

FLUJOS DE TRABAJO EN BASES DE DATOS MULTIUSUARIO

Capacidades que ofrecen las bases de datos espaciales para la implementación de flujos de trabajo en entornos multiusuario

Gabriel J. Corrales Jiménez, Ana L. Garita Fernández

Programa Maestría Profesional en Sistemas de Información Geográfica y Teledetección
Universidad Nacional
Universidad de Costa Rica
Costa Rica

Resumen. En este trabajo se realiza una introducción a la edición de bases de datos espaciales, haciendo una diferenciación entre la edición versionada y la no versionada. Específicamente se desarrollan dos flujos de trabajo sencillos de edición multiusuarios no versionada utilizando PostgreSQL en combinación con QGIS, y por otra parte ArcGIS, de manera que el lector pueda tener una visión comparativa de la utilización de ambos métodos. Finalmente, producto de esta comparación, se genera un análisis concluyente respecto a la aplicación de edición no versionada con ambos software.

Palabras clave: *edición multiusuario, edición no versionada, edición versionada, bases de datos espaciales, geodatabases, manejo de datos geoespaciales, SIG.*

Abstract - This paper introduces spatial database editing, making a differentiation between versioned and not versioned editing. Specifically, two simple non-versioned multiuser editing workflows are developed using PostgreSQL in combination with QGIS, and on the other hand utilizing ArcGIS, so that the reader can have a comparative view of the use of both methods. Finally, as a result of this comparison, a conclusive analysis is generated regarding the application of non-versioned editing with both software.

Keywords - *multi-user edition, non-versioned edition, versioned edition, spatial databases, geodatabases, geospatial data management, GIS.*

1. INTRODUCCIÓN

Los SIG (sistemas de información geográfica) son sistemas integradores de tecnología informática, personas e información geográfica para capturar, analizar, almacenar y representar datos georreferenciados [1].

Al ser el almacenamiento parte fundamental de los SIG, las bases de datos juegan un papel muy importante en estos sistemas, de ahí que la mayoría de los actuales SGBD (Sistemas Gestores de Bases de Datos) permitan la incorporación de información geográfica. De esta manera se originan las **bases de datos espaciales**, para almacenar entidades del mundo real que ocupan un lugar en el espacio y mantienen relaciones espaciales entre ellas y el resto de elementos del entorno [2].

Son las bases de datos espaciales las que permiten la edición de estos elementos en entornos de trabajo multiusuario. La edición multiusuario se puede llevar a cabo de dos formas, la **edición no versionada**, que realiza ediciones y actualizaciones dinámicas de las capas sin la utilización de diferentes

versiones de las tablas. Esta metodología se puede llevar a cabo mediante el uso de vistas, que son tablas virtuales que almacenan consultas como objetos [3], pero que en realidad sólo almacenan su definición y no los datos propiamente.

La edición versionada, por su parte, trabaja en diferentes versiones, aislando el trabajo en diferentes sesiones de edición que permiten la edición concurrente. Este método conserva una versión principal original o versión predeterminada [4].

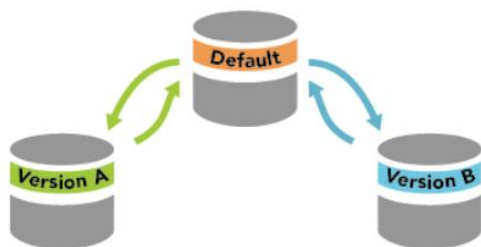


Fig. 1. Edición multiusuario versionada.

Fuente: [4]

La edición versionada es la forma más robusta de trabajo multiusuario. Requiere de una conexión con una base de datos relacional como IBM DB2, Informix, Oracle, PostgreSQL o Microsoft SQL Server, así como un administrador de bases de datos involucrado en el proceso, por lo que es el método más complicado de establecer [5].

En el desarrollo de este artículo se amplían los conceptos relacionados con la edición multiusuario en bases de datos espaciales y se realiza una comparación entre diferentes flujos de trabajo en entornos multiusuario no versionados: el trabajo con PostgreSQL junto con QGIS, y varias alternativas para el uso de conjuntos de datos distribuidos entre múltiples *geodatabases* de archivos en ArcGIS, *software* de ESRI, la empresa líder mundial en sistemas de información geográfica.

Esta comparación se llevó a cabo a través de la realización de un ejemplo práctico del flujo de trabajo entre PostgreSQL y QGIS. Mientras que, en el caso del flujo de trabajo en ArcGIS, el ejemplo mostrado se obtuvo a partir de investigación bibliográfica.

Este trabajo pretende, mediante esta comparación, distinguir los flujos de trabajo en edición no versionada, además de proveer a los lectores, a través de ejemplos prácticos, un paso a paso en cada ambiente, así como identificar las ventajas de cada uno de ellos.

2. DESARROLLO

Para realizar una buena gestión de datos espaciales es importante tener claro cuáles flujos de trabajo son útiles para los fines de almacenamiento y gestión de información geoespacial. Debido a esto, se realizó una comparativa de los flujos de trabajo en entornos multiusuario no versionados, utilizando PostgreSQL en conjunto con QGIS, y ArcGIS de ESRI.

Sin embargo, previo a la presentación de esta comparación, es pertinente ampliar sobre las capacidades de ArcGIS en la edición multiusuario versionada.

Lo primero que debe conocerse es el concepto de *geodatabase*, que representa un modelado de datos espaciales orientado a objetos, que permite la definición personalizada de los elementos mediante relaciones espaciales y topológicas. Las *geodatabases* almacenan los datos espaciales en una única base de datos relacional [6].

Existen tres tipos de *geodatabase* con diferentes características operativas en

ArcGIS. Dos de ellas permiten edición multiusuario no versionada (geodatabases personales y geodatabases de archivos) y la tercera permite la edición versionada (geodatabases corporativas).

Las **geodatabases personales** almacenan los conjuntos de datos dentro de un archivo de datos de Microsoft Access con un límite de tamaño de 2 GB. Las **geodatabases de archivos** son almacenados como carpetas y cada conjunto de datos se aloja como un archivo de hasta 1 TB. Las **geodatabases corporativas** pueden no tener límite de tamaño y cantidad de usuarios y se almacenan en bases de datos relacionales como Oracle, Microsoft SQL Server, IBM DB2, IBM Informix o PostgreSQL [6].

Como estas últimas permiten edición versionada se conocen también como **geodatabases multiusuario**. Esto es posible a través del uso de un producto de ESRI llamado **ArcSDE** (Spatial Database Engine), que funciona como intermediario o servidor de aplicaciones entre el SGBD y el cliente (arquitectura de tres niveles). ArcSDE también permite que las aplicaciones cliente se conecten directamente a servidores de datos con información espacial (arquitectura de dos niveles) [7].

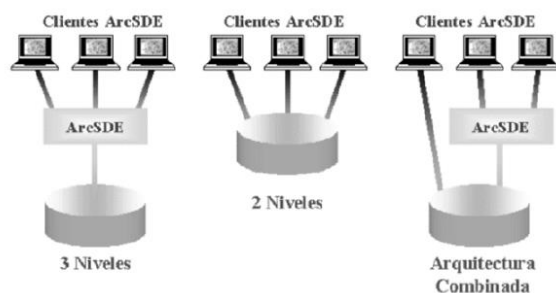


Fig. 2. Arquitectura de ArcSDE. Fuente: [7].

El tipo de edición versionada obtenido mediante ArcSDE es conocido como *versionado tradicional*. Otro tipo de

versionado es el que facilita el modelo de SIG web a través de servicios de entidades. Este es conocido como *versionado en rama* [8].

Para el versionado en rama, ESRI utiliza otro de sus productos, **ArcGIS Enterprise**, que consiste en un modelo SIG web con una arquitectura basada en servicios para permitir flujos de trabajo de edición multiusuario a través de capas de entidades web. Estas capas son conocidas también como servicios de entidades, y se comparten para visualizar, consultar y editar datos en la web [9].

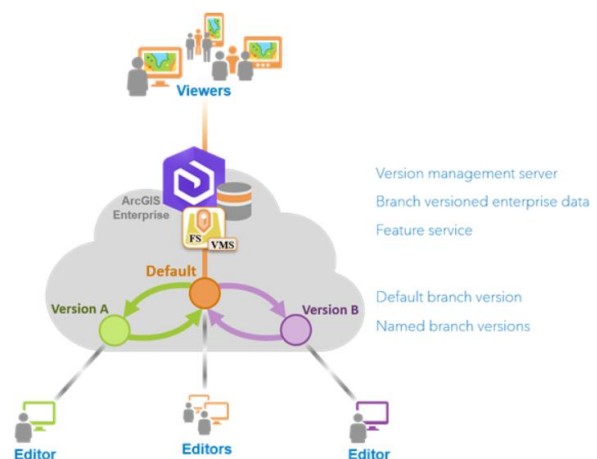


Fig. 3. Flujo de trabajo de versionado en rama (ArcGIS Enterprise). Fuente: [9].

Las ediciones versionadas mediante las geodatabases corporativas crean una estructura jerárquica de edición para los usuarios, en la que se conceden privilegios específicos de edición a través de vistas a la base de datos, permitiendo ambientes de trabajo aislados del total de los datos que la contienen, así como un control de la edición realizada previo a una publicación, identificación de conflictos y control de versiones [5].

Sin embargo, en este artículo no se detalla esta metodología de trabajo, debido a su complejidad, ya que se

considera como un tema avanzado que involucra la organización de personal de distintas áreas para su implementación [5].

Por lo tanto, a continuación se retoma la comparación de ediciones multiusuario no versionadas, entre el trabajo con PostgreSQL junto con QGIS, y varias alternativas para el uso de conjuntos de datos distribuidos en *geodatabases* de archivos de ArcGIS.

A. Flujo de trabajo en PostgreSQL y QGIS

En términos generales, este flujo de trabajo requiere de la instalación de PostGIS para agregar el elemento geoespacial a PostgreSQL, en conjunto con QGIS, que permite la edición de manera directa de los elementos, sin intermediarios, como se muestra en la siguiente figura:

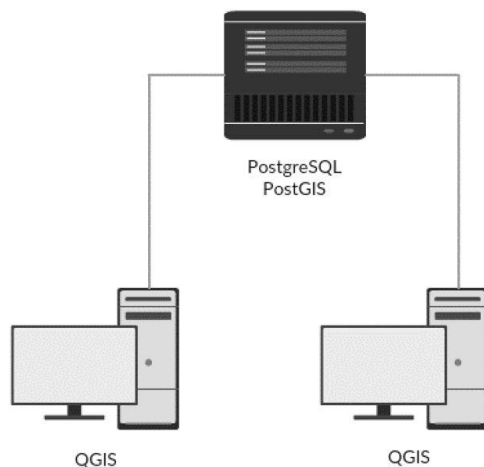


Fig. 4. Arquitectura cliente-servidor entre QGIS y PostgreSQL. Fuente: [10].

Utilizando el software pgAdmin de PostgreSQL, se crea una base de datos y se cargan los datos. Para este caso de estudio se cargó un respaldo de una base de datos llamada "Atlas", que contiene capas de las provincias, cantones y distritos de Costa Rica.

Mediante el uso de "Query Tool" es posible realizar consultas dirigidas al trabajo con múltiples usuarios, las cuales se pueden aplicar a la creación de vistas de edición, que consisten en selecciones de datos de interés mediante consultas, para ser mostrados a usuarios específicos. A la tabla consultada se le llama *tabla base*, y a la vista generada se le llama *tabla virtual*. Las vistas permiten ocultar información, mejorar el rendimiento y simplificar la administración de los permisos de los usuarios [3]. La figura 5 muestra un ejemplo de la creación de una vista para los cantones de tres provincias.

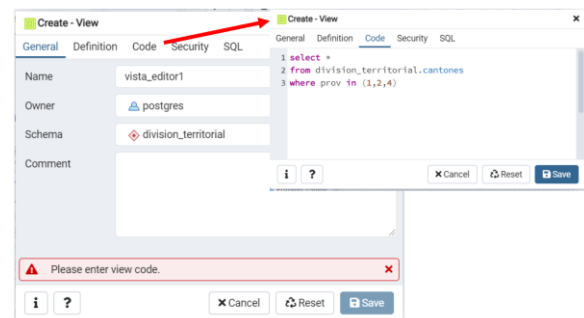


Fig. 5. Creación de vista de edición.

Estas vistas se almacenan en la base de datos y pueden ser consultadas o modificadas en cualquier momento, convirtiéndose en una alternativa para desplegar datos de la misma tabla a varios usuarios distintos o de varias tablas a un solo usuario.

En un programa para visualizar información geográfica como QGIS, es posible visualizar las vistas de edición realizadas, a través de una conexión a la base de datos que las contenga. La figura 6 muestra las vistas creadas en PostgreSQL y la visualización de estas desde la conexión a la base de datos en QGIS.

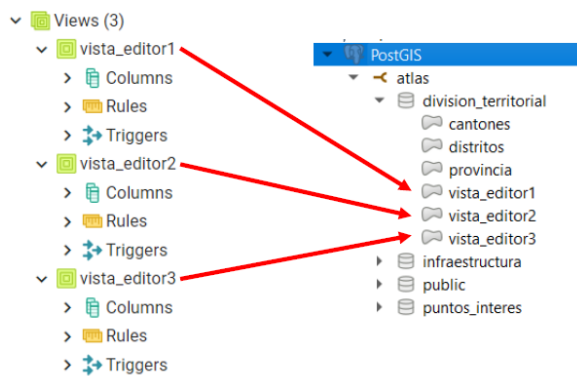


Fig. 6. Vistas en PostgreSQL y en QGIS.

Una vez que se tiene la información cargada en QGIS, se puede realizar ediciones según cada vista que se le asigne a los usuarios que tengan acceso a ella.



Fig. 7. Edición de vistas en la base de datos "atlas", en QGIS.

Esta metodología separa las áreas de trabajo, las aísla y no muestra la edición durante la ejecución.

B. Flujos de trabajo en ArcGIS

ArcGIS utiliza la figura de geodatabase para almacenar sets de datos de diversos tipos, administrándolos como si estuviesen en una carpeta o en una base de datos relacional, utilizándose para editar y automatizar el acceso a ellos [11].

Una geodatabase utilizada en ArcGIS es ideal para organizar e integrar datos de múltiples usuarios de manera no versionada. Sin embargo, una de las principales limitantes del trabajo desde una geodatabase en donde se tienen entidades utilizadas por varios usuarios es que no pueden ser editadas al mismo tiempo [5].

A continuación, se presentan cuatro posibles esquemas para organizar los esfuerzos de edición colaborativa no versionada. Es necesario aclarar que ninguno de ellos representa un flujo de trabajo que sea directo y sin necesidad de ediciones posteriores:

1. Acceso al directorio común:

Se debe verificar si todos los editores tienen acceso al mismo espacio de trabajo de forma fiable. Luego se debe crear una geodatabase para cada editor con esquemas coincidentes, y un conjunto común de mapas base, que permita ediciones individuales, mientras el supervisor mantiene la visibilidad durante todo el proceso. Adicionalmente, cada editor puede visualizar lo que otros editan mediante el otorgamiento de permisos de lectura únicamente. Trabajar de esta manera puede evitar el uso indebido o corrupción de archivos.

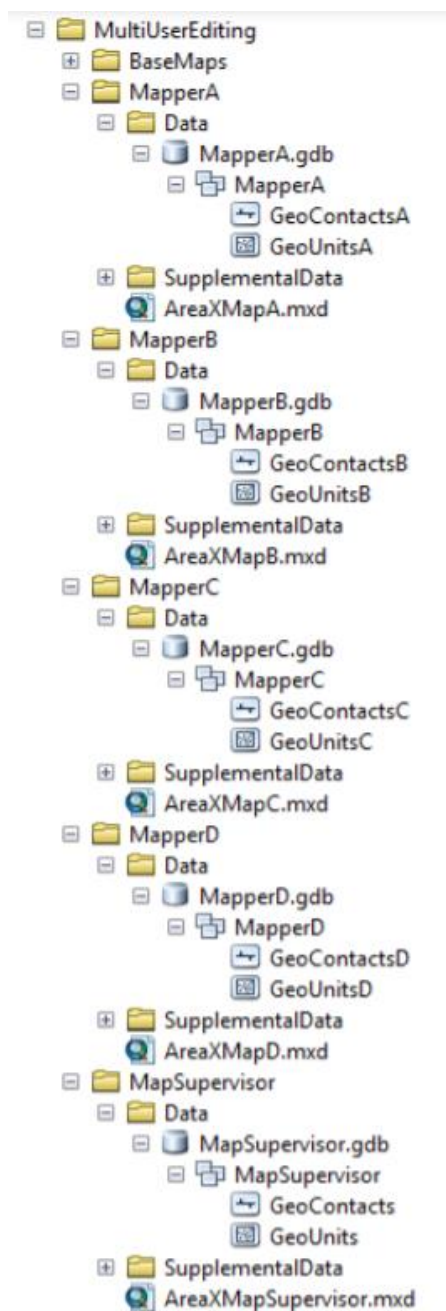


Fig. 8. Ejemplo de organización de acceso al directorio común. Fuente: [5].

2. Separación de elementos:

Esta modalidad es útil para cuando las áreas de trabajo son pequeñas o si el trabajo de edición a realizar se puede separar por tipo de morfologías, tipo de geometrías, u otros, siempre en la misma

área de trabajo. En tal caso se aplica el método de acceso al directorio común.

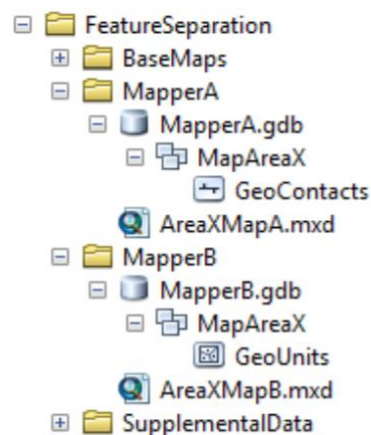


Fig. 9. Ejemplo de organización de un espacio de trabajo mediante separación de elementos. Fuente: [5].

3. Separación geográfica:

Este tipo de edición es funcional para áreas grandes, en las cuales se requiere que se divida geográficamente. Representa mayor rapidez al cubrir áreas de edición, sin embargo, es importante prestar atención a las zonas de límites entre editores.

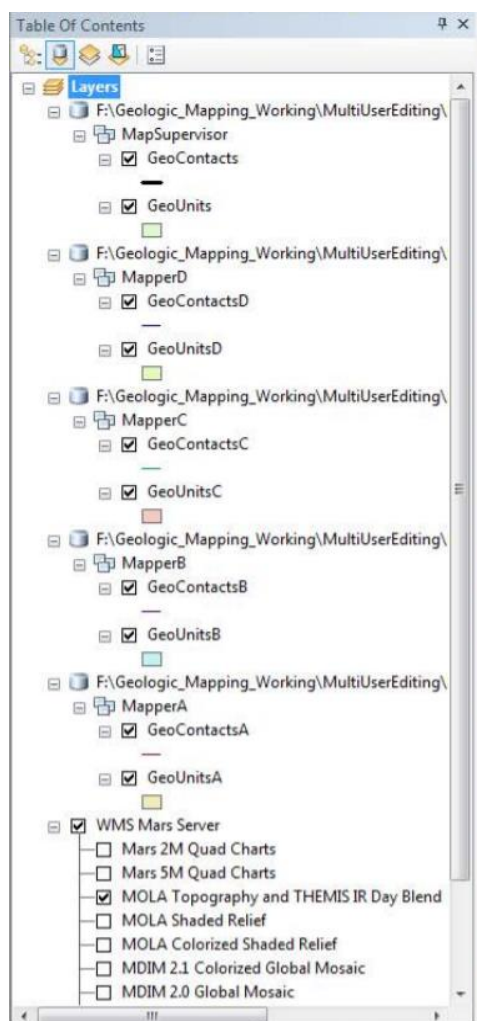


Fig. 10. Ejemplo de organización de un espacio de trabajo mediante separación geográfica.
Fuente: [5].

Al finalizar la edición, los datos pueden ser fusionados mediante la herramienta Fusionar.

4. Edición coordinada en espacios de trabajo aislados

Esto implica espacios separados de trabajo y una bitácora de control que indique información de cada edición. Este método es menos receptivo y propenso a problemas de edición individuales.

Los cuatro métodos anteriores suponen la utilización de una geodatabase. Para evitar problemas de bloqueos de esquemas y conflictos de edición, es

posible habilitar el seguimiento dando clic derecho en la entidad seguido de la selección "Manage", y posteriormente "Enable Editor Tracking". El resultado es un seguimiento de edición como se muestra en la siguiente figura.

OBJECTID	SHAPE	SHAPE_Len	created_user	created_date	last_edited_user	last_edited_date
5	Polyline	4400000	MAHUNTER	7/1/2016 4:57:18 P	MAHUNTER	7/1/2016 4:57:18 PM
6	Polyline	3300000	MAHUNTER	7/1/2016 4:57:18 P	MAHUNTER	7/1/2016 4:57:18 PM
7	Polyline	3300000	MAHUNTER	7/1/2016 4:57:18 P	MAHUNTER	7/1/2016 4:57:18 PM
8	Polyline	2200000	MAHUNTER	7/1/2016 4:57:18 P	MAHUNTER	7/1/2016 4:57:18 PM

Fig. 11. Seguimiento de edición. Fuente: [5].

3. CONCLUSIONES

- La edición multiusuario no versionada utilizando PostgreSQL como SGBD y QGIS como medio de edición y visualización de datos es más eficiente y sencilla que las distintas posibilidades de edición no versionada que permite ArcGIS mediante el uso de geodatabases de archivos, ya que, a diferencia de esta última, el flujo en PostgreSQL es directo y no requiere de ediciones posteriores.
- ArcGIS incluye opciones de edición versionado que requieren del uso de productos adicionales como ArcSDE o ArcGIS Enterprise, lo que implica un costo mayor del *software*, la necesidad de personal especializado y de un mayor nivel de organización para su implementación, en comparación con el flujo de trabajo mostrado en PostgreSQL.
- La alternativa de PostgreSQL implica independencia de funciones de almacenamiento en la base de datos y de edición y visualización de los datos espaciales en el software SIG utilizado, en este caso QGIS. Mientras tanto, en el caso de ArcGIS, al requerir un

intermediario entre el SGBD y el software de edición, no ofrece independencia de funciones.

- Las alternativas de código abierto como PostgreSQL y QGIS permiten la edición multiusuario de datos espaciales de manera satisfactoria y eficiente, por lo que el pago de licencias para acceder a *software* especializado que ofrezca estas posibilidades no es indispensable.
- Como recomendación para una futura investigación en esta misma línea, se puede llevar a cabo el flujo de trabajo de edición versionada mediante ArcSDE, así como estudiar en detalle la manera en que ArcGIS Online maneja las bases de datos y cómo se ha empezado a relacionar con el manejo BIM de proyectos que albergan información geoespacial, e investigar si existen alternativas en *software* gratuito que ofrezcan estas mismas posibilidades.

4. REFERENCIAS

[1] V. Olaya. "Sistemas de Información Geográfica", 2020.

[2] I. Del Bosque, C. Fernández, L. Martín-Forero, "Los sistemas de información geográfica y la investigación en ciencias humanas y sociales", 2012.

[3] D. Moisset, "PostgreSQL. Tutoriales Ya", Available: <https://www.tutorialesprogramacionya.com/postgresql/temarios/descripcion.php?inicio=50&cod=226&punto=68>, 2021.

[4] Descripción general del versionado. Available: <https://pro.arcgis.com/es/pro-app/latest/help/data/geodatabases/overview/overview-of-versioning-in-arcgis-pro.htm>, 2021.

[5] M. Hunter, "Best Practices for Multi-User Editing with ArcMap 10.3." MRCTR GIS Lab, USGS Astrogeology Science Center, Flagstaff, Arizona, Estados Unidos, 2017.

[6] D. Rojas, "Implementación del Sistema de Información Geoespacial del Instituto Mixto de Ayuda Social (SIG IMAS) para la administración de la información institucional de carácter espacial asociada al Área de Titulación en la Provincia de Limón", Escuela de Ingeniería Topográfica, Universidad de Costa Rica, 2018.

[7] J. Rodríguez, M. Asmal, V. Ramírez, "Análisis, diseño e implementación de un sistema de información geográfica para seguimiento de egresados de la Universidad Politécnica Salesiana", Carrera de Ingeniería de Sistemas, Universidad Politécnica Salesiana, Sede Cuenca, 2011.

[8] Tipos de versionado. Available: https://pro.arcgis.com/es/pro-app/latest/help/data/geodatabases/overview/versioning-types.htm#ESRI_SECTION1_6FA2CFB5F9484FF096740D653C674B5D, 2021

[9] Escenarios de versiones en rama. Available: <https://pro.arcgis.com/es/pro-app/latest/help/data/geodatabases/overview/branch-version-scenarios.htm>, 2021

[10] Z. Magyari-Sáska, "Developing and implementing multiuse, fully relational GIS database for desktop systems using open source technologies", Facultad de Geografía, Universidad Babeş-Bolyai, Rumania, 2015.

[11] F. Leonardo, "Mini curso de Geodatabase en ArcGIS. Ecuador". Available: <https://acolita.com/mini->

curso-manual-de-geodatabase-en-arcgis/, 2012.