

## Methodology of association of cadastral and electrical information using GIS and SQL tools: Case study in Quito Ecuador

## Metodología de asociación de información catastral y eléctrica mediante herramientas SIG y SQL: Caso de estudio Quito, Ecuador

 L. Jami<sup>1</sup>  C. Vallejo<sup>2</sup>  L. Godoy<sup>1</sup>  F. Vásquez<sup>1,2</sup>  J. L. Condo<sup>1</sup>

<sup>1</sup>Empresa Eléctrica Quito, Quito, Ecuador

E-mail: [liseth\\_dic@hotmail.com](mailto:liseth_dic@hotmail.com); [felipe\\_luis\\_6\\_d@hotmail.com](mailto:felipe_luis_6_d@hotmail.com); [hola\\_josecondo@yahoo.es](mailto:hola_josecondo@yahoo.es)

<sup>2</sup> Instituto de Investigación Geológico y Energético, Quito, Ecuador

E-mail: [catalina.vallejo@geoenergia.gob.ec](mailto:catalina.vallejo@geoenergia.gob.ec); [francis.vasquez@geoenergia.gob.ec](mailto:francis.vasquez@geoenergia.gob.ec)

### Abstract

This study develops a methodology for the unification of electricity and cadastral data through the combined application of GIS and SQL tools. The main spatial joining techniques used are: spatial intersection, entity joining by attributes, and proximity analysis. In addition, mainly SQL algorithms for data comparison and selection were applied. As a case study, we used cadastral and electricity information of the customers regulated by Empresa Eléctrica Quito (EEQ) within the urban area of Quito. It was possible to join 25% of the total data. Furthermore, it was found that the spatial distribution of the unified data is homogeneous within the study area, and is quantitatively representative in most of the areas. Finally, the results were checked by means of an on-site validation, thus verifying the validity of the proposed methodology. According to the results obtained, the joining methodology presented in this study can be replicated in other areas with similar characteristics to the data used, thus obtaining an important input for the definition of energy indicators for energy qualification, comparison, or categorization of buildings.

**Index terms**— Electricity consumption, residential buildings, cadaster, database merge, GIS, SQL.

### Resumen

En el presente estudio se desarrolla una metodología para la unificación de datos eléctricos y catastrales mediante la aplicación combinada de herramientas SIG y SQL. Las principales técnicas de unión espacial utilizadas son: intersección espacial, unión de entidades por atributos y análisis de proximidad. Además, fueron aplicados principalmente algoritmos SQL de comparación y selección de datos. Como caso de estudio se utilizó información catastral y eléctrica de los clientes regulados por la Empresa Eléctrica Quito (EEQ) dentro del casco urbano de Quito. Se logró unificar el 25 % del total de registros. Además, se encontró que la distribución espacial de los datos unificados es homogénea dentro del área de estudio, y es cuantitativamente representativo en la mayor parte de parroquias. Finalmente se comprobaron los resultados mediante una validación en sitio, con lo que se verificó la validez de la metodología planteada. En base a los resultados obtenidos, el procedimiento de unificación expuesto en este estudio, puede ser replicado en otras zonas cuya información presente características similares a los datos manejados y así obtener un insumo importante para la definición de indicadores energéticos para calificación energética, comparativa o categorización de edificaciones.

**Palabras clave**— Consumo eléctrico, edificios residenciales, catastro, unión de bases de datos, SIG, SQL.

Recibido: 30-06-2022, Aprobado tras revisión: 15-07-2022

Forma sugerida de citación: Aymacaña, L.; Vallejo, C.; Godoy, L.; Vásquez, F.; Condo, L. (2022). "Metodología de asociación de información catastral y eléctrica mediante herramientas SIG y SQL: Caso de estudio Quito, Ecuador". Revista Técnica "energía". No. 19, Issue I, Pp. 158-168

ISSN On-line: 2602-8492 - ISSN Impreso: 1390-5074

Doi: <https://doi.org/10.37116/revistaenergia.v19.n1.2022.532>

© 2022 Operador Nacional de Electricidad, CENACE



## 1. INTRODUCCIÓN

Según datos de la Agencia Internacional de Energía (IEA), entre 1976 y 2019 el consumo de energía se incrementó en un 92 % y la electricidad representó un tercio del uso de energía de los edificios en 2019 [1]. En Ecuador esta realidad no es diferente. En 2020, las edificaciones residenciales consumieron el 40,12 % de la electricidad facturada a nivel nacional, siendo el sector mayor consumo eléctrico [2]. Esto evidencia que en Ecuador el parque edificado es en un sector clave para reducir el consumo de electricidad.

El primer paso para la disminución del consumo eléctrico es conocer el rendimiento energético de los edificios y sus parámetros influyentes. Varios estudios [3][4][5][6][7] indican que el consumo eléctrico tiene una relación directa con las características constructivas de las edificaciones. Es por esta razón que, a escala urbana, la recolección, integración y transformación de datos representa un gran desafío.

Ma. J et al. [8], realizaron la integración de información, mediante SIG y big data para estimar el consumo eléctrico de edificios a escala urbana. Esto se facilitó considerando a la ubicación como variable clave en los SIG, ventaja que no existe en otro tipo de interface, y su uso para la minería de datos generó grandes beneficios como la conexión exitosa y confiable bases de datos georreferenciadas.

Mutani et al. [9], asociaron datos de energía a la información de las edificaciones en Turín, mediante una herramienta SIG. Al conocer las ubicaciones de los datos de ambas temáticas se intersecaron espacialmente y con ello la información quedó unificada, viabilizando el análisis del rendimiento energético

Ali. U et al. [10], desarrollaron una metodología, basada en datos, para optimizar la toma de decisiones, a escala urbana, para el reacondicionamiento energético de edificios residenciales en Irlanda. Los resultados mostraron que los modelos para el reacondicionamiento, basado en datos reales, disminuyen la cantidad de información requerida sin afectar la precisión de las predicciones respecto a la clasificación energética de los edificios.

Dall'O' et al. [11], desarrollaron una herramienta para crear una base de datos georreferenciada con información referente al rendimiento energético de edificios residenciales en Lombardi-Italia. Toda la información recopilada fue utilizada para alimentar la base de datos en la plataforma SIG obteniéndose de esta manera una herramienta para el análisis del sector edificado de una localidad.

En Ecuador, varios autores [12] [13] [3] han realizado estudios para determinar el rendimiento de viviendas en diferentes ciudades. Los trabajos se basan en la evaluación individual de seis o menos edificaciones. Los resultados obtenidos sirven como

información de referencia para tipologías y localidades similares; sin embargo, no es factible su escalamiento a toda una municipalidad o ciudad. Para esto se requiere el análisis de información generada por diferentes entidades.

Actualmente, gran parte de las entidades públicas en Ecuador generan información georreferenciada. Sin embargo, no se encuentran interrelacionadas y la precisión y sistema de referencia diverge entre sí. Es por esta razón que, no es posible fusionar directamente las bases de datos catastrales y del servicio eléctrico a través de una clave principal o mediante ubicación.

Por lo expuesto, el objetivo de este estudio es desarrollar una metodología que permita la unificación de bases de datos con información catastral y eléctrica mediante el uso de herramientas SIG y SQL. Para la aplicación de la metodología, se ha seleccionado una muestra representativa de edificaciones residenciales localizadas en el casco urbano del Distrito Metropolitano de Quito (DMQ). Esto debido a que, en 2020, la Empresa Eléctrica Quito (EEQ) fue la segunda empresa eléctrica con mayor consumo registrado, por parte del sector residencial, frente a otras distribuidoras del país y, fue la primera en la región sierra [2].

La muestra inicial estuvo conformada por 633 800 medidores de electricidad, y mediante la metodología desarrollada se unificó aproximadamente el 25 % de los registros con la información catastral correspondiente. La nueva base de datos está conformada de 157 334 registros de predios y medidores asociados. Además, se identificó que el porcentaje de asociaciones obtenidas se encuentra homogéneamente distribuido dentro de la zona de estudio y es cuantitativamente representativo en la mayor parte de parroquias. La metodología planteada consta de cuatro (4) fases.

En el presente trabajo se describen de manera preliminar al área de estudio y las fuentes de información; así como, el estado de los datos. Luego, el proceso de depuración de las bases de datos, un análisis de las relaciones de unión existentes, así como, la unificación y validación. Posterior, se muestran los resultados, donde se indican las relaciones de unión seleccionadas y el número de datos unificados de cada fase. Además, se muestra la distribución espacial de los datos unificados y su porcentaje de representatividad por parroquia. Finalmente, se exponen las conclusiones y recomendaciones después de la aplicación de la metodología planteada.

## 2. METODOLOGÍA

Esta sección describe la información concerniente al área de estudio y a los datos de insumo, sobre los cuales se realiza una depuración. Además, se analizan las posibles relaciones de unión predio – medidor a presentarse y en base a ello, se establecen las técnicas y herramientas adecuadas para la unificación. Finalmente,



se efectúa una validación en sitio de los resultados. En la Fig. 1 se indican los pasos a seguir desde la recopilación de datos hasta la validación de resultados.



Figura 1: Metodología de estudio

## 2.1. Área de estudio y fuentes de información

La metodología desarrollada en este estudio es aplicada en la zona correspondiente al casco urbano del DMQ, la cual incluye 32 parroquias urbanas y cuenta con una superficie de 18 860 has [14]. El área de estudio se muestra en la Fig. 2.

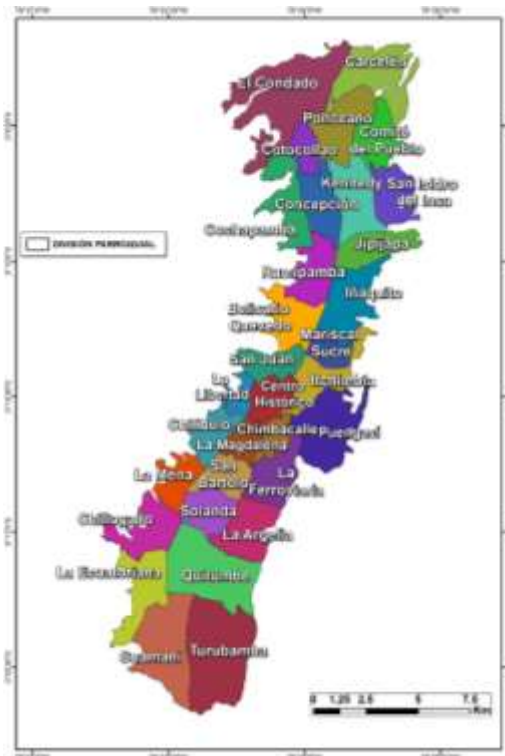


Figura 2: Área de estudio

La información utilizada procede de los organismos generadores competentes. La EEQ dispone de información eléctrica de los clientes regulados en el área de estudio. Por otra parte, la Secretaría de Territorio, Hábitat y Vivienda del Gobierno Autónomo Descentralizado del Distrito Metropolitano de Quito (GAD-DMQ) que dispone de información predial. Debido a que se trata de información sensible de la población, las bases de datos no son de acceso público, sin embargo, estas son administradas por las instituciones públicas antes citadas.

Los datos recopilados corresponden a dos geodatabases separadas, con archivos georreferenciados de tipo shapefile en formato vectorial. Se conoce que el catastro predial se encuentra a escala 1 : 1 000 y se rige a las normas de exactitud posicional establecidas por el

Instituto Geográfico Militar (IGM), por lo cual su precisión es no superior a  $0,3 \text{ m} \pm 0,03 \text{ m}$  [15][16]. Respecto a la cartografía de consumo eléctrico, se encuentra generada a escala de detalle, y los equipos utilizados para el levantamiento de información tienen una precisión superior respecto a la cartografía catastral.

El consumo eléctrico se encuentra representado como una cobertura tipo punto, cuyo sistema de referencia corresponde a WGS84 proyección UTM Zona 17 Sur. Posee 633 800 registros y 15 campos con información descriptiva. Por otro lado, el catastro predial se encuentra representado como una cobertura tipo polígono y su sistema de referencia corresponde a WGS84 proyección TMQ Zona 17 Sur. Posee 660 909 registros y 17 campos con datos descriptivos de cada predio.

Debido a que: a) la precisión y sistema de referencia de cada una de las coberturas diverge entre sí, b) existen registros duplicados y c) no existen campos en común (primary key), no es posible establecer una unificación directa.

## 2.2. Depuración de bases de datos

Antes de realizar cualquier tipo de operación entre capas, se considera pertinente tratar las diferencias sistemáticas existentes [17]. Por ello, la depuración de las bases de datos inicia con la homogenización de sus características espaciales. Los sistemas de referencia de la cobertura catastral y eléctrica son distintos, por lo que se realizó una transformación para que ambas adopten el sistema de referencia WGS84 Zona 17 Sur.

Posteriormente, se realizó una limpieza de los datos contenidos en cada registro de acuerdo a las características propias de cada campo. Para esto se analizó el dominio de cada variable. En los campos cuyo atributo corresponde exclusivamente a información tipo texto, se eliminaron los caracteres numéricos, y viceversa (sin incluir campos que poseen caracteres combinados). Además, se transformaron los caracteres de los registros de datos cualitativos de minúsculas a mayúsculas y se identificaron caracteres especiales para sustituirlos o eliminarlos.

Como resultado, los datos contenidos en la tabla de atributos de cada una de las coberturas presentan características espaciales que facilitan procesos geoespaciales y, además, la calidad de datos presenta una mejora que posibilita y favorece la interacción entre campos.

En esta etapa se identificó la existencia de objetos duplicados tanto de la base de datos catastral como eléctrica. Esto se refleja en la presencia de objetos sobrepuestos en una misma ubicación geográfica, lo cual, en el caso del catastro, indica la existencia de viviendas multifamiliares. Por ello, para definir las herramientas de unificación espaciales o por atributos, se requiere un análisis de las posibles relaciones de

unión predio – medidor (cobertura predial-cobertura eléctrica), que se encuentran como polígono y punto respectivamente y se indican en la siguiente sección.

**2.3. Identificación y selección de relaciones de unión**

Para identificar las posibles relaciones de unión entre predio (polígono) y medidor (punto) se realizó una intersección espacial entre ambas capas. Posteriormente se analizó el estado de correspondencia espacial y de atributos y se identificaron las siguientes relaciones de unión: “uno a uno”, “uno a varios”, “varios a uno” y “varios a varios”. En la Fig. 3 se indica gráficamente cada una de las relaciones.

En el caso de unión “uno a uno”, como se ve en la Fig. 3a se presenta un solo predio donde se encuentra ubicado un solo medidor. La fusión en este caso es sencilla, ya que se asigna la información del polígono al punto. La unificación es verificada a través de la coincidencia de atributos alfanuméricos de un campo en común entre ambas bases de datos, de no darse esta coincidencia, el registro fusionado queda descartado.

Bajo la relación de unión “uno a varios”, como en la Fig. 3b, existe un solo predio en el cual se encuentran ubicados varios medidores que registran distintos consumos. En este caso, el punto recibe la información del polígono y la unificación queda verificada si el atributo en común de ambas bases de datos coincide, de lo contrario el registro unificado es descartado.

La relación de unión “varios a uno” se indica en la Fig. 3c y Fig. 3d, se observa que para varios predios (que puede ser el caso de edificios de departamentos u otro tipo de vivienda multifamiliar), existe un solo medidor ubicado dentro de ellos. En este caso, el predio puede corresponder a un mismo propietario (Fig. 3c) o de distintos propietarios (Fig. 3d). Al ser del mismo propietario, la asignación de información del polígono al punto no es viable ya que para su verificación se requiere de visitas en sitio. Si los predios son de distintos propietarios, la asignación de información de polígono a punto es válida siempre y cuando exista la coincidencia de un atributo en común.

En el caso de unión “varios a varios”, como se observa en la Fig. 3e, Fig. 3f y Fig. 3g, existen varios predios y medidores en una misma ubicación geográfica. Respecto a predios, puede darse que uno (Fig. 3g) o varios (Fig. 3e y Fig. 3) corresponda a un mismo propietario. En relación a medidores, todos registran diferentes consumos, sin embargo, uno (Fig. 3g) o varios (Fig. 3e y Fig. 3f) corresponden a un mismo propietario. Si varios predios y medidores presentan un mismo propietario (Fig. 3e), no es posible establecer una fusión válida, para ello se requiere de visitas en sitio. Para aceptar una fusión, se debe presentar lo indicado en la Fig. 3f, es decir que uno o varios puntos debe pertenecer a un solo polígono, de lo contrario se requiere de visitas en sitio. El caso ideal, es indicado en la Fig. 3g, donde la información de punto y

polígono es única, por lo tanto, la asociación entre objetos es exitosa.

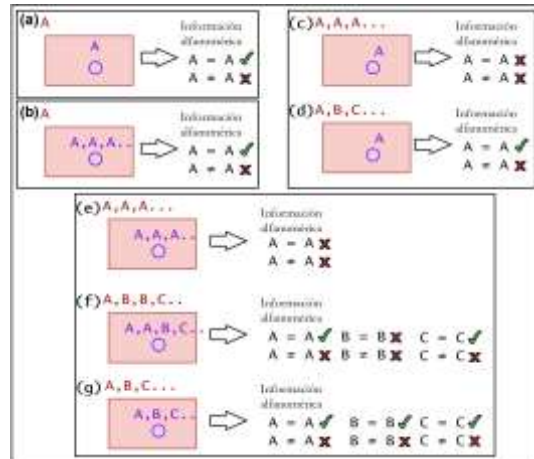


Figura 3: Relaciones de unión punto – polígono basado en su coincidencia espacial

Las relaciones antes mencionadas se basan en la coincidencia espacial entre punto y polígono. Sin embargo, es posible que exista correspondencia entre atributos alfanuméricos en común entre punto y polígono sin existir una correspondencia espacial entre ellos, obteniéndose una unión válida. En este caso, en base a los atributos en común de las capas, la única relación de unión punto - polígono posible encontrada, es de uno a uno.

La unión de los datos bajo las relaciones seleccionadas se realiza mediante métodos espaciales y de comparación de atributos en común, lo cual se indica en la siguiente sección.

**2.4. Unión de bases de datos**

Las relaciones de unión espaciales y alfanuméricas seleccionadas en la sección 2.3, se agrupan en tres casos: 1) un predio y uno o varios medidores, 2) varios predios (cuyo campo en común sea único) y uno o varios medidores, y 3) predio y medidor cuyo atributo alfanumérico de los campos en común sean coincidentes. La unión los casos mencionados se llevaron a cabo en cuatro fases esquematizadas en la Fig. 5 detalladas a continuación.

**2.4.1 Fase 1**

Para el caso donde existe un solo predio que espacialmente contiene a uno o varios medidores, inicialmente, se dividió la cobertura catastral en dos capas, una con predios duplicados y otra con predios únicos. Con esta última, se realizó una unión de atributos de polígono a punto basada en su correspondencia espacial. Se obtuvo una cobertura de salida, la cual es de tipo punto y contiene atributos de predios y medidores.

Posteriormente se realizó el proceso de verificación de objetos (PVO) unificados mediante la comparación



de información alfanumérica de los atributos en común de ambas coberturas. Aquellos registros cuya coincidencia fue idéntica se conservaron. Debido a que esta comparación se limita a la aceptación de coincidencias idénticas, se realizó también la programación de comandos SQL para la comparación de coincidencias entre caracteres de las entidades unificadas. En el presente estudio se aceptaron aquellos registros cuyos campos en común presenten una coincidencia idéntica y con cuatro (4) o más caracteres. Los objetos que cumplieron con estos dos tipos de comparaciones fueron aceptados, los restantes fueron descartados.

#### 2.4.2 Fase 2

En este caso, donde existen varios predios que espacialmente contienen a uno o varios medidores, interviene la capa del catastro que contiene registros duplicados.

Inicialmente se dividió la cobertura predial para obtener una capa con predios cuyo campo en común, con la cobertura de medidores, sea único. Con esta capa y con la cobertura de medidores que fueron descartados en la fase 1, se realizó una unión de atributos de polígono a punto basada en su coincidencia espacial. Se obtuvo una capa de salida tipo punto que contiene información de predio y medidor asociado. Luego, de la misma forma que en la fase 1, se realizó el PVO

#### 2.4.3 Fase 3 y Fase 4

La unificación en este escenario se lleva a cabo de un predio a un medidor donde exista una correspondencia de información alfanumérica de un campo en común entre ambas bases de datos.

Primero (fase 3), se realizó una unión por atributos con coincidencia idéntica mediante comandos de programación SQL tomando como clave principal o primary key a un campo en común entre la base de datos predial y de consumo eléctrico. Posteriormente, (fase 4) se realiza la comparación de coincidencias entre caracteres de las entidades unificadas, las cuales no presentaron coincidencia espacial.

Finalmente se llevó a cabo la verificación de las entidades unificadas mediante un análisis de proximidad entre los objetos asociados con el fin de descartar aquellas unificaciones cuya distancia de separación no sea tolerable.

En el presente estudio se ha establecido una tolerancia de 45 m. La distancia calculada está comprendida desde el centroide del predio hasta el medidor. Al tomar el centroide del polígono se debe

tener presente que existe un amplio rango de variación del tamaño de los predios, por lo que la distancia entre el predio y el medidor puede ser grande. Esto no significa que la unificación sea errónea. Para disminuir la ocurrencia de grandes distancias en predios grandes se debe medir desde el lado del predio más cercano al medidor. En la Fig. 4 se presenta un esquema del procedimiento realizado. Si la distancia señalada como M es menor a 45 m el registro asociado es verificado y aceptado.

#### 2.5. Análisis general de datos

La superficie de las 32 parroquias que intervienen dentro del estudio es distinta, como consecuencia, el número de datos difiere en cada parroquia. Por ello, se realizó un análisis de representatividad de los datos unidos por parroquia de acuerdo a la proporción de datos de cada una respecto al total.

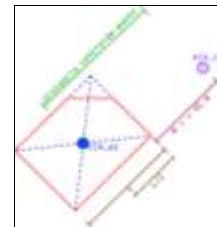


Figura 4: Cálculo de distancia predio – medidor

Se calculó el porcentaje de datos de la base de datos eléctrica completa por parroquia y se comparó con el porcentaje de datos de la base de datos unificada, así se identificaron aquellas parroquias cuya proporción de datos unificados es mayor, menor o igualitaria respecto a la base de datos completa.

Además de conseguir una proporción representativa en cada parroquia, es importante conocer si su distribución espacial es homogénea ya que no es conveniente que existan datos concentrados en determinadas zonas para futuras aplicaciones en materia de gestión energética. Al tener las coberturas georreferenciadas, con la ubicación de los datos se identificó la distribución espacial en cada parroquia.

Los diferentes análisis espaciales y entre atributos de cada fase, se realizaron utilizando diferentes tipos de software. Los procesos espaciales se realizaron con el software de Sistemas de Información Geográfica (SIG): QGIS. El análisis de comparación de campos por atributos en común, se lo realizó mediante la programación de comandos software estadístico: R-Studio. Los dos motores de análisis son de código abierto.

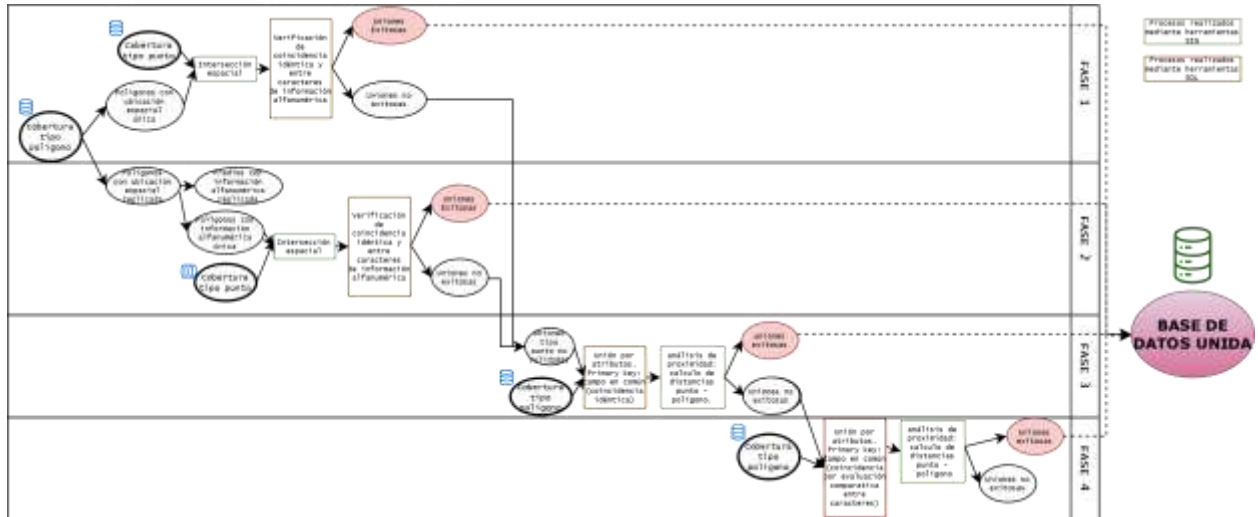


Figura 5: Proceso metodológico de unificación de datos punto - polígono

### 2.6. Validación

Una vez unificadas las bases de datos, se validaron los resultados obtenidos de las uniones predio – medidor con datos levantados en sitio. Del total de entidades asociadas por parroquia, se seleccionaron aquellas con uniformidad en la distribución espacial de los datos unidos. De estas parroquias se seleccionó una muestra aleatoria representativa del conjunto de datos asociados, los cuales fueron validados en sitio.

La toma de datos en sitio se llevó a cabo mediante aplicaciones móviles gratuitas para el levantamiento de información geográfica, donde principalmente se registró el número de medidor, número de casa y fotografía de la fachada de la muestra. Con estos datos, por un lado, se validó el medidor correspondiente al predio. Mediante el Código Único Eléctrico (CUEN) de cada medidor, se descargó la factura de cada uno y con ello se comparó que el número de medidor corresponda al número registrado en sitio. Por otro lado, se comparó la imagen de la fachada principal tomada en sitio, con la fotografía disponible en la cédula catastral del portal de datos del municipio de Quito. Se consideraron como válidas a aquellas uniones con coincidencia de número de medidor y fotografía de fachada.

## 3. RESULTADOS

### 3.1 Selección de relaciones de unión

Las relaciones espaciales y por atributos que implican dos o más datos a unirse con información alfanumérica en común, son más complejas de tratarse, ya que las entidades están representadas de manera distinta en cada base de datos. Esto genera incertidumbre al realizar la unificación. Por esta razón, en el presente estudio se seleccionaron las relaciones de unión espaciales: a, b, d, f, g y la relación alfanumérica de uno a uno (Fig. 3).

### 3.2 Unión de bases de datos

Para evaluar los resultados obtenidos luego de aplicar la metodología propuesta, se comparó el estado de los atributos de las coberturas de insumo frente a los atributos de la cobertura final unificada. Dentro de la cobertura catastral se han identificado predios en unipropiedad (viviendas unifamiliares) y en propiedad horizontal (viviendas multifamiliares). Se han obtenido varios casos de unión conforme las relaciones de unión existentes y las fases de unificación.

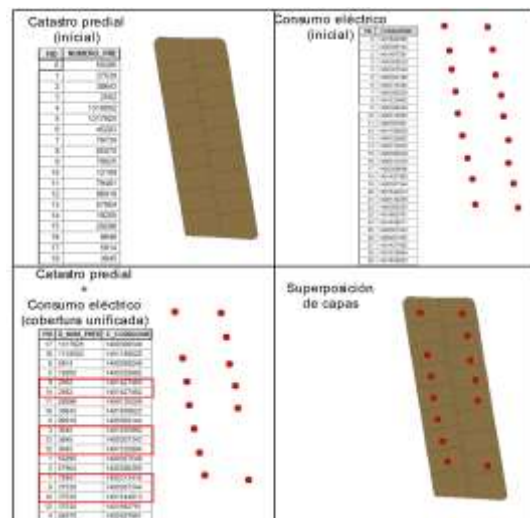


Figura 6: Unificación en la Fase 1 con relaciones “uno a uno” y “uno a varios”

#### 3.2.1 Fase 1

En la Fig. 6 se muestra un ejemplo de lo conseguido en la Fase 1. Inicialmente las coberturas se encuentran divididas y luego de la unión se obtiene una cobertura consolidada con los atributos de las bases de datos. Los objetos se encuentran asociados de “uno a uno” y de “uno a varios”, donde espacialmente cada medidor se



encuentra contenido en el predio asociado. Debido a que en esta fase se dividió a la cobertura de predios y se trabajó con aquellos predios unifamiliares, existe una mayor cantidad de medidores que predios.

En esta fase se obtuvo la mayor cantidad de datos unificados. El 14,4 %, que corresponde a 90 794 registros de viviendas unifamiliares, se unificaron exitosamente por el método de intersección espacial y verificación de correspondencia idéntica y entre caracteres de los campos en común entre la cobertura catastral y de consumo eléctrico. Los objetos que no se asociaron exitosamente ingresaron a la Fase 2 de unificación.

### 3.2.2 Fase 2

En el ejemplo de la Fig. 7 se indica lo conseguido en la Fase 2. La tabla de atributos de la capa resultante alberga objetos unidos bajo la relación “uno a uno”, donde cada medidor se encuentra ubicado dentro del predio asociado.

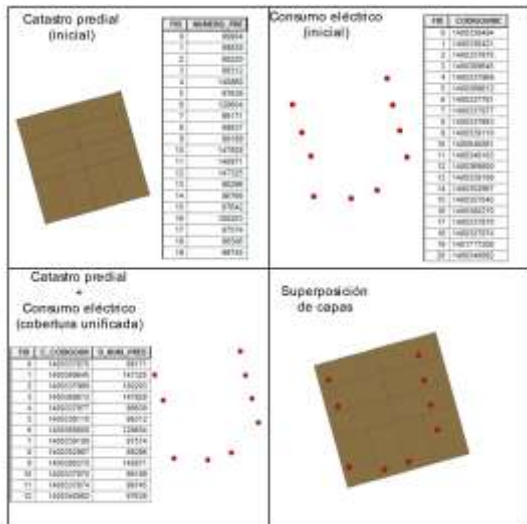


Figura 7: Unificación en la Fase 2 con relación “uno a uno”

Del número de datos restante no unificado (multifamiliares) que pasó a Fase 2, el 6,2 % fue unificado mediante el mismo método indicado en la Fase 1. La relación de unificación está basada en su coincidencia alfanumérica, son “uno a uno” y “uno a varios”. Los datos no unificados en la Fase 2 pasaron a Fase 3.

### 3.2.3 Fase 3

En la Fig. 8 se muestra el estado inicial y final de la unificación de la cobertura eléctrica y de catastro en la Fase 3. Se fusionaron a los objetos mediante campos en común, por lo que la relación de unión fue de “uno a uno”. Los medidores asociados a un predio no se encontraban en una misma ubicación espacial, por lo que se realizó un análisis de proximidad mediante el cálculo de distancias entre punto y polígono asociado.

Se consideraron como válidas a aquellas fusiones cuya distancia de separación se encontró dentro de un rango tolerable. Este procedimiento permitió la unión del 6,9 % de los datos no fusionados en la Fase 2. Los datos no unificados pasan a la Fase 4.

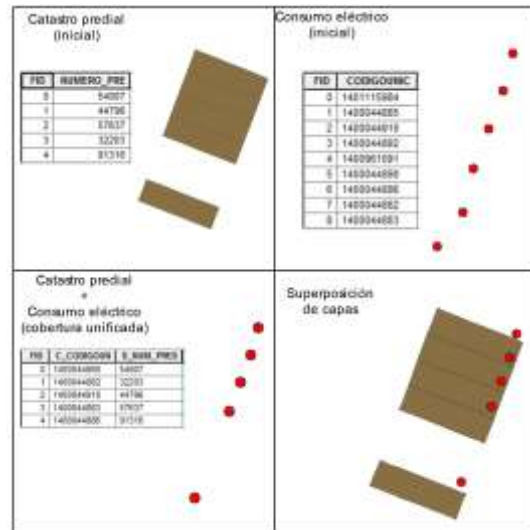


Figura 8: Unificación en la Fase 3 con relación “uno a uno”

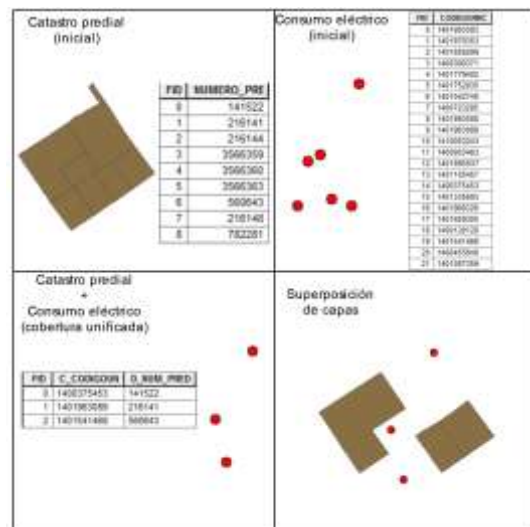


Figura 9: Unificación en la Fase 4 con relación de unión uno a uno

### 3.2.4 Fase 4

La Fase 4 corresponde a la última dentro del proceso de unificación. En la Fig. 9 se indica un ejemplo del estado de las bases de datos antes y después de la unificación. Se consiguió fusionar a aquellos registros donde se encontró diferencias menores entre los campos en común. Espacialmente, no existió una coincidencia espacial entre punto y polígono asociados, por lo que en estos casos se calculó la distancia existente para conservar o descartar registros fusionados. Finalmente, se consiguió unificar el 0,04 % de objetos no asociados resultante de la Fase 3.

Considerando la base de datos completa de medidores de energía eléctrica, en total se logró la unificación del 24,8 % correspondiente a 157 334 registros de medidores de energía eléctrica asociados a edificaciones.

### 3.3 Análisis general de datos asociados

Adicionalmente, se evaluó si la proporción de medidores asociados por parroquia es similar a la proporción de la base de datos de consumo eléctrico completa. Los resultados mostraron que las parroquias que no están siendo bien representadas en la base de datos unificada son principalmente Iñaquito y Quitumbe.

En el caso de la parroquia Iñaquito, en la base de datos completa de consumo eléctrico representa el 5,1 % mientras que en la base de datos unificada el 3,4 %. Para que en esta parroquia el número de datos unificados sea representativa, se espera que existan 8 040 datos unificados, el análisis muestra que existen 5 411, lo que indica que hay un faltante de 3 514 registros unificados. Por el contrario, las parroquias mejor representadas son Kennedy y El Condado ya que se lograron unificar más registros de los esperados.

### 3.4 Distribución espacial de datos asociados

Además de la cantidad de datos, es importante representar su distribución espacial. En la Fig. 10 se observa la presencia de datos alrededor de todas las parroquias. Existen ciertos vacíos que, en parroquias como El Condado, Concepción, Iñaquito, Quitumbe, Centro Histórico, entre otras, corresponden a áreas abiertas como zonas de recreación y parques. La parroquia que muestra una distribución espacial no uniforme, es decir mayores vacíos en su territorio es Turubamba, Comité del Pueblo y Turubamba.

### 3.5. Validación en sitio

Al revisar la distribución espacial de los datos unificados por parroquia, se encontró que La Kennedy de la zona Norte, e Iñaquito de la zona Centro – Sur, presentan mejor distribución espacial. Como se muestra en la Fig. 11, en las dos parroquias las entidades unificadas se distribuyen alrededor de toda la parroquia y fueron seleccionadas para a la validación en sitio. En la parroquia Iñaquito mediante la comparación del CUEN de cada medidor se validó el 20,91 % de datos, y por fachadas se validó el 95 %. En el caso de la parroquia La Kennedy, se validó el 61 % de entidades unificadas por medio del CUEN y el 92 % por la fachada de la edificación. Al realizar la toma de datos en esta parroquia, se encontró que el 32 % de medidores se hallaban en un lugar no visible para ser registrados por lo que de este porcentaje de datos no se tiene información.

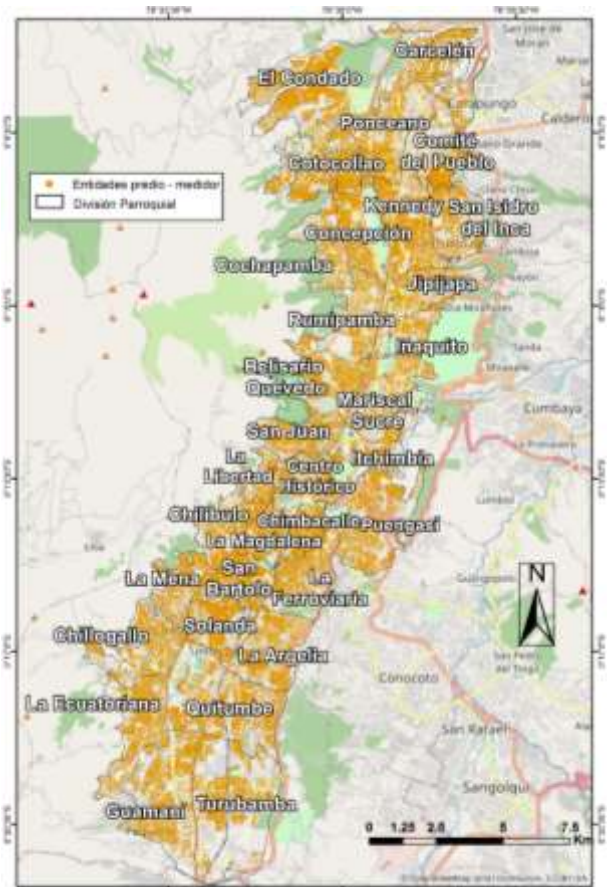


Figura 10: Distribución espacial de los datos unificados

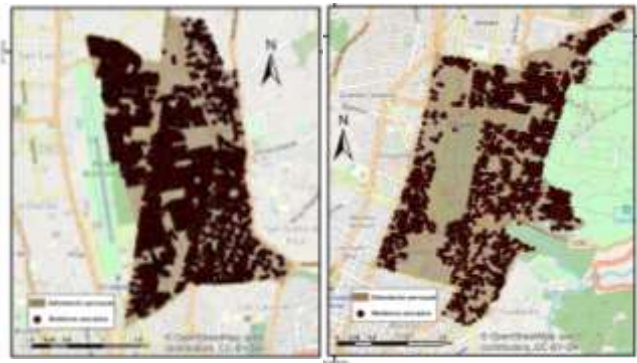


Figura 11: Distribución espacial de datos unificados de la parroquia Kennedy e Iñaquito

## 4. DISCUSIÓN

Lo ideal, para que una unificación sea sencilla, es contar con objetos cuya ubicación espacial corresponda a un solo registro en su tabla de atributos. De esta forma, se cuenta con la misma cantidad de objetos espaciales y registros alfanuméricos, con ello la relación de unión es de “uno a uno”. No obstante, no es el caso de los insumos manejados, siendo una limitante encontrada en la mayor parte de estudios relacionados a métodos de unificación de bases de datos.



El manejo de datos eléctricos y catastrales en Quito, implica la intervención de una gran cantidad de información que no puede ser unificada únicamente bajo la relación de “uno a uno”. Por lo que, el método de intersección espacial en combinación con los comandos de programación SQL de comparación idéntica y entre caracteres, presentó una mayor robustez en la unificación para predios registrados como viviendas unifamiliares. Sin embargo, en esta fase tuvieron intervención aquellos predios registrados como viviendas multifamiliares, donde es necesario tomar medidas alternas, computacionales o de trabajo de campo, para la unificación.

Los resultados obtenidos, con la aplicación de las herramientas de unión por atributos y comparación de campos en común, indican que la funcionalidad de las herramientas es la misma independientemente del orden de ejecución. Conociendo esto, es posible optimizar los procesos para que la carga computacional y el tiempo de ejecución, disminuya. La esquematización de los procesos en cada fase, contribuye a la aplicación de la metodología planteada. De esta forma es posible realizar cualquier tipo de mejora al proceso propuesto, lo cual no se puede contemplar en los escasos métodos de unificación encontrados en la literatura.

La unificación por atributos, mediante la comparación entre caracteres de los atributos en común de las dos bases de datos de insumo, se realiza mediante una comparación objeto a objeto, de la totalidad de datos de interés. La carga computacional para la ejecución de este proceso es alta, por lo que generalmente se hace referencia al alto grado de robustez del equipo a usar como condicionante para llevar a cabo cualquier tipo de operación donde intervenga una gran cantidad de datos. La presente metodología plantea dividir los datos por zonas y así disminuir el volumen de datos, lo que representa una alternativa para aplicar la herramienta sin que la capacidad computacional del equipo sea una limitante.

Los 157 334 registros unificados cuentan con información acerca del consumo eléctrico histórico desde enero de 2018 hasta diciembre de 2020 e información catastral del predio correspondiente, como área de construcción, número de pisos, área del terreno, ubicación, características constructivas de la envolvente, entre otros. De acuerdo con varios autores, la base de datos obtenida es el insumo más importante para el desarrollo de una diversidad de estudios enfocados en el comportamiento y rendimiento energético de edificaciones residenciales.

La aplicación combinada de herramientas SIG y comandos SQL representa un potencial instrumento para la unificación de bases de datos eléctricas y catastrales. Lo cual beneficia a las diferentes organizaciones dedicadas a la gestión del territorio y a su vez, a aquellas creadas para la administración de energía eléctrica. Con los resultados obtenidos, es

posible establecer indicadores o parámetros que permitan determinar políticas de eficiencia energética sobre el territorio.

## 5. CONCLUSIONES Y RECOMENDACIONES

Al aplicar la metodología planteada sobre la información de las bases de datos catastral y de consumo eléctrico del casco urbano de Quito, se logró unificar el 25 % de datos de consumo eléctrico a información predial en cuatro fases, lo que corresponde a 157 334 datos de 633 800 en total.

Para llevar a cabo el proceso de unificación entre dos coberturas tipo punto y polígono, los métodos SIG adecuados son: intersección espacial, unión por atributos y proximidad entre puntos. La funcionalidad estas herramientas SIG se potencializa con el uso combinado de herramientas SQL.

El tipo de vivienda es un factor condicionante para la aplicación de la metodología de unión. Al analizar las relaciones geométricas de unión existentes, se encontraron limitantes al unificar a aquellos objetos tipo polígono (predios) que se encuentran en propiedad horizontal o como viviendas multifamiliares debido a la incertidumbre existente ante la presencia de un solo medidor en varias unidades habitacionales.

Las parroquias Ñaquito y Quitumbe presentaron la menor cantidad de asociaciones, mientras que El Condado y Kennedy poseen la mayor cantidad, con el 6.5 % y 5.1 % respectivamente. Esto se dio debido a que en Ñaquito y Quitumbe existe una gran cantidad de viviendas unifamiliares donde no se pudo establecer una relación de unión válida.

La recolección de datos en sitio comprobó la validez de la metodología planteada en este estudio. Los errores encontrados en la validación se deben a posibles causas como *\_es humanos* en la toma de datos, cambio de propietario de los medidores o difícil acceso al medidor, dificultándose el levantamiento de información.

La metodología desarrollada puede ser replicada en otras áreas de estudio donde se requiera unificar bases de datos con características similares a la información de insumo utilizada.

Para perfeccionar la georreferenciación de la base de datos espacial de medidores de energía eléctrica, se recomienda realizar una revisión a escala grande de las zonas donde exista una cantidad menos representativa de registros unificados.

## AGRADECIMIENTOS

Este trabajo fue llevado a cabo gracias al financiamiento de la Empresa Eléctrica Quito (EEQ) y la cooperación de la Secretaría de Territorio, Hábitat y Vivienda del Gobierno Autónomo Descentralizado del Distrito Metropolitano de Quito. Los autores agradecen



a las instituciones por la gestión realizada para proporcionar la información necesaria para el desarrollo del estudio.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] Agencia Internacional de la Energía, “Seguimiento de edificios 2020 - Análisis - IEA.”
- [2] Agencia de Regulación y Control de Energía y Recursos Naturales no Renovables, “Estadística anual y multianual del sector eléctrico ecuatoriano, 2020,” 2021. [Online]. Available: <https://www.controlrecursosyenergia.gob.ec/wp-content/uploads/downloads/2021/06/Estadistica-2020-baja.pdf>
- [3] M. Baquero and F. Quesada, “Eficiencia energética en el sector residencial de la Ciudad de Cuenca, Ecuador,” *Maskana*, vol. 7, no. 2, pp. 147–165, 2016, doi: 10.18537/mskn.07.02.11.
- [4] M. Ghajarkhosravi, Y. Huang, A. S. Fung, R. Kumar, and V. Straka, “Energy benchmarking analysis of multi-unit residential buildings (MURBs) in Toronto, Canada,” *J. Build. Eng.*, vol. 27, no. October 2019, p. 100981, 2020, doi: 10.1016/j.jobe.2019.100981.
- [5] Y. J. Ahn and D. W. Sohn, “The effect of neighbourhood-level urban form on residential building energy use: A GIS-based model using building energy benchmarking data in Seattle,” *Energy Build.*, vol. 196, pp. 124–133, 2019, doi: 10.1016/j.enbuild.2019.05.018.
- [6] G. Mutani, R. Fontana, and A. Barreto, “Statistical GIS-based analysis of energy consumption for residential buildings in Turin (IT),” in *IEEE International Conference and Workshop in Obuda on Electrical and Power Engineering*, 2019, pp. 179–184. doi: 10.1109/CANDO-EPE47959.2019.9111035.
- [7] Y. Ma and G. Reynders, “Data-driven statistical analysis of energy performance and energy saving potential in the Flemish public building sector,” *J. Phys. Conf. Ser.*, vol. 1343, no. 1, 2019, doi: 10.1088/1742-6596/1343/1/012051.
- [8] J. Ma and J. C. P. Cheng, “Estimation of the building energy use intensity in the urban scale by integrating GIS and big data technology,” *Appl. Energy*, vol. 183, pp. 182–192, 2016, doi: 10.1016/j.apenergy.2016.08.079.
- [9] G. Mutani, R. Fontana, and A. Barreto, “Statistical GIS-based analysis of energy consumption for residential buildings in Turin (IT),” *CANDO-EPE 2019 - Proc. IEEE 2nd Int. Conf. Work. Obuda Electr. Power Eng.*, pp. 179–184, 2019, doi: 10.1109/CANDO-EPE47959.2019.9111035.
- [10] U. Ali et al., “A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings,” *Appl. Energy*, vol. 267, no. 114861, 2020, doi: 10.1016/j.apenergy.2020.114861.
- [11] G. Dall’O’, A. Galante, and M. Torri, “A methodology for the energy performance classification of residential building stock on an urban scale,” *Energy Build.*, vol. 48, pp. 211–219, 2012, doi: 10.1016/j.enbuild.2012.01.034.
- [12] J. Taco and L. Tipán, “Metodología para la determinación de indicadores de Eficiencia Eléctrica en la Zona Residencial,” *Rev. Técnica “Energía,”* vol. 16, no. 2, pp. 70–90, 2020, doi: 10.37116/revistaenergia.v16.n2.2020.354.
- [13] J. Espinoza, Q. Felipe, A. Calle, and O. Jessica, “Estándares de sustentabilidad para viviendas patrimoniales del Centro Histórico de Cuenca – Ecuador,” *Estoa*, vol. 8, no. 15, pp. 65–77, 2019, doi: 10.18537/est.v008.n015.a06.
- [14] Empresa Pública de Movilidad y Obras Públicas, “Plan Maestro de Movilidad para el Distrito Metropolitano de Quito 2009-2025,” 2009.
- [15] Ministerio de Desarrollo Urbano y Vivienda, “ACUERDO MINISTERIAL Nro. 017 - 20.”
- [16] Instituto Geográfico Militar, “Protocolo De Fiscalización Para Proyectos De Generación De Cartografía Base Con Fines Catastrales Escala 1 : 1 000, Obtenida a Través De Método Aerofotogramétrico, Digitalización 2D Sobre Ortofoto O Levantamiento Topográfico,” Quito, Ecuador, 2019.
- [17] A. Gruber, C.-H. Su, S. Zwieback, W. Crow, W. Dorigo, and W. Wagner, “Recent advances in (soil moisture) triple collocation analysis,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 45, pp. 200–211, 2016, doi: 10.1016/j.jag.2015.09.002.



**Lisseth Jami Aymacaña.-** Nació en Latacunga, Ecuador en 1994. Recibió su título de Ingeniera Geógrafa y del Medio Ambiente de la Escuela Politécnica del Ejército en 2019. Sus campos de investigación están relacionados con el uso de tecnologías y herramientas aplicables dentro de los Sistemas de Información Geográfica (SIG) para la gestión territorial y fenómenos o variables que influyen en la dinámica del territorio.



**Catalina Vallejo Coral.-** Nació en Quito, Ecuador en 1986. Recibió su título de Ingeniera Mecánica de la Escuela Politécnica Nacional, en 2011; de Master en Ciencias con especialidad en Ingeniería Energética del Tecnológico de Monterrey, México en 2017. Sus

campos de investigación están relacionados con la eficiencia energética y confort térmico edificaciones residenciales y comerciales.



**Francis David Vásquez Aza.-** Nació en Ibarra, Ecuador en 1991. Recibió su título de Ingeniero Mecánico de la Escuela Politécnica Nacional en 2016 y su título de Master en Sistemas y Tecnologías de Conversión de Energía de la Universidad Rovira i Virgili,

Tarragona-España en 2020. Sus campos de investigación están relacionados con la Eficiencia Energética en Edificaciones y las Energías Alternativas.



**Luis Felipe Godoy Vaca.-** Nació en Quito, Ecuador en 1989. Recibió su título de Ingeniero Mecánico de la Escuela Politécnica Nacional en 2016. Actualmente cursa una maestría en ingeniería Química y Energética en la Universidad Otto-von-Guericke en

Magdeburg, Alemania. Sus campos de investigación están relacionados con la simulación energética y la eficiencia energética.



**José Luis Condo León.-** Nació en Riobamba, Ecuador en 1992. Recibió su título de Ingeniería en Estadística Informática de la Escuela Superior Politécnica de Chimborazo, en 2020. Su campo de investigación esta con el análisis de datos multivariados, utilizando

técnicas estadísticas, en el ámbito de la energía.