




Methodological Proposal for Targeting Electric Subsidy "Tarifa De La Dignidad" in Residential Users of the Electric Companies of Ecuador

Propuesta de una Metodología para la Focalización del Subsidio Eléctrico "Tarifa De La Dignidad" en Usuarios Residenciales de las Empresas Eléctricas del Ecuador

 E.G. Novoa¹

 G.B. Salazar¹

 E.M. Buitrón²

 G.S. Salazar¹

¹Departamento de Energía Eléctrica, Escuela Politécnica Nacional, Quito, Ecuador
E-mail: edison.novoa@epn.edu.ec; gabriel.salazar@epn.edu.ec; gabriel.salazar01@epn.edu.ec

²Facultad de Ingeniería Civil y Ambiental, Escuela Politécnica Nacional, Quito, Ecuador
E-mail: eliana_buitron@hotmail.com

Abstract

This paper includes a proposal for targeting the dignity tariff subsidy on residential users of electric power distribution companies in Ecuador. For targeting, the grouping data technique is used based on technical, geographical, and socioeconomic user's data.

This paper describes the process necessary to collect, clean, and process the data for targeting the subsidy. Then, this paper shows the process to match the data and consolidate it into a single database of residential users. In a next phase, this paper defines a poverty equation that considers, in a weighted way, the technical and socioeconomic characteristics of the users and poverty equation allows calculating the poverty level of users. Then, the way to group users according to their poverty level is described, as well as the way to establish the best number of groups to be generated among users. Finally, this paper describes how to define which users and user groups should or should not receive the subsidy.

This paper has been applied to the residential users of the Quito District Metropolitan and, for these users, the results of the targeting are showed. For the study case the estimated savings that the State would receive for including the targeting methodology are shown too.

Index terms— Distribution power systems, data mining, clustering technique, spatial data analysis, subsidy, targeting.

Resumen

El artículo presenta una propuesta para la focalización del subsidio tarifa de la dignidad en los usuarios residenciales de las empresas de distribución de energía eléctrica del Ecuador. Para la focalización se utiliza la técnica de agrupación de datos considerando las características técnicas, geográficas y socioeconómicas de los usuarios.

El artículo describe el proceso necesario para recopilar, depurar y procesar los datos a fin de que puedan ser utilizados para la focalización del subsidio. Luego, se detalla el proceso necesario para emparejar los datos y consolidarlos en una sola base de datos de usuarios residenciales. En una siguiente fase, se define una ecuación de pobreza que considera, de manera ponderada, las características técnicas y socioeconómicas de los usuarios y que permite calcular su nivel de pobreza individual. Luego, se describe la forma de agrupar a los usuarios según su nivel de pobreza, así como la manera de establecer la mejor cantidad de grupos que se deben generar de entre los usuarios. Finalmente, el artículo presenta una propuesta metodológica general para focalizar el subsidio tarifa de la dignidad en el Ecuador.

El presente artículo presenta los resultados de la focalización del subsidio tarifa de la dignidad aplicada a los usuarios residenciales del Distrito Metropolitano de Quito.

Palabras clave— sistemas eléctricos de distribución de electricidad, minería de datos, técnicas de agrupación de datos, análisis espacial de datos, subsidios, focalización.

Recibido: 23-04-2023, Aprobado tras revisión: 14-06-2023

Forma sugerida de citación: Novoa, E.; Salazar, G.Y.; Buitrón, E.; Salazar, G.S. (2023). "Propuesta de una Metodología para la Focalización del Subsidio Eléctrico "Tarifa de la Dignidad" en Usuarios Residenciales de las Empresas Eléctricas del Ecuador".

Revista Técnica "energía". No. 20, Issue I, Pp. 16-25

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

Doi: <https://doi.org/10.37116/revistaenergia.v20.n1.2023.579>

© 2023 Operador Nacional de Electricidad, CENACE



1. INTRODUCCIÓN

Actualmente, en el Sector Eléctrico Ecuatoriano, el subsidio tarifa de la dignidad, emitido mediante decreto ejecutivo No. 451-A vigente desde el 12 de julio de 2007 y conforme con el procedimiento para la aplicación del subsidio, emitido por la ARCERNNR, define que: aquellos usuarios residenciales cuyos consumos mensuales de energía sean inferiores a 110 kWh – mes en las empresas distribuidoras de la región Sierra y 130 kWh – mes en las de la región Costa/Oriente e Insular pagarán un costo preferencial de energía consumida igual a 0,04 US\$/kWh y un cargo por comercialización de 0,70 US\$ por abonado al mes, puesto que dichos usuarios representan a los quintiles 1 y 2 de ingresos según la Encuesta de Condiciones de Vida elaborada por el Instituto Nacional de Estadísticas y Censos - INEC [1].

El subsidio tarifa de la dignidad se originó hace 15 años con dos objetivos: el primero, reducir el valor de la planilla eléctrica de los usuarios de las Empresas Eléctricas con escasos recursos económicos y, el segundo, reducir las pérdidas no técnicas de energía en las Empresas Eléctricas; sin embargo, en la práctica, el subsidio ha terminado beneficiando a usuarios pobres y no pobres; y este, es justamente el problema, que el subsidio no está direccionado para beneficio exclusivo de usuarios pobres. El subsidio tarifa de la dignidad no se ha actualizado a lo largo de los años y su aplicación nunca ha dependido de las características socio económicas de los usuarios. Por tanto, es importante focalizar el subsidio tarifa de la dignidad puesto que permitiría reducir el monto de dinero que se destina para este beneficio.

La asignación del subsidio eléctrico tarifa de la dignidad ha tenido problemas desde su creación, puesto que la falta de información socioeconómica de los usuarios residenciales se convertiría en una limitante para su aplicación, y, puesto que se escogió como criterio de asignación condicional al consumo de energía característico de las familias con quintiles de ingresos económicos más bajos del Ecuador. El problema es que no se consideró que dicho criterio condicional no es bidireccional y que, peor aún, no aplica necesariamente en el sentido inverso; es decir, el hecho de que un usuario consume poca energía eléctrica no implica que sea pobre [2]. La literatura económica recomienda a los gobiernos la eliminación de los subsidios generalizados puesto que dan señales equívocas a los usuarios finales y distorsionan el mercado; sin embargo, la medida de eliminar algún subsidio tiene un costo político y social muy alto, razón por la cual, para disminuir los subsidios, se debe usar mecanismos de focalización basados en su adecuada asignación [3].

La Agencia de Regulación y Control de Electricidad (ARCONEL actual ARCERNNR), a partir de la expedición de la LOSPEE, ha sido facultada para modificar y/o eliminar la asignación del subsidio tarifa de la dignidad y, para ello, ha focalizado el subsidio utilizando el método de frecuencia de consumo estricto a

11 meses [1]; sin embargo, este mecanismo sigue vinculado únicamente al consumo mensual de energía, parámetro que, por sí solo, no constituye un indicador determinante para definir la situación económica y social del consumidor final.

Ante los evidentes inconvenientes existentes durante la aplicación del subsidio tarifa de la dignidad surgen varias preguntas como: ¿se podría mejorar la asignación el subsidio? ¿es posible determinar geográficamente qué usuarios y qué zonas realmente requieren del subsidio eléctrico? Un ejemplo claro de que no se ha estudiado este tema, es el hecho de que no se puede saber con exactitud si el subsidio llega sólo a usuarios pobres, no se tienen cifras al respecto, nadie lo ha medido o ha investigado. Por tanto, al ser un trabajo especializado en el área de distribución de energía eléctrica, resulta un tema que merece estudiarse puesto que está estrictamente relacionado con el desarrollo comercial de las Empresas Eléctricas de distribución de energía eléctrica en el Ecuador [4].

El presente artículo plantea el desarrollo de una metodología general para la focalización del subsidio eléctrico tarifa de la dignidad que, en base al análisis de las características técnicas, geográficas y socioeconómicas de los usuarios, contribuya a mejorar la asignación del subsidio en las Empresas Eléctricas del Ecuador.

2. AGRUPACIÓN DE DATOS

La agrupación de datos o clustering es una herramienta de análisis de datos no supervisado que está enfocada en la búsqueda de grupos dentro de un conjunto de datos. A cada conjunto de datos se lo llama clúster y se caracteriza por poseer datos parecidos entre sí.

La agrupación de datos puede realizarse de diferentes maneras dependiendo del criterio de agrupación y definiendo de antemano el número de clústeres. Si el número de clústeres no es conocido la agrupación no tiene respuesta única para todos los casos y, más bien, tiene un valor específico para cada problema. Por tanto, en la mayoría de los casos, la técnica de clustering requiere de humanos expertos en el dominio de su aplicación [5].

El procedimiento de agrupación más conocido es K-means, desarrollado por Steinhouse en 1956 y popularizado por MacQueen en 1967. Este método conceptualmente es simple y se define de la siguiente manera: si μ_1, \dots, μ_K son los K centros de los clústeres, K-means los encuentra minimizando

$$J(\mu_1, \dots, \mu_K) = \frac{1}{n} \sum_{k=1}^K \sum_{x_i \in G_k} \|x_i - \mu_k\| \quad (1)$$

Donde G_k es el clúster de elementos con centro más próximo a μ_k , es decir:

$$G_k = \left\{ x_i : \|x_i - \mu_k\| = \min_{1 \leq j \leq K} \|x_i - \mu_j\| \right\} \quad (2)$$

La limitación de este método es que no da buenos resultados si dentro de los datos existen datos atípicos o si las agrupaciones tienen grandes diferencias en su orientación, tamaño y forma [6]. Existen variaciones en la técnica de agrupación puesto que la forma de calcular la similitud ($x_i - \mu_k$) puede realizarse de distintas maneras.

2.1. Evaluación de Grupos

Para evaluar los resultados de la agrupación de datos existen dos tipos de indicadores, el primero, evalúa parámetros internos de los grupos de datos y, el segundo, evalúa parámetros externos de los grupos de datos.

El Índice de Davies-Bouldin es uno de los indicadores que permite evaluar los parámetros internos de los grupos de datos resultantes de la técnica de agrupación y considera que existe una mejor agrupación si la distancia media de un grupo de datos, respecto a su centro, dividida para la distancia existente entre los centros de los grupos es la menor. La definición de dicho índice es la siguiente

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left\{ \frac{\bar{d}_i + \bar{d}_j}{d_{ij}} \right\} \quad (3)$$

Donde DB es el índice de Davies-Bouldin, k es el número de grupos, \bar{d}_i es la distancia promedio de los datos al centro del grupo i, \bar{d}_j es la distancia promedio de los datos al centro del grupo j y d_{ij} es la distancia entre los centros de los grupos i y j. La agrupación óptima para diferentes números de grupos, según este criterio, es aquella con el menor valor de DB [7], [8].

3. METODOLOGÍA

3.1. Recopilación y Procesamiento de Datos

En esta sección se ha revisado la información técnica, geográfica y socioeconómica de los usuarios residenciales del Ecuador que puede estar al acceso del personal de las Empresas Eléctricas. Una vez revisada la información disponible se ha determinado el uso de los siguientes datos:

- Datos técnicos: Consumo mensual de energía eléctrica por usuario en kWh.
- Datos geográficos: Ubicación georeferenciada de cada usuario, coordenadas (X,Y).
- Datos socioeconómicos: El valor básico referencial del metro cuadrado de suelo y los ingresos familiares mensuales promedio por sector.

La estructura de los datos mencionados se muestra en las Tablas 1, 2 y 3.

Tabla 1: Estructura de los Datos Técnicos

Código del Cliente	Consumo de Energía Mensual [kWh]
387320	90
...	...
1499703	75

Tabla 2: Estructura de los Datos Geográficos

Código del Cliente	Coordenada X	Coordenada Y
387320	780572.91	9981908.61
...
1499703	776101.09	9977057.11

Tabla 3: Estructura de los Datos Socioeconómicos

Identificador geográfico	Tipo de dato	Valoración del suelo [US\$/m ²]
250	Polígono	120
...
2303	Polígono	500

3.2. Recopilación y Procesamiento de Datos

El emparejamiento de datos se lo puede realizar mediante cualquier algoritmo de emparejamiento y/o mediante cualquier software que posea esta función. Para el artículo se ha seleccionado el uso del software Matlab, ArcGIS y Excel.

Para emparejar los datos técnicos y económicos primero se cargan los datos técnicos en Matlab como una matriz de m códigos de cada cliente y sus n consumos de energía. Luego, se cargan los datos geográficos como una matriz de p códigos de cada cliente y q atributos, en este caso, las coordenadas X y Y. Una vez cargados los datos técnicos y geográficos se los empareja utilizando como atributo común los códigos de los clientes. El resultado de este primer emparejamiento de datos se observa en la Tabla 4.

Tabla 4: Estructura de los Datos Técnicos y Geográficos

Código del Cliente	Coordenada X	Coordenada Y	Consumo de Energía Mensual [kWh]
387320	780572.91	9981908.61	90
...
1499703	776101.09	9977057.11	75

Luego, se debe ubicar espacialmente los datos económicos de valoración del suelo y de ingresos económicos. Dado que los datos de valoración del suelo se los puede encontrar directamente en archivos geográficamente referenciados no hace falta emparejarlos. Sin embargo, en el caso de los datos de ingresos económicos por núcleo familiar, es necesario primero inferir algunos datos para, luego, emparejarlos con su ubicación geográfica. Antes de ubicar espacialmente los datos de ingresos económicos por núcleo familiar primero se han promediado las respuestas de los encuestados dentro de un mismo sector a fin de inferir dicho resultado al valor de ingresos por zona. Los resultados de este emparejamiento deben tener la estructura mostrada en la Tabla 5.

Tabla 5: Estructura de los Datos Socioeconómicos y Geográficos

Identificador Geográfico	Tipo de Dato	DPA_ZONA	Ingreso Económico Promedio por Zona [US\$/mes]
1627	Polígono	170150200	750.11
...
...
...
2048	Polígono	170150272	320.43

Finalmente, es necesario emparejar los datos geográficos (X,Y), técnicos (kWh) y socioeconómicos (Valor del suelo e ingresos promedio) en un solo archivo o base de datos. Debido a que los datos provienen en distintos formatos y de distintas fuentes, es necesario primero enlazar los datos técnicos de consumo con la ubicación de los usuarios residenciales; luego, es necesario enlazar los datos socioeconómicos con su ubicación espacial para, al último, emparejar todos los datos en un solo archivo. Este último emparejamiento requiere del siguiente procedimiento:

1) Se debe generar un archivo tipo ráster con los datos de las áreas de valoración del suelo por metro cuadrado. Esta actividad se la puede realizar mediante la herramienta de conversión de datos de ArcGIS “Polygon to Raster” teniendo en cuenta que el valor de la rasterización es el valor del suelo por metro cuadrado.

2) Se debe generar un archivo tipo ráster con los datos de ingresos económicos promedio por zona. Esta actividad se la puede realizar mediante la herramienta de conversión de datos de ArcGIS “Polygon to Raster” teniendo en cuenta que el valor de la rasterización es el valor de ingresos económicos promedio por sector.

3) Una vez que los datos socioeconómicos se encuentran geográficamente ubicados, en un archivo tipo ráster, se los debe emparejar con los datos de ubicación de los usuarios residenciales y sus consumos; lo cual, se puede realizar mediante la herramienta de análisis espacial de ArcGIS “Extract Values to Points”.

3.3. Agrupación de Datos

3.3.1 Definición del nivel de pobreza

Para la agrupación de los usuarios se definirá primero una propuesta de ecuación de pobreza que considere, de manera ponderada, los criterios técnicos y socioeconómicos de los usuarios.

A continuación, se plantea la ecuación de pobreza en términos generales.

$$P(X) = w_1 \cdot X_1 + w_2 \cdot X_2 + w_3 \cdot X_3 + \dots + w_n \cdot X_n \quad (4)$$

P es la pobreza que depende de X datos o criterios normalizados, entre 0 y 1, ponderados con w pesos o costos; para la definición de la pobreza podrán existir n criterios, donde la suma de sus w_n pesos deberá ser siempre 1.

La normalización de datos, entre 0 y 1, de las variables X se realiza según la ecuación (5) y para su aplicación se debe conocer el valor máximo y mínimo ($X_{máx}$ y $X_{mín}$) de los datos originales.

$$X_{normalizada} = \frac{X_{original} - X_{mín}}{X_{máx} - X_{mín}} \quad (5)$$

La normalización de datos permite reducir el efecto de valores atípicos, extra tendenciales o demasiado grandes, volviéndolos adimensionales para el cálculo de la pobreza.

Una vez procesados los datos, considerando aquellos disponibles para este trabajo, la ecuación de pobreza puede definirse de la siguiente manera.

$$P(X_{criterios}) = w_{técnico} \cdot X_{técnico} + w_{suelo} \cdot X_{suelo} + w_{ingresos} \cdot X_{ingresos} \quad (6)$$

Donde:

$W_{técnico}$: Es la ponderación que se le asigna a los datos de consumo de energía en relación con la pobreza.

$X_{técnico}$: Son los datos normalizados de consumo de energía por usuario.

W_{suelo} : Es la ponderación que se le asigna a los datos de valoración del suelo en el que reside el usuario en relación con la pobreza.

X_{suelo} : Son los datos normalizados de valoración del suelo según la ubicación del usuario.

$W_{ingresos}$: Es la ponderación que se le asigna a los datos de ingresos económicos de los usuarios con relación a la pobreza.

$X_{ingresos}$: Son los datos normalizados de ingresos económicos por usuario.

Dado que los datos ya estarían normalizados se debe readecuar el cálculo de pobreza a fin de que tenga correspondencia con los resultados de pobreza esperados, 1 para usuarios pobres y 0 para usuarios no pobres. A continuación, se presenta una primera ecuación para el cálculo de pobreza por usuario propuesta para la focalización, la cual considera igual ponderación para cada criterio técnico, socioeconómico y geográfico.

$$P(X_{criterios}) = \frac{1}{3} * (1 - X_{técnico}) + \frac{1}{3} * (1 - X_{suelo}) + \frac{1}{3} * (1 - X_{ingresos}) \quad (7)$$

La ecuación (7) permite considerar que los usuarios con consumos de energía más bajos y que residen en sectores con valores de suelo e ingresos promedio más bajos son personas más pobres.

Si bien la ecuación (7) es una muy buena aproximación para el cálculo de pobreza es necesario ajustar los pesos de los criterios considerando la calidad de los datos, la relación de los datos con la pobreza y los objetivos que tiene la focalización. La siguiente ecuación muestra el cálculo de pobreza propuesto para la focalización.

$$P(X_{\text{criterios}}) = 0,1 * (1 - X_{\text{técnico}}) + 0,6 * (1 - X_{\text{suelo}}) + 0,3 * (1 - X_{\text{ingresos}}) \quad (8)$$

Para el criterio de consumo de energía se ha considerado un 10 % de la ponderación total puesto que, a pesar de que los datos son buena calidad, no se tiene certeza de si un menor consumo de energía está asociado con la existencia de un usuario pobre. Para el criterio de valoración del suelo se ha considerado un 60 % de la ponderación total puesto que la información es de calidad, tiene relación con la pobreza y cumple con el enfoque que se le quiso dar a la focalización. Finalmente, para el criterio de los ingresos económicos por núcleo familiar se le ha asignado solo un 30 % de la ponderación total puesto que, si bien es un criterio importante para la focalización del subsidio, los datos que se obtuvieron del INEC, en varios casos, no presentan valores de ingresos registrados por los encuestados. En cualquier caso, los pesos o ponderación de cada criterio podrán ser ajustados por los organismos Estatales encargados de ejecutar la focalización.

3.3.2 Definición de la técnica de agrupación

Para definir la manera de resolver el problema de focalización, planteado en este artículo, se han analizado tres principales aspectos:

- La naturaleza del problema
- La naturaleza de los datos
- La practicidad de la técnica de solución

La naturaleza del problema: la focalización del subsidio eléctrico requiere necesariamente de alguna técnica que permita definir qué usuarios sí deben recibir el subsidio y qué usuarios no deben recibirlo; esto, solo se puede hacer mediante la técnica de agrupación de datos o mediante la técnica de clasificación de datos. La diferencia entre la técnica de agrupación de datos y la técnica de clasificación de datos es que en la técnica de agrupación de datos no se requiere conocer, como dato de entrada, si un usuario es pobre o no pobre; contrario a la técnica de clasificación de datos. Para el caso del Ecuador, dado que de entrada no se conoce qué usuarios son pobres o no pobres, la solución al problema de focalización demanda necesariamente del uso de la técnica de agrupación de datos.

La naturaleza de los datos y la practicidad de la técnica de solución al problema: para la focalización del subsidio eléctrico en los usuarios residenciales del

Ecuador se han seleccionado variables técnicas y socioeconómicas que, por naturaleza, contienen valores solamente numéricos y que pueden ser tratados mediante arreglos de matrices o vectores.

Dado que la técnica de agrupación de datos k-means puede trabajar con variables de datos inmersas en matrices y vectores y, dado que, además solo requiere de la definición del número de grupos que se desean obtener de la agrupación, se ha decidido, por practicidad, realizar la focalización con dicha técnica de agrupación de datos. Para la configuración de la técnica k-means en Matlab se ha mantenido, como medida de la similitud predeterminada, el valor de la distancia euclídeana cuadrada entre datos.

En todo caso, para el presente artículo, no ha sido necesario realizar un análisis comparativo sobre el uso de los diferentes algoritmos de agrupación o sobre el uso de diferentes formas de calcular la similitud de los datos puesto que con cualquier algoritmo de agrupación y con cualquier forma de calcular la similitud de los datos siempre se puede llegar a obtener agrupaciones válidas de usuarios para la focalización del subsidio.

La técnica de agrupación definida y propuesta para la focalización del subsidio es kmeans y requiere, como datos de entrada, los valores resultantes de la ecuación de pobreza por cada usuario residencial y el número de grupos que se desean encontrar.

3.3.3 Evaluación y definición de grupos

Para la evaluación y definición de grupos se ha decidido calcular el índice de Davies-Bouldin para distintas cantidades de grupos de usuarios, al menos entre 2 y 20 grupos; donde la cantidad de grupos recomendada para la focalización será aquella que minimice el índice Davies-Bouldin.

3.3.4 Criterio de asignación

Una vez que se ha logrado determinar la mejor cantidad de grupos de usuarios para la focalización se debe definir qué grupos de usuarios deben o no recibir el subsidio. Existen dos maneras de definir qué grupos de usuarios, que actualmente reciben el subsidio, deberían a futuro recibir este beneficio. La primera forma, la más restrictiva, se enfoca en asignar el subsidio a los usuarios con mayor nivel de pobreza los “más pobres” y, la segunda forma, la menos restrictiva, se enfoca en dejar de otorgar el subsidio a los grupos de usuarios con menor nivel de pobreza los “menos pobres”.

Dado que el subsidio surgió inicialmente con el objetivo de favorecer a los usuarios de los dos quintiles con ingresos económicos más bajos en el Ecuador, es decir el 40 % de la población. Se propone, de manera similar, que para este trabajo se considere la focalización del subsidio para el 40 % de usuarios “más pobres” según su cálculo de pobreza.

El porcentaje de focalización lo podrían definir los organismos del Estado y/o las Empresas Eléctricas de acuerdo con los resultados que se esperasen obtener de la focalización.

4. RESULTADOS

En esta sección se presentan los resultados de aplicar la metodología propuesta para la focalización del subsidio tarifa de la dignidad en los usuarios residenciales de la Empresa Eléctrica Quito (EEQ) dentro del Distrito Metropolitano de Quito (DMQ). Para el caso de estudio los datos técnicos obtenidos corresponden a las planillas de consumo y a la geodatabase de la EEQ para el mes de enero de 2018; los datos de la Encuesta Nacional de Ingresos y Vivienda y la división político-administrativa del Ecuador corresponden a los datos del año 2010 y, finalmente, los datos económicos de valoración del suelo en el DMQ corresponden a los datos del año 2020.

4.1. Estructura de los Datos de Entrada

4.1.1 Datos técnicos

La EEQ posee dentro de su zona de concesión, la provincia de Pichincha y parte de la provincia del Napo, cerca de 1.048.575 usuarios. La cantidad de usuarios depurados que se han encontrado dentro del DMQ y que reciben el subsidio son cerca de 227.534 usuarios. A continuación, la Fig. 1 muestran la ubicación geográfica de dichos usuarios en un mapa.

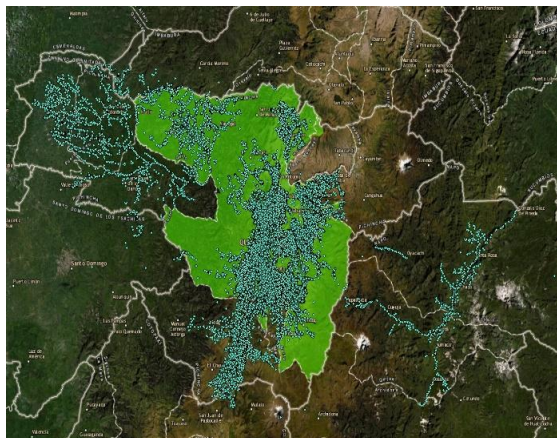


Figura 1: Usuarios de la Empresa Eléctrica Quito

4.1.2 Datos socioeconómicos

El primer dato socioeconómico definido para la focalización del subsidio es el valor del suelo por metro cuadrado. Este dato se lo encuentra en el mapa de áreas de valoración o áreas de intervención valorativas (AIVAs) del DMQ, dicho mapa está disponible en los datos abiertos del DMQ en [9]. A continuación, la Fig. 2 muestra un mapa térmico según la valoración del suelo en el DMQ, las gamas de color rojo representan los sectores donde el costo del suelo es más alto y las gamas

de color verde representan los sectores donde el costo del suelo es el menor.

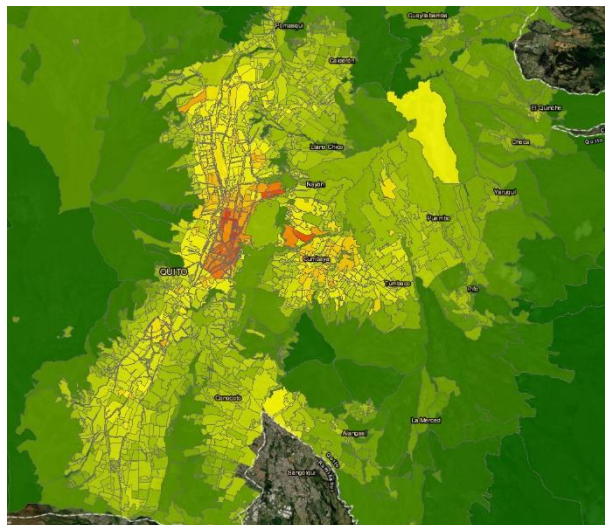


Figura 2: Mapa de Valoración del Suelo en el DMQ

El segundo dato socioeconómico corresponde a los ingresos económicos obtenidos del INEC; datos que se pueden manipular mediante el uso del software SPSS de IBM. A continuación, la Fig. 3 muestra el resultado del emparejamiento de datos socioeconómicos y geográficos en un mapa con escala térmica sobre los ingresos económicos, las gamas de color rojo representan las zonas con mayores ingresos económicos y las gamas de color verde representan las zonas con menores ingresos.

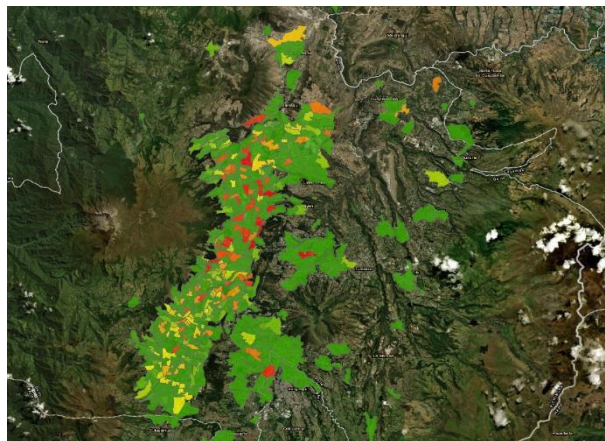


Figura 3: Mapa de la División Político-Administrativa del DMQ con Escala Térmica sobre los Ingresos

4.2. Agrupación de Datos

4.2.1 Agrupación de usuarios por su ubicación

Una primera forma de agrupar los datos es de acuerdo con su ubicación espacial, se ha realizado esta forma de agrupación utilizando la ubicación de los usuarios residenciales del DMQ, que actualmente reciben el subsidio tarifa de la dignidad, utilizando la técnica kmeans. A continuación, la Fig. 4 muestra los resultados de la agrupación de usuarios para el caso de 10 grupos.

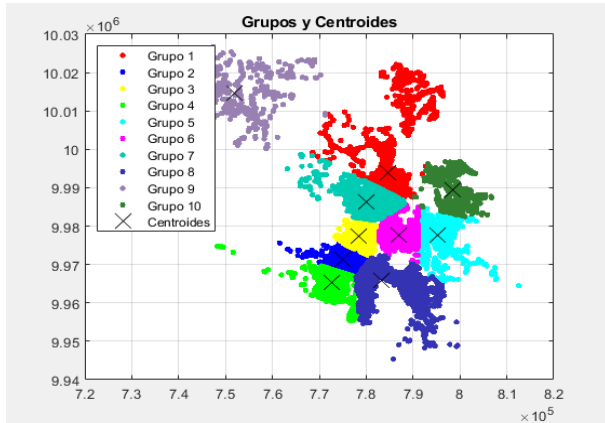


Figura 4: Agrupación Espacial de los Usuarios Residenciales del DMQ en 10 Grupos

4.2.2 Agrupación de usuarios por su consumo de energía eléctrica

Otra forma de agrupar a los usuarios, que actualmente reciben el subsidio, es de acuerdo con su consumo mensual de energía eléctrica. A continuación, la Fig. 5 muestra, por ejemplo, el resultado de generar 5 grupos de usuarios en el DMQ según sus consumos mensuales de energía eléctrica. Los usuarios se han coloreado con escala térmica, las gamas de color verde representan a los usuarios con consumos de energía más bajos y las gamas de color rojo representan a los usuarios con consumos de energía más altos.

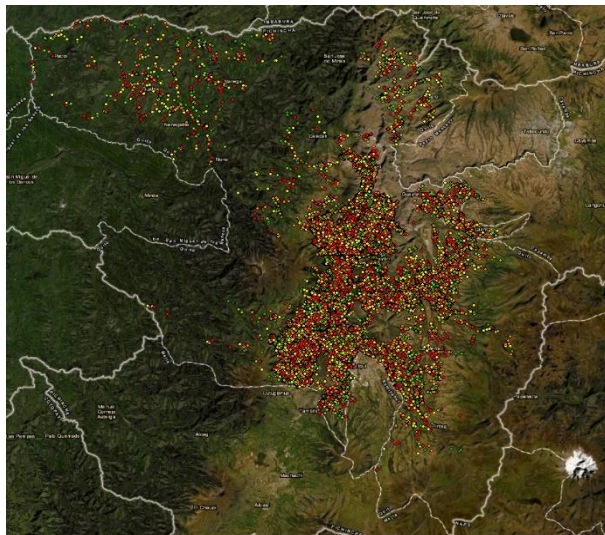


Figura 5: Agrupación de los Usuarios Residenciales del DMQ en 5 Grupos Según su Consumo de Energía Eléctrica

4.2.3 Agrupación de usuarios según su nivel de pobreza

De acuerdo con lo explicado metodología de este trabajo de investigación, para lograr considerar todas las características técnicas, socioeconómicas y geográficas de los usuarios, para la focalización del subsidio, se debe primero calcular el nivel de pobreza por usuario, 1 para usuarios pobres y 0 para usuarios no pobres; y, con dicho

resultado, realizar la agrupación sobre los valores calculados de pobreza. A continuación, la Fig. 6 muestra, por ejemplo, el resultado de formar 5 grupos de usuarios residenciales según su nivel de pobreza calculado.

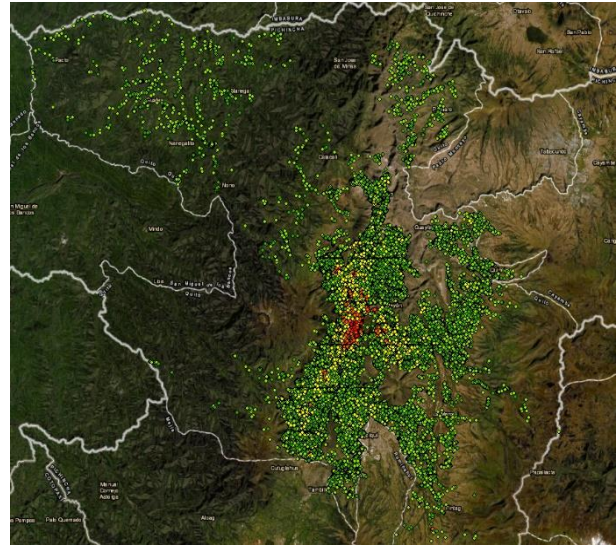


Figura 6: Agrupación de los Usuarios Residenciales del DMQ en 5 Grupos Según su Nivel de Pobreza Calculado

4.3. Evaluación de Grupos

Una vez que se ha definido que la agrupación de usuarios, según su nivel de pobreza, es la opción que permite considerar las características técnicas, geográficas y socioeconómicas de cada usuario se debe definir la mejor cantidad de grupos que se pueden generar en el universo de usuarios. Para esta actividad se ha realizado una evaluación utilizando el índice de Davies-Bouldin para distintas cantidades de grupos, al menos entre 2 y 20 grupos. A continuación, la Fig. 7 muestra la variación del índice de Davies-Bouldin para los distintos números de grupos. Como resultado de la evaluación se puede observar que el índice va reduciendo de valor a medida que se forman más grupos y aparece un óptimo local al minimizarse el índice de Davies-Bouldin a 0,50 cuando se forman 13 grupos.

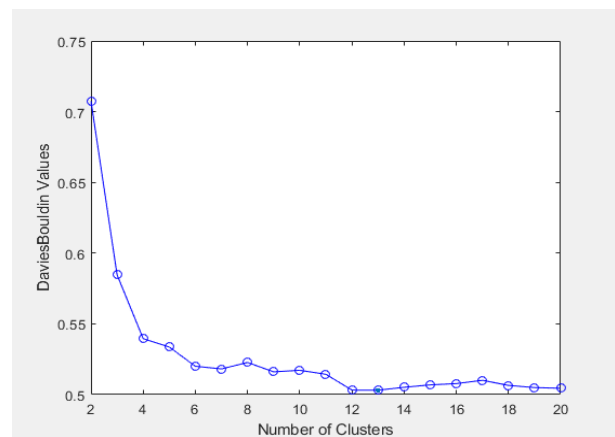


Figura 7: Índice de Davies-Bouldin para Diferentes Números de Grupos

4.4. Asignación del Subsidio

Una vez que se ha observado que, para la focalización del subsidio, la mejor forma de agrupar a los usuarios debe realizarse sobre el nivel de pobreza calculado y, una vez que se ha determinado que la mejor cantidad de grupos para agrupar a los usuarios es 13 grupos, se ha vuelto a agrupar a los usuarios del DMQ en 13 grupos, como lo muestra la Fig. 8 y la Tabla 6.

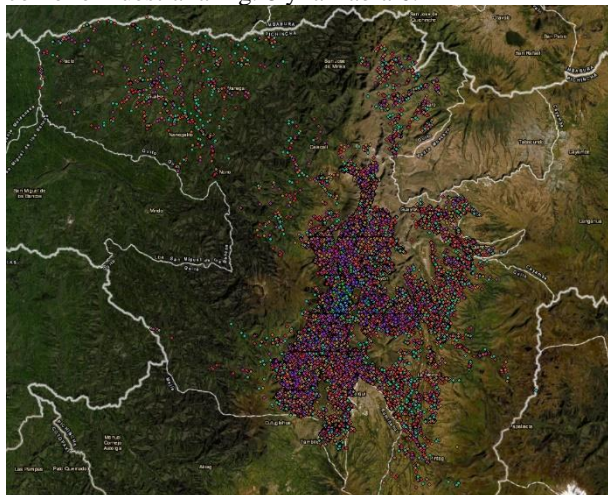


Figura 8: Agrupación de Usuarios del DMQ, en 13 Grupos, Según su Nivel de Pobreza

Tabla 6: Resultados de Centroides y Cantidad de Usuarios por grupo

Grupo	Centroide del grupo según su nivel de pobreza	Usuarios por Grupo
Grupo 1	0,39	405
Grupo 2	0,57	1.188
Grupo 3	0,70	2.297
Grupo 4	0,76	4.071
Grupo 5	0,80	9.201
Grupo 6	0,84	16.912
Grupo 7	0,86	27.409
Grupo 8	0,88	36.761
Grupo 9	0,90	36.237
Grupo 10	0,91	31.666
Grupo 11	0,93	26.900
Grupo 12	0,95	21.134
Grupo 13	0,97	13.353

El último paso que queda por realizar para la focalización del subsidio es, según el criterio más restrictivo, filtrar el 40 % de usuarios (91.014 usuarios) que se encuentren en los grupos con mayor nivel de pobreza. Según los resultados de la agrupación el 40 % de usuarios más pobres se encuentran dentro de los grupos 10, 11, 12 y 13; dado que los usuarios de un grupo no se pueden separar para cumplir exactamente con el 40 % de usuarios se ha incluido al grupo 10 con todos sus usuarios dentro de la lista de beneficiarios, dando un total de 93.053 usuarios que sí recibirían el subsidio. Para visualizar los resultados de la focalización se presenta, en la Fig. 9, los usuarios que no recibirán el subsidio en color rojo y los que sí seguirán recibiendo el subsidio en color verde.

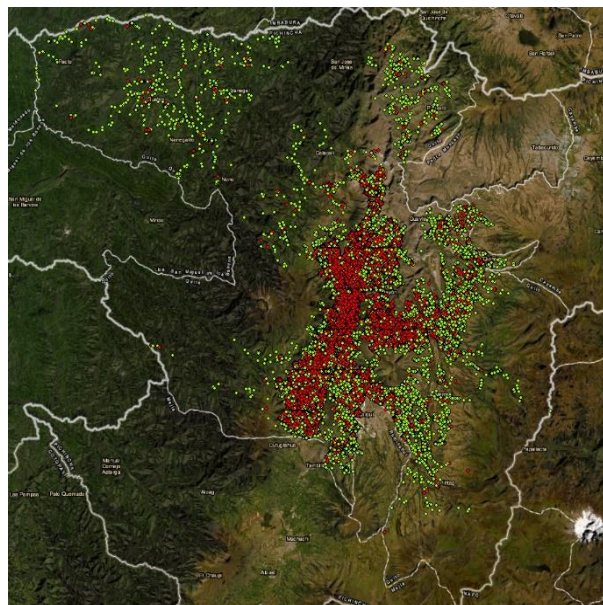


Figura 9: Mapa de Focalización del Subsidio en los Usuarios Residenciales del DMQ

4.5. Propuesta Metodológica

La propuesta metodológica para focalizar el subsidio tarifa de la dignidad considerando las características técnicas, geográficas y socioeconómicas de los usuarios residenciales de las Empresas Eléctricas del Ecuador, utilizando la técnica de agrupación de datos, se describe a continuación.

1. Recopilar los datos técnicos, geográficos y socioeconómicos de los usuarios residenciales de las Empresas Eléctricas del Ecuador.
2. Depurar los datos obtenidos y procesarlos por separado según los criterios de la sección 3.1.
3. Emparejar los datos para que cada usuario posea como atributos su ubicación, su consumo de energía y la valoración del suelo e ingresos económicos según el sitio en el que se ubica el usuario.
4. Definir una ecuación de pobreza que, para su cálculo, considere de manera ponderada los datos técnicos, geográficos y socioeconómicos de los usuarios.
5. Definir la ponderación o peso que se le dará a cada característica técnica o socioeconómica de los usuarios de acuerdo con la calidad de los datos y de acuerdo con la relación que tengan los datos con la pobreza.
6. Calcular el nivel de pobreza para cada uno de los usuarios residenciales según sus características técnicas, geográficas y socioeconómicas.
7. Determinar la mejor cantidad de grupos que se puede generar en la agrupación de usuarios residenciales minimizando, para ello, el índice de Davies-Bouldin.
8. Agrupar a los usuarios residenciales según su nivel de pobreza y en la cantidad de grupos que haya resultado del punto 7.

9. Ordenar a los grupos resultantes del punto 8 de acuerdo con el valor de su centroide de pobreza. Los grupos con un valor de centroide más cercano a 1 representan a los grupos de usuarios más pobres y los grupos con valor de centroide más cercano a 0 representan a los grupos de usuarios menos pobres.

10. Definir un porcentaje de focalización del subsidio.

11. Determinar la cantidad de grupos de usuarios y a los usuarios que, con mayor nivel de pobreza, ingresen dentro del porcentaje de focalización deseado.

12. En la base de datos del sistema comercial de cada Empresa Eléctrica se debe generar un atributo para cada usuario residencial que indique si es beneficiario o no del subsidio tarifa de la dignidad, dicho atributo se deberá llenar con los resultados de la focalización, según el código de cada cliente, para que así se lo considere como beneficiario o no beneficiario del subsidio en la facturación de energía eléctrica.

5. CONCLUSIONES Y RECOMENDACIONES

La agrupación de usuarios, en base a su ubicación, no constituye, por sí sola, una forma de focalizar el subsidio tarifa de la dignidad debido a que los grupos y la cantidad de grupos que resultan de la agrupación no presentan una relación observable con la pobreza. Por tanto, para focalizar el subsidio tarifa de la dignidad es necesario estructurar una base de datos que permita relacionar la ubicación de los usuarios residenciales con sus características de consumo de energía eléctrica y con sus características socioeconómicas de valoración del suelo e ingresos económicos, para, sobre dicha base de datos, calcular un nivel de pobreza por usuario y así agruparlos según su pobreza; la cantidad de grupos que se requiere para la focalización del subsidio se podría definir evaluando el índice de Davies-Bouldin.

AGRADECIMIENTOS

Un agradecimiento especial a la Escuela Politécnica Nacional por apoyar al desarrollo del presente artículo técnico.

REFERENCIAS BIBLIOGRÁFICAS

- [1] ARCONEL, «Agencia de Regulación y Control de Electricidad.» 08 10 2020. [En línea]. Available: <https://www.regulacionelectrica.gob.ec/tarifa-dignidad/>. [Último acceso: 08 10 2020].
- [2] J. Franco y G. Aragón, «Consortio de investigación económico y social.» 20 Octubre 2020. [En línea]. Available: <http://www.cies.org.pe/sites/default/files/files/articulos/economiasociedad/02-franco-aragon.pdf>. [Último acceso: 20 Octubre 2020].
- [3] Bhagwati, Jagdish y Ramaswami, “Domestic Distortions, Tariffs, and the Theory of Optimum

Subsidy,” *Journal of Political Economy*, vol. 71, n° 1, p. 7, 1963.

- [4] J. Franco y G. Aragón, “La focalización es relevante: propuesta de un esquema óptimo de subsidios al consumo eléctrico residencial e impactos sociales de su implementación,” Centro de Investigación de la Universidad del Pacífico, Lima, 2010.
- [5] E. Gallardo, “Técnicas de clustering aplicadas a la resolución de problemas de optimización combinatoria con restricciones espaciales y temporales, Leganés”: Universidad Carlos III de Madrid, 2018.
- [6] J. González, Métodos de clustering robustos, Buenos Aires: Universidad de Buenos Aires, 2019.
- [7] D. L. Davies y D. W. Bouldin, “A Cluster Separation Measure,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vols. 1, n° 2, pp. 224-227, 1979.
- [8] G. C. Cabrera-Celi, E. G. Novoa-Guaman and P. F. Vasquez-Miranda, "Design of secondary circuits of distribution networks using clustering and shortest path algorithms", 2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), pp. 1-6, sep. 2017.
- [9] Gobierno Abierto Quito, «Gobierno Abierto Quito,» Municipio Metropolitano de Quito, 10 10 2020. [En línea]. Available: http://gobiernoabierto.quito.gob.ec/?page_id=1122. [Último acceso: 10 10 2020].



Edison Novoa Guamán.- Nació en Quito, Ecuador en 1992. Recibió su título de Ingeniero Eléctrico de la Escuela Politécnica Nacional en 2015; de Máster en Ingeniería Eléctrica en Distribución de la Escuela Politécnica Nacional en 2021. Sus campos de investigación están relacionados con los sistemas eléctricos de distribución de energía.



Gabriel Salazar Yépez.- Nació en Quito, Ecuador en 1975. Recibió su título de Ingeniero Eléctrico de la Escuela Politécnica Nacional en 2000; de Doctor en Ingeniería Eléctrica de la Universidad Nacional de San Juan, Argentina, en 2005. Sus campos de investigación están relacionados con los sistemas eléctricos de generación, transmisión y distribución de energía.



Eliana Buitrón Hernández.- Nació en Carchi, Ecuador en 1992. Recibió su título de Ingeniera Civil de la Escuela Politécnica Nacional en 2017. Sus campos de investigación están relacionados con el análisis estructural, el estudio de comportamiento del suelo, el ensayo de materiales y la georreferenciación.



Gabriel Salazar Pérez.- Nació en Quito, Ecuador en 1997. Recibió su título de Ingeniero Eléctrico de la Escuela Politécnica Nacional en 2020; es estudiante egresado del Programa de Maestría en Electricidad Mención Redes Eléctricas Inteligentes de la Escuela Politécnica Nacional. Sus campos de investigación están relacionados con los sistemas eléctricos de generación, transmisión y distribución de energía.