DOCUMENTO METODOLÓGICO

Metodología para la recolección de información primaria

Convenio interadministrativo UPME No. CV-008-202, IPSE No.147-2021, suscrito entre la Unidad de planeación minero energética- UPME; el Instituto de planificación y promoción de soluciones energéticas para las zonas no interconectadas - IPSE y la Universidad del Cauca.

|  |  |  |
| --- | --- | --- |
| Un dibujo de una cara feliz  Descripción generada automáticamente con confianza baja | Símbolos | Universidad del Cauca | C:\Users\crstn\Downloads\Logotipo IPSE-02 (1).png |

PERS Cauca

Plan de energización Rural Sostenible

**Documento Metodológico**

**Metodología para la recolección de información primaria**

Sistematización, Redacción y Análisis

Maximiliano Bueno López

Director Académico del Proyecto

Andrés Mauricio Gómez Sánchez

Líder caracterización socioeconómica

Cristian Eduardo García Bermúdez

Profesional en Estadística

Luis Felipe Jiménez Chaves

Analista de datos

Juliana Isabel Sarmiento Castillo

Analista de datos

Plan de energización Rural Sostenible del Departamento del Cauca

PERS-CAUCA

Universidad del Cauca

UPME

IPSE

Popayán, Cauca

Colombia

Noviembre de 2021

# Introducción

El presente documento tiene como propósito básico explicar la metodología utilizada en el proceso de determinación de la muestra representativa para la población rural del Departamento del Cauca. Esto se realiza a partir del diseño de instrumentos estadísticos apropiados para el levantamiento de información primaria.

Este documento será la guía para el trabajo de campo que se desarrollará durante la ejecución del PERS Cauca y busca tener información en torno al aprovechamiento de la energía por fuente, tenencia y usos para diversos sectores, en especial, el residencial y otros como el comercial e industrial.

El documento consta de cuatro partes: i) los objetivos, los cuales muestran la relevancia del PERS Cauca en su contexto geopolítico; ii) el diseño de la muestra del proyecto; iii) la población objeto del estudio, que para el caso del departamento será de 275.985 viviendas existentes en el área rural del territorio y iv) el tamaño de muestra de 2.091 encuestas que se aplicarán en igual número de viviendas en las siete (7) regiones y 63 encuestas adicionales para unidades comerciales e industriales.

# Objetivos

En seguimiento a los objetivos generales del Plan de Energización Rural Sostenible (PERS) para el Departamento del Cauca, se establecen los lineamientos para la definición de una muestra representativa a nivel subregional que permita cumplir con los siguientes objetivos:

* Caracterizar el consumo básico de energía por uso y fuente en el sector residencial de zonas rurales representativas por medio de información primaria.
* Caracterizar y analizar las variables socioeconómicas de las viviendas rurales del departamento de cauca por medio de información obtenida a partir de encuestas primarias e información secundaria.
* Analizar la oferta de recursos energéticos en las regiones del departamento del Cauca.
* Determinar el consumo básico de subsistencia para el sector residencial.

Para fijar e implementar un número de encuestas adecuado en cada municipio y/o región de acuerdo a sus características de ruralidad y de cobertura energética, se emplearon métodos de muestreo que son ajustados de acuerdo con la experiencia de otros PERS departamentales (Nariño, Putumayo y Cundinamarca).

# Marco muestral

Para la definición del número adecuado de encuestas que permitan inferir sobre las características de uso y consumo de energía eléctrica en las viviendas rurales y/o cabeceras municipales del departamento del Cauca, se emplearon múltiples fuentes de información que permitieran tener una idea del comportamiento de los parámetros poblacionales de algunos indicadores claves sobre el tema. El siguiente esquema resume las fuentes y los temas que apoyaron para la estimación y asignación de la muestra por distintas unidades de desagregación.

Esquema . Fuentes de información para el marco muestral del PERS Cauca

Fuente: Elaboración PERS\_CAUCA

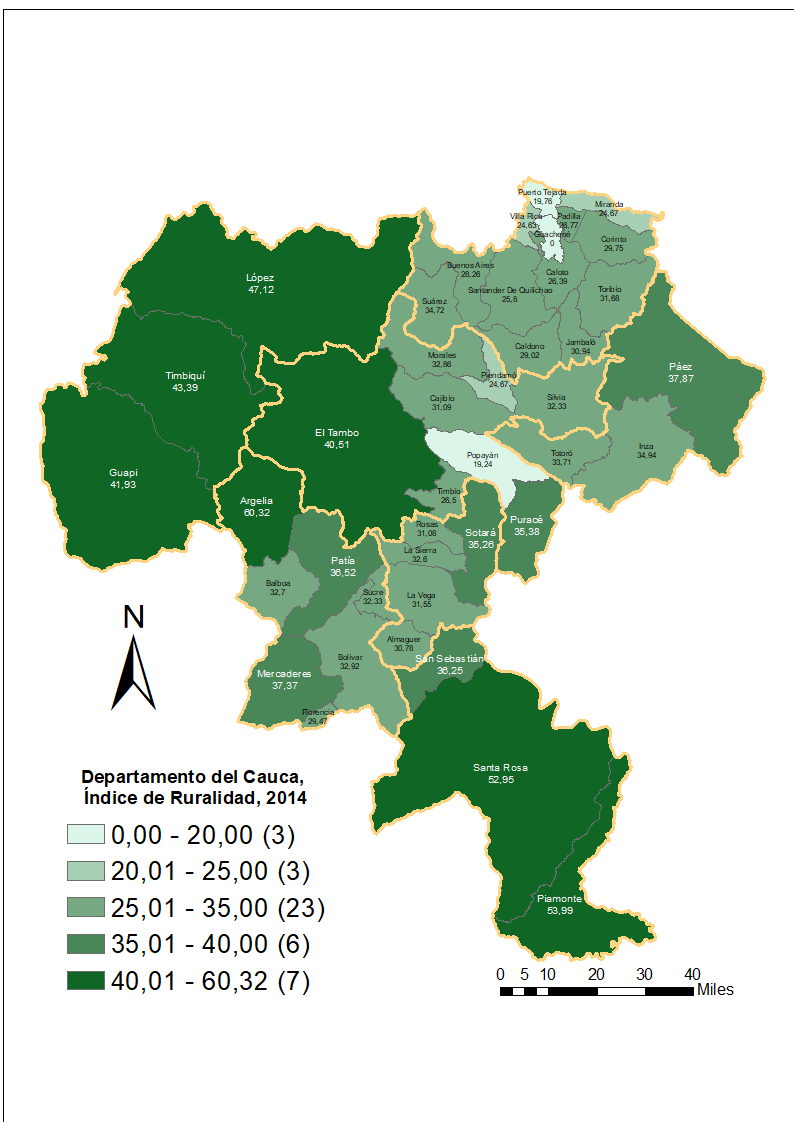
La información de las distintas fuentes de información se analizará de manera detallada en este documento, sin embargo, de manera a priori es importante precisar que se observan comportamientos no homogéneos a nivel municipal. Si se quisiera obtener indicadores representativos sobre el uso y consumo de energía eléctrica a nivel de las 42 demarcaciones territoriales que conforman el departamento del Cauca se tendría que optar por un tamaño y una asignación de encuestas suficientes para todos los municipios, este proceso implicaría incrementar los costos de operación para el trabajo de campo, así como elevar el tiempo de levantamiento y tabulación de la información.

Debido a esto, la variabilidad y otros indicadores de tendencia central se estimarán a nivel municipal, sin embargo, se dará prioridad a la asignación a partir de la conformación de siete subregiones para el departamento del Cauca, en cada subregión se identifican municipios con características peculiares en torno a la cobertura, condiciones climáticas y consumo energético, de los cuáles se recomienda enfocar el levantamiento de información primaria para tener representatividad a nivel departamental y subregional.

Por lo tanto, para tener mayor precisión en las estimaciones derivadas de la muestra y alcanzar a inferir sobre zonas o regiones con problemas de acceso a energía eléctrica en el departamento, el universo de estudio se centró en las viviendas rurales con distintas demandas de consumo energético distribuidas en los municipios y/o regiones del departamento, es relevante mencionar que esta delimitación dio prioridad a municipios con baja cobertura del servicio, pero también permitió que las regiones más pobladas del departamento tuvieran una participación notoria como el caso de la zona centro y norte.

Ahora bien, según distintas fuentes de información como el Censo Nacional Agropecuario 2016, el Departamento del Cauca es una de los territorios “más rurales” del país. A partir del índice de ruralidad se observa que en algunos municipios concentran una parte importante de viviendas con características rurales e incluso en sus cabeceras municipales, por lo tanto, en las demarcaciones con un índice de ruralidad superior al 40 por ciento, las viviendas ubicadas en cabeceras municipales se contemplan para la estimación de las encuestas.

Mapa . Departamento del Cauca, Índice de Ruralidad, 2014



Fuente: Informe PERS Cauca con base en PNUD Informe Nacional de Desarrollo Humano – 2011

Nota: No se tiene información para el municipio de Guachené (Zona Norte)

En síntesis, se tomará como universo de estudio 61.4 por ciento del total de viviendas proyectadas para el departamento del Cauca, la mayoría localizadas en los municipios de la zona centro (29.98%) y zona norte (28.41%). Sin embargo, en las otras cinco regiones el porcentaje de viviendas rurales y resto supera la participación relativa de viviendas ubicadas en cabeceras municipales (Gráfico 1). La siguiente tabla presenta el universo de viviendas por municipio y subregión.

Tabla . Universo de viviendas consideradas para el diseño muestral según municipio y/o región



Fuente: PERS Cauca con base en DANE, Censo de Población y Vivienda 2018 y Proyecciones de la Población

Gráfica . Distribución porcentual de las proyecciones de viviendas según región y contexto, 2020

Fuente: PERS Cauca con base en Censo de Población y Vivienda 2018 y Proyecciones de la Población de los municipios de Colombia.

# Diseño muestral

Para la estimación de la varianza se emplearon diversos indicadores de acceso a cobertura eléctrica o número de unidades suscritoras del servicio, en este apartado, los resultados se centran en los indicadores seleccionados para el tamaño de muestra optimo, sin embargo, en los siguientes segmentos se resumen las estimaciones para otros indicadores con distintas fuentes de información.

## Indicadores para estimar la varianza

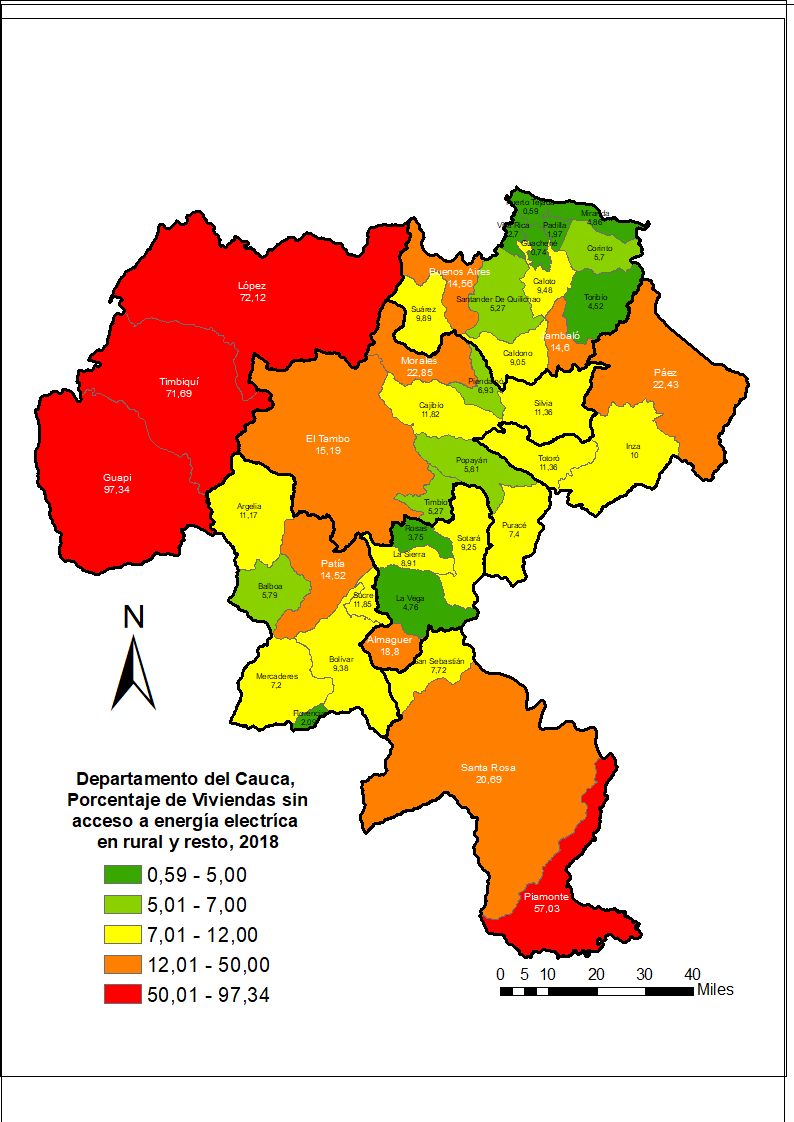
### Cobertura Rural Energía Eléctrica

Para aproximarnos a los parámetros poblacionales de las viviendas rurales se consideraron distintas fuentes de información que proporcionan la variabilidad del objeto de estudio (cobertura de energía eléctrica). Los indicadores empleados para estimar el tamaño de muestra óptimo fueron el porcentaje de viviendas sin acceso a energía eléctrica a nivel municipal a partir del Censo Nacional de Población y Vivienda 2018 y de la información del Plan Indicativo de Expansión de Cobertura de Energía Eléctrica (PIE) de 2018 de la Unidad de Planeación Minero-Energética.

En el departamento del Cauca, el acceso a energía eléctrica difiere según el contexto de residencia. El CNPV2018 estimó que 98.97% de las viviendas localizadas en cabeceras municipales y 86.35% de las localizadas en centros poblados y rurales dispersos tenían energía eléctrica en sus unidades residenciales. Sin embargo, el UPME proyectó para 2020 que la cobertura rural de energía eléctrica departamental fue de 83.6 por ciento. Si bien en el agregado departamental las estimaciones del DANE y UPME se asemejan, es importante, anotar que en algunos municipios persisten diferencias importantes.

La información del mapa 2 permite apreciar la heterogeneidad en cuanto al acceso a energía eléctrica en los municipios del departamento del Cauca con base en la información del DANE. En los tres municipios del departamento de la zona Occidente y Piamonte (Bota Caucana), la cobertura de acceso a energía eléctrica es limitada; en el caso de Guapi 97.34% de viviendas no tenían disponible este servicio. Ahora bien, en algunas regiones sobresalen algunos municipios como el caso de Morales (Centro), Páez (Oriente), Santa Rosa (Bota Caucana), donde aproximadamente una de cada cinco viviendas no tenía acceso a energía eléctrica. En contraposición, Popayán y Timbío, Piendamó en la Zona Centro; Bolívar en la Sur; Santander de Quilichao y Corinto en la zona norte, destacaron por bajos porcentajes de viviendas rurales sin acceso al servicio.

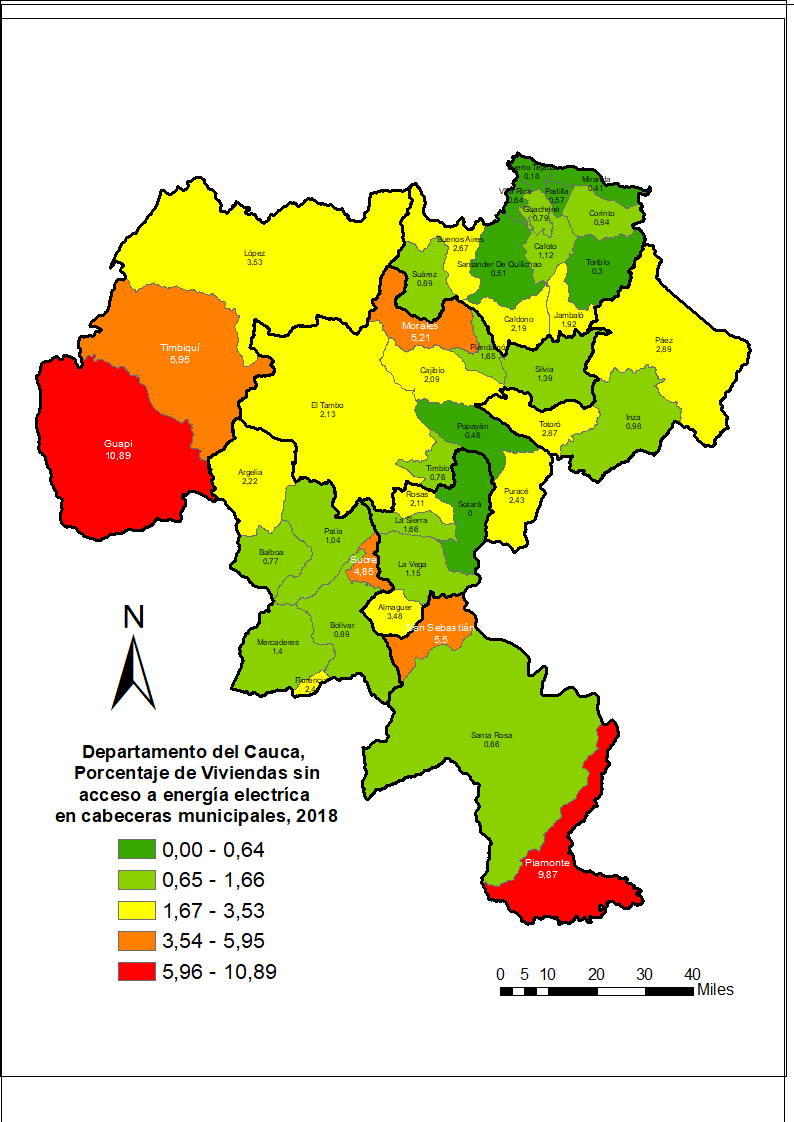
Mapa 2. Departamento del Cauca, Porcentaje de Viviendas rurales y resto sin acceso a energía eléctrica, 2018



Fuente: PERS Cauca con base en Censo Nacional de Población y Vivienda 2018

Por su parte, el mapa 3 muestra el porcentaje de viviendas sin acceso a energía eléctrica en cabeceras municipales. En general, si se compara con los sectores rurales y centros poblados en todos los municipios, el acceso a energía eléctrica en sus cabeceras municipales es alto, sin embargo, sobresalen Guapi y Piamonte con porcentajes cercanos al 10%, situación que justifica la inclusión de las unidades ubicadas en las cabeceras municipales de acuerdo con el índice de ruralidad al momento de distribuir el tamaño óptimo de encuestas.

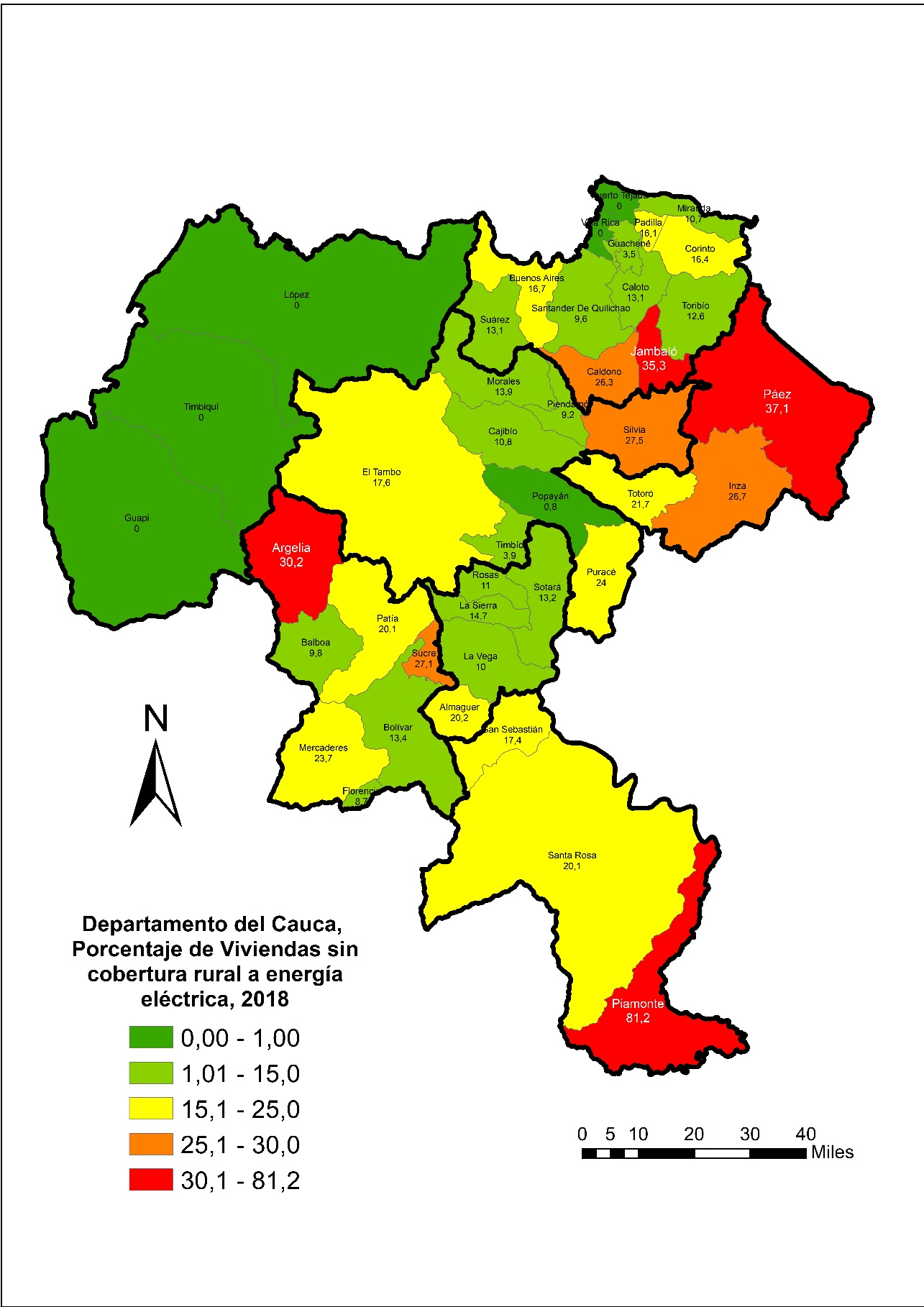
Mapa . Departamento del Cauca, Porcentaje de Viviendas ubicadas en cabeceras municipales sin acceso a energía eléctrica, 2018



Fuente: PERS Cauca con base en Censo Nacional de Población y Vivienda 2018

Ahora bien, al analizar el indicador de porcentaje de viviendas rurales sin cobertura de energía eléctrica según la información del PIEC se observan algunos municipios con baja cobertura como es el caso de Piamonte (81.2%), Páez (37.1%), Jambaló (35.3%) y Argelia (30.2%). En sentido contrario resaltan con cobertura total, los tres municipios de la zona occidente del departamento, estos resultados, si bien no son directamente comparables con los mapas 2 y 3 muestran ordenamientos disimiles en cuanto al acceso a cobertura de energía.

Mapa 4. Porcentaje de viviendas rurales sin cobertura de energía eléctrica, 2020

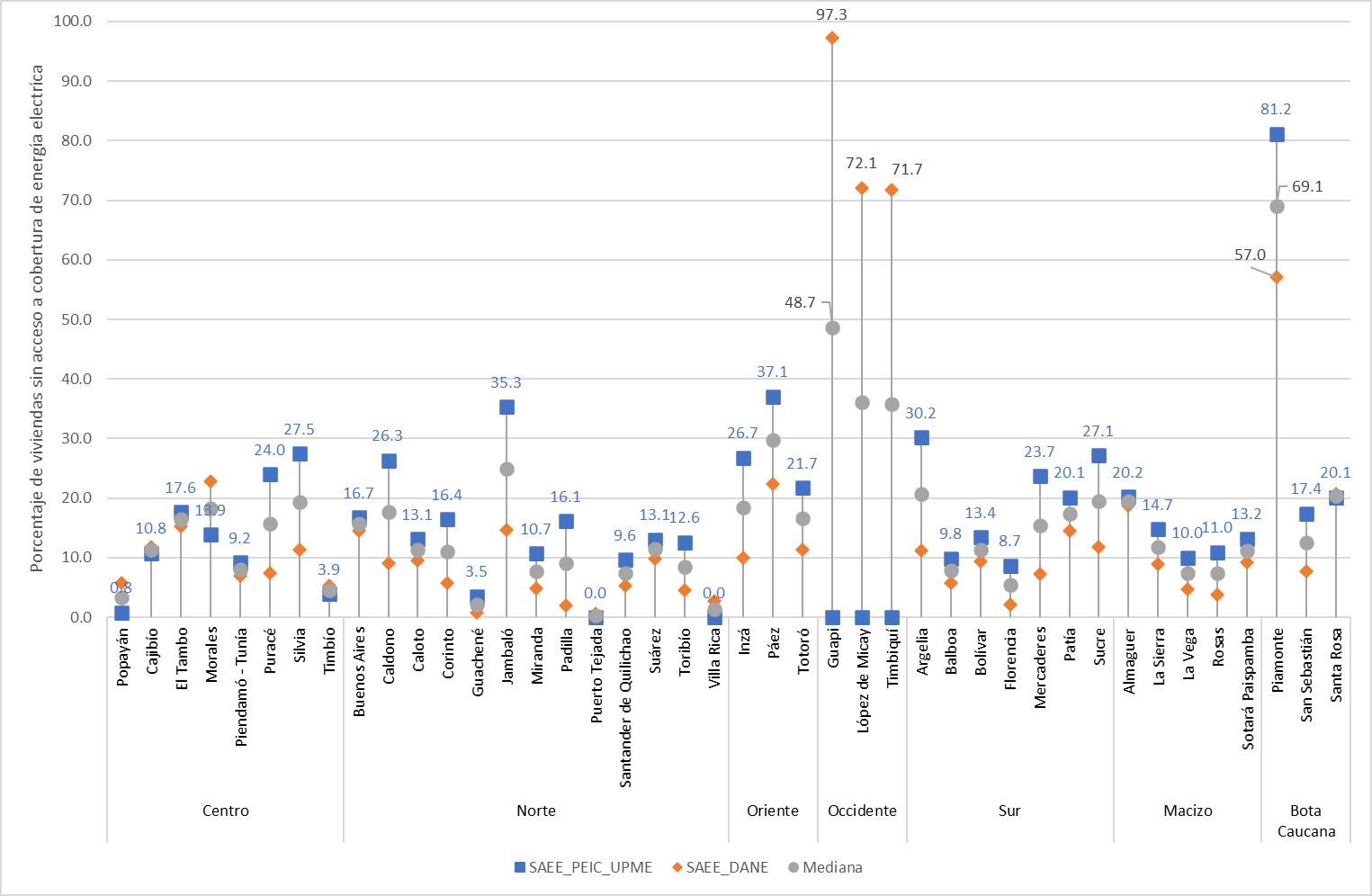


Fuente: PERS CAUCA con base en Plan Indicativo de Expansión de Cobertura de Energía Eléctrica (PIE) de 2018 de la Unidad de Planeación Minero-Energética.

Por tanto, con el objetivo de contemplar la heterogeneidad existente en este indicador a partir de ambas fuentes de información, se empleó la mediana entre ambas observaciones para la estimación de las varianzas apropiadas que permitan aplicar la estimación del tamaño muestral (Gráfica 2). Se considera importante la mediana para reducir la dispersión en el indicador de interés, en especial, para los municipios de la subregión occidente y otros particulares como Piamonte, .Jambaló y Páez.

Dicha heterogeneidad y la ubicación de algunos nodos territoriales con acceso limitado al servicio, obliga ajustar los tamaños de muestra de acuerdo con la heterogeneidad observada y la participación de viviendas rurales sin acceso al servicio sobre el total de cada región. Además, en algunos casos, como en el sector comercial e institucional conviene fijar un número de encuestas que permitan de algún modo su representatividad y lo más importante, obtener parámetros de consumo por fuente energética y una aproximación a su distribución y nivel de consumo en cada proceso de la actividad...

Gráfica . Porcentaje de viviendas rurales sin cobertura de energía eléctrica según distintas fuentes



Fuente: PERS CAUCA con base en Plan Indicativo de Expansión de Cobertura de Energía Eléctrica (PIE) de 2018 de la Unidad de Planeación Minero-Energética y Censo Nacional de Población y Vivienda 2018.

### Suscripción de unidades comerciales e industriales por región

Por otro lado, para las unidades comerciales e industriales con base en el Sistema Único de Información de Servicios Públicos (SUI), se obtuvo una aproximación de la distribución de estas unidades por municipios. En 2021 estaban suscritas en el SUI 15.815 unidades comerciales y 2.141 industrias, la mayoría concentradas en el norte y la capital del departamento del Cauca. Con está información se realizó una asignación proporcional de las unidades industriales y comerciales por subregión considerando el 1.5 por ciento de suscritores sobre el total del departamento.

Gráfica 3. Departamento del Cauca, Distribución de unidades comerciales e industriales según región de localización, 2021

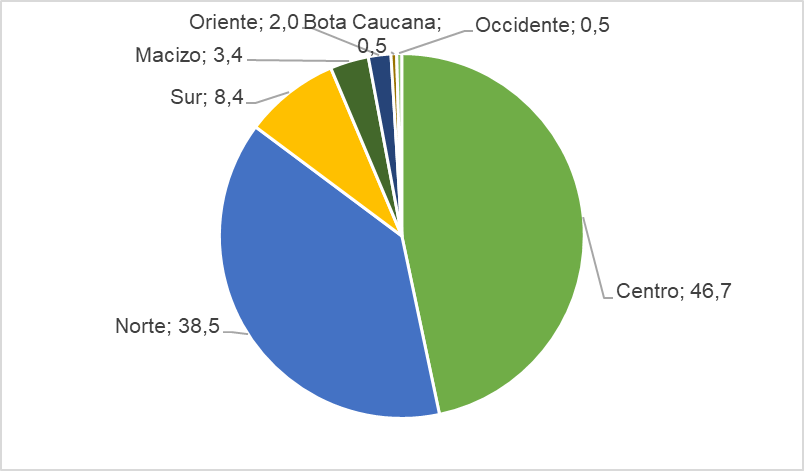
Gráfico, Gráfico de rectángulos

Descripción generada automáticamente

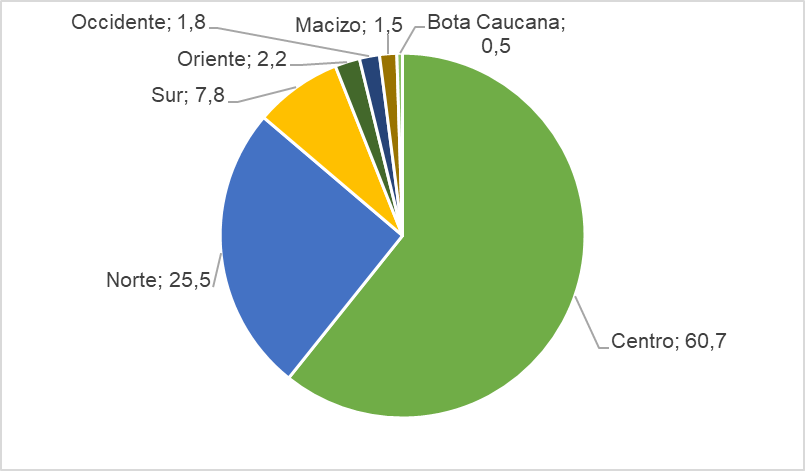
Fuente: Elaboración PERS-CAUCA con base en Sistema Único de Información de Servicios Públicos

Gráfica . Departamento del Cauca, Distribución de unidades comerciales e industriales según región de localización, 2021

**Comerciales**



**Industriales**



Fuente: Elaboración PERS-CAUCA con base en Sistema Único de Información de Servicios Públicos

## Definición del tamaño de muestra

De acuerdo con el indicador clave definido en el apartado anterior, se empleó un método de muestreo que permitiera contemplar la heterogeneidad en la falta de acceso a la energía eléctrica en los 42 municipios del departamento del Cauca, es importante destacar a su vez que la agregación y estimaciones finales se realizarán en términos de las regiones, si bien se utilizan los municipios como insumo para la obtención de la varianza y el planteamiento del tamaño de muestra, todos los resultados serán plasmados a nivel de las regiones. A continuación, se definen algunos elementos metodológicos del proceso.

### Tipo de muestreo

Muestra probabilística, estratificada, de conglomerados, multietápica:

* + Probabilística: Cada unidad de la población objetivo tiene una probabilidad conocida de selección mayor a 0.
  + Estratificada: Con el fin de ganar precisión de la muestra, se propone realizar una estratificación de acuerdo con las siete zonas geográficas: centro, norte, oriente, pacifico, sur, macizo y piedemonte amazónica, y de acuerdo con las particularidades en la distribución de viviendas en las áreas dispersas y en la cabecera municipal (para municipios con índice de ruralidad mayor al 40 %), para un total de 7 estratos muestrales.
  + Conglomerados: Unidad final de muestreo compuesta por el segmento (definir cuántas viviendas por cada uno, por ejemplo, la gran encuesta de hogares utiliza 10).
  + Multietápica
    - Unidad Primaria de Muestreo (UPM): para las cabeceras son grupos contiguos de acuerdo con la división política (ej: manzanas) del mismo municipio, que contiene mínimo 12 segmentos con un promedio de 10 viviendas contiguas (5 a 14 viviendas). En la zona rural, la UPM está conformada por una sección o sector censal.
    - Unidad Secundaria de Muestreo (USM): el segundo paso es seleccionar dentro de cada UPM, el número necesario de segmentos requeridos, los cuales corresponden a la USM.
    - Unidad terciaria de Muestreo (UTM): finalmente, la UTM corresponde a las viviendas que conforman cada USM seleccionada para conformar la muestra.

### Tamaño de la muestra

Teniendo en cuenta el muestreo planteado multietápico, el cálculo del tamaño de muestra se realizará por medio de la siguiente expresión:

Donde la varianza es calculada por medio de la expresión

n : es el tamaño de la muestra.

EM: Máximo margen de error absoluto aceptable.

Z: Nivel de confianza deseado.

: Peso de cada estrato en el muestreo

Valor calculado del diseño muestral acorde a la varianza. Se estiman a priori de la información disponible, para el caso particular serán los promedios y varianzas de la población objeto de estudio.

: Se hace que la representación del número de viviendas al interior de cada lado de manzana en cada estrato sea proporcional al número de ellas en el lado analizado (autoponderación) para ello suponemos constante la fracción de muestreo en la segunda etapa

Las proporciones y. resaltan los pesos correspondientes a la asignación por sub región y municipio respectivamente, los pesos son calculados de la información a priori con la se cuenta. Sin embargo, es de destacar que las estimaciones fueron tomadas como un estimador de la mediana entre la fuente del PERS y el DANE. La varianza general fue calculada de dichas fuentes de información por lo que se ajustó el tamaño de muestra a un total de 2091. Para la obtención de la muestra se deben considerar la desagregación de la población por regiones y dentro de las regiones los municipios, en ese sentido se tienen consideradaradas cada una de las etapas, los valores de refencia que se presentan en las formulas anteriores permiten obtener una estimación de la muestra teniendo en cuenta el diseño de muestreo, presentado en el literal anterior. Dentro de la estimación se consideran las varianzas intra region intra municipio y la contrucción de la varianza total, también se consideran los pesos de cada municipio en la región y de la región dentro del departamento, esto con la finalidad de garantizar la representatividad de cada una de las regiones.

Teniendo en cuenta los valores, fijando un nivel de confianza en el 10% y un margen de error permisible máximo del 10%, se obtiene un tamaño muestral de 2.526 encuestas; de las cuales 150 se asignarán para unidades residenciales y/o comerciales localizadas en puntos clave del departamento.

Es importante precisar que el ejercicio de estimación de la muestra, se realizó con varios indicadores que contemplan diferentes unidades de observación y fuentes de información; sin embargo, dado el interés de investigación, el tamaño muestral se asignó con base en la variabilidad del porcentaje de viviendas sin acceso a energía eléctrica en zonas rurales y centros poblados.

Tabla 2. Estimación del tamaño de muestra según diferentes fuentes de información

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Indicador y fuente de Información** | **n** | **Promedio** | **Varianza** | **Significancia** | **Nivel de Confianza deseado** |
| Porcentaje de viviendas sin acceso a energía eléctrica en zonas rurales y centros poblados **(DANE, Censo Nacional de Población y Vivienda 2018)** | **2526** | 15,44 | 426,43 | 0,10 | 1,96 |
| Componente de Servicios del Índice de Necesidades Básicas Insatisfechas **(Censo Nacional de Población y Vivienda 2018)** | **2561** | 7,67 | 107,38 | 0,10 | 1,96 |
| Suscriptores de un servicio de energía eléctrica **(Sistema Único de Información del sector de servicios públicos domiciliarios del país)** | **2635** | 13,59 | 559,27 | 0,10 | 1,96 |
| Porcentaje de viviendas sin Acceso a Energía Eléctrica **(DANE, Gran Encuesta Integrada de Hogares, 2018)** | **2477** | 7,35 | 56,90 | 0,10 | 1,96 |
| Porcentaje de viviendas sin acceso a Energía Eléctrica **(SISBEN, Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales, 2020)** | **3263** | 5,88 | 104,91 | 0,10 | 1,96 |
| Municipios con índices de ruralidad mayor a 40% y con porcentajes de cobertura a Energía eléctrica según fuente PERS | **2091** | **84,50** | **16,5581** | **0,10** | **1,96** |

Fuente: Elaboración PERS-CAUCA

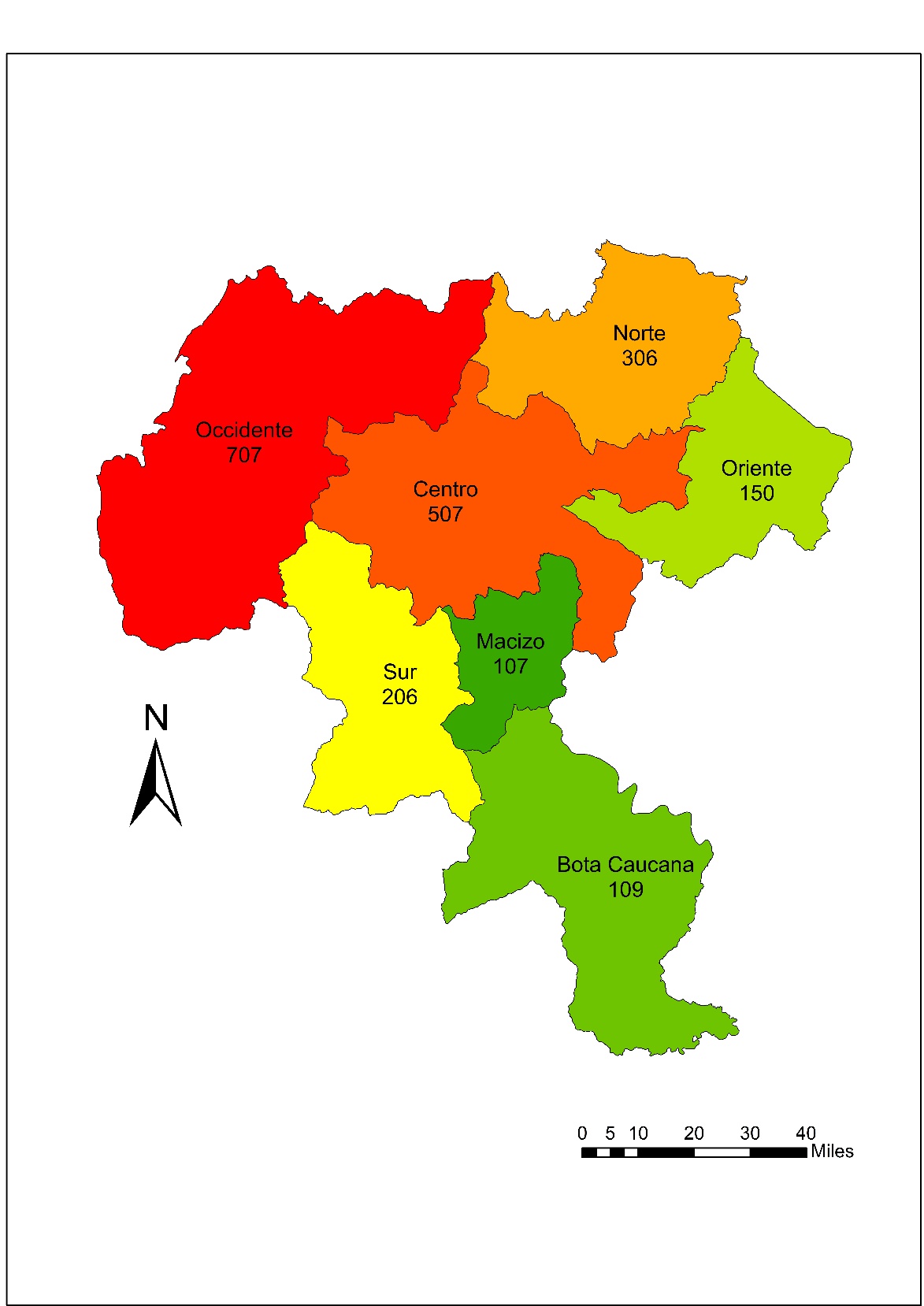
## Distribución de encuestas por cada Municipio o región

Para realizar la asignación de la muestra, se tomará como base la estimación de la proporción de viviendas sin acceso a energía por cada región, de esta manera se obtendrá los pesos asociados para cada una de las regiones del departamento del cauca. En cuanto a las unidades de los conglomerados, se planteará una estimación de la proporción de viviendas sin acceso a energía eléctrica en las zonas rurales de los municipios.

Para asignar la muestra, se hará uso de las probabilidades de Bayes, por lo tanto, se multiplicarán las proporciones y el tamaño de muestra, en este sentido se obtendrá la muestra específica para las regiones con baja cobertura al servicio y una participación considerable de viviendas rurales.

Los cálculos de las muestras se realizan por medio de la librería *Sampling* del software estadístico R-project. Una vez se tienen los pesos asociados la asignación proporcional de las estimaciones se realiza por medio de tablas de Excel.

Mapa 5. Departamento del Cauca, distribución de las encuestas por región



Fuente: Elaboración PERS-CAUCA

Ahora bien, de acuerdo con las estimaciones realizadas, es importante que en algunos municipios se logre obtener información para un número determinado de encuestas que permita inferir sobre el comportamiento del consumo energético al interior de cada región. La siguiente tabla muestra la distribución municipal con la referencia de los municipios claves.

Tabla . Distribución de encuestas según Unidades residenciales por región

****

Fuente: Elaboración PERS-CAUCA

En la Tabla se muestra la desagregación del factor de expansión a nivel regional, es de notar que al realizar los cálculos la suma total de los productos comprende el 99.87% de la población total, esto se debe al margen de error atribuible a las fuentes de información. El calculo de las probabilidades a nivel regional no se pueden realizar de la misma forma en las que se realizan a nivel municipal.

Tabla . Distribución de encuestas según Unidades comerciales e industriales por región



Fuente: Elaboración PERS-CAUCA

Con la finalidad de garantizar la representatividad de las regiones, es de suma importancia que se garantice que en aquellos lugares en donde se asignó más de 50 encuestas, estas se lleven a cabo, entendiendo que en las condiciones logísticas se podrán redistribuir de manera aleatoria aquellas asignaciones estadísticamente no significativas.

El Factor de expansión corresponde a la asignación de las probabilidades de elegir cada elemento de la población objeto de estudio según las características asociadas al objetivo del estudio, en este caso siempre se realizará como 1/Pr(Viviendas|No Energía). Hay que tener en cuenta que para el cálculo se utiliza el teorema de Bayes el cual estará anidado a la población objeto de estudio y a la asignación muestral.

Tabla . Distribución de encuestas según Unidades residenciales por municipios



Fuente: Elaboración PERS-CAUCA

Tabla . Distribución de encuestas según Unidades comerciales e industriales por municipios



Fuente: Elaboración PERS-CAUCA