaño y generación de la muestra de evaluación, directrices de trabajo en campo (p.ej. técnicas de levantamiento, casos especiales, etc.), procesado estadístico de los datos (p.ej. verificación de hipótesis básicas, tratamiento de valores atípicos, etc.), entre otros. Esta guía es un documento adecuado para orientar en detalle estos contenidos.
  - Medida de la calidad. La(s) medida(s) de la calidad a utilizar debe(n) quedar especificada(s) (p.ej. medida ID 45 de ISO 19157).
- Aseguramiento del proceso de evaluación. Una vez se dispone de todos los aspectos que conforman la definición del proceso de evaluación de la exactitud posicional, se deberá proceder a asegurar que realmente se aplicará de manera adecuada. Para ello se consideran diversos aspectos del aseguramiento, todos ellos interrelacionados. Éstos son:
  - Aseguramiento del personal. Consiste en disponer de personal cualificado. El personal que vaya a realizar los trabajos de campo (p.ej. observaciones GNSS) y gabinete (p.ej. análisis estadísticos), deberá disponer de la formación y capacitación adecuadas.
  - Aseguramiento del instrumental. Consiste en disponer de equipos físicos (p.ej. receptores GNSS) y de recursos computacionales (p.ej. programas de

cálculo) que sean los adecuados al método y exactitudes pretendidas. Todos los equipos y sistemas deben estar calibrados.

- Aseguramiento estadístico. Consiste en disponer de una muestra suficiente (cantidad), representativa (distribución) en la que las hipótesis estadísticas (aleatoriedad, independencia, ausencia de sesgo) y de calidad (mayor exactitud) se cumplan. En relación al tamaño de muestra a utilizar en la evaluación, se deberá tener especial cuidado en distinguir si se trata de una estimación, donde se debe aplicar la teoría de muestreos, o de un control de calidad, donde el tamaño de muestra vendrá condicionado por los riesgos de usuario y productor que se deseen asumir.
- Aseguramiento de los procesos. Todos los procedimientos estarán descritos y serán comprendidos por todo el personal involucrado con el objetivo de conseguir una evaluación fiable y económico. Dentro de este aseguramiento se incluye la planificación de los trabajos de campo (p.ej. rutas, fechas y horas de los trabajos de observación GNSS, etc.), según es usual en los proyectos de ingeniería.
- Trabajo de campo para la evaluación. Se ejecutarán los trabajos de campo consistentes en la realización de las observaciones precisas sobre el terreno de los objetos (puntos) que forman el CDR. Aspectos relevantes de este proceso son:
  - Identificación inequívoca de los puntos. No puede existir equivocación en la identificación en la referencia (p.ej. en la realidad o campo) de los puntos que forman el CDR, y además, deberán generarse evidencias de la identificación (p.ej. croquis, fotografías, etc.).
  - Observación. El método de levantamiento seleccionado (p.ej. estático rápido) deberá ejecutarse siguiendo un procedimiento estandarizado previamente por la organización, y tal que se asegure la obtención de la exactitud propia de ese método (p.ej. evitando el efecto multicamino e interferencias electromagnéticas).
- Trabajo de gabinete. Se ejecutarán los trabajos de procesado en gabinete consistentes en la obtención de las coordenadas precisas de los objetos (puntos) que forman el CDR. Aspectos relevantes de este proceso son:
  - Procesado de observaciones. Las observaciones de campo serán procesadas mediante una herramienta de cálculo validada y siguiendo un procedimiento estandarizado previamente por la organización, y tal que se asegure la obtención de exactitud propia de ese método e instrumental.
  - Identificación inequívoca en el producto a evaluar de los puntos homólogos a los CDR, que constituirán el CDE. Deberá asegurarse que no se comente equivocación.

- Verificación de prerequisites cartográficos. Se deberá verificar que las coordenadas del CDR y del CDE corresponden a un mismo sistema geodésico de referencia y proyección cartográfica.
- Verificación de hipótesis estadísticas sobre el error posicional. Se deberá verificar que se cumplen las hipótesis estadísticas que se requieran como base para la aplicación del método estadístico que se utilice para el análisis (p.ej. normalidad, homocedasticidad, atípicos, etc.). Se seguirá un procedimiento estandarizado previamente por la organización para este caso.
- Aplicación de unos de los MEEP estandarizados (p.ej. NMAS, EMAS; NSSDA).
- Análisis e informe. Se examinarán los resultados numéricos, para obtener salidas gráficas (p.ej. mapa de distribución de errores, etc.). Siguiendo un procedimiento estandarizado previamente por la organización para este caso (p.ej. una plantilla de resultados).

Estos aspectos son complementados en la guía de una manera más concreta con las sugerencias que se realizan en los apartados posteriores, dedicados a los aspectos de la ejecución, estadística y metacalidad.

En lo relativo a la ejecución, se incide especialmente a la identificación los puntos en el CDE y el CDR, incluyendo el caso de trabajo con MDE y con imágenes.

En referencia a la estadística se incide en: i) que la muestra de errores debe ser aleatoria, ii) que la normalidad de los errores no siempre se cumplirá y esto es crítico para métodos basados en esta hipótesis, iii) que los valores atípicos deben ser analizados de manera independiente, iv) que la presencia de sesgo relevante debe ser analizada, v) que la homocedasticidad debe ser verificada y vi) que los errores de las distintas componentes,  $x$ ,  $y$ ,  $z$ , deben ser independientes.

En cuanto a la metacalidad, se incide en la importancia de aplicar un método general adecuado y bien asegurado que permita ofrecer elementos de juicio relativos a la metacalidad de los resultados y sus procesos de obtención.

### ***Informe de la evaluación***

Finalmente, un aspecto relevante de toda evaluación es la manera de informar sobre los resultados obtenidos y, en esta línea, se adopta la perspectiva del informe de calidad, independientemente que se menciona en la norma internacional ISO 19157. Para ello en la Guía se propone el esquema que se presenta en la Tabla 3, que se inspira, en parte, en un ejemplo incluido en la norma UNE 148002 (2016). Se trata de un informe bastante exhaustivo y detallado que se propone vincular con aquellos aspectos que cada organización considere oportunos para obtener la conformidad respecto al método de evaluación que diseñe y aplique. A continuación se explica cada uno de los bloques de esta propuesta:

**Tabla 3**  
**Esquema de contenidos de un informe de calidad independiente**  
**para la evaluación de la exactitud posicional**

<i>Contenidos</i>	<i>Contenidos (cont.)</i>
1) <u>Identificación del producto de datos a evaluar</u> Nombre ID Productor Descripción cualitativa Propósito Especificaciones Exactitud de diseño (teórica)	5) <u>Resultados</u> Lista de errores definitiva (sin atípicos) Parámetros estadísticos básicos Diagrama circular de distribución de errores X, Y Diagrama de distribución de errores Z Histogramas de los errores Distribución espacial de los errores Distribución espacial de los atípicos Lista de coordenadas y errores Asignación del sesgo Resultados de las medidas y conformidad Resultados de los MEEP Interpretación de los resultados
2) <u>Aspectos generales de la evaluación</u> <u>Elemento de la calidad</u> Ámbito Medidas de la calidad Niveles de conformidad Método de evaluación	6) <u>Metacalidad de los resultados y procesos</u> Confianza Homogeneidad Representatividad
3) <u>Fuente de mayor exactitud y lista de coordenadas</u> Fuente de referencia Dimensión Exactitud de la referencia Recubrimiento: poblacional, temático, espacial Interoperabilidad Aseguramiento de la aleatoriedad de la muestra	7) <u>Fecha y firma del responsable</u> Fecha Firma
4) <u>Comprobación de las hipótesis estadísticas de los errores</u> Lista de errores Aleatoriedad Atípicos Normalidad Sesgos Independencia Homocedasticidad Interpretación de las comprobaciones	

- Identificación del producto de datos a evaluar. Reconoce el producto y describe los aspectos más relevantes (p.ej. contenido, propósito, especificaciones), así como las especificaciones propias de la exactitud posicional.
- Aspectos generales de la evaluación. Se establece la unidad de calidad de datos para identificar y comunicar de manera clara los elementos y ámbitos de la calidad que se evalúan. Igualmente se debe identificar la medida, niveles de conformidad (si los hubiera) y el método de evaluación utilizados de tal manera que no quede ambiguo ningún aspecto relevante (tal como se ha indicado en los distintos apartados de esta guía). Para ello, al objeto de no hacer demasiado extenso el informe, se aconseja resumir el método de evaluación e introducir referencias a documentos externos que aporten más detalles (por ejemplo, ejecución de trabajos en campo, en gabinete, etc.).
- Fuente de mayor exactitud (CDR) y lista de coordenadas. Esta parte de informe de evaluación tiene mucha importancia desde la perspectiva de la metacalidad, por ello se incluyen contenidos que han de permitir apreciar la calidad del CDR. Se trata pues de un apartado crítico dentro del informe. Se puede aportar información sobre: cuál es la fuente de referencia, si es planimétrica y/o altimétrica, su exactitud, el recubrimiento poblacional (tamaño de muestra), temático o espacial, si se ha comprobado la interoperabilidad (que el CDE y el CDR se encuentren en el mismo sistema de referencia), cómo se ha generado la muestra, etc.
- Comprobación de las hipótesis estadísticas sobre los errores. Esta parte del informe se centra en mostrar las evidencias relativas a la comprobación de que todas las hipótesis estadísticas, ya sean implícitas o explícitas, y que son requeridas por el/los MEEP aplicado/s, se satisfacen.
- Resultados. Incluye los resultados finales (cumple/no cumple o un valor de medida) que ofrece el método o métodos aplicados. También debe incluir las distribuciones de los errores y demás parámetros en las formas relevantes (espacial, histograma, etc.). Si existía sesgo y ha sido asignado se debe explicar. Finalmente, es conveniente una breve interpretación de todos los resultados de manera conjunta.
- Metacalidad de los resultados y procesos. En esta parte del informe se desarrollarán explicaciones justificativas relativas a los elementos de la metacalidad. Deberán estar basadas en hechos objetivos presentados en los apartados 2, 3, 4, 5 y 6 (Tabla 3).
- Fecha y firma del responsable. Toda evaluación debe tener asignada una fecha y un técnico responsable que ha de aprobarla.

## Conclusión

Se ha presentado la justificación de la necesidad de elevar el nivel de estandarización de los procesos de evaluación de la exactitud posicional y cómo ello ha motivado que en el seno del Proyecto de Asistencia Técnica del IPGH (PAT2018)

(Agenda del IPGH 2010-2020) “Propuesta de adopción de metodologías y procedimientos para la evaluación de la calidad de la información geográfica para los Estados Miembros del IPGH”, se haya abordado la realización de una propuesta de una guía genérica para la evaluación de la exactitud posicional de datos espaciales. La guía desarrollada no pretende ser de aplicación directa, pues cada organización tiene un saber hacer, y unas condiciones propias, que no permiten establecer un proceso común. No obstante, aun siendo genérica, esta propuesta de guía aborda de manera conjunta e interrelacionada temas que suelen generar disparidad de criterios y formas de trabajo, lo que origina variabilidad en los resultados y falta de posibilidad real de comparación. Se propone un método general para la evaluación de exactitud posicional, acompañado de aspectos más prácticos, y un entorno explicativo que permiten entender mejor qué y cómo se deben hacer las cosas en una evaluación de la exactitud posicional. Además, la guía ofrece un esquema, también general, de informe, que puede ayudar a entender mejor los procesos de evaluación de la exactitud tanto a los productores como a los usuarios y donde se han introducido contenidos de metacalidad. Para la elaboración de la guía se ha adoptado un método basado en evidencias, de tal manera que se tiene gran confianza en todo lo propuesto.

Son muchos los aspectos genéricos y detalles que podrían añadirse a la guía desarrollada, pero consideramos que es única en su concepción y que aporta valor al sector. Esperamos que tenga buena acogida y que pronto los usuarios nos remitan sus comentarios para seguir manteniéndola y mejorándola.

### ***In memoriam***

Queda en nuestra memoria y agradecimiento Edison Rojas, nuestro compañero y amigo, que solicitó este proyecto y que nos ha dejado sin ver este fruto, que también se debe a su esfuerzo y liderazgo.

### **Agradecimientos**

Al Instituto Panamericano de Geografía e Historia, especialmente a las comisiones de Cartografía y Geografía por financiar dentro del programa de Proyectos de Asistencia Técnica 2018 (Agenda del IPGH 2010-2020) el proyecto “Propuesta de adopción de metodologías y procedimientos para la evaluación de la calidad de la información geográfica para los Estados miembros del IPGH”.

### **Bibliografía**

Ariza-López, F.J., Xavier E., Chicaiza E.G., Buenaño X. (2017). “Métodos de evaluación de la calidad posicional en Hispanoamérica: análisis de la situación”, en *Revista Cartográfica*, núm. 94, pp. 65-88. Recuperado de <https://revista-sipgh.org/index.php/rcar/article/view/342>

- Ariza-López, F.J. (2002). “Calidad en la producción cartográfica”. RA-MA, Madrid.
- (2017). “Documento P1.2. Plan para el desarrollo de guías de implementación. Anexo: Guía para el desarrollo de Guías Prácticas”, Proyecto de Modernización de la Administración de Tierras en Colombia, Agencia de Implementación, Bogotá.
- (ed.) (2013). “Fundamentos de evaluación de la calidad de la Información Geográfica”, Universidad de Jaén.
- ASCE (1983). “Map Uses, scales and accuracies for engineering and associated purposes”, *American Society of Civil Engineers, Committee on Cartographic Surveying*, Surveying and Mapping Division, New York, USA.
- ASPRS (2015). “ASPRS Positional accuracy standards for digital geospatial data”, *Photogrammetric Engineering & Remote Sensing*, 81(3), 53 pp. Recuperado de [http://www.asprs.org/a/society/divisions/pad/Accuracy/Draft\\_ASPRS\\_Accuracy\\_Standards\\_for\\_Digital\\_Geospatial\\_Data\\_PE&RS.pdf](http://www.asprs.org/a/society/divisions/pad/Accuracy/Draft_ASPRS_Accuracy_Standards_for_Digital_Geospatial_Data_PE&RS.pdf)
- Congalton, R.G. and Green, K. (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd Edition, Lewis Publishers, Boca Raton.
- EPA (2002). “Guidance for Choosing a Sampling Design for Environmental DataCollection (EPA QA/G-5S)”, U.S. Environmental Protection Agency.
- EU (2012). “Mapping Guide for a European Urban Atlas”. Ref. Ares (2012)1348219 - 15/11/2012, European Commission.
- FGDC (1998). “FGDC-STD-007: Geospatial Positioning Accuracy Standards, Part 3. National Standard for Spatial Data Accuracy”, Federal Geographic Data Committee, Reston, USA. Recuperado de <https://www.fgdc.gov/standards/projects/accuracy/part3/chapter3>
- IPGH (1978). *Instituto Panamericano de Geografía e Historia: Especificaciones para mapas topográficos*, Panamá, Instituto Panamericano de Geografía e Historia.
- ISO (2014). “19115-1:2014 Geographic information - Metadata -Part 1: Fundamentals”, International Organization for Standardization.
- ISO (2007). ISO 19131:2007 Geographic information - Data product specifications. International Organization for Standardization.
- ISO (2013). ISO 19157:2013 Geographic information - Data quality. International Organization for Standardization.
- MinSalud (2014). “Guía Metodológica para la elaboración de Guías de Práctica Clínica con Evaluación Económica en el Sistema General de Seguridad Social en Salud Colombiano”, Ministerio de Salud y Protección Social, Santa Fé de Bogotá, Colombia.
- MPLMIC (1999). “Positional Accuracy Handbook: Using the National Standard for Spatial Data Accuracy to measure and report geographic data quality”, Minnesota



- Planning Land Management Information Center. Recuperado de [https://www.mngeo.state.mn.us/pdf/1999/lmic/nssda\\_o.pdf](https://www.mngeo.state.mn.us/pdf/1999/lmic/nssda_o.pdf)
- MSHA (2001). "Survey field procedures manual", The Division of Plats and Surveys, Maryland State Highway Administration.
- NN.UU. (2013). "Visión del Marco normativo de las Américas". Naciones Unidas. Comité permanente de datos geoespaciales. Recuperado de [https://uns-tats.un.org/unsd/geoinfo/RCC/docs/rcca10/E%20Conf\\_103\\_15\\_CPIDEA\\_MarcoNormativo\\_ESP\\_final.pdf](https://uns-tats.un.org/unsd/geoinfo/RCC/docs/rcca10/E%20Conf_103_15_CPIDEA_MarcoNormativo_ESP_final.pdf)
- Poggioli D. (2010). "Spatial data quality in SDI environment". LAP Lambert Academic Publishing.
- UNE (2016). "UNE 148002:2016 Metodología de evaluación de la exactitud posicional de la información geográfica", UNE, Madrid.
- USBB (1947). "United States National Map Accuracy Standards", U.S. Bureau of the Budget, Washington, USA.



# Admissible relative errors of the parcel area while creating a land parcel map base using an orthoimage as a reference

## *Determinación de errores admisibles en el área de las parcelas al ajustar un parcelario usando un ortomosaico de referencia*

Irene Balado\*  
Carlos López-Vázquez\*\*

*Received October 17, 2018; accepted February 20, 2019*

### Abstract

Currently the official land parcel map of Uruguay does not allow a suitable interoperability with other available geographic information. We are considering using a national high accuracy and high resolution orthomosaic in two ways: either to support the creation of a brand new land parcel map or to allow amending the existing one. In particular, the present work is part of a set of experiments which considers different techniques to amend the existing land parcel map. In order to perform a fair comparison it is necessary to have test cases produced with a reference procedure properly and objectively evaluated. The metric of success will be used as attainable goal for any other alternative. Using the orthomosaic in order to identify the apparent limits of parcels in the land parcel map we were able to modify its location, thus making the map interoperable with the image. In order to measure the accuracy of the fit, we used the relative difference between the area declared at the parcel map (taken as the

\* Laboratorio LatinGEO, IGM+ORT, Facultad de Ingeniería, Universidad ORT Uruguay, Uruguay, correo electrónico: [irenebalado@gmail.com](mailto:irenebalado@gmail.com). ORCID: <https://orcid.org/0000-0002-4989-4768>.

\*\* Laboratorio LatinGEO IGM+ORT, Facultad de Ingeniería, Universidad ORT Uruguay, Uruguay, correos electrónicos: [carloslopez@uni.ort.edu.uy](mailto:carloslopez@uni.ort.edu.uy); [carlos.lopez@ieee.org](mailto:carlos.lopez@ieee.org). ORCID: <http://orcid.org/0000-0002-8444-1510>.

reference value) respect to the one computed from the GIS. We found that the discrepancy between the areas computed from the initial and the final dataset w.r.t. the declared area decreases, but not up to expected level. Once filtered for outliers the remaining discrepancies follow a normal distribution, where its bias and standard deviation varies from (-0.65%, 11.24%) before editing to (1.72%, 8.59%) later. Despite the drastic change in planimetric accuracy the improvement of the area accuracy is marginal. Considering the admissibility criteria of the Uruguayan National Cadastral Office, for the legacy dataset only 39.4% parcel maps could be accepted, a value which improves up to 49.7% after the manual edition. It is still far from the goal of 100%, and the reasons for such pessimistic finding will be discussed.

Keywords: *Apparent Cadastre, Geometric Interoperability, Conflation, Land parcel map.*

## Resumen

El parcelario vectorial oficial de Uruguay actualmente no permite una adecuada interoperabilidad geométrica con el resto de los datos geográficos disponibles. Pronto existirá un ortomosaico nacional de imágenes de alta exactitud y resolución, el cual podría utilizarse de dos formas: o bien para apoyar la creación de un nuevo parcelario vectorial o bien para intentar corregir el existente. El presente trabajo es parte de experimentos que analizan diferentes técnicas de corrección geométrica semiautomática del parcelario existente. Para valorar lealmente la bondad de esos procedimientos, es necesario generar un caso de referencia con un procedimiento también de referencia y medir objetivamente la exactitud obtenida, la que luego se utilizará como meta en el proceso alternativo. Utilizando un ortomosaico controlado se ajustó el parcelario vectorial a los límites aparentes. Para medir la bondad del ajuste, se utilizó la discrepancia relativa de las áreas de la mensura y la calculada por el SIG. En este artículo se muestra que con las imágenes disponibles y los procedimientos utilizados la discrepancia entre las áreas de las parcelas derivadas del parcelario y las declaradas en las mensuras disminuye tras la edición, pero no tanto como se esperaba. Si se filtran los casos de parcelas con errores descomunales el resultado es de distribución normal, y sus parámetros (media, desviación estándar) pasan de (-0.65%, 11.24%) antes de la edición a (-1.72%, 8.59%) luego de ella. La conclusión es que la mejora obtenible tras este procedimiento de edición es marginal, al menos cuando el error es medido en relación al ajuste del área de la mensura. Con los criterios de DNC serían admisibles hoy apenas el 39.4% de las parcelas, subiendo ese valor en el nuevo parcelario a 49.7%, aún muy lejos del 100% que sería deseable. Se plantea una discusión sobre las posibles causas de este resultado pesimista.

Palabras clave: *catastro aparente, interoperabilidad geométrica, confluencia, parcelario.*

## Introduction

In Uruguay the Cadastre has covered all the country for more than a century (Guerra Daneri, 1999; Erba, 2005) with the sole exception of the public access areas which are administered separately. The cartographic base is the survey map, a document which describes the geometry of the parcel made by professional surveyors working as freelancers. The National Cadastre (Dirección Nacional de Catastro, hereinafter DNC) collects all the individual parcel maps into a single graphical document (the land parcel map) which is updated continuously (Erba, 2005) by adding any change as far as it happens.

Aside from its geometry, the attributes of the parcel includes a numerical id, owner, critical dates, etc., as well as side lengths and the area. In particular the area is a very important value, with effects on taxes. The declared area is computed by the acting surveyor using raw field data and not computed from the plan. Due to this, in Uruguay it is deemed to be a highly accurate number.

The current land parcel map of Uruguay (created in digital form from 1995 to 1999) holds nearly 250.000 rural parcels. According to Barreto *et al.* (2010) and Faure Valbi *et al.* (2010) it has a planimetric accuracy of the order of 200 m in the 95% confidence level, a value which precludes its joint use with other geographic information of similar scale. In addition to this problem (related with absolute coordinates) it is important to take into consideration the local regulations. In Uruguay the DNC does not pose any requirements on the planimetric accuracy of survey map. Instead they requires that to update any parcel map with a new one the discrepancy of the areas should be less than 5% and any parcel side length should not differ in more than 2% (Resolución 24/996) unless the surveyor provide good reasons to dismiss the old value. With the legacy land parcel map the situation today is that from its 250.000 rural parcels after comparing the computed area only 36% differs less than 5% w.r.t. its declared area. The accuracy of a land parcel map in terms of the parcel area is barely mentioned in the literature, which prefers the planimetric accuracy and do not mention neither the area nor length accuracy.

If the situation is that there is no legacy dataset to work with, the ideal procedure to have an accurate land parcel map with accurate absolute coordinates, area and side lengths would be to repeat the surveyors work in every parcel, using high accuracy technologies today available. In practice such approach is impossible, considering cost, time and even legal reasons. Once we discard the extreme solution, one possible technical alternative is to build the land parcel map from a high resolution and high accuracy orthoimage.

Under the assumption that the manual edition would be the ideal procedure for this (but not necessarily the cheapest), in this work we want to measure the accuracy that could be achieved with the process using metrics related with the parcel area and not merely the absolute coordinates. Under the assumption that a legacy land parcel

map, updated and completely operational is available, the outcome of this work can be used as a reference in order to compare with alternative procedures yet to imagine, including automatic image processing, mathematical massive transformation of the coordinates, building from scratch using individual land parcel maps, and any other alternative.

### State of the Art

Historical aerial imagery are very important in order to detect changes in ecological factors, forest, land use administration, coastal management (Ratcliffe, 2001; Nagarajan and Shenk, 2016). They are also important in order to develop a Land Administration System (Cadastré) offering a suitable base for either manual or automatic map making (Ali *et al.*, 2012).

A land parcel map is a live dataset, with daily changes that imply a costly, permanent and though task. Keeping pace with the cadastral regular updates poses a significant effort to other public services companies (electric power, waterworks, sewage, etc.) to assure interoperability with their own data, typically requesting to edit their own information in order to fit the land parcel map (Merrit and Masters, 1999).

It is possible to adjust the geometry of a land parcel map using an orthoimage either by automatic procedures or manually. For the former, as a first step it requires to properly match homologue objects that exist in both datasets. Ruiz-Lendínez *et al.* (2017) reported the process used to automatically identify control points related with crossing roads. The algorithm described improved the performance over earlier results, still with some difficulties in urban areas. Despite the detail, they do not offer any quantitative result in terms of area or positional accuracy. Trias-Sanz *et al.* (2007) went beyond that, because they presented a couple of algorithms able to, after identify borders in the image (likely related with parcel limits, roads or avenues, walls, etc.), matching such linear features with the equivalent limits existing in the land parcel map. Even though they offer some experimental results, they are not presented in terms of positional or areal accuracy but just in the matching stage. In the same topic, Kohli *et al.* (2017) recently quantified the probability of finding parcel limits as a function of the landscape type observed from high resolution images, concluding that such probability is highly dependent on the morphology but also in social issues. Thus, the possibility of building a land parcel map from an image is very dependent on the landscape.

In most of the papers on the topic the possibility to properly and accurately define the parcel limits from details in the orthoimage is taken for granted. Burns and Brown (1978) developed a statistical model based upon the binomial distribution to characterize the interpreter perception. The problem under research is the identification of lines and other geological discrete events in optic images. Their experiment required

overlapping a transparent film over an image, drawing the detected lines, and later scanning the film with a 0.5 mm resolution. Each pixel with a mark is denoted to 1 and otherwise it is set to 0. After processing at least three independent interpreters they numerically estimate the parameters of the binomial distribution. Afterwards, they use them to assign a probability to each pixel instead of mere counting black and white ones. Despite they characterized the probability that the true line indeed goes through a particular pixel, in the model they did not considered at all the positional error of the line. Unfortunately, to the author's best knowledge, this topic has been barely reported in the literature.

The method used in this paper has some links to the ones used by Ondulo and Kalande (2006) and Sengupta *et al.* (2016). They evaluated the discrepancies in terms of areas between the survey plan and the land parcel map considering the technical, legal and economical impacts. Ondulo and Kalande (2006) reported a parcel area discrepancy in some cases larger than 50% in a land parcel map build from aerial photographs from Kenya scanned, rectified, georeferenced and later projected in UTM using WGS 84. For further analysis they filtered out those cases which differ more than 20% from the recorded area value. They concluded that to lower such differences the planimetric accuracy of the land parcel map should not be larger than 2 m. To achieve that, they suggest using panchromatic images of 1 m resolution or multispectral images of 4 m resolution. They also reported that a clear direct relation exist between the relative error of area and the area itself.

Sengupta *et al.* (2016) also edited the position of the corners of the parcels in a legacy land parcel map using orthoimages as the reference. The corners were modified through an affine transformation taking into consideration the location of 10 to 15 control points found also in the image. In this case, the reference area of the survey plan was not produced by the surveyor himself using the field data book, but from the already built land parcel map, so its own accuracy is limited. The relative errors after the geometric transformation were almost always below 2%, but the errors before the transformation were not reported.

Srinivas *et al.* (2012) build a brand new land parcel map using an orthoimage at scale 1:10.000 and they reported the resulting accuracy in terms of length. The relative error of the image was of the order of 0.07%. The legacy and the new land parcel map have discrepancies in length of the order of 5%, but unfortunately they did not mention the area.

In this paper we took into consideration the literature in order to assess the relative error of the legacy land parcel map. We requested the opinion of some surveyors about the expected discrepancy between the geometric area of a survey plan and the numeric value stated there. In relatively modern survey maps, the computation is performed by the CAD so the error is negligible. In older works, the stated area was

calculated directly from the field data book and afterwards the map was drawn independently from the computations. The experts stated that relative error of the area might be not more than a few per cents. According to the present rules, in Uruguay a new survey plan of an already measured parcel might differ in area up to 5%, and in parcel side lengths up to 2% in order to be inserted in the system. Otherwise a separate explanation of the reasons of the discrepancy should be offered to the administration (Resolución 24/996).

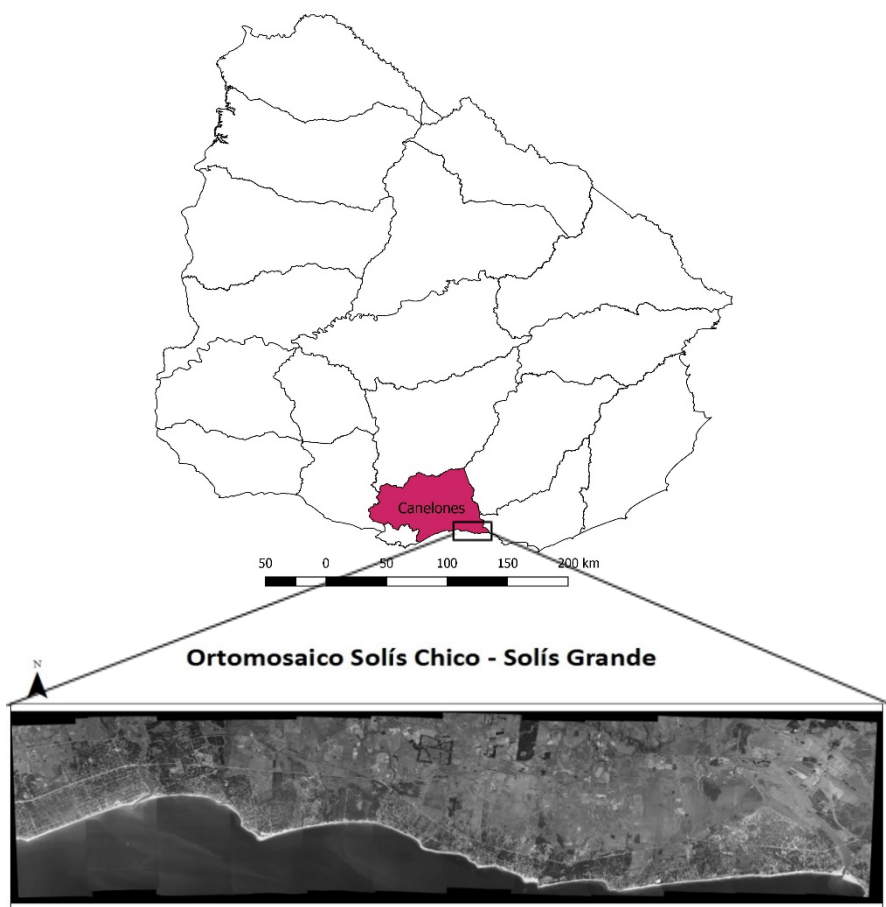
The initial goal of this project was to measure the accuracy of an improved land parcel map using the relative area error. We evaluated with such metric both a legacy parcel map as is and its modified version using an orthoimage as the reference.

### Data and methods

The reference base used to edit the land parcel map both at suburban and rural areas was an orthoimage scale 1:25.000. The flight by the Uruguayan Air Force took place in 1987 over the Canelones county along lines parallel to the coastline. It covered a geological structure associated to a paleorelief with heights not exceeding 20 m, leading to a smooth and undulated geomorphology (Goso *et al.*, 2017). Its pre-Holocen and Holocen sediments gave pace to valleys oriented NNW-SSE of fluvial cycles of the rivers Solís Chico and Solís Grande, used as East and West limits for the orthoimage (see Figure 1). The scanning process was performed to 15 micres in the Military Geographic Service (SGM) and the aerotriangulation was generated with the Trimble Inpho Match-AT (Aerial Frame Triangulation version 7.1.3) specifying a pixel size of 50 cm. After that a Digital Terrain Model (comprising a Digital Elevation Model and a Digital Surface Model) was built using 26 control points. Then, a second field work campaign was performed gathering 24 new independent control points, and with them the NSSDA (FGDC, 1998) planimetric accuracy was estimated as 3.2 m at the 95% confidence level. The parameters (accuracy and resolution) of this orthoimage are comparable to those of an ongoing national flight, which by contract requested images of 0.50 m resolution and 3.0 planimetric accuracy.

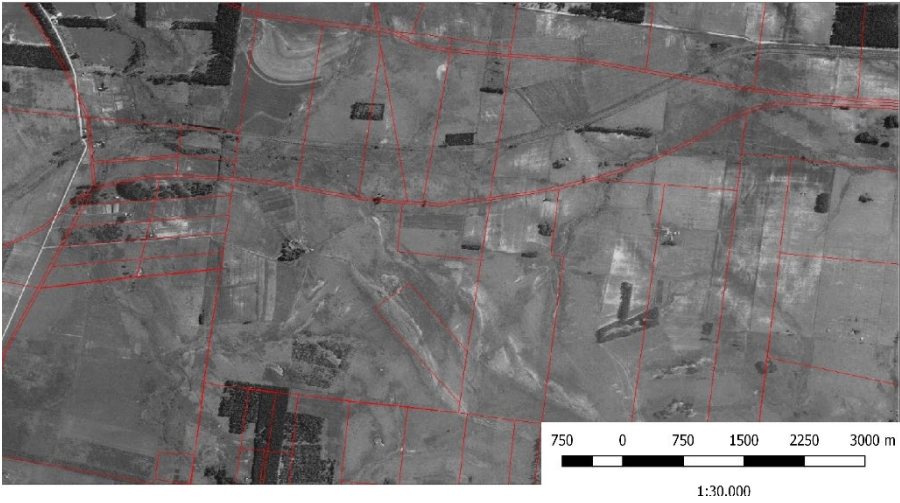
The legacy land parcel map for the rural area has known geometric problems. Its accuracy was computed for some areas giving values of the order of 200 m (Faure Valbi *et al.*, 2010; Barreto *et al.*, 2010), 276 m or even 136 m (López-Vázquez, 2014). Comprising nearly 250.000 parcels, only 36% of the survey plans has a declared area which differ less than 5% with the existing digital land parcel map. On the other side, nearly 15% has discrepancies larger than 50%, differences that can not be attributed only to graphical problems. From the national land parcel map only those parcels completely included in the image were considered further. For each parcel the polygon's area were evaluated before edition (Figure 2) and after edition (Figure 3), both using ArcGIS 10.0.



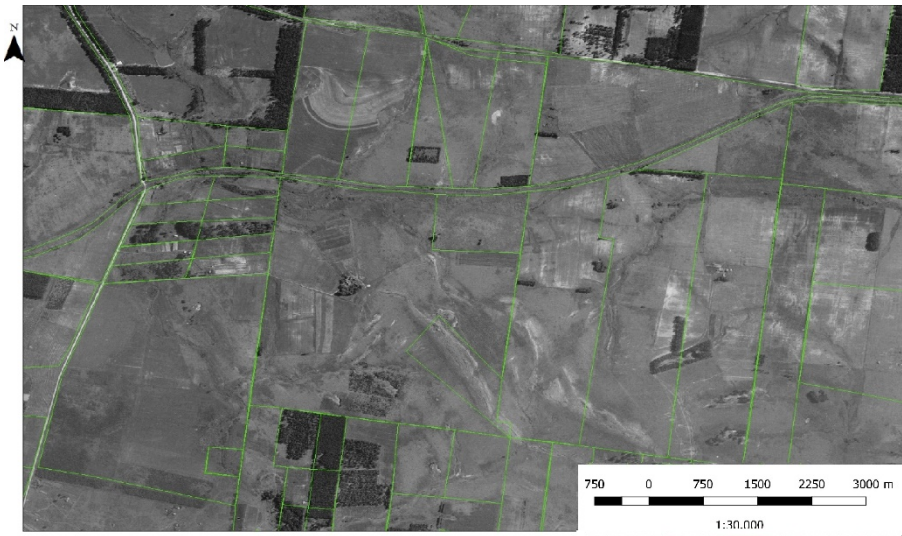


**Figura 1.** Location of the study area. The Canelones county is denoted in colour. The West and East limits of the orthoimage are the rivers Solís Chico and Solís Grande. Source: Own elaboration.

The edition to match the image was a manual process. After an interpretation of the parcel limits in the image we modified the location of the parcel corners. Thus, and without independent information, we should accept that the resulting planimetric accuracy is of the order of the one estimated for the orthoimage.



**Figura 2.** Illustration of part of the land parcel map before edition, with the image as background. Notice that there are displacements which easily exceed 200m. Source: Own elaboration.



**Figura 3.** Illustration of the final land parcel map, after manual edition to fit the image. Source: Own elaboration.

## Methodological Issues

Both systematic and random errors are unavoidable while digitizing land parcel maps (Tong *et al.*, 2005). While analyzing pre and post edition relative errors we found some large outliers. Such disagreement might be related to an identification problem of the proper survey plan, which despite having the same parcel number it has suffered some fusion/division later which was not adequately recorded in the land parcel map, or a yet unexplained error at the moment of digitizing. In any case, and for this work, such cases were removed from further consideration.

During the edition stage we found some difficulties associated to the visual interpretation of parcel limits (typically denoted by wires). In some cases we found wires, but they were not related with the parcel limit but with some internal ones, delimiting cattle areas, agricultural uses, forest, etc. In others, the rivers either removed the wires in a flood or we suspect that the wires have never existed, because it is very common that rivers that are also natural boundaries for the parcels are not limited by wires or walls. It is also possible that, after a parcel division, there is no effort to denote such new situation in the field. Finally, we have to mention the trivial case of some parcel limits that simply cannot be discerned in the image, due to the lack of contrast between textures, similar color or by rocky bed that precludes easy identification. In all such cases such parcels were removed from further analysis, because to find a solution might require a specialized effort not considered in the research.

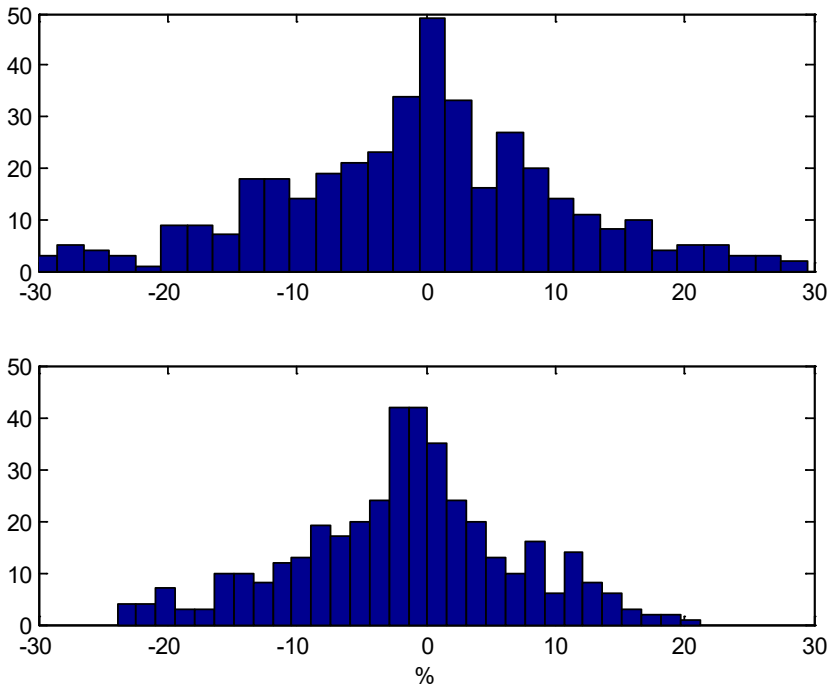
## Results

The results have shown that, given the available images and the procedures followed, we were able to diminish the overall discrepancy between the areas computed from the land parcel map geometry and those declared by the surveyors in the survey plan. For the experiment we have considered nearly 9600 ha, being 80% of them rural parcels. Individual parcel areas range from 293 to 0.018 ha in rural areas and from 13.5 to 0.017 ha in urban areas.

The universe of relative errors of the area is a mix between a normal distribution and other with more extreme values. If we remove all those cases with an absolute relative error larger than 25% (a discrepancy which safely we can attribute to non-geometric problems) both the initial and final relative error verify the Kolmogorov test of normality with a 95% confidence level. The initial universe comprises 477 parcels; after the filtering only 399 remain of the initial dataset and 415 of the final one after edition (Figure 4). The parameters of the normal population for the initial land parcel map were -0.65% as bias and 11.24% for the standard deviation, while for the edited version they were -1.72% and 8.59% respectively. The associated histograms are presented in Figure 5.



**Figure 4.** Location of those parcels with extreme discrepancies in area relative error. In green we show those parcels which relative area error belongs to the interval  $[-25\%,+25\%]$  and in red those with larger values. Non coloured cases denotes parcels without enough information to perform the computations.  
Source: Own elaboration.



**Figure 5.** Histogram of the relative area errors before (top) and after (bottom) the edition (in %).  
Source: Own elaboration.

## Discussion

Before starting the research, we anticipated small differences between the legacy land parcel map parcel areas and those values declared in the survey plan, and that such differences could be diminished drastically after the edition. Both proved to be false. Some large discrepancies, exceeding 25%, forced us to remove some parcels from further analysis because it was assumed that the differences are not due to geometrical problems. Once the outliers were removed, we confirmed through the Kolmogorov test at the 95% confidence level that the remaining relative errors belong to a normal distribution. We estimated the standard deviation before and after the edition, rendering values of 11.24% and 8.59% respectively. With such results at hand, we confirmed that the relative error in area diminishes after the edition, but not in the expected amount.

In Uruguay, the regulations state that a new survey plan for a already existing parcel might differ from early values in no more than 5% in area, and up to 2% in lengths in order to be accepted by the DNC without further explanation (Resolución 24/996). That rules offered a second metric to assess the improvements after the edition.

The planimetric accuracy of the image (according to FGDC, 1998) is of 3.2 m, slightly larger than the recommended value of Ondulo and Kalande (2006) which suggest 2 m. The pixel size of the image was 0.50 m, while Sengupta *et al.* (2016) used a similar orthoimage with pixel size 0.45 m. We might suspect that both the orthoimage accuracy and resolution might explain the remaining relative error of the parcel areas, but there is not enough information in the literature to check. Notice that the planimetric accuracy of the land parcel map was drastically improved, decreasing from something of the order of 200 m down to 3.2 m, but the accuracy in terms of area did not improve as much.

The Apparent Cadastre, edited after interpretation of the orthoimage only has planar coordinates X and Y, and thus the height was dismissed. According to Rodríguez (2000) this might explain having a difference with the area from the image (which neglects the height) with the area directly computed from the field data book, which considers the terrain height.

## Conclusions

Most of the literature related with cadastral accuracy improvement put focus on the planimetric error. In this work we attempted to quantify the impact of a particular methodology, namely manual edition with interpretation of a background reference image, in terms of the relative error of the parcel area. After filtering out outliers, we confirmed that the initial relative error follows a normal distribution, with standard deviation 11.24%. Taking into consideration the uruguayan regulations, if the land

parcel map was used as the reference, only 39.4% of the survey plans will be admissible by the DNC. Despite coming from a small sample, this value is representative of the situation for the whole country, which accounts for 36%. The workaround of using the apparent parcel limits inferred from the image was not as efficient as expected, just lowering the standard deviation of the resulting population from 11.24% to 8.59%. According to DNC rules, the rate of admissible parcels against the edited land parcel map increased in turn to 49.7%. The planimetric accuracy was not measured independently of the image, but was deemed to be drastically better than the legacy values (3.2 m vs. 200 m).

There are a number of parcels which differ too much from their geometric counterpart, a difference that was not solved after the edition. Thus, in order to perform a Cadastral Upgrade Project, a first step must identify and solve through a careful analysis by surveyors of the available documents before going into the edition.

Finally, it should be mentioned that since the Decree 318/995 the surveyor must include in the survey plan a table with coordinates in a local, orthogonal system of every corner of the parcel. That is good, but might not be enough. One possible improvement for the future could be that, aside of the polyester analog map already requested, the surveyor should provide a digital file with the parcel corners in absolute coordinates. That way it will be possible (and hopefully easier) to check parcel side lengths against the neighbors, as well as to check the area against earlier information, all during the evaluation process of the survey map.

## Acknowledgements

This work could not have been done without the support of the Director and personnel of the Geodesy and Topography Division and the Restitution and Photogrammetry Division of the Military Geographic Service of Uruguay. We want to mention also the support of Mr. Ing. Agrim. Hebenor Bermúdez, who provided references and valuable advice for the task.

## References

- Ali, Z.; Tuladharb, A. and Zevenbergen, J. (2012). "An integrated approach for updating cadastral maps in Pakistan using satellite remote sensing data", *International Journal of Applied Earth Observation and Geoinformation*, vol. 18, 386-398. <https://doi.org/10.1016/j.jag.2012.03.008>
- Barreto, L.; Bermúdez, H.; Blanco, D.; Di Leoni, A.; Faure, J.; Méndez Baillo, R.; Pérez Rodino, R.; Sánchez, M.; Striewe Dieste, E. (2010). "Evaluación de la Calidad de la Exactitud Geométrica Absoluta del Parcelario Rural Digital Vectorial del Departamento de Lavalleja", *Agrimensura*, vol. 44, 2-7.

- Burns, K.L. and Brown, G.H. (1978). "The Human Perception of Geological Lineaments and other Discrete Features in Remote Sensing Imagery: Signal Strengths, Noise Levels and Quality", *Remote Sensing of Environment*, vol. 7, 163-176. [https://doi.org/10.1016/0034-4257\(78\)90045-7](https://doi.org/10.1016/0034-4257(78)90045-7).
- Decreto 318/995 (1995). "Registración de Firma y Título Profesional en la Dirección General del Catastro Nacional. Planos de Mensura", Uruguay. Recuperado de <http://cort.as/-1694>.
- Erba, D.A. (2005). "Historia del Catastro Territorial en Latinoamérica: los países del Conosur", *CT Catastro*, no. 53, 41-57.
- Faure Valbi, J.; Blanco, D.; Bermúdez, H.; Méndez-Baíllo, R.; Barreto, L. y Striewe, E. (2010). "Evaluación de la calidad de la exactitud geométrica absoluta del parcelario rural digital vectorial del Departamento de Lavalleja", Proyecto de Instituto de Agrimensura, Facultad de Ingeniería, Universidad de la República (unpublished), 37 pp.
- FGDC (1998). "Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy", *Federal Geographic Data Committee*, FGDC-STD-007.3-1998, 28 pp. <https://www.fgdc.gov/standards/projects/FGDC-standards-projects/accuracy/part3/chapter3>.
- Goso, C.; Faure, J.; Pratto, D.; Barreto, L.; Picchi, D.; Scaglia, F.; Paris, A.; Samaniego, L.; Ubilla, D. y García, G. (2017). "Vulnerabilidades Geoambientales de la costa del Departamento de Canelones", *Nuevas miradas a la problemática de los ambientes costeros. Sur de Brasil, Uruguay y Argentina*, Goso, César (comp. y rev.), DIRAC-Facultad de Ciencias, Cap. 8, 172-196.
- Guerra Daneri, E. (1999). "La estructura Catastral del establecimiento rural", *Revista de la Facultad de Derecho* (2a época), no. 16, 177-195. Recuperado de <https://revista.fder.edu.uy/index.php/rfd/article/view/312>.
- Kohli, D.; Bennett, R.; Lemmen, C.; Asiama, K. and Zevenbergen, J. (2017). "A Quantitative Comparison of Completely Visible Cadastral Parcels Using Satellite Images: A Step towards Automation", FIG Working Week 2017 Surveying the world of tomorrow - From digitalisation to augmented reality Helsinki, Finland, May 29-June 2, 14 pp.
- López-Vázquez, C. (2014). "Mejora de Exactitud Posicional: ejemplo para el caso del CONEAT", II Jornadas de Cartografía, 30-31 octubre, Montevideo, Uruguay, 10 pp.
- Merrit, R. and Masters, E. (1999). "Digital cadastral upgrades-A Progress report", *Proceedings of the First International Symposium on Spatial Data Quality*, Hong Kong, 180-188.

- Nagarajan, S. and Schenk, T. (2016), "Feature-Based registration of historical aerial images by Area minimization", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, 15-23.
- Nagarajan, S. and Schenk, T. (2016), "Feature-Based registration of historical aerial images by Area minimization", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, 15-23. <https://doi.org/10.1016/j.isprsjprs.2016.02.012>.
- Ondulo J.D. and Kalande W. (2006). "High Spatial Resolution Satellite Imagery for PID Improvement in Kenya" Shaping the Change XXIII FIG Congress, Munich, Germany, October 8-13, 9 pp.
- Ratcliffe, J.H. (2001). "On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units", *Int. J. Geographical Information Science*, vol. 15, no. 5, 473-485. <https://doi.org/10.1080/13658810110047221>.
- Resolución 24/996 (1996). "Uniformización de normas de planos de mensura, copias, formatos, escalas", Dirección Nacional de Catastro, Uruguay. Recuperado de <http://cort.as/-1iPz>.
- Rodríguez, I. (2000). "La imagen digital en el Catastro, aplicaciones en la realización de la cartografía catastral", *Catastro CT*, no. 38, 61-76.
- Ruiz-Lendínez, J.J.; Maćkiewicz, B.; Motek, P. and Stryjakiewicz, T. (2017). "Method for an automatic alignment of imagery and vector data applied to cadastral information in Poland", *Survey Review*, 12 pp. <https://doi.org/10.1080/00396265.2017.1388959>.
- Sengupta, A.; Lemmen, C.; Devos, W.; Bandyopadhyay, D.; Van der Veen, A. (2016). "Constructing a seamless digital cadastral database using colonial cadastral maps and VHR imagery: an Indian perspective", *Survey Review*, vol. 48, No. 349, 258-268. <https://doi.org/10.1179/1752270615Y.0000000003>.
- Srinivas, P.; Raghu Venkataraman, V. and Jayalakshmi, I. (2012). "Digital Aerial Orthobase for Cadastral Mapping", *J Indian Soc Remote Sens*, vol. 40, no. 3, 497-506. <https://doi.org/10.1007/s12524-011-0183-2>.
- Tong, X.; Shi, W.; Liu, D. (2005). "A Least Squares-Based Method for Adjusting the Boundaries of Area Objects", *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 2, 189-195.
- Tong, X.; Shi, W.; Liu, D. (2005). "A Least Squares-Based Method for Adjusting the Boundaries of Area Objects", *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 2, 189-195. <https://doi.org/10.14358/PERS.71.2.189>.
- Trias-Sanz, R.; Pierrot-Deseilligny, M.; Louchet, J. and Stamon, G. (2007), "Methods for Fine Registration of Cadastre Graphs to images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, vol. 29, no. 11, 1990-1201 <https://doi.org/10.1109/TPAMI.2007.1108>.



# Atlas escolares para la educación geográfica de niños y jóvenes

## *School atlases for the geographical education of children and young people*

José Jesús Reyes Nunez\*

*Recibido 10 de enero de 2019; aceptado 19 de febrero de 2020*

### **Resumen**

El objetivo de este artículo es presentar de manera concisa el surgimiento y desarrollo de los atlas escolares desde el siglo XVIII hasta nuestros días. Una vez definido su significado, se presentan brevemente las tareas que se realizan durante la edición de atlas de este tipo, así como algunos ejemplos. En esencia, las etapas de la creación de un atlas escolar no han variado mucho, sin embargo los soportes en que se publican, sí. Somos testigos de la evolución de los soportes: de las ediciones impresas en papel primero, después los soportes electrónicos (CD), y actualmente la web. Se ofrecen también ideas sobre las posibilidades que pueden brindar las aplicaciones hechas para dispositivos móviles y qué nuevas investigaciones se realizan para diversificar y optimizar el uso de los atlas escolares en el futuro.

Palabras clave: *atlas escolar, cartografía escolar, web cartografía, tecnología móvil.*

### **Abstract**

The main aim of this article is to present concisely the beginning and development of school atlases from the 18th century to the present day. After defining the school atlases, those tasks that are faced during the edition of atlases of this type are briefly presented, as well as some examples. In essence, the stages of creating a school atlas

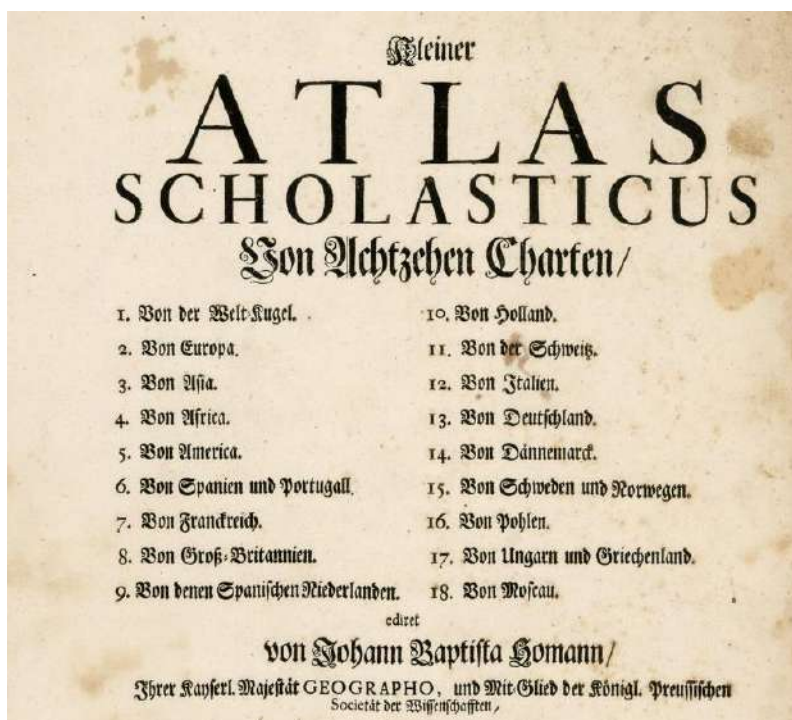
\* Universidad Eötvös Loránd, Hungría, correo electrónico: [jesusreyes@caesar.elte.hu](mailto:jesusreyes@caesar.elte.hu).  
ORCID: <http://orcid.org/0000-0001-6080-8667>.

have not changed markedly, changing only the medium in which they are published. We can witness how they developed from paper-printed editions first to electronic media (CD) and currently to the web. Ideas are also given on the options that applications made for mobile devices can offer and what new research is done to diversify and optimize the use of school atlases in the future.

Key words: *school atlas, school cartography, web cartography, mobile technology.*

### Introducción histórica

Desde hace siglos se han usado mapas en la enseñanza, para complementar e ilustrar los conocimientos impartidos en asignaturas relacionadas con la Geografía, pero también en otras que precisan de la información espacial en su contenido, como por ejemplo tradicionalmente ha sido la Historia. Dos de los primeros atlas expresamente publicados para ser usados por alumnos son obra del cartógrafo alemán Johann

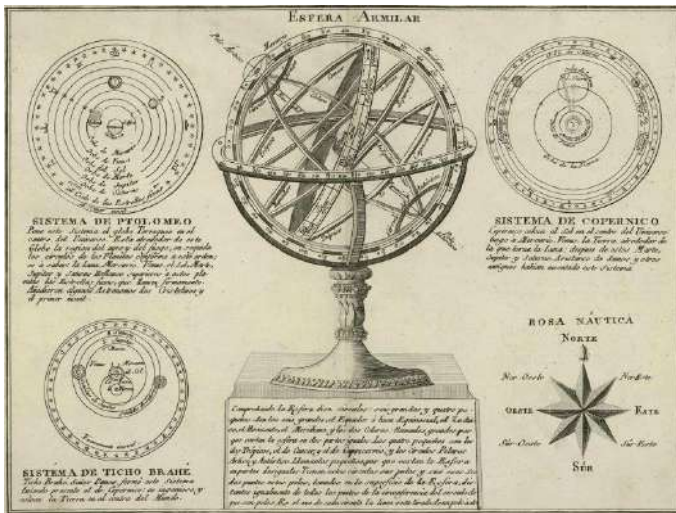


**Figura 1.** Portada del *Kleiner Atlas Scholasticus* de J. B. Homann con una lista de los mapas incluidos. Fuente: <https://www.raremaps.com/gallery/detail/22948/title-page-kleiner-atlas-scholasticus-homann-heirs>

Baptist Homann. El primero vió luz pública en 1710 bajo el título de *Kleiner Atlas Scholasticus*, contando con 26 mapas a color (Figura 1).

El segundo atlas fue publicado en Nurenberg en 1719 y su título en latín es *Atlas Methodicus* con el subtítulo “Explorandis Juvenum Profectibus in Studio Geographico...”. En él se incluyeron 14 mapas para representar los países de Europa, Asia, África y América, también presenta cuatro representaciones del sistema solar según las ideas de Ptolomeo, Tycho Brahe, Copérnico y Descartes. Como detalle interesante, California aparecía dibujada como una isla en el mapa de América, y aparece también en la lista de nombres geográficos entre los nombres de las islas.

Durante el siglo XVIII se continuaron publicando atlas para las escuelas. En 1774, en Lyon (Francia), Jean-Marie Bruyset publicó su *Atlas des enfans, ou nouvelle méthode pour apprendre la géographie...*, con 24 mapas caracterizados por tener un contenido simplificado, que incluso fue traducido al inglés. A finales de ese mismo siglo, en 1792 sería publicado el primer atlas escolar en lengua española en la ciudad de Madrid, España. Su autor fue don Tomás López y se titulaba *Atlas Elemental Moderno o colección de mapas para enseñar a los niños Geografía con una idea de la Esfera*. El atlas, que contaba con un texto introductorio de 24 páginas y 27 mapas, se caracterizó por el grado de detalle de los mapas, que enriquecía el ejemplo del *Atlas Metódico de Homann*, incluyendo una introducción que ofrecía conocimientos de geografía astronómica y seguidamente dos páginas que estaban dedicadas a representaciones gráficas relacionadas con este tema (Figura 2).



**Figura 2.** Una de las páginas ilustrando el texto de la introducción del *Atlas Elemental Moderno* de don Tomás López.

Fuente: <http://bdh-rd.bne.es/viewer.vm?id=1110&page=19>

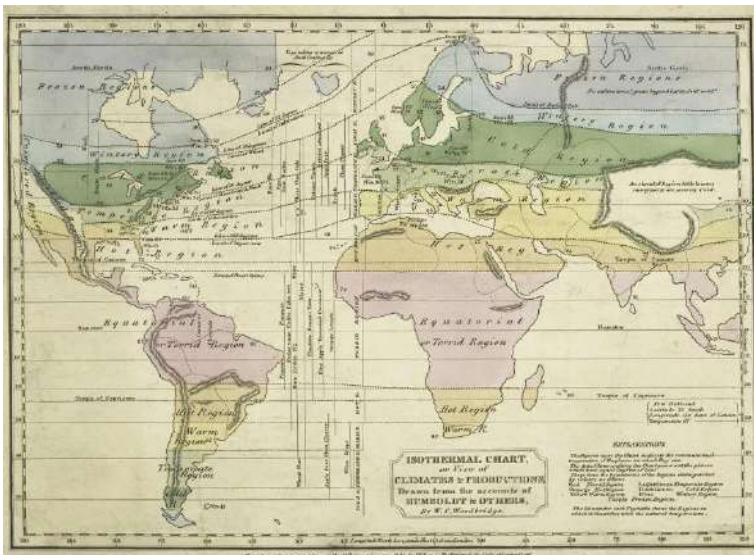
En 1795 es publicado el primer atlas escolar en los Estados Unidos: *The General Atlas for Carey's Edition of Guthrie's Geography Improved*, aunque según Patton (1999) los mapas de este atlas ya habían sido publicados anteriormente por Mathew Carey en atlas dedicados a los adultos.

El siglo XIX constituyó un período de vital importancia en el desarrollo de lo que posteriormente llamaríamos cartografía, y en especial de la publicación de los atlas escolares, fue cuando el uso de estos atlas se extendió con mayor fuerza en muchos países. Como ejemplo cito solamente dos países ubicados en dos continentes distintos: entre los países del continente europeo selecciono a Francia, que se destacó por la variedad y profesionalidad de los mapas temáticos que se incluían en los libros de texto. El autor más reconocido internacionalmente fue el profesor, geógrafo, economista e historiador Pierre Émile Levasseur (1828-1911), quien posteriormente fuera nombrado presidente honorario de la Sociedad Geográfica Francesa. Levasseur no se limitó a ilustrar y complementar los textos de sus libros con mapas temáticos, sino también creó atlas escolares que se caracterizaron porque en su introducción contenía una presentación gráfica y textual muy detallada de los conceptos elementales relacionados con la astronomía, los mapas y la geografía.

Otro país con un desarrollo muy diversificado de los atlas escolares fue Estados Unidos. Durante el siglo XIX se fundaron casas editoriales que se dedicaran a publicar este tipo de materiales, y nombres como Sidney E. Morse, Nathaniel G. Huntington, A. N. (Jessie) Olney y William C. Woodbridge, entre otros, se identificaron con los atlas escolares por ellos publicados. Mencionaría tres detalles interesantes, que corroboran el auge que tuvieron los atlas escolares estadounidenses en este siglo: después de su viaje por América del Sur y Central, Alejandro von Humboldt hizo las mediciones necesarias para crear el primer mapa del mundo de isotermas, que presentó la variación de la temperatura en la Tierra. Este mapa fue dibujado por primera vez en 1823 por William C. Woodbridge (Figura 3) y sería publicado en su atlas escolar de 1826 (*School Atlas designed to accompany Woodbridge's Rudiments of Geography*), que antecediéndose a su publicación en Europa, que ocurre años más tarde, con el famoso *Atlas Berghaus* de 1838. Esta obra de Woodbridge se considera un hito en la historia de la cartografía escolar estadounidense, ya que por primera vez incluía tres mapas temáticos: el ya mencionado de isotermas, otro de animales del mundo y uno de moral y política.

En los Estados Unidos también se publicó por primera vez un atlas escolar hecho por una mujer, que al mismo tiempo fue el primer atlas histórico escolar publicado en ese país: el *Ancient Atlas* de Emma Willard, quien era una maestra devenida la primera mujer cartógrafa estadounidense y que fue publicado en New York en 1828. Por último, mencionaría el atlas escolar publicado por Jesse Olney, cuyo título original en inglés es *School Atlas to Accompany the Practical System of Geography for Schools and Families* y que tuvo un total de 91 ediciones en un período de 21 años,

entre 1829 y 1860 (Patton, 1999). Todo este desarrollo impetuoso propio del siglo XIX se hace mucho más interesante cuando leemos lo que el cartógrafo Erwin Raisz escribió en su subcapítulo dedicado a los atlas escolares en su célebre libro de texto *Cartografía General* en 1948: “Los atlas escolares no son comunes en los Estados Unidos, ya que los libros de texto de Geografía contienen todos los mapas necesarios” (Raisz, 1948: 218). Catorce años más tarde confirmaría esta afirmación en su libro *Principios de Cartografía* (Raisz, 1962: 102).

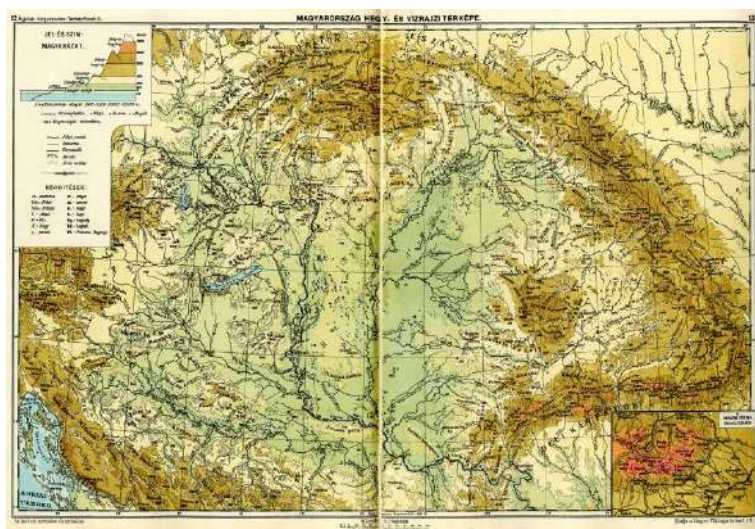


**Figura 3.** “Mapa del mundo de isoterma”, realizado por William C. Woodbridge en 1823.  
Fuente: [https://commons.wikimedia.org/wiki/File:Woodbridge\\_isothermal\\_chart.jpg](https://commons.wikimedia.org/wiki/File:Woodbridge_isothermal_chart.jpg)

El siglo XIX también trajo consigo la fundación y desarrollo de firmas cartográficas o casas editoriales especializadas en temas geográficos, cuya actividad se extendería también al siglo XX. Entre ellos tenemos, por ejemplo, el *Atlas General Vidal-Lablache: histoire et géographie*, publicado por primera vez en 1894 en París, Francia, y que tendría nuevas ediciones en el siglo siguiente, o los atlas escolares publicados en Alemania y Austria por firmas como Gotha & Justus Perthes (que publicó por primera vez en 1888 el *Methodischer Schulatlas* de Sydow y Wagner y cuya última edición fue la número 23 en 1944) o la editorial Hölzel que desde 1861 hasta el día de hoy continúa publicando su *Kozenn-Atlas*.



Otras naciones más pequeñas también trataron y tratan de mantener la publicación de atlas escolares para sus escuelas. Como ejemplo característico podría citar a Hungría, cuyo primer atlas escolar fue publicado en el año 1800 en la ciudad de Debrecen, a unos 230 km de la actual Budapest. Su título fue *Nuevo Atlas Escolar para alumnos de primaria* y su autor el profesor Ézsaiás Budai. El atlas contenía solo 12 mapas y no incluía un mapa de Hungría: el país estaba representado solamente en un mapa regional. Durante los siguientes 90 años los atlas escolares húngaros eran traducciones de atlas alemanes o austríacos. En 1890 un conocido cartógrafo, Manó Kogutowicz, funda el Instituto Geográfico de Hungría que, a pesar de su nombre, era una compañía privada que se dedicaría principalmente a la publicación de atlas, mapas murales y esferas para las escuelas. Hasta la Primera Guerra Mundial, el Instituto publicó siete *atlases* escolares geográficos y tres atlas escolares históricos, que tuvieron numerosas ediciones, aparte de 37 mapas murales físicos, 24 mapas murales de provincias húngaras, 13 mapamundis murales históricos y 10 mapas murales históricos de Hungría, sin contar la producción de mapas ciegos para tareas escolares y globos terrestres para las aulas (Klinghammer, 2001). Su labor no fue solo reconocida en el país, sino también a nivel internacional, ganando medallas en exposiciones y congresos internacionales. Podemos afirmar que Kogutowicz y su instituto sentaron las bases del desarrollo de la cartografía escolar moderna en Hungría (Figura 4).



**Figura 4.** Versión publicada en 1906 del primer mapa físico del valle de las Cárpatos hecho con curvas de nivel en el Instituto Geográfico de Hungría.

## Definiendo los atlas escolares

En la cartografía podemos encontrar que la definición de atlas ha ido modificándose con el paso del tiempo. Citaría tan solo un ejemplo válido para los atlas tradicionales impresos: en 1948 el cartógrafo norteamericano de origen húngaro Erwin Raisz en su libro *Cartografía General* definió un atlas como una colección de mapas, generalmente en un solo volumen, publicado en un mismo idioma, con simbolización y proyecciones uniformes, pero no necesariamente con la misma escala (p. 216). En 1962, Raisz ofreció una definición muy similar, con la única diferencia de que ya no se limita a hablar de un atlas como una publicación en un solo volumen, sino en varios volúmenes o con hojas sueltas que permiten una actualización más rápida.

Con la llegada de la era digital a la cartografía, se consideró que la definición de atlas también debía ser adaptada a esa nueva tecnología, utilizándose un estilo de redacción que puede ser aplicado tanto a los atlas impresos como a los electrónicos o digitales. En 1993, Olev Koop define un atlas como “una colección sistemática y coherente de datos geográficos, en formato análogo o digital, representando un área particular y/o uno o más fenómenos geográficos, basada en una narrativa junto con herramientas para la navegación, búsqueda de información, análisis y representación” (Koop, 1993).

Si visitamos la página web de la Comisión de Atlas de la Asociación Cartográfica Internacional (ACI, ICA por sus siglas en inglés) encontramos varias definiciones, entre ellas la que fue redactada por el profesor checo Vít Voženilek (2014): “la cartografía actual considera un atlas como una serie de mapas compilados con un objetivo, que han sido sistemáticamente organizados de acuerdo con el contenido temático, la extensión espacial y el punto de vista temporal, y ensamblados usando un lenguaje cartográfico común”.

Estas definiciones también son aplicables a los atlas escolares, aunque estos tienen sus características muy propias, por ejemplo su publicación se realiza casi siempre en un solo volumen. Una definición de atlas escolares podemos encontrarla en el libro electrónico *The World of Maps*, publicado por la ACI en su página web: “los atlas escolares son aquellos que introducen a los estudiantes en los aspectos físicos y socioeconómicos de la geografía mundial” (Ormeling, 2014). Erwin Raisz no define los atlas escolares en sus libros, pero en su opinión estos atlas deben cumplir tres funciones: ser fuente de información, dar imágenes claras y fáciles de recordar de los patrones geográficos, así como estimular el interés de los alumnos (Raisz, 1948). Se puede afirmar que la diferencia entre un atlas escolar y otros tipos está determinada fundamentalmente por la edad y conocimientos del “público” o los “usuarios” al que está dirigido, o sea de los alumnos, lo que significa que en el momento de crear los mapas se debe considerar cuáles métodos de representación utilizar, así como cuál es el contenido que los alumnos estudian en dependencia del nivel de enseñanza que cursan.

### ¿Qué principios deben regir la publicación de un atlas escolar?

Al ser el atlas escolar una obra para ser usada por alumnos, su redacción debe cumplir requisitos específicos. Uno de ellos es la selección del lenguaje gráfico usado para la representación de datos. Patrick Wiegand en su libro *Learning and teaching with maps* enfatiza acertadamente que los métodos usados para la representación de datos debe adecuarse a las edades a la que se dirige el atlas, pero al mismo tiempo precisa ser variado (Wiegand, 2006). Eso significa que deben utilizarse (y deben adaptarse) métodos de diferentes tipos (símbolos proporcionales, coropletas, diagramas, etc.), para que los alumnos vayan adquiriendo experiencia en la lectura e interpretación de los datos representados en los mapas. Otro requisito tan importante como el anterior es que su contenido debe incluir los conocimientos que se orientan en los currículos nacionales, para que el atlas pueda ser considerado un verdadero complemento de los libros de texto que utilizan los alumnos, siendo atlas creados para complementar o enriquecer aquellos conocimientos que son ubicados espacialmente (o, para usar una denominación más moderna, que son georeferenciados) y que se imparten en las escuelas de nivel primario y secundario. De esta manera se puede inferir que es muy importante que en cada país se pueda producir sus propios atlas escolares, aunque por diferentes razones existen aún países en los que no se publican y por ello en las escuelas se hace uso de aquellos publicados por firmas o editoriales cartográficas de otros países, pero que son internacionalmente reconocidas por la calidad de sus trabajos. Esto es posible porque existen conocimientos básicos generales que se pueden considerar comunes en todos los sistemas educativos y que usualmente son incluidos en los atlas escolares de cualquier país, por ejemplo los mapas físicos y políticos de los continentes. Pero si analizamos al detalle su contenido, podremos constatar que hay, por lo menos un tema de los currículos nacionales que no pueden presentar en absoluto o de manera totalmente satisfactoria: la información geográfica, económica, social, cultural, etc. más detallada que se relaciona con el país en cuestión y sus regiones, provincias, etc. También podemos encontrar excepciones: cuando una firma cartográfica de un país publica un atlas escolar, cuya venta está específicamente dirigida otro país, por ejemplo la serie de atlas escolares que la editorial Oxford University Press publica periódicamente para ser vendido en la India: en las 160 páginas de la tercera edición del *Oxford Student Atlas for India* encontramos 32 mapas dedicados al subcontinente indio, además de ofrecer como anexo un mapa de la India (Oxford University Press, 2019).

La redacción de los mapas para los atlas escolares es un proceso muy delicado, que pone a prueba la experiencia del cartógrafo: hay que crear mapas que pueden parecer simplificados si los comparamos con los de un atlas mundial hecho para el público en general, pero ese proceso de selección del contenido y su generalización cartográfica siempre debe realizarse tomando en cuenta las indicaciones que ofrecen los currículos nacionales, sincronizando el contenido de los mapas con el de los libros



de texto. No es posible incluir una descripción, ni tan siquiera abreviada, de cada una de las fases del proceso de redacción de los mapas. En su lugar, analicemos brevemente una de las tareas que constituye un momento clave en la edición del atlas: la selección de los nombres geográficos que aparecerán en los mapas y la toma de decisión en cómo escribir correctamente esos nombres. Este es un proceso también complejo, en el cual los cartógrafos cuentan con la colaboración de especialistas en Letras. Dentro de esta tarea, un ejemplo específico es el uso de exónimos. Los exónimos son aquellos nombres empleados en un idioma determinado para aquellos objetos geográficos situados fuera del área, donde ese idioma es oficial y que se diferencian del nombre original. Por ejemplo: en el idioma español Nueva York y Londres son los exónimos de New York y London. La práctica usual en los atlas escolares es tratar de utilizar exónimos siempre que es posible y en algunos países incluso, junto con el exónimo se coloca el nombre en el idioma original entre paréntesis, para ofrecer una mayor información al estudiante principalmente de nivel secundario. Al mismo tiempo, las Naciones Unidas recomiendan disminuir el uso de los exónimos (Tichelaar y Ormeling, 2009), aunque en la práctica cada país mantiene la tendencia a usarlos. En algunos países se hacen listas oficiales de exónimos que se pueden utilizar en los mapas, pero ello no es una práctica usual en todos los países.

La estructura de los atlas publicados para su uso en las escuelas tiene en general características comunes: la mayoría de estos atlas comienzan con una presentación gráfica de aquellos conocimientos básicos que se relacionan con la geografía y los mapas. El contenido a mostrar es muy variado y depende de los grados para los que se publica el atlas. Estas introducciones gráficas generalmente comienzan con conocimientos de carácter astronómico, uno de ellos puede ser la ubicación de la Tierra en el Sistema Solar y el movimiento de la Tierra alrededor del Sol y sobre su propio eje. Hay atlas en los que también se muestran gráficamente otros conceptos relacionados con las ciencias de la tierra: estructura interna del planeta Tierra, estructura y composición de la atmósfera, etc. Esto puede ser seguido por conocimientos elementales de cartografía matemática: la presentación de algunas proyecciones cartográficas y nociones relacionadas con el sistema de coordenadas geográficas. A continuación se puede incluir los métodos de representación del relieve, partiendo desde una imagen satelital hasta los principios básicos de las curvas de nivel. Más tarde se hace la introducción al concepto de escala en los mapas, presentando un territorio a diferentes escalas. Y finalmente se puede utilizar esta sección del atlas para presentar detalladamente la leyenda o las leyendas de sus mapas.

El orden de presentación de los mapas también puede variar en los atlas escolares. Se puede afirmar que la secuencia de mapas en un atlas escolar sigue un orden lógico basado en tres aspectos esenciales: la orientación en tiempo y espacio, la geografía de las geoesferas (litosfera, hidrosfera, atmósfera, clima y zonas geográficas)

y la geografía de las sociedades (población, asentamiento y economía) (Kocsis, Márton y Jordan, 2019). El profesor brasileño Marcello Martinelli (2008) propone el siguiente orden para un atlas publicado según un currículo nacional: presentación de nociones básicas, colección de mapas partiendo del espacio más próximo al alumno a escala relativamente grande hasta los mapas físico y político del país, continuando con los mapas por continentes, para concluir con los mapas físico, político y temáticos del mundo, y otro de los husos horarios. Por ejemplo, los atlas escolares publicados en Hungría siguen un orden muy similar a este, pero el *Atlas Geográfico Escolar Rumano* (Mândruț, 2018) sigue un orden inverso de los mapas, partiendo desde el mundo para finalizar con los mapas político y físico del país, al igual que algunos de los atlas escolares editados en Bulgaria en los últimos 15 años (Bandrova, 2006).

Los atlas escolares tampoco se limitan a ser complemento de los libros de texto relacionados con la Geografía. Un ejemplo típico de ello son los atlas escolares históricos, que se publican para las clases de Historia a diferentes niveles en muchos países. Pero hay también ejemplos que pueden ser considerados curiosos por su singularidad e incluso rareza: en Hungría se publica un *Atlas Escolar de Literatura* para los grados 5 al 8 de las escuelas primarias (56 pp.) y otro para las escuelas secundarias (72 pp.), que contienen la información mapificada de los conocimientos que se imparten en la asignatura de Literatura, desde la vida de los escritores y poetas hasta los lugares donde se desarrolla la trama de los libros que se estudian, además de incluir también mapas sobre la historia de la literatura.

Un caso excepcional y único en la publicación de atlas escolares puede considerarse la experiencia que se está desarrollando en Brasil, a partir de la década de los noventa. Considerando la extensión territorial de país y la diversidad de sus regiones, los educadores brasileños decidieron comenzar la publicación de los llamados atlas escolares municipales, que como su nombre lo indica, contiene información textual y gráfica muy detallada sobre un municipio del país. El primer atlas escolar municipal se publicó en 1994 en Ijuí (Rio Grande do Sul) con la coordinación de la profesora Helena Callai (Lastoria, 2007) y los últimos atlas se han publicado en Pirenópolis (2018) y Distrito Federal (2019). No se tratan de atlas escolares tradicionales formados fundamentalmente por mapas. Tal vez sería más correcto caracterizarlos como un libro de texto profusamente ilustrado y acompañado también de numerosos mapas. El atlas en ocasiones también incluye ejercicios para ser contestados por los alumnos, o para investigar por cuenta propia, su formato es variado, así como la impresión, que puede ser a color o en blanco y negro. Su característica más meritoria es que ofrece conocimientos propios de la región donde vive el alumno, así como la masividad de la colección, ya que se han publicado numerosos atlas de este tipo en los 25 años transcurridos desde sus inicios.

## Atlases escolares para la educación del futuro

### *Atlases escolares electrónicos*

Los atlases escolares tampoco podían ser una excepción del impacto ocasionado por la revolución digital que la aparición de las computadoras personales generó a partir de la década de los ochenta. En los años noventa, se comenzaron a desarrollar investigaciones y publicarse los llamados atlases electrónicos (también conocidos como digitales). Las primeras versiones digitales de atlases (incluyendo también los escolares), en muchos casos se publicaban en un disco compacto que acompañaba la versión impresa, utilizando las nuevas herramientas que ofrecían las técnicas digitales. Estos son aquellos atlases que en 1996 Kraak y Ormeling identificaban como “view-only atlases” (atlases solo para visualización o lectura) y que definían como una extensión de los atlases impresos, brindando nuevas opciones propias del mundo digital (acceso rápido y “al azar” de los mapas a través de un menú, opción de visualizar en pantalla una parte aumentada del mapa, etc.) pero usando solamente mapas estáticos. Otro tipo de atlases electrónicos o digitales son los llamados interactivos, que permitían que el usuario decidiera cómo deseaba representar los datos a los que tenía acceso, por ejemplo cambiando los colores con que se representaban los datos o clasificando esos datos de diferentes maneras. Por último, también se desarrollaron los atlases analíticos, que como su nombre indica, permitían operaciones más complejas con los datos a través de los mapas o combinando los diferentes datos para crear mapas nuevos (Kraak y Ormeling, 1996).

Los discos compactos (CD) que acompañaban a los atlases escolares generalmente incluían materiales multimedia como vídeos de temas cartográficos (historia de los mapas, proyecciones, tipos de mapas, etc.), animaciones para mostrar diferentes conceptos geográficos o astronómicos (movimiento de rotación y traslación de la Tierra), mapas en formato *raster* (imagen) para su uso en las clases o por los alumnos en sus tareas individuales, juegos interactivos educativos y tablas de datos geográficos, económicos, etc. De esta manera, los discos compactos convertían al atlas escolar en un material educativo más íntegro, que brindaba múltiples opciones novedosas al alumno (y también al maestro) para enriquecer sus conocimientos. El uso de discos compactos para complementar el atlas se mantuvo hasta los primeros años de la década del 2010, cuando ya estaba ganando popularidad entre las generaciones más jóvenes el uso de los dispositivos móviles y en especial de los teléfonos móviles inteligentes (smartphones). Adaptándose a las más novedosas tecnologías, los atlases escolares también sustituyeron los discos compactos por una opción más sencilla que combina el uso de los dispositivos móviles con Internet: los materiales multimedia se almacenan en un servidor con acceso en la web, y se crearon aplicaciones que permitían conectar el atlas con el dispositivo móvil. Esta solución se expande sin mayores obstáculos, no solo por su facilidad para usarla, sino porque

también significa un ahorro material y financiero para las firmas cartográficas. Así, en la actualidad se hace usual encontrar en una esquina de la portada de un atlas escolar un texto breve pidiendo que se descargue una aplicación o presentando también un código QR para tener acceso a más mapas, ejercicios interactivos, animaciones, videos, juegos, etc.

### ***Atlases escolares y tecnologías móviles***

La opción de usar aplicaciones con un dispositivo móvil para acceder a materiales anexos a un atlas escolar puede ser considerada tan solo una especie de introducción a la verdadera adaptación de estos atlases al llamado mundo móvil. Para que eso ocurra, debemos ofrecer una respuesta cartográfica y tecnológicamente bien fundamentada a la pregunta: ¿cómo pueden servirse los atlases escolares de Las nuevas soluciones que se desarrollan para la cartografía móvil (por ejemplo los *Location-based services*, LBS)?

Actualmente, en diferentes tiendas virtuales, se pueden comprar atlases supuestamente creados para los dispositivos móviles. Si los analizamos, veremos que en muchas ocasiones se trata de “rediseños” y solo en contados casos realmente “reeditados” para los dispositivos móviles. La mayoría está compuesto solamente por mapas estáticos (a menudo versiones en formato *raster* de los mapas de un atlas), haciendo un uso mínimo de las nuevas opciones brindadas por esos dispositivos, limitándose a diseñar un entorno adaptado al tamaño pequeño de la pantalla y a aprovechar la ventaja que significa el uso de la pantalla táctil para la selección de opciones o simplemente “hojear el atlas”.

Pero un atlas escolar solo puede ser llamado “móvil” si es capaz de aprovechar las ventajas que ofrece esta tecnología y que no pueden ofrecer los atlases publicados en otros soportes electrónicos, por ejemplo en discos compactos. Reichenbacher (2001) definió claramente lo que debe ser la cartografía móvil: “es aquella que trata de las teorías y tecnologías de la visualización cartográfica dinámica de los datos espaciales y su uso interactivo en dispositivos portátiles en cualquier momento y en cualquier lugar bajo la consideración especial del contexto actual y de las características de los usuarios”.

Si nos basamos en esta definición, entonces se puede afirmar que la edición y diseño de un atlas escolar para dispositivos móviles debe satisfacer tres condiciones concretas (Reyes, 2013):

1. Posicionamiento en espacio y tiempo (*positioning in space and time*): la primera condición que posibilita la ejecución de las dos siguientes. Este posicionamiento es totalmente automático, cuando el dispositivo móvil dotado de GPS determina la situación geográfica de la persona que lo usa así como la fecha y la hora (a tiempo real) de la tarea que está solucionando en ese momento.

2. Reconocimiento del contexto (*context awareness*): después del posicionamiento, el atlas escolar puede identificar y coleccionar las informaciones relacionadas con el medio ambiente (contexto) en que se encuentra el usuario (el alumno). Estos datos pueden ser colectados por los llamados “geomotores” (*geo-engines*), que fueron definidos por Reichenbacher (2001). Los datos pueden ser utilizados conjuntamente con aquellos datos más generales que son almacenados previamente en las bases de datos que utiliza el atlas, por ejemplo, datos meteorológicos de una región determinada que han sido recogidos durante años u otras informaciones relacionadas con el contexto actual y que pueden ser culturales, históricas, sociales, etc.
3. Representación adaptativa (*adaptive representation*): solución gráfica para representar los datos seleccionados durante el reconocimiento del contexto. Estos deben ser representados en un mapa dinámico, cuyo contenido temático puede ser automáticamente modificado según la variación de los datos a tiempo real. Esto es lo que Safiza (2007) llama un modelo cartográfico flexible (*flexible cartographic model*) característico de las aplicaciones de tipo LBS.

*¿Hasta qué punto precisamos usar de un modelo cartográfico flexible en un atlas escolar?*

El uso de un modelo cartográfico flexible no tiene que ser de carácter obligatorio para todos los mapas de un atlas escolar móvil, porque hay escalas y temas que tradicionalmente aparecen en los atlas escolares y no necesitan de una actualización a tiempo y espacio real. Un ejemplo de este tipo de mapas puede ser un mapa físico del mundo u otro que presente el relieve e hidrografía de un continente, una región o de un país.

Los modelos cartográficos flexibles pueden ser utilizados para complementar y actualizar la información representada en algunos mapas temáticos a mayor escala después del reconocimiento del contexto del alumno. Incluso podemos permitirnos ir mucho más allá de los límites y objetivos tradicionales de los atlas escolares al aplicar las soluciones que brinda la tecnología móvil: el uso de los geomotores definidos por Reichenbacher (2011) nos ofrece la base tecnológica necesaria para expandir, multiplicar las funciones del atlas escolar y convertirlo en un material educativo interactivo mucho más completo. Este material lúdico no incluiría solo mapas, sino también ejercicios y juegos interactivos que serían automáticamente adaptados al contexto del alumno a través de su dispositivo móvil. Podríamos llamarlos geoejercicios adaptivos (*adaptive geo-exercises*), que combinarían las informaciones obtenidas a tiempo y espacio real con aquellas que fueron previamente almacenadas en la base de datos del atlas.

Para ejemplificar su uso podríamos imaginar una clase de Geografía en una escuela secundaria, en la que los alumnos estudian una lección sobre el clima. Si combinamos los datos climáticos previamente almacenados en nuestro atlas escolar (que pueden ser datos de temperatura, precipitaciones o presión atmosférica en diferentes ciudades del mundo) con los datos obtenidos a tiempo real sobre el estado del tiempo en la ciudad donde vive el alumno, entonces los ejercicios se le plantearán usando ambos tipos de datos, cobrando mayor actualidad y ayudándolo a una mejor comprensión de la enseñanza. Esta realmente no es una propuesta “futurística”, porque el desarrollo de la tecnología móvil no solo nos brinda estas opciones en el presente, sino incluso son actualmente utilizadas en aplicaciones didácticas creadas específicamente para niños y jóvenes.

Es lamentable que las grandes firmas cartográficas actualmente no se muestren interesadas en desarrollar los atlas escolares en esta dirección. En años anteriores se experimentaron algunas soluciones, por ejemplo una firma muy reconocida internacionalmente comenzó a comercializar su *Atlas Mundial* como una aplicación para dispositivos móviles, que conjuntamente con las soluciones tradicionales permitía también, por ejemplo, obtener datos a tiempo real del estado del tiempo en un país seleccionado por el usuario. Desdichadamente solo desarrollaron esta aplicación durante algunos años, y la última versión de este fue publicada en el 2015.

### ***Atlases escolares y los servicios de mapas en la web***

Una posible alternativa a la falta de atlas escolares móviles puede ser el uso de los servicios de mapas en la web para representar el contenido que se enseña en las clases. De cierta manera podemos afirmar que la aparición en el mercado de Google Maps y Google Earth en el año 2005 significó “un antes y después” en la historia de la cartografía contemporánea. Los dos servicios de la firma Google revolucionaron el uso tradicional de los mapas por los usuarios: los mapas de ciudades y turísticos en el sentido más clásico de la palabra, quedaron prácticamente sustituidos por estas dos aplicaciones en un período de tiempo relativamente corto. Si hasta ese momento la cartografía digital y la tecnología SIG era un mundo limitado a los especialistas y a algunos pocos interesados, principalmente Google Maps pero también Google Earth convirtieron los mapas digitales, las imágenes satelitales y las herramientas de búsqueda más sencillas en un fenómeno de masas: es la primera vez en la historia de la cartografía que se pone en manos de los usuarios (es decir, en manos de un público con un conocimiento sobre el uso de los mapas que puede ser considerado promedio o por debajo del promedio) un producto prácticamente gratis y de funcionamiento fácil de aprender para satisfacer casi todas sus necesidades relacionadas con la orientación espacial.

El uso cada vez más masivo de las aplicaciones desarrolladas por la firma Google y otras firmas que surgirían después del 2005 (OpenStreetMap en el 2006, Bing Maps

en el 2009 así como Apple Maps y HereWeGo en el 2012) traerían consigo el surgimiento de nuevos conceptos o la redefinición de otros en la cartografía de la web: información geográfica voluntaria (*volunteered geographic information*, VGI), contenido generado por el usuario (*user generated content*), geografía de masas (*mass geography*), cartógrafos laicos (*lay cartographers*), datos a grande escala o masivos (*big data*) entre otros.

Su influencia no podía dejar de sentirse también en la educación. Solo los temores de los maestros de mayor edad al uso de las nuevas tecnologías o aquellos con una formación informática insuficiente, así como las limitaciones materiales de una escuela podían obstaculizar la expansión de su uso en las clases de Geografía u otras asignaturas. Profesores de educación superior e investigadores en numerosos países han estudiado y aún hoy continúan estudiando las opciones que brinda su uso en las escuelas. Algunos ejemplos:

- Proyecto “La conciencia global en la escuela - Comprender el mundo con mapas” (*Global Awareness in School – Understanding the World with Maps*) ejecutado por el Departamento de Geografía de la Universidad de Viena (Austria) y cuyo objetivo principal fue experimentar el estudio del uso de Google Earth para representar datos temáticos en mapas creados para las escuelas (Kriz, 2008).
- Proyecto “Museo de Globos Virtuales”, desarrollado por el Departamento de Cartografía y Geoinformática de la Universidad Eötvös Loránd en Budapest, Hungría (Figura 5). Iniciado en el año 2007, actualmente este museo contiene los modelos virtuales de 152 globos terráqueos, estelares, lunares y de otros planetas. El más antiguo de ellos data de 1507. El objetivo de este proyecto es permitir el uso interactivo (rotación, ampliación, etc.) de los globos digitalizados y visualizados en Google Earth (Gede *et al.*, 2013). La página web donde se puede acceder la versión en inglés del museo es: <http://terkeptar.elte.hu/vgm/?lang=en>
- Proyecto “Geotecnologías y recursos de multimedia en la enseñanza de la cartografía: percepción socioambiental del Rio Alcântara en el municipio São Gonçalo (Rio de Janeiro)”. Investigación realizada para una tesis de maestría en la Universidad Estatal de Rio de Janeiro, cuyo objetivo fue organizar una encuesta con la participación de alumnos de una escuela del barrio de São Gonçalo en la ciudad de Niterói, para preguntar a los vecinos sobre temas relacionados con el medio ambiente al mismo tiempo que recolectaban informaciones sobre aquellos lugares en el barrio de interés histórico, social, turístico, etc. En la siguiente etapa del proyecto se enseñó a los alumnos cómo representar en mapas estas informaciones usando Google Earth (Barros de Sousa, 2014).

Estos proyectos nos permiten vislumbrar las posibilidades que brindan los servicios de mapas en la web para la aparición de los futuros atlas escolares que utilicen como base cartográfica estas aplicaciones. Mi opinión es que tanto estos servicios de



**Figura 5.** Uno de los 152 globos que contiene el Museo de Globos Virtuales.

Fuente: <http://terkeptar.elte.hu/vgm/?lang=en>

mapas (y principalmente Google Maps por su gran popularidad entre los usuarios laicos) pueden ser utilizados en la cartografía escolar, pero no deben (y aún no pueden) sustituir totalmente los mapas de un atlas escolar editado especialmente para determinados grados de la enseñanza primaria o secundaria. Esta afirmación está basada en que la cartografía temática prescribe que el contenido de las bases cartográficas se determine en dependencia del tema que se representa en un mapa y esta es una opción que los productos Google aún no ofertan, ya que solamente ofrece ocho tipos de bases cartográficas que se diferencian entre sí por el diseño, no permitiendo al usuario decidir cuáles capas (*layers*) son usadas o no en nuestro mapa.

El uso de los servicios de mapas también puede ser útil en otras tareas inherentes a los atlas escolares, ya en el siglo XIX era tradición incluir una introducción para representar y explicar gráficamente diferentes conceptos relacionados con los mapas. Un ejemplo típico es la descripción gráfica sobre cómo varía la representación de un territorio desde las grandes escalas a las más pequeñas. Este es un tema que Google Maps o Google Earth puede asumir por completo y de forma mucho más efectiva e interactiva que la cartografía tradicional. En un atlas escolar tradicional se presenta este proceso a través de un territorio muy conocido dentro del país o tomando como punto de partida una escuela “imaginaria”, para ir aumentando a través de diferentes fragmentos de mapas el territorio que se muestra y así ir disminuyendo también la escala de los mapas que se utilizan en esta representación. Los productos Google pueden cumplir esta misma tarea automáticamente, ya que después que el alumno

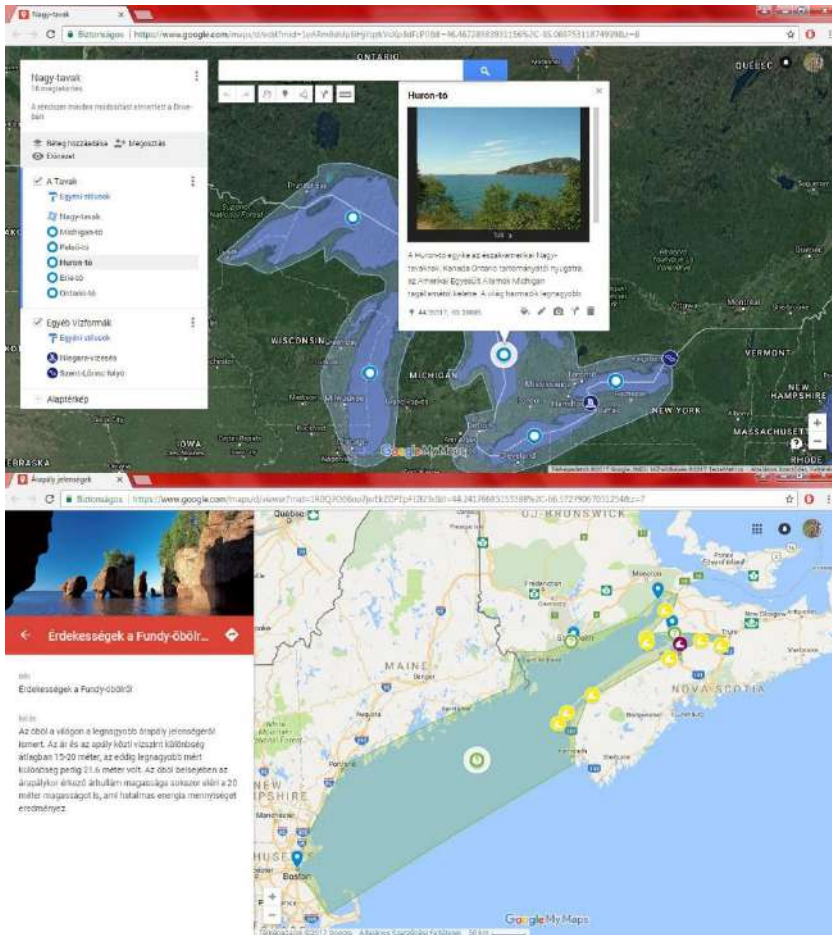


identifica su ubicación en el mapa (por ejemplo, dónde está su escuela) puede disminuir interactivamente la escala para mostrar cómo varía la representación a medida que aumenta el territorio a visualizar en la pantalla (Figura 6).



**Figura 6.** La parte superior muestra cómo se ilustra la variación de la escala en un atlas escolar húngaro de 1890 y en la parte inferior vemos el mismo proceso en Google Earth.

El uso de los servicios de mapas en la cartografía escolar (y fundamentalmente de Google Maps o Apple Maps) es ventajoso también por otros motivos más prácticos: a partir del año 2005 el diseño de estas aplicaciones se ha hecho familiar para millones de usuarios en todo el mundo (incluyendo las generaciones más jóvenes que aún estudian en las escuelas), además de ser una plataforma empleada en numerosas aplicaciones cartográficas desarrolladas para dispositivos móviles. Todo ello contribuye a hacer más rápido y sencillo el aprendizaje del uso de un atlas escolar por los usuarios. Esta ventaja fue aprovechada por Reyes y Kiss (2018) en un proyecto que permitió ver cómo estudiantes de noveno grado de una escuela secundaria húngara podían utilizar Google Maps y la aplicación Google My Maps para hacer ellos mismos sus mapas, después de haber estudiado en clases la unidad dedicada a la Geografía del agua. Se dedicó un turno de clases (45 minutos) para enseñar el uso de esa aplicación, al final de la cual los alumnos se organizaron en parejas y como tarea para la casa hicieron de manera independiente 13 mapas relacionados con ese tema (Figura 7).



**Figura 7.** Mapas de temas relacionados con la Geografía del Agua hechos en Google Maps por alumnos húngaros (Reyes y Kiss, 2018).

Cuando hablamos de servicios de mapas en la web, tampoco se puede olvidar un segmento de los Sistemas de Información Geográfica que comienza a cobrar también popularidad: los llamados GIS Online, cuyo representante más notorio es la página ArcGIS Online (<https://www.arcgis.com/index.html>). Esta página web ofrece la posibilidad a cualquier persona de registrarse y usar de manera gratuita una versión muy simplificada del sistema ArcGIS, diseñada para su uso en la nube web y por cualquier persona que desee crear un mapa sencillo basándose en datos temáticos. ArcGIS Online también tiene un llamado “Atlas Vivo del Mundo” (*Living Atlas of the World*) que nos permite el acceso gratuito a numerosos mapas de diferentes temas

que pueden ser utilizados en la educación. ArcGIS Online incluye también una opción que tal vez sea la más conocida: la creación de *story maps*, que son materiales multimedia interactivos que se realizan usando diferentes diseños. Utilizando estos *story maps* se puede incluso crear un atlas escolar en línea, como quedó demostrado con uno de los proyectos de investigación desarrollado por universidades españolas en los últimos años: “Atlas Digital Escolar: aprender Geografía en Educación Secundaria con ArcGIS Online” (Figura 8) cuya idea general se accede en la siguiente dirección: <http://atlasescolar.maps.arcgis.com/apps/MapJournal/index.html?appid=77ae3efc94174a2fb216abda32b564f4> (González *et al.*, 2016).



Figura 8. Portada del Atlas Digital Escolar.

### A manera de epílogo

Las investigaciones no se limitan a las opciones anteriormente descritas y ya se buscan nuevos horizontes para experimentar soluciones aún más novedosas. Yendo más allá de la web, se comienza a estudiar cómo aplicar las tecnologías más modernas en la cartografía, abriendo el camino para su implementación en la educación. En el 2019 tuve la oportunidad de escuchar las ideas expuestas por un joven investigador búlgaro, Nikolai Yonov, sobre las posibilidades que la realidad aumentada (*Augmented Reality*, AR) nos ofrece para complementar la visualización de un atlas escolar con informaciones temáticas (datos, imágenes, modelos tridimensionales, videos, sonidos, etc.) recogidas y actualizadas a tiempo real (Yonov, 2019). ¿Podrá aplicarse esta tecnología de manera masiva en un futuro cercano? Debemos confiar en que se pueda dar una respuesta positiva a esta interrogante en los próximos años.

## Bibliografía

- Anderson, J., Carrière, J. y LeSann, J. (2003). “A Pilot Electronic School Atlas of Quebec, Canada”, *International Research in Geographical and Environmental Education*, 12(4): 383-390. <https://doi.org/10.1080/10382040308667552>
- Bandrova, T. (2006). *Atlas de Geografia y Economía para 11 y 12 grados*, Sofía. Datamap.
- Barros de Sousa, I. (2014). “Geotecnologias e recursos de multimídia no ensino de cartografia: percepção socioambiental do Rio Alcântara no Município de São Gonçalo/RJ”, tesis de grado, Universidade do Estado do Rio de Janeiro.
- Gede, M., Ungvári, Zs. and Zentai, L. (2013). *Virtual Globes Museum 2.0 – Adding the Power of Community*, *Proceedings of the 26th International Cartographic Conference*, 600 pp., Dresden.
- Miguel González, R., Lázaro, M.L., Gil, J., Buzo, I. y Moreno, C. (2016). *Atlas Digital Escolar: aprender Geografía con ArcGIS Online*. <https://doi.org/10.14198/GeoAlicante2015.68>
- Kocsis, K., Márton, M. y Jordan, P. (2019). “Geographic/Thematic Content”, in *The Atlas Cookbook*. Recuperado de <https://atlas.icaci.org/awards-and-publications/atlas-cookbook/>
- Koop, O. (1993). “Electronic Atlas Production Tools”, in Klinghammer, I., Zentai, L. y Ormeling, F. (eds.), *Proceedings of the Joint ICA Seminar on Electronic Atlases*, pp. 129-137.
- Kraak, M.-J. and Ormeling, F. (1996). *Cartography: visualization of spatial data*, Essex, Longman.
- Kriz, K. (2008). “Cartographic Design Issues utilizing Google Earth for Spatial Communication”, en *Proceedings of the 6th ICA Mountain Cartography Workshop*, Lenk, Switzerland, pp. 129-131.
- Lastória, A.C. (2007). “A cartografia escolar e a concepção de atlas escolar municipal”, en *Dialogus*, 3: 111-126. Recuperado de <http://sites.ffclrp.usp.br/laife/atlas.htm>
- Martinelli, M. (2008). *Um atlas geográfico escolar para o ensino-aprendizagem da realidade natural e social*. *Portal da Cartografia*, 1(1): 21-34. Recuperado de <http://www.uel.br/revistas/uel/index.php/portalcartografia>
- Mândruț, O. (2018). *Atlas Geografic Școlar*, Bucarest, Corint.
- Ormeling, F. (1996). “Functionality of electronic school atlases”, *Seminar on Electronic Atlases II*. Recuperado de <http://lazarus.elte.hu/hun/doktoran/ferjan/1997-electronic.doc>
- (2014). “Atlases”, in Ormeling, F. y Rystedt, B. (eds), *The World of Maps*. Recuperado de <https://icaci.org/publications/the-world-of-maps/the-world-of-maps-english/>

- Oxford University Press (2019). *Oxford Student Atlas for India*, Third Edition, Oxford.
- Patton, J.C. (1999). "The American School Atlas: 1784-1900", *Cartographic perspectives*, 33(2): 4-32.
- Raisz, E. (1948). *General Cartography*, New York, McGraw-Hill.
- (1962). *Principles of Cartography*, New York, McGraw-Hill.
- Reichenbacher, T. (2011). "The World in your pocket - towards a Mobile Cartography", in *Proceedings of ICC 2011*. Recuperado de [http://icaci.org/files/documents/ICC\\_proceedings/ICC2001/icc2001/file/f16005.pdf](http://icaci.org/files/documents/ICC_proceedings/ICC2001/icc2001/file/f16005.pdf).
- Reyes, J; Kiss, J. (2018). "Ideas y experiencias sobre el uso de servicios de mapas web en la educación", en *Boletim Paulista de Geografia* (99): 86-101, São Paulo.
- Reyes Nuñez, J.J. (2013). "Smartphone-based school atlas?", *Cartographica*, 48(2): 126-133. University of Toronto Press.
- Safiza, S.B.K.B. (2007). "Cartographic representation on small devices", RGI 149 Report no. 2, Delf University of Technology.
- Tichelaar, T. y Ormeling, F. (2009). *Geographical names and education: school world atlases*. Recuperado de [http://www.stagn.de/static/UNGEGN/2009/11\\_Geographical%20names%20and%20education.pdf](http://www.stagn.de/static/UNGEGN/2009/11_Geographical%20names%20and%20education.pdf)
- Voženilek, V. (2014). "Aspects of the Thematic Atlas Compilation", in Brus, J., Vondrakova, A. y Voženilek, V. (eds.), *Modern Trends in Cartography: Selected Papers of CARTOCON 2014*, Heidelberg. Editorial Springer, pp. 3-12.
- Wiegand, P. (2006). "Learning and teaching with maps", New York, Routledge.
- Yonov, N. (2019). "School Atlas with Augmented Reality", *Proc. Int. Cartogr. Assoc.*, 2: 150. <https://doi.org/10.5194/ica-proc-2-150-2019>



# Columbus: aportes de la teledetección y análisis de coste anisotrópico a un debate histórico

*Columbus: contributions of remote sensing and anisotropic cost analysis to an historical debate*

Ángel M. Felicísimo\*

*Recibido 05 de noviembre de 2019; aceptado 10 de diciembre de 2019*

## Resumen

La ruta seguida por la primera expedición de Cristóbal Colón es incierta debido a la ausencia o poca fiabilidad de esta información en el diario del almirante. Los actuales datos procedentes de teledetección y su análisis mediante SIG pueden aportar nueva información mediante análisis de los campos de viento estacionales en la zona. Los resultados del análisis de seis años muestran que el régimen de vientos es homogéneo y que configura dos corredores o pasillos de bajo coste relativo para la navegación a vela con las condiciones de los barcos del siglo XV. De estos dos corredores uno conduce con exactitud al destino al que se cree que llegó Colón el 12 de octubre de 1492 y se presenta como la ruta más probable, algo más al Sur que la propuesta hasta el presente en la bibliografía. Los SIG se muestran como una herramienta capaz de aportar información inédita a partir de datos globales, tomados en este caso por el escaterómetro QuikSCAT, aplicando métodos de análisis espacial poco usados en la literatura científica.

Palabras clave: *Cristóbal Colón, análisis de coste anisotrópico, viento, QuikSCAT, Sistemas de Información Geográfica.*

\* Universidad de Extremadura, España, correo electrónico: [amfeli@unex.es](mailto:amfeli@unex.es).  
ORCID: <https://orcid.org/0000-0002-9953-8614>.



## Abstract

The route of the Christopher Columbus' first expedition is uncertain due to the absence or unreliability of this information in the admiral's diary. Current remote sensing data and GIS analysis can provide new information through analysis of the seasonal wind fields in the area. The results of the six-year analysis show that the wind regime is homogeneous and that it configures two relatively low-cost corridors for sailing with the conditions of 15th century ships. One of these two corridors leads exactly to the destination where Columbus is believed to have arrived on October 12th 1492 and is presented as the most likely route, somewhat further south than that proposed so far in the literature. The GIS are shown to be a tool capable of providing unprecedented information from global data, taken in this case by the QuikSCAT catterometer, applying methods of spatial analysis rarely used in the scientific literature.

Key words: *Christopher Columbus, anisotropic cost analysis, wind, QuikSCAT, Geographic Information Systems.*

## Introducción

La navegación marítima es una actividad humana que tiene varios miles de años de antigüedad. Las técnicas de navegación permitieron la colonización atravesando las barreras que suponían los océanos, antes absolutas. Partiendo de las balsas simples, apenas manejables, se consiguió un control progresivo de la navegación mediante la incorporación progresiva de elementos estructurales como remos, velas, diferentes formas de timón, etc. Algunas barcas ya ciertamente complejas se han datado en más de seis mil años en tumbas egipcias, aunque parece que se limitaban a la navegación fluvial. Estos elementos, el perfeccionamiento de los cascos y el aprendizaje sobre técnicas de orientación permitieron la navegación marítima a distancias más lejanas de las costas. Los fenicios dominaban el Mediterráneo en el primer milenio a.C. con barcos dotados de remos y vela, los cretenses tenían comercio marítimo mil años antes y la primera batalla naval documentada históricamente fue hacia el año 1200 a.C. entre chipriotas e hititas. En el otro lado del mundo, en los océanos Índico y Pacífico, la navegación oceánica permitió colonizar los archipiélagos de la Polinesia, cubriendo distancias de miles de km con canoas de uno o más cascos; por ejemplo, la cultura lapita llegó a Samoa y Tonga entre los años 1300 y 900 a.C. en una dispersión que les alejó hasta 6 000 km desde su origen probable en el sudeste asiático. Debe mencionarse necesariamente la navegación vikinga que sobre el siglo IX llegó a Islandia, a Groenlandia en el siglo X y de la que se han encontrado restos de un asentamiento en L'Anse aux Meadows, en Terranova (McGovern, 1982).

Puede decirse que la edad de oro de la navegación oceánica fue en la segunda mitad del siglo XV, no solo con los viajes colombinos, sino con la apertura de rutas



al Océano Índico por Bartolomeu Dias o Vasco de Gama, y la primera circunnavegación terrestre por la expedición Magallanes-Elcano que partió de España en 1519. Estos grandes viajes fueron posibles gracias a adelantos técnicos como la brújula —antiguo invento chino cuya primera mención en Europa es el siglo XII—, el astrolabio —conocido en el mundo islámico desde el siglo VIII y llegado a Europa hacia el XII—, o la corredera para estimar la velocidad —finales del siglo XV o inicios del XVI.

En este contexto, el objetivo del presente trabajo es presentar nueva información relativa a uno de los viajes marítimos más famosos de la historia: el de la expedición dirigida por Cristóbal Colón en 1492, que supuso el descubrimiento de América para el continente europeo. Esta información es el resultado de análisis de datos tomados por satélites y procesados mediante Sistemas de Información Geográfica (SIG) y es un ejemplo de cómo datos actuales pueden aportar algo de luz a acontecimientos pasados.

Este tipo de análisis ya fue utilizado por nuestro grupo de investigación en trabajos de naturaleza biológica sobre dispersión a larga distancia (Muñoz & Felicísimo, 2006) y sobre migraciones de aves oceánicas (Felicísimo *et al.*, 2008; Gonzalez-Solis *et al.*, 2009) y ha conseguido aportar una evidencia sólida sobre el control del viento en los fenómenos de dispersión y migración a grandes distancias.

En este caso es necesaria una introducción sobre las características de la navegación en el siglo XV, las del primer viaje de Colón y las deficiencias en la documentación del mismo, derivadas de la ausencia de instrumentos para trazar la derrota de los barcos en aguas abiertas. Los datos siguientes se han tomado de Martín López (2019), que analiza y comenta la principal fuente disponible sobre los días del primer viaje colombino. Es el llamado “diario de a bordo de Colón”, aunque en realidad es un manuscrito copia del original—o de una copia del original, actualmente, ambos perdidos—, realizada por Bartolomé de las Casas hacia 1530.

La primera cuestión relevante para este trabajo es que no se conocen con detalle las características concretas de los barcos que se usaron en la expedición. Aunque suele hablarse de tres carabelas, de nombres “Santa María”, “Pinta” y “Niña”, la opinión más aceptada en estos momentos es que Colón armó una nao, la “Santa María”, nave capitana, y dos carabelas: la “Pinta” y la “Niña”. El aspecto más importante de esta cuestión es que la nao era una embarcación incapaz de navegar con vientos laterales por lo que su rumbo era esencialmente el del viento en popa, con ángulos de ceñida muy cerrados. La carabela tenía más capacidad en la navegación aprovechando el viento lateral, pero se sabe por el mencionado diario de Colón que los tres barcos fueron juntos, por lo que los rumbos cortando el viento estaban limitados a las capacidades de la nao Santa María.

La segunda cuestión es sobre la derrota de los barcos y, en concreto, sobre la determinación de longitud, latitud y distancia recorrida. Ya se ha comentado las limitaciones de los instrumentos en esta época, pero a esto se añade la escasez y poca

fiabilidad de los datos que Colón incluye en su diario. Hay numerosas referencias al uso de la brújula, pero Martín López (2019) señala las grandes dificultades para mantener el rumbo algo que “solo era posible realizar con viento favorable y de dirección constante”. Por otra parte, no hay datos de latitud ni longitud. Sobre la longitud no hay sorpresa ya que su determinación con cierta fiabilidad solo fue posible varios siglos más tarde, con la disponibilidad de cronómetros suficientemente exactos. La latitud podría haberse determinado con el cuadrante —el sextante no existía— pero solo se mencionan medidas en tres ocasiones y en todas ellas los datos se consideran pésimos, hasta el extremo de que Colón decidió dejar de tomarlos.

Finalmente, la distancia recorrida es prácticamente una incógnita ya que la corredera no se había inventado aún y, aunque hay estimaciones en el diario, su fiabilidad es muy dudosa. Por añadidura, Colón reconoce en su diario que miente sistemáticamente para no alarmar a las tripulaciones por las grandes distancias recorridas. A partir de esta información, se han planeado hipótesis sobre la ruta de esta primera expedición. En la Figura 1 se puede ver la trayectoria mostrada en un mapa de la Biblioteca Virtual Miguel de Cervantes (Fundación Biblioteca Virtual Miguel de Cervantes, s.f.) donde el viaje transcurre desde La Gomera (28° N) hasta arribar al supuesto punto de llegada, la isla de San Salvador (24° N).



**Figura 1.** Trayectoria del primer viaje a América de Cristóbal Colón.

Fuente: Elaboración propia a partir del mapa de la Biblioteca Virtual Miguel de Cervantes; base OSM.

En resumen, la trayectoria o derrota de las naves es incierta y, aunque hay hipótesis diversas, la verificación de las mismas no es posible con los datos disponibles. Los datos de los que partimos en este trabajo son pocos pero suficientes para obtener

algunos resultados: 1) el punto de salida de las naves fue la isla de La Gomera, en las Islas Canarias y 2) las naves no tenían capacidad para realizar ceñidas significativas y debían mantener el viento por la popa con pocas desviaciones.

A partir de estas premisas, la idea del presente trabajo es realizar simulaciones de trayectorias sobre campos de viento en la misma zona y en las mismas fechas. El objetivo es mostrar la ruta más probable, compatible con las restricciones de entrada. Para ello es necesario disponer de datos de viento y de métodos de análisis que muestren la mayor o menor dificultad para desplazarse sobre ellos en función de su acimut y velocidad. El método adecuado para esta labor se denomina análisis de coste anisotrópico y será explicado posteriormente.

## Material y métodos

### *Sistemas de Información Geográfica*

El análisis de coste anisotrópico se ha realizado con el Sistema de Información Geográfica ArcInfo (ESRI Inc.). El motivo es que cuando comenzamos a trabajar con este tipo de análisis, existían varias opciones para trabajar con análisis de coste isotrópico, pero solo ArcInfo tenía herramientas para el anisotrópico. Durante los primeros trabajos, comenzados en el año 2003, se desarrolló y depuró una cantidad notable de código para este tipo de procesamiento. Aunque parte de él ha cambiado, especialmente el de lectura y transformación de los ficheros originales, lo hemos conservado durante más de una década completamente funcional, aunque hace unos años que dejó de comercializarse. Las rutinas fueron programadas en el lenguaje propio de ArcInfo, AML y, con modificaciones menores, son las que se han utilizado para este trabajo.

La generación de mapas e imágenes y algunos cálculos de asignación de valores de coste a localizaciones se ha realizado con QGIS 3.10.

### *Datos de viento*

Los datos de viento provienen del escaterómetro SeaWinds portado por el satélite QuikSCAT, desarrollado por el JPL de la NASA para medir la velocidad y la dirección de los vientos de la superficie del océano. El primer QuikSCAT fue lanzado el 19 de junio de 1999 y posteriormente se han sucedido varios satélites portadores de instrumentos similares<sup>1</sup>. El instrumento SeaWinds es un radar especializado (escaterómetro) que mide la velocidad y dirección de los vientos a una altura nominal de 10 m sobre la superficie marina (JPL, Jet Propulsion Laboratory, 2001).

Todos los productos estándar de QuikSCAT están en formato HDF. Los archivos HDF de QuikSCAT contienen un conjunto de 16 SDS —*scientific datasets*—; cada

<sup>1</sup> <http://www.eumetrain.org/data/4/438/navmenu.php?tab=2&page=3.0.0>

uno de ellos es una matriz rectangular de 1440 píxeles de este a oeste y 720 píxeles de norte a sur  $0.25^\circ$  en una proyección estándar “plate carrée” con  $0.25^\circ$  de resolución espacial en el Ecuador.

Para este trabajo se han usado datos de los años 1999 a 2005 correspondientes a los períodos que coinciden con la expedición colombina. Los datos se descargaron del PO-DAAC (Physical Oceanography Distributed Active Archive Center) donde están disponibles gratuitamente<sup>2</sup>. Los años más recientes han sido descartados para evitar la posible influencia de los últimos años muy cálidos en los patrones espaciales de trayectorias del viento, algo no verificado pero potencialmente posible.

El procesamiento de este tipo de datos no es inmediato. Los valores de una imagen clásica de teledetección son datos escalares, distribuidos en una escala lineal, que pueden ser procesados por métodos estadísticos convencionales. Sin embargo, los píxeles del QuikSCAT almacenan información vectorial ya que representan velocidad y acimut del viento mediante los componentes  $u$  y  $v$  de cada vector, proyecciones de cada vector sobre los ejes X e Y, cada uno en un SDS diferente. De acuerdo con la convención adoptada para los datos de QuikSCAT, una dirección de viento de  $0^\circ$  implica un flujo hacia —no desde— el Norte.

El primer paso en el procesamiento de los vectores QuikSCAT para este trabajo ha sido la extracción de los SDS que contienen los valores  $u$  y  $v$  de dichos vectores y su transformación a un formato utilizable por el SIG utilizado que en este caso fue ESRI *grid*.

El procesamiento de este tipo de datos en un SIG, presenta algunas dificultades que es necesario superar. La primera de ellas es que la cobertura diaria de QuikSCAT no es completa, efecto de la trayectoria orbital. Las áreas sin datos no pueden aceptarse porque rompen la continuidad espacial necesaria para el análisis de la trayectoria anisotrópica. Como esas áreas vacías cambian diariamente, la solución adoptada fue generar una síntesis de los días de duración del viaje calculando la media de los valores de  $u$  y  $v$ , aprovechando que los vientos son muy constantes en estas latitudes. Tras este proceso, se calcularon los componentes acimut y módulo para cada año ya que son entradas necesarias para el cálculo de coste.

La segunda dificultad es que, aunque todos los SIG implementan opciones para realizar álgebra de mapas, los vectores de viento no pueden ser analizados con estadística ni operaciones convencionales porque están distribuidos en una escala circular. Parte del código desarrollado lo fue para poder operar con este tipo de datos y poder aplicar estadística y operaciones específicas para datos circulares (Fisher, 1995; Mardia, 1972). Estos procedimientos incluyen rutinas para la correcta proyección del acimut en función de la proyección que, en este caso, dada la zona de trabajo

<sup>2</sup> <https://podaac.jpl.nasa.gov/>

y las propiedades que deben analizarse, ha sido una Mercator, cilíndrica y tangente al Ecuador.

### ***Análisis de coste anisotrópico***

El análisis de coste anisotrópico es un procedimiento que estima el coste acumulativo de moverse desde una celda origen determinada a cada una de las otras celdas en un modelo raster. El nuestro caso se analiza el coste de moverse sobre una superficie de fricción determinada por los campos de viento —velocidad y acimut en cada tesela o celda de la superficie. El coste de moverse con un acimut determinado —por ejemplo, hacia el oeste— sobre una celda es el inverso de la velocidad del viento multiplicada por un “factor horizontal” que penaliza progresivamente las desviaciones angulares respecto al acimut del viento. En el caso de los barcos, a igualdad de velocidad del viento, el menor coste es cuando se mueven con viento de popa y el mayor con viento de proa.

Una vez disponibles los valores de acimut y velocidad en cada celda, el análisis de coste anisotrópico se desarrolla definiendo una celda origen (isla de La Gomera) a partir de la cual construye una superficie de coste donde se refleja el coste acumulado mínimo para llegar a cada celda del modelo desde la celda origen. Este proceso supone un análisis completo de las rutas potenciales hasta cada celda desde el origen, descartándolas todas salvo la que presente el valor de coste acumulado mínimo. Este procedimiento se ha realizado para cada año del periodo estudiado.

### ***Síntesis de costes en los puntos de llegada potenciales***

La parte final del análisis ha sido la valoración del coste acumulado para una serie de puntos de llegada alternativos o potenciales. Para ello se ha distribuido una cadena de puntos cercanos a las costas de América y asignado a cada uno el coste acumulado calculado en la fase anterior. La visualización conjunta de estas cadenas (una por año) permite valorar la homogeneidad o heterogeneidad de los resultados y la localización, en su caso, de “pasajes” de bajo coste relativo, donde los barcos hubieran podido llegar con más facilidad en el escenario de los campos de viento de cada año. Todo ello se refleja en los mapas de síntesis finales.

## **Resultados**

### ***Datos de entrada***

Los datos de entrada para el análisis de coste anisotrópico son, como se ha mencionado en el apartado anterior, los modelos raster de velocidad y acimut del viento. La Figura 2 muestra un ejemplo de la distribución de la velocidad del viento en la zona correspondiente al año 1999. El resto de años muestra una limitada variabilidad como se podrá ver en los mapas finales.



Figura 2. Velocidad media del viento en el periodo de estudio del año 1999.

La Figura 3 muestra la distribución de acimutes del viento para el mismo año 1999. Dado que es una variable de distribución circular, hay una discontinuidad en la codificación de tonos entre los 359 y los 0°. Se dispone de estos modelos para cada año del periodo 1999-2005; en ellos, la tierra emergida tiene valor *nodata* y actúa como barrera de resistencia infinita.



Figura 3. Acimutes del viento en el periodo de estudio del año 1999.

### Coste acumulado y ventanas de llegada

Los resultados finales se reflejan mediante mapas de coste acumulado con punto de origen en la isla de La Gomera. La Figura 4 muestra ese modelo para el año 1999. En cada modelo se ha añadido una cadena de puntos cercanos a la costa americana para los cuales se ha transferido el valor de coste de la celda correspondiente y codificado el resultado en colores. El objetivo es ver con más claridad lo que podríamos llamar “ventanas” o zonas de arribada preferentes, con mínimos de coste que suponen una llegada más probable a esa zona de costa si sólo actuara el viento. En la Figura 5 se muestra los resultados para todos los años analizados y la media aritmética.

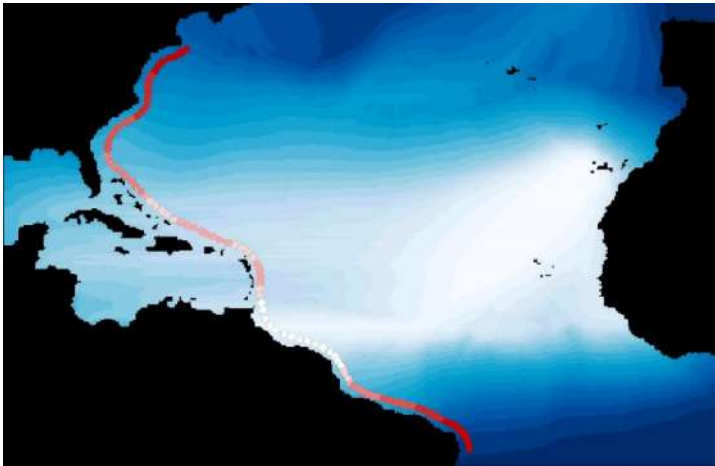


Figura 4. Modelo de coste y “ventanas” de llegada para el año 1999. Las zonas de llegada más probables en función del viento son los círculos de tonos más claros.

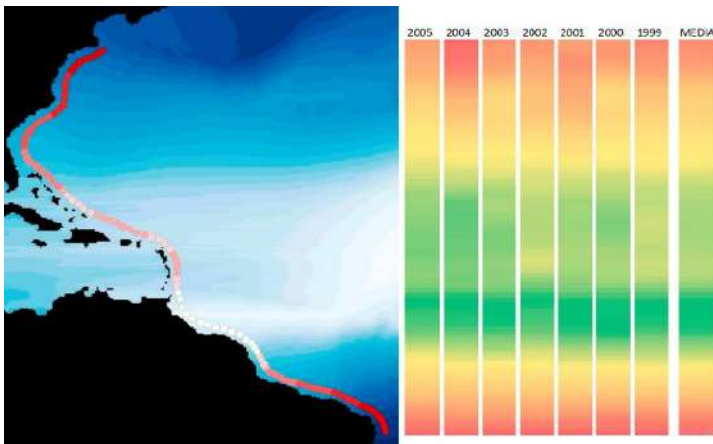
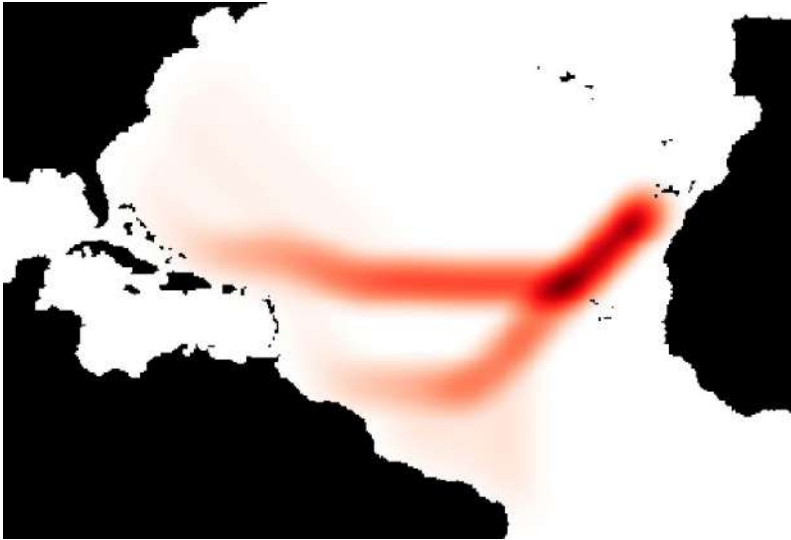


Figura 5. Coste acumulado en diferentes puntos de llegada a la costa americana para los años 1999 a 2005 y su media aritmética. Las zonas en verde son la “ventanas” de llegada de mínimo coste.

Finalmente, se han trazado los caminos de mínimo coste a cada punto de la cadena para reflejar si las rutas siguen un mismo trazado o si hay alternativas diferentes en función del viento; la Figura 6 muestra los resultados.



**Figura 6.** Caminos de coste mínimo para la llegada a las costas de América con la isla de La Gomera como punto de origen.

## Discusión

Los resultados anteriores muestran varias cuestiones de interés para el tema de debate: la ruta de Colón en su viaje a América. La primera es que los caminos posibles están limitados para naves cuya capacidad de viajar con el viento de costado era muy limitada. Si se asume esta limitación, los modelos de coste utilizados asignan al océano los costes que definen las rutas potenciales menos costosas para una embarcación que solo dependiera del viento como fuerza de impulsión (Figura 4). La segunda es que los resultados en los años analizados son bastante homogéneos (Figura 5), se muestran dos “ventanas” de llegada a latitudes diferentes. La más estable se sitúa sobre los 4 y 10° de latitud norte, por encima de la Zona de Convergencia Intertropical y entre la Guayana Francesa y Trinidad y Tobago. La segunda, menos intensa y con algo más de variabilidad interanual, está sobre los 22 a 25° N, en la zona entre Islas Turcas y Caicos y Bahamas. El cálculo explícito de los caminos de mínimo coste hacia América permite elaborar la Figura 6, se observan dos rutas alternativas que podrían haberse seguido en función de las limitaciones de la navegación asumidas. Igualmente se muestra que a la salida de La Gomera, el rumbo más adecuado es el SO y que, llegando a las latitudes de las Islas de Cabo Verde, se presentan dos



alternativas: seguir hacia el SO o virar directamente hacia el oeste y atravesar el Atlántico sin variar este rumbo. En las escasas anotaciones de su diario, Colón mantuvo que su rumbo era hacia el oeste por lo que en este escenario, la expedición colombina debió tomar la primera ruta de estas dos alternativas. No se menciona, sin embargo, las primeras etapas hacia el SO que en el contexto del coste, serían prácticamente obligadas. En cualquier caso, la ruta directa al Oeste, tomada desde La Gomera, aparece improbable en este análisis dada la constancia de vientos al SO y la gran extensión de la zona de calmas o vientos débiles que se extiende al oeste de las Islas Canarias (Figura 2). Cabe mencionar un trágico ejemplo de la segunda ruta hacia las costas de Venezuela: en mayo del año 2006, se encontró una embarcación con una docena de cadáveres a bordo que intentó llegar desde las costas de África a probablemente, las Islas Canarias, un camino de inmigración a Europa relativamente común. Tras quedar a la deriva, fueron arrastrados por el viento y las corrientes durante días, fueron encontrados a unas 80 millas al Sur de Barbados, exactamente en el lugar de llegada de la segunda ruta mostrada en este trabajo.

Queda finalmente por comentar que el punto de llegada de la expedición colombina se ha situado en la isla de San Salvador, en las Bahamas, que con su latitud de 24° N queda exactamente en la salida de la primera ruta mostrada en la Figura 6.

En conclusión, en este trabajo se muestra un análisis que utilizando datos de un escaterómetro portado por el satélite QuikScat y cálculos de coste sobre los campos de viento oceánicos, aporta información sobre la ruta posible de la primera expedición de Colón, apoyando la hipótesis de una ruta latitudinal relativamente alta, pero no la mostrada en la Figura 1 ya que se incluiría una primera etapa con rumbo SO que, con los datos actuales, es prácticamente obligada para los barcos de la época. Este trabajo puede servir como ejemplo de utilidad de este tipo de análisis, ya demostrada en trabajos previos en otros contextos (véase la “Introducción”) aunque, lógicamente, no es una prueba definitiva que pueda resolver la incógnita que se aborda. El principal problema es que los vientos pudieron haber cambiado en cinco siglos y hacer inaplicables los resultados actuales. Es cierto que el régimen general de los alisios en el Atlántico Norte ya es conocido y se usa en la navegación en los siglos siguientes a este viaje y también es conocido que la ruta de vuelta fue, coherentemente con el régimen general de vientos en la zona, por latitudes mucho más elevadas. No puede descartarse sin embargo, variaciones locales o temporales que desplazarían la ruta real de los barcos respecto a las modelizadas en este análisis pero la uniformidad de los años estudiados aporta cierta garantía a la solidez del resultado.

## Bibliografía

- Felicísimo, A.M., Munoz, J. & Gonzalez-Solis, J. (2008). “Ocean Surface Winds Drive Dynamics of Transoceanic Aerial Movements”, en *Plos One*, 3(8). <https://doi.org/10.1371/journal.pone.0002928>

- Fisher, N.I. (1995). *Statistical analysis of circular data* (Repr. 1. paperback ed). Univ. Press.
- Fundación Biblioteca Virtual Miguel de Cervantes (s. f.). *Cristóbal Colón en la Biblioteca Virtual Miguel de Cervantes*. Biblioteca Virtual Miguel de Cervantes. Recuperado 18 de enero de 2020. Recuperado de [http://www.cervantesvirtual.com/portales/cristobal\\_colon/](http://www.cervantesvirtual.com/portales/cristobal_colon/)
- Gonzalez-Solis, J., Felicísimo, A., Fox, J.W., Afanasyev, V., Kolbeinsson, Y. & Munoz, J. (2009). "Influence of sea surface winds on shearwater migration detours", *Marine Ecology Progress Series*, no. 391, pp. 221-230. <https://doi.org/10.3354/meps08128>
- JPL, Jet Propulsion Laboratory, J. (2001). *QuikSCAT Science Data Product. User's Manual. Version 2.2* (D-18053; p. 95). JPL, Jet Propulsion Laboratory. <ftp://ftp.scp.byu.edu/docs/ProjDocs/QSUG4-4.pdf>
- Mardia, K.V. (1972). *Statistics of directional data*, Academic Press.
- Martín López, J. (2019). *El viaje del Descubrimiento. Notas y comentarios al diario de Colón*, 2a ed., Centro Nacional de Información Geográfica. Recuperado de <http://www.ign.es/web/resources/acercaDe/libDigPub/ViajeDescubrimiento.zip>
- McGovern, T.H. (1982). "The Vinland Adventure: A North Atlantic Perspective", *North American Archaeologist*, 2(4). Recuperado de <https://journals.sagepub.com/doi/abs/10.2190/6HN5-AC9T-TWVJ-QDXM>
- Muñoz, J. & Felicísimo, A.M. (2006). Wind as a long distance dispersal vehicle in the Southern Hemisphere. XVII International Botanical Congress Abstracts, XVII International Botanical Congress Abstracts, 74.

# Geospatial data quality (ISO 19157-1): evolve or perish

Calidad de datos geospaciales (ISSO 19157-1): evoluciona o muere

Francisco Javier Ariza López\*  
Pablo Barreira González♦  
Joan Masó Pau•  
Alaitz Zabala Torres■  
Antonio Federico Rodríguez Pascual▲  
Gonzalo Moreno Vergara○  
José Luis García Balboa□

*Received December 15, 2019; accepted January 29, 2020*

## Abstract

Given the circumstance that the process for the revision of the international standard ISO 19157 is currently open, this article presents a critical reflection on its content, application and some challenges posed by the new types of data (e.g. big data, BIM data, etc.), that also have a geospatial component and to which, therefore, this international standard can be applied as well. Proposals are put *forward* going along

- \* Universidad de Jaén, España, e-mail: [fjariza@ujaen.es](mailto:fjariza@ujaen.es).  
ORCID: <https://orcid.org/0000-0002-5204-3630>.
- ♦ Instituto Geográfico Nacional, España, e-mail: [pbarreira@mitma.es](mailto:pbarreira@mitma.es).  
ORCID: <https://orcid.org/0000-0002-1305-4847>.
- CREA, Fac Ciències, UAB, España, e-mail: [joan.maso@uab.cat](mailto:joan.maso@uab.cat).  
ORCID: <http://orcid.org/0000-0002-2983-4629>.
- Universitat Autònoma de Barcelona, España, e-mail: [Alaitz.Zabala@uab.cat](mailto:Alaitz.Zabala@uab.cat).  
ORCID: <https://orcid.org/0000-0002-3931-4221>.
- ▲ Centro Nacional de Información Geográfica, España, e-mail: [afrodriguez@fomento.es](mailto:afrodriguez@fomento.es).  
ORCID: <https://orcid.org/0000-0002-0111-1349>.
- Instituto Geográfico Nacional, España, e-mail: [gmvergara@mitma.es](mailto:gmvergara@mitma.es).  
ORCID: <https://orcid.org/0000-0003-4717-4561>.
- Universidad de Jaén, España, e-mail: [jlbalboa@ujaen.es](mailto:jlbalboa@ujaen.es).  
ORCID: <https://orcid.org/0000-0002-3109-5888>.

three lines of improvement, on the one hand the consideration of new data quality elements and on the other, the reinforcement of the interoperability of this international standard with other standards related to data quality, and finally various improvements (e.g. standardization of evaluation methods, clearly introducing the life cycle, improvement of the definition of metaquality, etc.) of the standard, which come from experience.

Key words: *geospatial data quality, ISO 19157-1, revision.*

## Resumen

Dado que actualmente está abierto el proceso de revisión de la norma internacional ISO 19157-1, este artículo presenta una reflexión crítica sobre su contenido, aplicación y algunos desafíos planteados por los nuevos tipos de datos (p.ej. *big data*, datos BIM, etc.), que también tienen un componente geoespacial y los que, por tanto, esta norma internacional también podría aplicarse. Las propuestas se presentan siguiendo tres líneas, por un lado, la consideración de nuevos elementos de calidad de datos, por otro, el refuerzo de la interoperabilidad de este estándar internacional con otros estándares relacionados con la calidad de datos y, finalmente, varias mejoras más específicas que se basan en la experiencia (p.ej. estandarización de métodos de evaluación, introducción del ciclo de vida del producto de datos, mejora de la definición de la metacalidad, etc.).

Palabras clave: *calidad de datos geoespaciales, ISO 19157-1, revisión.*

## Introduction

The economic relevance of geographic information (GI) and geospatial services has been valued in numerous studies in the last decade (e.g. PIRA, 2000), all of which indicate large returns and added value to society (e.g. reduction of travel times, reduction of emissions of contaminating gases, etc.). For instance, one of the latest studies (αlphaβeta, 2017) estimates that the global consumer's benefits from geospatial services are more than US\$550 billion annually. GI is also recognized as a tool for good governance by many organizations (e.g. United Nations, the World Bank, International Federation of Surveyors, etc.). A very popular article published in *The Economist* has stated that "The world's most valuable resource is no longer oil, but data"<sup>1</sup> and an in-depth study conducted by Hahmann and Burghardt (Dresden University of Technology) in 2012 has shown that at least 78% of the information we usually manage is geospatial information.<sup>2</sup>

<sup>1</sup> <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

<sup>2</sup> <https://www.tandfonline.com/doi/abs/10.1080/13658816.2012.743664>

Taking into account the view that defines GI as a kind of model of the real world (Burroughs, 1986), its quality aspects can be considered a key element because they describe the quality of such a model and its relationship with reality, or more precisely with the universe of discourse, which is the part of the real world that includes everything of interest for a geospatial data product (ISO 19101).

The concept of quality of geographic information, as we know it today, was introduced into the international geospatial agenda only 40 years ago. It was in 1982 when the first studies began under the auspices of the American Congress of Surveying and Mapping. Thus, a proposal for a standard was created (Moellering, 1987) which referred to suitability for use, quality reports and five categories of quality elements (lineage, positional accuracy, accuracy of attributes, logical consistency and completion). It is relevant to indicate that we are still working with these elements of quality, with almost no changes. Around this time the International Cartographic Association promoted a book edited by Stephen C. Guptill and Joel L. Morrison under the title "Elements of Spatial Data Quality" (Guptill & Morrison, 1995). This manual was the catalyst for greater international concern regarding this subject in university and research fields.

The first international standard considering geospatial data quality was probably the DIGital Geographic Exchange STandard (DIGEST),<sup>3</sup> whose first version was issued in 1992 by the Digital Geographic Information Working Group which was established in 1983. This proposed standard, somewhat modified, was adopted as Federal Information Processing Standard 173 by the National Institute of Standards and Technology (NIST, 1994). DIGEST has become a NATO standardization agreement (STANAG 7074). After that, the European experimental standard<sup>4</sup> on geospatial data quality was approved in 1998 by the CEN/TC 287, containing a complete quality model for geographic information, and four years later the first version of ISO 19113:2002 (ISO, 2002) on quality principles was published, followed by another international standard (IS) on quality evaluation procedures (ISO 19114, 2003) and a technical specifications definition on quality measures (ISO/TS 19138) (ISO 2006). Finally, in 2013 the three documents were fused in a unique IS for geospatial data quality (ISO 19157) (ISO 2013), which was amended in 2018 for describing data quality using coverages. Finally, in August 2019 ISO Technical Committee 211 resolved to revise this IS, and today it is currently under revision. Therefore, although there is a solid theoretical framework and an ample bibliography regarding this matter, we believe that geospatial data quality is not fully implemented in production processes yet, probably due to several reasons, among others:

<sup>3</sup> [http://www.dgiwg.org/dgiwg/htm/documents/historical\\_documents.htm](http://www.dgiwg.org/dgiwg/htm/documents/historical_documents.htm)

<sup>4</sup> ENV 12656:1998, Geographic Information - Data description - Quality.

- Geographic information production processes are arduous, expensive and take a long time to be completed, therefore sometimes the approach of producing the best possible product with the available budget is adopted. Taking quality into account is neither cheap nor simple.
- There are very popular solutions which do not pay too much attention to data quality. Some geospatial non-official data providers such as Virtual Globes (e.g. Google Earth) and volunteered geographic information (VGI) (e.g. Open Street Map) are widely used since the cartographic democratization<sup>5</sup>, while their quality is unknown and/or unquantified and in spite of their approach based on “take it as it is”. This situation is balanced out by great usability and quality of service, and sometimes openness and global coverage.
- Not only data producers but also data users and brokers have some kind of “quality immaturity” and there is not much geospatial data quality demand. This situation is probably due, among other causes, to the fact that there are many quality indicators and measures which makes the comparison of different datasets not easy.
- Most GIS tools have not taken data quality into consideration until now.
- The transition from “the best possible quality” to a level of quality that “fits for purpose” has not been completed by geospatial data producers and users.
- Probably the international standard ISO 19157:2013 is relatively recent, and its complete application will take a while in a sector with slow production processes and big inertia.
- On the other hand, new data sources and techniques are invading us with huge amounts of geospatial data which are frequently updated (e.g. satellite imagery, UAV, LIDAR, etc.). A very high updating frequency and the fact of being sensor data seems to leave little room to consider quality assessment.

In this context, it would be also extremely important to have an optimized quality framework based on the best possible geospatial data quality IS, and the final purpose of this article is to contribute to improving the technical content of ISO 19157 as much as possible.

ISO 19157:2013 is a good and complete standard covering all elements of geospatial data quality under a unique and consistent approach. It is better structured than the former ISO 19113, ISO 19114 and ISO/TS 19138 documents, includes sharp UML models, and introduces metaquality and some interesting considerations in the annexes about criteria on how to apply it and how to mix different quality elements.

Nevertheless, ISO 19157:2013 is a complex standard which has been thoroughly revised, and a quick analysis shows some gaps. For instance, in general there are few examples and no examples for important contents (e.g. metaquality), the user has no

<sup>5</sup> [3] <https://www.poynter.org/reporting-editing/2005/holovaty-wins-10000-batten-award-for-chicago-crime-org/>

chance to create its own quality elements, a model for quality reports is missing, raster and image quality is not taken enough into account, several measures have definition problems, etc.

The ISO/TC211 has already held some meetings in which this issue has been discussed (e.g. sessions in Toulouse (France) and Omiya (Japan)). Recently (Malta, January 2020) Eurogeographics has organized a meeting where the process as well as suggestions received by TC211 from the 28 experts who are collaborating in the review have been presented.

This paper presents some ideas coming from experts nominated by the Spanish technical committee 148 of UNE (previously known as AENOR). Our objective is multiple, on the one hand to indicate the greatest weaknesses and deficiencies in the data quality model of ISO 19157 and its application, and on the other hand to indicate a set of current opportunities and challenges in relation to data that require location and the quality of this data. Finally, we indicate a set of improvements that should be adopted in the new version of ISO 19157 if we really want it to be applied in an extensive, intense and correct way. So this article aims to examine in-depth those and other problems of this IS and try to propose solutions to improve it in the complex arena of the current geographic data ecosystem. In the next section a critical analysis of ISO 19157:2013 is presented, trying to identify weak points and areas of potential improvement and proposing solutions, new approaches and future lines of progress. Section 4 is devoted to the new challenges facing geospatial quality due to the new types of data, like the aforementioned ones, which probably requires a new approach based on some kind of “quick and big” quality information. As a consequence of sections 3 and 4, a set of proposals for the revision of ISO 19157:2013 is presented in section 5, with the aim of enhancing the applicability and usefulness of the standard. In some points a quite complete idea about how to update the standard is provided, while in other cases just some ideas and concepts are included. Finally in section 6, some conclusions summarizing the contributions are outlined.

### **Critical analysis of International Standard ISO 19157 and its application**

Quality should be used by producers and users. One natural way is to include data quality exigencies in the specifications. Another way is to include the data quality results in the metadata, and for this reason we will start this section talking about specifications and metadata. Ariza-López and Rodríguez-Pascual (2018) presented a small study (in April 2018), consulting the information available on 19 websites of National Mapping Agencies of the American continent and they found that only 11 times ( $\approx 58\%$ ) metadata of the available data and services were published, that only one organization (5.3%) publishes quality information of its data beyond the lineage and that on only 6 occasions (32%), was descriptive information of the available geographic data products published, although on 4 occasions it was labeled

“Technical Documentation” and in one case presented as specifications. In addition, this is not unique to the American continent. A simple review of the sections dedicated to quality in the implementation rules of INSPIRE also give, in many cases, a sense of insufficiency. These simple studies indicate that there are problems with the inclusion of quality aspects in the specifications and metadata of geospatial data products (ISO 19131 and ISO 19115-1, respectively), or rather, that there are problems with the use of the quality framework proposed by ISO 19157.

The original ISO 19115 metadata standard is the base for the metadata records included in Spatial Data Infrastructures and clearinghouses catalogues that collect descriptions of geospatial data products. As data quality elements are an integral part of the metadata model we should expect that most producers provide a comprehensive description of the different components of the data quality in their metadata records, but this is not the common case. An analysis of the metadata harvested by an old version of the GEOSS<sup>6</sup> portal reveals that most datasets include no data quality indicators and if they do they rarely go beyond positional accuracy (Zabala *et al.* 2013).

ISO 19115 has been always blamed for being long and complex. Indeed, the standard is very comprehensive, with more than a hundred properties that can potentially be populated. Its complexity lies in the difficulty in separating the dataset description into so many properties, making the creation of a metadata record tedious and time-consuming. Despite these difficulties, individual properties are well defined and relatively easy to understand and populate if the information is at hand, with one exception: data quality. Actually, a quality model is included in ISO 19115 but this document alone does not provide enough details. ISO 19115:2003 mentions the 15 subclasses and some properties needed to specify quality measures but it does not include any concrete measures. We need to refer to ISO 19157:2013 to discover a list of about 80 quality measures, each one with the methodology and statistical analysis needed to extract the result. In our experience, the ISO metadata standard is very popular among practitioners but they rarely have access to ISO 19157, and so data quality remains relatively unknown. This could be remedied by metadata tools providing the necessary alternatives and information to the user, but the current metadata editors do not develop data quality in the necessary level of detail. Unwittingly, TC211 might have made the situation worse by removing the data quality element from ISO 19115-1:2014 and delegating its definition to ISO 19157. There is a need to make ISO 19157 known among the community as well as to metadata editor developers.

ISO 19157 organizes the quality measures into six classes representing mainly the components of the information (spatial, thematic, temporal, logical, etc.) that are

<sup>6</sup> <https://www.earthobservations.org/geoss.php>



subdivided into 15 subclasses (data quality elements) which define what is measured (omissions, commissions, absolute accuracy, topological consistency, etc.). Conceptually, there is no reason why this could not be extended to other aspects of quality (e.g. redundancy, quality of metadata, quality of service). However, the design rules used to create the subclasses by generalization make them difficult to extend once they are encoded in a data format such as XML. This limits the extensibility to each revision of the standard. In fact, included in ISO 19157 was a new class for usability as an effort to extend the scope of the IS beyond the producer perspective into the user perspective. However, usability is described in a confused manner as an aggregation of producer's measurements and conformance to requirements, which might not be the best approach. Instead, a new model for user-created quality in the form of feedback could be better.

The measures included in ISO 19138:2006 Annex D were collected among the ones commonly used by mapping agencies. The last revision of the list, now conforming annex D of ISO 19157:2013, included a few new additions. Meanwhile geospatial information has become popular in other sectors, making the current scope quite limited. One example of a possible extension of scope is the big data world, where modeling is regularly used to investigate and predict Earth variables at regional or global level. These models are validated in different ways that are not necessary compatible with the current list of measures. The need for other quality measures is also observed in the emerging crowdsourcing, citizen science or other non-authoritative data products. The metadata records are also under scrutiny and new measures are proposed to evaluate the quality and completeness of their descriptions. Another gap is introduced by the very same ISO 19157 when providing an example for metaquality. The example used in Annex E.3 introduces a confidence quality measure called "safety factor". Unfortunately, ISO 19157 Annex D does not contain the description of this measure, obfuscating the usefulness of the example and leaving the reader with incomplete information on how to report metaquality. Ideally, ISO19157 should be encoded as an extendable list of quality measures. Annex H of ISO 19157 provides some clues on how an extensible catalogue of measures should be built as a registry. However, the annex ignores the possibility of making the measure available as a dynamic ontology which includes the current Annex D as well as other measures from other standards and best practices (e.g. ISO 8000, ISO/IEC 25012, ISO/TR 21707, GUM<sup>7</sup>, VIM<sup>8</sup> guides). Ideally, this ontology would link to actual examples on how to use the measures in real live cases. Even if new quality measures seem necessary, we have to be careful in increasing the list too much: the more measures we have the more difficult it becomes to compare the quality of datasets with different origins.

<sup>7</sup> [https://www.bipm.org/utis/common/documents/jcgm/JCGM\\_100\\_2008\\_E.pdf](https://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf)

<sup>8</sup> [https://www.bipm.org/utis/common/documents/jcgm/JCGM\\_200\\_2012.pdf](https://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2012.pdf)

Data quality measures are based on applied statistics. The same statistic can be applied to different measures used in different components and with different thresholds or parameters. This is particularly true for assessing uncertainties as described in ISO 19157 Annex G. The same statistic is applied in several quality measures of annex D, making the selection of the right quality measure unnecessarily confusing. In our opinion current measures do not make enough effort in separating the domain of numbers (e.g. the individual uncertainties of each feature) from the statistical expression or mathematical metrics that will be applied to the domain in order to summarize the quality into a measurable indicator (e.g. standard deviation), adding an unnecessary level of confusion. By separating the domain from the metrics, applications and metadata editors could focus on the latter while experts and practitioners could concentrate on preparing and specifying the domain.

At present, some users as well as producers consider that ISO 19157 is too focused on quality evaluation. The relation of the data quality measurements with the life cycle management of the data product is introduced in Annex B, enumerating which stages of the life cycle quality evaluation should be performed and this could be reported as part of the quality evaluation description as metaquality. More emphasis should be placed on this topic. Only a mature and easy to use quality standard that considers all steps in the life cycle of a product will result in a certification of geospatial data products incorporated into industrial procedures.

Another problem presented by this model, as formulated and applied, is the grain. The model presents problems working at the instance level. This is problematic in the case of aiming to ensure traceability over instances and also to derive global quality values from data products generated by aggregation of instances from various sources (Ariza-López & Rodríguez-Pascual, 2018).

Finally, the revision and improvement of this international standard should be considered in a broad sense that also includes aspects of the quality of geoservices and producers. Let us think that many data are offered through geoservices and that the quality of the data producer, as an organization, affects the quality of your product (the data). Ariza-López & Rodríguez-Pascual (2018) present an analysis with this more global perspective of the challenges of quality in geospatial data.

### **New types of data: the challenges**

Since The Economist (2017) published a story titled, "The world's most valuable resource is no longer oil, but data", the sentence "data is the new oil" has become a usual way to indicate the value of data. In this way, the so-called data economy (Wikipedia, 2020) is a main concern of world and regional institutions (e.g. UN-EAPD 2019 and EU 2020). Geospatial data and these types of data indicated below, and many other types of data, are part of this data economy. From our point of view, ISO 19157 is focused on geospatial data but with a classical producer

perspective. For this reason, and thinking about the data economy, we consider that new perspectives must be included in the revision of this international standard in order to guarantee a higher level of application and convergence between different highly related data-types, at least considering the following types:

- Big data are “high-volume, high-velocity and/or high-variety information assets which demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation” (Gartner, 2020). There is no specific type of big data. The relationship of big data with geospatial data is clear (Eldawy and Mokbel, 2013), many of the geospatial data (e.g. remote sensing images) are true sets of big data. The relationship is so obvious that many authors talk about geospatial big data (Lee and Kang, 2015; Robinson *et al.*, 2017). Quality issues are pointed out as an important challenge in geospatial big data (Robinson *et al.*, 2017; Lee and Kang 2015; Li *et al.*, 2016).
- Building information model (BIM) data sets are model-based geometric information, enriched thematically, semantically and relationally which, managed with the right software tools, allow a more efficient management of buildings and facilities (Ariza-López *et al.*, 2019). BIM data are very similar to geospatial data and also deal with geospatial data because they must be integrated into a geographical framework (the actual location of the building) environment (the surrounding geographical-topographic reality), and also collect the presence, dimensions, positions and exact attributes of the elements of interest. BIM data are directly related to 3D geospatial city data. Puyan *et al.* 2017 highlight the interest and importance of the quality of BIM data for facility management purposes and present examples of errors in BIM data that are very similar to those that occur in geospatial data. In this way, Puyan *et al.* (2017) show the close links between geospatial data and BIM data, Song *et al.* (2017) indicate the need and benefits of the integration of BIM and GIS and Ariza-López *et al.* (2019) develop BIM data quality controls based on geospatial data quality elements.
- Volunteered geographic information (VGI) (Goodchild, 2007) is a kind of participative/collaborative geospatial data where citizens, often untrained and regardless of their expertise and background, create geographic information. The quality of VGI has been a hot topic from the beginning of this trend, and there are a lot of papers dealing with this topic. The need for specific data quality elements, metrics and methods is clearly pointed out in Gusminia *et al.* (2017), Degrossi *et al.* (2018), Senaratne (2017).
- Statistical data are those produced by statistical agencies. Official statistics data are grouped into topics (e.g. economy, population, international trade, etc.) (SDMX 2009). These data have different levels of aggregation ranging from microdata (e.g. the income of a person) to an added value for a country (e.g. gross domestic product). There is a clear confluence between statistical data and geospatial data so that many statistical organizations (e.g. INEGI in Mexico, IECA in Andalucía (Spain), Eurostat in the European Union, etc.), are producing

geospatially enabled-statistical data and micro data. The current trend is that all statistical data have a location. There are already grids with geospatialized statistical data for some regions (Eurostat, 2020). There is a conceptual framework in place to include the geographical component (UN-ISGI, 2018; Moström *et al.*, 2019). Data quality and quality management are relevant issues for official statistics (UN-SD, 2019).

- Earth observation data and images. These are clearly geospatial data. They are also big data. Unfortunately, ISO 19157 is difficult to apply directly to this type of data. With the aim of process simulations in the Earth monitoring scope (e.g. climate change), there is a great concern about the quality of images and derived products (e.g. essential climate variables)<sup>9</sup>. The Committee on Earth Observation Satellites developed the project Quality Assurance Framework for Earth Observation (QA4EO) where a series of key guides provide a quality assurance framework for images and related processes, but ISO 19157 is not applied.
- Geolinked data. Linked Data are defined as “structured data which is interlinked with other data so it becomes more useful through semantic queries; It builds upon standard Web technologies such as HTTP, RDF and URIs, but rather than using them to serve web pages only for human readers, it extends them to share information in a way that can be read automatically by computers” (Wikipedia, 2020). Geolinked data or Linked Geodata<sup>10</sup> consists of enriching the web with geospatial data. Geospatial data are linked to other data and any other type of data can be linked to a position given by geospatial data. Some national mapping agencies (e.g. OS in UK, IGN in Spain) already offer linked geospatial data. The relevance of linked geospatial data is clearly indicated in López-Pellicer *et al.*, (2011). Working with the graph established by the links adds a degree of complexity to aspects of data quality. Zaveri *et al.* (2014) carry out an exhaustive work of compilation and organization of the numerous dimensions and measures that can be applied to evaluating the quality of the linked data. There have also been initiatives focused on quality dimensions and measures for geospatial linked data (GeoKnow, 2012).
- IoT data are data produced by a “system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers (UIDs) and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction” (Wikipedia, 2020). IoT systems are related to digital twins and mirror spaces. Lot data refer to data derived from sensors (e.g., humidity, rain, heat stroke, temperature, etc.) that monitor real-world situations, and actuators (e.g. stepper motors, control valves, switches, etc.) that can modify real-world situations. Location powers the analytic capacity of IoT data-based systems. As indicated by Karkouch *et al.* (2016), data quality is crucial to gaining user engagement and acceptance of the IoT paradigm and services. There are several studies on the

<sup>9</sup> <http://www.qa4ecv.eu>

<sup>10</sup> <http://linkedgeodata.org>

quality of IoT systems (Ahmed *et al.*, 2019), but the paper by Karkouch *et al.* (2016) focuses on the quality of the IoT data. In this paper, some of the established data quality dimensions are assailable to categories of geospatial data quality elements or geospatial data quality elements. It also points out several problems (e.g. outlier management) that are very similar to those inherent to management in the geospatial data domain.

### **Proposals for the revision of ISO 19157-1**

As explained in previous sections, we can assert that there is a need for major changes which comes from: i) the adoption of new perspectives in terms of data, ii) the need for a greater interoperability with other ISO international standards and, iii) the experience acquired in the application of ISO 19157.

#### ***Candidate data quality elements***

In order to adopt new perspectives new data quality elements need to be defined, and this opportunity must be opened up as well (as in ISO 19113). Some examples of new data quality elements can be trust dimensions for open and linked data (Zaveri *et al.*, 2013), the quality of free text for descriptive texts included in metadata records (Ureña-Cámara *et al.*, 2019) and quality elements proposed for images, photogrammetric flights or other spatial gridded products (Ariza-López, 2013).

Additionally to this approach, we can consider that a DQ-Element can be split into sub-elements, however, this makes the implementation and extensibility of ISO 19157 more complex and it does not add much value. Nevertheless, creating an attribute in a DQ-element to include the quality subclass name and defining a code list makes it easier without losing any functionality. This will imply the need for new “quality categories” or dimensions, new measures, and so on.

The current version of ISO 19157 is more focused on vector data, leaving gridded data as something not frequently used. This fact makes the rule difficult to apply to this kind of geospatial data product. There is a need to place the focus on other data types (e.g. LIDAR, BIM, geo-linked data, etc.).

In order to overcome this obstacle several changes need to be undertaken so that ISO 19157 can be applied to raster data correctly, covering all its aspects.

- I Data compression: data represented in a grid is always a consequence of generalization. In this sense, raster formats tend to compress data in order to represent variables in a more discrete way, when most times they are continuous. To cover this issue a new DQ\_Element called Compression will be very useful.
- II Many raster products are obtained from satellite images or photogrammetric flights. These scenes present some deficiencies most of the time: pixels failures,

rows/columns with no data, the presence of clouds or other weather conditions that hinder the extraction of information, shadows, etc. This is the case not only for specifications but also for DQ assessment of a product.

- III Representing physical variables through raster data: many remote sensing products which are geospatial data locate physical variables (such as temperature, pressure...) in space. Since these variables are most times completely different, the IS should leave open the chance of proposing specific DQ measures and elements for this kind of information.
- IV Orthophotos are another usual example of gridded data that should be able to have its own DQ elements and measures. In this sense, measures that cover the percentage of censored data in the orthophoto or the mosaic, mosaic cuts, or the degree of geometric and radiometric matching among adjacent orthophotos should be applied to this type of product.
- V Photogrammetric flight planning assessment, as part of the lifecycle of a product (e.g. orthophoto), is of relevance for producers and needs to be considered in this sense as well (Ariza-López, 2013).
- VI Measures for thematic classification: in this sense there are several spatial metrics which come from the world of landscape metrics that can present some clues to understanding the degree of accuracy in terms of thematic classification. These metrics describe the shape, distance between patches of the same category, fragmentation of the category and so on, that could be applied to the classification obtained and the real world. The comparison of these metrics could contribute to a better DQ assessment. Also, alternatives to the confusion matrix and derived parameters (e.g. Kappa index) should be considered as they are present in remote sensing studies (e.g. ice Coefficient, Relative bias, among others (Padilla *et al.*, 2015)).
- VII Data derived from Lidar is becoming one of the main sources in geospatial products. ISO 19157 should cover this specific type of data as well as others, not only in terms of positional accuracy but also thematic accuracy and redundancy, which plays a key role in this kind of data.

Some examples of possible new elements taken from official documents are presented in Tables 1.a, b, c. Thus, the possibility of developing new elements should be left open. This opportunity existed in ISO 19113 but was eliminated in ISO 19157:2013, which goes against the backwards compatibility that should be sought with the new versions.

**Table 1a**  
**Example of a new quality element (Case 1)**

---

<i>Name:</i>	Radiometric discontinuity
<i>Definition:</i>	Closeness of the radiometric values of homologous pixels of two images in a common area, is defined as: “S’il y a effectivement eu un mosaïquage, la classe de précision sera exprimée par la différence de valeur radiométrique par canal tolérée sur les raccords entre images ne correspondant pas à un linéament, divisée par la radiométrie maximale de l’image et exprimée sous forme de pourcentage”
<i>Justification:</i>	When creating mosaics of images, the presence of radiometric discontinuities is a common circumstance derived from many different circumstances and reasons. This situation is undesirable. The presence of a radiometric discontinuity is an aesthetic and an exploitation problem
<i>Source:</i>	Ministère de l’équipement, des transports, du logement, du tourisme et de la mer. Arrêté du 16 septembre 2003 portant sur les classes de précision applicables aux catégories de travaux topographiques réalisés par l’Etat, les collectivités locales et leurs établissements publics ou exécutés pour leur compte

---

**Table 1b**  
**Example of a new quality element (Case 2)**

---

<i>Name:</i>	Integrity
<i>Definition:</i>	Is defined for aeronautical data as: a degree of assurance that an aeronautical data and its value has not been lost or altered since the data origination or authorized amendment
<i>Justification:</i>	This aspect is of great relevance for nautical and airspace security. The data quality model of the International Civil Aviation Organization considers this issue. This issue can be of interest for other uses and purposes (e.g. homeland security, army, medical urgencies, fiscal data, etc.)
<i>Source:</i>	ICAO (2010). ANNEX 15 to the Convention on International Civil Aviation. Aeronautical Information Services

---

**Table 1c**  
**Example of a new quality element (Case 3)**

---

<i>Name:</i>	Geometric fidelity
<i>Definition:</i>	Geometric fidelity is the measure defined as: that any real world alignment or shape, when viewed at the source survey scale, must be accurately reflected in the data to the required specification
<i>Justification:</i>	Real world objects (e.g. buildings) can be registered in a dataset without their exact and true relationships with their surroundings. It is necessary to have an assessment of the number of such types of objects in the dataset This information is relevant for the producer (quality of his processes or supplies) and for users
<i>Source:</i>	OS (2007). TOPO-96 Data quality. Ordnance survey

---

### ***Interoperability with other International Standards***

In order to achieve greater interoperability with other ISO international standards or documents (e.g. ISO 8000, ISO/IEC 25012, ISO/TR 21707) and third-party documents (e.g. VIM, GUM, etc.), which are interrelated to ISO 19157, means the necessity of a great effort of coordination in concepts, terms, perspectives, etc. Along this line, a well-defined ontology is missing for the quality of geospatial data. This ontology should be compatible with other ontologies of other application domains (e.g. dqv<sup>11</sup> of the W3C), and allow greater interoperability between different types of data.

Clear examples are the terms accuracy and uncertainty, which are not appropriately defined within ISO 19157, implying the need for a complete revision of many of the terms and measures proposed in it. Another example entails quantitative quality assessment, where a distinction between estimation and control is needed. Estimation means the precise determination of a parameter value and its confidence interval. Control (quality control) means taking a decision about the acceptance or rejection of a previously-stated hypothesis within a statistical framework of assumed risk.

Another issue entails the relationship between ISO 19157 and ISO 19131, for which to many of those that have applied both, the differences are not always completely clear. In the data specifications (according to ISO 19131) of many national mapping agencies there are usually no quality specifications on the data, and in some cases evaluation results appear when they are actually metadata. This is completely incorrect. More clarification is required, maybe an informative annex can help.

### ***Experience applying ISO 19157***

Finally, the experience acquired in applying ISO 19157 shows us the bulk of the changes that we consider is relevant to introduce in its revision.

### ***Distinction between the quality of a dataset and the uncertainty in individual attributes of a feature***

The scope element allows narrowing the scope of data quality to a particular attribute but unfortunately it is rarely implemented. We believe this is partly because it is not clearly explained in the current version of ISO 19157. In addition, in the world of Big Data and unstructured records where the abundance/redundancy of information about a single property can be used as a quality test, the distinction mentioned above is becoming more important. This is particularly true in case as of citizen science of

<sup>11</sup> <https://www.w3.org/TR/vocab-dqv/>



VGI, where redundancy in acquisitions and expert validations are used as quality indicators. A paradigmatic example is the identification of pheno-phases in a certain location and time, a situation in which redundant and coherent measures increase the veracity of each individual measurement that need to be quantified. Current measures in ISO 19157 are still applicable but they require a better definition of the individual inputs to assess.

*The quality assessment methods and their standardization*

The use of the same, well specified, data quality elements, scopes and measures does not ensure the compatibility between the results of two assessments if different quality assessment methods were applied. Even if these methods have place in the

**Table 2**  
**Example of a template for methods standardization**

<i>Line</i>	<i>Component</i>	<i>Description</i>
1	Method identifier	Unique identifier within a namespace
2	Name	Name is the name of the method
3	Purpose	A description of the purpose of the quality assessment method
4	Method type	Indication of the method type (direct or indirect)
5	Result type	Indication of the result type (quality estimation or quality control)
6	Description	A general description of the method
7	Source	Identification of an explaining source(s), if exists
8	Detailed description	Description of the assessment
8.1	Full inspection based	Explanation of the full inspection process
8.2	Sampling based	Explanation of the sampling based process
8.2.1	Sample Scheme	Explanation of the sample scheme
8.2.2	Sample Size	Calculation of the sample size
8.2.3	Sample collection	Explanation of the sample collection process
8.3	Resources	Description of resources to be used
8.3.1	Instrumental	Specifications about instruments
8.3.2	Human	Specifications of skills
8	Measures	Identification of the standardized measures to be used
10	Procedure	A complete identification and explanation of the steps of the evaluation method
10.x	Step x	Explanation of each step of the procedure

data model we consider that quality assessment methods should be better and they require a very similar treatment to measures (standardized measures of annex D of ISO 19157). This implies the definition of a list of components and also a catalog of standardized methods. Methods must be well defined in order to be understandable, replicable and their results interoperable. A standardized quality assessment method must offer an unambiguous definition, and a structure of the methods extended in a way that is presented in Table 2. Also, the methods must be established in the data product specifications (following ISO 19131), because product specifications must include quality requirements and how (the method by which) their achievement is measured.

In the current ISO model, conformity is a binary value that should also be associated with a methodology. There is a need for presenting a relationship with the data quality assessment methods are the concepts of conformity level and the quality control decision. The first refers to the minimum good quality level (e.g. at least 90%) or the maximum bad quality level (e.g. at most 5%) that the user is willing to accept. A standardized measure, the units of the measurement (e.g. m or mm) and the value are the key elements for defining a conformity level. The second refers to how an acceptance/rejection decision is taken in a quality control. This decision must be taken by comparing a result of a standardized quality assessment method versus a conformity level by means of a given rule where producers' and users' risk are previously established. The conformity level must be established in the data product specifications (following ISO 19131) as well as the quality control decision rules.

### *Metaquality and its use*

There are very scarce references to the actual use of metaquality. One of the problems is that there is a lack of clear examples. An informative annex added to the new version of the IS could be adequate. In this sense, we would like to mention the use of metaquality developed in UNE 148002:2016 (UNE 2016) (see Figure 1), where metaquality elements (confidence, homogeneity and representativity) are more detailed. In order to cover several approaches to studying metaquality; for example, they divide the representativity into several topics: spatial, temporal, thematic, participative and global. They introduce the concepts of qualitative and quantitative confidence as well, which could also be interesting for the revision of ISO 19157.

### *Report and its standardization*

ISO 19157 indicates that in order to provide more details than those reported as metadata, a standalone quality report may additionally be created. The structure for this standalone quality report remains unspecified. This situation leads to reports being very different and resulting in a clear problem in terms of interoperability, for

“In this standard confidence on the results of a PQC process is determined by two complementary aspects:

- Qualitative. The rigorous application of the methods is the main guarantee of trust from a qualitative perspective. This aspect must be ensured by the participation of experts in the quality of geographic information in work teams and by the requirements stated in section 9 of this standard.
- Quantitative. Effective enforcement of the following aspects is the basis of trust from the quantitative perspective: sample size, randomness, independence of the control process and greater accuracy of the reference SDS. These aspects should be ensured by compliance with the requirements of sections 8 and 9 of this standard.

In this standard homogeneity of the results of a PQC is determined by:

- Production of the controlled SDS. The homogeneity of the controlled SDS is beyond the scope of this standard; however it should be noted that it can be a critical aspect in the case of SDS where numerous persons or organizations have intervened, where concur diverse backgrounds, knowledge, skills, etc., or different work methodologies are applied (e.g. OpenStreetMap).
- The extension of the control process. For PQC processes dilated in space or in time appropriate quality management measures shall be taken in order to ensure homogeneity of the PQC process at all times. Key elements to ensure homogeneity are, among others: documented procedures, the establishment of standards in education and training of personnel involved including verification mechanisms to ensure consistent processes, etc.

In this standard representativity of the result of a PQC should be evaluated from multiple perspectives. Since the assessment is based on sampling, the representativity should be:

- Considered in relation to the following aspects:
  - *Space. The spatial representativeness of the sample by its effective spatial distribution compared to the actual spatial distribution of the population.*
  - *Time. The temporal representativeness of the sample by its effective temporal distribution compared to the actual temporal distribution of the population.*
  - *Theme. The thematic representativeness of the sample by its effective thematic distribution of categories and attributes compared to the actual thematic distribution of the population.*
  - *Participation. In the case of studies with the participation of various organizations (e.g. national series) or individuals (e.g. OpenStreetMap), has the same sense as previous cases but related to the participation issue in the population.*
  - *Global. Refers to global representativeness as an interpretation of all the partial representativeness given above being considered in a specific PQC.*
- Evaluated with appropriate techniques, some techniques are applicable:
  - Visual comparison of histograms and distribution functions of the sample and the population.
  - *Adherence tests between the curves representing the distribution functions of the sample and the population (e.g. by means of the Kolmogorov-Smirnov test for continuous cases and Chi2 for discrete cases)”.*

Figure 1. Translation of section 10 of the Spanish UNE 148002:2016 standard.

data quality comparison and for data certification. To deal with this problem a well-defined and flexible quality report is needed.

### *Quality along the lifecycle.*

Geospatial data quality, as presented in ISO 19157, is not understood as something that must take place along a product lifecycle. This idea (data quality along the lifecycle, in a process) must be presented and explained in the early beginning of the IS since it is one of the keys to understanding the concept of quality related to geospatial data. In our experience, most of those who apply the 19157 standard, or who wish to apply it, only focus on the quality of the final products. They do not anticipate that ISO 19157 should be applied throughout the entire product life cycle. We believe that this problem has several causes, being one of them that the IS does not sufficiently explain it and does not present example about this. In addition, ISO 19157 scope needs to be widened to include such as photogrammetric and topographic production processes. Currently, it difficult for specialists in these fields to apply the standard.

Therefore, following this line, at least two additions are required in the revision of ISO 19157: on the one hand the idea of geospatial data products' life cycle, and on the other hand a model of how to apply quality throughout this life cycle. For the first addition there are several proposals for geospatial data product lifecycles, for instance the United States Geological Survey<sup>12</sup> considers five phases (plan, acquire, process, analyze and preserve and publish/share). A proposal for quality throughout the lifecycle is that of the *Arbeitsgemeinschaft der Vermessungsverwaltungen* (AdV, 2002) which was adopted and improved by the *Instituto de Estadística y Cartografía de Andalucía* (IECA, 2011). As an example, a new version of this model is presented in Figure 2, which includes the Plan Do Check Act perspective (Deming's Cycle) using the stages of the USGS lifecycle (plan, acquire, process, analyze, preserve, publish) and the quality assessments proposed by the AdV and improved by IECA (2011). The Cross-cutting elements (describe, manage quality and Backup Secure) of the USGS data lifecycle are also labeled. This model includes the main quality management functions (quality control, quality improvement, quality assurance and quality deployment):

- Q1: Assessment of the basic model against general and strategic guidelines.
- Q2: Assessment of the application model and specifications against the basic model.
- Q3: Assessment of the application model and specifications against specific requirements.
- Q4: Assessment of the data product against its logical consistency rules.

<sup>12</sup> <https://www.usgs.gov/products/data-and-tools/data-management/data-lifecycle>

- Q5: Assessment of the data product against the real world.
- Q6: Assessment of the data product performance for analysis and uses.
- Q7: Assessment of the product continuous improvement process.

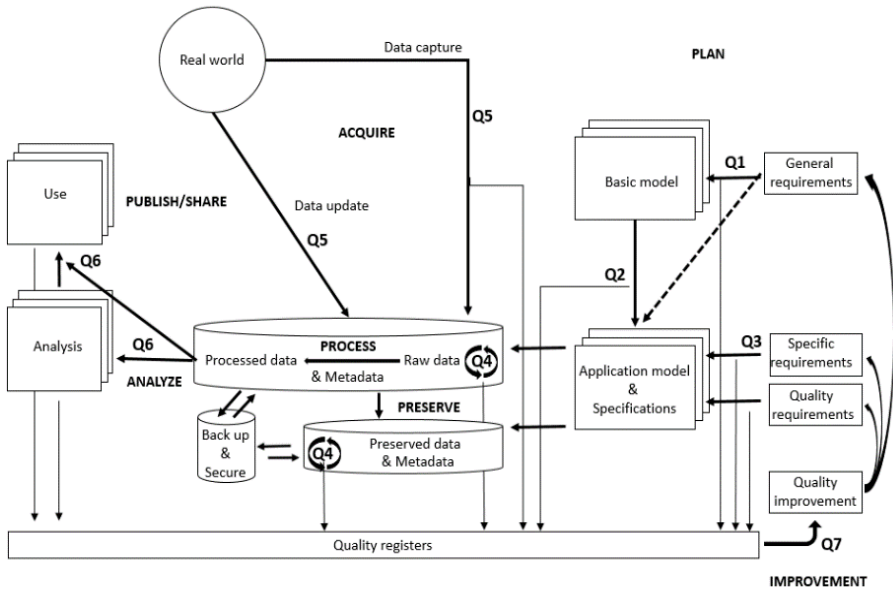


Figure 2. Example model for quality throughout the cifecycle.

*Unify UncertML and ISO 19157 to improve Annex D.*

We identified some conceptual overlapping in the list of quality measures presented in annex D. This is due to the use of similar statistical concepts applied to different quality elements but using the same measure from the statistical point of view. A simplification in the way the list of measures in Annex D is presented is necessary in order to increase comprehension. We propose a way to achieve this without losing anything, using already existing alternatives: UncertML 2.0<sup>13</sup> and QualityML.<sup>14</sup>

UncertML 2.0 provides a semantic description of statistics that can be used to compute uncertainties. A priori, it seemed simple to associate those to ISO 19157 quality measures but in practice it was not. By extending the concepts in UncertML

<sup>13</sup> <https://web.archive.org/web/20161029215725/http://www.uncertml.org/>

<sup>14</sup> <http://www.qualityml.org/>

a strategy to better normalize measures of ISO 19157 by making the list of measures more compact was achieved.

In essence, the concept of statistics is extended to include other quality metrics used to compute the result of each quality measure value when applied to a certain domain. QualityML provides a matrix of the combinations of indicators, measurements, domains and metrics most commonly used. The main idea behind this structure is to unlink measures, domains and metrics descriptions in order to maximize the generalization of descriptions and increase coherence among several measures using the same metrics (even with different domains), or several quality indicators using the same measures.

In fact, ISO 19157 already introduces the concept of data quality basic measure to avoid the repetitive definition of the same concept. There are data quality measures that have certain commonalities. Two principle categories of *data quality basic measures* are listed in the Annex: i) the uncertainty-related data quality basic measures (e.g. the LE50 basic measure used in linear error probability [id 33], time accuracy at 50% significance level [id. 55] or attribute value uncertainty at 50% significance level [id. 69]), and ii) the counting-related data quality basic measures which are based on the concept of counting errors or correct items.

QualityML goes one step beyond this generalization effort in the ISO 19157 basic measures, and groups one describing the same metric but with different parameters. For example, all the measures regarding "half length of the interval" are grouped in a single general metric called Half-lengthConfidenceInterval, which includes a parameter to describe the confidence level (or probability) of the true value being between the lower and the upper limit. Level has to be in the range [0,1]. This QualityML metric (Half-lengthConfidenceInterval) includes several ISO 19157 basic measures such as LE50 (and thus measures with id. 33, 55 and 69) but also LE68.3 (used in standard linear error, id. 34), time accuracy at 68,3% significance level (id. 54) and attribute value uncertainty at 68.3% significance level (id. 68), LE90, LE95 etc. All these ISO 19157 measures are grouped in a single QualityML with a single parameter "level" to identify the significance level. The advantage of this generalization is not only the increase of coherence in the quality measures and metrics description, but also the possibility of describing any other confidence level interval in a standardized way.

This is done in QualityML not only for uncertainty-related data quality basic measures describing one dimensional random variables (Z, using "Half-length Confidence Interval", examples above) but also for uncertainty-related data quality basic measures describing two dimensional variables and for counting-related data quality basic measures. More details about this approach can be found in Zabala & Maso (2016).

### *User feedback in the broader sense.*

Gray knowledge about the data can be as useful as quality indicators to assess "fit for purpose". A structured way to include experiences of users of products that report applications not initially foreseen as well as problems and workarounds should be included in this IS (or in another document related to it). Users can produce valuable and complementary metadata about resources structured in feedback about each resource they are interested in, have used, etc. There are plenty of elements that can be included in a feedback item about a resource (or a group of them) such as ratings, comments, usage, related publications, additional lineage steps, quality elements or significant events description. All these user feedback metadata elements complement producer metadata and add value to the dataset descriptions. It also helps increasing users' engagement as they can see a real opportunity to create a community and establish social links on a geospatial portal around the datasets they are interested in. The data producers may also take advantage of this situation, being able to respond to users' demands in creating new versions of the resources or answering their concerns as new feedback items (related to the previous ones).

The work developed on previous projects such as GeoViQua,<sup>15</sup> CHARMe<sup>16</sup> and Melodies<sup>17</sup> and on the OGC Geospatial User Feedback (GUF) working group, led to the approval of the OGC Geospatial User Feedback Conceptual Model Standard (Masó & Bastin, 2016) using the ISO schemas as a baseline. Geospatial User Feedback is metadata that is predominantly produced by the consumers of geospatial data products as they use and gain experience with those products. This standard complements existing metadata conventions whereby documents recording dataset characteristics and production workflows are generated by the creator, publisher or curator of a data product. As a part of metadata, the GUF data model reuses some elements of ISO 19115-1:2014 but not the general structure. This selective use of ISO metadata elements was intended to prioritize future interoperability with developing ISO metadata models, and would allow an easy integration into the new version of ISO 19157.

## **Conclusions**

Geospatial data are relevant for the decision making of many daily activities and large investments, and therefore the quality of this data is important. Thus, ISO 19157 is a significant IS in the domain of geospatial data. ISO 19157 is currently in a revision process and therefore it is desirable that all interested parties are aware of this and

<sup>15</sup> <http://www.geoviqua.org/>

<sup>16</sup> <https://www.copernicus.eu/en/characterisation-metadata-enable-high-quality-climate-applications-and-services>

<sup>17</sup> <https://www.the-iea.org/projects/melodies-research-project/>

their ability to propose improvements. This document presents the contributions from a Spanish experts group.

In our view, ISO 19157 is a standard that has adequately fulfilled its function as a model to define, quantify and report quality, although its application has not been as widespread as desirable and when applied, has not been without problems. Many of the application problems are not particular of ISO 19157, they are typical of all standards in general, and this is a situation that requires a specific analysis that is outside the scope of this study.

Since technology and data availability have changed tremendously in recent years, we consider this revision of ISO 19157 appropriate. The challenges presented by the new types of available data as well as certain parts of traditional production processes that were not adequately covered with the ISO 19157:2013 version makes this revision essential if we want the proposed quality model that is continuously being applied. We need a model that ensures the highest degree of interoperability in the definition (conceptualization), quantification and reporting of data quality. For this reason the convergence with other standards and the inclusion of new dimensions of quality and new quality elements are aspects that we consider critical.

Additionally, the experience in the application of the model offers us a set of very specific guidelines focused on the field of spatial data production, such that they can make the application of the IS better understood, more efficient, powerful and versatile.

Therefore, we consider that the review is a good opportunity to improve this IS, otherwise the new types of data will bring other proposals for data quality that will surely require more additional efforts. There is no other option than to evolve or perish.

We are convinced that quality aspects are going to be more and more important in the near future for the proper use of geographic data and correct decision-making based on them in order to face the global challenges threatening the survival and well-being of humankind in the long term, as expressed in the 17 Sustainable Development Goals defined by the United Nations.

## References

- alphaBeta (2017). "The economic impact of geospatial services: how consumers, businesses and society benefit from location-based information". [https://www.alphabeta.com/wp-content/uploads/2017/09/GeoSpatial-Report\\_Sept-2017.pdf](https://www.alphabeta.com/wp-content/uploads/2017/09/GeoSpatial-Report_Sept-2017.pdf).
- AdV (2002). Erarbeitung eines Qualitätssicherungssystems für die Geodaten des amtlichen Vermessungswesens. Grundsätze für Qualitätskriterien und standardisierte Prüfverfahren für die Anwendung des AFIS-ALKIS-ATKIS-



- Basisschemas bei der Entwicklung der Anwendungsschemata. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland.
- Ahmed, B.S.; Bures, M., Frajtak, K.; Cerny, T. (2019). "Aspects of Quality in Internet of Things (IoT) Solutions: A Systematic Mapping Study", *IEEE Access*, vol. 7, pp. 13758-13780. <https://doi.org/10.1109/ACCESS.2019.2893493>.
- Ariza-López, F.J.; Rodríguez-Pascual, A.F. (2018). "Calidad en datos geográficos, geoservicios y productores de datos: análisis crítico", *Revista Cartográfica*, núm. 97, pp: 105-143. <https://doi.org/10.35424/rcar.v0i97.178>.
- Ariza-López F.J. (ed.) (2013). *Fundamentos de evaluación de la calidad de la información geográfica*, Servicio de Publicaciones de la Universidad de Jaén.
- Ariza-López, F.J.; Rodríguez-Avi, J.; Reinoso-Gordo, J.F.; Ariza-López, Í.A. (2019). "Quality Control of "As Built" BIM Datasets Using the ISO 19157 Framework and a Multiple Hypothesis Testing Method Based on Proportions", *ISPRS Int. J. Geo-Inf.*, 8, 569.
- Burrough, P.A. (1986). "Principles of Geographic Information Systems for Land Resource Assessment", *Monographs on Soil and Resources Survey* No. 12, Oxford Science Publications, New York.
- Castro, L., Porto, J.; dos Santos, R.; Zipf, A. (2017). A taxonomy of quality assessment methods for volunteered and crowdsourced geographic information. *Transactions in GIS*, 22(2): 542-560. <https://doi.org/10.1111/tgis.12329>.
- Degrossi, L.C., Porto de Albuquerque, J., Dos Santos Rocha, R., & Zipf, A. (2018). "A taxonomy of quality assessment methods for volunteered and crowdsourced geographic information", *Transactions in GIS: TG*, 22(2): 542-560. <https://doi.org/10.1111/tgis.12329>.
- Eldawy, A.; Mokbel, M.F. (2013). *The Era of Big Spatial Data: A Survey*, Foundations and Trends in Databases, 6(3-4): 163-273, Hanover, USA, Now Publishers Inc.
- EU (2020). "Building a European data economy". <https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy>
- Eurostat (2020). "Population grids". <https://bit.ly/2TwpC7>.
- Gartner (2020). <https://www.gartner.com/en/information-technology/glossary/big-data>
- GeoKnow (2012). "Metrics for Linked Geospatial Information. Deliverable 3.4.1.GeoKnow - Making the Web an Exploratory for Geospatial Knowledge".
- Goodchild, M. (2007). "Citizens as Sensors: The World of Volunteered Geography", *GeoJournal*, 69: 211-221. <https://doi.org/10.1007/s10708-007-9111-y>.
- Guptill, S.C.; Morrison, J.L. (1995). *Elements of Spatial Data Quality*, Pergamon Press.

- Gusminia, M.; Jabeurb, N.; Karama, R.; Melchioria, M.; Renso, C. (2017). "Reputation evaluation of georeferenced data for crowd-sensed applications", *The 8th International Conference on Ambient Systems, Networks and Technologies (ANT 2017)*, May 16-19, Madeira, Portugal.
- IECA (2011). NTCA 01-002. "Modelo Aseguramiento Calidad productos de IG en Andalucía". Recuperado de [https://www.juntadeandalucia.es/institutodeestadisticaycartografia/ieagen/sea/ntca/01\\_modelos/01002\\_Modelos\\_Aseguramiento\\_Calidad.pdf](https://www.juntadeandalucia.es/institutodeestadisticaycartografia/ieagen/sea/ntca/01_modelos/01002_Modelos_Aseguramiento_Calidad.pdf)
- ISO (2002). "Geographic information - Quality principles"
- ISO (2003). ISO 19114:2003 Geographic information - Quality evaluation procedures
- ISO (2006). ISO/TS 19138:2006 Geographic information - Data quality measures
- ISO (2013). ISO 19157:2013 Geographic information - Data quality
- ISO (2014). ISO 19115-1:2014 Geographic information - Metadata - Part 1: Fundamentals
- Karkouch, A.; Mousannif, H.; Al Moatassimeand, H.; Noel, T. (2016). "Data Quality in Internet of Things: A state-of-the-art survey", *Journal of Network and Computer Applications*, 73: 57-81. <http://dx.doi.org/10.1016/j.jnca.2016.08.002>.
- Lee, J.-G., Kang, M. (2015). Geospatial Big Data: Challenges and Opportunities. *Big Data Research*, 2(2): 74-81. <https://doi.org/10.1016/j.bdr.2015.01.003>.
- Li, S., Dragicevic, S., Anton, F., Sester, M., Winter, S., Coltekin, A.; Cheng, T. (2016). "Geospatial Big Data Handling Theory and Methods: A Review and Research Challenges", *ISPRS Journal of Photogrammetry and Remote Sensing*, 115: 119-133. <https://doi.org/10.1016/j.isprsjprs.2015.10.012>.
- López-Pellicer, F.; Vilches-Blazquez, L.; Zarazaga, F.J.; Muro-Medrano, P.; Corcho, O. (2011). "The Delft Report: Linked Data and the challenges for geographic information standardization", in *Jornadas Ibéricas de Infraestructuras de Datos Espaciales 2011 (JIIDE 2011)*, November 9-11, 2011, Barcelona, Spain.
- Masó, J.; Bastin, L. (2016). *OGC Geospatial User Feedback Standard. Conceptual model*. OGC 15-097r1. In Internet: <http://docs.opengeospatial.org/is/15-097r1/15-097r1.html>.
- Moellering, H. (1987). A draft proposed standard for digital cartographic data. National Committee for Digital Cartographic Data Standards. <https://doi.org/10.3133/ofr87308>.
- Moström, J.; Hedeklint, K.; Dysterud, M.; Engelién, E.; Bloch, V.; Tammisto, R., Lillethun, A. (2019). "GSGF Europe - Implementation guide for the Global Statistical Geospatial Framework in Europe". Proposal from the GEOSTAT 3 project. <https://bit.ly/2Ts8PEG>.

- NIST (1994). *Spatial Data Transfer Standard (SDTS). Federal Information Processing Standards Publication 173-1*, National Institute of Standards and Technology.
- Padilla, M., Stehman, S.V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Chuvieco, E. (2015). "Comparing the Accuracies of Remote Sensing Global Burned Area Products using Stratified Random Sampling and Estimation", *Remote Sensing of Environment*, 160: 114-121. <https://doi.org/10.1016/j.rse.2015.01.005>
- PIRA (2000). Commercial exploitation of Europe's public sector information, Final Report, Pira International, European Commission, Directorate General for the Information Society.
- Puyan, A.; Zadeh, G.W.; Cavka, H.B.; Staub-French, S.; Pottinger, R. (2017). "Information Quality Assessment for Facility Management", *Advanced Engineering Informatics*, 33: 181-205. <https://doi.org/10.1016/j.aei.2017.06.003>.
- Robinson, A.C.; Demšar, U.; Moore, A.B.; Buckley, A. Jiang, B.; Field, K., Sluter, C.R. (2017). "Geospatial big data and cartography: research challenges and opportunities for making maps that matter", *International Journal of Cartography*, 3 (sup1): 32-60. <https://doi.org/10.1080/23729333.2016.1278151>.
- SDMX (2009). List of subject-matter domains.
- Senaratne, H.; Mobasher, A.; Loai Ali, A.; Capineri, C.; Haklay, M. (2017). "A review of volunteered geographic information quality assessment methods". *International Journal of Geographical Information Science*, 31(1). 139-167 <https://dx.doi.org/10.1080/13658816.2016.1189556>
- Song, Y.; Wang, X.; Tan, Y.; Wu, P.; Sutrisna, M.; Cheng, J.C.P.; Hampson, K. (2017). "Trends and Opportunities of BIM-GIS Integration in the Architecture, Engineering and Construction Industry: A Review from a Spatio-Temporal Statistical Perspective", *ISPRS International Journal of Geo-Information*, 6(12): 397. <https://doi.org/10.3390/ijgi6120397>.
- Sun, J.; Harrie, L.; Jensen, A.; Eriksson, H.; Tarandi, V.; Uggla, G. (2018). *Description of geodata quality with focus on integration of BIM-data and geodata*. Smart Built Environment. Available online in: <https://bit.ly/3clKusH>.
- UNE (2016). UNE 148002:2016 *Metodología de evaluación de la exactitud posicional de la información geográfica*.
- UN-EAPD (2019). "Data Economy: Radical transformation or dystopia?", *Frontier Technology Quarterly*, January 2019. <https://bit.ly/2TcnmVQ>.
- UN-ISGI (2018). *Global Statistical Geospatial Framework: Linking Statistics and Place*. United Nations Expert Group on the Integration of Statistical and Geospatial Information.
- UN-SD (2019). *United Nations National Quality Assurance Frameworks Manual for Official Statistics* (UN NQAF Manual). UN Statistics Division

- Ureña-Cámara, M.A.; Nogueras-Iso, J.; Lacasta, J.; Ariza-López, F.J. (2019). "A method for checking the quality of geographic metadata based on ISO 19157", *International Journal of Geographical Information Science*, 33(1): 1-27. <https://doi.org/10.1080/13658816.2018.1515437>
- Wikipedia (2020a). "Data economy". [https://en.wikipedia.org/wiki/Data\\_economy](https://en.wikipedia.org/wiki/Data_economy)
- Wikipedia (2020b). "Internet of things". [https://en.wikipedia.org/wiki/Internet\\_of\\_things](https://en.wikipedia.org/wiki/Internet_of_things)
- Wikipedia (2020c). "Linked data". [https://en.wikipedia.org/wiki/Linked\\_data](https://en.wikipedia.org/wiki/Linked_data)
- Zabala, A.; Masó, J. (2016). *Testbed-12 Imagery Quality and Accuracy Engineering Report*. OGC 16-050, in Internet: <http://docs.opengeospatial.org/per/16-050.html>.
- Zabala, A.; Riverola, A.; Serral, I.; Díaz, P.; Lush, V.; Masó, J.; Pons, X.; Habermann T. (2013). Rubric-Q: Adding Quality-related Elements to the GEOSS Clearinghouse Datasets. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (IEEE J-STARS), 6(3): 1676-1687. <https://doi.org/10.1109/jstars.2013.2259580>.
- Zaveri, A.; Rula, A.; Maurino, A.; Pietrobon, R.; Lehmann, J.; Auer, S. (2014). "Quality Assessment Methodologies for Linked Data: A Survey", *Semantic Web Journal*. <http://www.semantic-web-journal.net/content/quality-assessment-linked-data-survey>.

# Reproducibilidad en AGILE: experiencias, logros y recomendaciones

## *Reproducibility in AGILE: experiences, achievements, and recommendations*

Carlos Granell\*

Barbara Hofer<sup>▲</sup>

Daniel Nüst<sup>♦</sup>

Frank O. Ostermann<sup>•</sup>

Rusne Sileryte<sup>■</sup>

*Recibido 3 de enero de 2020; aceptado 30 de enero de 2020*

### Resumen

Este artículo describe las experiencias, actividades realizadas, recursos generados y recomendaciones para la promoción e incentivación de prácticas en investigación reproducible en el campo de la ciencia de la información geográfica. Aunque el artículo se centra en la comunidad y conferencia AGILE (Association of Geographic Information Laboratories in Europe), creemos que los recursos y lecciones aprendidas pueden ser extrapolables a comunidades y asociaciones científicas afines en otras regiones, como puede ser Latinoamérica. En este sentido, hacemos especial hincapié

\* Universitat Jaume I de Castellón, España, correo electrónico: [carlos.granell@uji.es](mailto:carlos.granell@uji.es).

ORCID: <http://orcid.org/0000-0003-1004-9695>

▲ University of Salzburg, Austria, correo electrónico: [barbara.hofer@sbg.ac.at](mailto:barbara.hofer@sbg.ac.at).

ORCID: <http://orcid.org/0000-0001-7078-3766>

♦ University of Münster, Alemania, correo electrónico: [daniel.nuest@uni-muenster.de](mailto:daniel.nuest@uni-muenster.de).

ORCID: <http://orcid.org/0000-0002-0024-5046>.

• University of Twente, Enschede, Países Bajos, correo electrónico: [f.o.ostermann@utwente.nl](mailto:f.o.ostermann@utwente.nl).

ORCID: <http://orcid.org/0000-0002-9317-8291>

■ Delft University of Technology, Países Bajos, correo electrónico: [r.sileryte@tudelft.nl](mailto:r.sileryte@tudelft.nl)

ORCID: <http://orcid.org/0000-0002-8245-3016>

en la descripción de las directrices para la redacción de artículos reproducibles propuestas en el seno de AGILE, así como en una serie de recomendaciones dirigidas principalmente a asociaciones y organizadores de conferencias científicas para la adopción y promoción paulatina de dichas prácticas. Se trata pues de una llamada a la comunidad latinoamericana de la ciencia de la información geográfica a que consideren los recursos aquí detallados con el fin de incentivar la investigación reproducible en sus comunidades científicas.

Palabras clave: *reproducibilidad, investigación reproducible, directrices para artículos reproducibles, ciencia de la información geográfica, AGILE.*

## Abstract

This article describes the experiences, activities carried out, resources generated and recommendations for the promotion of practices in reproducible research in the field of geographic information science. Although the article focuses on the AGILE (Association of Geographic Information Laboratories in Europe) community and conference, we believe that the resources and lessons learned can be extrapolated to related scientific communities and associations in other regions, such as Latin America. In this regard, we place special emphasis on the description of the guidelines for reproducible articles proposed within AGILE, as well as on a series of recommendations directed mainly to associations and organisers of scientific conferences for the gradual adoption and promotion of these practices. It is therefore a call to the Latin American GIScience community to consider the resources detailed here in order to encourage reproducible research in their scientific communities.

Key words: *reproducibility, reproducible research, guidelines for reproducible papers, GIScience, AGILE.*

## Introducción

Este artículo reflexiona sobre la aplicación en general de la investigación reproducible en las Ciencias de la Información Geográfica y, en particular, se centra en el caso de la conferencia internacional AGILE (<https://agile-online.org/>). No se trata de un artículo donde se discute la definición del término reproducibilidad científica, y su importancia para la ciencia moderna y nuestra comunidad investigadora. Respecto a la primera cuestión, la definición, existe una extensa bibliografía científica al uso (véase, por ejemplo, Barba, 2018). Respecto a la segunda, su importancia, no podemos más que subrayar de nuevo la relevancia de la reproducibilidad como un pilar básico de la ciencia. Otros autores también han escrito mucho sobre ello incluyendo un reciente artículo en la *Revista Cartográfica* (Ballari, 2018), por lo que no vamos a incidir de nuevo en esta cuestión, aunque sí que repasaremos brevemente algunos trabajos recientes que se centran explícitamente en la reproducibilidad en la Ciencia

de la Información Geográfica (CIG). A estas alturas, está fuera de toda duda para el lector que la investigación reproducible, la ciencia abierta, la reproducibilidad computacional y replicación son imprescindibles para la buena práctica y la integridad en la investigación y ciencia actual.

Este artículo relata la experiencia de un pequeño grupo de investigadores en la CIG (los autores) para la promoción e incentivación de prácticas en investigación reproducible en la comunidad AGILE. Describimos nuestra historia, las acciones y actividades llevadas a cabo y los logros conseguidos y nos centramos, especialmente, en uno de los recursos generados —las directrices para artículos reproducibles—, que se ponen en práctica por primera vez en la conferencia AGILE 2020. De ningún modo nuestra intención es elevar nuestro ego personal con este artículo. Nuestra intención es mucho más prosaica, regida por la simple idea de compartir experiencias y resultados tanto positivos como negativos.

Queremos que el relato de nuestras actividades y experiencias sirvan de ejemplo de primera mano para aquellos lectores interesados en la investigación reproducible, que las hagan suyas y las adapten a sus entornos. Creemos firmemente que la fórmula para incentivar progresivamente las prácticas en investigación reproducible en una comunidad y/o conferencia en la CIG es exportable y aplicable a otras situaciones y regiones (Granell *et al.*, 2018). Por ello, este artículo incluye una serie de recomendaciones para la puesta en marcha de la investigación reproducible, que puede aplicarse a diversos contextos, como por ejemplo dentro de un grupo de investigación, en un departamento, o incluso a comunidades relacionadas con la CIG de ámbito nacional o supranacional.

Si el primer objetivo es transmitir nuestras experiencias y logros obtenidos en la comunidad AGILE, el segundo objetivo de este artículo es una “llamada” a la comunidad CIG de América Latina en cuanto a la investigación reproducible. En especial, creemos que América Latina está bien posicionada para la adopción de prácticas reproducibles en la investigación, ya que la cultura de lo abierto (acceso abierto, software abierto, etc.) ha estado tradicionalmente bien enraizada en el mundo académico y educativo.

A continuación, introducimos a grandes rasgos el concepto de reproducibilidad en la investigación, citando algunos trabajos recientes en el campo de la CIG, así como recursos docentes relacionados. Luego, enumeramos las actividades realizadas en el marco de nuestra iniciativa para la promoción de las buenas prácticas en investigación reproducible tanto en la comunidad AGILE como en su conferencia anual. Uno de los recursos generados de especial relevancia para la conferencia AGILE son las directrices para autores para la elaboración de artículos reproducibles, que se tratan en una sección aparte. Finalmente, terminamos el artículo con una serie de recomendaciones, dirigidas a los distintos actores que forman parte del proceso de investigación científica (investigadores, universidades, entidades editoras, entidades

de política científica, etc.) y que, desde su posición e intereses particulares, pueden influir significativamente en la promoción de hábitos y buenas prácticas en la investigación reproducible.

### Breve reseña sobre la reproducibilidad en CIG

En esta sección definimos el término reproducibilidad. Luego, indicamos algunos trabajos recientes o de especial relevancia en la aplicación o estudio de aspectos reproducibles a la CIG, que dan fe del creciente interés de la comunidad investigadora de la CIG en las prácticas reproducibles. Finalmente, a modo informativo, añadimos algunos recursos docentes que se centran en la reproducción desde la óptica de la educación y de la realización de materiales docentes. Este artículo no trata la reproducibilidad como concepto en la docencia, pero definitivamente será un aspecto clave en el futuro próximo para que la próxima generación de científicos y científicas adquieran las prácticas de investigación reproducibles como parte natural de sus actividades diarias.

Nuestro interés en reproducibilidad se centra en las publicaciones en conferencias (como AGILE) y, por lo tanto, nos basamos en la definición adoptada en trabajos previos (Ostermann y Granell, 2017; Nüst *et al.*, 2018). Otros trabajos como Stodden *et al.* (2014), Barba (2018), Nüst *et al.* (2018) y la colección de artículos y editoriales de *Nature* (<https://www.nature.com/collections/prbfkwmwvz>) abordan en detalle la literatura en cuanto a las diferentes interpretaciones del concepto de reproducibilidad, desde disciplinas y áreas del conocimiento científico dispares. En este trabajo, nos ceñimos a la siguiente definición: “un documento reproducible garantiza que un revisor o lector pueda recrear el flujo de trabajo computacional de un estudio o experimento, incluidos los conocimientos previos y el entorno computacional. El primero implica que el argumento científico es comprensible y sólido. El último requiere una descripción detallada del software y los datos usados, y que ambos están disponibles de forma abierta.” (Nüst *et al.*, 2018: 7).

Con respecto a la literatura científica focalizada en la reproducibilidad en la CIG, Richardson *et al.* (2015) ya alertaba de las barreras a la reproducibilidad en la CIG con respecto a la revelación de la localización y otros datos confidenciales de participantes en un proyecto de investigación o experimento. Sin embargo, ha sido a partir de 2018 cuando la comunidad de la CIG ha puesto el ojo con mayor atención a las prácticas de investigación reproducible. Por ejemplo, nuestro trabajo previo, Ostermann y Granell (2017) y Nüst *et al.* (2018) ya ponen el foco en la reproducción y replicación de estudios en la CIG, bien en los datos geoespaciales generados por el usuario o en trabajos de investigación en el marco de la conferencia AGILE, respectivamente. El último de ellos lo trataremos en detalle en la siguiente sección. En la línea de la reproducción de estudios que procesan datos generados por el usuario,



Cerutti *et al.* (2019) examina la utilidad de los sistemas de gestión del flujo de trabajos científicos para reproducir tres estudios que investigan las redes sociales y los incendios forestales en Australia.

Gahegan (2019), reflexiona sobre las dificultades de la reproducción desde el punto de la eCiencia (*eScience* en inglés), donde el campo de la geo-computación es determinante, es decir, la aplicación de análisis computacionales a casos de usos donde las dimensiones espacial y temporal son claves. El autor sugiere distintas aproximaciones para la mejora de la reproducción de flujos de trabajo geo-computacionales. Relacionado con la compartición de análisis computacionales, Konkol *et al.* (2020) comparan exhaustivamente una serie de herramientas y aplicaciones para la publicación y ejecución de flujos de trabajos computacionales de forma reproducible.

Otro hilo de trabajos relacionados se centra en el plano teórico en lugar de aspectos prácticos en el uso de herramientas de ayuda a la reproducción, como en los ejemplos anteriores. Los trabajos de Kedron y colegas (Kedron, 2018 y Kedron *et al.*, 2019), se contextualizan en la investigación relacionada con la rama de la geografía humana o económica. Los autores discuten la incertidumbre que va aparejada a las distintas conceptualizaciones e interpretaciones en la difusión y comunicación de los resultados de los análisis geográficos. Todo ello, sin lugar a dudas, afecta seriamente a la reproducción y replicación de estudios y análisis en el campo de la geografía humana y económica. Otro ejemplo destacable es la reflexión de Kray *et al.* (2019) con respecto a la investigación reproducible en el campo de la geo-informática. Los autores, se alejan un tanto de la aproximación teórica de los dos ejemplos anteriores, para reflexionar sobre los retos y ventajas de la reproducción en la geo-informática, y en la CIG en general, donde la tecnología y las ciencias de la computación son fundamentales.

Terminamos esta reseña con un guiño a una serie de recursos docentes sobre la reproducción en la ciencia en general y aplicada a la CIG. Cabe destacar la iniciativa FORRT (*Framework for Open and Reproducible Research Training*, en inglés) (FORRT, 2019) que teoriza sobre las dimensiones y factores a tener en cuenta para la incorporación de prácticas reproducibles en el currículum docente. Crüwell *et al.* (2018) mantienen una lista debidamente curada de recursos y artículos relevantes que van desde la problemática inicial hasta cuestiones de análisis estadístico que minan la reproducción. Aunque la lista de recursos tiene un cierto sesgo hacia la psicología, debido a que la reproducibilidad ha sido tradicionalmente un aspecto crítico en esta disciplina, resulta relevante para el diseño y confección de un hipotético currículum sobre la reproducibilidad y replicabilidad de la ciencia.

Si miramos a la CIG, ya empiezan a surgir recursos docentes, en los cuales la reproducción ocupa una posición predominante. Por ejemplo, como parte de la visión que integra la ciencia de los datos y la CIG (Singleton y Arribas-Bel, 2019), Arribas-Bel (2019) describe un curso completo que incluye módulos didácticos, junto con

código y datos, en una mezcla única que aplica técnicas de análisis de datos desde la perspectiva geográfica, atendiendo por supuesto a aspectos de reproducibilidad, especialmente con el uso de Jupyter Notebooks (Millman y Pérez, 2014).

### **Iniciativa de reproducibilidad en AGILE: experiencias y actividades**

En esta sección resumimos las actividades y acciones realizadas en los tres últimos años, desde el primer encuentro del grupo en la conferencia de AGILE 2017 (mayo 2017) hasta las actividades en curso y planteadas para la próxima conferencia de AGILE en 2020. No se trata únicamente de trazar una línea temporal de los acontecimientos, la cual resulta por sí sola interesante, sino de recalcar el hecho de que un grupo reducido y apasionado en la promoción de las prácticas reproducibles en la investigación puede influir de manera considerable en un grupo mucho mayor, la comunidad AGILE, hasta el punto de introducir cambios estructurales en la próxima conferencia, en cuanto a la ciencia abierta y prácticas reproducibles de los trabajos de investigación enviados. El camino no ha sido fácil, y nuestros objetivos iniciales no eran ni mucho menos tan ambiciosos. Como veremos en las actividades realizadas en el resto de esta sección, en todo momento nuestro contexto de actuación se limitó a la comunidad AGILE, y no a otras comunidades CIG en otras latitudes y regiones. Delimitar un campo de actuación manejable y de acorde a las posibilidades, pero a la vez lo bastante ambicioso, ha sido sin duda uno de los factores de éxito de nuestra iniciativa (Granell *et al.*, 2018).

El mensaje que queremos transmitir es que un pequeño grupo puede llevar a cabo grandes cambios, hasta el punto de influenciar la comunidad de la cual forma parte. Esperamos que las actividades y recursos aquí relatados, junto con las recomendaciones expuestas en la próxima sección, surjan de ejemplo para que (pequeños o grandes) colectivos en Latinoamérica impulsen el cambio hacia prácticas de reproducibilidad en sus propias comunidades de ámbito regional o nacional, o incluso, supra nacional, en el ámbito de la investigación en la Ciencia de la Información Geográfica.

#### ***Origen y contexto***

Nuestras actividades nacen y se desarrollan en el ámbito de AGILE. Al tratarse de una asociación europea, cabe incluir unas breves líneas sobre sus objetivos y trayectoria para aquellos que la desconozcan. En primer lugar, AGILE es una asociación que agrupa actualmente a unos 100 miembros, provenientes de diversos países europeos. Un miembro es un grupo de investigación o departamento de universidad cuya investigación, tanto teórica como aplicada, se desarrolla en torno a la CIG. Como asociación, AGILE lleva a cabo diversas acciones para sus miembros con el fin de

incentivar la comunicación y la transferencia de conocimiento. Destaca la conferencia homónima, que se celebra anualmente de forma ininterrumpida desde 1998. La conferencia AGILE cumple ya 23 ediciones en 2020 y por ello se puede considerar una de las más longevas en su campo.

Los autores de este artículo son miembros de la Asociación y la conferencia ha sido un evento fundamental desde el inicio de sus carreras investigadores, hasta el punto que podríamos considerarla como nuestra *alma mater*, si tal calificativo fuera aplicable a conferencias tal como lo es para la universidad de formación. De lo anterior se desprende una motivación compartida del grupo para mejorar AGILE y regenerar la conferencia tradicional (hasta 2019) de forma que la reproducibilidad y los datos abiertos sean características de los artículos de la conferencia (para la próxima edición de 2020).

Por lo tanto, una de las claves de éxito fue la acertada elección de la comunidad. Podríamos haber elegido nuestras universidades o nuestros grupos de investigación. Sin embargo, nuestra motivación, la red de contactos, e incluso, un cierto sentimiento para mejorar el evento científico que nos ha visto crecer y desarrollarnos científicamente, fueron sin lugar a dudas factores decisivos a la hora de optar por AGILE como objetivo. Asimismo, creemos que estos mismos factores pueden ser perfectamente válidos para la elección de comunidades objetivo en otras regiones como en América Latina.

A continuación desglosamos la lista de acciones (<https://o2r.info/reproducible-agile/>) y como han crecido en complejidad, en cuanto a los resultados obtenidos, y en eficiencia, a la hora de estimular cambios estructurales en la conferencia AGILE.

### ***Acciones realizadas***

El primer encuentro del grupo tuvo lugar en la conferencia AGILE 2017 (Wageningen, Países Bajos), que justo marcaba el vigésimo aniversario de la conferencia y, por esa razón, el evento volvía al país donde se celebró la primera edición. En el taller previo a la conferencia Reproducible Geosciences Discussion Forum (<https://o2r.info/reproducible-agile/2017/>), discutimos el tema de la investigación reproducible en general y si la “crisis de reproducibilidad” que se anunciaba en economía (Ioannidis, Stanley y Doucouliagos, 2017), neurociencia (Button *et al.*, 2013), y en otros tantos campos del conocimiento (Ioannidis, 2005) también había llegado a la CIG. El intercambio de opiniones llevó rápidamente al hecho de que la reproducibilidad, la reproducibilidad computacional, replicación, y otros términos relacionados (véase Barba, 2018), no parecían ocupar un lugar destacado en la agenda de la comunidad de AGILE. Como resultado, el taller se reorientó hacia el intercambio de conocimiento y la colaboración entre el grupo con el fin de profundizar más sobre el

estado de reproducibilidad de los trabajos de investigación publicados en la conferencia AGILE y explorar cómo podría mejorar y extenderse la conferencia con buenas prácticas de investigación reproducible.

De esta primera toma de contacto cabe destacar dos acciones fundamentales. 1) marcó la senda de la colaboración del grupo. La distribución geográfica del grupo hacía difícil las reuniones presenciales, a excepción de los encuentros anuales en cada edición de la conferencia, por lo que organizamos teleconferencias periódicas (que siguen todavía en curso) para debatir y consensuar próximas acciones y repartir responsabilidades. 2) acordamos la redacción de un artículo para la siguiente conferencia AGILE (2018), en el cual responderíamos a las siguientes preguntas: ¿Qué aspectos hay que tener en cuenta para reproducir un artículo concreto? ¿Cuál es el estado de reproducibilidad de los mejores artículos presentados en edición pasadas de la conferencia AGILE? ¿Qué podría hacer AGILE (como asociación) para fomentar la reproducibilidad en las próximas ediciones de la conferencia?

Nüst *et al.* (2018) plasmaron algunas de las respuestas para las preguntas anteriores. El trabajo analizó el nivel real de reproducibilidad de un conjunto de artículos (32) nominados para las sesiones de mejores trabajos (tanto completos como cortos) de las conferencias AGILE realizadas entre 2010 y 2017 (ambos inclusive). Con respecto a la primera pregunta, propusimos una serie de criterios —datos, pre-procesamiento, métodos, entorno computacional, y resultados—, y asignamos un nivel de reproducibilidad entre 0 (no reproducible) y 3 (completamente reproducible). De esta forma, calculamos fácilmente el nivel de pre-reproducibilidad (Stark, 2018) de cada trabajo, sopesando el nivel de cada criterio basado en la lectura de cada artículo. Los resultados arrojaron una valoración pobre en cuanto a los niveles de reproducibilidad, coincidiendo en gran medida con la opinión de los propios autores de los trabajos analizados (Nüst *et al.*, 2018). Respecto a la tercera pregunta del párrafo anterior, planteamos una serie de medidas y recomendaciones para atajar el problema de la falta de reproducibilidad, que comentaremos más adelante en la sección “Recomendaciones”.

Llegó la siguiente conferencia AGILE 2018 (Lund, Suecia) y realizamos el taller Reproducible Research Publications (<https://o2r.info/reproducible-agile/2018/>). Con una asistencia más bien mediocre, presentamos los resultados de nuestro análisis (Nüst *et al.*, 2018) para poner de relieve las carencias en reproducibilidad de los trabajos presentados en conferencias pasadas de AGILE. El caso de estudio despertó el interés de los participantes, pasando inmediatamente de un plano conceptual a uno práctico y experimental. El concepto de reproducibilidad en sí es sencillo, lo difícil es responder a la pregunta ¿por dónde empiezo para adquirir prácticas reproducibles en mi investigación? La segunda parte del taller, versó justo en la reproducción por parte de los participantes del flujo de trabajo computacional depositado en un repositorio abierto (Nüst, 2018) asociado al análisis publicado (Nüst *et al.*, 2018).

Los resultados de la reproducción fueron dispares y desiguales, dependiendo del conocimiento previo de los participantes, desde expertos hasta principiantes, en el lenguaje utilizado (R) y en las herramientas (RStudio). Sin embargo, lejos de ser un fracaso, la experiencia fue fructífera, ya que demostró que la reproducción de un trabajo por un investigador no experimentado no es una tarea imposible. Algunos de los participantes, sin el conocimiento previo en el lenguaje R, lograron reproducir en menos de dos horas un compendio de investigación bien preparado (Nüst *et al.*, 2017), es decir, un paquete reproducible que integra código y datos. En la parte negativa, la instalación del entorno computacional adecuado para la ejecución del código y las dependencias entre paquetes de software, como apuntado posteriormente por Konkol *et al.* (2019), fueron las mayores dificultades encontradas por los participantes. Aunque parezca banal, es una lección valiosísima para los autores de artículos reproducibles: si deseas que los no expertos puedan reproducir tu trabajo, no asumas que saben lo que es fácil de hacer (para el autor).

El taller finalizó con un debate abierto. Los participantes admitieron que la reproducción manual es una experiencia de aprendizaje eficaz siempre y cuando se parta de un trabajo reproducible bien preparado. Conocer de primera mano los problemas con los que se encuentran a menudo otros investigadores al reproducir un trabajo, es una ayuda invaluable para motivar los cambios de hábitos en un investigador en cuanto a la adopción de buenas prácticas de investigación reproducibles. No todos los problemas son previsible, por supuesto. Hubo consenso en admitir que una sala llena de gente con conocimientos e intereses diversos, y con sistemas operativos y versiones de software variopintos, encontrará siempre alguna dificultad en la reproducción que el autor no pudo prever con antelación.

Una de las apreciaciones más destacables del taller, fue que el concepto de reproducción no es difícil de entender. Todo el mundo lo entiende y aprecia su importancia para la ciencia. Sin embargo, parece ser que su simplicidad conceptual correlaciona negativamente con la complejidad experimental. Todos los participantes expresaron serias dudas en cuanto por dónde empezar para adquirir hábitos reproducibles en su trabajo de investigación diario. Aun reconociendo que no existen proyectos de investigación “típicos”, y que por lo tanto la reproducción está sujeta a las idiosincrasias de cada proyecto o trabajo de investigación, hubo un clamor popular por la disponibilidad de materiales educativos, guías o directrices para ayuden a los investigadores en la aplicación de la investigación reproducible.

Nosotros escuchamos e hicimos de las directrices nuestro siguiente reto.

Con una aportación económica de la asociación AGILE, los autores junto con un grupo reducido de expertos internacionales en ciencia abierta, curación de datos y reproducibilidad (véase <https://o2r.info/reproducible-agile/initiative/>), nos reunimos tres días (abril, 2019) en Delft (Países Bajos) para abordar la creación de unas directrices para autores y revisores para la elaboración de artículos reproducibles para

AGILE. Las directrices deberían ayudar tanto a los autores, para preparar los manuscritos, como a los revisores, para evaluarlos. Las directrices deberían ser lo bastante genéricas para cubrir la mayoría de los tipos de trabajos enviados a la conferencia AGILE, pero al mismo tiempo lo bastante detalladas para que sirvieran de ayuda a los autores, incluyendo por ejemplo recomendaciones en cuanto a repositorios de datos (p.e. Zenodo, etc.) y de código (p.e. OSF, GitLab), flujos de trabajo transparentes (p. ej., Notebook en Binder), normas de citación y tipo de licencias de software y datos. Posteriormente, en la sección “Directrices para la redacción de artículos reproducibles”, las describimos en detalle.

El tercer taller de la serie se realizó en la Conferencia AGILE 2019 (Limassol, Chipre), con el título de “Reproducible Research” (<https://o2r.info/reproducible-agile/2019/>). Con un registro de asistentes mayor que en la edición anterior, el taller se centró en la difusión de las directrices recién publicadas (Nüst *et al.*, 2019), reforzado con la presentación de dichas directrices en la sesión de posters (Hofer *et al.*, 2019). En definitiva, por un lado, sopesamos que tal las directrices habían sido recibidas por parte de la comunidad AGILE. Por otro, seguimos enfatizando en el taller la parte práctica de la reproducibilidad, planteando sesiones de reproducción de trabajos para niveles básico y avanzado.

Recientes actividades ponen de manifiesto el calado de las prácticas en investigación reproducibles en la comunidad AGILE. Por ejemplo, la quinta edición, celebrada en Tartu (Estonia) en noviembre de 2019, dedicó, por primera vez, un día completo a la reproducibilidad. Además, el grupo (los autores) seguimos adelante con la cuarta edición de talleres sobre reproducibilidad para AGILE 2020 (<https://o2r.info/reproducible-agile/2020/>).

### ***Políticas realizadas***

Las actividades en el plano técnico (la serie de talleres, las directrices, etc.) son necesarias, pero no suficientes para llevar a cabo cambios estructurales de calado. Sin la apuesta clara de la Asociación, en su compromiso por la integración de la reproducibilidad en las conferencias AGILE, todo el trabajo realizado hasta el momento corría el serio peligro de quedar en papel mojado.

Por fortuna, el Consejo apoyó la reproducibilidad. El siguiente paso lógico fue instaurar oficialmente dichas directrices en AGILE, como parte fundamental de la próxima conferencia en el 2020, lo cual acarreó cambios substanciales en la estructura organizativa de la conferencia:

- La obligatoriedad para los autores de ajustarse a las directrices para la redacción de artículos reproducibles enviados a la conferencia AGILE (Nüst *et al.*, 2019).

- La creación de un nuevo Comité de Reproducibilidad, para asegurar que todos los manuscritos enviados cumplieran las directrices de artículos reproducibles para AGILE.
- La creación de un servicio de discusión en línea para ayudar a los autores en la adaptación de las directrices en sus trabajos.
- La creación de unas recomendaciones para los revisores (Nüst *et al.* 2020), para tener en cuenta los aspectos relacionados con la reproducibilidad en la revisión de artículos.
- La transformación lógica de la conferencia AGILE a una conferencia de acceso abierto. Para la edición de 2020, los artículos largos se publicarán en Springer con licencia en acceso abierto, mientras que los artículos cortos lo harán con Copernicus Publications, un reputado editor de acceso abierto.

Muchos más cambios se entreven para futuras ediciones de la conferencia AGILE, con el fin de convertirla en un referente de reproducibilidad en el campo de la CIG. Pero como en todo, los cambios radicales deben ser progresivos, sobre todo para dar tiempo a la comunidad a interiorizar los hábitos y prácticas en investigación reproducible, resumidos someramente en las directrices que comentamos a continuación.

### **Directrices para la redacción de artículos reproducibles**

En esta sección nos centramos en las directrices para la redacción de artículos reproducibles para la conferencia AGILE. El recurso de referencia es la versión inglesa (Nüst *et al.*, 2019a), la cual se actualiza periódicamente a media que se incorporan comentarios, aclaraciones y ejemplos de los propios autores. También hay disponible una versión traducida al castellano (Granell, 2019).

Las directrices no son exhaustivas en cuanto a los pasos a realizar. Existen excelentes recursos que describen reglas y pasos concretos para que los datos y análisis computacionales asociados a un trabajo científico sean reproducibles (p.e. Sandve *et al.*, 2013; Wilson *et al.*, 2017, Rule *et al.*, 2019). Sin embargo, las directrices aportan un valor adicional con respecto a estos recursos más generalistas, en cuanto que se centran en las características propias de los trabajos de investigación en la CIG, aportando ejemplos específicos a esta disciplina.

¿Qué elementos o aspectos son los más destacables en las directrices? Por su utilidad práctica, el primero es sin lugar a dudas la lista de verificación de la reproducibilidad de un artículo previo al envío. Los autores/as pueden comprobar fácilmente si algún elemento de la lista todavía requiere su atención. Resulta un recurso muy útil y rápido para asegurar que los artículos enviados cumplen con todos los requisitos pertinentes a la reproducibilidad.

Otro elemento destacable es el contenido de la nueva sección “Disponibilidad de Datos y Software”. Con ella forzamos a los autores/as a indicar dónde se encuentran

publicados los datos y el código, para que otros investigadores/as puedan acceder a esos recursos. Evidentemente, la reproducibilidad implica necesariamente el acceso público a los recursos asociados a una investigación. Esta sección por lo tanto permite aglutinar en un único lugar toda la información necesaria para acceder a los datos y código utilizados en el propio artículo. En caso de que existan restricciones, como por ejemplo que los datos utilizados se rigen por una licencia restrictiva que no permite su acceso y compartición, entonces sugerimos que se indique también los motivos. El objetivo de esta sección es que sea homogénea a todos los artículos y que permita fácilmente evaluar el nivel de pre-reproducibilidad (Stark, 2018) de un artículo simplemente leyendo la información de la sección. Las directrices incluyen una plantilla para el contenido de la misma, que puede tomarse como base para otros tipos de documentos como por ejemplo trabajos de maestrías.

Como los datos y el código son elementos clave para la reproducción de un trabajo, el documento de directrices los trata en dos apartados independientes. El apartado de datos se refiere a la información de entrada en una investigación; a los datos intermedios y resultados finales (p.e. mapas, tablas, graficas), que normalmente forman parte del artículo. El segundo apartado engloba tanto el código (o paquete de software, biblioteca, etc.) pertinente al análisis computacional realizado, como al entorno necesario para ejecutar dicho código.

Ambos apartados siguen una estructura similar. Organizamos las sugerencias a los autores en una tabla de dos dimensiones. Por un lado, respondemos los cuestionamientos de: qué, dónde y cómo; y por otro, planteamos para cada una de las tres preguntas anteriores, requisitos mínimos, que todos los autores deben cumplir, y también requisitos ideales. La reproducción de un artículo de investigación no es una cuestión de todo o nada, sino que hay un espectro de posibilidades entre los dos extremos. Por lo tanto, queremos transmitir la idea que la redacción de un artículo reproducible es un proceso gradual, de menos (mínimo) a más (ideal), y reconocemos que los autores sin experiencia se fijen con los requisitos mínimos y, a medida que adquieren experiencia y destrezas en el uso de herramientas y aplicaciones de soporte a la reproducibilidad, tengan en consideración las sugerencias ideales.

Destaca también que ambos apartados incluyen una serie de ejemplos (“que pasa si”) en el contexto de la CIG, que aportan sugerencias concretas a los autores para cada uno de los supuestos planteados. Por ejemplo, si los datos no están debidamente anonimizados y se publican con el nombre de los autores, damos información de cómo crear enlaces anónimos a los datos depositados en repositorios para permitir el proceso de revisión por el sistema de doble ciego. Para que las directrices fuesen un recurso conciso y breve para la comunidad investigadora, mantenemos un sitio web y wiki asociado (<https://osf.io/phmce/wiki/home/>) donde se incluyen muchos otros ejemplos para que los autores reconozcan su situación particular y sepan cómo actuar para hacer que sus artículos sean (más) reproducibles.



## Recomendaciones

Sin lugar a dudas, adquirir un nuevo hábito requiere esfuerzo y persistencia. A diferencia de hábitos que se centran únicamente en el individuo como la adquisición de hábitos saludables, la ciencia actual ya no se limita a la labor de un investigador individual, sino que se ve afectada por un ecosistema de actores con diferentes intereses que moldean el proceso científico. Desde investigadores, grupos de investigación, universidades, asociaciones, pasando por las editoriales que publican las revistas académicas, hasta las agencias públicas y privadas que subvencionan la investigación, y los comités de selección y promoción que evalúan los méritos científicos, todo ellos influyen de alguna manera en cómo se realiza la labor científica diaria. Por lo tanto, para mejorar significativamente la reproducibilidad en la investigación, se deben realizar cambios en cada uno de estos actores, con mejoras en los planes de estudio educativos, en los procesos de laboratorio, la política científica de las universidades, las normas de publicación de artículos de las editoriales de revistas académicas y en los criterios de las agencias de financiación, de modo que los esfuerzos requeridos para llevar a cabo las buenas prácticas en la investigación reproducibles sean reconocidos y debidamente valorados.

Es esta sección no abordamos cada uno de estos actores, sino que nos centramos en recomendaciones para las asociaciones relacionadas con la CIG, que tienen el poder para influir positivamente en una comunidad amplia de investigadores/as en la promoción de prácticas de investigación reproducible. Este aspecto es especialmente relevante en Latinoamérica, por ejemplo, con el Instituto Panamericano de Geografía e Historia (IPGH, <https://www.ipgh.org/>), ya que puede jugar un papel trascendental a la hora de promocionar la investigación reproducible en la región. Para autores individuales, remitimos a la discusión en Nüst *et al.* (2018), que a su vez incluye una buena selección de recursos para autores.

Por lo tanto, ¿qué pueden hacer las asociaciones relacionadas con la CIG para fomentar la reproducibilidad? Un prerequisite clave es que las propias asociaciones científicas como IPGH u otras, que proveen soporte para la publicación científica ya sea por medio de revistas académicas o conferencias, reconozcan el papel primordial de las prácticas de investigación reproducible para la comunidad de la CIG. Otra recomendación obvia es que las publicaciones sean en acceso abierto, y con licencias permisivas para los propios autores y para el resto de la comunidad investigadora. En este sentido, las publicaciones periódicas del IPGH ya se licencian en “Creative Commons”, los artículos están en abierto, y el sistema de gestión y publicación de revistas se basa en código abierto (Open Journal System), lo cual allana el camino hacia las siguientes recomendaciones.

Las directrices para autores son un recurso fundamental para sentar las bases de una conferencia o revista reproducible. Como hemos visto en la sección anterior, unas directrices deben incluir instrucciones claras sobre cuándo, cómo y dónde publicar

material complementario (datos, código) al artículo escrito. Además, las directrices deben sensibilizar a los autores para que hagan visible la información relacionada con la reproducibilidad para los revisores y otros investigadores, por ejemplo, mediante una sección específica del artículo al estilo de la sección de “Disponibilidad de datos y software”. No hace falta empezar de cero. Las directrices para la redacción de artículos reproducibles (Nüst *et al.*, 2019a; Granell, 2019) están a disposición de todas las asociaciones y comunidades científicas que editen publicaciones periódicas o que promueven conferencias para que las personalicen a sus particularidades, con el fin de promover la investigación reproducible entre la comunidad de autores. Cabe destacar aquí el trabajo de Liu y Salganik (2019) que describen con sumo detalle su experiencia como editores de un número especial centrado únicamente en artículos reproducibles. Los retos, problemas, y las recomendaciones que plantean los autores son sin duda una guía práctica de gran ayuda para editores y organizadores de conferencias.

Relacionado con las directrices, las asociaciones tienen a su disposición otros elementos que puede llevar a cabo sin demasiado esfuerzo, pero con un impacto considerable en cuanto a la motivación y reconocimiento de la labor de los autores de artículos reproducibles. Uno de estos elementos son los premios para los mejores artículos reproducibles, similares a los tradicionales premios a los mejores artículos científicos. Esta práctica ya existe en algunas comunidades y conferencias (véase Nüst *et al.*, 2018), y poco a poco se va instaurando. En el caso de la conferencia AGILE 2020, se pondrá en marcha una sesión especial para los tres mejores trabajos reproducibles, con el fin de visibilizar y reconocer el esfuerzo de los autores y que sirvan de ejemplo para la comunidad AGILE.

Otro elemento son los “badges” o insignias. Kidwell *et al.* (2016) demuestran que las insignias han tenido un efecto positivo en la publicación de datos asociados a artículos publicados en la revista *Psychological Science*. Otras asociaciones de renombre como ACM (*Association for Computing Machinery*) disponen de un amplio rango de insignias para reconocer aspectos distintivos de los artículos, como por ejemplo la disponibilidad de datos en abierto o que el artículo sea reproducible (<https://www.acm.org/publications/policies/artifact-review-badging>). En el ámbito de las CIG, Nüst *et al.* (2019b) presentaron recientemente un sistema de insignias integrado en motores de búsqueda de literatura científica como Google Scholar.

Con estas básicas recomendaciones, dirigidas a distintos actores que forman parte del proceso de investigación científica, queremos destacar que la promoción de hábitos en investigación reproducible no es una odisea imposible. Se trata más bien de una cuestión de voluntad y que todos los actores implicados expresen su convencimiento por la instauración progresiva de estas prácticas. Los investigadores y grupos de investigación pueden hacer esfuerzos de forma individual, pero para que despegue

definitivamente la cultura de la ciencia abierta y reproducible, todos los demás actores, desde agencias de política científica hasta universidades y asociaciones científicas, deben apostar definitivamente por la investigación reproducible para que la próxima generación de investigadores e investigadoras desarrolle su trabajo de forma transparente, íntegra y honesta.

## Conclusiones

En este artículo hemos presentado nuestras experiencias, actividades y logros en la promoción de recomendaciones para la investigación reproducible, contextualizada a la comunidad AGILE. Destacamos sobre todo las directrices para autores de artículos reproducibles, un recurso disponible para cualquier investigador interesado en reproducibilidad para que pueda adoptarlas a su propio entorno, ya sea de forma individual, en su grupo de trabajo, departamento, universidad, e incluso en la comunidad científica de la cual es miembro.

Este artículo también es una llamada de la comunidad latinoamericana, para que haga suya los recursos descritos en este trabajo y, sobretodo, promocióne de forma activa la investigación reproducible en la CIG en su entorno local. Pero es también una llamada a colaboración, para que se establezcan sinergias entre la comunidad AGILE y asociaciones similares en la región latinoamericana. Necesitamos una apuesta clara por la ciencia abierta, la reproducción y replicación, así como por los datos y códigos en abierto y públicamente accesibles, como prevención a una crisis de reproducibilidad en nuestra disciplina.

## Agradecimientos

Los autores agradecen la predisposición de los miembros del Consejo, así como el apoyo en interés de la comunidad AGILE, con respeto a las acciones y recursos generados para la promoción de la investigación reproducible en AGILE. También agradecen la ayuda financiera de AGILE para la redacción de las directrices de artículos, reproducible (<https://bit.ly/2vxlspv>). Carlos Granell ha sido parcialmente financiado por el programa Ramón y Cajal del Ministerio de Ciencia e Innovación de España (referencia RYC-2014-16913). Daniel Nüst ha sido financiado por el proyecto “Opening Reproducible Research” (<https://o2r.info>) de la Fundación de Investigación Alemana (DFG, referencia PE1632/17-1).

## Bibliografía

Ballari, D. (2018). “Reproducibilidad científica: ¿Qué es y por qué debemos interesarnos en geo-ciencias?” *Revista Cartográfica*, 97: 147-155. <https://doi.org/10.35424/rcar.v0i97.179>

- Barba, L.A. (2018). "Terminologies for reproducible research", *ArXiv*, preprint arXiv: 1802.03311.
- Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J. & Munafò, M.R. (2013). "Power failure: why small sample size undermines the reliability of neuroscience", *Nature Reviews Neuroscience*, 14(5): 365-376.
- Cerutti, V., Bellman, C., Both, A., Duckham, M., Jenny, B., Lemmens, R.L.G. & Ostermann, F.O. (2019). "Improving the reproducibility of geospatial scientific workflows: the use of geosocial media in facilitating disaster response", *Journal of Spatial Science*. <https://doi.org/10.1080/14498596.2019.1654944>.
- Crüwell, S., van Doorn, J., Etz, A., Makel, M.C., Moshontz, H., Niebaum, J.C. & Schulte-Mecklenbeck, M. (2018, November 16). "7 Easy Steps to Open Science: An Annotated Reading List". <https://doi.org/10.31234/osf.io/cfzyx>.
- FORRT (2019, December 13). Introducing a Framework for Open and Reproducible Research Training (FORRT)". <https://doi.org/10.31219/osf.io/bnh7p>.
- Gahegan, M. (2019). "Reproducible Geocomputation: an open or shut case?", *Geo-Computation 2019*, 1-4 pp. <https://doi.org/10.17608/k6.auckland.9870089.v1>
- Granell, C., Nüst, D., Ostermann, F.O. & Sileryte, R. (2018). "Reproducible Research is like riding a bike", *PeerJ Preprints*, 6: e27216v1. <https://doi.org/10.7287/peerj.preprints.27216v1>
- Granell, C. (2019). "Directrices para artículos reproducibles". <https://doi.org/10.17605/OSF.IO/MF9BE>. Traducción en castellano del recurso <https://doi.org/10.17605/OSF.IO/CB7Z8>
- Hofer, B., Broman, K.W., Granell, C., Graser, A., Hettne, K., Daniel Nüst, D. & Teperek, M. (2019). "Reproducible Publications at AGILE Conferences – Proposed Guidelines for Authors and Reviewers", en Kyriakidis, P., Hadjimitsis, D., Skarlatos, D. & Mansourian, A. (eds.), *Accepted Short Papers and Posters from the 22nd AGILE Conference on Geo-information Science*, Limassol, Chipre, Editorial, Stichting AGILE.
- Ioannidis, J.P.A. (2005). "Why most published research findings are false", *PLOS Medicine*, 2(8): 124.
- Ioannidis, J.P.A., Stanley & T.D., Doucouliagos, H. (2017). "The power of bias in economics research", *The Economic Journal*, 127(605), F236-F265.
- Kedron, P. (2019). "Can Reproducible and replicable research facilitate causal explanation in Geography?", en Kyriakidis, P., Hadjimitsis, D., Skarlatos, D. & Mansourian, A. (eds.) (2019). *Accepted Short Papers and Posters from the 22nd AGILE Conference on Geo-information Science*. Limassol, Chipre. Editorial, Stichting AGILE.
- Kedron, P., Frazier, A.E., Trgovac, A.B., Nelson, T. & Fotheringham, A.S. (2019). "Reproducibility and Replicability in Geographical Analysis", *Geographical Analysis*. <https://doi.org/10.1111/gean.12221>.

- Kidwell, M.C., Lazarević, L.B., Baranski, E., Hardwicke, T.E., Piechowski, S., Falkenberg, L.-S. & Nosek, B.A. (2016). “Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency”, *PLOS Biology*, 14(5), e1002456. <https://doi.org/10.1371/journal.pbio.1002456>
- Konkol, M., Kray, C. & Pfeiffer, M. (2019). “Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study”, *International Journal of Geographical Information Science*, 33(2): 408-429. <https://doi.org/10.1080/13658816.2018.1508687>.
- Konkol, M., Nüst, D. & Goulier, L. (2020). “Publishing computational research -- A review of infrastructures for reproducible and transparent scholarly communication”, *ArXiv*, preprint arXiv: 2001.00484.
- Kray, C., Pebesma, E., Konkol, M., & Nüst, D. (2019). “Reproducible Research in Geoinformatics: Concepts, Challenges and Benefits”, en Timpf S. *et al.* (eds). *14th International Conference on Spatial Information Theory (COSIT 2019)*, (8:1-8:13). <https://doi.org/10.4230/LIPICs.COSIT.2019.8>
- Liu, D.M. & Salganik, M.J. (2019). “Successes and Struggles with Computational Reproducibility: Lessons from the Fragile Families Challenge”, *Socius: Sociological Research for a Dynamic World*, 5, 237802311984980. <https://doi.org/10.1177/2378023119849803>.
- Millman, K.J. & Pérez, F. (2014). “Developing open-source scientific practice”, en Stodden, V., Leisch, F. & Peng, R.D. (eds.). *Implementing Reproducible Research*, Editorial Chapman and Hall/CRC, pp. 169-204.
- Nüst, D., Konkol, M., Schutzeichel, M., Pebesma, E., Kray, C., Przybytzin, H. & Lorenz, J. (2017). “Opening the Publication Process with Executable Research Compendia”, *D-Lib Magazine*, 23(1-2). <https://doi.org/10.1045/january2017-nuest>.
- Nüst, D. (2018). Reproducibility Package for “Reproducible research and GIScience: an evaluation using AGILE conference papers” (Version 6-fixed) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.2575062>.
- Nüst, D., Granell, C., Hofer, B., Konkol, M., Ostermann, F.O., Sileryte, R. & Cerutti, V. (2018). “Reproducible research and GIScience: an evaluation using AGILE conference papers”, *PeerJ*, 6:e5072. <https://doi.org/10.7717/peerj.5072>
- Nüst, D., Ostermann, F.O, Sileryte, R., Hofer, B., Granell, C. Teperek, M. & Hettne, K. (2019a). *AGILE Reproducible Paper Guidelines*, versión 2, julio 2019. <https://doi.org/10.17605/OSF.IO/CB7Z8>
- Nüst, D., Lohoff, L., Einfeldt, L., Gavish, N., Götza, M., Shahzeib Tariq Jaswal, S.T. & van Eek, A. (2019b). “Guerrilla Badges for Reproducible Geospatial Data Science”, en Kyriakidis, P., Hadjimitsis, D., Skarlatos, D. & Mansourian, A. (eds.), *Accepted Short Papers and Posters from the 22nd AGILE Conference on Geoinformation Science*, Limassol, Chipre, Editorial: Stichting AGILE. <https://doi.org/10.31223/osf.io/xtsqh>

- Nüst, D. *et al.* (2020). Reproducibility Review at AGILE 2020 Conference. Recuperado de <https://osf.io/vq34n/>
- Ostermann, F.O. & Granell, C. (2017). "Advancing science with VGI: reproducibility and replicability of recent studies using VGI", *Transactions in GIS*, 21(2): 224-237. <https://doi.org/10.1111/tgis.12195>.
- Richardson, D.B., Kwan, M.-P., Alter, G. & McKendry, J.E. (2015). "Replication of scientific research: addressing geoprivacy, confidentiality, and data sharing challenges in geospatial research", *Annals of GIS*, 21(2): 101-110. <https://doi.org/10.1080/19475683.2015.1027792>.
- Rule, A., Birmingham, A., Zuniga, C., Altintas, I., Huang, S. C., Knight, R. & Rose, P.W. (2019). "Ten simple rules for writing and sharing computational analyses in Jupyter Notebooks", in *PLoS computational biology*, 15(7). <https://doi.org/10.1371/journal.pcbi.1007007>
- Sandve, G.K., Nekrutenko, A., Taylor, J. & Hovig, E. (2013). "Ten simple rules for reproducible computational research", *PLoS computational biology*, 9(10). <https://doi.org/10.1371/journal.pcbi.1003285>
- Singleton, A. & Arribas-Bel, D. (2019). "Geographic Data Science", *Geographical Analysis*. <https://doi.org/10.1111/gean.12194>
- Stodden, V., Leisch, F. & Peng, R.D. (eds) (2014). *Implementing Reproducible Research*, Boca Ratón, FL, EEUU. CRC Press.
- Stark, P. B. (2018). "Before reproducibility must come preproducibility", *Nature*, 557(7706): 613-614.
- Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L. & Teal, T.K. (2017). "Good enough practices in scientific computing", *PLoS computational biology*, 13(6). <https://doi.org/10.1371/journal.pcbi.1005510>

# La situación de la producción científica latinoamericana en el área de la Ciencia de Información Geográfica

*The state of affairs of the latin-american GIScience  
scientific production*

Carlos López-Vázquez\*  
Miguel-Ángel Bernabé-Poveda\*\*

*Recibido 6 de enero de 2020; aceptado 30 de enero de 2020*

## Resumen

Se ha planteado el objetivo para la *Revista Cartográfica* de captar un mínimo de 45 artículos por año, con el fin de ser considerado su ingreso al catálogo SciELO. Para valorar lo razonable del objetivo se ha intentado por primera vez estimar el volumen de la producción científica de los países miembros del Instituto Panamericano de Geografía e Historia en el área de la Ciencia de la Información Geográfica. Se realizó un análisis bibliométrico de los artículos publicados en un conjunto representativo de revistas internacionales, filtrando aquellos cuyos autores son de países de la región. Se excluyó a los Estados Unidos, en razón de su peso relativo en el área. El objetivo principal es duplicar la recepción de trabajos para publicar ser publicados en la *Revista Cartográfica*, aunque pareciera difícil de cumplir en un corto plazo. Como alternativa, se considera interesar a autores fuera de la región, quienes típicamente toman primero en cuenta la categorización de un revista, misma que no se ha logrado con *Revista Cartográfica*, volviéndose un círculo vicioso de difícil solución. Subs-

\* Universidad ORT URUGUAY, Uruguay, correos electrónicos: [carloslopez@uni.ort.edu.uy](mailto:carloslopez@uni.ort.edu.uy); [carlos.lopez@ieec.org](mailto:carlos.lopez@ieec.org). ORCID: <http://orcid.org/0000-0002-8444-1510>

\*\* Universidad Nacional de Catamarca, Argentina, correo electrónico: [ma.bernabe@gmail.com](mailto:ma.bernabe@gmail.com). ORCID: <https://orcid.org/0000-0003-4502-3524>

diariamente se ha realizado un análisis discriminado por países que puede ser utilizado en el futuro para identificar grupos y líneas activas de investigación en la región.

Palabras clave: *Ciencia de la Información Geográfica, cienciometría, IPGH, SciELO, Latindex.*

## Abstract

The goal for *Revista Cartográfica* to publish a minimum of 45 articles per year has been raised in order to be considered to enter the SciELO catalog. To assess the feasibility of the objective, an estimate of the volume of scientific production of the Pan American Institute of Geography and History member countries in the area of Geographic Information Science has been attempted for the first time. A bibliometric analysis of the articles published in a representative set of international journals was performed, filtering those whose authors are from countries in the region. The US was excluded because of its high relative weight in this scientific area. The main conclusion is that the goal of capturing part of these works and thus more than duplicating the material published today in *Revista Cartográfica* seems difficult to achieve in the short term. It will be necessary to interest authors from outside the region, who typically first take into account the journal's categorization, a categorization that has not yet been achieved, thus forming a vicious circle of difficult solution. Subsidiarily, a country-discriminated analysis has been carried out that could be used in future works to identify active research groups and research lines in the region.

Key words: *Geographic Information Science, PAIGH, SciELO, Latindex.*

## Introducción

El Instituto Panamericano de Geografía e Historia (IPGH), es un organismo gubernamental panamericano, creado en 1928 y vinculado desde 1949 con la Organización de Estados Americanos (OEA), en carácter de organismo especializado. Su actividad técnica se organiza en cuatro comisiones, a saber Cartografía (creada en 1941), Historia (1946), Geografía (1946) y Geofísica (1969). A falta de una definición expresa de su área de competencia, la Comisión de Cartografía del Instituto ha asumido que las áreas hoy consideradas dentro de la Ciencia de la Información Geográfica (CIG en lo sucesivo) serían las que al presente mejor la describiría, sin perjuicio de reconocer que hay una intersección importante pero parcial con otras ramas de la ciencia bien establecidas como la Geodesia.

Según reza su estatuto, el IPGH tiene por objeto "...Fomentar, coordinar y difundir los estudios cartográficos, geofísicos, geográficos e históricos y los relativos a las ciencias afines de interés para América...". Este trabajo apunta a apoyar la misión de difusión, la cual en el área específica de la CIG se lleva adelante mediante la *Revista Cartográfica* (RCA en lo sucesivo), creada en 1952 como revista anual. Desde esos



comienzos, y en un proceso de mejora continua, se fueron ajustando métodos y definiciones y finalmente se pasó en 2016 a dos números anuales (unos 18 a 20 artículos por año) permitiendo así que RCA sea incluida en Latindex (<http://www.latindex.org>). Ahora se está considerando cumplir con los requisitos para ingresar en SciELO (<http://www.scielo.org>), un sistema colectivo iberoamericano para publicación electrónica cooperativa de revistas científicas en Internet que tuvo su origen en Brasil. Tanto Latindex como SciELO son catálogos colectivos que recogen los nombres de publicaciones periódicas de la región que cumplen ciertas características. Básicamente se plantean requisitos de tipo editorial que, aunque no implican al contenido, obligan al editor a garantizar un grado de calidad en la operativa. Hay versiones mundiales equivalentes: por ejemplo, el Journal Citation Reports (JCR en lo sucesivo), SCOPUS, SCIMAGO, etc. Una utilidad (tal vez secundaria) de estos catálogos es ofrecer criterios de forma independiente para valorar la seriedad de una revista, y por lo tanto medir indirectamente el mérito de publicar en ellas. Eso es utilizado por los organismos de promoción de la ciencia en cada país para asignar fondos, promocionar investigadores, definir áreas prioritarias, etc. Estos organismos asignan menos puntaje por publicar en una revista de la que se ignora si tiene un Comité Editorial o si lo tiene no se conocen sus integrantes, tiene periodicidad desconocida, etc., todos aspectos muy técnicos, por lo que para su dilucidación se recurre a servicios de catalogación especializados. No es el objetivo aquí discutir los requerimientos de cada uno de estos catálogos; sí se puede decir que en la región latinoamericana se considera que Latindex es un nivel mínimo a cumplir, y que SciELO constituye un escalón superior. Una vez alcanzada esa clasificación regional, la siguiente etapa será ingresar a un catálogo internacional tipo JCR, SCOPUS, SCIMAGO o similar.

En lo sucesivo cuando se aluda a *la comunidad científica de la región* se estará aludiendo a *la comunidad científica de la región en la temática CIG*. Como se verá, esa comunidad no está claramente identificada, así como tampoco hay un criterio estricto para hacerlo. Se define la *región* como el conjunto de los países miembro del IPGH exceptuando Estados Unidos, por razones que se presentarán más adelante. Entre otros requisitos formales, para que una revista pueda ser considerada por SciELO, se necesita asegurar regularmente un volumen mínimo de artículos por año. Ese mínimo no es especificado centralmente, sino que cada país define cual será. En el caso de México (lugar donde se publica RCA) se ha manejado seguir el criterio de Brasil, que al presente especifica un mínimo de 45 artículos por año (recomendado 55) para revistas del área de Ciencias de la Tierra (Anon, 2017: 14). Ello es sensiblemente superior a la actual producción recogida en RCA que es de unos 20 artículos anuales repartidos en dos números. También la frecuencia debería ser modificada, recomendando publicar un número cada dos meses o como máximo tres meses, pero este aspecto no será considerado aquí.

La tasa de rechazo de RCA es del orden del 25%, en línea con lo habitual para este tipo de revistas. Lo crítico para lograr alcanzar 45 artículos anuales es asegurar un volumen superior y sostenible de trabajos enviados por los autores a su consideración. De ninguna manera se está considerando disminuir los requerimientos de aceptación. Esta investigación tuvo como propósito valorar la factibilidad de esta iniciativa en relación a los requisitos presentes de SciELO. Para ello debe estimarse el volumen total y distribución por países de la producción académica en el mundo que a) puede atribuirse al área de CIG y b) fue realizada por investigadores instalados en los Estados Miembros del IPGH. Como resultado secundario puede ofrecerse un diagnóstico del estado de situación de la CIG en la región.

Este artículo está organizado en varios apartados. En Antecedentes se analizarán trabajos similares, que en particular definen listados de revistas de interés para el área CIG. Posteriormente se discutirán los Datos y Métodos aplicados, sus limitaciones y características. A continuación se presentarán los Resultados, discriminando en producción por países y finalmente se presentarán las Conclusiones junto con un esbozo de los trabajos futuros.

### Antecedentes

La delimitación temática de la CIG (nombre introducido por Goodchild, 1992) ha sido abordada por varios autores pero aún no puede decirse que sea un tema cerrado (Di Biase *et al.* 2006; Prager y Plewe, 2009; Kemp, Kuhn y Brox, 2013; Reitsma, 2013; Blaschke y Merschdorf, 2014; Biljecki, 2016). Las ciencias nuevas son, en muchos casos, una mezcla de temas reclasificados de otras ciencias consolidadas más la incorporación de nuevos temas surgidos indirectamente de avances tecnológicos o de otros eventos externos (Xie *et al.*, 2017). Sin perjuicio de aceptar la existencia de revistas de ámbito general, las revistas específicas caracterizan a las ciencias establecidas (Yuan, 2017). Se asume implícitamente que el Editor y sus asesores filtran estrictamente los trabajos que se reciben de acuerdo con una línea establecida, y que esa línea además se mantiene en el tiempo. En esas hipótesis hay una asociación directa entre el área científica y un conjunto dado de revistas que cubren la temática. Debe dejarse constancia de que la irrupción de nuevos temas y soluciones, así como el desarrollo de nuevas tecnologías, para atender a viejos temas, pone en tensión ese modelo ya que, o bien hay que decidir si se extiende el ámbito temático de la revista, o bien si se aceptan trabajos de temáticas que *a priori* no hubieran sido de interés.

Para conocer qué revistas publican trabajos del ámbito de la CIG se analizaron varios trabajos:

- Apoyándose en expertos y siguiendo el proceso de encuestas Delphi, Caron *et al.* (2008) construyeron un primer listado de las revistas más importantes del área

CIG con el fin de orientar a los investigadores a dónde remitir sus trabajos. Existían trabajos previos metodológicamente similares, pero no contemplaban específicamente el área CIG. Mayormente (pero no exclusivamente) el listado incluye revistas en idioma inglés, sin restricción expresa a un ámbito geográfico. El listado de la Tabla 1, considera 46 revistas organizadas en cuatro grupos que los autores indican son de importancia decreciente. Las revistas tienen sede en Estados Unidos, Europa Occidental y Canadá. No son necesariamente especializadas en CIG sino que su presencia en el listado debe interpretarse como que habitualmente aceptan trabajos de CIG; este detalle es de particular importancia, y será retomado luego. La utilidad de este listado es ponderar el mérito de los autores al publicar en esas revistas.

- Kemp, Kuhn y Brox (2013) elaboraron también un listado de revistas de CIG utilizando un procedimiento Delphi similar al de Caron *et al.* (2008). La clasificación se organizó en tres grupos: a), que contenía los títulos líderes en el área, b) y c), cada uno con menor importancia que el previo. La diferencia entre estos dos últimos era debida a cumplir o no algunas características formales (comité de evaluación internacional o no; uso de al menos dos revisiones independientes o no, etc.). La encuesta se planteó a individuos con título de PhD y que hubieran publicado en al menos cinco revistas diferentes del área. La misma fue contestada mayoritariamente por individuos de Europa y Estados Unidos, seguramente sesgando así los resultados. Los autores ponen énfasis en destacar que el ranking resultante no debe tomarse como definitivo sino como orientativo, a falta de mejores y más profundos trabajos que lo validen. Como elemento adicional al de revistas, los autores ofrecen un ranking de Conferencias.
- En un interesante enfoque Scarletto (2014) construyó un listado de revistas de CIG a partir de las citas encontradas en solamente cuatro de ellas en el período 2008-2010: (1) *International Journal of Geographical Information Science* (IJGIS en lo sucesivo), (2) *Annals of the Association of American Geographers*, (3) *Cartography and Geographic Information Science*, and (4) *Cartographic Journal*. Los números entre paréntesis corresponden a la categoría asignada a ellas por Caron *et al.* (2008). Tras restringirlo a publicaciones en inglés el resultado fue un listado de 2 070 revistas. Tras analizar las citas, la autora concluye que las revistas pueden clasificarse en tres zonas, contando la más influyente con 23 títulos y la segunda con 190. El principal mérito que tiene este trabajo es que, una vez elegidas las cuatro revistas de base, el proceso para elegir los títulos más influyentes es objetivo, a diferencia del seguido por Caron *et al.* (2008) o Kemp, Kuhn, y Brox (2013). Desafortunadamente se descartaron revistas por no publicarse en idioma inglés, y se utilizó apenas un período de tres años para construir el ranking. El solape con el listado de Caron *et al.* (2008) es importante, como reconocen Melo y Queiroz (2019).

**Tabla 1**  
**Producción acumulada y anualizada de países miembros del IPGH**  
**(excepto Estados Unidos) discriminada por revista según listado de Caron *et al.* (2008)**

Revista		Promedio anualizado		Acumulado hasta 2019		
		09-14	14-19	2009	2014	2019
Primera categoría	International Journal of Geographic Information Science	2.0	4.0	15	25	45
	International Journal of Remote Sensing	5.2	37.0	271	297	482
	Photogrammetric Engineering & Remote Sensing	2.0	2.2	24	34	45
	Computers and Geosciences	7.6	19.6	86	124	222
	Transactions in GIS	0.8	3.2	9	13	29
	Geoinformatica	2.2	1.6	10	21	29
	Geomatica	1.2	0.0	0	6	6
Segunda categoría	Cartography and Geographic Information Sciences	0.2	0.8	2	3	7
	Environment and Planning B	1.0	0.4	0	5	7
	IEEE Transactions Geoscience and Remote Sensing	2.0	14.6	52	62	135
	Remote Sensing of Environment	10.4	25.2	106	158	284
	Computers, Environment and Urban Systems	0.8	2.6	8	12	25
	Annals of the Association of American Geographers	0.2	1.2	12	13	19
	URISA Journal					
	Environment and Planning A	1.0	3.2	8	13	29
	Landscape Ecology	5.2	14.2	29	55	126
	Cartographica	0.0	0.4	1	1	3
	Journal of Geographic Information and Decision Analysis					
	Marine Geodesy	0.8	0.4	6	10	12
	Revue Internationale de Géomatique					
	Journal of Geographical Systems	0.0	1.4	0	0	7
Professional Geographer	0.6	1.4	11	14	21	
Tercera categoría	Transactions of the Institute of British Geographers	0.2	0.4	1	2	4
	Earth Surface Processes and Landforms	3.8	11.2	17	36	92
	Progress in Human Geography	0.2	0.6	3	4	7
	Spatial Cognition and Computation	0.2	0.8	3	4	8
	Mapping Sciences and Remote Sensing	0.0	0.0	4	4	4
	Canadian Journal of Remote Sensing	0.6	2.4	40	43	55
	Cartographic Journal	0.6	1.6	1	4	12
	Canadian Geographer	0.2	0.4	1	2	4
	Geoscience Canada	0.0	0.6	2	2	5
	Cybergeo	1.6	2.2	6	14	25
	International Journal of Urban and Regional Research	2.2	3.2	29	40	56
	Applied Geography	7.2	16.0	8	44	124
	Surveying and Land Information Science	0.0	0.0	1	1	1
	Earth Observation and Remote Sensing	0.0	0.0	0	0	0
	Cartographic Perspectives	0.2	0.0	0	1	1
	Acta Cartographica					
	Remote Sensing Reviews	0.0	0.0	4	4	4
	Espace Géographique	0.2	0.2	7	8	9
	Geofocus					
	Mappemonde	0.0	0.0	7	7	7
Geographical Analysis	1.0	1.8	0	5	14	
Geocarto International	0.0	5.4	14	14	41	
IV	Geomatique Suisse					
	Géospacial Solutions	0.0	0.0	3	3	3
<b>Totales</b>		<b>61.4</b>	<b>180</b>	<b>801</b>	<b>1108</b>	<b>2008</b>

**Fuente:** Elaboración propia en base a SCOPUS.

Nota: Si no hay artículos pero la revista está en SCOPUS se le indica con un cero, y se deja la celda en blanco cuando la revista no está catalogada en SCOPUS.

- En un trabajo reciente, Biljecki (2016) hace un estudio cuantitativo (De Solla Price, 1978) específico a revistas de CIG. Del total de 20 revistas que finalmente considera, 12 están en el listado de Caron *et al.* (2008). En ese proceso señala varias dificultades. Una primera es que el autor considera que como aún no hay una definición precisa del ámbito de acción de la CIG ello lleva a que artículos de estas temáticas aparezcan en revistas de áreas relacionadas pero que no pueden clasificarse como “revistas de CIG”. Estas incorporan trabajos de CIG pero no constituyen su ámbito exclusivo. Entre otros, los resultados presentados incluyen caracterizar la producción discriminada por país así como por revista. El autor finalmente señala que el 5% de los países son responsables del 76% de la producción científica en el área. En cuanto a la distribución por países, Estados Unidos es dominante (31.3% del total) y de la región el único que aparece entre los primeros 15 es Brasil, con el 1.1% del total. Finalmente, el autor presenta resultados de producción bibliográfica en relación a la población y al PBI.

Existen trabajos estrechamente relacionados con los previos, pero que ignoran la distinción entre CIG y SIG (Sistemas de Información Geográfica).

- Tian *et al.* (2008) realizaron un extenso trabajo cubriendo artículos que contengan “GIS” o “Geographical Information Systems” publicados en el período 1997-2006. La base de datos utilizada fue la *Science Citation Index* (SCI) suministrada por el *Institute for Scientific Information (ISI) Web of Science*. Los cinco países más productivos para SIG son Estados Unidos, Reino Unido, Canadá, Alemania y China; el primero tiene mucha distancia con los siguientes, por eso se ha considerado el trabajo por separado en el resto del trabajo. En relación a las revistas, señalan que el 15% del total de trabajos fueron publicados en 10 de ellas (de un total de 1 918), siendo IJGIS la más activa.
- Melo y Queiroz (2019) analizan artículos publicados en un conjunto de 20 revistas en el período 2006-2017, constituyendo esas 20 revistas una parte de las seleccionadas por Caron *et al.* (2008) y por Scarletto (2014). El criterio de selección ha sido que publiquen trabajos del área CIG o que mencionan a los SIG. Los autores confirman que la producción académica mundial del área conjunta CIG más aquellos que trabajos que usan SIG ha crecido a un ritmo sostenido de 1.8% anual, oscilando entre 170 y 250 artículos al año. Ese valor es consistente con el reseñado por Tian *et al.* (2008) para el período previo 1997-2006 y con los resultados actualizados que se presentarán más adelante.
- Por último, y en un trabajo que podría considerarse como de ámbito regional (haciéndolo comparable con el presente) Li *et al.* (2016) analizan la producción china en cuatro revistas nacionales de topografía, geodesia, y áreas afines para el período 2003-2013. Del análisis de los 5 863 artículos con participación de

5 822 autores resulta que el 59% de los trabajos son del área Cartografía, Geodesia o SIG.

El uso de revistas específicas no es la única alternativa. Otra posibilidad sería delimitar la producción del área a través del uso de palabras clave, idea que Blaschke y Merschdorf (2014) recomiendan considerar con mucha precaución. Según ellos, los investigadores en CIG aún no han acordado una terminología precisa y los conceptos principales tal vez no están suficientemente establecidos, todo lo cual pone en duda la validez práctica de este enfoque.

### Datos y métodos

En lo sucesivo se asumirá que la producción mundial en el área CIG está limitada a lo publicado en revistas del listado especificado por Caron *et al.* (2008). Ello es válido a los fines de acotar el volumen de producción académica regional por la que podría aspirar a competir RCA. Como un segundo criterio interesaba identificar aquellos trabajos publicados en esas revistas por autores que establecían como afiliación a alguna institución localizada en un país miembro del IPGH. Ambos criterios de búsqueda son algo arbitrarios y tienen varios inconvenientes:

- El hecho de publicar en esas revistas no implica automáticamente que la temática del artículo sea de CIG. Por lo tanto, lo que se obtendrá es una estimación optimista de la producción de la región en el tema CIG.
- No se puede valorar fácilmente la importancia de la contribución de los nacionales de los países miembros a un artículo: en algunos casos serán los principales responsables pero en casos extremos quizá aparecen como coautores simplemente por haber suministrado los datos para el artículo.
- Un tercer inconveniente, que quedará claro al discriminar la producción por países de la región, es que trabajos con autores de dos o más diferentes países miembros serán contabilizados tantas veces como corresponda, lo que exagera la producción conjunta. En contrapartida, es habitual que autores de la región realicen estudios en países de fuera de ella y publiquen con la filiación accidental que posean, siendo invisible su relación (si la hubiera) con el país de origen y por tanto reduciendo la producción regional. Se asumirá que ambas desviaciones se compensan mutuamente.
- Por último, debe señalarse que para asignar a un país el mérito de una publicación hay dos criterios. Podría usarse la última filiación conocida del autor, y atribuir a ese país toda la producción del mismo. Ello reconocería que la capacidad de producción científica está con el autor. Otro criterio es razonar que la publicación se dio en un cierto contexto, con recursos aportados por el país y que de alguna forma

los mide. En ese caso la filiación se mantiene tal cual aparece en la publicación original, y es el criterio que se adoptará aquí.

Las revistas fueron identificadas a través de su(s) ISSN. La información se extrajo de la base SCOPUS, la cual es muy extensa pero al presente (pasados más de 10 años) todavía no incluye una media docena de las revistas del listado de Caron *et al.* (2008). Entre las ausentes en particular interesa vivamente el caso de *Geofocus International Review of GI Science and Technology*, la cual a la fecha (2019) sí está incluida en el listado de *Journal Citation Reports* (JCR).

*Geofocus* acepta artículos en castellano, por lo que su ausencia seguramente tiene un impacto no despreciable en las estimaciones de la producción regional en CIG a realizar en este trabajo. Esta revista, que tiene una media de 18 artículos por año divididos en dos números anuales, que la mayoría de sus artículos están en castellano y que, como se ha anotado antes, está incluida en JCR pero no en SciELO, es potencialmente un competidor importante para RCA. Los autores que podrían publicar en castellano en RCA (que además acepta trabajos en inglés, francés y portugués) probablemente prefieran considerar primero hacerlo en *Geofocus* al estar catalogada en un listado de prestigio como JCR.

En su artículo, Caron *et al.* (2008) clasificaron las revistas en cuatro categorías de acuerdo a su relevancia académica, reflejada en los colores de la Tabla 1. Puesto que las conclusiones de este trabajo se extraerán de lo publicado en ese conjunto de revistas, es importante comentar lo dispar del volumen publicado en alguna de ellas. Por ejemplo, en la Tabla 1 puede verse que los autores de la región publicaron en *International Journal of Remote Sensing* 10 veces más de lo que hicieron en *IJGIS*. Si se asume que la primera no es una revista especializada en CIG pero la segunda sí, la consecuencia de considerar conjuntamente lo publicado en ambas es que la componente CIG se diluye irremediamente. Lo mismo ocurre con otras revistas de interés más general: *Remote Sensing of Environment* recibió más de seis veces artículos de autores regionales que *IJGIS*.

### **Caracterización de la producción regional**

La caracterización de la capacidad productiva de la región en el área CIG puede hacerse de tres maneras:

1. Identificando los autores CIG de la región y caracterizando su producción en cualquier revista o,
2. localizando los artículos CIG basado en su contenido pero publicados en cualquier revista y reteniendo aquellos con autores de la región o,
3. identificando un conjunto de revistas representativas y analizando la producción con autores de la región. Este fue el procedimiento finalmente seguido.

Aunque las alternativas (a) y (b) no son de implementación directa se discutirán más adelante. La consulta a SCOPUS se realizó en diciembre de 2019 de dos maneras:

- Para cada revista del listado, se contabilizó el número de trabajos con al menos un autor cuya filiación estuviera en alguno de los países miembros. Dado el (los) ISSN correspondientes de cada revista, manualmente se descargó un fichero en formato CSV y el resto del procesamiento se hizo localmente.
- Para un país miembro, se contabilizó el número de trabajos en cualquiera de las revistas del listado con al menos un autor con afiliación en ese país. Como se encontraron algunas inconsistencias con los datos previos se descartó utilizarlos.

El uso de un banco de datos dinámico como SCOPUS tiene sus dificultades, porque tras reiterar más tarde las mismas consultas algunos resultados no son repetibles. Se notó también que hubo revistas del listado que inicialmente no aparecían, luego aparecían y finalmente desaparecían nuevamente tiempo más tarde, por lo que se optó por utilizar únicamente la primera de las consultas mencionadas, e inferir de los datos descargados todo lo demás.

### ***Identificación de países***

Tanto SCOPUS como otras bases del mismo tipo ocasionalmente tienen problemas con la afiliación de los autores; la construcción de estos servicios rescata del pasado analógico datos para su procesamiento en el presente, y el proceso de digitalización tiene sus debilidades. Puede verse que por ejemplo muchos trabajos de las universidades de Estados Unidos simplemente no declaran explícitamente el país. También ocurre con algunas universidades europeas y con muchas empresas y organismos oficiales de investigación. Otros problemas son las diferentes formas de codificar los nombres: en ocasiones en los registros figuran abreviaturas como “Engl” y en otras “England”. Por lo tanto el volumen de producción por país debe ser considerado como conservador, pudiendo haber más trabajos que no fueron correctamente imputados.

### ***Identificación de autores***

Otro problema no menor es la dificultad que existe en asignar correctamente la autoría a un autor basado simplemente en el nombre, problema ya reconocido por Sun y Mason (2001) también trabajando en el área de CIG. Hasta hace muy poco tiempo no existía alternativa para bases como SCOPUS más que intentar interpretar el nombre (con sus abreviaturas) y asignarle a un único identificador la autoría de varios trabajos considerando a la vez la afiliación, tal vez la temática, etc.



El proceso funciona bien para nombres y apellidos infrecuentes, pero necesariamente falla en otros casos. Ello se agrava con las ambigüedades en la forma de auto-denominarse, que pueden cambiar con el paso del tiempo. La situación a futuro va en vías a mejorar drásticamente, en la medida que se populariza el uso de identificadores personales como ORCID ([www.orcid.org](http://www.orcid.org)) que resuelven estas ambigüedades (Lane, 2010; Haak *et al.*, 2012).

### ***Producción por país***

La estadística de producción por autor, institución, país, etc. puede hacerse de varias formas (Biljecki, 2016):

- En una, se cuentan las apariciones de un autor, en particular en los artículos publicados, independientemente del número de coautores. Ello tiene el inconveniente de que ciertos colectivos que acuerdan publicar conjuntamente están sobrerrepresentados en el total. Una primera contramedida para ello es contabilizar únicamente a los autores que están en primer lugar, bajo la hipótesis de que ese puesto se reserva para el autor más importante. No siempre es así, y en ciertos trabajos colectivos se señala explícitamente que el listado de autores está en orden alfabético. Aunque se especifique dentro del cuerpo del artículo, SCOPUS no lo contabiliza.
- Otro criterio es ponderar la aparición de un autor en particular con el inverso del número de coautores. La ventaja de este enfoque es que la suma de la contribución asignada a todos los autores iguala al número de artículos, pero igual que antes ignora cualquier diferencia entre los aportes de cada autor.
- Aún pueden combinarse los criterios anteriores: Kumar y Khormi (2012), asignaban un peso diferente entre el 0.75 para el primer autor y 0.25 para repartir equitativamente con el resto.

En el trabajo que aquí se presenta se ha ponderado el aporte a cada país teniendo en cuenta el inverso del número de coautores.

## **Resultados**

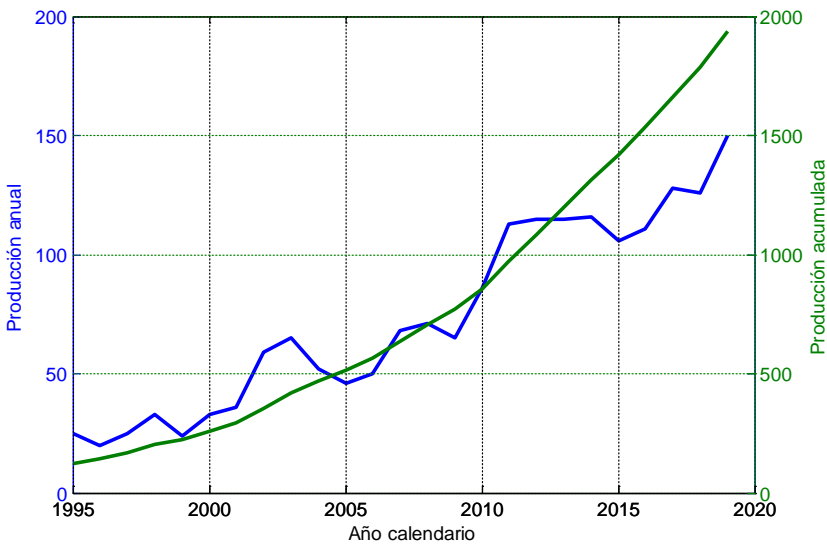
### ***Publicaciones de la región***

La Figura 1 muestra para el período 1995-2020 simultáneamente la producción anual y la acumulada de los países miembros del IPGH excluyendo a Estados Unidos. Por razones prácticas la búsqueda separó a ese país, ya que si bien pertenece al IPGH tiene un volumen de publicaciones que excede al menos un orden de magnitud (10 veces) el del resto de países del IPGH. Por los errores sistemáticos ya mencionados

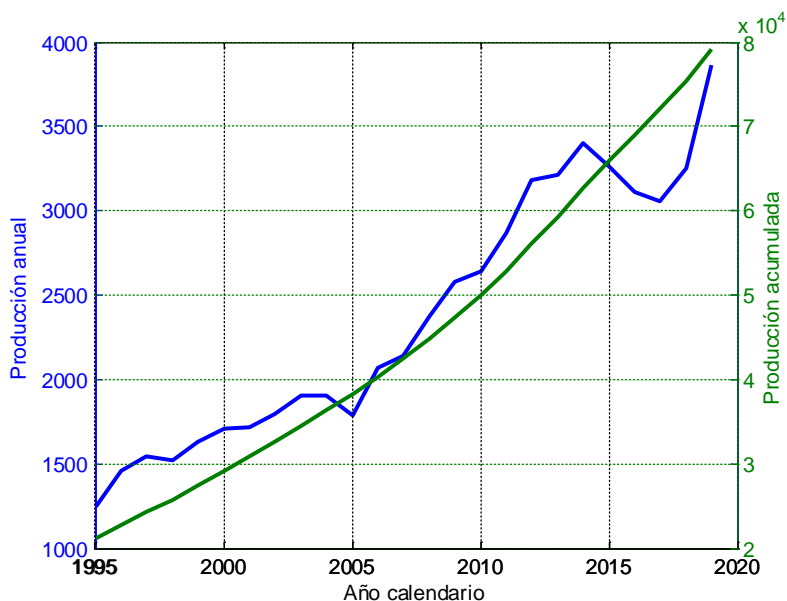
ésto es incluso un valor conservador, ya que hay muchos trabajos que siendo de Estados Unidos no fueron correctamente imputados a ese país. Independientemente de ello, y para los objetivos de este artículo, la consideración de incluir a dicho país, impediría valorar adecuadamente la situación de la región.

La evolución reciente de la situación en la región que se exhibe en la Figura 1 puede considerarse como positiva. Hay un sostenido aumento de la producción anual, que en 2011 excedió por primera vez los 100 trabajos y que actualmente alcanzó un pico de 150 por año. Por las razones ya expuestas no está claro qué parte de esa producción puede clasificarse como CIG, pero de igual manera el resultado sirve como cota superior. El aporte total *acumulado* de la región ha evolucionado fuertemente en el período, desde 99 artículos publicados hasta 1995 a cerca de 2 000 al presente. El crecimiento es más que lineal, lo que puede interpretarse como, que la comunidad está efectivamente en un proceso de expansión.

Para contrastar, se ha construido un gráfico equivalente (Figura 2) para la producción académica mundial anual en el mismo conjunto de revistas especificado. La curva muestra un crecimiento más próximo a lineal, correspondiente a una comunidad más consolidada. La producción es del orden de los 4 000 artículos por año, por lo que la región con 150 artículos en el mismo periodo año contribuye con menos del 4% del total.



**Figura 1** Producción anual y acumulada en el área CIG para países miembros del IPGH (a excepción de Estados Unidos).



**Figura 2** Producción mundial, anual y acumulada, en las mismas revistas consideradas en la Figura 1.

### *Possible valoración de países*

Los méritos relativos de cada país/autor/institución debido a sus publicaciones tomarán la forma de un puntaje, y los países con mayor puntaje son los más destacados. El problema es elegir la forma de calcular el puntaje, para lo que en este trabajo se han usado dos alternativas:

- Para los gráficos, se han contado simplemente el número de artículos.
- Para las tablas que se presentan a cada país se le asigna un puntaje por cada artículo publicado equivalente a  $1/n$ , siendo  $n$  el número de coautores del mismo.

En la Tabla 2 se presenta el resultado para la región. La producción por año calendario muestra demasiadas variaciones, por lo que se presentan resultados anualizados obtenidos para dos quinquenios sucesivos.

Con base en la información de la Tabla 2, en la Tabla 3 se han organizado bloques de países. Con el paso del tiempo Brasil sigue siendo responsable de cerca de la mitad de la producción regional, pero ha caído el peso relativo de los siguientes tres productores grandes y aumentado el del resto. No se presenta aquí, pero la contabilidad

directa sin ponderar por el número de coautores no arroja resultados muy diferentes, ya que típicamente los números se dividen entre tres sugiriendo una proporción entre autores locales y externos. Parece ser que esta manera de repartir disminuye mucho el mérito del grupo de los tres que le siguen a Brasil (que tal vez están mejor conectados y colaboran mucho con extranjeros) mientras que aumenta levemente el mérito de Brasil y bastante el todos los demás, ya que publican con menos autores de extrazona.

**Tabla 2**  
**Producción acumulada y anualizada discriminada por países miembros del IPGH**  
**(exceptuando a Estados Unidos)**

	Promedio anualizado		Acumulado hasta 2019		
	2009-2014	2014-2019	2009	2014	2019
Paraguay (PY)	0.00	0.00	0.09	0.09	0.09
Belize (BZ)	0.00	0.07	0.00	0.00	0.33
Guatemala (GT)	0.00	0.00	0.31	0.31	0.31
Haití (HT)	0.00	0.00	0.50	0.50	0.50
República Dominicana (DO)	0.00	0.03	1.08	1.08	1.25
El Salvador (SV)	0.05	0.07	0.50	0.75	1.12
Nicaragua (NI)	0.00	0.09	0.44	0.44	0.89
Honduras (HN)	0.03	0.07	0.14	0.31	0.64
Uruguay (UY)	0.46	0.28	4.46	6.75	8.17
Bolivia (BO)	0.17	0.19	3.25	4.08	5.02
Panamá (PA)	0.17	0.51	0.71	1.57	4.10
Venezuela (VE)	0.13	0.22	4.98	5.63	6.74
Costa Rica (CR)	0.43	0.06	6.04	8.17	8.45
Perú (PE)	0.40	1.10	2.08	4.09	9.58
Ecuador (EC)	0.46	0.89	4.16	6.45	10.89
Colombia (CO)	1.82	2.92	13.85	22.94	37.53
Chile (CH)	5.55	4.75	26.63	54.36	78.12
Argentina (AR)	3.16	3.21	37.87	53.69	69.77
México (MX)	6.34	4.97	57.41	89.12	113.95
Brasil (BR)	14.91	17.66	149.89	224.46	312.77
<b>Total IPGH</b> <b>(exceptuando Estados Unidos)</b>	<b>34.08</b>	<b>37.08</b>	<b>314.38</b>	<b>484.80</b>	<b>670.21</b>

**Fuente:** Elaboración propia en base a SCOPUS.

Nota: El conteo es ponderado, agregando a cada país una cuota parte de la unidad teniendo en cuenta el número de coautores.

**Tabla 3**  
**Producción anualizada y organizada por grupos, expresados en porcentaje a la producción total de la región**

	<i>Promedio anualizado</i>		<i>Acumulado hasta 2019</i>		
	<i>2009-2014</i>	<i>2014-2019</i>	<i>2009</i>	<i>2014</i>	<i>2019</i>
BZ, DO, GT, HT, HN, SV, NI y PY	0.24	0.89	3.06	3.48	5.13
BO, CO, CR, EC, PE, PA, UY y VE	11.83	16.61	39.53	59.68	90.48
AR, CH y MX	44.17	34.87	21.91	197.17	261.84
BR	43.76	47.63	149.89	224.46	312.77

**Fuente:** Elaboración propia en base a SCOPUS.

### **Análisis y discusión**

Hay varias observaciones que pueden hacerse con las diferentes tablas y figuras presentadas.

#### ***Producción***

De la Figura 1, puede inferirse que, con fuertes oscilaciones año con año, hay una sostenida aceleración (excediendo la aproximación lineal) en la producción global de la región que recién alcanzó en 2019 un pico histórico de 150 trabajos anuales. Estos resultados están repartidos muy desigualmente en términos de países. Las Tablas 2 y 3, que recogen los valores totales por país acumulados hasta el año de referencia, muestra que en la región (definida como IPGH sin Estados Unidos):

- BR ha publicado casi la mitad del total (46.7%)
- Tres países (AR, CH y MX) colectivamente han aportado un 39.0%
- Ocho países (BO, CO, CR, EC, PE, PA, UY y VE) cerca del 13.5%
- Otros ocho países (BZ, DO, GT, HT, HN, SV, NI y PY) el 0.8%

A modo de referencia en la Tabla 4 se hace una comparación internacional, en la que puede verse que, globalmente, cada cinco años hubo un 61% más de producción. En este caso se contaron simplemente artículos, sin penalizar por el número de coautores. Para darle perspectiva a este resultado debe notarse que España sola es sistemáticamente más que toda Iberoamérica, y a su vez Francia casi triplica a esa producción. Canadá casi cuadruplica la producción de la región, mientras se confirma que Estados Unidos genera del orden de 10 veces lo de la región.

Como notables desempeños en el período pueden señalarse el grupo de Ecuador y Colombia que han venido aumentando su ya importante presencia en cada quinquenio. Se puede agregar Panamá al listado de los que duplican, pero con menos peso porque su producción absoluta es menor.

**Tabla 4**  
**Producción acumulada y anualizada de países del IPGH (sin Estados Unidos) como bloque y comparación con España, Francia, Canadá y Estados Unidos**

	Promedio anualizado		Acumulados hasta este año		
	2009-2014	2014-2019	2009	2014	2019
IPGH sin					
Estados Unidos	108.8	133.4	781	1 325	1 992
España	134.8	167.2	792	1 466	2 302
Francia	217.4	273.4	2 725	3 812	5 179
Canadá	314.6	377.0	5 419	6 992	8 877
Estados Unidos	1 069.8	1 340.4	18 275	23 624	30 326

**Fuente:** Elaboración propia en base a SCOPUS.

### **Multiautoría**

Debe recordarse que la metodología seguida impide ver detalles; la mera presencia de un único autor con afiliación en un país ya define un aporte. Con esos datos se podría haber construido tablas de autores, instituciones o países, pero dados los objetivos de este trabajo solamente se presentan discriminadas por países. Los resultados exhibidos dependen de la forma de valorar los méritos. No se presentan aquí, pero si se contabiliza simplemente el número de artículos hay un dominio importante de autores de Brasil, seguido muy de lejos por México. La existencia de grupos grandes (tanto de coautores como alumnos) parece tener un efecto importante en este resultado. Si, con el fin de filtrar ese efecto, se contabiliza en cambio el número de artículos del país del primer autor, la tabla tiene también varios cambios.

### **Publicaciones de CIG en revistas no especializadas**

Algunas revistas del listado tienen un volumen de artículos mucho mayor que el resto. El área CIG no es tan activa como para que aporte miles de artículos al año, como parecieran sugerir estas estadísticas mundiales. Con algunas escasas excepciones (tal vez aquellas revistas que en su título mencionan CIG expresamente) en el resto y tras analizar algo más detenidamente los contenidos puede verse que es difícil sostener que por haber sido publicado allí un artículo trate siempre sobre CIG. El objetivo de

Caron *et al.* (2008) era sugerir revistas apropiadas para publicar sobre CIG, pero en general ello no implicaba una especialización temática de las mismas sino apenas una adecuada receptividad.

### ***Futuro***

Si bien en este trabajo se ha usado un listado prefijado de revistas, en futuros trabajos se podrían considerar al menos dos alternativas, no exentas de problemas:

- La más simple sería identificar primero a los autores y luego rastrear su producción en cualquier revista. Además de la dificultad en identificar por anticipado a los autores, este enfoque presume que a lo largo de su vida útil el autor se mantendrá siempre apegado al área CIG cosa que no es cierta.
- Otra posibilidad es clasificar a los artículos de CIG a través de su contenido, independientemente de la revista y del autor, buscando relaciones basadas en algún tipo de palabras clave. SCOPUS dispone como atributo las palabras suministradas por el autor, pero también otras que se generan automáticamente. Esta sería la línea más promisoría si se decidiera cambiar la actual alternativa, faltando especificar el criterio de búsqueda y detalles asociados.

### ***Nuevos aspectos a tener en cuenta***

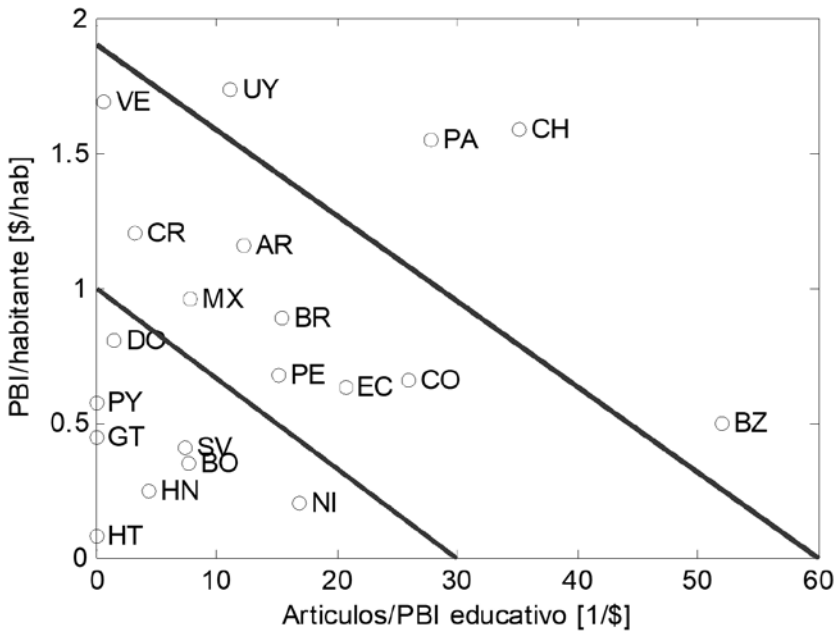
Si bien el objetivo de este trabajo no era establecer distinciones entre países, además de la producción fue posible explorar aspectos adicionales como la productividad académica per cápita o relacionarla con la inversión en investigación. Para ello se cuenta con información secundaria que se puede utilizar.

En la Figura 3 se relacionó la producción en relación al PBI y la población. En ella pueden distinguirse tres bandas: la de países con menor producción absoluta en la Tabla 2 se identifican aquí directamente con países que tienen simultáneamente un bajo PBI/habitante y una baja productividad en términos de artículos/PBI educativo. En estos casos la baja producción (medida por el número de artículos) ocurre simultáneamente con una baja productividad (medida como el cociente entre número de artículos y otra magnitud, como el número de habitantes o el PBI). En la Figura 3 se distingue luego una banda central, en la que coexisten países de gran tamaño y producción (Brasil, Argentina y México) junto con otros de tamaño medio (Venezuela, Perú y Colombia) así como algunos más pequeños (Costa Rica y Ecuador). Fuera de esta segunda región se encuentra Belice (con una productividad aparentemente alta pero asociada con un puñado de artículos), Chile y Panamá (que tienen alto PBI/habitante y además una alta productividad) y Uruguay, con alto PBI/habitante

pero relativamente baja productividad. Los datos corresponden a la producción ponderada y promediada del último quinquenio, mientras que los del PBI son de 2018 (excepto para Venezuela, para el que se usó el último dato disponible de 2014).

## Conclusiones

Si se deseara cuantificar la producción de la región en temas de CIG, la estrategia de contabilizar simplemente artículos publicados en un listado de revistas tiene limitaciones metodológicas importantes ya que la hipótesis de que esas revistas tocan exclusivamente temas de CIG no se sostiene. No hay estudios sistemáticos al respecto, y entre otras razones puede argumentarse que la propia área no está rigurosamente delimitada.



**Figura 3.** Productividad de los países en relación a la inversión educativa y al PBI por habitante.

Sin embargo, pueden establecerse relaciones que permiten dar cotas superiores para los valores de interés. La producción de España en ese listado de revistas es del orden de 170 artículos por año, mientras que lo publicado en *GeoFocus* (revista española especializada en el área) está en el eje de los 20 artículos también por año, constituyéndose en un  $20/(170+20)*100=10\%$  del total. Si se asume que el ámbito natural de operación de RCA está en los países miembro del IPGH puede establecerse



un paralelismo con la situación de *GeoFocus*, quien estando especializada en CIG capta un 10% del volumen de las revistas del listado. La producción de toda la región IPGH, calculada en las mismas condiciones, es de 134 artículos por año (en promedio). Si se admite que la parte de CIG es del orden del 10% del total, entonces se concluye que la fracción de CIG generada en la región sería de 13 a 14 artículos al año, que sumado a lo hoy publicado por RCA queda aún muy por debajo de lo requerido por SciELO. El cálculo previo es muy aproximado, y tiene carencias metodológicas ya expresadas. Además no hay que olvidar que RCA hoy no tiene el atractivo que tiene *GeoFocus*, que ya está catalogada en JCR.

La hipótesis de que los trabajos de CIG se publican únicamente en un listado de revistas especializadas no se sostiene. Una alternativa posible sería identificar autores de la región que trabajen en área CIG, y rastrear sus publicaciones personales en un conjunto abierto de revistas. Esto tiene el inconveniente de requerir la generación de la lista inicial de nombres, y un problema de fondo debido al hecho de que un investigador puede eventualmente entrar o salir del área temática sin previo aviso.

Como otra posibilidad, y siempre para futuros trabajos, se podría intentar determinar la pertenencia al área CIG no a nivel de revista sino de artículo, lo que una vez establecidos los criterios objetivos para ello llevará necesariamente a expandir la búsqueda fuera de un listado cerrado de revistas. Las alternativas que se imaginan para ello incluyen construir y utilizar un conjunto de palabras clave (propuestas por cada autor o generadas por SCOPUS), o siguiendo a Scarletto (2014) construir el conjunto con los trabajos que citan uno o varios emblemáticos que hayan sido previamente declarados como de CIG.

Puesto que el IPGH, como asociación de estados nacionales incluye prácticamente a toda América (con la excepción notable de Canadá, Cuba, Jamaica y los pequeños países del Caribe) puede considerarse que las estadísticas realizadas (en las que se excluyó deliberadamente a Estados Unidos a pesar de ser miembro) son representativas del total ya que la producción acumulada de Cuba y Jamaica hasta diciembre de 2019 era de apenas 12 y 14 trabajos respectivamente, mientras que el resto de la zona Caribe aportaba solamente 9 trabajos.

El desafío de incluir RCA en SciELO logrando los 45 artículos por año que el índice considera cantidad mínima, más que duplicando así el actual volumen debe descartarse una vez que se ha mostrado que el volumen de producción de toda la región en las revistas especializadas en CIG de más prestigio (y aun incluyendo a varias no tan prestigiosas) es del orden de 134 artículos por año. No es realista aspirar a captar 25 trabajos de esa producción de alto nivel, adicionales a los 20 ya publicados anualmente, con una revista que no aparece aún en catálogos de excelencia internacionales como SciELO, SCOPUS o JCR. Como ventaja frente a otras revistas internacionales puede argumentarse la flexibilidad de RCA, que recibe trabajos en cuatro idiomas incluyendo el castellano, pero no es la única revista del área CIG que lo hace.

Si se excluye a Estados Unidos, la comunidad latinoamericana activa en CIG no es hoy tan grande como para aportar en forma sostenida suficiente material para satisfacer ese volumen de producción. Basado simplemente en el crecimiento vegetativo, sin otras medidas de promoción ni cambios en las prioridades académicas, podría esperarse alcanzar ese volumen recién para 2050.

## Bibliografía

- Anónimo (2017). "Criterios, política y procedimientos para la admisión y la permanencia de revistas científicas en la Colección SciELO Brasil", Scientific Electronic Library Online, 31 pp. Recuperado de <https://tinyurl.com/uc7mdyp>
- Biljecki, F. (2016). "A scientometric analysis of selected GIScience journals", *International Journal of Geographical Information Science*, 30(7): 1302-1335. <https://doi.org/10.1080/13658816.2015.1130831>
- Blaschke, T. and Merschdorf, H. (2014). "Geographic information science as a multidisciplinary and multiparadigmatic field", *Cartography and Geographic Information Science*, 41(3): 196-213. <https://doi.org/10.1080/15230406.2014.905755>
- Caron, C.; Goyer, D., Roche, S. and Jaton, A. (2008). "GIScience journals ranking and evaluation: an international Delphi study", *Transactions in GIS*, 12(3): 293-321. <https://doi.org/10.1111/j.1467-9671.2008.01106.x>
- De Solla Price, D. (1978). *Scientometrics*, 1(1): 3-8, Editorial statements.
- Di Biase, D.; DeMers, M., Johnson, A., Kemp, K., Taylor Luck, A., Plewe, B. and Wentz, E. (eds) (2006). *Geographic information science and technology body of knowledge*. Washington, D.C., Association of American Geographers, 174 pp.
- Glänzel, W., & Schubert, A. (1985). "Price distribution: An exact formulation of Price's 'square root law'", *Scientometrics*, 7(3-6): 211-219. <https://doi.org/10.1007/BF02017147>
- Goodchild, M.F. (1992) "Geographical information science", *International Journal of Geographical Information Systems*, 6: 31-45. <https://doi.org/10.1080/02693799208901893>
- Haak, L.L.; Fenner, M., Paglione, L., Pentz, E. and Ratner, H. (2012) "ORCID: a system to uniquely identify researchers", *Learned Publishing*, 25: 259-264. <https://doi.org/10.1087/20120404>
- Kemp, K., Kuhn, W. and Brox, C. (2013). "Results of a survey to rate GIScience publication outlets", Technical report, 8 pp.
- Kumar, L. and Khormi, H.M. (2012). "Hotspots, research productivity and collaboration networks in remote sensing and GIS in Australia from 1991 to 2010", *Journal of Spatial Science*, 57(1): 101-113.
- Lane, J. (2010). "Let's make science metrics more scientific", *Nature*, 464: 488-489. <https://doi.org/10.1080/14498596.2012.686360>

- Li, J., Liu, H., Guo, W. and Yu, A. (2016). "Visual analysis based on the data of Chinese surveying and mapping journals", *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, III(4): 41-48.  
<https://doi.org/10.5194/isprs-annals-III-4-41-2016>
- Melo, A.V.F. and Queiroz, A.P. (2019). "Bibliometric Mapping of Papers on Geographical Information Systems (2007-2016)", *Bulletin of Geodetic Sciences*, 25(3): e20190015. <https://doi.org/10.1590/s1982-21702019000300015>
- Prager, S.D. and Plewe, B. (2009). "Assessment and Evaluation of GIScience Curriculum using the Geographic Information Science and Technology Body of Knowledge", *Journal of Geography in Higher Education*, 33:S1, S46-S69.  
<https://doi.org/10.1080/03098260903034012>
- Price, D.J. de Solla (1971). *Little Science-Big Science*, Columbia University Press, New York and London, 4th ed., 118 pp.
- Reitsma, F. (2013). "Revisiting the 'Is GIScience a science?' debate (or quite possibly scientific gerrymandering)", *International Journal of Geographical Information Science*, 27(2): 211-221. <https://doi.org/10.1080/13658816.2012.674529>
- Scarletto, E.A. (2014). "Mapping the Literature of GIS", *College & Research Libraries*, 75(2): 179-201.
- Sun, S. and Manson, S.M. (2011). "Social network analysis of the academic GIScience community", *The Professional Geographer*, 63: 18-33.  
<https://doi.org/10.5860/crl12-389>
- Tian, Y.G., Wen, C. and Hong, S. (2008). "Global scientific production on GIS research by bibliometric analysis from 1997-2006", *Journal of Informetrics*, 2: 65-74. <https://doi.org/10.1016/j.joi.2007.10.001>
- Xie, Y.; Efteliogl, E.; Ali, R.Y.; Tang, X.; Li, Y.; Doshi, R. and Shekhar, S. (2017). "Transdisciplinary Foundations of Geospatial Data Science", *ISPRS International Journal of Geo-Information*, 6(12), 24 pp.  
<https://doi.org/10.3390/ijgi6120395>
- Yuan, M. (2017). "30 years of IJGIS: the changing landscape of geographical information science and the road ahead", *International Journal of Geographical Information Science*, 31(3): 425-434.  
<https://doi.org/10.1080/13658816.2016.1236928>



# Sinfonía del Nuevo Mundo: el modelaje espacio-temporal de HGIS de las Indias, una infraestructura SIG para la América hispana borbónica\*

*The New World Symphony: the spatio-temporal modeling of HGIS of the Indies; a GIS infrastructure for Bourbon Hispanic*

Werner Stangl\*

Paul Stangl\*\*

*Recibido 2 de marzo de 2019; aceptado 25 de mayo de 2019*

## Resumen

Este artículo aborda las dificultades básicas del modelado de una infraestructura SIG para la historia de la Hispanoamérica borbónica, donde se busca reconstruir una docena de “niveles” territoriales y el desarrollo de poblaciones dentro de América. El trabajo pretende incluir la región comprendida desde California y Luisiana hasta la Patagonia, para un período de 108 años (1701-1808). Por la multiplicidad de incertidumbres en diferentes aspectos (espaciales, temporales, cualitativos), resulta particularmente difícil diseñar un modelo de las entidades territoriales y su cambio a través del tiempo. Se explica cómo hemos intentado aislar diferentes tipos de incertidumbres, creando cronologías parciales en lugares incluidos dentro de la base de datos. Concretamente, analizaremos la forma como fueron ensambladas las diferentes partes y se integraron a una geodatabase con una cronología de validez única, basada en

\* Este artículo se enmarca en el proyecto de las Indias HGIS, número P 26379-G18, apoyado por el Fondo Científico de Austria (FWF).

\* Yale University, CT, USA, correo electrónico: [werner.stangl@yale.edu](mailto:werner.stangl@yale.edu).  
ORCID: <https://orcid.org/0000-0002-7871-4201>.

\*\* Universidad de Graz, Austria, correo electrónico: [paul.stangl@edu.uni-graz.at](mailto:paul.stangl@edu.uni-graz.at).

búsquedas SQL sobre las cronologías parciales. Además, presentaremos el *workflow* implementado para posibilitar la conversión automática de simples tablas en *shape-files* a través del interfaz de nuestro sitio-web. A modo de conclusión, será presentado el *workflow* y un gráfico sinóptico del sistema que visualiza cómo las diferentes partes interactúan para formar una infraestructura estable para un SIG espacio-temporal, el cual fácilmente puede reimplementarse para otras áreas y períodos históricos.

Palabras clave: *SIG histórico, modelaje de base de datos, composiciones espacio-temporal, incertidumbre.*

### Abstract

This article outlines the basic difficulties for the modelling of a historical GIS data infrastructure on Bourbon Spanish America which aims at providing reconstructions of a dozen territorial “levels” and settlements across the double continent, from California and Louisiana to Patagonia, for a period of 108 years (1701-1808). Due to the manifold sources of uncertainty at many levels it is particularly demanding to successfully model the spatial embodiment of territorial entities and their change over time. In this text, we explain how we isolate different types of uncertainty and create different chronologies across the database. In more detail, we will show how we reassemble the different parts and integrate them into a geodatabase with a single chronology based on queries on the different partial chronologies. Also, we show how this workflow has been implemented to allow automatic creation of shapefiles from simple tables through the interface of our homepage. Instead of a narrative conclusion, we present a synoptic graphic visualizing how the different parts and processes provide a stable framework for a GIS data infrastructure covering a complete period of time —a framework that can be easily be re-implemented for different fields, regions, and periods.

Key words: *historical GIS, database modelling, space-time composites, uncertainty.*

### Introducción

Dentro de los estudios históricos, el uso de tecnologías SIG enfrenta un sinnúmero de desafíos y discusiones. Las polémicas, de carácter epistemológico y metodológico, han sido analizadas y ampliamente abordadas en algunos trabajos de carácter compilatorio (Gregory y Allen, 2007; Bodenhamer *et al.* 2010; Lünen y Travis, 2013). Más allá de estos debates, las dos problemáticas fundamentales para la construcción de una infraestructura geoespacial histórica son: 1) El modelado de los cambios en la “espacialidad” de las entidades históricas (territorios, vías de comunicación, o luga-

res, es decir, polígonos, líneas y puntos, entre otros) y; 2) La conservación de coherencia ante los distintos tipos de “incertidumbre” causada por fuentes históricas de densidad variable, incompletas, confusas y erróneas.

*HGIS de las Indias* es un sistema que ha sido calificado como un “atlas digital” para la historia de la América borbónica. Sin embargo, dentro de su propia lógica puede ser entendido como una infraestructura, un tipo de “contenedor espacio-temporal” que permitirá representar la espacialidad de los documentos producidos en la época y la producción historiográfica elaborada con base en estos. Para cumplir dicha función, el modelaje tiene que ser capaz de ofrecer una interpretación de la situación de cualquier documento o dato en un momento determinado. Para *HGIS de las Indias*, hemos elegido una granulación a nivel de año: cada fecha de validez en la base de datos se refiere a la situación a finales de año, por lo que situaciones inestables que existieron sólo por un par de meses no se representan. Como infraestructura, se reconstruye (con base en la granulación temporal referida) diversos tipos de territorios, como audiencias, obispados, provincias y partidos, así como 12 000 lugares de la época. También fue objeto de interés poder representar datos tabulares<sup>1</sup> con la localización o extensión correspondientes al dato. Este reto, como veremos, era el más complejo de cumplir.

Resulta evidente que nuestra interpretación de *límite*, más allá de las críticas que puede despertar el uso mismo del concepto, no es inmune a errores ante la gran extensión geográfica estudiada (desde Vancouver Island hasta la isla de Chiloé, y del Misisipi al Río de la Plata). La situación también resulta compleja porque existe una oscilante calidad y disponibilidad de descripciones y mapas de la época. En efecto, muchos autores han mezclado y confundido conceptos. Por ejemplo, resulta célebre el caso de Juan de Velasco (1842 tomo III: 33-34), quien describió la “Tenencia de los Pastos” (un concepto claramente vinculado con la administración) como una división perteneciente a la provincia de Popayán. En la lista de poblaciones que conformaban la “Tenencia” agrupó pueblos que cultural, étnica y lingüísticamente pertenecían a “los pastos” pero que, al encontrarse al sur del puente Rumichaca, en realidad eran parte del corregimiento de Ibarra, dentro de la provincia de Quito. Como consecuencia de tal error, los pueblos reagrupados por Velasco han sido mal descritos en cuadros demográficos de la provincia de Popayán y otros documentos de tipo administrativo para esa provincia (Herrera, 2009; Durán y Díaz, 2012).

Otros autores tienen todavía otras concepciones de qué era “la provincia de los pastos” (Cieza, 1554; Alcedo, 1786-1789). Dentro de tal interpretación, la provincia no se correspondía totalmente con las listas de colección de tributos y censos de población, pero tampoco era independiente de dicho eje de organización. Lo anterior ha provocado ciertos equívocos entre quienes tratan de reconstruir la etnografía de los

<sup>1</sup> Así como posiblemente referencias geográficas en documentos XML, pero para esto no hay aplicación concreta.

pastos, como por ejemplo Schortman y Urban (1992) o Cárdenas-Arroyo (1996). Las razones de las discrepancias entre las fuentes no son explícitas, y sólo se detectan a través de una contrastación profunda de la documentación disponible. Así, al inicio de la labor reconstructiva, teníamos una interpretación deficiente y poco clara de lo que era la provincia de “Los Pastos”, visión que poco a poco fue precisándose. Sin embargo, resulta concebible que aún no se haya llegado al final del proceso.

De hecho, en un censo de 1797 aparece por primera y única vez la población de “El Castigo” como parte de Los Pastos. Se trata de un palenque de cimarrones (esclavos fugitivos) que mantuvo una relación difícil con las autoridades de Popayán. Simplemente no disponemos de información sobre si el área de El Castigo desde siempre había sido entendida como parte de Los Pastos o si su inclusión en el censo ocurrió por un cambio territorial de “Los Pastos”. Si la segunda hipótesis es correcta, tampoco es posible establecer si anteriormente era parte de Barbacoas. No obstante, tenemos el siguiente indicio: el nombre de El Castigo se debe a la ejecución de 84 rebeldes sindaguas en 1635, una tribu relacionada con Barbacoas. En 1733, el cabildo de la ciudad de Popayán negoció con el palenque el primer acuerdo del mundo colonial. Si se hallaran más documentos, como una carta de un cura, se podría llegar a hipótesis totalmente nuevas.

Por estas razones es necesario diseñar la base de datos de una manera que facilite correcciones posteriores, sin tener que ajustar un gran número de archivos u objetos espaciales (polígonos). A esto se suma la incertidumbre sobre el desarrollo institucional. Cito una descripción de Peter Gerhard, generalmente considerado como autoridad para la geografía histórica de México, sobre la jurisdicción de Ciénega de los Olivos, en Nueva Vizcaya al norte de México:

*While there is evidence on an alcalde mayor at Ciénega de los Olivos in 1659, other documents imply that the valley was attached to the jurisdiction of Sta. Bárbara as late as 1731. By the early 1760s the alcaldía mayor of S[an] José y Ciénega de los Olivos seems to have included not only the valley of that name but also Nonoava and the mountain missions (Gerhard, 1982, p. 182; cursivas nuestras).*

Dicho en otras palabras, no sólo existen incertidumbres sobre los límites exactos entre jurisdicciones vecinas, sino también sobre las maneras de nombrarlas e incluso sobre su existencia. Para el interés del presente artículo, no nos dedicaremos a analizar los obstáculos a la hora de definir territorialidades indianas ni a las críticas teóricas que tiene que superar o integrar un SIG histórico como el que compone HGIS de las Indias. No todas estas incertidumbres conceptuales pueden ser representadas en una base de datos. Más bien se abordan en otras publicaciones, más monográficas y con la posibilidad de argumentación narrativa (Stangl, 2015; Stangl, 2017; Stangl, 2018). Para el diseño de nuestra base de datos no sólo tratamos de reducir el grado y



número de incertidumbres, sino introducir principios rectores metodológicos para traducir información cartográfica o textual, como la de Peter Gerhard, en términos de IT (*information technology*). Lo que queremos elaborar aquí es el diseño de la base de datos para poder trabajar con incertidumbres sobre todo cronológicas, así como los modelos y procesos necesarios para crear composiciones espacio-temporales a partir de diferentes aspectos de la base de datos. En otras palabras: componer una sinfonía del Nuevo Mundo, como insinúa el título.

### La base de datos en los tiempos de la incertidumbre

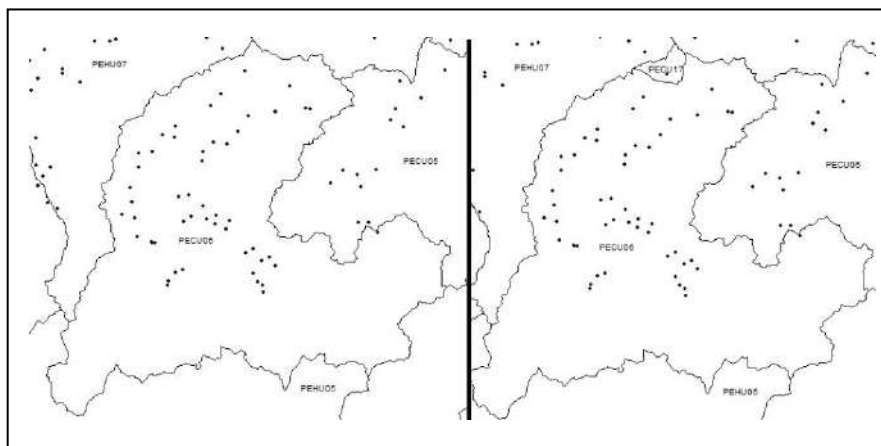
El aspecto más importante para lograr estas metas fue el de aislar fuentes de incertidumbre en diferentes partes de la base de datos. Para cada entidad de lugar o territorio, existe una simple entrada con un ID, los nombres y variantes que tenía a través del tiempo y un tipo genérico o general al que corresponde (“provincia”, “partido”, “obispado”, “audiencia”, “pueblo”, “villa”, “parcialidad”, entre otros). Cada entidad tiene un campo para indicar a cuál de las once “regiones genéricas” pertenece; los distritos pequeños y lugares además tienen un campo para indicar la “provincia genérica”; y para los lugares hay todavía más campos “genéricos” para el partido y el curato al que corresponden. Esta información sirve esencialmente para tener una segunda forma de diferenciar entre las entidades, más allá del simple ID, que es una secuencia de letras completamente abstracta (por ejemplo: JUQUPOPS para la provincia de los Pastos). Con los genéricos, tenemos por lo menos un elemento geográfico independiente de aspectos cronológicos, el cual no cambia. Así, para diferenciar entre “JUGDDUSB” (=“Santa Bárbara, Jurisdicción, GDJ [=Guadalajara], Durango”) y “PTGDCLSB” (=“Santa Bárbara, Partido, GDJ [=Guadalajara], California”); o entre “6000992” (=“Palpas, PER, Tarma, Cajatambo, Gorgor”) y “6000985” (=“Palpas, PER, Tarma, Cajatambo, Churín”).

El otro elementoacrónico y abstracto es la geometría de polígonos bajo el paradigma LCG (*least common geometry*). Este concepto es muy robusto para formar la base SIG de un sistema complejo espacio-temporal: se diseñó para crear el SIG histórico de Bélgica en la década de los noventa (De Moor y Wiedemann, 2001), al poco tiempo ingresó en manuales y libros generales de *best practice* (Ott y Swiaczny, 2001; Gregory y Ell, 2007), y también se empleó con éxito en HGIS Germany (Dietze, Wachtendorf y Zipf, 2007). Aunque el paradigma LCG todavía funciona muy bien (sobre todo en cuanto a la implementación y la interacción de la infraestructura con el mundo exterior y la visualización como la colaboración), dentro de nuestro proyecto tenemos un entorno y posibilidades técnicas muy diferentes. A esto lo llamamos “Web 2.0”.

HGIS de las Indias tiene que enfrentar dos variables que la mayoría de SIG históricos existentes no se ven en la obligación de considerar: el aspecto “colonial” y la unificación de patrones para período comparativamente temprano. La mayoría de

SIG-históricos nacionales, particularmente los que ensayan una reconstrucción territorial, están centrados en los siglos XIX y XX, y por lo tanto pueden hacerse uso de datos y mapas con una fiabilidad casi incuestionable, mientras que para el siglo XVIII la documentación conservada es parcial y carece de la misma homogeneidad. Además, la conocida complejidad de las “jurisdicciones solapadas”, esto es, la administración dividida en cuatro ramos conceptualmente independientes y las idiosincrasias del régimen colonial hispanoamericano (Pietschmann, 2003), imposibilitan un modelaje en clave jerárquica entre cada uno de los doce niveles territoriales (descritos en Stangl, 2017), los cuales deben modelarse por separado.

Para nuestro sistema, los polígonos LCG se definen por compartir la misma información en todos los aspectos de la base de datos durante el período estudiado. Cada vez que encontramos que un pedazo difiere en algún nivel reconstruido, aunque sea por un sólo año, hay que crear un polígono propio. Cuando se reconstruyeron los niveles territoriales a lo largo de un período de 108 años, resultaron 1 500 polígonos-LCG (contados hasta el momento). Cada vez que se crea un nuevo polígono, copiamos la información del “polígono madre” y cambiamos sólo los detalles de información en los que difiere. Es obvio que estos polígonos son abstractos y no representan ninguna entidad histórica. En el ejemplo del gráfico 1, podemos fijarnos en el pueblo de Lucuchanga, que en lo civil siempre había sido parte de Aymaraes, pero eclesiásticamente perteneció al curato de Huancarama en la vecina provincia de Andahuaylas. Por esta misma razón, Lucuchanga estuvo integrado a la diócesis de Huamanga en vez de la de Cuzco, por lo que debió formarse un nuevo polígono LCG (PECU17) para poder representar la mencionada diferencia en el modelo. No importa que se trate de una variación en un solo aspecto.



**Figura 1.** Polígonos-LCG de Aymaraes y Aymaraes-Lucuchanga.

Como se había señalado, el elemento de la cronología hace parte del diseño en diferentes partes. Primero, hay tablas con sus instancias cualitativas temporales tanto para los lugares del *Gazetteer*<sup>2</sup> como para las *Entidades* territoriales. Así, “JUNEYUCA, Campeche, Jurisdiccion, NES” existe como “jurisdicción de Campeche” entre 1701 y 1786 y como “subdelegación de Campeche” entre 1787 y 1808; y el lugar “1000226, Campeche, NES” existe como “Villa de Campeche” entre 1701 y 1776 y como “ciudad de Campeche” a partir de 1777. Con respecto a los lugares hay una variedad de aspectos que influyen en la definición de sus instancias. En efecto, una nueva instancia no sólo se introduce cuando una villa se convierte en ciudad, sino también cuando un lugar se erige en curato o si por un traslado a otro sitio (como la ciudad de Guatemala tras el terremoto de 1773), cambian las coordenadas. La cronología se modela de forma idéntica en cada tabla del sistema, con dos campos (START\_ y END\_) y la incertidumbre se expresa en campos separados, con un número reducido de marcadores como “not\_before”, “not\_after”, “exact” o “ca”.<sup>3</sup>

La Tabla de *Cabeceras* nos informa de que, por ejemplo, el lugar 1000226 era cabecera de su jurisdicción (JUNEYUCA) en todo el período de 1701 a 1808. Además, hay una “tabla relacional” (*Territorios*) que define la relación entre los polígonos y las entidades territoriales en los doce niveles. Continuando el ejemplo de Campeche, los polígonos “NEYU01” y “NEYU02” pertenecen a JUNEYUCA sólo entre 1701-1786, el polígono NEYU13 es parte de JUNEYUCA entre 1701 y 1808. En la Figura 2 muestra las diferentes tablas y sus relaciones en la base de datos, mantenida en MS-Access.

La mayor desventaja de LCG es que los objetos son muy abstractos, ya que los polígonos no representan objetos históricos y los identificadores también son genéricos. Un simple cambio de límite puede desencadenar una pequeña avalancha de nuevas entradas en la base de datos, como veremos en el siguiente apartado. Berman (2013) ha calificado el trabajo necesario en estas operaciones como “esmerado” y “laborioso”. Berman también ha señalado que “el cuidado necesario para mantener consistente las relaciones de los partículas LCG con los objetos geográficos-históricos es enorme”. Además, llama la atención la complejidad del proceso para situar cientos o miles de objetos a lo largo de distintas décadas. La mayor ventaja y fortaleza del diseño es que es posible aislar los diferentes tipos de incertidumbre. Dándole a cada uno su propio lugar:

<sup>2</sup> Gazetteer llamamos aquella parte del sistema que representa los lugares y poblados, no los territorios. En adelante, expresiones subrayadas y en itálicas corresponden a tablas, nombres de archivos o feature classes, expresiones solamente subrayadas corresponden a campos de tablas. Cabe advertir que en la base de datos concreta, algunos nombres de campos y valores difieren de los usados aquí. Para este artículo preferimos títulos más expresivos, mientras que en la batalla diaria de implementación, algunos nombres provisionales se atascaron por ya emplearse en diferentes procesos.

<sup>3</sup> Una explicación más completa de estos campos y los valores marcadores se encuentra en <https://www.hgis-indias.net/dokuwiki/doku.php?id=hgis:cronologia>

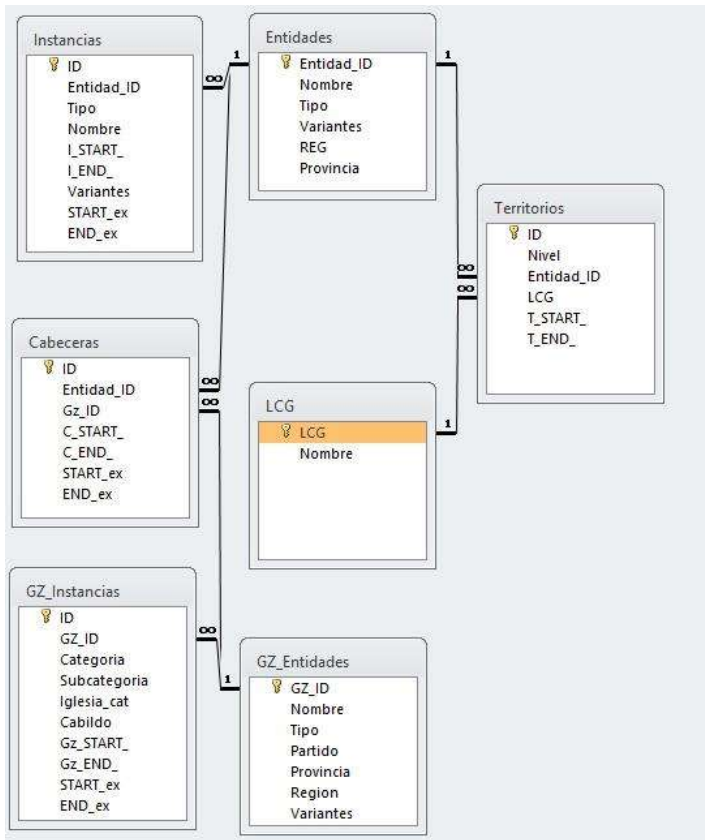


Figura 2. Modelo RDBM espacio-temporal de HGIS de las Indias.

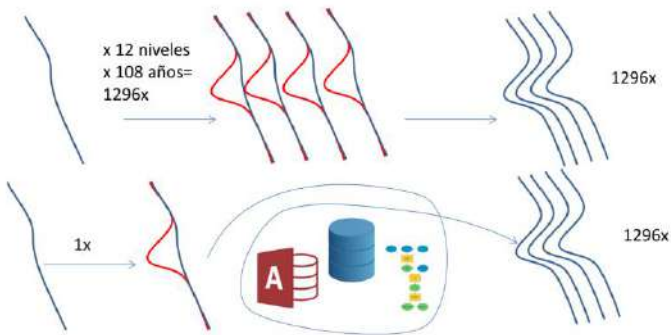


Figura 3: Esbozo de un ajuste de límites sin RDB/LCG (arriba) y con RDB/LCG (abajo).

- Las dudas sobre límites se ciñen a los polígonos LCG,<sup>4</sup>
- Las dudas sobre la pertenencia de un área a una entidad en determinado período son ubicadas en la “tabla relacional” de *Territorios*.
- Las dudas sobre el desarrollo institucional pasan a formar parte de la tabla de *Instancias*.

En la Figura 3, esboza la ventaja de esta estructura con el ejemplo de un simple ajuste de límite. Imaginemos que encontramos que una hacienda hasta ahora no considerada, situada al oeste del límite original, en realidad perteneció a las mismas entidades administrativas de los lugares situados al este. Sin el concepto de LCG y una base de datos relacional intermedia (para gestionar el sistema), deberíamos hacer la misma corrección 1 296 veces. Bajo el paradigma LCG, se hace el ajuste una vez y el resto lo deben manejar procesos y rutinas.

### “Una tabla para gobernarlas a todas”

La información, así organizada, genera mayores beneficios cuando se combina. Para crear un SIG en línea, es imprescindible presentar las diferentes facetas conjuntamente. Un problema esencial es la coherencia de las diversas cronologías de cada parte de la base de datos. Inventamos una situación simple para ilustrar dicha dificultad (Tabla 1).

En las tablas, algunas columnas contienen identificadores únicos para referirse a: entidades territoriales (Entidad\_ID); los polígonos a los que corresponden (LCG); y las capitales de las entidades (Cabecera\_ID). Los campos “\*\_START\_” y “\*\_END\_” indican las fechas de validez de cada entrada. En el ejemplo de la Tabla 1, se tienen diferentes combinaciones de información que tienen sentido para 1701-1745, 1746-1750 y 1751-1808. Un inconveniente grave para este esfuerzo es que ni siquiera ArcGIS procesa bases de datos relacionales con aspectos cronológicos divergentes en la parte geométrica y una tabla relacionada (porque la relación no puede modelarse usando una consulta compleja sino sólo en la identificación a partir de un campo llave).

Para solucionar esto, podríamos pensar en la siguiente estrategia: crear una interfaz poderosa con la posibilidad de que los usuarios aportaran *queries*. Basado en las búsquedas de los usuarios, el sistema podría producir los *output* necesarios para reprogramar la interfaz. Aunque es cierto que esta estrategia es muy flexible, para nuestro caso debemos descartarla por varias razones. Primero, puede afectar el

<sup>4</sup> En un ámbito en el que los linderos son difíciles de fijar y donde las fronteras son imprecisas, sería preferible individualizar el análisis a segmentos de límites en forma de líneas, dándoles atributos de precisión. Esto hubiese significado aumentar el número de objetos espaciales básicos de 1 500 polígonos a 10 000 segmentos de límite, implicando una multiplicación por siete de los elementos que deben controlarse y mantenerse coherentes.

**Tabla 1**  
**Tablas “Territorios”, “Instancias”, “Cabeceras”, “Gz\_Instancias”**

<i>TERRITORIOS</i>	<i>LCG</i>	<i>Entidad_ID</i>	<i>T_START_</i>	<i>T_END_</i>
	XY1	Una	1701	1745
	XY1	Otra	1746	1808
<i>INSTANCIAS</i>	<i>Entidad_ID</i>	<i>Título</i>	<i>I_START_</i>	<i>I_END_</i>
	Una	Corregimiento	1701	1808
	Otra	Corregimiento	1701	1750
	Otra	Gobernación	1751	1808
<i>CABECERAS</i>	<i>Cabecera_ID</i>	<i>Entidad_ID</i>	<i>C_START_</i>	<i>C_END_</i>
	C200	Una	1701	1808
	C300	Otra	1701	1730
	C100	Otra	1731	1808
<i>GZ_INSTANCIAS</i>	<i>Cabecera_ID</i>	<i>Tipo</i>	<i>Gz_START_</i>	<i>Gz_END_</i>
	C300	Villa	1701	1720
	C300	Ciudad	1721	1808
	C200	Ciudad	1701	1808
	C100	Villa	1701	1720
	C100	Ciudad	1721	1808

rendimiento del sistema en línea. De hecho, una base de datos de muchas tablas consumiría demasiados recursos y retrasaría la aparición de los resultados. Segundo, en una interfaz tan compleja las búsquedas sensibles sólo podrían producirse si el usuario sabe cómo funciona la base de datos que está detrás del sistema. Tercero, no existe actualmente ninguna aplicación que pudiera servir como punto de partida para un SIG en línea de esta índole. Para ello, sería necesario todo un equipo capaz de desarrollar un software pertinente. Finalmente, lo más problemático es que todo dependería de la interfaz, pues ningún producto “tiene sentido por sí mismo”. En otras palabras, y para usar un ejemplo, no existiría un objeto que pueda determinar la “extensión del corregimiento de Toluca en 1800”. Este solo podría reconstruirse formulándole a la interfaz la siguiente pregunta: “¿cuál era la extensión del corregimiento de Toluca en 1800?”

Para conseguir objetos concretos, es necesario procesar y juntar todos los componentes de la base de datos. Dentro del presente artículo y dentro del producto que construimos, resulta novedoso el *workflow* que media entre las diferentes partes de la base de datos relacional y que, progresivamente, lleva al *output* final. Como se verá,

la creación de una cronología lógica y consistente sobre un número de tablas relacionadas (donde cada una tiene su propia cronología), es un desafío considerable que merece ser documentado para poder reproducirse con mayor facilidad.

Primero, se crea una tabla maestra (*Dumtable*) que combina las diferentes tablas individuales por un inner join, y por lo tanto son combinados todos los diferentes campos cronológicos. En el ejemplo hipotético de la Tabla 2, tendríamos:

“Dumtable”

LCG	Entidad_ID	T_START	T_END	Titulo	I_START	I_END	C_ID	C_START	C_END	Tipo	Gz_START	GZ_END
XY1	Una	1701	1745	C	1701	1808	C200	1701	1808	Ciudad	1701	1808
XY1	Otra	1746	1808	C	1701	1750	C300	1701	1730	Villa	1701	1720
XY1	Otra	1746	1808	C	1701	1750	C300	1701	1730	Ciudad	1721	1808
XY1	Otra	1746	1808	G	1751	1808	C300	1701	1730	Villa	1701	1720
XY1	Otra	1746	1808	G	1751	1808	C300	1701	1730	Ciudad	1721	1808
XY1	Otra	1746	1808	C	1701	1750	C100	1731	1808	Villa	1701	1720
XY1	Otra	1746	1808	C	1701	1750	C100	1731	1808	Ciudad	1721	1808
XY1	Otra	1746	1808	G	1751	1808	C100	1731	1808	Villa	1701	1720
XY1	Otra	1746	1808	G	1751	1808	C100	1731	1808	Ciudad	1721	1808

Obviamente, no todas estas entradas tienen sentido, por lo que tenemos que limpiar la tabla eliminando todas las entradas ilógicas. No interesan aquellas entradas en las que los demás pares de cronología no tienen intersección con el período T\_START\_ a T\_END.

En MS-Access SQL, eliminamos estas líneas así:

```
DELETE Dumtable.*, T_END_, I_END_, T_END_, C_END_
FROM Dumtable
WHERE (((T_END_)<(I_START_)) OR (((I_END_)<(T_START_)) OR
(((T_END_)<(C_START_)) OR (((C_END_)<(T_START_))));
```

El resultado de esta operación es:

Tabla 3  
Dumtable ajustado

LCG	Entidad_ID	T_START	T_END	Titulo	I_START	I_END	C_ID	C_START	C_END	Tipo	Gz_START	GZ_END
XY1	Una	1701	1745	C	1701	1808	C200	1701	1808	Ciudad	1701	1808
XY1	Otra	1746	1808	C	1701	1750	C100	1731	1730	Ciuda	1721	1808
XY1	Otra	1746	1808	G	1751	1808	C100	1731	1808	Ciudad	1721	1808

Se nota, como es lógico, que desaparecen todas las entradas con la cabecera “C300” de la entidad “Otra” que había en la Tabla 1 porque dejan de ser cabecera antes de que el polígono formara parte de la provincia; desaparecen las entradas con “C100” en calidad de villa, porque ya era ciudad en 1746, cuando empieza a ser cabecera.

En este punto, haría falta definir el período en que la información combinada es válida. Es decir, calcular una nueva cronología en la que hay una sola entrada correcta para cada momento. Para esto, introducimos dos nuevos campos Dump\_START\_ y Dump\_END\_ que en un principio tienen el mismo valor como T\_START\_ y T\_END\_, pero que se cambian con esta búsqueda SQL:

```
UPDATE Dumptable SET (Dump_START_) = (I_START_), (Dump_END_) =
(I_END_)
WHERE (((T_START_)<(I_START_)) OR ((T_END_)>(I_END_)));
```

Lo mismo se hace con C\_START\_, Gz\_START\_, C\_END\_ y Gz\_END\_ (aunque en este ejemplo, estas *queries* no cambian nada).

**Tabla 4**  
**Dumptable final**

LCG	Entidad_ID	T_START_	T_END_	Título	I_START_	I_END_	...	Dump_START	Dump_END
XY1	Una	1701	1745	C	1701	1808	...	1701	1745
XY1	Otra	1746	1808	C	1701	1750	...	1746	1750
XY1	Otra	1746	1808	G	1751	1808	...	1751	1808

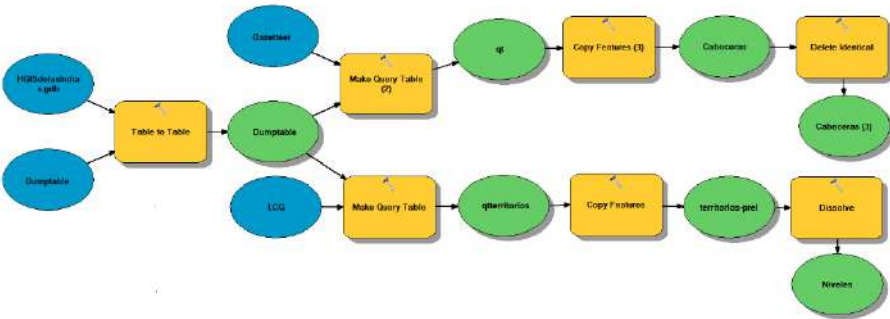
### SIG: el moderno Prometeo

A la tabla ajustada (que es la esencia del sistema), habría que darle cuerpo e insuflarle un límite espacial. Para esto, importamos el *Dumptable* en ArcGIS y hacemos una *query table* con el *feature class* LCG sobre Dumptable.LCG= LCG.LCG. El resultado de esta operación (Figura 4) todavía tendría el inconveniente de que cada polígono LCG (recordamos: en abstracto) seguiría existiendo como objeto. Para combinar aquellos LCG que comparten la misma información, faltaría un *dissolve* (eliminación de límites internos).

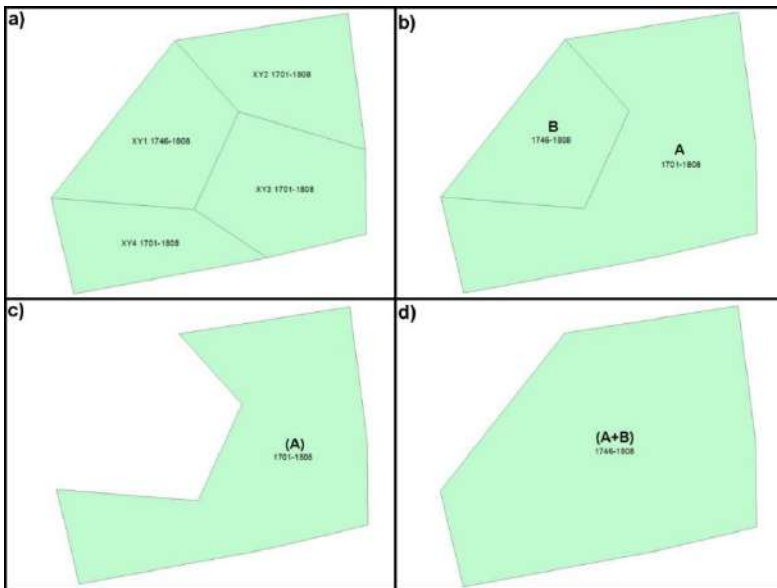
Sin embargo, este simple *dissolve* sólo es una solución parcial (*feature class: Territorios-prel*). La siguiente serie de gráficos ayuda a comprender la problemática. El Gráfico 5a muestra el resultado del *query table* de cuatro polígonos diferentes, y el gráfico 5b el resultado de *Territorios*, hecho sobre Dump\_START\_, Dump\_END\_, Entidad\_ID y Nivel idénticos. El polígono A pertenece a la entidad “Otra” desde



1701 a 1808; el polígono B desde 1746 a 1808. Pero lo que realmente queremos como resultado es un polígono que represente a la provincia entera en cada momento. En para 1746 a 1808, con la extensión “A+B” (Gráficos 5c y 5d). Así, para 1746 a 1808 todavía tenemos un límite interno que hay que eliminar.

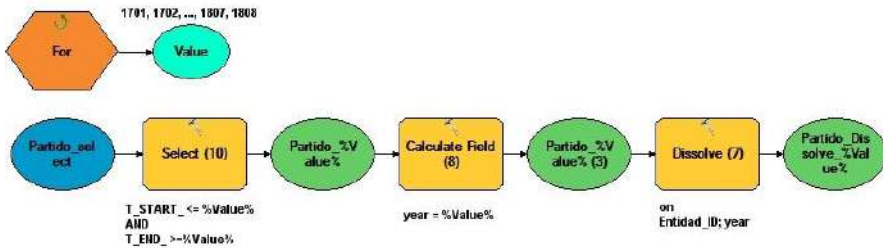


**Figura 4.** Modelo “Dissolve” en ArcGIS model builder para convertir territorios y cabeceras en una geodatabase.



**Figura 5:** Ilustración de diferentes resultados del workflow (a/b) y resultado deseado (c/d).

La solución al problema descrito es una búsqueda SQL con parámetro iterativo para cada año, agregando un campo de cronología nuevo (que llamamos *year*), luego repitiendo el *dissolve* en cada nivel por separado sobre valores idénticos en *year* y Entidad\_ID (Figura 6).<sup>5</sup>



**Figura 6:** Modelo “Dissolve 2” para nivel “Partido” en ArcGIS model builder.

El paso final sería realizar un *dissolve* sobre Entidad\_ID, Título, Nivel y Shape\_Area en años consecutivos. No es posible hacer el *dissolve* sólo sobre una entidad y un área porque entonces se perdería un cambio de título (de corregimiento a alcaldía mayor...). Así mismo, es posible que después de un cambio político vigente por algunos años, el área de una entidad vuelva a su estado anterior. Sin considerar la limitación “en años consecutivos”, entonces, se produciría un resultado erróneo. Lógicamente, la nueva *feature class* tendría que disponer nuevamente de campos START\_ y END\_, tomando los valores de los años extremos de los polígonos disueltos (=que no cambian). No es posible modelar este proceso directamente con *model builder* porque requiere de varios iteradores y para ello debe programarse directamente en Python.<sup>6</sup>

El resultado es parecido, pero no igual al deseado. Ahora tenemos 45 polígonos idénticos para la entidad “Otra” (de extensión “A”) para 1701, 1702,... hasta 1745; y 63 otros polígonos idénticos (de extensión “A+B”) para 1746, 1747,... hasta 1808. Al final, los 108 *feature classes* de cada nivel se reúnen en un *master feature class* (*Todosniveles*). Así se haría en todos los niveles de reconstrucción, para todas las entidades territoriales, lo cual daría como resultados unos 200 000 polígonos aproximadamente. Lo anterior resultaría excesivo, ya que se multiplican polígonos iguales con excepción del campo *year*, pero con la ventaja de poder hacer extractos fáciles para cada año.<sup>7</sup>

<sup>5</sup> Hay que hacer esto para cada nivel por separado pues, si se incluye Nivel en el dissolve, la geodatabase resultante se infla demasiado y tampoco sería suficiente el RAM del servidor.

<sup>6</sup> Puede verse el código empleado en Stangl (2019).

<sup>7</sup> En la página [www.hgis-indias.net](http://www.hgis-indias.net), se pueden descargar shapefiles para cada uno de los doce niveles y para cada un año.

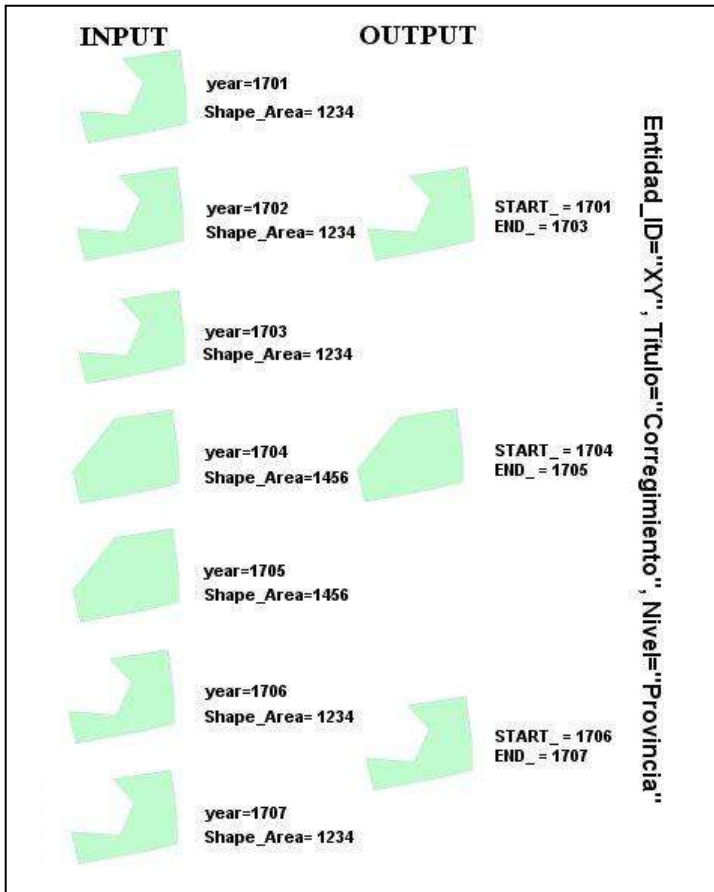


Figura 7. Esquema del proceso “Todosniveles\_dissolved”.

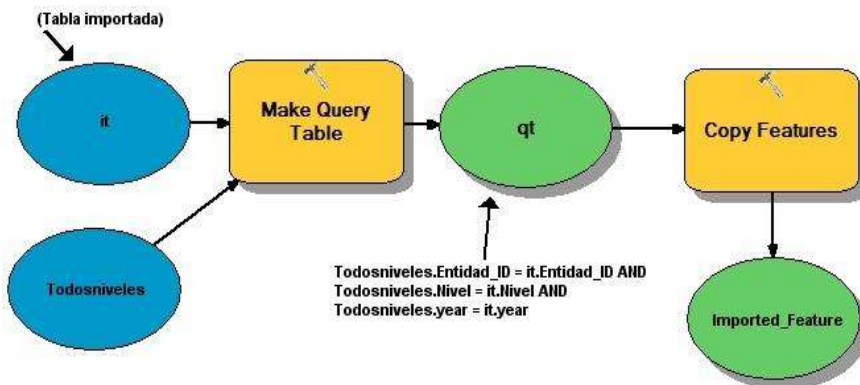
El mayor problema del modelo arriba descrito (con y sin el último proceso), sería el recálculo necesario siempre que se realicen ajustes en alguna parte de la base de datos, ya sea en los límites, en la asignación de atributos cualitativos o cronológicos en una parte del RDBM o del gazetteer. Para este recálculo existen rutinas escritas en Python que automatizan este proceso cada tres meses y que producen una nueva versión de HGIS de las Indias.

Con este núcleo tenemos la base para posibilitar la “especialización” de los datos tabulares externos, agregados por parte de los usuarios, aunque no sean integrados en la aplicación web del SIG. Hay muy pocas convenciones a las que tiene que sujetarse una tabla para poder agregarse a HGIS de las Indias, lo cual facilita su creación:

- Tiene que ser una tabla Excel o un .csv;
- si se refiere a lugares del *gazetteer*, tiene que contener los campos *START\_*, *END\_* para la validez y como identificador el *GZ\_ID* correspondiente para el dato;
- si se refiere a lugares del *gazetteer*, tiene que contener los campos *START\_*, *END\_* para la validez y como identificador el *GZ\_ID* correspondiente para el dato.

Para hacerlo más claro, en este modelo el parámetro *year* es necesario porque tenemos que definir el polígono que usamos (“el polígono de la entidad “XY”, en el nivel de “Provincia” tal como existió en “1750”), mientras que *START\_* y *END\_* regulan la validez del dato (“Entre 1750 y 1760 la provincia XY producía 1 000 marcos de plata”).

De modo análogo al ajuste de las diferentes cronologías (y que se ha descrito arriba), creamos un modelo que agrega la información de la tabla a la *feature class* *Todosniveles*. El siguiente gráfico muestra el modelo y la búsqueda SQL asociada (Figura 8).



**Figura 8:** Modelo para agregación de tablas externas (territorios) a HGIS de las Indias.

## Conclusiones

El diseño de la base de datos como infraestructura espacio-temporal no se ciñe al ámbito histórico concreto estudiado, sino que puede adaptarse fácilmente a otras áreas y temáticas complejas con un alto grado de cambio a través del tiempo o incertidumbres en los datos. Se espera que la detallada iteración de los problemas planteados y los diferentes *workflows* puedan servir de guía para otros proyectos que implican retos similares o problemas metodológicos. Estos pueden estar en la naturaleza de las fuentes, en la función de infraestructura o en la necesidad de tener que acoplar marcos cronológicos independientes, pero referentes a entidades espacio-temporales. Como se ha visto arriba, la base SQL para agregar datos temporales no

es excesivamente complicada. A partir de los procesos delineados aquí un ingeniero de desarrollo de software debería ser fácilmente capaz de programar un *plug-in* en ArcGIS o QGIS, para que cronologías independientes en el futuro se puedan manejar de forma más natural. Con esta capacidad, podrían relacionarse las tablas a la *geodatabase* preexistente, sin tener que realizar otros *shapefiles* o *feature classes*.

En el campo de la historiografía colonial, las principales pruebas para el modelo analizado en este artículo son: primero, su puesta en práctica en una aplicación interactiva concreta; segundo, el interés que podría generar en la comunidad de historiadores de la Hispanoamérica colonial y la capacidad efectiva de HGIS de las Indias de integrar nuevos datos; y, tercero, su compatibilidad con otras herramientas e infraestructuras digitales.

La realización del modelo se ha hecho en forma de SIG en línea usando una aplicación *open-source* llamada CMV-App Viewer.<sup>8</sup> Esta última ha sido adaptada a las necesidades del proyecto, incluyendo un regulador de cronología, alta modularidad para combinar diferentes capas, y una visualización comparativamente rápida.

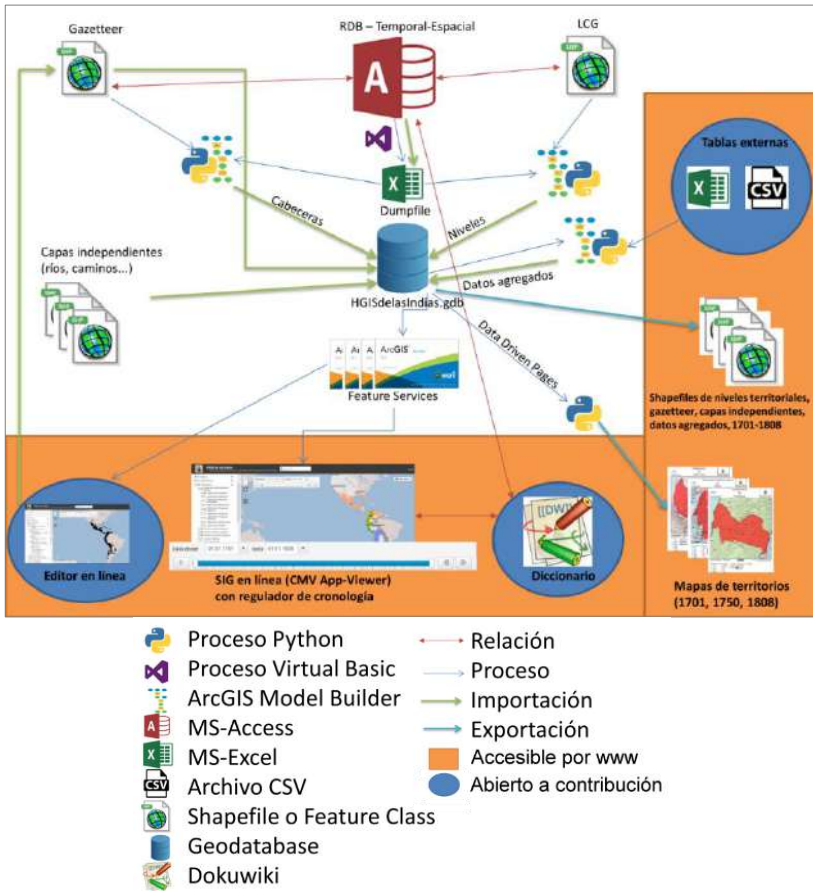
La capacidad de integrar datos externos de colegas se ha realizado a través del elemento tal vez más innovador de HGIS de las Indias: una rutina diaria del modelo para agregar tablas externas que permite al usuario registrado subir su tabla bien formada a la página-web y convertirla en un *shapefile* descargable.<sup>9</sup> Sin embargo, puede resultar difícil el motivar a la comunidad de colegas para usar los identificadores únicos de las entidades de HGIS de las Indias, incluso desde el momento de compilar sus datos. Esperamos que sea cuestión de tiempo, pero seguramente hace falta promoción para dejar claras las ventajas de enlazar diferentes *datasets* a través de una infraestructura común, y que valdría la pena realizar este esfuerzo “extra”. Aun así, sabemos que se han usado geometrías de HGIS de las Indias para la cartografía de varios artículos y libros —ya en los primeros dos años de su existencia.

Además, existen diferentes cooperaciones y proyectos que han planeado sumarse a la iniciativa y ser el germen para una ampliación del sistema a los siglos XVII y XVI. Nos encontramos cooperando intensamente con el equipo del proyecto “Digging into Early Colonial Mexico” de la Universidad de Lancaster, guiado por Patricia Murrieta Flores. Su grupo de trabajo extrae información estructurada de las relaciones geográficas novohispanas del siglo XVI —entre otros objetivos— y reconstruye la geografía escondida en esos textos.<sup>10</sup> Adicionalmente, en 2018 se ha llevado a cabo un avance de proyecto llamado “LatAm Gazetteer”, junto con Ben Brumfield (Austin), Gimena del Río (Buenos Aires) y otros, para explorar la creación de un gazetteer histórico

<sup>8</sup> <https://cmv.io/>.

<sup>9</sup> Idealmente con un archivo *.txt* homónimo con metadata Dublin-Core. Además, con subir una tabla, se acepta el uso de los datos bajo licencia Creative Commons BY-SA 4.0.

<sup>10</sup> <https://www.lancaster.ac.uk/digging-ecm/>.



**Figura 9.** El universo de HGIS de las Indias.

general de América Latina basado en el ya mencionado diccionario de Alcedo (1786-1789) para la plataforma Pelagios.<sup>11</sup>

A partir de “LatAm Gazetteer” se ha logrado la inclusión de nuestro *gazetteer* en dos infraestructuras digitales más generales. Una de ellas es la herramienta Recogito de Pelagios para la edición colaborativa de textos, donde ya se pueden identificar topónimos a través de la información proveída por HGIS de las Indias.<sup>12</sup> De forma similar, hemos sido incluidos en el “World Historical Gazetteer” diseñado por Karl

<sup>11</sup> <http://commons.pelagios.org/2019/02/final-report-on-latam-a-historical-gazetteer-of-colonial-latin-america-and-the-caribbean/>.

<sup>12</sup> <https://recogito.pelagios.org/>.

Grossner para el World History Center de la Universidad de Pittsburgh.<sup>13</sup> Se trata de una nueva plataforma con la meta de integrar diferentes gazetteers históricos y nuestra base de datos sirve de ejemplo implementación de datos espacio-temporales de terceros. Dentro del “World Historical Gazetteer” no sólo se incluyen los lugares, sino también los territorios. Y todo permitiendo la definición de validez temporal en los objetos.

Es en este reto donde tal vez mejor puede juzgarse la idoneidad de la estructura fundamental delineada en este artículo. Para ello se siguen matices prácticos y universales que facilitan: 1) la interoperabilidad con; 2) la integración en; 3) el ajuste a; y 4) la transformación en otros ámbitos. Para concluir, presentamos aquí una partitura gráfica de la “sinfonía del Nuevo Mundo”, la cual sintetiza el *continuum* espacio-temporal que es HGIS de las Indias.

## Referencias

- Alcedo, A. (1786-1789). *Diccionario geográfico-histórico de las Indias Occidentales o América*, 5 vols. Madrid, imprentas varias.
- Berman, M.L. (2013). “Modeling and Visualizing Historical GIS Data”, *Tiempos modernos* núm. 26 Recuperado de: <http://www.tiemposmodernos.org/tm3/index.php/tm/article/view/332/374> .
- Bodenhamer, D.; Corrigan, J.; y Harris, T.M. (2010). *The Spatial Humanities: GIS and the Future of Humanities Scholarship*, Bloomington, Indiana University Press.
- Cárdenas-Arroyo, F. (1996). “Frontera Arqueológica vs. Frontera Etnohistórica: Pasos y Quillacingas en la Arqueología del Sur de Colombia”, en Chantal Caillavet et al., *Frontera y poblamiento: estudios de historia y antropología de Colombia y Ecuador*, Lima, Institut Français d'études Andines. Recuperado de <http://books.openedition.org/ifea/2498>
- De Moor, M. & Wiedemann, T. (2001), “Reconstructing territorial units and hierarchies: A Belgian example”, *History and Computing* 13, pp. 71-97.
- Dietze, L.; Wachtendorf, C. & Zipf, A. (2007). “GIS-gestützte Erfassung und Web-basierte Visualisierung historischer Geodaten im Historischen GIS Germany”, Recuperado de: <https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/hgis-germany.agit2007.neu.pdf> .
- Durán y Díaz, J. (2012). *Estado General de todo el Virreinato de Santafe de Bogotá en el presente año de 1794*, Bogotá, Banco de la República.
- Gerhard, P. (1982). *The North Frontier of New Spain*, Princeton, Princeton University Press.

<sup>13</sup> <http://whgazetteer.org/> .

- Herrera Ángel, M. (2009). *Popayán: la unidad de lo diverso. Territorio, población y poblamiento en la provincia de Popayán, siglo XVIII*, Bogotá, Universidad de los Andes.
- Lünen, A. & Travis, C. (2013). *History and GIS. Epistemologies, Considerations and Reflections*, New York, Springer.
- Ott, T. & Swiaczny, F. (2001). *Time-Integrative Geographic Information Systems: Management and Analysis of Spatio-Temporal Data*, New York, Springer.
- Pietschmann, H. (2003). “Los principios rectores de organización estatal en las Indias”, en Annino, A. & Guerra, F. (eds.), *Inventando la Nación, Iberoamérica, siglo XIX*, México DF, Fondo de Cultura Económica.
- Schortman, E. & Urban, P. (1992). *Resources, Power, and Interregional Interaction*, New York-London, Springer.
- Stangl, W. (2015), “Scylla and charybdis 2.0: reconstructing colonial Spanish American territories between metropolitan dream and effective control, historical ambiguities and cybernetic determinism”, en *Culture & History Digital Journal*, 4(1), <http://dx.doi.org/10.3989/chdj.2015.008> .
- (2017), “¿Provincias y partidos o gobiernos y corregimientos? Los principios rectores del desordenamiento territorial de las Indias y la creación de un sistema de información histórico-geográfico”, en *Jahrbuch für Geschichte Lateinamerikas* 54: 157-210.
- , (2018) ‘The Empire Strikes Back’? HGIS de las Indias and the Postcolonial Death Star”, en *International Journal of Humanities and Arts Computing*, 12, 2:138-162.
- (2019), “Code: Create dissolved temporal-scoped polygons”, Harvard Dataverse, V1. <https://doi.org/10.7910/DVN/A3DTAW>
- Velasco, J. (1842). *Historia del Reino de Quito en la América meridional. Historia moderna*, Quito, Gobierno.



# Megaciudad Buenos Aires: Cartografía de su última expansión y conurbación mediante el procesamiento digital de imágenes satelitales nocturnas

*Megacity Buenos Aires: Cartography of its latest expansion and conurbation with digital processing of nocturnal satellite images*

Gustavo D. Buzai\*  
Eloy Montes Galbán\*\*

*Recibido 20 de diciembre de 2019; aceptado 30 de enero de 2020*

## Resumen

El procesamiento digital de imágenes satelitales nocturnas y el uso de Sistemas de Información Geográfica confirma que la última expansión y conurbación de la Megaciudad Buenos Aires se produjo con el cambio de siglo. Esto se verifica al analizar un período de dos décadas (1992-2012) con centralidad en el año 2002. Si bien la estabilización del crecimiento de la Megaciudad Buenos Aires comenzó hace cinco décadas, el análisis temporal de imágenes y la realización de cartografía temática permitieron representar espacialmente un nuevo proceso de conurbación. Este proceso se desarrolló en cinco ejes de expansión incorporando otras ciudades de la región metropolitana cuya distancia media al centro principal es de 61 km. Las imágenes satelitales nocturnas brindan nuevas perspectivas del espacio urbano al utilizarse los conceptos de megaciudad, expansión urbana, conurbación e *infraestructura gris*. A través de los elementos que brinda el área urbana consolidada, la urbanización difusa

\* Universidad Nacional de Luján, Argentina, correo electrónico: [gdb@unlu.edu.ar](mailto:gdb@unlu.edu.ar).  
ORCID: <https://orcid.org/0000-0003-4195-5324>.

\*\* Universidad Nacional de Luján, Argentina, correo electrónico: [emontesgalban@gmail.com](mailto:emontesgalban@gmail.com).  
ORCID: <https://orcid.org/0000-0002-4075-4873>.

y la mayor densidad y proporción de las vías de comunicación dan la posibilidad de captar configuraciones actuales y disponer de elementos para captar tendencias espaciales futuras.

Palabras clave: *Megaciudad, Buenos Aires, expansión urbana, imágenes satelitales nocturnas, Cartografía Digital.*

## Abstract

Digital processing of nocturnal satellite images and the use of Geographical Information Systems confirms that the last expansion and conurbation of the Megacity Buenos Aires occurred with the turn of the century. This is verified by analyzing a period of two decades (1992-2012) with centrality in 2002. Although the stabilization of the growth of the Megacity Buenos Aires began five decades ago, the temporal analysis of images and the realization of thematic cartography allowed a spatial representation of a new conurbation process. This process was developed in five axes of expansion incorporating other cities in the metropolitan region whose average distance to the main center is 61 km. The nocturnal satellite images provide new perspectives of the urban space when the concepts of megacity, urban expansion, conurbation and gray infrastructure are used. Through the elements provided by the consolidated urban area, diffuse urbanization and the greater density and proportion of communication channels we have the possibility of capturing current configurations and count with elements to foresee future spatial trends.

Key words: *Megacity; Buenos Aires, urban sprawl, nocturnal satellite images, digital cartography.*

## Introducción

De acuerdo al último reporte *The World's Cities* publicado por Naciones Unidas (UN, 2018) hay actualmente 33 megaciudades en el mundo, aglomeraciones urbanas que superan los 10 000 000 de habitantes. América Latina cuenta con seis: São Paulo, Brasil (21 650 000), Ciudad de México, México (21 580 000), Buenos Aires, Argentina (14 967 .000), Río de Janeiro, Brasil (13 293 000), Bogotá, Colombia (10 574 000) y Lima, Perú (10 391 000).

En estas megaciudades pueden verificarse los mayores efectos de la globalización en lo económico al orientarse hacia los mercados globales, en lo político a través de políticas de desregulación, en lo social en la polarización socio-espacial y en lo ecológico a la mayor probabilidad de riesgos (Borsdorf y Coy, 2009). Desde un punto de vista físico son las mayores extensiones de superficie urbana construida, aglomeraciones de miles de km<sup>2</sup> que se encuentran estructuradas con la totalidad de elementos tradicionales y actuales que permiten su funcionalidad. Desde sus comienzos

hasta hoy y desde el centro hasta sus bordes combina estructuras en permanente evolución.

Uno de los aspectos donde más ha centrado la atención respecto a la evolución espacial de las megaciudades son sus extensas periferias. Inicialmente la extensión física de la ciudad central, luego el proceso de conurbación que la fusiona físicamente con ciudades cercanas y finalmente la mejora en las vías de comunicación para extenderla y permitir la creación de barrios cerrados destinado a grupos sociales ligados a las actividades económicas del CBD (*Central Business District*) como centro de primer orden en la jerarquía de centros intraurbanos. Finalmente, algunas megaciudades desarrollan *exociudades* dentro de la trama urbana periférica, ciudades exteriores con semi-autonomía y, de esta manera su espacio exterior comienza a funcionar fragmentado en polígonos delimitados que concentran empleos de nivel terciario y cuaternario con viviendas y servicios esenciales en un paisaje general propio del CBD.

En este contexto, el presente trabajo tiene por objetivo analizar la última gran expansión que se produjo en la Megaciudad Buenos Aires, situación con la cual hizo su ingreso al siglo XXI. Corresponde a un amplio proceso de conurbación en los bordes más alejados fusionando cinco ciudades tradicionalmente separadas y con ello la Megaciudad Buenos Aires amplía su superficie y cantidad de habitantes en un salto cualitativo y cuantitativo de gran significancia.

Este aspecto que pudimos comenzar a delinear a partir del papel preponderante que tomó la *infraestructura gris*<sup>1</sup> en la región, de la evidencia empírica visual a partir de trayectos realizados y vuelos fotográficos en el marco de un proyecto de investigación previo (Buzai y Lanzelotti, 2019) pudimos comprobarlo a través del uso de técnicas geoinformáticas.

El procesamiento digital de imágenes satelitales nocturnas, junto al modelado cartográfico y la cartografía digital síntesis realizada mediante un Sistemas de Información Geográfica (SIG) nos permitió comprobar que la conurbación se había producido entre los años 1992 y 2002 y consolidado en 2012, llevando el límite más extenso de la megaciudad a 90 km del CBD de la ciudad central.

En los puntos siguientes se presenta el *marco teórico* que sustenta conceptualmente la investigación y los desarrollos de modelos urbanos para el estudio de las grandes ciudades de América Latina, el *marco metodológico* en donde se presenta una revisión documental de la utilización de imágenes satelitales nocturnas para el estudio del crecimiento urbano, características de los datos utilizados, las técnicas asociadas para su tratamiento y aspectos cartográficos, y finalmente se presenta la

<sup>1</sup> La *infraestructura gris* o infraestructura construida son sistemas de construcción humana (artificiales) que brindan soporte a la comunidad como las carreteras y otros sistemas de transporte como el de gestión de aguas pluviales y diversos servicios públicos (Benedict y McMahon, 2006).

aplicación, en la cual se incluye la cartografía producida mediante el tratamiento de las imágenes satelitales nocturnas y el mapa final, síntesis de la comprobación y base para la interpretación del nuevo proceso de expansión.

## **Perspectiva sistémica en focalización espacial**

### ***Marco conceptual***

Las distribuciones espaciales no se producen de manera aleatoria, sino que pueden asociarse a comportamientos generales que permiten llegar a la formulación de modelos y leyes para explicar la organización del territorio. Desde un punto de vista sistémico los estudios geográficos intentan descubrir regularidades que permitan comprender el comportamiento espacial.

Conocer estas leyes, da la posibilidad de realizar predicciones teóricas respecto de las configuraciones espaciales futuras en el marco de una Geografía Aplicada prospectiva a partir de considerar relaciones causales que permiten modelar escenarios futuros.

La investigación científica tiene como objetivo estudiar elementos específicos de la realidad y, principalmente, sus espacios relacionales. De esta manera, una tarea inicial es realizar una clara demarcación del campo de estudio a partir de definir totalidades estructuradas como sistemas.

Un marco de abordaje es la teoría general de los sistemas (TGS) formulada por Bertalanffy (1968) como construcción global que avanza empíricamente en el estudio de elementos y relaciones con el objetivo de encontrar similitudes modelísticas en estructuras de diferentes recortes de la realidad. De esta manera se ponen en evidencia aspectos matemáticos que pueden ser utilizados para estudiar y describir comportamientos en diferentes escalas. Los estudios geográficos centran su análisis en la escala humana y, dentro de ella, en el nivel espacial.

Un sistema se explica a través de su estructura general. La TGS permite unir diferentes elementos que provee la especialización científica y brinda posibilidades para un análisis global. El abordaje en el estudio de la organización del territorio permite obtener resoluciones a través del uso de las matemáticas como lenguaje de la ciencia, y la geometría como lenguaje de la forma espacial.

La consideración de especificidades que emergen en diferentes escalas fue la base para la formulación de la teoría de los sistemas complejos (TSC) sistematizada por García (2006). Este avance demostró importantes capacidades en dos niveles principales del análisis geográfico: en el estudio de la organización territorial y en la construcción de conocimientos en base al alcance epistemológico de los diferentes paradigmas de la Geografía.

Estas capacidades resultan posibles porque concibe una realidad organizada de manera estratificada a partir de niveles con funcionamientos semi-autónomos. Puede

considerarse que entre lo infinitamente grande (espacio sideral) y lo infinitamente pequeño (espacio microscópico) está la escala humana (Rosnay, 1977) donde la Geografía pone atención. Al aplicar la TSC al estudio de la organización territorial se realizan abordajes que podrían desembocar en tres niveles, uno *focal* que sería el de resolución propio de la temática (el espacio geográfico), uno *supra-focal* (escala mundial) de mayor amplitud y uno *infra-focal* (escala personal) de mayor nivel de detalle.

El paradigma cuantitativo que apoya el análisis espacial con SIG brinda la base metodológica para el nivel focal. Sus estudios abordan en aspectos concretos y visibles, por lo que, como actividad científica, se basa en los supuestos centrales del *materialismo sistémico* (Bunge, 2012) que establece que la realidad existe independientemente de cómo se la piense (*realismo*), que puede ser estudiada a través de su materialidad (*materialismo*), que estos elementos tienen relaciones sistémicas (*sistemismo*) y que la forma de mayor aptitud para aprehenderlo es la ciencia (*cientificismo*) en el marco de la Ciencia Aplicada y, en nuestro caso, de la Geografía Aplicada con intención de captar una nueva instancia en la conformación de la Megaciudad Buenos Aires.

### **Marco metodológico**

La Geografía como disciplina científica y, más específicamente, la Geografía Urbana, permite observar de qué manera los métodos y técnicas que se aplican al estudio del espacio urbanizado fueron evolucionando a la par de los avances geotecnológicos. La Geografía forma parte de la actual revolución geotecnológica especialmente en lo referente a la generación, acceso y manejo de información geográfica digital (Montes Galbán, 2017).

El paradigma de la *Geografía Cuantitativa* avanza en la década de 1980 hacia la *Geografía Automatizada* (Dobson, 1983) en un proceso continuo de incorporación de las técnicas geográficas de base racionalista y cuantitativa al ambiente informático. Esta evolución, que ya lleva muchas décadas (Buzai, 2015a), logra gran estabilidad a partir de la década del 2000, momento en que los diferentes *software* adquieren excelente compatibilidad con la finalidad de compartir información geográfica (Buzai y Ruiz, 2012).

Las tecnologías digitales en la actualidad presentan una gran variedad de posibilidades de aplicación y los SIG toman el lugar de *pivote* de integración para facilitar la obtención de resultados a partir de la convergencia creativa.

La valorización generalizada de estas aplicaciones ha sido muy importante y su prestigio creció al mismo tiempo que la mayoría de las ciencias muestran interés en la incorporación de la dimensión espacial en sus estudios.

La transformación del mundo real en un modelo digital con posibilidades de ser trabajado mediante procedimientos computacionales, exige una serie de operaciones

de estandarización que finalizan al nivel de *byte*. Mediante estas transformaciones, que se encuentran compuestas por procesos de fragmentación y estandarización de la información espacial, todo objeto geográfico puede definirse digitalmente a través de una geometría particular (en estructura vectorial: punto, línea y polígono. En estructura raster: celda), una localización precisa en el espacio absoluto ( $x$ - $y$  o geográficas), una serie de atributos (en estructura vectorial: variables. En estructura raster: número digital) y su existencia en un punto del eje temporal.

En sistema vectorial, la concreción de estos aspectos se realiza a través de la creación de *bases de datos alfanuméricas* y *bases de datos gráficas* y, para ello, existe *software* específico destinado al tratamiento de datos georreferenciados. Los primeros orientados a los procedimientos numéricos (*planillas de cálculo*, *programas de análisis estadístico*) y los segundos a la geometría (*diseño asistido por computador*, *cartografía automatizada*). Mediante esta organización se aplican procedimientos del análisis multivariado.

En el sistema *raster* el número digital (característica clasificatoria de la entidad) se ubica matricialmente en el mismo nivel de la celda que representa, por lo tanto la geometría y el número único son inseparables. Mediante esta organización se aplican procedimientos del procesamiento digital de imágenes satelitales y el modelado cartográfico a través del álgebra de mapas.

Actualmente los SIG combinan la totalidad de posibilidades de procesamiento de datos y, en este sentido, podemos ver como en el presente trabajo son un claro nexo entre el procesamiento digital de imágenes satelitales nocturnas de áreas urbanas y la representación en cartografía temática.

Considerando su aplicación al estudio del espacio geográfico como sistema complejo vemos que se apoyan en una focalización espacial. Todo análisis en el cual intervengan cuestiones sociológicas, económicas, políticas o culturales son plausibles de tratamiento siempre considerando su manifestación espacial. La aplicación apoya el materialismo sistémico porque aborda la realidad desde un punto de vista empírico, a través de su materialidad en relaciones modelísticas y sistémica, con metodología propias de la ciencia aplicada.

## **Modelos urbanos hacia la megaciudad en América Latina**

### ***Modelos urbanos***

Cuando los modelos urbanos clásicos (anillos de Burgess, 1925; sectores de Hoyt, 1939; núcleos múltiples de Harris y Ullman, 1945) son utilizados para explicar la realidad de las ciudades de América Latina, si bien los patrones geométricos de organización generan similares configuraciones espaciales el contenido correspondiente al mapa social es diferente.

Esta situación la había señalado inicialmente Schnore (1965) surgiendo que no era posible aplicar con éxito el modelo secuencial propuesto por Sjoberg (1960), el cual basa su análisis en el cambio de organización espacial que lleva el paso de la ciudad pre-industrial a la ciudad industrial. Las ciudades de América Latina presentan un mapa social inverso al de las ciudades de América Anglosajona (Buzai, 2014, 2015b) y este sería el aspecto distintivo sobre el cual se basarían los nuevos modelos urbanos.

Los modelos surgidos de las ciudades de América Latina muestran una inicial ciudad compacta al estructurarse en anillos y sectores y, en esta configuración, Griffin y Ford (1980) destacan la principal característica del mapa social: condiciones socio-económicas favorables que disminuyen hacia la periferia. En este proceso, notable durante la segunda mitad del siglo XX, el CBD concentra la mayor parte de la oferta de bienes y servicios de alta especificidad como centro principal. Asimismo, las actividades del comercio al por menor se expanden de manera lineal a partir de alguna importante vía de circulación (*spine*), situación que es acompañada por el desplazamiento sectorial de grupos sociales de alto nivel socioeconómico sobre una base de tres anillos que surgen desde el centro.

Una década y media después Ford (1996) amplía el modelo a partir de incorporar nuevos procesos en el centro como la *gentrificación* y en la periferia como la aparición de una avenida de circunvalación (tomado del modelo *Periférico* de Ciudad de México), la aparición del Mall en el extremo de la expansión lineal comercial (*spine*) y el parque industrial en diferente ubicación periférica. La experiencia muestra que el modelo en su versión inicial se ajusta muy bien al caso de las ciudades de tamaño intermedio y su modificación lo asimila a las grandes ciudades.

Modelos orientados hacia la estructura espacial de las grandes ciudades de América Latina estuvieron formulados inicialmente por Bähr (1976) y Mertins (1980) que dieron como resultado final la confluencia presentada en Bähr y Mertins (1982) y, junto a ellos, también debe ser considerado el aporte realizado por Borsdorf (1982) y su actualización en Borsdorf (1994). Tomando como referencia la aglomeración de Buenos Aires, Janoschka (2002) propone el modelo de *ciudad de islas*, una estructuración espacial que corresponde a la visión de archipiélagos urbanos formados por islas de riqueza en mares de pobreza (Berry, 1985; Borsdorf y Coy, 2009).

En la ciudad fragmentada se destaca el notable impacto de los flujos migratorios en el crecimiento poblacional y la expansión física de la ciudad. En su evolución histórica pueden verse varias estructuras superpuestas: 1) Iniciales anillos concéntricos a partir del CBD, 2) estructura moderna sectorial de expansión por ejes de crecimiento y 3) estructura periférica celular formada principalmente por asentamientos informales y proyectos gubernamentales de viviendas, apareciendo incipientemente barrios cerrados de alto nivel económico. Surgen extensas y heterogéneas periferias como el rasgo más notable de la megaciudad.

### *Periferias urbanas actuales*

El espacio periurbano, la franja marginal urbano-rural (Carter, 1983) es uno de los más complejos de la ciudad, ya que allí se superponen los servicios ambientales que llegan del ámbito rural y los servicios urbanos que se extienden desde el centro de la ciudad, ambos en una proporción de sus máximas posibilidades.

La *ciudad compacta* evolucionó hacia la *ciudad difusa* y al archipiélago urbano (Indovina, 2009) producto de la acelerada expansión urbana (*urban sprawl*) lejana al centro principal de la ciudad. Desde mediados del siglo pasado, esta expansión fue característica de las ciudades norteamericanas a partir de amplios loteos periféricos, la construcción de autopistas y la masividad en la propiedad automotor. La proporción relativa en la distribución poblacional permitió considerar una *contraurbanización* (Berry, 1976) y aparecen manchones periféricos que generan algunas problemáticas sociales al ser el soporte de una segregación poblacional (Whyte, 1958).

El proceso de expansión urbana se desarrolló de manera ininterrumpida hasta la actualidad (Mitchell, 2001) generando diferentes visiones socio-espaciales de un espacio geográfico sobre el cual resulta complicado establecer límites precisos demarcatorios entre lo urbano y rural.

El caso de Buenos Aires es paradigmático y fue tomado como ejemplo de la evolución de las grandes ciudades de América Latina en el marco de la globalización y del avance del liberalismo económico en la región. Si bien la aglomeración había comenzado un proceso de estabilidad en su valor superficial a partir de la década de los setenta (Buzai y Baxendale, 1998) a partir de mediados de la década de los noventa comienza a tomar el modelo de ciudad norteamericana ante la aparición de las urbanizaciones cerradas (*country clubs*) como polígonos residenciales y la creación de autopistas que las conectan con la ciudad central, la Ciudad Autónoma de Buenos Aires. Esto permite ver el desplazamiento suburbano de los grupos de alto nivel socio-económico que comienzan a segregarse (Buzai, 1992) y a generar sus principales actividades diarias en espacios cerrados vinculados por autopistas. La totalidad de espacios comienzan a ser *no-lugares* con paisajes similares en cualquier importante ciudad del mundo que experimente los mismos procesos (Auge, 2017).

En cuanto al mapa social, Buenos Aires presenta el modelo de ciudad pre-industrial en su momento de máxima industrialización en 1945 y toma el modelo de ciudad industrial cuando es evidente su disminución industrial a partir de 1990 (Buzai y Marcos, 2012). Se aprecian características de *suburbia* (Bauer, 1993) en una mezcla de proyectos de monoblocks de gestión pública y el impulso de los desarrolladores privados en espacios habitacionales y empresariales.

Este crecimiento urbano genera efectos irreversibles al afectar la resiliencia de los ecosistemas de entorno y que son de utilidad para la sostenibilidad del hábitat humano. La aglomeración de Buenos Aires produce efectos negativos sobre la llanura pampeana y el bosque ribereño, espacios que merecerían cuidado por sus servicios



ambientales, diversidad de especies y variados paisajes. Estos efectos pueden verse en la *huella paisajística* (Morello *et al.*, 2001) con clara evidencia a distancias de la ciudad central de entre 35 y 70 km.

Desde un punto de vista espacial, en esta franja se verifica una notable irregularidad y fragmentación en los bordes de la aglomeración, aspecto que pudo ser estudiado mediante la geometría fractal con la finalidad de captar la verdadera forma y patrones de crecimiento urbano (Batty y Longley, 1994), geometría aplicada a la aglomeración de Buenos Aires por Buzai, Lemarchand y Schuschny (1998) con fines de modelización de su expansión con autómatas celulares.

La fragmentación espacial y social en las periferias urbanas actuales es uno de los signos distintivos de la *posmetrópolis* (Soja, 2000).

## Metodología

### *Imágenes satelitales nocturnas: revisión documental*

Cuando se efectúa una primera revisión documental del uso de imágenes satelitales nocturnas en Geografía, es posible acceder a algunas publicaciones de divulgación científica general, que desde hace más de dos décadas muestran su aplicación en proyectos de investigación a nivel planetario, este es el caso de los mapas publicados por la National Geographic (1998, 2004). Estos mosaicos del planeta se construyeron en base a datos satelitales de luces nocturnas reunidos a lo largo de un año, en el cual se pueden identificar las regiones del planeta con grandes áreas urbanas mostrando un intenso brillo de luz blanca, por otro lado, también destacan los grandes incendios, las locaciones con combustiones de gas y la pesca nocturna.

De igual manera, está el “Proyecto Huella Humana” desarrollado por la *Wildlife Conservation Society* (WCS) y el *Center for International Earth Science Information Network* (CIESIN), su resultado permitió representar en un mapa global los niveles de influencia antrópica en el planeta (especialmente sobre la vida salvaje y tierras vírgenes), uno de los factores considerados para la construcción de este mapa, fueron las redes de energía, que se midieron a través de las imágenes satelitales que perciben luces por la noche, puesto que es un indicador adecuado para la identificación de la infraestructura energética (National Geographic, 2005).

Las investigaciones que utilizan imágenes satelitales nocturnas y que buscan dar cuenta de las actividades humanas son cada vez más abundantes y diversas entre las cuales podemos mencionar aplicaciones en una importante variedad de temas que contemplan la relación entre la sociedad y el medio físico-natural (Montes Galbán, 2019). Específicamente podemos nombrar aplicaciones urbanas en Henderson *et al.* (2003) quienes mencionan las dificultades de establecer un umbral de intensidad lumínica para su definición; en esta línea, otros trabajos intentaron realizar validaciones con otras fuentes satelitales (Small *et al.*, 2005; Potere *et al.*, 2009). A pesar de las

complejidades y limitaciones metodológicas se continua avanzando en el tema para la identificación, delimitación y monitoreo del crecimiento de las áreas urbanas a partir de las imágenes nocturnas, incrementándose los aportes científicos provenientes de diferentes continentes, América (Mitchell, 2001; Amaral *et al.*, 2006; Pérez Machado *et al.*, 2014; Martín Oriol, 2016), Asia (Roy Chowdhury *et al.*, 2012; Liu & Leung, 2015; Xin, *et al.*, 2017; Kamarajuggedda *et al.*, 2017), Europa (Nel.lo *et al.*, 2016; Nel.lo *et al.*, 2017; Checa *et al.*, 2017; Checa & Nel. lo, 2018) y estudios comparativos entre regiones del planeta (Zhou *et al.*, 2015).

Los antecedentes de la referida temática en Argentina y en el área que atañe a la presente investigación no son abundantes. Se puede mencionar una propuesta metodológica a escala nacional que tuvo como objetivo identificar las áreas intermedias entre los usos urbano y rurales, con este fin se consideraron entre los materiales las imágenes satelitales nocturnas (Quaíni, 2014); un antecedente más reciente es el trabajo desarrollado por Martín Oriol (2016), quien realizó un estudio comparativo entre las urbes de Buenos Aires y Lima, analizando los cambios en la forma urbana a través de imágenes nocturnas de satélite; por último, desarrollamos una primera aproximación utilizando imágenes del satélite *Suomi NPP* (compuestos anuales) ofrecidas por la *National Oceanic and Atmospheric Administration* (NOAA) para estimar el porcentaje de superficie iluminada en los partidos de la cuenca del río Luján (Montes Galbán, 2019), y evaluar la expansión y densificación del Gran Buenos Aires en el período 2012-2019 (Montes Galbán, 2020).

### **Información básica – Fuente de datos satelitales**

Los datos primarios utilizados en la investigación provienen de tres fuentes. La principal está referida a las imágenes de la Tierra por la noche producidas por el Grupo de Observaciones de la Tierra (*Earth Observation Group* —EOG por sus siglas en inglés), que forma parte del Centro Nacional de Datos Geofísicos (NGDC por sus siglas en inglés) de la Administración Nacional Oceánica y Atmosférica de los Estados Unidos (NOAA por sus siglas en inglés). Dichas imágenes son una composición que es posible gracias a los sensores denominados OLS (Operational Linescan System - Sistema Operacional en Línea) a bordo de la serie de satélites DMSP del Programa de Satélites Meteorológicos del Ministerio de Defensa de los Estados Unidos (*Defense Meteorological Satellite Program*).<sup>2</sup>

Los datos que presenta el EOG disponibles para su descarga en su versión 4<sup>3</sup>, son cuadrículas de 30 segundos de arco, que abarcan desde -180 a 180 grados de longitud

<sup>2</sup> La creciente demanda de los productos de iluminación nocturna por una amplia variedad de usuarios (científicos de uso del suelo, geógrafos urbanos, ecólogos, astrónomos, demógrafos, economistas y científicos sociales) ha propiciado el desarrollo de algoritmos para la generación de productos de luces nocturnas globales de alta calidad (Elvidge *et al.*, 2017).

<sup>3</sup> Entre la variedad de productos que el Centro Nacional de Datos Geofísicos pone a disposición para su descarga están los compuestos anuales DMSP-OLS Nighttime Lights Time Series en su versión 4

y -65 a 75 grados de latitud. Dichas cuadrículas tienen una resolución aproximada de 760 x 760 metros (en la latitud que corresponde al área estudiada) y están disponibles en formato GeoTIFF (Tabla 1). Para los efectos de la presente aplicación se descargaron las imágenes con una cobertura a intervalos temporales de diez años: 1992, 2002 y 2012.

**Tabla 1**  
**Características del DMSP-OLS y las imágenes utilizadas**

<i>Características</i>	<i>Descripción</i>
Cuantificación radiométrica	6 bits
Valores de luminosidad	0 a 63
Saturación	En zonas urbanas
Tamaño de pixel	30 segundos de arco
Calibración radiométrica	Sin calibración
Bandas espectrales	1 (Infrarrojos términos 10-12 um)
Sensibilidad a la radiación	Pancromática 500 – 900 nm
Hora nocturna de paso	19:30h UTC
Cobertura de las imágenes descargadas	-180 a 180 grados longitud / -65 a 75 grados de latitud
Serie temporal (trabajados)	1992 - 2002 - 2012

**Fuente:** Elaboración propia en base a Elvidge *et al.* (2013) y Checa *et al.* (2017).

La segunda fuente de datos empleada fueron imágenes del satélite Landsat 5 TM, escena Path: 225 Row: 84, (05 de enero de 1992) descargadas del Geoservidor *Earth Explorer* del *United States Geological Survey* (USGS, 2019).

Por último, para el desarrollo de las composiciones cartográficas fue necesario descargar las siguientes capas en formato *shapefile* (.SHP) provenientes del banco de datos abiertos del Instituto Geográfico Nacional (IGN, 2019):

- Departamentos / Partidos
- Provincias.

(<https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>). Los archivos utilizados en la investigación corresponden a los nombrados como “avg\_vis limpio” los cuales contienen las luces de ciudades, pueblos y otros sitios con iluminación persistente, incluidas las llamaradas de gas y descartando los eventos efímeros, como incendios. Se identificó el ruido de fondo y se reemplazó con valores de cero. Los valores de los datos oscilan entre 1 y 63, siendo este último el máximo nivel de luz (NOAA, 2019).

### Tratamiento de la información

En esta etapa de la investigación se tuvieron en cuenta como referencia los trabajos desarrollados por el Grupo de investigación sobre Energía, Territorio y Sociedad de la Universitat Autònoma de Barcelona (Nel.lo *et al.*, 2016; Checa *et al.*, 2017; Checa & Nel.lo, 2018).

El tratamiento digital de las imágenes y la elaboración de la cartografía se llevó a cabo en los *softwares* *TerrSet* versión 18.31 (Clark University, 2017) y *QGIS* versión 3.4.12 (QGIS Development Team, 2019). Una vez descargados los datos, se procedió a configurar el proyecto bajo el EPSG: 4326 (Coordenadas Geográficas – Datum WGS 84'), esto permitió tener toda la base de datos espacial unificada bajo un mismo sistema de referencia de coordenadas que facilitó la posterior combinación y análisis entre diferentes bases de datos espaciales. El paso siguiente fue realizar el recorte de las imágenes a la correspondiente área de estudio para lo cual fue necesario definir un cuadrante con las coordenadas extremas.

Como el objetivo de la investigación gira en torno al conocimiento de la extensión de la superficie urbana a partir de la superficie iluminada y, como no todos los píxeles con valores mayores que cero corresponden a las áreas de cobertura urbana, se estimó un valor mínimo (intensidad lumínica) de referencia ligado a diversas construcciones materiales humanas (Checa & Nel.lo, 2018)<sup>4</sup> que en el presente trabajo se denominan de manera amplia infraestructura gris. Para concretar este paso, era necesario contar con una capa de cobertura urbana que permitiera realizar una superposición sobre las imágenes nocturnas y poder aproximar un valor mínimo como umbral de luminosidad urbana. Para extraer la cobertura urbana se realizó un proceso de clasificación no supervisada (ISOCCLUS)<sup>5</sup> para lo cual se especificaron 15 categorías a través de tres iteraciones sobre siete bandas de la imagen *Landsat 5 TM*, dicha técnica se encuentra dentro de los procedimientos de análisis por agrupamiento (cluster) y se ha utilizado de manera satisfactoria con objetivos similares (Da Silva *et al.*, 2014).

Una vez lograda la cobertura urbana se realizó un proceso de extracción de los valores mínimos (Estadísticas de zona) y posteriormente se promediaron para llegar a establecer el valor del umbral de luz urbana en 44.41.

A partir del valor de referencia se delimitó la superficie de luminosidad urbana en las imágenes de 1992, 2002 y 2012, tomándose en consideración aquellos píxeles que tuvieran valores  $\geq 44.41$ .

Las imágenes resultantes de los procesos anteriores permitieron derivar los principales resultados de la investigación. Primero, se vectorizaron para luego calcular la

<sup>4</sup> Esto no quiere decir, que los usos urbanos no se puedan extender en aquellas áreas con niveles de luz por debajo del valor de referencia establecido, como aclara Checa & Nel.lo (2018) el concepto de luz urbana debe entenderse en términos de intensidad más que de exclusión.

<sup>5</sup> "ISOCCLUS es un clasificador no supervisado auto-organizado iterativo basado en un concepto similar a la reconocida rutina ISODATA de Ball y Hall (1965) y a las rutinas de grupo como los procedimientos media-H y media-K" (Eastman, 2012).

superficie (km<sup>2</sup>) de luminosidad urbana. Esto a su vez, también permitió obtener los valores relativos de incremento de la superficie para cada una de las décadas analizadas.

La secuencia de pasos metodológico-técnico se encuentran explicitados en el diagrama de flujo de la Figura 1.

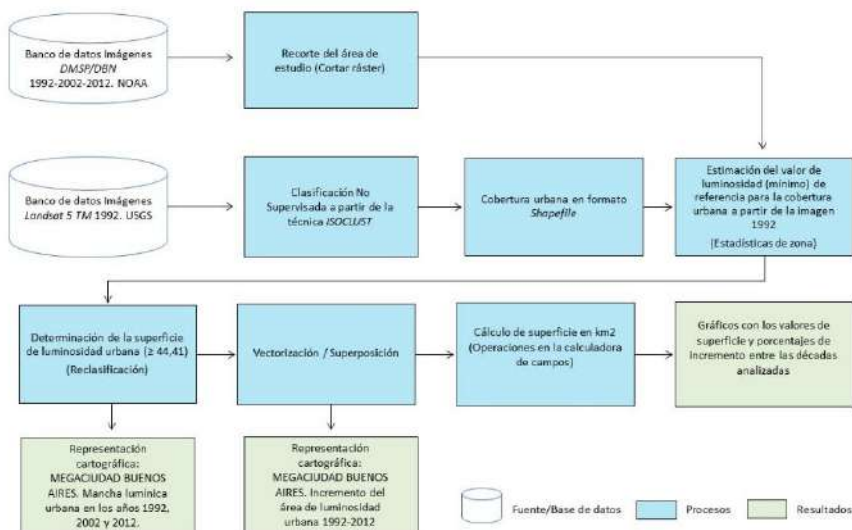


Figura 1. Diagrama de flujo de procedimientos.

## Aplicación

### Área de estudio

La gran mayoría de los estudios que intentan abarcar espacialmente a la principal área urbana argentina lo hacen a partir de considerar el concepto de Gran Buenos Aires (GBA) (INDEC, 2003) que incluye a la Ciudad Autónoma de Buenos Aires como ciudad central y 24 municipios de la Provincia de Buenos Aires que la rodean y forman una amplia superficie de conurbación. Considerando su vinculación con las áreas urbanas cercanas y separadas de la aglomeración central, la Dirección Provincial de Ordenamiento Urbano y Territorial (DPOUyT, 2007) la amplía, anexando 16 municipios a la tradicional definición de GBA, llevando su extensión a lo que fuera considerada la entidad funcional superando su definición física (aglomeración) y sus decenas de definiciones legales (municipios).

Aunque son semicírculos por haberse desarrollado en el frente costero del río de la Plata, se considera que el GBA tiene un área central y se extiende mediante la superposición de tres anillos y ocho sectores de crecimiento representado por sus diferentes ejes de expansión (Buzai y Marcos, 2012). Su ampliación actual se produciría en un cuarto anillo, aunque a esa distancia el área urbana se desarrolla principalmente de manera sectorial.

Considerando la mayor amplitud, el presente trabajo considera la extensión de un radio teórico de 90 km desde el CBD y la última representación en escala 1:400 000. De esta manera queda definida una superficie que permite captar la extensión urbana en plenitud.

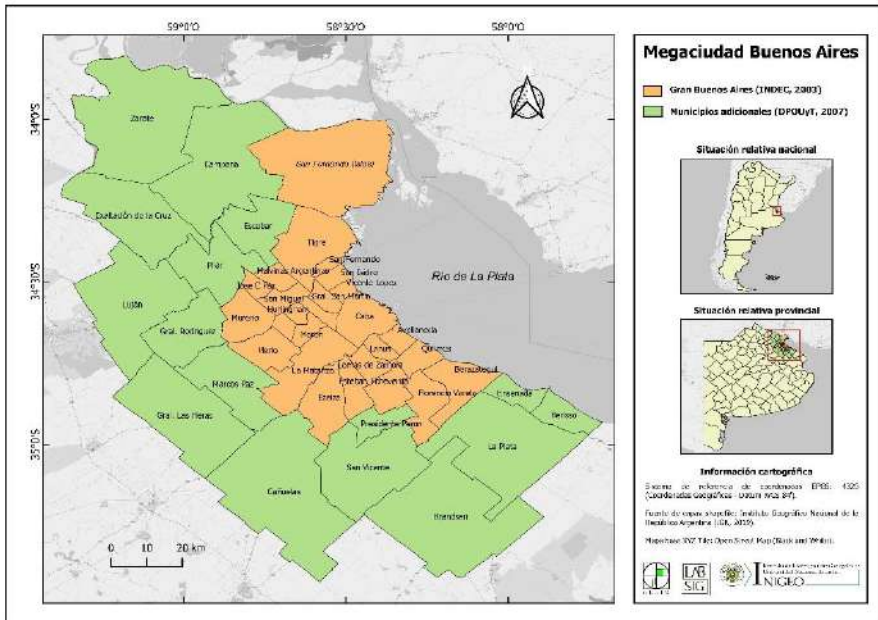


Figura 2. Área de estudio: Unidades espaciales municipales.

### Expansión y conurbación: 1992, 2002 y 2012

El procesamiento digital de las imágenes satelitales nocturnas brindó como resultado las composiciones cartográficas presentadas en la Figura 3, en donde los valores que

definen la aglomeración adquieren colores claros que representan las mayores luminosidades indicadoras de la infraestructura gris.<sup>6</sup>

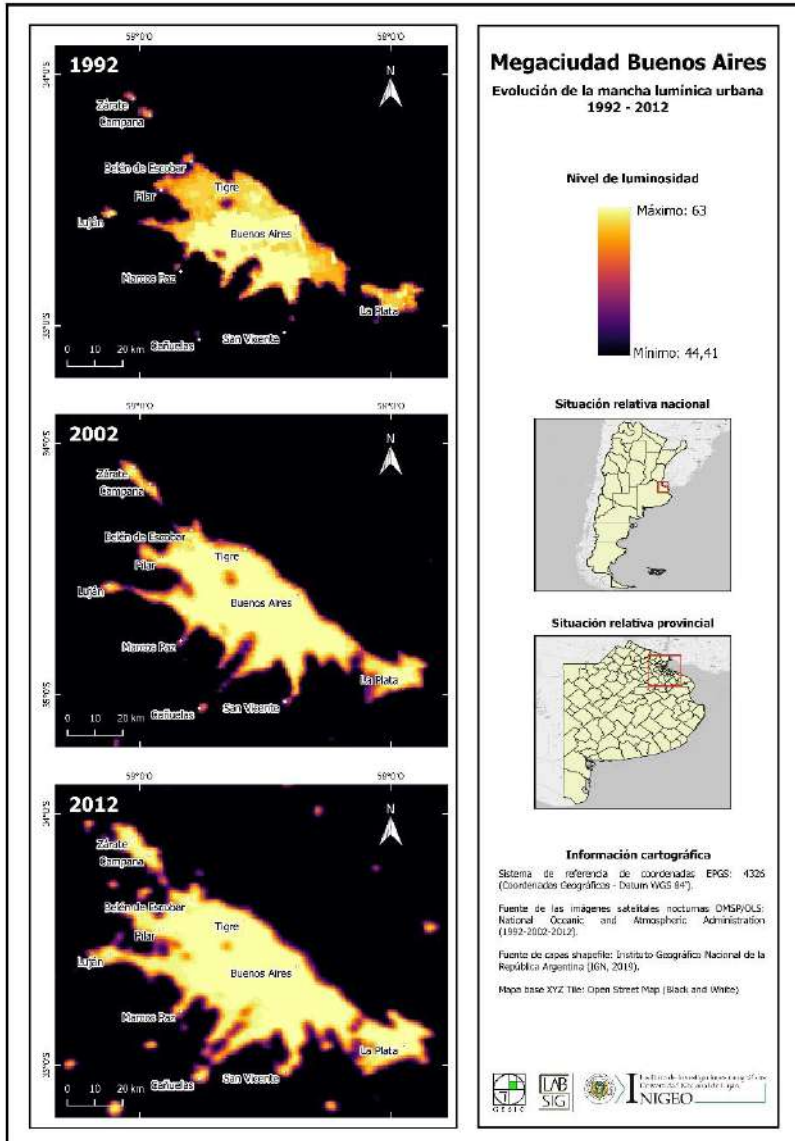


Figura 3. Procesamiento de imágenes satelitales nocturnas.

<sup>6</sup> En este rango de valores es posible distinguir los niveles de actividad humana más altos, que para el caso del presente estudio se focaliza principalmente en las coberturas urbanas.

En el año 1992, se aprecia un área urbana compacta con sectores lineales de crecimiento (denominados *tentáculos*). Todas las áreas urbanas cercanas a la Megaciudad Buenos Aires se encuentran claramente separadas. Esta situación que cambia en 2002 cuando se extienden la totalidad del crecimiento urbano lineal generando el último proceso de expansión con conurbación de sus ciudades cercanas para incorporarlas en la aglomeración central.

Estas anexiones se consolidan en el año 2012 y se convierten en la base para generar la cartografía síntesis, a partir de la superposición de capas de los diferentes años mediante el modelado cartográfico con SIG. De esta manera el mapa síntesis se convierte en la fuente empírica del proceso analizado y ámbito principal de interpretación.

### Cartografía síntesis

El mapa final de la Figura 4, muestra el vínculo entre el análisis geográfico y la Cartografía. La primera apoyándose en los conceptos centrales del análisis espacial que se representan: 1) *localización* en la ubicación específica de cada localidad en el espacio absoluto, 2) *distribución espacial* en el reparto de todas las localidades en el área de estudio y la forma de la mancha urbana total, 3) *asociación espacial* entre la representación urbana y la base topográfica y 4) *evolución espacial* a partir de superponer los diferentes años (*chess map*) y generar valores de superficie para los diferentes períodos.

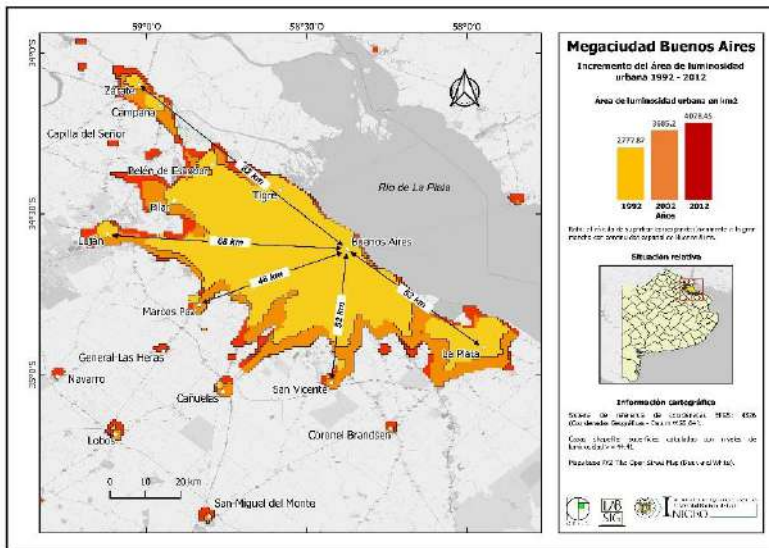


Figura 4. Megaciudad Buenos Aires. Cartografía síntesis de su último proceso de conurbación



Los colores utilizados corresponden a una variable de tipo ordenada ya que se establece una relación de sucesión temporal en los diferentes años con mayor intensidad hacia la actualidad, espacialmente hacia la periferia extrema.

La aglomeración en 1992 es completa y luego se le anexas las superficies existentes hasta el 2002 y hasta el 2012. Contemplando el análisis realizado por Bertin (1988), el *color* se utiliza en el sentido de la variable visual *valor* ya que estos colores no solamente representan diferencias sino que principalmente establecen un ordenamiento para la sucesión temporal.

El color amarillo presenta la aglomeración en 1992, una gran ciudad compacta y áreas urbanas a su alrededor que quedan incluidas en la mancha urbana en diferentes momentos. En la década 1992-2002, en naranja, se anexas ciudades en cuatro ejes: Luján (oeste), Marcos Paz (sudoeste), San Vicente (sur) y Gran La Plata (sudeste) y en la década 2002-2012, en rojo, en un eje: Zárate-Campana (Noroeste). Este último es el más lejano, allí la Megaciudad Buenos Aires llega a una expansión de 90 km dentro de un promedio de 61 km para la totalidad de anexiones. Los valores poblacionales incluidos se presentan en la Tabla 2. El incremento de la superficie lumínica urbana fue de un 32.69% en la década 1992-2002, mientras que, para la década de 2002-2012 fue de un 10.67%.

**Tabla 2**  
**Población de las recientes áreas urbanas conurbadas**

<i>Área urbana</i>	<i>Rango</i>	<i>Población</i>
Gran La Plata*	Área urbana conturbada	787.294
Luján	Ciudad de tamaño intermedio	67.266
Marcos Paz	Ciudad de tamaño intermedio	50.460
San Vicente	Ciudad pequeña	21.411
Zárate-Campana	Ciudad doble (aglomeración)	185.382
	Total	1.111.813

\* La Plata, Berisso y Ensenada.

**Fuente:** Elaboración propia en base al Censo Nacional de Población, Hogares y Viviendas, 2010.

De acuerdo a los reportes presentados por Naciones Unidas (UN, 2014, 2018) la población de la Megaciudad Buenos Aires evolucionó en las últimas décadas, en millones de habitantes, de la siguiente manera: 2000 (12.50), 2010 (15.00) y 2018 (14.97) con proyección al 2030 (16.50). Esto indica que de acuerdo a la última

conurbación en el año 2010 ya estaría a superando los 16 millones de habitantes<sup>7</sup> en una superficie aproximada de 4 000 km<sup>2</sup> con una densidad de población: 4 000 hab/km<sup>2</sup>.

### Consideraciones finales

El trabajo realizado constituye un aporte del análisis espacial basado en el procesamiento digital de imágenes satelitales nocturnas y SIG aplicados al estudio de la estructura del espacio geográfico y su evolución, poniendo su foco de atención en la expansión urbana para el descubrimiento de situaciones actuales y tendencias futuras.

Las manifestaciones espaciales de diversas relaciones sociales brindan la posibilidad de abordar el estudio de la materialidad empírica en la escala humana de un sistema complejo. El espacio geográfico manifestado por la superficie terrestre es el *nivel focal* y allí son aplicados los principales conceptos del análisis espacial mediante el uso de las actuales geotecnologías.

En este marco la confluencia entre teoría, metodología y técnicas de análisis espacial permitieron llegar a una novedosa comprobación sobre la evolución espacial reciente de la Megaciudad Buenos Aires.

Durante el cambio de siglo pudo verificarse que su último crecimiento se transformó en un notable proceso de conurbación que incorporó en la aglomeración central seis áreas urbanas que rodeaban sus bordes a distancias de entre 42 km y 82 km del CDB, de esta manera, la Megaciudad Buenos Aires experimentó una nueva expansión con sectores de conurbación.

Las imágenes satelitales nocturnas, permiten obtener una novedosa visión de la aglomeración a partir de mostrar claramente el espacio adaptado y sistema de flujos bajo el concepto de *infraestructura gris*. Surge claramente la materialidad empírica en su máxima expresión, motivo por el cual mediante su uso fue posible tener una clara perspectiva de la tendencia del alcance urbano. Esto brinda nuevas dimensiones a la megaciudad en distancias, superficies y valores poblacionales, además de mostrar su dinamismo ante un crecimiento urbano mixto, que se hace evidente ante la evolución conjunta del crecimiento por urbanizaciones cerradas (dispersa) y la conurbación (compacta) en la periferia extrema.

La Megaciudad Buenos Aires es un importante ejemplo del desarrollo urbano de una de las mayores áreas urbanas de América Latina en el marco socio-económico y político de la actual globalización.

<sup>7</sup> Coincidente con la aproximación de City Population (2019) donde consignan la Megaciudad Buenos Aires (incluido La Plata) con 16.3 millones de habitantes.

## Bibliografía

- Amaral, S., Monteiro, A.M., Camara, G. & Quintanilha, J.A. (2006). "DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon", *International Journal of Remote Sensing*, 27(5): 855-870. [DOI:10.1080/01431160500181861](https://doi.org/10.1080/01431160500181861)
- Auge, M. (2017). *Los no lugares. Espacios del anonimato*, Gedisa, Barcelona.
- Bähr, J. (1976). "Neuere Entwicklungstendenzen lateinamerikanischer Grosstädte", *Geographische Rundschau*, 28(4): 125-133.
- Bähr, J. & Mertins, G. (1982). "A model of the social and spatial differentiation of Latin American Metropolitan cities", *Applied Geography and Development*, 19: 22-45.
- Batty, M. & Longley, P. (1994). *Fractal Cities*, Academic Press, London.
- Bauer, I. (1993). "Le 'suburbia', sommes-nous concernés?", *Urbanisme*, 1: 67-88.
- Benedict, M.A. & McMahon, E.T. (2006). *Green Infrastructure. Linking Landscapes and Communities*, Washington, Island Press.
- Berry, B.J.L. (1976). *Urbanization and counterurbanization*, New York, Arnold.
- (1985). "Islands of renewal in seas of decay", in Peterson, P.E. (ed.), *The New Urban Reality*, Washington, The Brookings Institution, pp. 69-96.
- Bertalanffy, L. von (1968). *General System Theory*, George Brazillier, New York.
- Bertin, J. (1988). *La gráfica y el tratamiento gráfico de la información*, Madrid, Taurus Ediciones.
- Borsdorf, A. (1982). "Die lateinamerikanische Grosstadt. Zwischenbericht zur Diskussion um ein Modell", *Geographische Rundschau*, 34: 498-501.
- Borsdorf, A. (1994). "Die Stadt in Lateinamerika. Kulturelle Identität und urbane Probleme", *Geographie und Schule*, 89: 3-12.
- Borsdorf, A. & Coy, M. (2009). "Megacities and Global Change: Case Studies from Latin America", *Die Erde*, 140(4): 1-20.
- Bunge, M. (2012). *A la caza de la realidad. La controversia sobre el realismo*, Barcelona, Gedisa.
- Burgess, E.W. (1925). "The growth of the city: an introduction to a research project", in R.E. Park, E.W. Burgess & R.D. McKenzie (ed.), *The City*, Chicago, University of Chicago Press, pp. 47-62.
- Buzai, G.D. (1992) "El desplazamiento suburbano de los grupos sociales de alto nivel económico: el caso de Buenos Aires", *Análisis Geográfico*, 4(8): 25-34.
- (2014). *Mapas Sociales Urbanos*, Buenos Aires, lugar editorial.
- (2015a). "Sistemas de Información Geográfica, 50 años, *Boletín Red GESIG*, 1: 2-3.

- (2015b). “Urban models in the study of Latin American cities”, *Innsbrucker Geographische Studien*, 40: 271-288.
- Buzai, G.D. & Baxendale, C.A. (1998). “Buenos Aires (1869-1991). La geometría urbana como representación de una historia económica y sociodemográfica”, *Signos Universitarios*, 18(34): 71-88.
- Buzai, G.D. & Lanzelotti, S.L. (dir.) (2019). *Atlas de Geografía Humana de la cuenca del río Luján*, Luján, Instituto de Investigaciones Geográficas.
- Buzai, G.D., Lemarchand, G.; & Schuschny, A. (1998).” Aplicación de la geometría fractal al estudio del medio ambiente y las geociencias”, en Matteucci, S.D.; Buzai, G.D. (eds.), *Sistemas Ambientales Complejos: herramientas de análisis espacial*, Buenos Aires, Eudeba, pp. 347-364.
- Buzai, G.D. & Marcos, M. (2012). “The social map of greater Buenos Aires as empirical evidence of urban models”, *Journal of Latin American Geography*, 11(1): 67-78.
- Buzai, G.D. & Ruiz, E. (2012). “Geotecnósfera. Tecnologías de la Información Geográfica en el contexto global del sistema mundo”, *Anekumene*, 4: 88-106.
- Carter, H. (1983). *El estudio de la Geografía Urbana*, Madrid, IEAL.
- Checa, J., & Nel.lo, O. (2018). “Urban Intensities. The Urbanization of the Iberian Mediterranean Coast in the Light of Nighttime Satellite Images of the Earth”, *Urban Science*, 2(4): 1-17. <https://doi.org/10.3390/urbansci2040115>.
- Checa, J., Gomà, A., López, J., Martín, J., & Nel.lo, O. (2017). “Energía y proceso de urbanización en España. Nuevas perspectivas de investigación a partir de la imagen satelital nocturna de la tierra”, en M. Zaar, M. Vasconcelos, & H. Capel (eds.), *La electrificación y el territorio. Historia y futuro*, Barcelona, Universitat de Barcelona, pp. 1-20. Recuperado de <http://www.ub.edu/geocrit/Electr-y-territorio/ChecaGoma.pdf>
- City Population (2019). *Major Agglomerations of the World*. Recuperado de <https://www.citypopulation.de/world/Agglomerations.html>
- Clark University. (2017). TerrSet. Versión 18.31. Massachusetts, Clark Labs.
- Da Silva, C.J., Insaurralde, J.A., & Cardozo, O.D. (2014). “Cartografía de coberturas del suelo mediante sensores remotos, de la ciudad de Resistencia”, Argentina (2013). *Revista Geográfica Digital*, 11(21): 1-18. <https://doi.org/10.30972/geo.11212103>
- Dobson, J.E. (1983) “Automated Geography”, *The Professional Geographer*, 35(2): 135-143.
- DPOUyT (2007). *Lineamientos Estratégicos para la Región Metropolitana de Buenos Aires*, La Plata, Dirección Provincial de Ordenamiento Urbano y Territorial de la Provincia de Buenos Aires.

- Elvidge, C.D., Baugh, K., Zhizhin, M. & Chi Hsu, F. (2013). “Why VIIRS data are superior to DMSP for mapping nighttime lights”, in *Proceedings of the Asia-Pacific Advanced Network*, 35: 62-69. <https://doi.org/10.7125/APAN.35.7>
- Elvidge, C.D., Baugh, K., Zhizhin, M., Chi Hsu, F. & Ghosh, T. (2017). “VIIRS night-time lights”, *International Journal of Remote Sensing*, 38(21): 5860-5879. <https://doi.org/10.1080/01431161.2017.1342050>
- Eastman, R.J. (2012). *Idrisi selva. Guía para SIG y procesamiento de imágenes*, Massachusetts, Clark University.
- Ford, L. (1996). “A new and Improved Model of Latin American City Structure”, *Geographical Review*, 86(3): 437-440.
- García, R. (2006). *Sistemas complejos*, Barcelona, Gedisa.
- Griffin, E. & Ford, L. (1980). “A model of Latin American city structure”, *Geographical Review*, 70(4): 397-422.
- Harris, C.D. & Ullman, E.L. (1945). “The nature of cities”, *The Annals of the American Academy of Political and Social Sciences*, 242: 7-17.
- Henderson, M., Yeh, E.T., Gong, P., Elvidge, C. & Baugh, K. (2003). “Validation of urban boundaries derived from global night-time satellite imagery”, *International Journal of Remote Sensing*, 24(3): 595-609. <https://doi.org/10.1080/01431160304982>
- Hoyt, H. (1939). *The Structure and Growth of Residential Neighborhoods in American Cities*, Washington, Federal Housing Administration.
- Instituto Geográfico Nacional —IGN (2019). *Datos abiertos. Capas SIG*. Recuperado de <https://www.ign.gob.ar/NuestrasActividades/InformacionGeoespacial/CapasSIG>
- Instituto Nacional de Estadística y Censos —INDEC (2003). *¿Qué es el Gran Buenos Aires?*, Buenos Aires, Publicaciones del INDEC. Recuperado el 15 de junio de 2019, de <https://web.archive.org/web/20101113161116;http://www.indec.gov.ar/nuevaweb/cuadros/4/folleto%20GBA.pdf>
- (1 de noviembre de 2019). *Estadísticas de población*. Recuperado de <https://www.indec.gov.ar/>
- Indovina, F. (2009). *Dalla città diffusa all'archipelago metropolitano*, Milano, Franco Angeli.
- Janoschka, M. (2002). “Stadt der Inseln” Buenos Aires: Abschottung und Fragmentierung als Kennzeichen einer neuen Stadtmodells“, *RaumPlanung*, 101: 65-70.
- Kamarajugedda, S.A.; Mandapaka, P.V. & Lo, E.Y. (2017). “Assessing urban growth dynamics of major Southeast Asian cities using night-time light data”, *International Journal of Remote Sensing*, 38(21): 6073-6093. <https://doi.org/10.1080/01431161.2017.1346846>

- Liu, L. & Leung, Y. (2015). "A study of urban expansion of prefectural-level cities in South China using night-time light images", *International Journal of Remote Sensing*, 36: 5557-5575. <https://doi.org/10.1080/01431161.2015.1101650>
- Martin Oriol, J. (2016). "Observing changes in urban form through night satellite images in Buenos Aires and Lima", en O. Nel.lo & R. Mele (eds.), *Cities in the 21st Century*, London and New York, Routledge, pp. 93-95.
- Mertins, G. (1980). "Typen inner- und randstädtischer Elendsviertel in Grosstädten des andinene Südamerika", *Lateinamerika Studien*, 7: 269-295.
- Mitchell, J.G. (2001). "La expansión urbana", *National Geographic*, 9(1): 82-105.
- Montes Galbán, E. (2017). "Ideas sobre tecnociencia y experimentación digital en geografía", *Vientos del Norte*, 2: 7-17. Recuperado de <http://editorial.unca.edu.ar/Publicacione%20on%20line/CD%20VIENTOS%20017/PDF/N5%20Vol%20II/1-Ideas%20sobre%20tecnociencia.pdf>
- (2019). "Cuenca. Luminosidad nocturna", en G.D. Buzai & S.L. Lanzelotti (dir.), *Atlas de Geografía Humana de la cuenca del río Luján*, Luján, Instituto de Investigaciones Geográficas, pp. 168-169.
- (2020). "Expansión y densificación urbana del Gran Buenos Aires (2012-2019) a partir de imágenes satelitales nocturnas", *Revista Geográfica Digital*, en prensa.
- Morello, J., Buzai, G., Baxendale, C.A., Matteucci, S.D., Rodríguez, A., Godagnone, R.E. & Casas, R.R. (2000). "Urbanización y consumo de tierra fértil", *Ciencia hoy*, 18(65): 50-61.
- National Geographic (1998). *Población y recursos* [Mapa]. 1:54.109.440, Washington, D.C., National Geographic Maps.
- (2004). *La tierra de noche* [Mapa]. 1:46.360.000, Washington, D.C., National Geographic Maps.
- (2005). *África la huella humana* [Mapa]. 1:14.636.000. Washington, D.C., National Geographic Maps.
- Nel.lo, O., López Redondo, J., Martín Oriol, J. & Checa Rius, J. (2016). *La luz de la ciudad. El proceso de urbanización en España a partir de las imágenes nocturnas de la Tierra*, Barcelona, Universitat Autònoma de Barcelona.
- Nel.lo, O., López, J., Martín, J. & Checa, J. (2017). "Energy and urban form. The growth of European cities on the basis of night-time brightness", *Land Use Policy*, 61: 103-112. <https://doi.org/10.1016/j.landusepol.2016.11.007>
- National Oceanic and Atmospheric Administration —NOAA (2019). *Version 4 DMSP-OLS Nighttime Lights Time Series*. Recuperado de <https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>

- Pérez Machado, R.P.; Vizeu Barrozo, L., Luchiari, A. & Small, C. (2014). "Identification of the urban continuum of Sao Paulo Metropolitan Agglomeration with people, lights and land use. First results of a methodological approach", *Annual Conference of the IGU Urban Commission Poznan*, IGU, pp. 1-42. Recuperado de [https://www.researchgate.net/publication/266202035\\_Identification\\_of\\_the\\_urban\\_continuum\\_of\\_Sao\\_Paulo\\_Metropolitan\\_Agglomeration\\_with\\_people\\_lights\\_and\\_land\\_use\\_First\\_results\\_of\\_a\\_methodological\\_approach](https://www.researchgate.net/publication/266202035_Identification_of_the_urban_continuum_of_Sao_Paulo_Metropolitan_Agglomeration_with_people_lights_and_land_use_First_results_of_a_methodological_approach)
- Potere, D.; Schneider, A., Angel, S. & Civco, D. L. (2009). "Mapping urban areas on a global scale: which of the eight maps now available is more accurate?", *International Journal of Remote Sensing*, 30: 6531-6558. <https://doi.org/10.1080/01431160903121134>
- QGIS Development Team (1 de noviembre de 2019). *QGIS. Un Sistema de Información Geográfica libre y de Código Abierto*. Recuperado de <https://qgis.org/es/site/>
- Quaíni, K. (2014). *Procedimiento metodológico para la generación de mapas de interfase urbano-vegetación rural*, Córdoba, Proyecto BPR-BID.
- Rosnay, J. de (1977). *El macroscopio. Hacia una visión global*, Madrid, AC.
- Roy Chowdhury, P.K., Maithani, S. & Dadhwal, V.K. (2012). "Estimation of urban population in Indo-Gangetic Plains using night-time OLS data", *International Journal of Remote Sensing*, 33: 2498-2515. <https://doi.org/10.1080/01431161.2011.616550>
- Schnore, L.F. (1965). "On the spatial structure of cities in the two Americas", en Ph. M. Hauser and L.F. Schnore (eds.), *The Study of Urbanization*, New York, John Wiley & Sons.
- Sjoberg, G. (1960). *The Preindustrial City. Past and Present*, Glencoe, Free Press.
- Soja, E. (2000). *Posmetropolis. Critical studies of cities and regions*, Oxford, Blackwell.
- Small, C., Pozzi, F. & Elvidge, C.D. (2005). *Spatial analysis of global urban extent from DMSP-OLS night lights*, 96: 277-291. <https://doi.org/10.1016/j.rse.2005.02.002>
- United Nations —UN (2014). *Urban Agglomerations 2014*, New York, Department of Economic and Social Affairs, Population Division.
- (2018). *The World's Cities in 2018 – Data Booklet*. New York, Department of Economic and Social Affairs, Population Division.
- United States Geological Survey —USGS (2019). *Earth Explorer*, Recuperado de <https://earthexplorer.usgs.gov/>

- Xin, X., Liu, B., Di, K., Zhu, Z., Zhao, Z., Liu, J., Zhang, G. (2017). “Monitoring urban expansion using time series of night-time light data: a case study in Wuhan, China”, *International Journal of Remote Sensing*, 38(21): 6110-6128. <https://doi.org/10.1080/01431161.2017.1312623>
- Whyte, W. (1958). “Urban Sprawl”, in *The Editors of Fortune. The Exploding Metropolis: A Study of the Assault on Urbanism and How Our Cities Can Resist It*, New York, Anchor Books.
- Zhou, Y., Smith, S.J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B. & Elvidge, C.D. (2015). “A global map of urban extent from nightlights”, 10(5): 1-11. <https://doi.org/10.1088/1748-9326/10/5/054011>



# RESEÑAS





# Manual of Digital Earth

Guo, H., Goodchild, M.F. & Annoni, A. (eds.), Singapore, Springer, 2020, 852 pp.

ISBN: 978-981-32-9914-6. DOI <https://doi.org/10.1007/978-981-32-9915-3>

Luis M. Vilches-Blázquez  
Instituto Politécnico Nacional, México

El concepto de la Tierra Digital (*Digital Earth*) fue acuñado por primera vez en el libro de Al Gore titulado “*Earth in the Balance*” (Gore, 1992). En 1999, la Tierra Digital fue concebida como una representación tridimensional del planeta con múltiples resoluciones que permitirían buscar, visualizar y poner en valor grandes cantidades de información georreferenciada en entornos físicos y sociales. Dicho sistema permitiría a los usuarios navegar a través del espacio y tiempo, acceder a datos históricos, así como a predicciones futuras, y su uso sería promovido, de igual forma, por científicos, responsables políticos y niños (Gore, 1999).

Este concepto fue motivado por la idea de que las preguntas complejas no se pueden responder desde un solo dominio, sino que se necesitan múltiples disciplinas (Janowicz & Hitzler, 2012). Así, la Tierra Digital trajo consigo la visión de una plataforma común que apoyaría la cooperación nacional e internacional para el desarrollo sostenible global y que conllevaría un nuevo punto de crecimiento económico y bienestar social (International Society for Digital Earth, 2012).

Tras cumplirse dos décadas de este concepto (*Digital Earth*), se han logrado muchos avances, definiendo estándares, implementando prototipos, popularizando productos de la industria y creando aplicaciones a través de iniciativas en todo el mundo (Yang *et al.*, 2010). Muchos de los elementos no solo están disponibles y son utilizados diariamente por cientos de millones de personas en todo el mundo (Craglia *et al.*, 2008), sino que se están convirtiendo en un elemento clave para promover una comprensión y análisis global (Guo, 2016). De esta manera, la consolidación y avance de los elementos de la Tierra Digital están contribuyendo a la implementación exitosa de la Tierra del Futuro (*Future Earth*) (Songnian & Jun, 2017).

Tomando en cuenta este contexto, este manual, elaborado con la participación de más de 100 colaboradores y editores de 18 países, aborda diversas preguntas con respecto a la relevancia de la Tierra Digital, su futuro y su potencial para apoyar el desarrollo científico y el progreso social. Así, el manual presenta la siguiente estructura:

- El capítulo de introducción proporciona una visión general de la Tierra Digital y su marco de referencia, su respuesta global, evolución y relación con iniciativas científicas globales, así como la respuesta a desafíos globales que son detallados en capítulos posteriores.
- La Parte I, presenta uno de los puntos claves en este libro, asociado a una visión general sobre conceptos, teorías y métodos fundamentales relacionados con la Tierra Digital, como son: los Sistema de Mallas Globales y Discretas (*DGGS*), sensores remotos, sistema global de navegación por satélite (*GNSS*), infraestructuras de información geoespacial, *zoom* continuo y transformación multi-escala, *Big Data* y computación en la nube, inteligencia artificial y técnicas de aprendizaje automático, Internet de las cosas (*IoT*) y redes sociales. Además, esta parte describe algunas tecnologías para el procesamiento de datos geoespaciales (computación de alto rendimiento y geoprocesamiento en línea y distribuido), y proporciona detalles sobre geovisualización y realidad extendida (realidad virtual, aumentada y mixta) con ejemplos realmente interesantes. Cada uno de estos temas está organizado como un capítulo independiente.
- La Parte II describe diversas aplicaciones de la nueva generación de sistemas de la Tierra Digital que tienen como objetivo satisfacer la demanda de intercambio de información para su contribución en diferentes escenarios. Por lo tanto, un gran número de estos capítulos ofrecen evidencia del potencial, limitaciones y desafíos de la Tierra Digital y, más específicamente, sobre su nueva visión denominada *Big Earth Data* (Guo, 2016). Así, se describe el nuevo entorno para generar, compartir e integrar datos, información y conocimientos multidimensionales, multi-escalares y multitemporales en apoyo de los Objetivos de Desarrollo Sostenible y, más concretamente, en aquellos escenarios relacionados con el cambio climático, mitigación de desastres y ciudades y comunidades sostenibles. Además, esta parte reúne tres capítulos de especial interés que discuten el patrimonio digital, ciencia ciudadana y aspectos económicos. El primero muestra cómo se utilizan los medios digitales para preservar, proteger, estudiar y presentar el patrimonio natural y cultural. En el segundo, se aborda una visión general del concepto de ciencia ciudadana, sus prácticas tecnológicas e impactos sociales, así como sus desafíos, mientras el último se centra en el valor económico de la Tierra Digital que, según los autores, depende de perspectivas asociadas con para quién, con qué propósito y con cuándo.
- La Parte III, aglutina las contribuciones de una serie de regiones a los objetivos de la Tierra Digital. De este modo, el manual destaca el establecimiento de la *Infraestructura Europea de Datos Espaciales*, los programas *Galileo* y *Copernicus*, las plataformas basadas en la nube y el panorama de la ciencia ciudadana en el caso europeo. Con respecto a las propuestas australianas, se discuten diversas iniciativas nacionales (por ejemplo, *Geoscape*, *Australian Geoscience Data Cube*

o *Digital Earth Australia*) y se proporcionan casos de uso en varios dominios, lecciones aprendidas y oportunidades para este país. En China, el desarrollo de las tecnologías y aplicaciones de la Tierra Digital se ha basado en estrategias y políticas, respaldadas por tecnologías de red e información (por ejemplo, 5G, IPv6, infraestructura de computación en la nube o plataformas de *Big Data*), que se han utilizado en algunas ciudades y aplicaciones. Finalmente, una serie de datos e iniciativas muestran una consolidada comunidad en Rusia que presenta incluso antecedentes históricos.

- La Parte IV de este manual proporciona una visión general de las estructuras actuales de enseñanza y aprendizaje donde la Tierra Digital está presente y destaca la importancia de un marco de aprendizaje integrado para una mejor preparación de (futuros) estudiantes en el ámbito geoespacial. Además, se tratan los códigos éticos asociados con la (geo) privacidad y la dignidad humana, basándose en los escenarios de la legislación de Europa y Estados Unidos, como referencia, aunque se realizan menciones a la legislación y directrices del continente africano y China.
- Finalmente, el libro concluye con la revisión del pasado y presente de la Tierra Digital, seguido de la presentación de un conjunto de desafíos y tendencias futuras asociadas con la nueva era del *Big Earth Data*. Adicionalmente, los autores especulan sobre cómo la Tierra Digital puede evolucionar en los próximos años en función de los impulsores científicos, avances tecnológicos, adopción de aplicaciones y la construcción de una comunidad (virtual y física) relevante.

En conclusión, recomiendo este libro a todos los lectores de la *Revista Cartográfica* y a aquellos investigadores, estudiantes y tomadores de decisiones relacionados o interesados en la Tierra Digital, así como al conjunto de la comunidad geoespacial. Las más de 800 páginas que componen esta obra recogen el contenido más actualizado, a la fecha, y uno de los compendios más detallados sobre esta disciplina promovido por la Sociedad Internacional para la Tierra Digital (*International Society for Digital Earth - ISDE*). Por todo lo anterior, este libro se convierte en una obra de referencia y, en consecuencia, merece un lugar en cada biblioteca universitaria y en las estanterías de aquellos profesionales que trabajan en el ámbito geoespacial.

## Bibliografía

- Craglia, M., Goodchild, M.F., Annoni, A., Camara, G., Gould, M., Kuhn, W. & Parsons, E. (2008). "Next-generation Digital Earth. A position paper from the Vespucci Initiative for the Advancement of Geographic Information Science", *International Journal of Spatial Data Infrastructure Research*, 3(3): 146-167.
- Gore, A. (1992). *Earth in the balance: ecology and the human spirit*, Houghton Mifflin, Boston.

- (1999). “The Digital Earth: Understanding Our Planet in the 21st Century” *Photogrammetric Engineering and Remote Sensing*, 65(5): 528.
- Guo, H. (2016). “Digital Earth and Future Earth”, *International Journal of Digital Earth*, 9(1): 1-2. <https://doi.org/10.1080/17538947.2015.1135667>
- International Society for Digital Earth (2012). Statutes of the International Society for Digital Earth. <http://www.digitalearth-isde.org/statues>
- Janowicz, K. & Hitzler, P. (2012). The Digital Earth as Knowledge Engine (editorial). *Semantic Web Journal*, 3(3): 213-221. <https://doi.org/10.3233/SW-2012-0070>
- Songnian, L. & Jun, C. (2017). Supporting future earth with global geospatial information, *International Journal of Digital Earth*, 10(4): 325-327. <https://doi.org/10.1080/17538947.2016.1275832>
- Yang, C., Raskin, R., Goodchild, M. & Gahegan, M. (2010). “Geospatial Cyberinfrastructure: Past, Present, and Future”. *Computers, Environment and Urban Systems*, 34 (4): 264-277. <https://doi.org/doi:10.1016/j.compenvurbsys.2010.04.001>

# Análisis espacial con R: Usa R como un Sistema de Información Geográfica

Jean-François, Mas, European Scientific Institute, 2018, 151 pp. ISBN: 978-608-4642-66-4. [Disponible en http://eujournal.org/files/journals/1/books/JeanFrancoisMas.pdf](http://eujournal.org/files/journals/1/books/JeanFrancoisMas.pdf)

Elena, Chicaiza Mora  
R-Ladies, Quito, Ecuador

El libro titulado *Análisis espacial con R: Usa R como un Sistema de Información Geográfica*, realiza una breve introducción de los principales conceptos para el manejo del lenguaje y entorno de programación R (tipos de objetos y operaciones básicas) para posteriormente acercar al lector al uso de la librería o paquete *sf*, para datos espaciales en formato vector a través de sus principales funciones para lectura, escritura y análisis. Y la librería *raster* para datos en formato *raster*.

El libro aborda, desde una perspectiva práctica y aplicativa con un lenguaje de fácil entendimiento, los primeros pasos para iniciar con el manejo de R en aplicaciones de análisis espacial; para ello, es necesario que los usuarios tengan conocimientos básicos de Sistemas de Información Geográfica, es un material de referencia en español para quienes no están familiarizados con la programación en R.

En los primeros capítulos, el autor de una forma sencilla, explica el funcionamiento de R a través de la instalación, versiones, paquetes y la interface más popular, como es R-Studio, para luego hacer una clara descripción de la estructura de los datos como objetos y objetos espaciales en R, así como sus principales operaciones básicas.

La organización de los objetos espaciales en R, y su representación espacial es abordada desde los dos modelos de datos geográficos básicos como son: vector y *raster*, con el uso de los paquetes *sf* y *raster* respectivamente. Es importante destacar la recomendación del uso del paquete *sf*, y no el paquete *sp* que fue el primero y más difundido. Esta recomendación se debe al lanzamiento del paquete *sf* en el año 2017, que presenta una estructura de los datos más simple y poco a poco desplaza el uso del paquete *sp*.

Con respecto a la importación y exportación de datos espaciales, se presentan los formatos más conocidos tanto para vector, como es el caso del formato *shape*, y

*Geodatabase* de ESRI y de *raster* con diferentes formatos de imagen y la opción de escoger la codificación de los datos.

En los capítulos 5 y 6 se presentan algunas operaciones básicas vectoriales, como: intersección, cálculo de distancias, áreas de influencia (*buffers*), agregación espacial. En el caso de las operaciones básicas *raster* que se abordan en el libro están: creación de mosaicos, recorte, reclasificación de valores, agregación espacial. Estas funciones se realizan a través de la librería *raster*. Adicionalmente, se muestran ejemplos de ejecución de métodos de re-muestreo (*resample*) como: vecino más cercano, bilineal; y operaciones de filtrado espacial. Por otro lado, se muestran operaciones de rasterización y de álgebra de mapas.

El capítulo 7, titulado “Análisis geoestadístico: detección de *hot spots*”, describe el uso de la librería *spdep* para la definición y visualización de un atributo (ponderaciones espaciales). Se presenta un estudio de caso de estadística espacial, sin embargo no se aborda el tema de geoestadística propiamente, pues en R existen las librerías *geor*, *geostat*, *Rgeostats*.

El análisis de sensores remotos, es abordado en el capítulo 8, iniciando con la importación de un conjunto de bandas. La visualización y el pre-tratamiento, así como la posibilidad de realizar fusión entre bandas multiespectrales y pancromática. El cálculo de índices espectrales, como SAVI, NDVI también es mostrado.

Por su parte, la clasificación supervisada y no supervisada de imágenes satelitales se la pueda realizar a través de la aplicación de la librería *raster*. El libro describe brevemente otros paquetes con los cuales es posible procesar diferentes fuentes de sensores remotos, incluyendo LiDAR.

En el capítulo 9, se presentan funciones para la creación de mapas sencillos, principalmente usando las librerías: *maptools* y *RColorBrewer*. Y en el capítulo 10 se hace una breve reseña del uso de R con QGIS y Dinámica EGO.

Finalmente, en la sección de Anexos, se presenta un apartado en donde se aborda el uso de los conceptos y la lógica de la librería *sp* para la gestión de datos espaciales, dado que esta librería ha sido usada extensivamente en la comunidad de usuarios de R en los últimos años.

De manera general, el libro presenta una estructura de fácil entendimiento para el manejo de datos espaciales con R, desde un punto bastante operativo, sin centrarse en algoritmos y conceptos de análisis espacial, buscando siempre, permitir al lector no desanimarse en el aprendizaje del entorno de programación que tiene R y motivando la profundización del aprendizaje a través de las comunidades y redes colaborativas.



# REVISTA cartográfica

---

## Definición de la Revista

La *Revista Cartográfica* se publica desde 1952, es una publicación de la Comisión de Cartografía del Instituto Panamericano de Geografía e Historia un Organismo Especializado de los Estados Americanos (OEA).

El objetivo de la Revista es contribuir a la difusión del conocimiento en las áreas de cartografía, geodesia e información geográfica en general; promoviendo un enfoque que contribuya a compartir investigaciones, estudios y publicaciones en las áreas de conocimiento que son de interés para la revista.

La Revista publica artículos científicos (investigación, revisión y temáticos) y reseñas en los cuatro idiomas oficiales del IPGH (español, inglés, francés y portugués). Los artículos que se propongan para su publicación deberán ser originales y no estar publicados o propuestos para tal fin en otra revista.

Desde 2016 tiene una periodicidad semestral (enero-junio / julio-diciembre). La recepción de artículos está abierta permanentemente.

---

## Proceso de evaluación por pares

En atención a las buenas prácticas editoriales la *Revista de Cartográfica* los artículos que cumplan los requisitos formales y de calidad serán evaluados por pares académicos externos a la institución editora. Se aplicará el sistema de evaluación “doble ciego”.

---

## Secciones

### Artículos:

- Artículo de investigación: documentos que presentan resultados de nuevas investigaciones, u otras contribuciones originales que sean de aporte a las ciencias y disciplinas de interés para la Revista.
- Artículo de revisión: documentos que analicen, sistematicen e integren resultados de material ya publicado sobre un área del conocimiento, con el objeto de divulgar los avances y las tendencias de desarrollo en dicha área. Requiere una cuidadosa y rigurosa revisión bibliográfica del tema, un artículo de revisión no debería considerar menos de 60 referencias directas e indirectas. Solo para áreas de conocimiento de desarrollo muy reciente se consideran artículos de revisión con un número menor de referencias.
- Artículos temáticos: documentos que respondan a una convocatoria de la Revista para conformar un número temático coordinado por un destacado especialista como editor invitado.

### Reseñas:

Documentos que presentan un análisis crítico de libros u otras monografías que sean significativa para las ciencias y disciplinas de interés de la revista, que hayan sido publicados en los últimos tres años y que cuenten con ISBN. De forma excepcional se aceptarán reseñas de libros no actuales, clásicos, que sean significativos o de referencia para análisis y estudios actuales.

---

## Norma para autores

Los artículos y reseñas se enviarán únicamente en formato digital a través de <https://www.revistasipgh.org/index.php/rcar>. Por lo tanto, es necesario que el autor este registrado de forma correcta y que disponga de usuario y contraseña.

Los archivos de texto se enviarán en formato Word, utilizando las plantillas correspondientes para que cada tipo de documento aceptado por la Revista.

Para asegurar la evaluación ciega del artículo, el archivo debe enviarse sin datos ni metadatos que permitan la identificación del autor. No se debe incluir nombre, afiliación, correo electrónico u organismo financiador en el documento. Estos datos se recogen de manera separada en la plataforma OJS. Los documentos que no cumplan los requisitos de anonimato serán retirados del proceso de evaluación.

Los Artículos deberán tener una extensión máxima de 25 páginas, incluyendo figuras, mapas, gráficos, tablas, citas y bibliografía. Deberán incluir un resumen de 300 palabras máximo, en el idioma del artículo y en inglés.

Estructura de los Artículos, en el siguiente orden:

- Título del trabajo en el idioma original
- Título del trabajo en el segundo idioma (en inglés para artículos en español, portugués o francés). Para artículos en inglés se indicará el título en español.
- Resumen (100-300 palabras) y las palabras claves (3-5) en el idioma en que esté redactado el artículo
- Resumen (100-300 palabras) y las palabras claves (3-5) en el segundo idioma (en inglés para artículos en español, portugués o francés). Para artículos en inglés se incluirá el resumen en español.
- Texto del artículo, con las tablas, figuras, mapas, etc.
- El cuerpo del texto seguirá preferentemente la estructura: introducción, metodología, resultados y discusión de resultados
- Bibliografía (que empezará en una página nueva) en formato APA 6ª edición

Nota: para la adecuada selección de las palabras clave se recomienda emplear el Tesauro de la UNESCO - Ciencias de la Tierra

Las reseñas tendrán una extensión máxima de 8.000 caracteres con espacios incluidos, y deben incluir la siguiente información:

- Título de la obra
- Nombre y apellidos del autor/a del libro (o dir./coord., en su caso)
- Año, ciudad y editorial
- N° de páginas
- ISBN edición impresa o ISBN edición digital, según corresponda.
- Nombre del autor/a de la reseña e institución
- Texto de la reseña

Las figuras, fotografías, mapas, gráficas, cuadros y tablas deberán ubicarse dentro del texto. Las imágenes, fotografía y mapas se deben enviar por separado en formato JPG o PNG, con una resolución mínima de 300 dpi/ppp (o superior de acuerdo con el tamaño de la figura) sin compresión y a color. Para la impresión del número de la Revista se modificarán a escala de grises, siempre que no afecte la correcta comunicación de la información, en particular los mapas.

Las citas y referencias bibliográficas se harán según las Normas APA. Se recomienda el uso de gestores de referencias bibliográficas, como Zotero o Mendeley. Se debe incluir enlace DOI o enlace web en todos los casos que sea posible.

Para asegurar el cumplimiento de todas las normas de tipografía y maquetación, es obligatorio el uso de las plantillas que se indican para cada tipo de contribución disponible en:

[www.revistasipgh.org/index.php/rcar/Directrices](http://www.revistasipgh.org/index.php/rcar/Directrices)

María Ester González

Editora

Departamento de Geografía, Facultad de Arquitectura, Urbanismo y Geografía

Universidad de Concepción

Victoria 486-490, CP 4030000, Concepción, Chile

Correos electrónicos:

[editor\\_revista\\_cartografica@ipgh.org](mailto:editor_revista_cartografica@ipgh.org) | [maria.ester.gonzalez@ipgh.org](mailto:maria.ester.gonzalez@ipgh.org) | [mariaesgonzalez@udec.cl](mailto:mariaesgonzalez@udec.cl)

No se devolverá el material enviado.

---

### **Función editorial del Instituto Panamericano de Geografía e Historia**

El IPGH publica seis revistas, impresas y distribuidas desde México. Estas son: *Revista Cartográfica*, *Revista Geográfica*, *Revista Geofísica*, *Revista de Historia de América*, *Antropología Americana* y *Revista de Arqueología Americana*.

Se invita a todos los investigadores y profesionales de las áreas de interés del IPGH: cartografía, geografía, historia, geofísica y ciencias afines, a que presenten trabajos de investigación para que sean publicados en nuestras revistas científicas.

Si requiere mayor información, favor de comunicarse a:

Departamento de Publicaciones

Secretaría General del IPGH

Ex-Arzobispado 29, Colonia Observatorio, 11860 Ciudad de México, México

Tels.: (+52-55) 5277-5888 / 5515-1910

Correo electrónico: [publicaciones@ipgh.org](mailto:publicaciones@ipgh.org)

Edición del  
Instituto Panamericano de  
Geografía e Historia realizada en su  
Departamento de Publicaciones  
Ex Arzobispado 29, Colonia Observatorio  
11860, Ciudad de México, México.  
Teléfono 5277-5791 5277-5888  
[publicaciones@ipgh.org](mailto:publicaciones@ipgh.org)  
2020

# ESTADOS MIEMBROS DEL INSTITUTO PANAMERICANO DE GEOGRAFÍA E HISTORIA

## **Argentina**

## EL IPGH, SUS FUNCIONES Y SU ORGANIZACIÓN

### **Belice**

El Instituto Panamericano de Geografía e Historia (IPGH) fue fundado el 7 de febrero de 1928 por resolución aprobada en la Sexta Conferencia Internacional Americana que se llevó a efecto en La Habana, Cuba. En 1930, el Gobierno de los Estados Unidos Mexicanos construyó para el uso del IPGH, el edificio de la calle Ex Arzobispado 29, Tacubaya, en la Ciudad de México.

### **Bolivia**

### **Brasil**

### **Chile**

En 1949, se firmó un convenio entre el Instituto y el Consejo de la Organización de los Estados Americanos y se constituyó en el primer organismo especializado de ella.

### **Colombia**

### **Costa Rica**

El Estatuto del IPGH cita en su artículo 1o. sus fines:

### **Ecuador**

1) Fomentar, coordinar y difundir los estudios cartográficos, geofísicos, geográficos e históricos, y los relativos a las ciencias de interés para América.

### **El Salvador**

2) Promover y realizar estudios, trabajos y capacitaciones en esas disciplinas.

### **Estados Unidos de América**

3) Promover la cooperación entre los Institutos de sus disciplinas en América y con las organizaciones internacionales afines.

### **Guatemala**

Solamente los Estados Americanos pueden ser miembros del IPGH. Existe también la categoría de Observador Permanente, actualmente se encuentran bajo esta condición: España, Francia, Israel y Jamaica.

### **Haití**

### **Honduras**

El IPGH se compone de los siguientes órganos panamericanos:

### **México**

1) Asamblea General

### **Nicaragua**

2) Consejo Directivo

3) Comisión de :

### **Panamá**

Cartografía (Costa Rica)

Geografía (EUA)

### **Paraguay**

Historia (México)

Geofísica (Ecuador)

### **Perú**

4) Reunión de Autoridades

### **República Dominicana**

5) Secretaría General (México, D.F., México)

### **Uruguay**

Además, en cada Estado Miembro funciona una Sección Nacional cuyos componentes son nombrados por cada gobierno. Cuentan con su Presidente, Vicepresidente, Miembros Nacionales de Cartografía, Geografía, Historia y Geofísica.

### **Venezuela**



Modelización y preedicción espacio-tiempo de la irradiancia solar global a corto plazo mediante redes neuronales artificiales y geoestadística *Federico Vladimír Gutiérrez-Corea, Miguel Ángel Manso-Callejo y Francisco Serradilla-García* • Patrones especiales asociados a la infestación comunitaria por vectores de la enfermedad de Chagas *Gabriela Ramírez-Hernández, Jean François Mas y Janine M. Ramsey* • Propuesta de una guía para la evaluación de la exactitud posicional de datos espaciales *Francisco Javier Ariza-López, José Luis García-Balboa, Joselyn Robledo Ceballos, José Rodríguez-Avi y Rosario Casanova* • Admissible relative errors of the parcel area while creating a land parcel map base using an orthoimage as a reference *Irene Balado y Carlos López-Vázquez* • Atlas escolares para la educación geográfica de niños y jóvenes *José Jesús Reyes Nunez* • Columbus: aportes de la teledetección y análisis de coste anisotrópico a un debate histórico *Ángel M. Felicísimo* • Geospatial data quality (ISO 19157-1): evolve or perish *Francisco Javier Ariza López, Pablo Barreira González, Joan Masó Pau, Alaitz Zabala Torres, Antonio Federico Rodríguez Pascual, Gonzalo Moreno Vergara y José Luis García Balboa* • Reproducibilidad en AGILE: experiencias, logros y recomendaciones *Carlos Granell, Barbara Hofer, Daniel Nüst, Frank O. Ostermann y Rusne Sileryte* • La situación de la producción científica latinoamericana en el área de la Ciencia de Información Geográfica *Carlos López-Vázquez y Miguel Ángel Bernabé-Poveda* • Sinfonía del Nuevo Mundo: el modelaje espacio-temporal de HGIS de las Indias, una infraestructura SIG para la América hispana borbónica *Werner Stangl y Paul Stangl* • Megaciudad Buenos Aires: Cartografía de su última expansión y conurbación mediante el procesamiento digital de imágenes satelitales nocturnas *Gustavo D. Buzai y Eloy Montes Galbán* • *Manual of Digital Earth* *Luis M. Vilches-Blázquez* • Análisis espacial con R: Usa R como un Sistema de Información Geográfica *Elena Chicaiza Mora*

ISSN 0080-2085