

**DESARROLLO Y COMPARACIÓN DE ALGORITMOS PARA LA
CLASIFICACIÓN Y MODELACIÓN DE LA DEMANDA
ELÉCTRICA DE LOS MUNICIPIOS DE NARIÑO**



**PRESENTADO POR:
GABRIEL E. NARVAEZ MORALES
CAMILO ERNESTO PAZMIÑO**

**UNIVERSIDAD DE NARIÑO
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE ELECTRÓNICA
SAN JUAN DE PASTO
2015**

**DESARROLLO Y COMPARACIÓN DE ALGORITMOS PARA LA
CLASIFICACIÓN Y MODELACIÓN DE LA DEMANDA
ELÉCTRICA DE LOS MUNICIPIOS DE NARIÑO**



**PRESENTADO POR:
GABRIEL E. NARVAEZ
MORALES
CAMILO ERNESTO PAZMIÑO**

**TRABAJO DE GRADO DE GRADO PARA OPTAR POR EL TÍTULO
DE INGENIERO ELECTRÓNICO**

**ASESOR:
Mg. WILSON OLMEDO ACHICANOY MARTÍNEZ**

**UNIVERSIDAD DE NARIÑO
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE ELECTRÓNICA
SAN JUAN DE PASTO
2015**

NOTA DE RESPONSABILIDAD

“La Universidad de Nariño no se hace responsable por las opiniones o resultados obtenidos en el presente trabajo y para su publicación priman las normas sobre el derecho de autor”.

Acuerdo 1. Artículo 324. Octubre 11 de 1966, emanado del honorable Consejo Directivo de la Universidad de Nariño.

NOTA DE ACEPTACIÓN

Firma del Presidente del Jurado

Firma Jurado 1

Firma Jurado 2

San Juan de Pasto, 4 de junio de 2015

AGRADECIMIENTOS

Agradecemos a:

Primero, nos gustaría agradecer sinceramente a nuestro asesor de trabajo de grado, Mg Wilson Olmedo Achicanoy Martínez, su esfuerzo, dedicación y buenos consejos han sido fundamentales para nuestra formación como investigadores.

Agradecemos también a la planta docente y al grupo de investigación GIIEE cuyo apoyo ha sido imprescindible para forjar las bases de nuestra vida profesional.

Finalmente agradecemos a nuestras familias por su apoyo incondicional, en especial durante estos duros años de formación profesional.

CONTENIDO

| | |
|---|----|
| INTRODUCCIÓN | 13 |
| 1. CARACTERÍSTICAS DEL DEPARTAMENTO DE NARIÑO | 16 |
| 2. CLASIFICACIÓN | 20 |
| 2.1. Variables..... | 20 |
| 2.2. Pre-procesamiento de la información | 23 |
| 2.3. Métodos de clasificación | 25 |
| 2.3.1. Método de agrupamiento por Mixturas de Gaussianas | 27 |
| 2.3.2. <i>K-Nearest Neighbors</i> (KNN) | 27 |
| 2.3.3. <i>K-Means Clustering</i> | 28 |
| 2.3.4. Redes Neuronales Artificiales RNA..... | 28 |
| 2.4. Pruebas..... | 29 |
| 2.4.1. Método de agrupamiento por Mixturas de Gaussianas | 30 |
| 2.4.2. <i>K-Means Clustering</i> | 33 |
| 2.4.3. Redes Neuronales Artificiales RNA..... | 35 |
| 3. MODELAMIENTO | 43 |
| 3.1. Modelamiento individual por municipio..... | 44 |
| 3.1.1. Suavizado y descomposición | 44 |
| 3.1.2. Modelo | 49 |
| 3.1.3. Pronóstico..... | 56 |
| 3.2. Modelamiento por Clusters..... | 58 |
| 3.2.1. Relación entre población y suscriptores residenciales..... | 59 |
| 3.2.2. Relación entre suscriptores y consumo residencial..... | 61 |
| 4. DISCUSIÓN Y CONCLUSIONES | 65 |
| 5. RECOMENDACIONES Y TRABAJO FUTURO | 67 |
| REFERENCIAS | 68 |
| ANEXO 1. Características de los municipios agrupados en 14 clusters con el método de mixtura de Gaussianas..... | 70 |
| ANEXO 2. Características de los municipios agrupados en 8 clusters con el método k-means..... | 71 |
| ANEXO 3. Características de los municipios agrupados en 9 clusters con el método de RNA...72 | |
| ANEXO 4. Características de los municipios de la Región 1 de la clasificación con el método de RNA..... | 74 |
| ANEXO 5. Características de los municipios de la Región 2 de la clasificación con el método de RNA..... | 75 |
| ANEXO 6. Características de los municipios de la Región 3 de la clasificación con el método de RNA..... | 75 |
| ANEXO 7. Características de los municipios de la Región 4 de la clasificación con el método de RNA..... | 76 |
| ANEXO 8..... | 77 |
| INTRODUCCIÓN | 80 |
| 1. Módulo de Clasificación | 81 |
| 1.1. Barra de Menú..... | 82 |
| 1.1.1. Módulo de Clasificación | 82 |
| 1.1.2. Módulo de Predicción | 83 |

| | | |
|--------|---|----|
| 1.1.3. | Análisis de Registro Histórico..... | 83 |
| 1.2. | Lista de Selección de Municipio..... | 84 |
| 1.3. | Total Habitantes | 85 |
| 1.4. | Total Viviendas | 85 |
| 1.5. | Producción Agrícola..... | 85 |
| 1.6. | Producción Pecuaria..... | 86 |
| 1.7. | Cuadro de Mapas..... | 86 |
| 1.8. | Botón Calcular | 87 |
| 1.9. | Clasificación Exitosa..... | 88 |
| 1.10. | Botón Limpiar | 89 |
| 2. | Módulo de Predicción | 90 |
| 3. | Módulo Análisis de Registro Histórico | 94 |
| 3.1. | Tratamiento de la información..... | 94 |
| 4. | Bibliografía | 98 |

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1. División geográfica de Nariño | 16 |
| Figura 2. Departamento de Nariño por subregiones. | 19 |
| Figura 3. Clasificación de los municipios de Nariño mediante <i>Gaussian Mixture Clustering</i> | 31 |
| Figura 4. <i>Silhouette Graph</i> . 14 Clusters. | 33 |
| Figura 5. Clasificación de los municipios de Nariño mediante <i>k-means Clustering</i> | 34 |
| Figura 6. <i>Silhouette Graph</i> . 8 Clusters. | 35 |
| Figura 7. <i>Self-organized map</i> SOM..... | 36 |
| Figura 8. Clasificación de los municipios de la ZI según su producción Agropecuaria. | 38 |
| Figura 9. Clasificación de la Región 1 del <i>clustering</i> primario. | 39 |
| Figura 10. Clasificación de la Región 2 del <i>clustering</i> primario. | 40 |
| Figura 11. Clasificación de la Región 3 del <i>clustering</i> primario. | 41 |
| Figura 12. Clasificación de la Región 4 del <i>clustering</i> primario. | 43 |
| Figura 13. Filtros aplicados a series de consumo. (a) Consumo industrial de La Unión-Filtro MAF-M. (b) Consumo residencial de Pasto-Filtro MF..... | 46 |
| Figura 14. Descomposición multiplicativa de la serie de consumo industrial de La Unión. (a) Tendencia y <i>Seasonality</i> . (c) Irregular. | 48 |
| Figura 15. ACF, PACF, x_t vs. x_{t-1} y x_t vs. x_{t-2} de la señal diferenciada del consumo residencial de La Unión. (a) ACF y PACF. (b) x_t vs. x_{t-1} y x_t vs. x_{t-2} | 51 |
| Figura 16. Parámetros del modelo ARIMA (2, 1, 1). | 52 |
| Figura 17. Parámetros del modelo ARIMA (2, 1, 2) x (12, 12). | 53 |
| Figura 18. Parámetros del modelo ARIMA (1, 1, 2)/GARCH (2, 1)/ARCH (2, 1)..... | 54 |
| Figura 19. Pronóstico a 120 meses del consumo industrial de La Unión. (a) ARIMA (2, 1, 1). (b) ARIMA (2, 1, 2) x (12, 12). (c) ARIMA (1, 1, 2)/GARCH (2, 1)/ARCH (2, 1)..... | 58 |
| Figura 20. Ecuación lineal entre población y suscriptores residenciales. | 60 |
| Figura 21. Coeficientes de la ecuación lineal entre población y suscriptores residenciales..... | 60 |
| Figura 22. Comportamiento de los suscriptores residenciales en el tiempo. | 61 |
| Figura 23. Ecuación cuadrática entre suscriptores y consumo residencial. | 62 |
| Figura 24. Coeficientes de la ecuación cuadrática entre suscriptores y consumo residencial..... | 63 |
| Figura 25. Comportamiento del consumo energético en el tiempo. | 63 |
| Figura 26. Comparación entre la proyección con un modelo ARIMA (2, 1, 1) y la proyección obtenida por suscriptores y población. | 64 |

LISTA DE TABLAS

| | |
|--|----|
| Tabla 1. Municipios de Nariño..... | 17 |
| Tabla 2. Principales productos agrícolas de Nariño. | 24 |
| Tabla 3. Características de los municipios en el <i>Cluster 7</i> | 32 |
| Tabla 4. Características de los municipios de los <i>clusters 7 y 8</i> | 35 |
| Tabla 5. Patrones de ACF y PACF. | 50 |
| Tabla 6. Resultados de las pruebas Ljung-Box Q y Engle del modelo ARIMA (2, 1, 1)..... | 53 |
| Tabla 7. Resultados de las pruebas Ljung-Box Q y Engle del modelo ARIMA (2, 1, 2) x (12, 12). | 55 |
| Tabla 8. Resultados de las pruebas Ljung-Box Q y Engle del modelo ARIMA (1, 1, 2) /GARCH (2, 1)/ARCH (2, 1). | 56 |
| Tabla 9. Pruebas AIC/BIC. | 56 |

LISTA DE ANEXOS

| | |
|--|----|
| ANEXO 1. Características de los municipios agrupados en 14 clusters con el método de mixtura de Gaussianas..... | 70 |
| ANEXO 2. Características de los municipios agrupados en 8 clusters con el método k-means..... | 71 |
| ANEXO 3. Características de los municipios agrupados en 9 clusters con el método de Redes Neuronales Artificiales | 72 |
| ANEXO 4. Características de los municipios de la Región 1 de la clasificación con el método de Redes Neuronales Artificiales | 74 |
| ANEXO 5. Características de los municipios de la Región 2 de la clasificación con el método de Redes Neuronales Artificiales | 75 |
| ANEXO 6. Características de los municipios de la Región 3 de la clasificación con el método de Redes Neuronales Artificiales | 75 |
| ANEXO 7. Características de los municipios de la Región 4 de la clasificación con el método de Redes Neuronales Artificiales | 76 |
| ANEXO 8. MANUAL DE USUARIO..... | 77 |

RESUMEN

La clasificación y modelación del consumo de energía eléctrica para una región en particular ofrece la posibilidad de predecir su demanda y asegurar el suministro necesario para el consumo industrial, comercial y residencial de dicha zona. También ofrece la posibilidad de realizar un análisis que permita llevar soluciones de energía eléctrica a los lugares donde ésta, por diferentes razones, no se presta.

Para realizar una adecuada clasificación de los municipios del departamento de Nariño según la demanda de energía eléctrica, se compararon varios algoritmos de agrupamiento automático estudiados en el área del aprendizaje de máquinas. Los resultados de cada algoritmo fueron analizados para seleccionar el más adecuado para los objetivos de esta investigación. La herramienta de *clustering* basada en el entrenamiento de Redes Neuronales Artificiales (RNA) se seleccionó como la de mejor desempeño y mayores funciones para alcanzar los resultados planteados.

Por el tipo de información con que se cuenta, la clasificación se realiza en dos etapas: la primera, utiliza la información que se organizó en variables categóricas. Este proceso agrupa los municipios de la zona interconectada de Nariño en 4 *clusters*, los cuales se tratan como regiones en este trabajo. Posteriormente, dentro de cada región, se realiza una clasificación interna que separa en diferente número de *clusters* cada una de ellas.

La modelación del consumo energético para los *clusters* se realizó con modelos univariados ARIMA, en donde se utilizaron las series de registro histórico de consumo para su modelamiento y predicción. La base de datos de las series de consumo energético tienen periodicidad mensual, y la información pertenece al período comprendido entre enero 2003 a septiembre de 2014. En el proceso de la información se usaron filtros, y técnicas descomposición de tipo aditiva y multiplicativa. También se hizo un modelamiento usando las variables de población y suscriptores residenciales, con el que se logró una predicción a 7 años del consumo y se comparó los resultados obtenidos con ambos métodos.

ABSTRACT

Classification and modeling of electric energy consumption for a particular region offer the possibility of predict the demand, and ensure the supply needed for the industry, commerce and residential consumption. Also, it gives the possibility of make an analysis to allow bringing solutions of electric energy to places where this services is not provided.

For an appropriate classification of Nariño Department municipalities based on the electric energy demand, many algorithms of clustering, studied in machine learning, were analyzed. Results of each algorithm were compared searching for the right selection for the objectives of this investigation. Clustering tool based on neural networks training, was selected as the one with the best performance to achieve the planned results.

Because the kind of information we have the classification is done in two phases: Firstly, we use the information of categorical variables. This process gives 4 clusters, which are denominated as regions in this paper. Thus, inside every region an internal classification is made to separate every region in different number of clusters.

Modeling of electricity consumption for the clusters is done with univariate ARIMA models, where the time series of historical consumption were used for modeling and prediction. The time series of the electrical demand for Nariño municipalities have monthly periodicity, the initial time is January/2003 and the final time is September/2014, in addition, modeling of electrical demand is done with filters and decomposition techniques for signal processing. Were used too population and residential subscribers time series, with this model, a seven years prediction of electric energy consumption is achieved, and the results obtained with both models are compared.

INTRODUCCIÓN

La creciente dependencia a los aparatos electrónicos en la vida cotidiana, el trabajo y las actividades diarias, hacen necesario que se garantice el suministro de energía con calidad y de forma constante. Surge entonces, la necesidad de clasificar y modelar la demanda eléctrica para poder prever el consumo futuro y ofrecer un servicio adecuado para los usuarios y mejorar la rentabilidad para los comercializadores, distribuidores y transmisores de la misma.

La proyección de la energía eléctrica depende de predicciones basadas en consumos anteriores que no son muy exactas puesto que no tienen en cuenta un gran número de variables: energéticas, sociales, económicas, entre otras. Estas influyen en el potencial de consumo. La clasificación de la demanda de energía eléctrica y la caracterización del consumo por regiones tiene importancia en el diseño de estrategias para la implementación de sistemas que satisfagan la demanda de las diferentes zonas donde éstas se ejecuten.

El nivel y precisión de planificación que se proponga va a depender de los datos que se utilicen en la toma de decisiones. Frente a la disponibilidad de datos más precisos, los resultados serán más fiables. Sin embargo, bien conviene acotar que uno de los mayores problemas que se plantea en la resolución de cualquier ejercicio de electrificación rural en países en vías de desarrollo se debe a la carencia de datos o a la disponibilidad de los mismos. Entendiendo esto, la metodología puede ser una herramienta bastante adecuada para efectuar un análisis hasta un nivel de pre- factibilidad o inclusive hasta de factibilidad (Rojas Zerpa, 2012).

Una gran cantidad de modelos han sido estudiados en la comunidad internacional y se han aplicado a diferentes series de demanda de acuerdo a la necesidad de cada país. Estos métodos se pueden clasificar a grandes rasgos en tres grupos: El primer grupo se basa en modelos estadísticos (regresión, series de tiempo y/o econométricos), los cuales modelan la serie de demanda en función de sus valores pasados (modelos univariados) o en función de algunas variables exógenas como el tiempo o factores socio-económicos, como es el caso de la estimación por variables económicas (Benavente, Galetovic, Sanhueza, & Serra, 2005), (Egelioglu, Mohamad, & Guven, 2001).

El segundo grupo reúne los modelos basados en inteligencia artificial. Esta técnica ha sido usada en los últimos años y se han elaborado estudios donde se comparan los resultados con métodos tradicionales. Dentro de las técnicas usadas en este grupo se encuentran las redes neuronales artificiales, modelos neuro –

difusos y modelos híbridos. El tercer grupo son modelos empíricos que dependen del juicio y la intuición humana (Rueda, Velásquez Henao, & Franco Cardona, 2011), (Medina Hurtado & García Aguado, 2005).

En Colombia, la UPME¹, que es la entidad encargada de realizar la planeación y análisis del sistema energético, elabora un pronóstico de la demanda de energía mensual y anual del país. Sin embargo, este modelo, aborda de manera muy general las necesidades propias de cada región por lo cual quedan grandes interrogantes en cuanto a la forma de aplicación de las metodologías y los resultados obtenidos. Adicionalmente, en la literatura científica hay muy poca información relevante, con lo cual no se puede tener conclusiones contundentes que permitan determinar cuál es el mejor modelo a seguir de acuerdo con el pronóstico requerido (Unidad de Planeación Minero Energética, Subdirección de Planeación Energética - Grupo de Demanda Energética, 2013).

En el país existen zonas que carecen de interconexión con las redes eléctricas nacionales lo cual implica una problemática social y económica pues dificulta el acceso a servicios de educación, salud, comunicación, factores productivos a los que este servicio permite suscribirse. Dentro de las comunidades afectadas por esta problemática, el Departamento de Nariño cuenta con una vasta zona, que debido a diferentes tipos de dificultades: falta de carreteras, presencia de grupos armados ilegales, desidia del estado, entre otras, se ve imposibilitada a superar dicha situación. Actualmente se desarrollan múltiples iniciativas encaminadas a solucionar esa problemática, puesto que esto implica, oportunidades de educación y de trabajo para un gran número de familias (Benavente, Galetovic, Sanhueza, & Serra, 2005).

Actualmente se está desarrollando el proyecto Plan de Energización Rural Sostenible para el Departamento de Nariño (PERS-Nariño)², el cual busca evaluar los lineamientos de política energética e identificar, formular y estructurar proyectos integrales y sostenibles en el corto, mediano y largo plazo como parte de la planeación en el país, así como identificar la problemática particular de las zonas rurales del departamento. Adicionalmente se está llevando a cabo el proyecto denominado “Análisis de Oportunidades Energéticas con Fuentes Alternativas en el Departamento de Nariño –ALTERNAR–”³, este proyecto es la primera fase de un programa a largo plazo que pretende contribuir a la definición de nuevos modelos

¹ <http://www1.upme.gov.co/>

² <http://sipersn.udenar.edu.co:90/sipersn/>

³ <http://190.254.4.127:90/alternar/>

de gestión energética para mejorar la cobertura en zonas apartadas y/o no interconectadas de Nariño. El fin de esta iniciativa es proponer un esquema de aprovechamiento integral y sostenible de los recursos locales para la generación de energía alternativa mediante la investigación en tecnologías de redes inteligentes y microrredes aplicadas a zonas rurales, abordando los retos del cambio climático, el uso racional de la energía y la solución de las necesidades socioeconómicas propias de cada región.

Este trabajo se divide en dos secciones; en la primera se presenta una identificación y clasificación de escenarios de consumo apropiados para los municipios de Nariño, considerando la información que pueda caracterizarlos, asociarlos en diferentes grupos y modelar su consumo. La clasificación de la demanda de energía eléctrica y la caracterización del consumo por regiones tiene importancia en el diseño de estrategias para la implementación de sistemas que satisfagan la demanda de las diferentes zonas donde estas se ejecuten.

En la segunda sección se encuentra modelación y proyección del consumo energético basándose en la clasificación que se realiza en la primera parte. La importancia de esta fracción del trabajo radica en conocer el potencial del consumo para establecer cuanta energía debe asegurarse para satisfacer las necesidades de las regiones que se estudian.

1. CARACTERÍSTICAS DEL DEPARTAMENTO DE NARIÑO

El Departamento de Nariño está ubicado al suroccidente de Colombia limitando al norte con el departamento de Cauca, al sur con la República de El Ecuador, al oriente con el departamento de Putumayo y al occidente con el Océano Pacífico. Nariño es atravesado de sur a norte por la cordillera de Los Andes y cuenta además con un pequeño porcentaje de territorio ubicado en la Región Amazónica. Por sus muchas y muy diversas características el Departamento es clasificado en subregiones. La división más general que se realiza es la geográfica donde se encuentran tres grandes subregiones: Llanura del Pacífico, Cordillera de Los Andes y la Vertiente Amazónica (Viloria de la Hoz, 2007) (Figura 1).

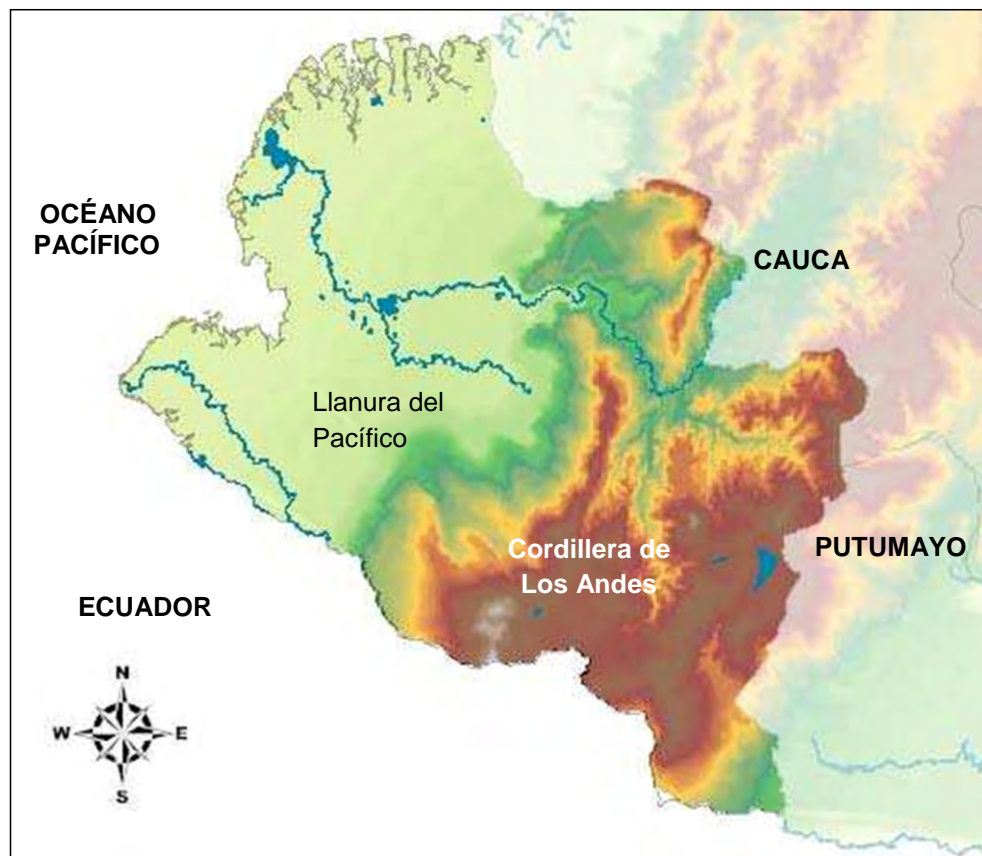


Figura 1. División geográfica de Nariño.
Fuente: IGAC

La Llanura Pacífica representa el 52% del territorio departamental. La subregión de la cordillera de los Andes ocupa el 46% del territorio departamental y su relieve montañoso alcanza alturas cercanas a los 5000 m.s.n.m. El 2% restante corresponde a la vertiente Amazónica, la cual se caracteriza por su vegetación selvática y de altas precipitaciones. Ésta última subregión está vinculada económicamente al departamento de Putumayo (Viloria de la Hoz, 2007).

El Departamento de Nariño tiene una extensión de 33.268 km², que corresponde al 2.9% del territorio nacional. Lo conforman 64 municipios, que debido a las particularidades del territorio, se encuentran ubicados a diferentes altitudes y presentan características significativamente diferentes, tanto en aspectos geográficos, como económicos y sociales. Este hecho repercute en el análisis de la demanda de energía eléctrica de cada municipio, puesto que introduce variables de diferentes tipos al estudio, y que se tratan más adelante. Del total de los municipios del departamento, 54 pertenecen exclusivamente a la zona interconectada a la red eléctrica nacional⁴ (ZI), 6 pertenecen a la zona no interconectada (ZNI) y 4 se encuentran en las dos zonas (Achicanoy Martínez, 2014). En la Tabla 1 puede verse la lista del total de municipios de Nariño y si estos pertenecen a la zona interconectada, no interconectada o a las dos zonas.

Tabla 1. Municipios de Nariño.

| MUNICIPIO | ZONA | MUNICIPIO | ZONA |
|------------------|-------------|------------------|-------------|
| ALBAN | ZI | GUALMATAN | ZI |
| ALDANA | ZI | ILES | ZI |
| ANCUYA | ZI | IMUES | ZI |
| ARBOLEDA | ZI | IPIALES | ZI |
| BARBACOAS | ZI/ZNI | LA CRUZ | ZI |
| BELEN | ZI | LA FLORIDA | ZI |
| BUESACO | ZI | LA LLANADA | ZI |
| CHACHAGÜI | ZI | LA TOLA | ZNI |
| COLON GENOVA | ZI | LA UNIÓN | ZI |

⁴ El Sistema Interconectado Nacional (SIN) está compuesto por las plantas y equipos de generación, la red de interconexión, las redes regionales e interregionales de transmisión, las redes de distribución y las cargas eléctricas de los usuarios (Artículo 11 de la Ley 143 de 1994).

| | | | |
|----------------------|--------|---------------|--------|
| CONSACA | ZI | LEIVA | ZI |
| CONTADERO | ZI | LINARES | ZI |
| CORDOBA | ZI | LOS ANDES | ZI |
| CUASPUD | ZI | MAGÜI PAYAN | ZI/ZNI |
| CUMBAL | ZI | MALLAMA | ZI |
| CUMBITARA | ZI | MOSQUERA | ZNI |
| EL CHARCO | ZNI | NARIÑO | ZI |
| EL PEÑOL | ZI | OLAYA HERRERA | ZNI |
| EL ROSARIO | ZI | OSPINA | ZI |
| EL TABLÓN DE GÓMEZ | ZI | PASTO | ZI |
| EL TAMBO | ZI | POLICARPA | ZI |
| FRANCISCO PIZARRO | ZNI | POTOSI | ZI |
| FUNES | ZI | PROVIDENCIA | ZI |
| GUACHUCAL | ZI | PUERRES | ZI |
| GUAITARILLA | ZI | PUPIALES | ZI |
| RICAUARTE | ZI | SANDONA | ZI |
| ROBERTO PAYAN | ZI/ZNI | SANTA BÁRBARA | ZNI |
| SAMANIEGO | ZI | SANTA CRUZ | ZI |
| SAN ANDRÉS DE TUMACO | ZI/ZNI | SAPUYES | ZI |
| SAN BERNARDO | ZI | TAMINANGO | ZI |
| SAN LORENZO | ZI | TANGUA | ZI |
| SAN PABLO | ZI | TUQUERRES | ZI |
| SAN PEDRO DE CARTAGO | ZI | YACUANQUER | ZI |

Nariño es un departamento cuyo territorio puede denominarse mayormente rural con altos niveles de pobreza y diferencias económicas y sociales muy marcadas entre los distintos municipios que lo conforman. En su mayoría, la economía departamental depende del sector agropecuario. A pesar de la infraestructura de transporte que se ha construido en las últimas cinco décadas, lo cierto es que todavía continúa marginado del resto del país. De acuerdo con la Gobernación de Nariño: “históricamente, una de las mayores causas del atraso del Departamento ha sido su aislamiento y falta de vías de comunicación. Hoy se sigue padeciendo de este mismo mal” (Gobernación de Nariño, 2004). La red vial de carreteras de Nariño es de 6.500 kilómetros, concentrándose en gran medida en la zona andina del departamento. Por el contrario, en la subregión del Pacífico existen pocas vías, siendo la excepción la transversal Pasto-Tumaco, la cual se encuentra en unas condiciones regulares y los ramales

Junín-Barbacoas y Barbacoas-Magüí Payán, vías terciarias en pésimo estado (Viloria de la Hoz, 2007).

Administrativamente Nariño se divide, como se ve en la Figura 2, en 13 subregiones: Sanquianga, Pacifico Sur, Telembí, Pie de Monte Costero, Exprovincia de Obando, Sabana, Abades, Occidente, Cordillera, Centro, Juanambú, Rio Mayo y Guambuyaco. Estas Subregiones fueron definidas en el plan de desarrollo departamental “Nariño Mejor” 2012-2015, como una manera de clasificación para consolidar un desarrollo endógeno con base a las capacidades económicas, sociales, institucionales e iniciativas diferenciales en las subregiones y municipios del departamento (Castillo Burbano & Jurado, 2014).



Figura 2. Departamento de Nariño por subregiones.
Fuente: Gobernación de Nariño, Secretaría de Planeación Departamental. Elaborado por Gerardo Ramos, Marzo 2012.

2. CLASIFICACIÓN

La identificación y clasificación de escenarios de consumo energético para el Departamento de Nariño, por las características antes mencionadas, necesita la recopilación de diversa información. Se debe contar con series de tiempo representativas de los consumos energéticos, que permitan realizar un análisis de su comportamiento y establecer su tendencia, además se requiere datos sociales y económicos de los municipios que permitan plasmar una clasificación más acertada según las características propias de cada región.

2.1. Variables

Debido a la heterogeneidad existente en el departamento, son muchas las variables que podrían adoptarse para clasificar los municipios; sin embargo, debido a la dificultad para conseguir información detallada y confiable, y tras los análisis realizados y las pruebas hechas utilizando herramientas informáticas, se seleccionan algunas de estas variables como las más representativas para el propósito de esta investigación. Los resultados de dichas pruebas se evaluaron directamente por el grado de fiabilidad de los resultados de la clasificación con los métodos escogidos para tal fin, a través de las técnicas que se muestran más adelante (Sección 2.4).

Respecto al consumo de energía eléctrica, se cuenta con una base de datos disponible en la plataforma del Sistema Único de Información de Servicios Públicos SUI⁵. En dicha plataforma se encuentra disponible información de consumo mensual de energía eléctrica en kilovatios hora al mes (kWh/mes), discriminada para cada municipio de la zona interconectada (ZI) y, a su vez, por cada una de las empresas que prestan el servicio en cada uno de ellos. Adicionalmente, la información reportada se clasifica por estrato y por tipo de suscriptor, es decir, se diferencia el consumo residencial, comercial e industrial. Además de la información antes descrita, en esta plataforma se encuentra información sobre el número de suscriptores para cada municipio organizada de manera similar a la información sobre el consumo. Los datos adquiridos para la presente

⁵ Sistema Único de Información de Servicios Públicos SUI: <http://www.sui.gov.co>.

investigación corresponden al periodo comprendido entre enero de 2003 y septiembre de 2014.

Para la clasificación de los escenarios de consumo de energía eléctrica, la información social y económica en los municipios juega un importante papel, y pese a que no existe una fórmula que relacione directamente los indicadores económicos con la forma en que las personas hacen uso de la electricidad, se evidencia, por medio de un análisis de varianza y covarianza de la información disponible, que están relacionados, tal y como se discute más adelante. Los índices con los que se cuenta para la investigación, y que fueron facilitados por el CEDRE⁶ de la Universidad de Nariño, son: Índice de Ruralidad, Índice de Miseria y Necesidades Básicas Insatisfechas.

- **Índice de Ruralidad (IR):** La relación entre distancia y densidad poblacional es inversamente proporcional, es decir, a mayor distancia y menor densidad poblacional mayor es el IR; inversamente, a menor distancia y mayor densidad menor es el IR. Este valor oscila entre 0 y 100. Así, si el valor resultante entre distancia y densidad poblacional se acerca a 100, más alto es el IR y entre más cercano a 0, más bajo el IR (Castillo Burbano & Jurado, 2014).

El IR permite relacionar aspectos geográficos y económicos en tanto aborda el comercio, transporte y demás interacciones con los centros urbanos. En materia de energía eléctrica, relaciona los posibles costos que podrían derivar de la expansión en infraestructura para la interconexión a la red nacional.

- **Índice de Miseria:** Consiste en la suma de la tasa de desempleo y la tasa de inflación, al suponer que ambos constituyen costos económicos para la región.

Varios estudios establecen como un parámetro de mucha importancia la intensidad de pago en las comunidades a las que

⁶ Centro de Estudios de Desarrollo Regional y Empresarial de la Facultad de Ciencias Económicas y Administrativas de la Universidad de Nariño.

quiere proveerse una solución en el suministro de energía. Teniendo en cuenta lo anterior, analizar las oportunidades de sostenimiento económico en dichas comunidades es primordial para conseguir buenos resultados en el modelo que se propone.

- **Necesidades Básicas Insatisfechas (NBI):** La metodología de NBI busca determinar, con ayuda de algunos indicadores simples, si las necesidades básicas de la población se encuentran cubiertas. Los grupos que no alcancen un umbral mínimo fijado son clasificados como pobres. Los indicadores simples seleccionados son: viviendas inadecuadas, viviendas con hacinamiento crítico, viviendas con alta dependencia económica y viviendas con niños en edad escolar que no asisten a la escuela (Departamento Administrativo Nacional de Estadística DANE, 2015). La electricidad juega un papel determinante en la ponderación de las NBI de cada municipio, razón por la cual se incluye como una variable importante para el proceso de clasificación de las poblaciones.

Por otra parte, al considerarse que la economía del Departamento depende en gran medida de la agricultura, la ganadería, la pesca y otras actividades relacionadas, como puede evidenciarse en los Informes de Coyuntura Económica Regional (ICER), que para el año 2013 reporta una participación del 14.3% en el PIB del departamento donde es superado únicamente por la participación del sector servicios y el comercio (Departamento Administrativo Nacional de Estadística, 2014). Se incluye información sobre estos ítems en los parámetros de clasificación. Estas variables, aunque no son totalmente sociales, tienen fuerte injerencia en las otras variables sociales y económicas de la población; pues está estrechamente relacionada con la capacidad actual y potencial de los beneficiarios de un proyecto para generar ingresos. También considera la gestión productiva actual y potencial de las comunidades en su interacción con los recursos naturales, agentes del mercado, patrones de ingresos (estacional, ingreso monetario y no monetario). Tiene efectos trascendentes en la sostenibilidad de los proyectos, pues influye directamente en la capacidad de pago, en la pobreza y en los riesgos financieros que puede tener un proyecto. Estas variables influyen estrechamente en otras variables e indicadores económicos como: nivel de desempleo, estabilidad de los recursos, inmigración, niveles de salarios y

otras (Arriaza, 2005) que a pesar de no ser utilizadas directamente en este estudio, son variables de alto impacto en el comportamiento de consumo de las personas.

2.2. Pre-procesamiento de la información

La información disponible fue recolectada de diferentes fuentes y no toda es del mismo tipo, para el caso de los reportes de consumo de energía eléctrica y cantidad de suscriptores, por ejemplo, se trata de series de tiempo; mientras que los índices de tipo socioeconómico antes mencionados, provistos por el CEDRE de la Universidad de Nariño, se presentan como datos transversales correspondientes al año 2011. De esta manera, toda la información utilizada para la clasificación pertenece a este mismo año y debido a que la cantidad de datos disponibles es más completa en relación con otros años.

Por otro lado, teniendo en cuenta que se debe manejar escalas de valores similares para una correcta clasificación, información como la cantidad de suscriptores o el total del consumo de energía eléctrica en cada municipio, son datos que presentan grandes y marcadas diferencias en sus escalas dependiendo del tamaño de cada municipio o la ubicación geográfica de éste. Por consiguiente, los datos correspondientes a estas variables deben escalarse de forma tal que todos los municipios se encuentren en un rango similar. Por lo anterior, los datos recolectados son separados en grupos, el primero es la información de producción agropecuaria de cada municipio. Esta información es tratada como variables de tipo categórico debido a que se estima su importancia en la economía del municipio en lugar de usar cantidades. Esto debido a que el rango de producción varía significativamente según el tamaño y población de cada municipio. Más adelante se explica de forma más detallada a que corresponde y como es utilizada la información de producción agropecuaria para la clasificación. El segundo grupo corresponde a las variables que se asumen como cuantitativas; dentro de este grupo se encuentra la cantidad de personas en cada vivienda, y para su cálculo se utiliza la información de suscriptores (viviendas) reportado por las empresas prestadoras del servicio de energía eléctrica y la cantidad de personas por municipio reportada por el DANE; la cantidad promedio de personas por vivienda se obtiene al dividir la cantidad de habitantes de un municipio entre la cantidad de viviendas. Se

tiene además, los índices socioeconómicos a los cuales se les realiza el proceso de normalización, es decir, se toma el mayor valor dentro del conjunto de datos de cada variable y se divide cada valor del conjunto entre esa cantidad.

El número de productos agrícolas que se produce en el Departamento, debido a la geografía nariñense, es bastante amplio. No obstante existen algunos que se producen en más municipios o en cantidades significativamente mayores, pero al igual que en el caso del consumo de energía, no se puede tratar esa información de manera directa como fue obtenida, por ello, fueron seleccionados 34 de los principales productos cultivados en el Departamento (Tabla 2) teniendo en cuenta que la producción de estos fuese comparativamente más amplia que la de otros cultivos en el municipio, y a la vez que dicha producción fuera relevante dentro de la producción total en el Departamento (Ministerio de Comercio, Industria y Turismo, 2013). De ahí se tiene que por cada municipio, se haya escogido entre uno y cinco productos agrícolas como los más representativos, expresando la información obtenida como un vector con variables indicativas o binarias, es decir, en el caso de que un producto agrícola sea significativo para la producción de un municipio este tendrá un valor alto o uno, en caso contrario tendrá un valor bajo o cero. En total el vector de producción agrícola para cada municipio está conformado por 34 valores binarios. La información de producción agrícola de Nariño se encuentra disponible en la plataforma web Trazas Nariño⁷.

Tabla 2. Principales productos agrícolas de Nariño.

| | |
|------------------|------------------|
| Arveja | Mango |
| Banano | Maní |
| Borojó | Mora |
| Café | Palmito |
| Caña panelera | Papa |
| Cebada | Papa amarilla |
| Cebolla cabezona | Papaya |
| Cebolla junca | Piña |
| Chiro | Plátano cafetero |
| Cítricos | Plátano costa |

⁷ Sistema de Trazabilidad Agrícola de la Gobernación de Nariño: <http://trazasnarino.org>.

| | |
|---------|--------------|
| Coco | Remolacha |
| Fique | Repollo solo |
| Frijol | Sandia |
| Haba | Tomate árbol |
| Lechuga | Tomate mesa |
| Lulo | Trigo |
| Maíz | Zanahoria |

Bajo un esquema similar se organiza la información de producción pecuaria, tal como la producción de leche, el sacrificio de ganado para consumo de carne y la producción piscícola y pesquera. Se pensó utilizar el índice de Producto Interno Bruto (PIB) en lugar de la producción agropecuaria debido a que provee información más completa sobre la influencia de la producción local sobre la demanda de energía eléctrica, sin embargo, no fue posible conseguir información sobre este indicador discriminada para cada municipio. Por lo tanto se consideró realizar el análisis con la información disponible.

2.3. Métodos de clasificación

En muchos casos en los que es necesario extraer y analizar características de un conjunto de datos, la programación directa puede resultar muy complicada al tratar de especificar la forma correcta en que esto debe hacerse. Los sistemas de aprendizaje ofrecen un método alternativo para superar esta situación. A través de la información extraída de los datos de muestra, dichos sistemas son capaces de adaptarse e inferir una solución para esas tareas. En general, el campo del análisis de patrones estudia sistemas que usan la metodología de aprendizaje para descubrir patrones en los datos. Dentro del campo del aprendizaje de máquinas, existen dos formas principales de aprendizaje: supervisado: se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo que determina la respuesta que debería generar la red a partir de una entrada determinada; y no supervisado: la salida representa el grado de familiaridad o similitud entre la información que se le está presentando a la entrada y las informaciones que se le han mostrado hasta entonces. Está constituido por un conjunto de reglas que dan a la red la habilidad de aprender

asociaciones entre los patrones que ocurren en conjunto (Alpaydin, 2010) (Llanos, Azuaje, & Romero, 2013) .

El análisis *Cluster*, también llamado análisis de segmentación o análisis taxonómico, crea grupos o ramas de información o *clusters*. El Agrupamiento (*Clustering*) es el método más importante y ampliamente usado de aprendizaje no supervisado: El análisis *Cluster* ayuda a descubrir la organización interna del conjunto de datos a través de la clasificación de la información en *clusters*. Esta palabra indica la separación en grupos con características similares. Intuitivamente, la división en *clusters* puede ser definida por la semejanza interna de los datos del clúster y las diferencias externas entre *clusters*. De ahí que la información se divide en un número de grupos compuestos por datos similares en los que existen elementos distintivos. Esta metodología es ampliamente usada en el análisis estadístico multivariado y en el aprendizaje de máquinas. El número de *clusters* no está especificado. Como soluciones exactas, siempre es difícil encontrar computacionalmente aproximaciones efectivas a través de procedimientos de relación (Shawe-Taylor & Cristianini, 2004).

Una gran variedad de algoritmos de *clustering* han surgido en los últimos años, los cuales se pueden clasificar en:

- Métodos Jerárquicos
 - Algoritmos Aglomerativos
 - Algoritmos Divisivos
- Métodos de Particionado y Recolocación
 - *Clustering* Probabilístico
 - Métodos de los k-vecinos (*kmedoids*)
 - Métodos de las k-medias
 - Algoritmos Basados en Densidad
 - *Clustering* de Conectividad Basada en Densidad *Density-Based Connectivity Clustering*
 - *Clustering* basado en Funciones de Densidad
- Métodos Basados en Rejillas
- Métodos Basados en la Co-Ocurrencia de Datos Categóricos
- *Clustering* Basado en Restricciones
- Algoritmos para Datos de Grandes Dimensiones
- *Clustering* Subespacial
- Técnicas de *Co-Clustering*

A continuación se presentan los algoritmos utilizados en la presente investigación:

2.3.1. Método de agrupamiento por Mixturas de Gaussianas

Este método es utilizado con mucha frecuencia en el agrupamiento de datos. Pertenece a los llamados métodos jerárquicos, los cuales tienen por objetivo agrupar *clusters* para formar uno nuevo o bien separar alguno ya existente para dar origen a otros dos, de tal forma que, si sucesivamente se va efectuando este proceso de aglomeración o división, se minimice alguna distancia o bien se maximice alguna medida de similitud.

Las mixturas Gaussianas son combinaciones de distribuciones normales o funciones de Gauss. Una mixtura de k Gaussianas puede por lo tanto ser escrita o ser vista como una suma de densidades de Gaussianas. Una mixtura de Gaussianas con ciertos valores de pesos se ve de la forma:

$$gm(x) = \sum_{k=1}^K w_k * g(\mu_k, \Sigma_k)(x) \quad (1)$$

Donde $g(\mu_k, \Sigma_k)$ es la función de densidad de probabilidad Gaussiana de cada grupo y w_k son pesos (valores positivos cuya suma es igual a 1). Al variar el número de Gaussianas K , los pesos w y los parámetros de cada de las funciones de densidad μ y Σ , las mixturas de Gaussianas pueden ser usadas para describir algunas Funciones de Densidad de Probabilidad Complejas. Los parámetros de la función de densidad de probabilidad son el número de Gaussianas, sus factores de peso y los parámetros de cada función Gaussiana tales como la medida μ y la matriz de covarianza Σ (Oropeza Rodríguez, 2006).

2.3.2. K-Nearest Neighbors (KNN)

El análisis de vecinos más próximos (*nearest neighbors* en inglés) es un método para clasificar casos basándose en su parecido a otros casos. En el aprendizaje automático, se desarrolló como una forma de reconocer

patrones de datos sin la necesidad de una coincidencia exacta con patrones o casos almacenados. Los casos parecidos están próximos y los que no lo son están alejados entre sí. Por lo tanto, la distancia entre dos casos es una medida de disimilitud. Los casos próximos entre sí se denominan “vecinos.” Cuando se presenta un nuevo caso (reserva), se calcula su distancia con respecto a los casos del modelo. Las clasificaciones de los casos más parecidos – los vecinos más próximos – se cuadrán y el nuevo caso se incluye en la categoría que contiene el mayor número de vecinos más próximos. Se puede especificar el número de vecinos más próximos que deben examinarse; este valor se denomina k^8 . A pesar de ser un método automático de clasificación, este algoritmo requiere de un objetivo definido para la ubicación de las variables respecto de una observación, tal como sucede con los métodos de aprendizaje supervisados.

2.3.3. K-Means Clustering

En palabras simples, es un algoritmo para clasificar o agrupar objetos basándose en atributos o características en un número de ramas k , donde k es un número entero positivo. La agrupación se realiza al minimizar la suma de los cuadrados de las distancias entre los datos y el centroide del correspondiente clúster (Teknomo, 2007).

$$J = \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (2)$$

Donde $\|x_j - \mu_i\|^2$ corresponde a la medida de la distancia elegida entre un punto del conjunto de observaciones x y el centro del *cluster* μ .

2.3.4. Redes Neuronales Artificiales RNA

En los últimos años la computación neuronal ha emergido como una tecnología práctica con aplicaciones exitosas en muchos campos. La mayoría de esas aplicaciones tienen que ver con el reconocimiento de

⁸ Disponible en línea: http://www-01.ibm.com/support/knowledgecenter/SSLVMB_21.0.0

patrones haciendo uso de arquitecturas de red con realimentación tal como el perceptrón multicapa y la red de función de base radial (Bishop, 1995). El agrupamiento es otra excelente aplicación para las redes neuronales. Este proceso involucra agrupar los datos por sus similitudes. Los algoritmos para realizar esta tarea son diversos, por ejemplo se puede realizar *clustering* con el método de mapas de auto-organización (SOM por sus siglas en inglés), los cuales aprenden a clasificar información de acuerdo a la distribución y topología de los vectores de entrada⁹.

2.4. Pruebas

Al no existir una fórmula o método específico que permita determinar el número de *clusters* que agrupe de mejor manera la información que se desea clasificar, se debe recurrir a la realización de pruebas y la comparación de los resultados de cada una de éstas para la selección de la mejor solución (ensayo y error). El punto de partida seleccionado para la realización de las pruebas de clasificación es la división administrativa existente que separa a Nariño en 13 subregiones. La selección de esta división obedece únicamente a la toma de una referencia para los resultados y no como un objetivo de clasificación. Se resalta que la mencionada clasificación del Departamento es de tipo geográfica y que aunque algunos de los municipios que integran cada subregión comparten algunas características, aquellas que influyen la demanda de energía varían entre municipios que pertenecen a la misma subregión geográfica. En este caso se cuenta con dos tipos de variables diferentes: categóricas y numéricas. Debido a esto la clasificación se realiza en dos etapas: la primera utiliza las variables nominales que corresponden a la producción agropecuaria significativa de cada municipio. La segunda etapa consiste en el agrupamiento de los municipios según sus índices socioeconómicos y cantidad promedio de personas por vivienda. Los resultados que a continuación se presentan para los modelos de agrupamiento por Mixturas de Gaussianas y K-Means corresponden a los grupos obtenidos tras realizar todo el proceso de clasificación, es decir, los resultados tras las dos etapas de la clasificación. En el caso de los resultados de las RNA, se realiza una explicación más completa del proceso, pues éste método se utiliza finalmente para la clasificación de los municipios.

⁹ Disponible en: <http://www.mathworks.com/>

Tradicionalmente, la validación del *clustering* se ha clasificado en 3 categorías: externa, interna y relativa. La validación externa asume el conocimiento de la partición correcta de los datos. Es decir, la verdadera partición es conocida (lo que se conoce como *ground truth* o *standard gold*). La validación interna no necesita el conocimiento de la partición correcta ya que se basa únicamente en los datos agrupados. Es decir, evalúa la partición con base en los datos y las distancias entre ellos. La ventaja principal de esta aproximación es que no requiere el conocimiento de la partición correcta por lo que se puede utilizar en aplicaciones reales y no sólo en experimentos de laboratorio. Su desventaja principal es que evalúa la partición sobre los mismos datos que se han utilizado para construirla. El término validación relativa no está claramente definido, algunos autores lo usan para la comparación de varias particiones, pero sólo mediante índices de validación interna. Otros autores la limitan a la comparación de particiones obtenidas mediante la variación de los parámetros de un mismo algoritmo (Gurrutxaga Goikoetxea, 2010).

Para determinar cuan bien ubicados se encuentran los elementos dentro de cada *cluster*: existe un valor conocido como coeficiente de silueta (*silhouette*). Éste utiliza el cuadrado de la distancia euclidiana para indicar la proximidad de los puntos dentro de un *cluster* y compararla con la distancia hacia los puntos en otro *cluster*. Esta medida se encuentra entre +1 y -1. Los valores cercanos a +1 indican una buena clasificación. Por su parte, los valores próximos a 0 indican que no existe diferencia en si un punto se ubica en un *cluster* o en otro. Finalmente, los valores que se aproximan a -1 indican que los puntos probablemente se asignaron dentro del grupo erróneo.

2.4.1. Método de agrupamiento por Mixturas de Gaussianas

El método de mixturas de Gaussianas organiza los datos de forma jerárquica, requiere la selección de parámetros para la clasificación y es necesario que se tenga gran conocimiento en el objetivo de la clasificación para interpretar los resultados. En este caso, se realizaron algunas pruebas con los parámetros: media y covarianza, seleccionados mediante la variación de valores iniciales obtenidos mediante una técnica de tipo aleatorio. Por otra parte, la elección del número de grupos requiere que se haga varias pruebas con el fin de seleccionar la mejor solución. Para este

método se tiene que los mejores resultados se obtuvieron con 14 *clusters* según las pruebas de validación realizadas usando el método de silueta. En el Anexo 1 se encuentra el compendio de los 14 *clusters*; los municipios que los conforman y algunas de las características de cada uno de ellos.

La Figura 4 corresponde a la clasificación de los municipios dentro del mapa geográfico del Departamento, cada *cluster* tiene un color característico con el que se identifican los municipios que lo conforman.

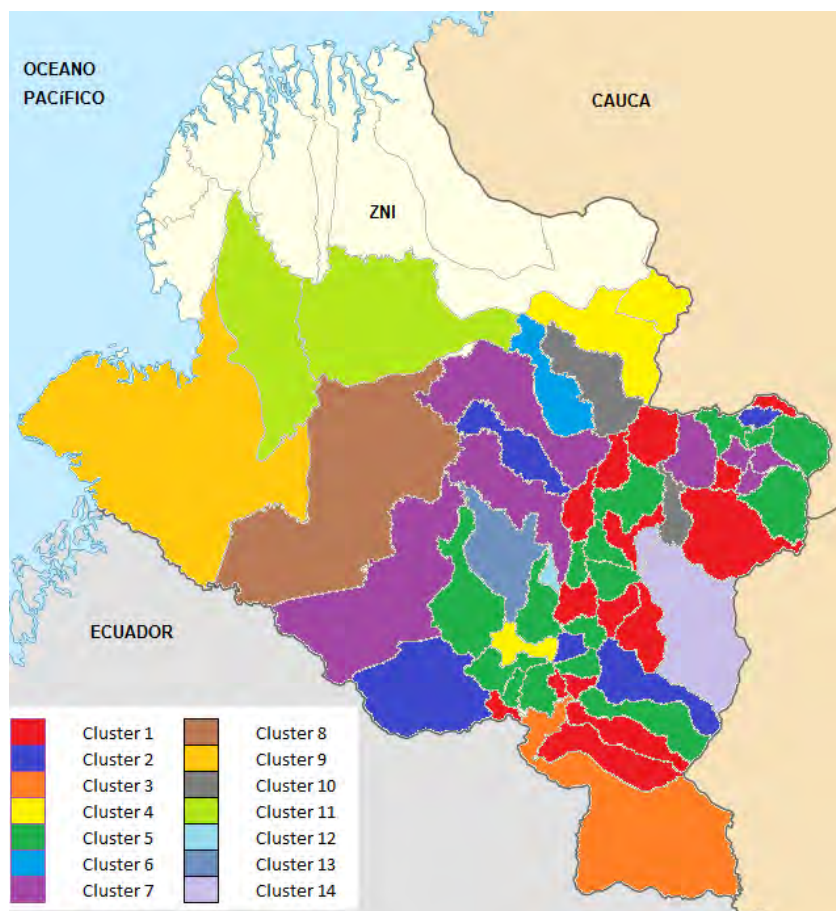


Figura 3. Clasificación de los municipios de Nariño mediante *Gaussian Mixture Clustering*.

Si bien la clasificación realizada por el algoritmo, vista en el mapa (Figura 4), dista de la clasificación administrativa que se toma como referencia en esta investigación, se puede observar que los municipios que ahora conforman las nuevas 14 subregiones comparten algunas propiedades, tal como se ve en la Tabla 3. Allí se toma como ejemplo el *cluster 7* para mostrar algunas características de los municipios que son agrupados por este algoritmo. A pesar de algunas excepciones, la clasificación obtenida

agrupa municipios con consumos de energía eléctrica per cápita similares según los datos suministrados por las empresas prestadoras del servicio a través de la plataforma del SUI.

Tabla 3. Características de los municipios en los *Cluster 2 y 7*.

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Natalidad | Pob_total | Consumo Residen. [kWh] | Consumo Comercial [kWh] | Consumo Industrial [kWh] | Subregión | PersXviv | ConsPerCap [kWh] | CLUSTER |
|---------------------|----------------|-----------|---------|------|-----------|-----------|------------------------|-------------------------|--------------------------|----------------|----------|------------------|---------|
| COLON_GENOVA | 1750 | 38.24 | 3.06 | 4.05 | 10.24 | 9959 | 1186005 | 788.50 | | Río Mayo | 4.89 | 119.09 | 2 |
| CUMBAL | 3100 | 51.51 | 2.55 | 3.7 | 11.65 | 34858 | 4403992 | 4814.28 | 19814.15 | Exprovincia de | 5.65 | 126.34 | |
| FUNES | 2300 | 55.46 | 3.39 | 4.17 | 10.32 | 6685 | 909421 | 12742.87 | 2778 | Exprovincia de | 4.38 | 136.04 | |
| LA_LLANADA | 2300 | 52.61 | 2.52 | 3.61 | 10.77 | 6127 | 755036 | 37030.25 | 17971.56 | Guambuyaco | 5.87 | 123.23 | |
| OSPINA | 2850 | 40.52 | 3.15 | 4.08 | 6.55 | 8547 | 1060157 | 266.67 | 1561.22 | Sabana | 4.86 | 124.04 | |
| ALBAN | 1935 | 29.89 | 2.92 | 3.82 | 5.39 | 20980 | 1776298 | 3857.25 | | Río Mayo | 6.92 | 84.67 | 7 |
| LOS_ANDES | 1588 | 55.93 | 3.47 | 4.23 | 5.97 | 18804 | 1602493 | 1633.40 | 8595 | Guambuyaco | 6.90 | 85.22 | |
| RICAUARTE | 1181 | 57.8 | 3.7 | 4.19 | 14.45 | 17090 | 1462803 | 6716.55 | 3388.75 | Pie de Monte | 5.19 | 85.59 | |
| SAMANIEGO | 1750 | 43.69 | 3.14 | 4.06 | 7.13 | 49906 | 4320418 | 6031.35 | 4844.18 | Abades | 7.16 | 86.57 | |
| SAN_BERNARDO | 2100 | 35.38 | 3.3 | 4.05 | 4.85 | 17116 | 1378722 | 1351.33 | | Río Mayo | 8.63 | 80.55 | |
| SAN_LORENZO | 2150 | 44.21 | 3.22 | 4.05 | 10.02 | 19261 | 1762450 | 6583.25 | 152.16 | Juanambú | 4.39 | 91.50 | |
| SAN_PEDRO_DE_CARTAG | 1750 | 39.95 | 3.14 | 4.04 | 11.02 | 7348 | 658683 | 419.21 | 2960 | Juanambú | 5.55 | 89.64 | |

No obstante a lo anterior existen casos como los del municipio de Sapuyes, el cual es clasificado en el mismo cluster que los municipios de Leiva y El Rosario. En este caso los índices socio-económicos tienen una fuerte influencia en la selección puesto que comparten una importante similitud en ese aspecto. Sin embargo, en materia de condiciones de demanda de energía eléctrica, se puede establecer que la clasificación realizada puede mejorarse.

Al aplicar el gráfico de silueta a esta clasificación, ver Figura 4, los resultados muestran que el agrupamiento ubica 7 *clusters* con un solo elemento. A la vez, cuatro elementos clasificados toman valores negativos, esto debido a que la medida de las distancias entre los puntos que conforman dichos *clusters* con respecto a algunos puntos de otros *clusters* no son significativamente menores, en otras palabras, no existe la suficiente cohesión entre los puntos de un mismo *cluster*, lo que implica que probablemente se encuentren ubicados en el *cluster* incorrecto.

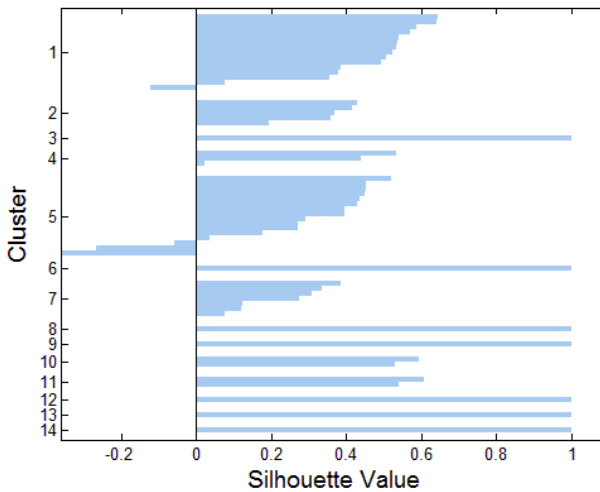


Figura 4. *Silhouette Graph*. 14 Clusters.

2.4.2. K-Means Clustering

La clasificación realizada con este algoritmo, para el caso de esta investigación, presenta resultados bastante interesantes, lo que demuestra que para el tipo de datos con que se cuenta, los métodos de Particionado y Recolocación funcionan mejor que aquellos que realizan una clasificación jerárquica. El mapa de Nariño con la clasificación realizada por el algoritmo de agrupamiento *k-means* en ocho *clusters* es presentado en la Figura 7.

En este caso es interesante analizar cómo están conformadas las subregiones, puede verse que algunas de las características de los municipios ubicados en diferentes *clusters* se encuentran en rangos muy similares, especialmente el valor correspondiente a consumo de energía per cápita, el cual es tomado como referencia para comprobar la validez del modelo. Lo anterior puede evidenciarse en los datos de la tabla 4, en la cual se resume algunas de las variables propias de los municipios asignados a los *clusters* 7 y 8.

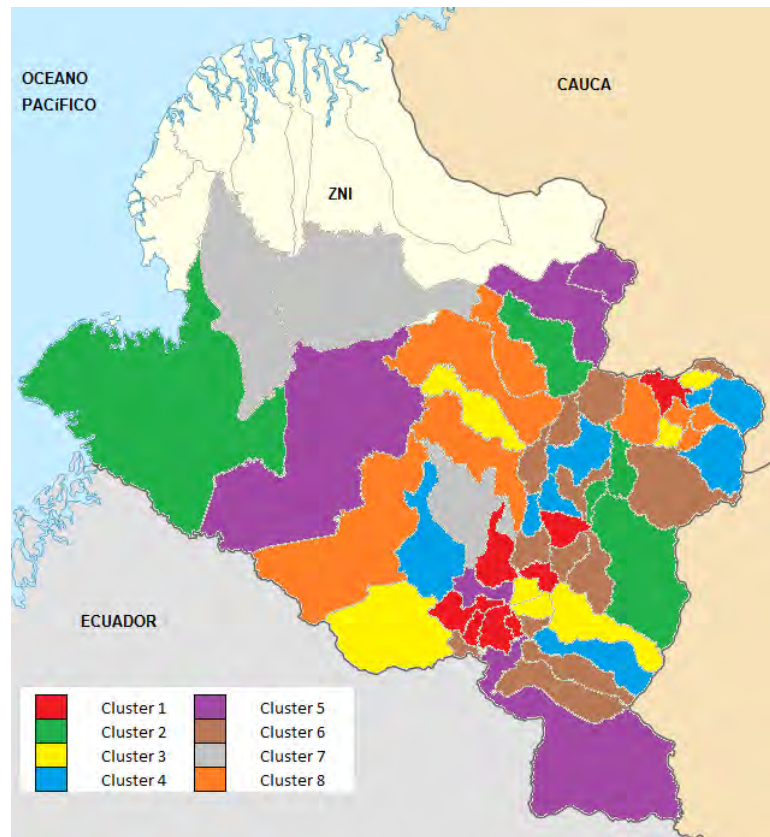


Figura 5. Clasificación de los municipios de Nariño mediante *k-means Clustering*.

Al igual que con el algoritmo de *Mixturas de Gaussianas* aquí observamos algunos casos en que la clasificación realizada no es enteramente satisfactoria para los objetivos con los que se realiza, citamos el cluster 6, en el cual se agrupan los municipios de Buesaco, Potosí, Córdoba y algunos otros. La revisión de las características de éstos municipios indica grandes brechas en varias materias como los de producción agropecuaria entre ellos, una vez más los índices socio-económicos juegan un papel determinante para la clasificación que aquí se alcanzó.

La Figura 8 presenta la clasificación de los municipios, mostrando el número de elementos en cada *cluster* con su correspondiente valor de fiabilidad en la clasificación realizada. Los resultados obtenidos con ocho *clusters* entregan resultados bastante interesantes, pero se debe destacar que hay algunos elementos con valores negativos que si bien son muy cercanos a cero, demuestran que es posible que hayan sido agrupados de manera errónea.

Tabla 4. Características de los municipios de los *clusters* 7 y 8.

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Natalidad | Pob_total | Consumo Resid | Consumo Comercial | Consumo Industrial | Subregión | PersXviv | ConsXpers |
|----------------|----------------|-----------|---------|------|-----------|-----------|---------------|-------------------|--------------------|--------------|----------|-----------|
| MAGUI_PAYAN | 40 | 60.83 | 3.46 | 4.41 | 5.2 | 19822 | 1357591 | | | Telebí | 8.68 | 68.49 |
| PROVIDENCIA | 2300 | 34.65 | 3.48 | 4.17 | 4.98 | 12644 | 677131 | 5651.00 | | Abades | 5.71 | 53.55 |
| ROBERTO_PAYAN | 24 | 59.2 | 3.65 | 4.29 | 7.05 | 20132 | 1230507 | | | Telebí | 6.35 | 61.12 |
| SANTA_CRUZ | 2400 | 49.79 | 3.72 | 4.22 | 3.54 | 24886 | 847780 | 2413.60 | 390 | Abades | 9.24 | 34.07 |
| ALBAN | 1935 | 29.89 | 2.92 | 3.82 | 5.39 | 20980 | 1776298 | 3857.25 | | Río Mayo | 6.92 | 84.67 |
| CUMBITARA | 1600 | 50.9 | 4.61 | 4.61 | 7.43 | 13590 | 1408626 | 8714.99 | | Cordillera | 8.59 | 103.65 |
| LOS_ANDES | 1588 | 55.93 | 3.47 | 4.23 | 5.97 | 18804 | 1602493 | 1633.40 | 8595 | Guambuyaco | 6.90 | 85.22 |
| RICARTE | 1181 | 57.8 | 3.7 | 4.19 | 14.45 | 17090 | 1462803 | 6716.55 | 3388.75 | Pie de Monte | 5.19 | 85.59 |
| SAMANIEGO | 1750 | 43.69 | 3.14 | 4.06 | 7.13 | 49906 | 4320418 | 6031.35 | 4844.18 | Abades | 7.16 | 86.57 |
| SAN_BERNARDO | 2100 | 35.38 | 3.3 | 4.05 | 4.85 | 17116 | 1378722 | 1351.33 | | Río Mayo | 8.63 | 80.55 |
| SAN_LORENZO | 2150 | 44.21 | 3.22 | 4.05 | 10.02 | 19261 | 1762450 | 6583.25 | 152.16 | Juanambú | 4.39 | 91.50 |
| SAN_PEDRO_DE_C | 1750 | 39.95 | 3.14 | 4.04 | 11.02 | 7348 | 658683 | 419.21 | 2960 | Juanambú | 5.55 | 89.64 |

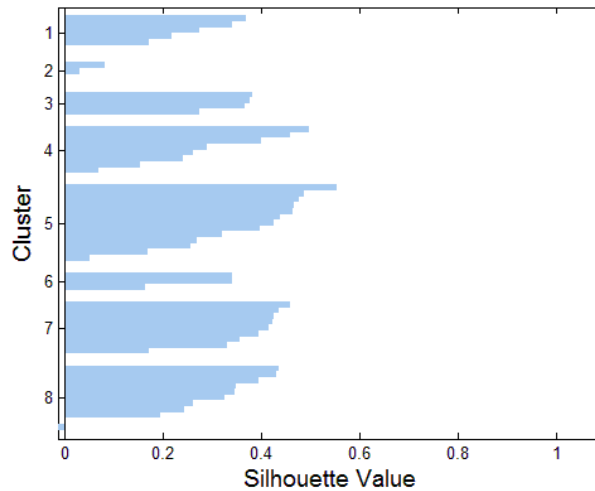


Figura 6. *Silhouette Graph*. 8 Clusters.

2.4.3. Redes Neuronales Artificiales RNA

Dentro del área de las RNA, existen diferentes métodos para realizar el proceso de clasificación, uno de los más utilizados por su buen desempeño y sencillez es el de mapas de auto-organización (SOM por sus siglas en inglés). Este tipo de RNA cuenta con una capa y una o más neuronas organizadas en una red. Los SOM aprenden tanto la organización como la distribución de los vectores de entrada que están siendo entrenados. Las

distancias entre neuronas son calculadas de acuerdo con algunas funciones específicas; en este caso se utiliza la más común de éstas: la función de distancia *link*, la cual es simplemente el número de enlaces o pasos que debe tomarse para alcanzar la neurona en consideración. La Figura 9 representa un ejemplo de la topología de una red de tipo SOM.

Como antes se mencionó, los datos utilizados para la clasificación son de dos tipos diferentes; para el primer caso: los datos categóricos, una red primaria es entrenada dando como resultado cuatro regiones principales, de diferente tamaño, con características particulares pero asociadas por la relevancia de la producción agropecuaria. El mapa presentado en la Figura 8 permite identificar esta primera división de la cual se desprende el resultado final de la clasificación realizada para esta investigación.

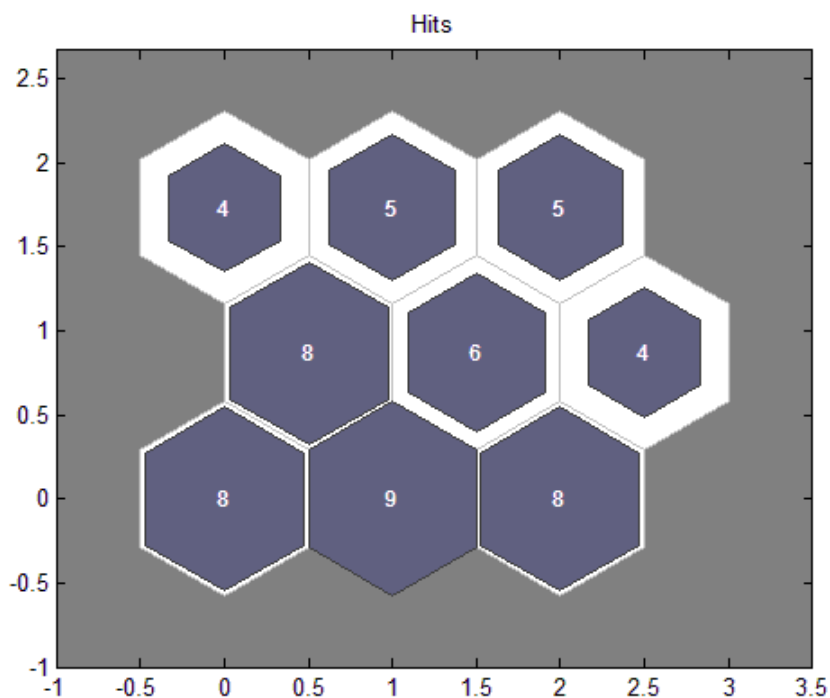


Figura 7. Self-organized map SOM.

Posteriormente, los índices socioeconómicos, junto con la cantidad de personas que en promedio habitan en una vivienda en cada municipio, son utilizados para realizar una clasificación más fina de los comportamientos que pueden influenciar en el tipo de consumo de cada población en el departamento de Nariño. El agrupamiento final que se obtiene tras este

proceso se muestra para cada uno de los *clusters* obtenidos en la clasificación primaria. En la Figura 9 se observa los *clusters* obtenidos para la primera región (*cluster* 1 de la clasificación primaria). De la misma manera se presenta los resultados del agrupamiento para de los municipios pertenecientes a cada región en las Figuras 10, 11 y 12.

La clasificación primaria de los municipios de Nariño, salvo pocas excepciones divide al Departamento según la geografía. Las cuatro regiones se marcan por su ubicación, este fenómeno es comprensible debido a que la obtención de productos agrícolas, así como los conseguidos de fuentes animales, dependen de características geográficas como temperatura, altitud, entre otros. Los cuales se comparten entre los municipios pertenecientes a las mismas regiones geográficas. Uno de los casos en que más separación geográfica se observa entre municipios que pertenecen al mismo *cluster* es el de Buesaco (*Cluster* 2 de la Figura 8). A pesar de que este municipio se ubica en el extremo opuesto al de los demás municipios que integran este *cluster*, sus características de producción presentan mayor afinidad con esos municipios en comparación a los que se encuentran ubicados a su alrededor, siendo la producción de carne, pesca y cítricos los elementos que más influyen en este agrupamiento. El mismo caso se presenta para los municipios de La Cruz, Colón, San Pedro de Cartago y Belén pertenecientes al *Cluster* 3 en la Figura 8 que al igual que Buesaco, a pesar de encontrarse alejados de la ubicación de los demás municipios que integran el *Cluster*, presentan mayor afinidad con estos en los comportamientos de producción que con los municipios vecinos.

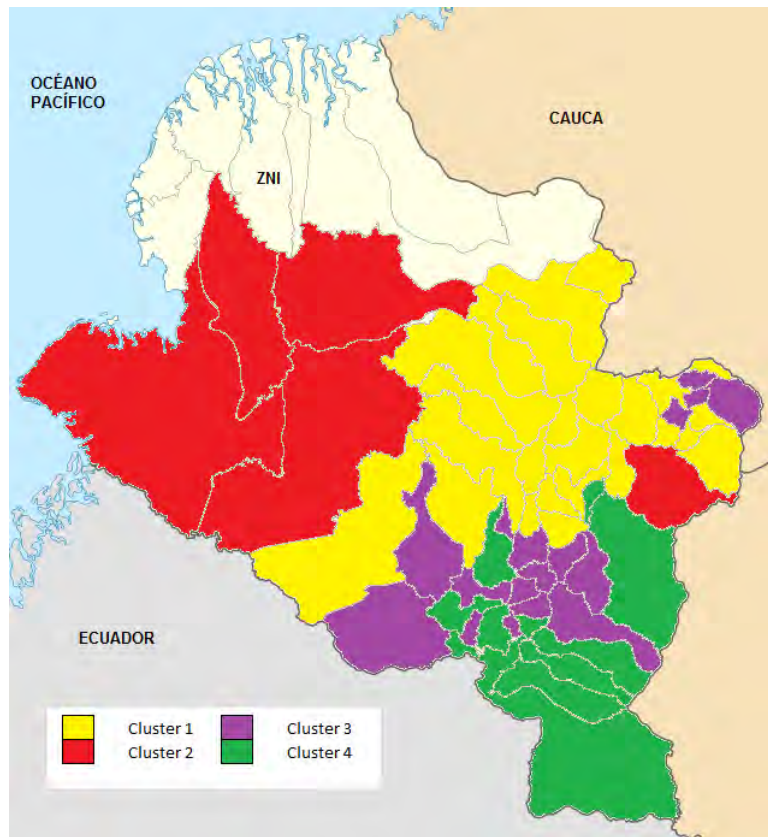


Figura 8. Clasificación de los municipios de la ZI según su producción Agropecuaria.

La primera división realizada permite separar en cuatro grandes grupos los municipios de la ZI de Nariño, sin embargo para la finalidad de la investigación se necesita un proceso que permita integrar más características influyentes en el comportamiento, o posible comportamiento, de consumo de energía eléctrica. Para ello, dentro de cada uno de las regiones obtenidas en el proceso anterior se realiza una clasificación utilizando los índices socioeconómicos propios de cada municipio. La Figura 9 muestra el agrupamiento de los municipios de la primera región en seis *clusters* identificados por diferentes colores. A diferencia de los resultados obtenidos en la clasificación primaria en los cuales el agrupamiento era claramente más geográfico, en este caso encontramos que los municipios que se asocian pueden estar separados por municipios que pertenecen a otro *cluster* o incluso muy alejados unos de otros, no obstante a lo anterior los municipios que conforman cada rama del proceso de clasificación comparten enormes similitudes en sus características y comportamientos que influyen la demanda de energía.

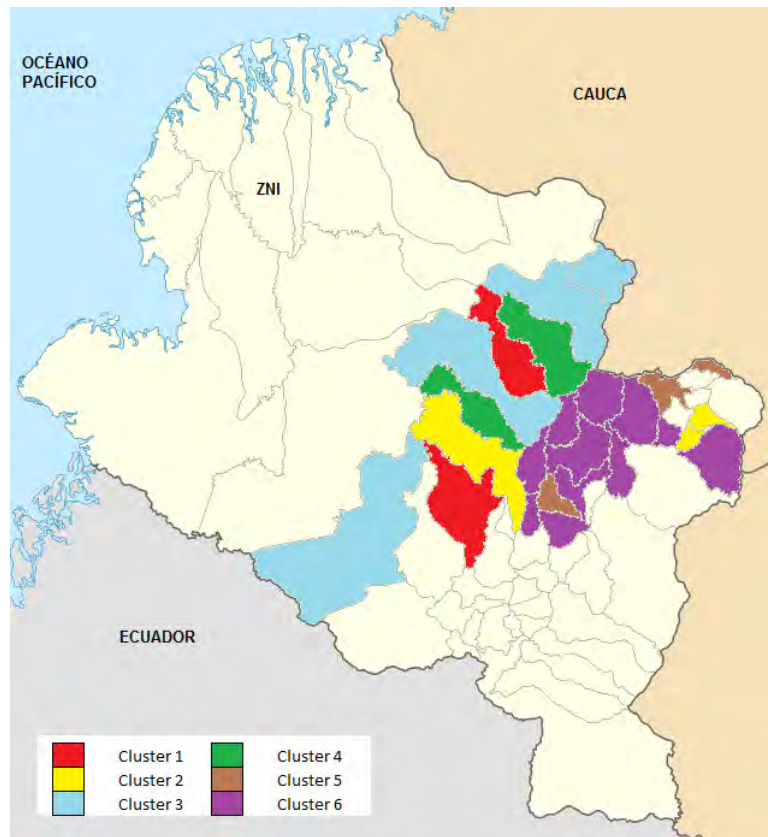


Figura 9. Clasificación de la Región 1 del *clustering* primario.

El cluster 5 se toma como ejemplo para demostrar la validez del modelo de clasificación. En ese grupo se encuentra los municipios de San Pablo, La Unión y Sandoná. Estos tres municipios, además de tener el café, el plátano cafetero, la caña y la muy baja explotación pecuaria como características comunes en el nivel de producción, cuentan con índices socio-económicos muy similares razón por la cual son agrupados en el mismo cluster.

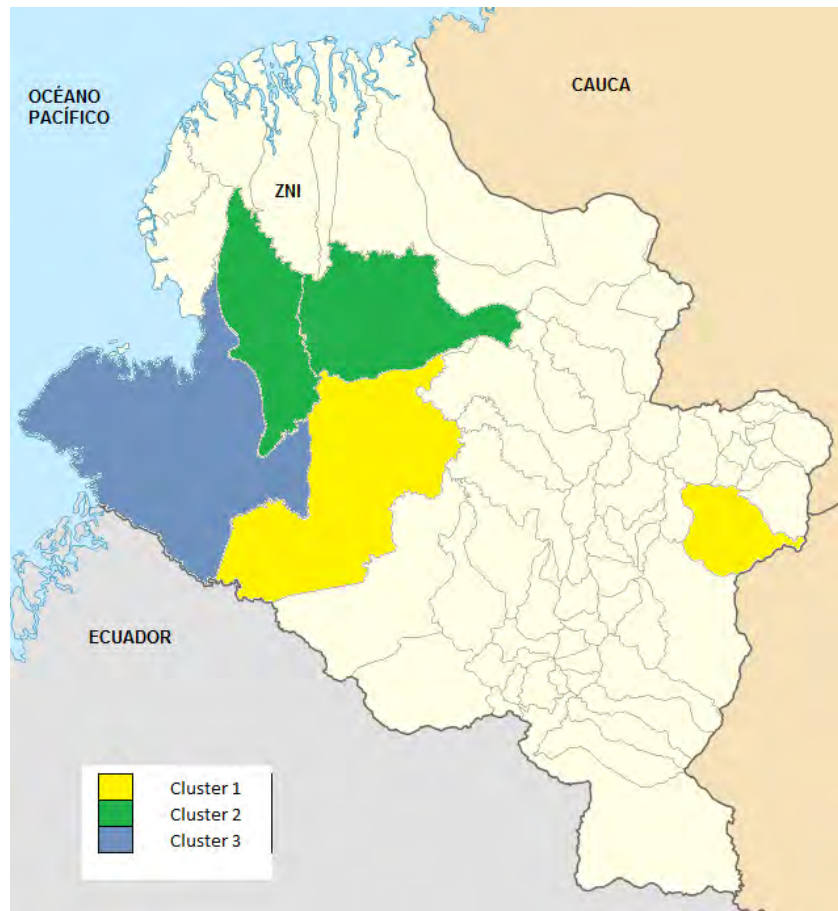


Figura 10. Clasificación de la Región 2 del *clustering* primario.

En la Figura 10 se encuentra representada la clasificación de la segunda región obtenida en el proceso de *clustering* primario; este es el caso con menos municipios y menos *clusters*. Tan solo cinco municipios se ubican en esta región o rama de la primera clasificación, cuatro de ellos ubicados en la zona de la llanura del Pacífico y el último en la zona Andina del Departamento. La clasificación final realizada a esta región la separa en tres *clusters*, uno de ellos conformado por un solo municipio: Tumaco. Las características geográficas y de producción del municipio de Tumaco lo colocan en esta región, pero éste es separado finalmente debido a que a diferencia de los otros municipios de la misma zona, ese municipio tiene una economía más influyente en el Departamento, además de contar con mayor comercio y un comportamiento más urbano en comparación a los municipios que lo rodean. Por su parte, Buesaco pese a su distancia con Barbacoas es asociado a éste debido a la gran cantidad de características

compartidas tanto en la parte de producción agropecuaria como en el tema socioeconómico, donde los índices de Miseria y NBI presentan valores muy cercanos. Por último, los otros dos municipios que se asocian en el *Cluster* 2 de la Figura 10 son Magüi Payán y Roberto Payán. En este último caso la interpretación de la clasificación de estos dos municipios en el mismo grupo es bastante sencilla puesto que los dos presentan más similitudes que diferencias.

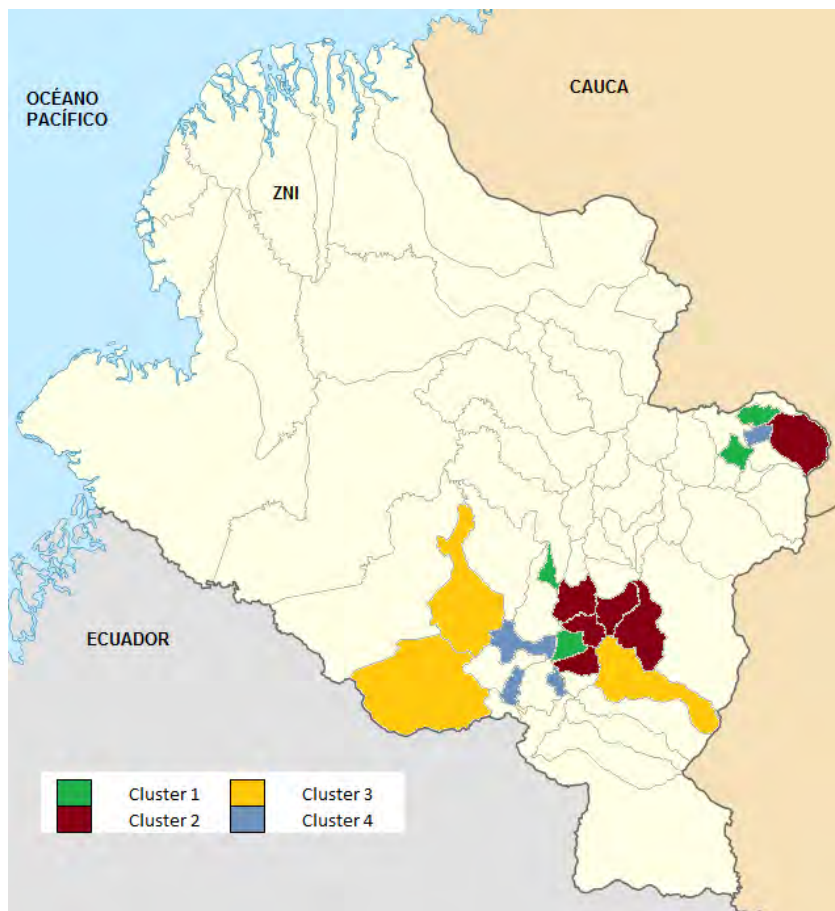


Figura 11. Clasificación de la Región 3 del *clustering* primario.

El agrupamiento interno de la tercera región de la clasificación primaria se muestra en la Figura 11. Al igual que en el caso de la primera región, se observa que los municipios que conforman los *clusters* están dispersos, lo cual se explica por las similitudes en materia económica entre los municipios de cada rama.

El caso donde mayor dispersión geográfica se observa entre los municipios pertenecientes al mismo cluster es para el que contiene a Guaitarilla, Iles,

Imues, La Cruz, Tangua y Yacuanquer. Para estos municipios los índices de ruralidad y NBI son bastante cercanos así como la variable que representa el promedio de personas que habitan en cada vivienda. La producción de maíz, papa y trigo enlazan a estos municipios dentro de la región en la que se ubican.

Finalmente, la última región de la clasificación primaria, se clasifica en cinco *clusters* los cuales se muestran en la Figura 12. Dentro de los municipios que se encuentran en la rama de la clasificación primaria se tiene que aquí se ubican tres de las principales municipalidades del Departamento: San Juan de Pasto, capital de Nariño; Túquerres y el municipio fronterizo Ipiales. La forma en que son clasificados estos municipios permite explicar el criterio adoptado para presentar este modelo para la tipificación de los municipios por su demanda de energía.

Los tres son puntos de convergencia de personas y negocios de las regiones que los rodean. Las particularidades de esta región son compartidas por las municipalidades antes mencionadas a pesar de no estar ubicadas una junto a la otra. Ipiales y Túquerres pertenecen al mismo *cluster*, mientras San Juan de Pasto se ubica en otro diferente debido a que sus condiciones como capital de departamento influyen en sus características socioeconómicas diferenciándolo de los otros dos, tal como sucede en el caso de Tumaco dentro de la región a la que éste pertenece. A nivel de producción agrícola los tres municipios tienen en común la papa, la papa amarilla y la zanahoria; en explotación pecuaria se puede decir que los municipios presentan alta producción de leche. Por último, la diferencia que separa a Pasto de Túquerres e Ipiales son los índices socio-económicos, los cuales para el caso de Pasto presentan una marcada diferencia, no solo con los dos municipios en cuestión, sino con todos los del Departamento.

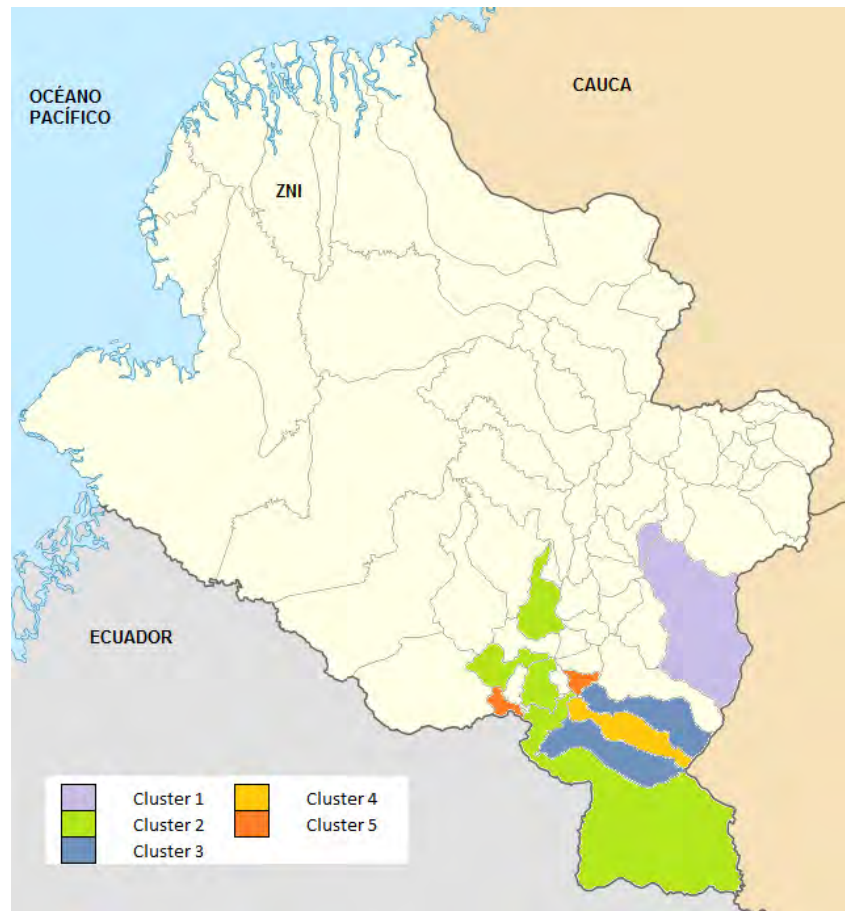


Figura 12. Clasificación de la Región 4 del *clustering* primario.

3. MODELAMIENTO

La predicción del consumo eléctrico se divide en dos partes, una para los municipios que cuentan con registro histórico de consumo, en la que se usa modelos ARIMA para realizar el pronóstico, y otra sección para los grupos (*clusters*) en donde se utiliza información de proyección de la población ¹⁰ para realizar el pronóstico del consumo y se comparará estos resultados con modelos ARIMA para observar las diferencias o similitudes obtenidas en la aplicación de ambos métodos.

¹⁰ <http://www.dane.gov.co/>

3.1. Modelamiento individual por municipio

Para la predicción del consumo eléctrico de los municipios, que cuentan con series de tiempo, se tiene en cuenta solo la variable de consumo energético, la información de consumo se encuentra organizada como se explica en la sección 4.1. Las series de tiempo que representan el consumo energético tienen notorios datos atípicos y se deben realizar acciones para eliminar dicha información errónea y obtener una mejor señal para su tratamiento. Una de las fallas más notorias en la información es que no se reportan datos de consumo ni de suscriptores para los meses de septiembre y noviembre del año 2012. Para poder realizar un tratamiento adecuado de las señales, se debe tener en cuenta su evolución y las componentes que la afectan; a continuación se muestra un resumen de éstas:

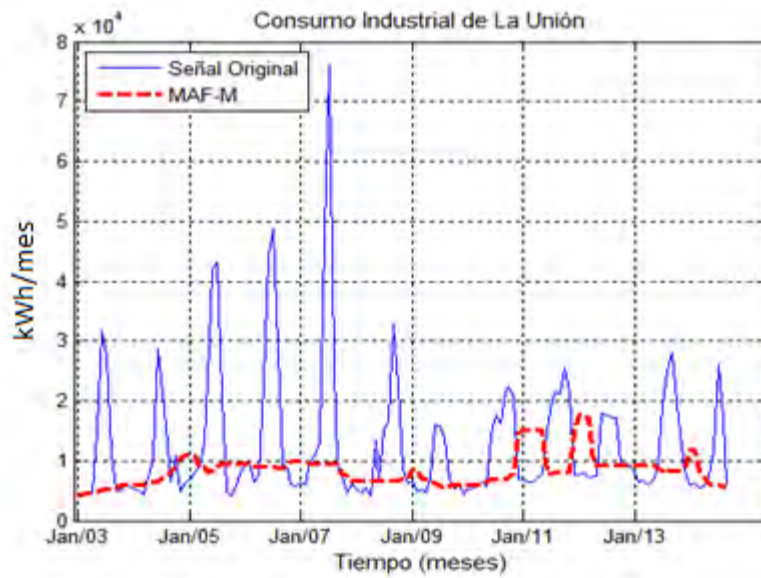
- Tendencia: patrón de comportamiento de largo plazo.
- *Seasonality*: patrón de repetición de frecuencia que puede ser anual, semestral, trimestral, mensual, etc.
- Componente estocástica irregular: porción imprevisible del comportamiento de la serie, no sigue ningún patrón.
- *Outliers*: datos atípicos que están por fuera de la tendencia.
- Cambios abruptos: cambios en las características que se producen muy rápido respecto al período de muestreo.
- Varianza no constante o heterocedasticidad: varianza de residuos no constante. (Fernández Sampedro, 2006)

3.1.1. Suavizado y descomposición

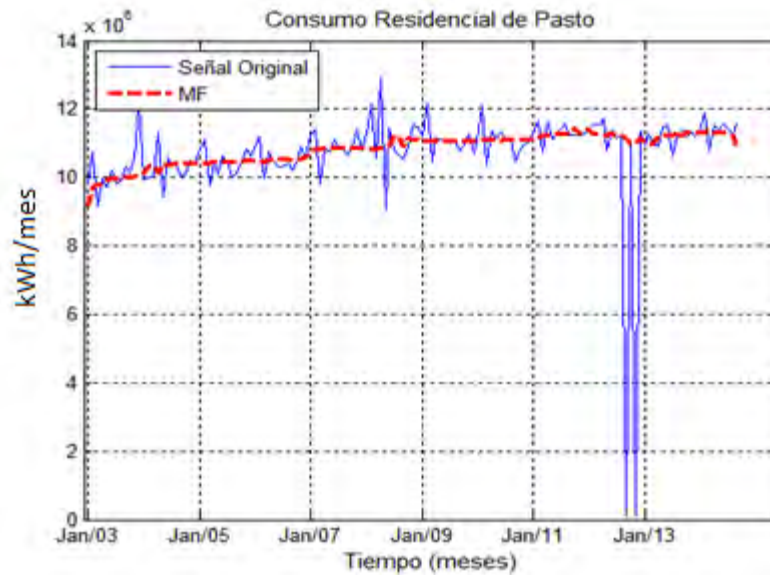
Las técnicas de suavización de las series de tiempo son uno de los primeros pasos en el tratamiento de las señales, el objetivo de suavizar las series es observar la tendencia, obtener información y visualizar los comportamientos tales como *seasonality* y eliminar cambios bruscos y *outliers*, el suavizado también ayuda a definir los posibles modelos para la predicción de la demanda. En este trabajo se usaron filtros de promedio móvil con ventanas de tipo variable, para series que no presentan *seasonality*, y filtros de promedio móvil con ventanas de tiempo variable que dependen de la longitud del patrón de repetición, para series con *seasonality*.

Se realizaron pruebas con cinco tipos de filtros: (1) filtro de promedio móvil MAF (*Moving Average Filter*), (2) filtro de promedio móvil trimestral MAF-Q (*Moving Average Filter - Quarterly*), (3) filtro de promedio móvil anual MAF-M (*Moving Average Filter - Monthly*), (4) filtro de promedio ponderado exponencial EWMAF (*Exponential Weighted Moving Average Filter*), y (5) filtro mediana MF (*Median Filer*) (Achicanoy Martínez, 2014). El diseño de los filtros se basa en la construcción de una serie de variables a partir de los datos originales, realizando para ello el promedio de estos datos, por ejemplo, la construcción del filtro MAF-M consiste en sustituir el dato de cada mes por el promedio de los datos de los últimos 12 meses, de esta manera para calcular el MAF-M para el mes de enero de 2010, se calcula el promedio de los doce meses comprendidos entre febrero de 2009 y enero de 2010. El filtro MF con ventana variable es de gran ayuda en la eliminación de *outliers*, éste se aplica por lo general a series sin patrones de repetición o patrones de repetición pequeños.

Cada tipo de filtro presenta ventajas como también desventajas y depende del tipo de serie y el objetivo que se quiera alcanzar aplicar uno de los filtros mencionados, por ejemplo si la serie en estudio no presenta *seasonality* muy marcada, podría utilizarse un filtro MF, en otros casos sería más conveniente utilizar un promedio móvil de mayor orden dependiendo del orden de la repetición, es decir si se observa una repetición anual es conveniente un filtro MAF-M, si la repetición es trimestral lo mejor es el filtro MAF.Q. En el caso del filtro MAF al seleccionar un orden de filtro grande se tendrá mayor influencia de los datos antiguos y en contrapartida un orden bajo dará mayor influencia a los datos recientes de la serie. El filtro EWMAF, se basa en la atenuación de los valores de la serie de tiempo, obteniendo el promedio de estos de manera exponencial; es decir, los datos se ponderan dando un peso mayor a las observaciones más recientes y uno menor a las más antiguas, por tanto este filtro es de utilidad cuando se quiere dar una alta ponderación a los datos más recientes de la serie. (Quispe Llanos, Técnicas Estadísticas de Predicción, 1998).



(a) Consumo industrial de La Unión-Filtro MAF-M.



(b) Consumo Residencial de Pasto-Filtro MF con tamaño de ventana 13.

Figura 13. Filtros aplicados a series de consumo. (a) Consumo industrial de La Unión-Filtro MAF-M. (b) Consumo residencial de Pasto-Filtro MF.

La Figura 13 (a) muestra el consumo industrial del municipio de La Unión. En esta serie se observan comportamientos de *seasonality* claramente definidos y se aplica un filtro MAF-M para obtener la tendencia la serie. En la Figura 13 (b) se dibuja el comportamiento del consumo residencial de Pasto al cual se le aplicó un filtro MF con tamaño de ventana igual a 13 (en

este filtro se elige generalmente un tamaño de ventana impar y no debe ser mayor a 13 que corresponde al número de meses que tiene un año más uno), se observan patrones de *seasonality* entre enero de 2005 y enero de 2011, aunque son de menor amplitud a los de la Figura 13 (a); el filtro muestra la tendencia del consumo, elimina los *outliers* y los cambios abruptos presentes en septiembre y noviembre de 2012, como se mencionó anteriormente esta información no se encuentra en los reportes del SUI.

El consumo industrial de La Unión presenta un comportamiento particular con *seasonality* muy marcada que no se ve reflejado de forma tan notoria en los consumos residenciales de los municipios de Nariño, por lo tanto se usará a manera de ejemplo para modelar su consumo. Para obtener mayor información acerca de los parámetros de *seasonality* de esta serie, se puede aplicar técnicas de descomposición que sirven para la identificación de los períodos de repetición asociados a comportamientos de *seasonality*; la descomposición puede ser por método aditivo para series con *seasonality* constante o por método multiplicativo para series con *seasonality* variable en el tiempo. Los pasos a seguir para realizar la descomposición de la serie son: (i) caracterización de la tendencia, (ii) eliminación de la tendencia a la serie original, (iii) extracción de los patrones de *seasonality* y (iv) evaluación de la componente irregular.

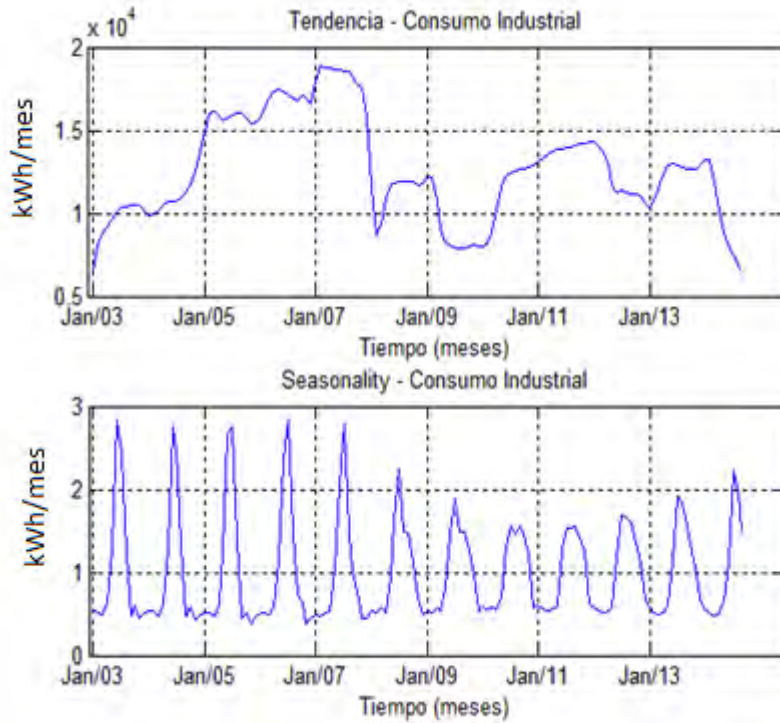
En la Figura 14 se muestra el resultado de la descomposición por multiplicación aplicada a la serie de consumo industrial de La Unión, se decide usar este tipo de descomposición ya que la *seasonality* no se mantiene constante. La descomposición por multiplicación se realiza de la siguiente manera:

$$y_t = T_t S_t I_t \quad (3)$$

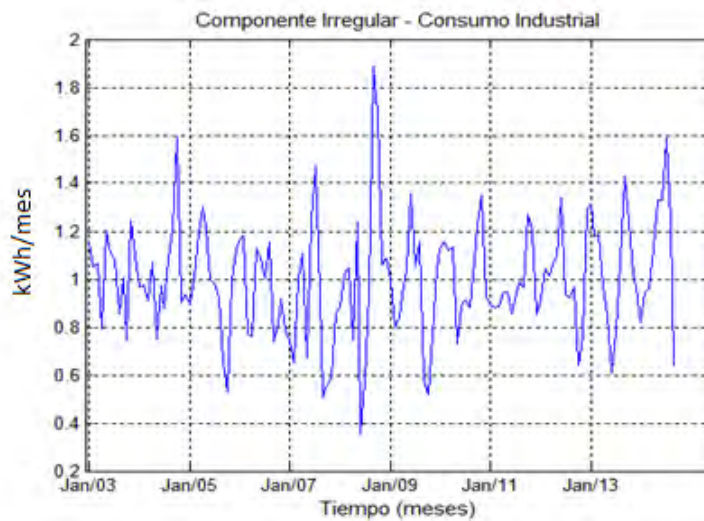
Donde y_t es la observación en el instante t , T_t es la tendencia determinística, S_t es la componente determinística de *seasonality* e I_t es la componente estocástica irregular.

En la Figura 13 (a) se puede ver la señal original del consumo industrial de La Unión, en la Figura 14 (a) se muestra la tendencia obtenida con el filtro MAF-M, y se traza la *seasonality*, en la cual se observa un ciclo de repetición cada 12 meses, es decir, tiene un patrón de *seasonality* anual y

en la Figura 14 (b) se mira la componente irregular de la serie que se despeja de la ecuación (3) obteniendo $\hat{I}_t = y_t / (\hat{T}_t \hat{S}_t)$.



(a) La Unión: Tendencia y *Seasonality*.



(b) La Unión: Componente irregular.

Figura 14. Descomposición multiplicativa de la serie de consumo industrial de La Unión. (a) Tendencia y *Seasonality*. (c) Irregular.

3.1.2. Modelo

Las series de consumo energético de los municipios de Nariño presentan diferentes comportamientos, como *seasonality*, tendencia, cambios abruptos y varianza no constante, por lo tanto para buscar un modelo que se ajuste a cada serie en particular se realiza la construcción de modelos ARIMA (*AutoRegresive Integrated Moving Average*), que deriva de sus tres componentes AR (Autoregresivo), I (Integrado) y MA (Medias Móviles). Este modelo requiere una sola variable, en este caso se usará el consumo de energía eléctrica en el tiempo. El modelo ARIMA permite describir un valor como una función lineal de datos anteriores y errores debidos al azar, además puede incluir un componente cíclico o *seasonality*. (Arce & Mahía, 2011). (de la Fuente Fernández, 2011).

Un modelo ARIMA (P, D, Q) permite describir una serie de observaciones después de que hayan sido diferenciadas D veces, a fin de extraer las posibles fuentes de no estacionariedad¹¹, toma los valores de la serie como los datos pasados y construye el modelo que mejor se ajuste a su comportamiento. Para series mensuales es conveniente contar por lo menos con 6 años completos de información (de la Fuente Fernández, 2011); la calidad del modelo depende totalmente de la calidad de los datos de la serie. Los parámetros P, D y Q son enteros no negativos que corresponden a las componentes AR, I y MA, respectivamente. (Uriel Jiménez, 1992)

La identificación del modelo consiste en encontrar los valores (P, D, Q) más apropiados a la serie temporal, si esta presenta tendencia, lo primero que se debe hacer, es convertirla en estacionaria mediante una diferenciación de orden D. Una vez diferenciada la serie, una buena estrategia consiste en comparar las funciones de autocorrelación ACF (*Autocorrelation Function*) y la función de autocorrelación parcial PACF (*Partial Autocorrelation Function*), aunque no existe un procedimiento general para obtener un modelo definitivo a partir de ACF y la PACF, en la mayoría de los casos se puede obtener la formulación de un posible

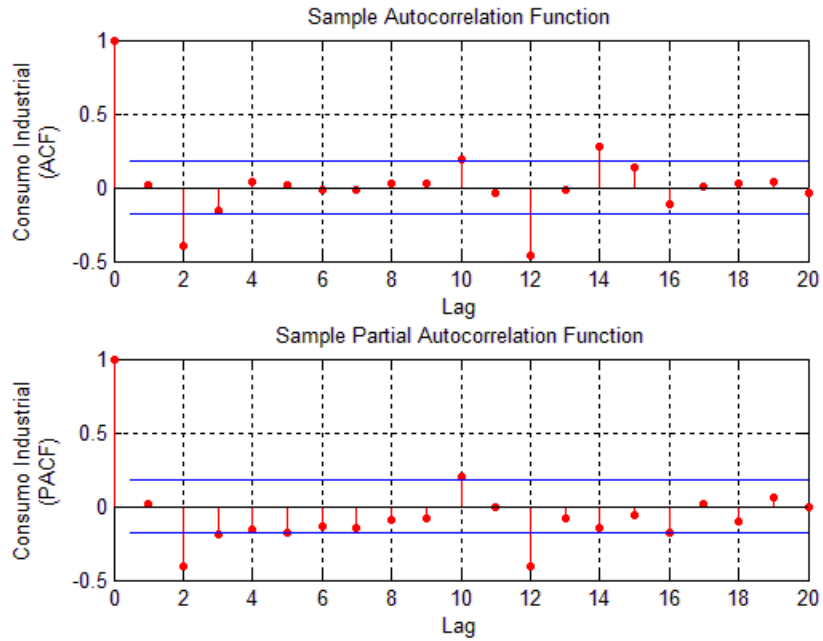
¹¹ Estacionariedad: Una serie de tiempo es fuertemente estacionaria, si su distribución conjunta es invariante en el tiempo (todos los momentos de la distribución no dependen del tiempo). Una serie de tiempo es débilmente estacionaria si la media, la varianza y la covarianza son independientes del tiempo.

modelo a través de los patrones característicos de estas gráficas. En la tabla 5 se muestra un resumen sobre los posibles modelos teniendo en cuenta los patrones combinados de ACF y PACF.

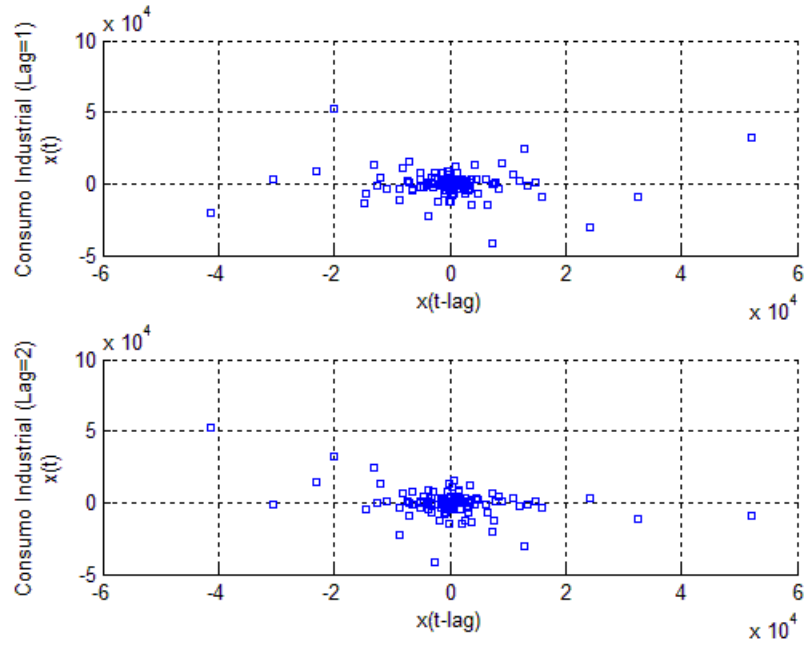
Se realizan las pruebas de la ACF y la PACF a la señal diferenciada del consumo industrial de La Unión, los resultados se observan en la Figura 15. Los patrones de la ACF y la PACF de la Figura 15 (a), sugieren un modelo con términos AR (12) y MA (12). La Figura 15 (b) muestra que no hay correlación entre las observaciones actuales y las observaciones con *lags* 1 y 2, que es lo que se requiere para una serie estacionaria, es decir, mediante la diferenciación se logró convertir la serie en una de tipo estacionaria.

Tabla 5. Patrones de ACF y PACF.

| ACF | PACF | Posible modelo |
|---|---|---------------------|
| Patrón sinusoidal o de decrecimiento que converge a cero. Alternación de signos positivos y negativos | Valores significativos en los primeros P <i>lags</i> | AR de orden P |
| Valores significativos en los primeros Q <i>lags</i> | Patrón sinusoidal o de decrecimiento que converge a cero. Alternación de signos positivos y negativos | MA de orden Q |
| Patrón sinusoidal o de decrecimiento que converge a cero. Alternación de signos positivos y negativos | Patrón sinusoidal o de decrecimiento que converge a cero. Alternación de signos positivos y negativos | Términos de AR y MA |



(a) ACF y PACF.



(b) x_t vs. x_{t-1} y x_t vs. x_{t-2} .

Figura 15. ACF, PACF, x_t vs. x_{t-1} y x_t vs. x_{t-2} de la señal diferenciada del consumo residencial de La Unión. (a) ACF y PACF. (b) x_t vs. x_{t-1} y x_t vs. x_{t-2} .

Se realizaron pruebas con tres tipos de modelos: (i) ARIMA (P, D, Q),

para series sin *seasonality*, (ii) ARIMA (P, D, Q) x (P_s, Q_s) (ARIMA multiplicativo para series con *seasonality*) y (iii) ARIMA (P, D, Q)/GARCH/ARCH (ARIMA compuesto) para series con varianza no constante. Se escogieron estos modelos debido a que se ajustan a los diferentes comportamientos de las series de Nariño. Continuando a manera de ejemplo con el consumo industrial de La Unión se muestran a continuación los resultados de la aplicación de los tres modelos ARIMA a dicha serie.

- ARIMA (P, D, Q): se realiza un test AIC/BIC (The Pennsylvania State University, Department of Statistics, 2013), para obtener los mejores valores de los parámetros del modelo, los cuales resultaron ser P=2, D=1, Q=1, es decir, el modelo presenta dos términos AR, una diferenciación, un término MA, y un término constante. Los parámetros se muestran en la Figura 16. De acuerdo a la estadística t, todos los parámetros excepto la constante son significativos al 5% (para ser significativos al 5% los coeficientes deben estar por fuera del intervalo crítico (-1.96, 1.96)).

| ARIMA(2,1,1) Model: | | | |
|--|-------------|----------------|-------------|
| ----- | | | |
| Conditional Probability Distribution: Gaussian | | | |
| Parameter | Value | Standard Error | t Statistic |
| ----- | ----- | ----- | ----- |
| Constant | 4.87996 | 7.50536 | 0.650197 |
| AR{1} | 1.08574 | 0.111438 | 9.74301 |
| AR{2} | -0.211599 | 0.0907904 | -2.33063 |
| MA{1} | -1 | 0.073707 | -13.5672 |
| Variance | 1.84463e+06 | 0.000231082 | 7.98258e+09 |

Figura 16. Parámetros del modelo ARIMA (2, 1, 1).

Se realiza la prueba de Ljung-Box Q de autocorrelación de los residuos y la prueba de Engle de heterocedasticidad de los residuos (The Pennsylvania State University, Department of Statistics, 2013), ambas para los *lags* 5, 10 y 15, los resultados se muestran en la tabla 6. La prueba de Ljung-Box Q, consiste en analizar si se presenta autocorrelación antes de cada *lag*, en dicho

caso se rechaza (valores de decisión igual a cero), en esta prueba la autocorrelación se rechaza solo para valores inferiores al *lag* 5. En la prueba de Engle, sí la varianza de los residuos no es constante antes de cada lag, se acepta (valores de decisión igual a uno), en este caso se acepta para los *lags* 10 y 15. Estos resultados indican que el modelo no se ajusta a la serie, ya se debe tener en cuenta los comportamientos de *seasonality*.

Tabla 6. Resultados de las pruebas Ljung-Box Q y Engle del modelo ARIMA (2, 1, 1).

| Variable | Ljung-Box Q (Autocorrelación) | | | Engle (heterocedasticidad) | | |
|----------|-------------------------------|---------|---------|----------------------------|---------|---------|
| Lag | 5 | 10 | 15 | 5 | 10 | 15 |
| Decisión | 0 | 1 | 1 | 0 | 1 | 1 |
| pValue | 0.1847 | 0.0001 | 0.0003 | 0.1642 | 0.0000 | 0.0012 |
| Stat | 7.5209 | 34.9271 | 41.5797 | 7.8584 | 38.5133 | 37.2013 |
| cValue | 11.0705 | 18.3070 | 24.9958 | 11.0705 | 18.3070 | 24.9958 |

- ARIMA Multiplicativo (P, D, Q) x (P_s, Q_s) : las pruebas AIC/BIC arrojaron como mejores resultados los parámetros; P=2, D=1, Q=2, P_s=12, Q_s=12, período del patrón de repetición s=12, los parámetros del modelo se muestran en la Figura 17, el cual presenta dos términos AR, dos términos MA, una diferenciación y no incluye constante, según la estadística *t* los términos son significativos al 5%, excepto del patrón de *seasonality* relacionado al termino AR.

```

ARIMA(2,1,2) Model Seasonally Integrated with Seasonal AR(12) and MA(12) :
-----
Conditional Probability Distribution: Gaussian

Parameter      Value      Standard      t
              Value      Error      Statistic
-----
Constant      0          Fixed      Fixed
AR{2}         0.565083   0.267742   2.11055
SAR{12}       0.118464   0.143348   0.82641
MA{2}         -0.814567   0.213047   -3.82341
SMA{12}       -0.880122   0.136719   -6.43745
Variance      2.60103e+06 250788     10.3715
    
```

Figura 17. Parámetros del modelo ARIMA (2, 1 ,2) x (12, 12).

Se realizan las pruebas de autocorrelación y heterocedasticidad, ambas para los *lags* 5, 10 y 15, los resultados se plasman en la

tabla 7, de la cual se lee, que la prueba de autocorrelación se rechaza solo para valores inferiores al *lag* 5, lo cual indica según este modelo una posible *seasonality* cada 4 meses. La prueba de heterocedasticidad indica que la varianza es constante solo para el *lag* 5. Aunque no se logra obtener varianza constante para todos los *lags* y que no exista autocorrelación en los residuos, este modelo tiene en cuenta la *seasonality* de la serie, lo cual puede influir en una mejor predicción de la serie que el modelo anterior.

- ARIMA Compuesto (GARCH/ARCH): los parámetros obtenidos con la prueba AIC/BIC para este modelo fueron; ARIMA (1, 1, 2)/GARCH (2, 1)/ARCH (2, 1), los parámetros se muestran en la Figura 18, los valores de los coeficientes resultan significativos al 5%, a excepción del término de la constante del modelo.

| ARIMA(1,1,2) Model: | | | |
|--|------------|----------------|-------------|
| ----- | | | |
| Conditional Probability Distribution: Gaussian | | | |
| Parameter | Value | Standard Error | t Statistic |
| ----- | | | |
| Constant | -0.144672 | 3.11009 | -0.046517 |
| AR{1} | 0.844683 | 0.0203167 | 41.5759 |
| MA{1} | -0.736691 | 0.120632 | -6.10692 |
| MA{2} | -0.263309 | 0.123086 | -2.13923 |
| | | | |
| GARCH(2,1) Conditional Variance Model: | | | |
| ----- | | | |
| Conditional Probability Distribution: Gaussian | | | |
| Parameter | Value | Standard Error | t Statistic |
| ----- | | | |
| Constant | 2865.98 | 1.43021e-05 | 2.00389e+08 |
| GARCH{2} | 0.991195 | 0.0829714 | 11.9462 |
| ARCH{1} | 0.00880507 | 0.0111187 | 0.791915 |

Figura 18. Parámetros del modelo ARIMA (1, 1, 2)/GARCH (2, 1)/ARCH (2, 1).

Se realizan las pruebas de autocorrelación y heterocedasticidad, ambas para los *lags* 5, 10 y 15, los resultados se plasman en la tabla 8, de la cual se lee, que las pruebas de autocorrelación y

heterocedasticidad se rechazan antes de los *lags* evaluados, es decir no se presenta autocorrelación y la varianza es constante en todos los *lags*, por lo tanto, el modelo con mejor ajuste para la serie de consumo industrial del Municipio de La Unión, es el ARIMA Compuesto, que tiene en cuenta la varianza no constante de la serie.

Se realiza una prueba AIC/BIC a los tres modelos para comprobar cuál ostenta mejor desempeño para este caso. La prueba Criterio de Información de Akaike (AIC) sirve para medir la calidad relativa de un modelo estadístico y la prueba Criterio de Información Bayesiano (BIC) es un criterio para la selección de modelos entre un conjunto finito (Akaike, 1974).

Los resultados se muestran en la tabla 9, se escoge el que presente menores valores AIC/BIC, que resulta ser, como se esperaba según lo expuesto anteriormente, el modelo ARIMA Compuesto. Aunque el modelo ARIMA Multiplicativo (2, 1, 2) x (12, 12) tiene en cuenta la *seasonality* de la serie, se obtiene mejores resultados, según la tabla 9, con un ARIMA (2, 1, 1), el cual no tiene en cuenta la *seasonality*. Las pruebas para los tres modelos se realizaron bajo las mismas condiciones, es decir, se aplicó un filtro MF, con tamaño de ventana igual a 13 para realizar la estimación de los tres modelos, se puede cambiar el tamaño de la ventana o el tipo de filtro para cada modelo, en busca de mejores resultados.

Tabla 7. Resultados de las pruebas Ljung-Box Q y Engle del modelo ARIMA (2, 1, 2) x (12, 12).

| Variable | Ljung-Box Q (Autocorrelación) | | | Engle (heterocedasticidad) | | |
|-----------------|-------------------------------|---------|---------|----------------------------|---------|---------|
| | 5 | 10 | 15 | 5 | 10 | 15 |
| Lag | 5 | 10 | 15 | 5 | 10 | 15 |
| Decisión | 0 | 1 | 1 | 0 | 1 | 1 |
| pValue | 0.4825 | 0.0041 | 0.0329 | 0.5035 | 0.0021 | 0.0209 |
| Stat | 4.4806 | 25.7706 | 26.5159 | 4.3257 | 27.5665 | 28.1089 |
| cValue | 11.0705 | 18.3070 | 24.9958 | 11.0705 | 18.3070 | 24.9958 |

Tabla 8. Resultados de las pruebas Ljung-Box Q y Engle del modelo ARIMA (1, 1, 2) /GARCH (2, 1)/ARCH (2, 1).

| Variable | Ljung-Box Q (Autocorrelación) | | | Engle (heterocedasticidad) | | |
|-----------------|-------------------------------|---------|---------|----------------------------|---------|---------|
| | 5 | 10 | 15 | 5 | 10 | 15 |
| Lag | 5 | 10 | 15 | 5 | 10 | 15 |
| Decisión | 0 | 0 | 0 | 0 | 0 | 0 |
| pValue | 0.5196 | 0.8659 | 0.9662 | 0.9911 | 0.9996 | 1.0000 |
| Stat | 4.2101 | 5.3602 | 6.6694 | 0.5263 | 1.1931 | 2.0684 |
| cValue | 11.0705 | 18.3070 | 24.9958 | 11.0705 | 18.3070 | 24.9958 |

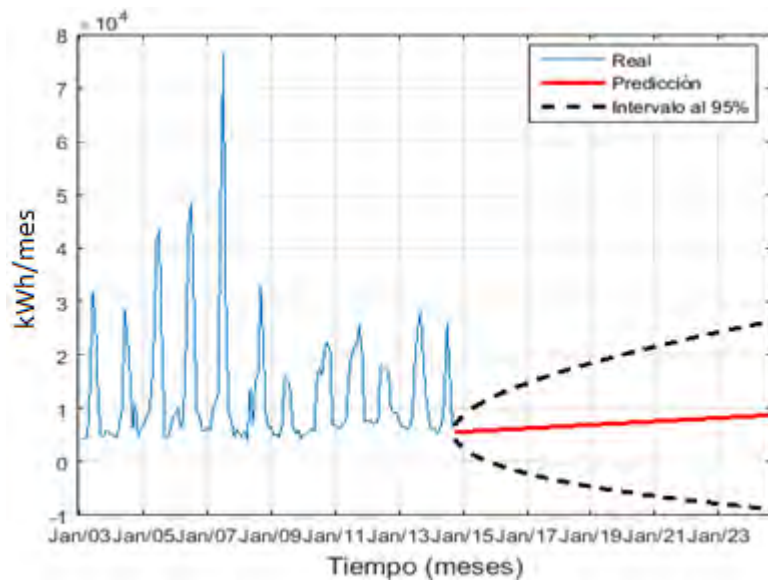
Tabla 9. Pruebas AIC/BIC.

| Variable | ARIMA (P, D, Q) | ARIMA Multiplicativo | ARIMA Compuesto |
|------------|-----------------|----------------------|-----------------|
| AIC | 2432 | 2450 | 2241 |
| BIC | 2447 | 2465 | 2262 |

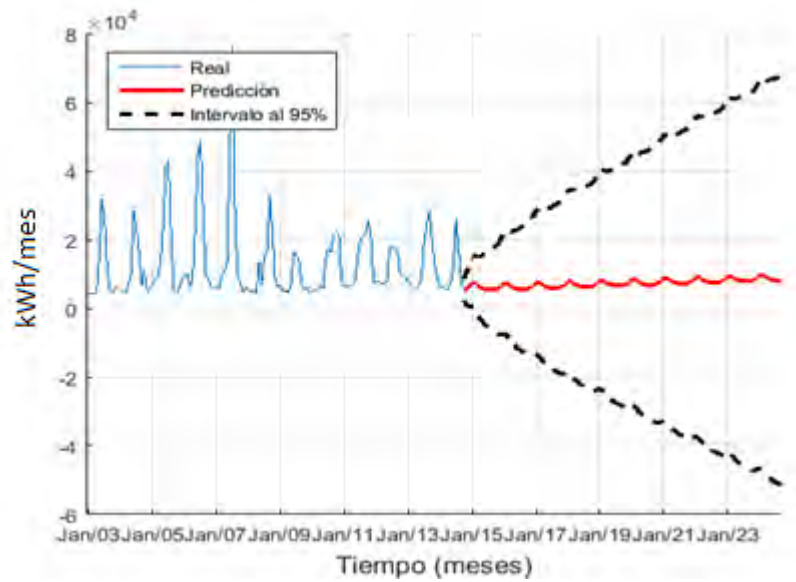
3.1.3. Pronóstico

Debido a la periodicidad mensual de la información, se puede estimar un modelo en que el corto plazo es un mes (Benavente, Galetovic, Sanhueza, & Serra, 2005). A manera de ejemplo se presenta, en la Figura 19, el pronóstico del consumo industrial de La Unión a 120 meses con cada uno de los tres modelos expuestos, el número de meses que se elige para la predicción no influye en la construcción del modelo. En general, para elegir el modelo más robusto, se escoge el que presente menores valores AIC/BIC.

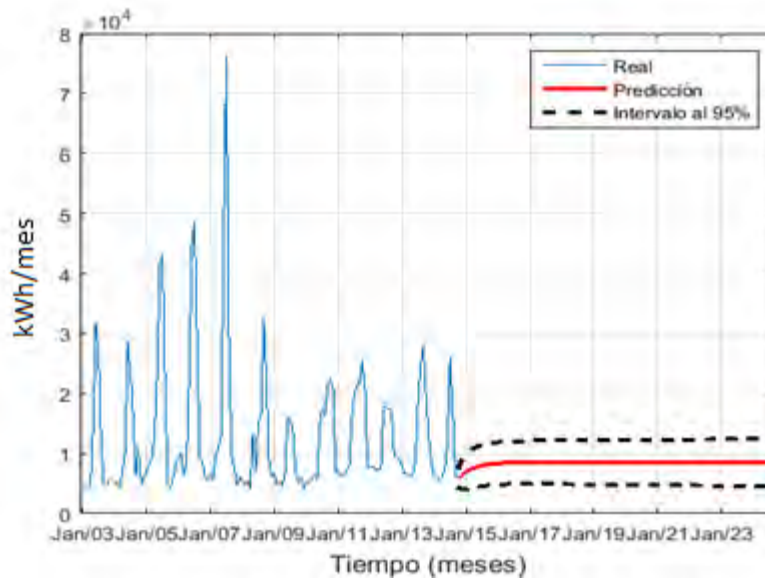
Como se mencionó anteriormente, el mejor modelo para el consumo industrial de La Unión, es el ARIMA Compuesto, el cual se lo puede ver en la Figura 19 (c), donde se muestra cuál ha sido el consumo real, la predicción y el intervalo de confianza al 95%. Efectivamente la Figura 19 (c) muestra de mejor manera la tendencia del consumo, en comparación con las Figuras 19 (a) y (b), y como lo indica la tabla 9, aunque el modelo ARIMA Multiplicativo tiene en cuenta la *seasonality* se obtuvo un mejor resultado con el ARIMA (P, D, Q).



(a) Predicción consumo industrial de La Unión con ARIMA (2, 1, 1).



(b) Predicción consumo industrial de La Unión con ARIMA (2, 1, 2) x (12, 12).



(c) Predicción consumo industrial de La Unión con ARIMA (1, 1 ,2)/GARCH (2, 1)/ARCH (2, 1).

Figura 19. Pronóstico a 120 meses del consumo industrial de La Unión. (a) ARIMA (2, 1, 1). (b) ARIMA (2, 1, 2) x (12, 12). (c) ARIMA (1, 1 ,2)/GARCH (2, 1)/ARCH (2, 1).

3.2. Modelamiento por *Clusters*

Una vez se cuenta con la clasificación de los municipios agrupados en *clusters*, se procede a realizar la predicción del consumo de energía eléctrica para los mencionados *clusters*, para esto se tienen en cuenta tres variables principales: población, número de suscriptores y consumo de energía eléctrica, estas variables están representadas en series de tiempo. Las variables de consumo energético y suscriptores se encuentran organizadas como se explica en la sección 4.1. La variable población contiene información anual del número de habitantes de cada municipio de Nariño desde el año 1985 hasta el año 2020; para los años 1985-2005 las series de población se obtienen por estimación y de 2005 a 2020 por proyecciones realizadas por el DANE teniendo en cuenta el censo realizado en 2005¹². Debido a que las series de consumo energético y suscriptores poseen información a partir del año 2003, la información de

¹² www.dane.gov.co

población se tiene en cuenta a partir de ese mismo año y se descarta información anterior a dicho año.

3.2.1. Relación entre población y suscriptores residenciales

El primer paso consistió en realizar una regresión entre la población y el número de suscriptores residenciales en cada *cluster*, con el fin de conocer cómo se relacionan estas dos variables, es decir, cómo se ha visto afectado el número de suscriptores respecto al crecimiento poblacional. Una vez se obtiene la relación entre población y suscriptores, se procede a evaluar cómo será el comportamiento del número de suscriptores hasta el año 2020, esto debido a que se cuenta con la predicción de la población hasta ese año como se explicó antes, de esta manera se obtiene una predicción a mediano plazo (7 años), a partir de las proyecciones poblacionales.

A continuación se describe el proceso realizado paso a paso para la clasificación primaria 1, *cluster* 5, el proceso es el mismo para todos los *clusters* de cualquier de las regiones de la clasificación primaria. Este *cluster* contiene los siguientes municipios: La Unión, San Pablo y Sandoná. Para el tratamiento de la información de suscriptores y población, se utiliza el promedio los municipios contenidos en el *cluster*.

Se hicieron pruebas con tres tipos de regresiones: lineal, exponencial y cuadrática, para escoger la regresión que mejor se ajuste a las variables se realiza una comparación entre los errores cuadráticos medios RMSE (*Root Mean Square Deviation*) y se selecciona la de menor RMSE.

Al realizar las regresiones se encontró que la mejor relación, es decir, la de menor RMSE, entre los suscriptores residenciales y la población para este *cluster*, es una ecuación lineal, ésta se dibuja en la Figura 20, en la cual se observa la relación punto a punto entre los suscriptores y la población, y la curva resultado de la regresión. La ecuación se dibuja en la Figura 21, en esta se muestran los valores de los coeficientes p con 95% de límites de confianza, la variable x representa la población, con esta ecuación se procede a evaluar cuál será el comportamiento de los suscriptores residenciales en un tiempo de 18 años (2003-2020), teniendo en cuenta que se conoce la proyección de la población para dicho periodo, como se indica en la Figura 22.

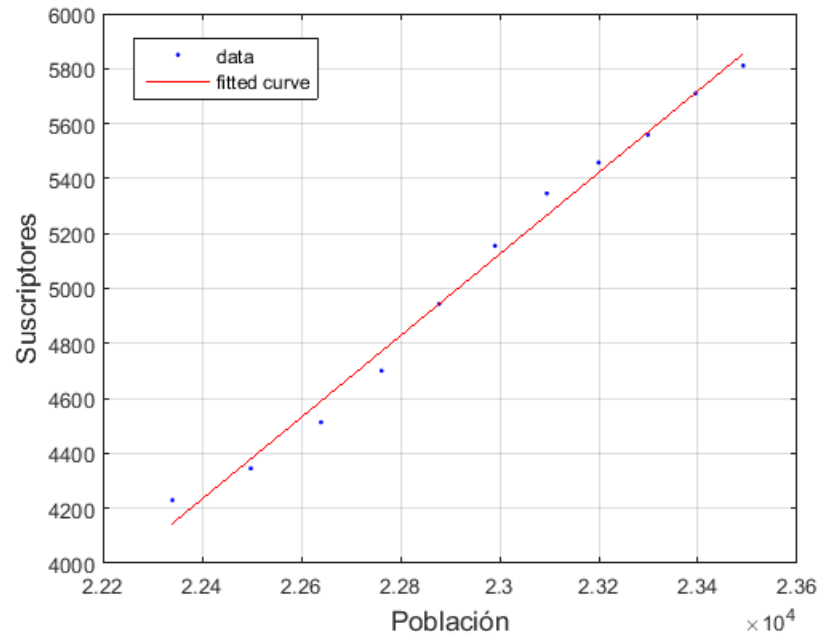


Figura 20. Ecuación lineal entre población y suscriptores residenciales.

```

Suscriptores =

Linear model Poly1:
Suscriptores(x) = p1*x + p2
  where x is normalized by mean 2.296e+04 and std 375.3
Coefficients (with 95% confidence bounds):
p1 =      557.6  (510.7, 604.5)
p2 =      5069  (5025, 5114)

```

Figura 21. Coeficientes de la ecuación lineal entre población y suscriptores residenciales.

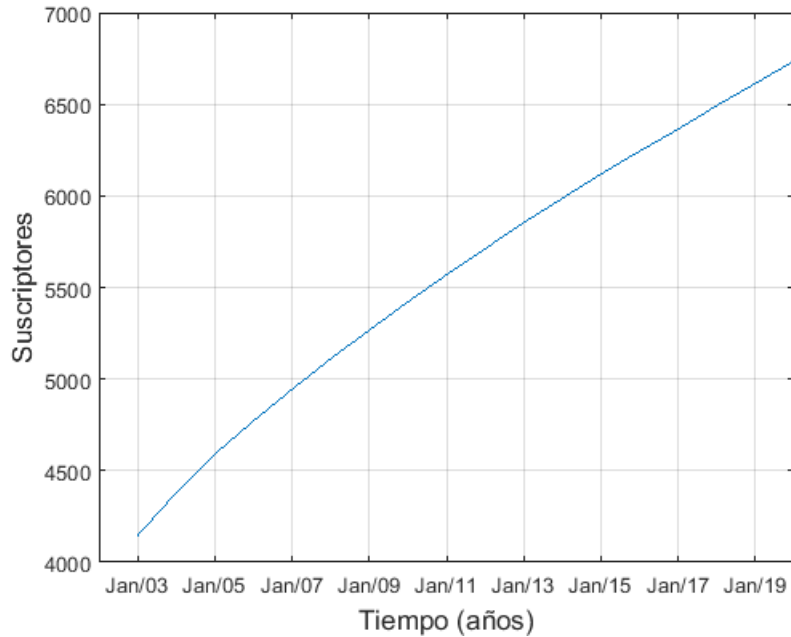


Figura 22. Comportamiento de los suscriptores residenciales en el tiempo.

Para este *cluster* se observa un comportamiento creciente de los suscriptores en el tiempo, al pasar de aproximadamente 4000 suscriptores residenciales en enero de 2003 a cerca de 7000 para enero de 2020. Desde enero de 2003 hasta diciembre de 2013 se muestra en la Figura 22 la relación entre suscriptores y población, y a partir de enero de 2014, se muestra la proyección teniendo en cuenta la población para los años 2014-2020. Se realiza esta proyección al reemplazar el valor de x en la ecuación de la Figura 21 con la población para dichos años.

3.2.2. Relación entre suscriptores y consumo residencial

El segundo paso, consiste en realizar una regresión entre los suscriptores y el consumo residencial, esta regresión se realiza con la información mensual de suscriptores y consumo de enero 2003 a diciembre de 2013, es decir se cuenta con 132 datos, una vez obtenida la ecuación que describe el comportamiento se evalúa cuál será la proyección para los próximos 7 años (hasta 2020), ya que se cuenta con esta información tal como se describió en la sección anterior.

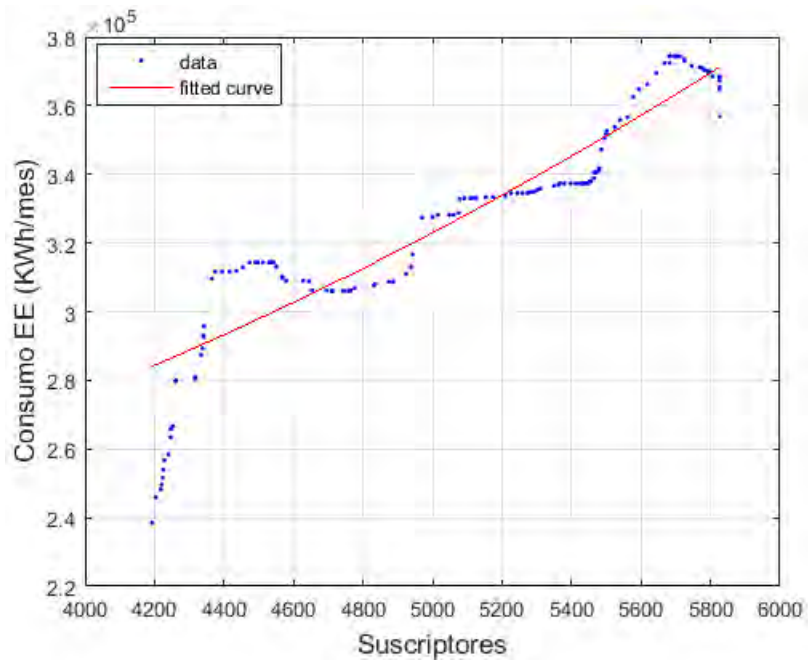


Figura 23. Ecuación cuadrática entre suscriptores y consumo residencial.

En este caso se obtuvo una regresión cuadrática entre suscriptores y consumo para el *cluster* en estudio, la cual se dibuja en la Figura 23. Los coeficientes de la regresión se muestran en la Figura 24, donde x representa los suscriptores residenciales. En esta ecuación se reemplaza el valor de x por el valor de los suscriptores obtenidos en la sección anterior, teniendo en cuenta el crecimiento poblacional, y se logra de esta manera obtener una proyección del consumo energético hasta el año 2020.

El comportamiento del consumo residencial se muestra en la Figura 25, donde se divide con una línea punteada el consumo en enero de 2014, donde empieza la proyección teniendo en cuenta, el número de suscriptores obtenidos para estos años.

```

Consumo =

Linear model Poly2:
Consumo(x) = p1*x^2 + p2*x + p3
where x is normalized by mean 5070 and std 537.2
Coefficients (with 95% confidence bounds):
p1 =      1790  (-629.6, 4210)
p2 =  2.907e+04 (2.716e+04, 3.098e+04)
p3 =  3.268e+05 (3.238e+05, 3.298e+05)

```

Figura 24. Coeficientes de la ecuación cuadrática entre suscriptores y consumo residencial.

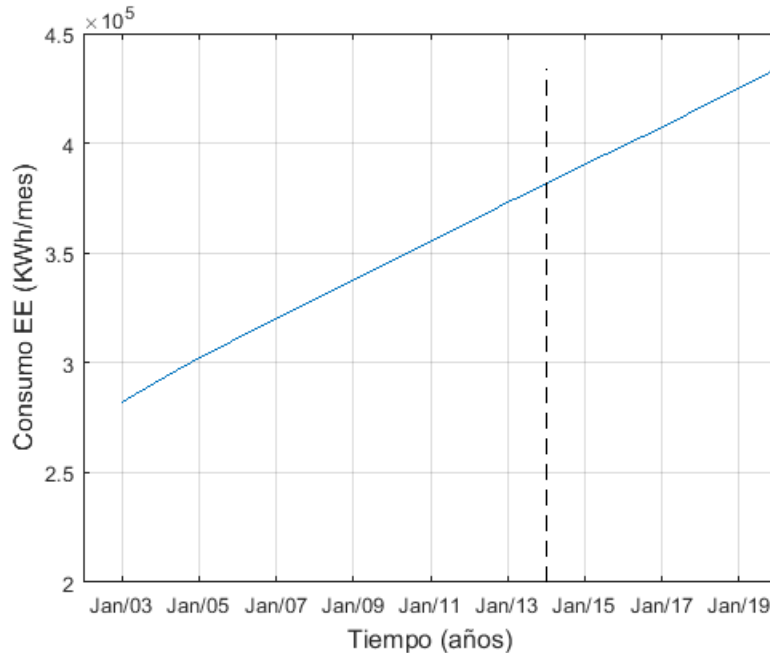


Figura 25. Comportamiento del consumo energético en el tiempo.

Se realiza la predicción para el *cluster*, usando un modelo ARIMA (2, 1, 1), ésta se muestra en la Figura 26, también se dibuja la predicción obtenida por medio de los suscriptores y la población. La predicción por medio del modelo ARIMA se realizó a 14 años, durante los 7 primeros años de esta predicción (igual tiempo de proyección que el anterior método), se observa que su comportamiento es muy similar. Se puede usar el primer método para predicciones de hasta máximo 7 años, que es el tiempo del cual se tiene información de la población, y para realizar un pronóstico a largo plazo, se puede utilizar cualquiera de los modelos ARIMA expuestos

en la sección 3.1.2.

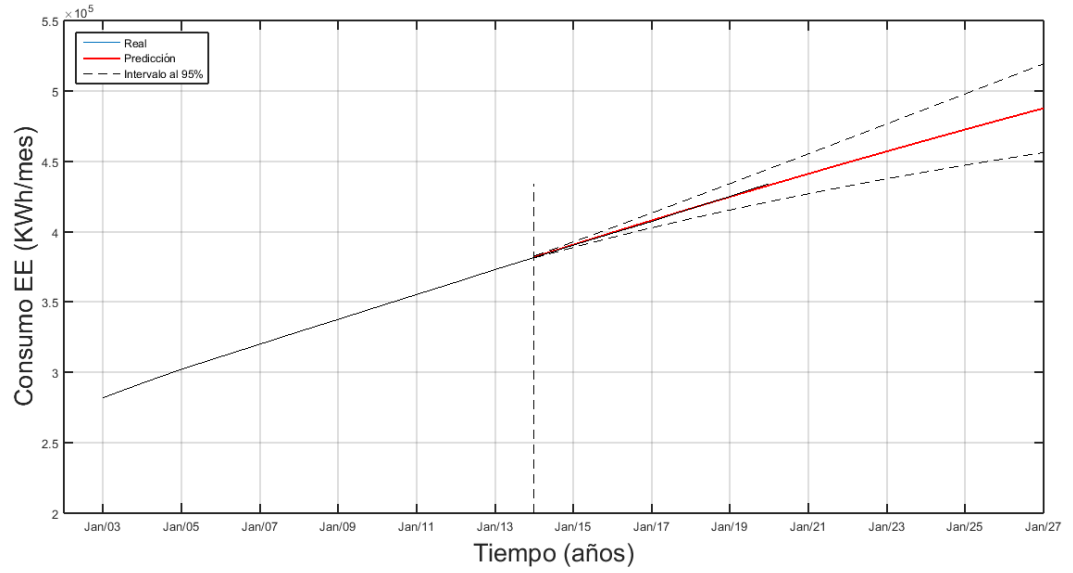


Figura 26. Consumo energético histórico, proyección con un modelo ARIMA (2, 1, 1) y los intervalos de confianza y la proyección obtenida por suscriptores y población.

4. DISCUSIÓN Y CONCLUSIONES

La demanda de energía eléctrica es un campo de estudio bastante amplio y que ha sido abordado desde diferentes áreas del conocimiento. Al revisar la literatura, se hace notorio que para trabajar en este tema es necesario hacerlo desde un ángulo particular. La presente investigación se enfoca en las herramientas provistas por métodos de inteligencia artificial, pues son muchos los estudios que hablan sobre sus ventajas sobre otros métodos, los estadísticos (regresión, series de tiempo y/o econométricos) por ejemplo.

Las herramientas computacionales, más específicamente, métodos de inteligencia artificial permiten clasificar a las poblaciones del departamento de Nariño a través de la implementación de diferentes algoritmos los cuales presentan resultados muy interesantes al introducir variables de diferente naturaleza, en este caso variables económicas, sociales y relacionadas con el consumo de energía eléctrica.

La clasificación de los municipios de Nariño, tomada como referencia inicialmente en esta investigación se realizó a través del Plan Nacional de Desarrollo (PND) en el año 2010. Dicha clasificación se realizó para facilitar procesos administrativos del Departamento. Principalmente, el agrupamiento obedece a características geográficas. A diferencia del proceso realizado en el PND, la clasificación de los municipios de Nariño según su demanda eléctrica tiene en cuenta las marcadas diferencias entre ellos. Se determinó que para la finalidad de este trabajo, las variables más significativas que muestran influencia en la demanda son: producción agrícola, la producción pecuaria, necesidades básicas insatisfechas (NBI), ruralidad y miseria. Además, se considera la cantidad promedio de personas que habitan en cada casa. Los resultados de la clasificación distan de la referencia debido a que muchos municipios geográficamente vecinos no comparten características socioeconómicas, las cuales influyen diversos comportamientos de las poblaciones de cada municipio, entre ellas, la demanda de energía eléctrica.

Un caso de estudio interesante que debe considerarse, no solo en el aspecto de demanda eléctrica; son los municipios ubicados en la Llanura del Pacífico, en muchos de los resultados de clasificación puede

observarse la conformación de uno o más grupos en los que los municipios de la mencionada región son puestos en clusters marcadamente diferenciados del resto del departamento. Especialmente, es el caso de Magüi Payán y Roberto Payán. Para el caso de esta investigación, ésta situación puede presentarse por las condiciones de escasa infraestructura y lejanía entre las cabeceras que conforman la región, además de la escasa o inexistente malla vial sumado a la situación de orden público por la que atraviesa esta zona del departamento.

El principal inconveniente al modelar registros históricos de consumo consiste en la calidad de la información, si no se cuenta con una serie confiable y con cantidad suficiente de información el modelamiento del consumo resulta ineficiente. En este sentido cabe aclarar que la base de datos de consumo energético descargada de la plataforma del SUI fue actualizada este año, y se debió descargar de nuevo toda la información recopilada, ya que la nueva información presentaba diferencias con la que se había descargado anteriormente.

El modelamiento del consumo energético por *clusters* incluye variables como el número de suscriptores y la población que permiten una mejor aproximación a su comportamiento, y por ende una proyección más acertada. Además, con la información agrupada en *clusters* se facilita modelar el comportamiento para cualquier municipio de Nariño, aunque éste no pertenezca a la zona interconectada, buscando el *cluster* que mejor se ajuste a las particularidades del municipio o región en estudio.

La predicción de la demanda eléctrica ha sido abordada desde distintos ángulos, en algunos casos se ha tenido en cuenta sólo una variable, estudiando solamente el consumo energético en el tiempo, y en otros casos se ha incluido variables exógenas, por ejemplo variables socio-económicas, y se han utilizado distintos métodos, como los estadísticos o tradicionales, y métodos más avanzados como los de inteligencia artificial, que presentan un desempeño más robusto en la modelación y predicción. En la bibliografía se encontró modelamientos de consumo con redes neuronales artificiales, sin embargo la información utilizada, corresponde a periodos de tiempo de 24 horas, por lo tanto, se pueden hacer modelos de consumo diario. Se encontró también métodos estadísticos, en los que usaron variables exógenas como el PIB, la inflación, el crecimiento poblacional y el consumo de diferentes fuentes de energía, e inclusive se

tuvo en cuenta la huelga del Cerrejón en 2013, para realizar el modelamiento y la predicción.

En este trabajo se planteó realizar una clasificación con variables sociales y económicas que puedan caracterizar los municipios en subregiones, por medio de métodos de inteligencia artificial, y realizar modelos de demanda que se ajusten a los comportamientos particulares de las clasificaciones obtenidas. La modelación del consumo se realizó con métodos estadísticos, debido a la escasa confiabilidad de la información, si se contara con un período de muestreo por horas, se podría realizar modelos por métodos de inteligencia artificial, con el fin obtener un modelo de demanda no solo mensual, sino también semanal y diaria, diferenciada entre fines de semana, festivos y días laborales. Además no se incluyó variables como el PIB y la inflación para el modelamiento, como lo hace la UPME, ya que esta información se encuentra sólo para el Departamento de Nariño, y no segregada por municipios, por lo tanto, esa información se podría incluir en un trabajo futuro para realizar un modelo general para el Departamento.

5. RECOMENDACIONES Y TRABAJO FUTURO

Uno de los mayores inconvenientes que se presentaron para la realización de este trabajo fue la escasa y, en algunos casos, poco confiable información. Las series de tiempo que corresponden al consumo y a la cantidad de suscriptores, presentan datos atípicos y en periodos de tiempo poco eficientes para la realización de las pruebas que este tipo de estudio demandan. Un sistema más eficiente de medición que involucre el manejo de la información en tiempo real puede facilitar los procesos de modelamiento.

Proyectos como el de “Análisis de Oportunidades Energéticas en el Departamento de Nariño” (ALTERNAR) o el “Plan de Energización Rural Sostenible” (PERS) de Nariño, involucran un trabajo de acercamiento a las comunidades desde el cual se puede obtener información relevante para el cálculo de las variables de tipo socioeconómico con el cual se pueda construir un modelo más completo. Así mismo, un experto en modelos econométricos puede sugerir la inclusión de más parámetros, con los que pueda construirse modelos multivariados.

REFERENCIAS

- Achicanoy Martínez, W. O. (2014). *Modelo de demanda de energía eléctrica para El Departamento de Nariño (Enfoque univariable clásico)*. San Juan de Pasto: Plan de Energización Rural Sostenible para el Departamento de Nariño PERS-Nariño.
- Akaike, H. (Diciembre de 1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, AC-19(6), 716-723.
- Alpaydin, E. (2010). *Introduction to Machine Learning*. London: The MIT Press.
- Arriaza, H. (2005). *Metodología para la inclusión de variables sociales en la formulación, ejecución y administración de proyectos de energía rural*. Quito: OLADE.
- Benavente, J. M., Galetovic, A., Sanhueza, R., & Serra, P. (2005). ESTIMANDO LA DEMANDA RESIDENCIALPOR ELECTRICIDAD EN CHILE: EL CONSUMO ES SENSIBLE AL PRECIO. *Cuadernos de Economía*, 42, 31-61.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Birmingham: Clarendon press - Oxford.
- Castillo Burbano, Á. M., & Jurado, J. A. (2014). *Caracterización social y económica del Departamento de Nariño*. San Juan de Pasto: Plan de Energización Rural Sostenible del Departamento de Nariño PERSN.
- Centro de Estudios de Desarrollo Regional y Empresarial CEDRE. (2012). *Indicadores socio-económicos Departamento de Nariño*. Pasto.
- de la Fuente Fernández, S. (2011). *SERIES TEMPORALES: MODELO ARIMA*. Madrid: Universidad Autónoma de Madrid.
- Departamento Administrativo Nacional de Estadística DANE. (7 de Abril de 2015). *Estadísticas Sociales: Necesidades Básicas Insatisfechas*. Obtenido de www.dane.gov.co
- Egelioglu, F., Mohamad, A., & Guven, H. (2001). Economic Variables and Electricity Consumption in Northern Cyprus. *Energy*, 26(4), 355-362.
- Fernández Sampedro, J. Á. (2006). *Descomposición tradicional de series temporales*. Madrid: Universidad Autónoma de Madrid.
- Garre, M., Cuadrado, J. J., & Sicilia, M. Á. (2005). *Comparación de diferentes algoritmos de clustering en la estimación de coste en el desarrollo de software*. Madrid: Universidad de Alcalá.

- Garrido Puerto, A. (2011). *MODELO DE PREVISIÓN GEOGRÁFICO Y TEMPORAL DE LA DEMANDA ELÉCTRICA*. Madrid.
- Gobernación de Nariño. (2004). *Informe de Gestión*. San Juan de Pasto.
- Gobernación de Nariño. (20 de 11 de 2014). *Sistema de Trazabilidad Agrícola*. Obtenido de <http://www.trazasnarino.org>
- Gurrutxaga Goikoetxea, I. (2010). *Aportaciones a la clasificación no supervisada y a su validación. Aplicación a la seguridad informática*. Donostia.
- Llanos, D., Azuaje, A., & Romero, I. M. (4 de Junio de 2013). *Redes Neuronales*. Obtenido de <http://redesneuronal.blogspot.com/>
- MathWorks Inc. (2015). *MathWorks*. Obtenido de <http://www.mathworks.com>
- Medina Hurtado, S., & García Aguado, J. (2005). Protección de la demanda de energía en Colombia mediante un sistema de inferencia difuso neuronal. *Energética*(33), 15-24.
- Ministerio de Comercio, Industria y Turismo. (2013). *Departamento de Nariño*. Bogotá D.C.: Oficina de estudios económicos.
- Oropeza Rodríguez, J. L. (2006). Algoritmos y métodos para el reconocimiento de voz en español mediante sílabas. *Computación y Sistemas*, 270-286.
- Penn State. (s.f.). *psu*. (Eberly College of Science) Recuperado el Septiembre de 2014, de <https://onlinecourses.science.psu.edu/stat510/>
- Quispe Llanos, R. (1998). *Técnicas Estadísticas de Predicción*. Lima: Universidad Inca Garcilaso de la Vega.
- Rojas Zerpa, J. C. (2012). *Planificación del suministro eléctrico en áreas rurales de los países en vías de desarrollo: un marco de referencia para la toma de decisiones*. Zaragoza: Universidad de Zaragoza.
- Rueda, M., Velásquez Henao, J., & Franco Cardona, C. (2011). AVANCES RECIENTES EN LA PREDICCIÓN DE LA DEMANDA DE ELECTRICIDAD USANDO MODELOS NO LINEALES. *Dyna*(167), 36-43.
- Secretaría de Agricultura y Medio Ambiente - Gobernación de Nariño. (2014). *TRAZAS Sistema de Trazabilidad Agrícola*. Recuperado el Diciembre de 2014, de <http://trazasnarino.org/>
- Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel Methods For Pattern Analysis*. Cambridge: Cambridge University Press.

- Superintendencia de Servicios Públicos. (2010). *Sistema Único de Información de Servicios Públicos SUI*. Obtenido de <http://www.sui.gov.co>
- Tabares, H., & Hernández, J. (2008). Mapeo curvas típicas demanda de energía eléctrica del sector residencial, comercial e industrial de la ciudad de Medellín, usando redes neuronales artificiales y algoritmos de interpolación. *Revista Facultad de Ingeniería Universidad de Antioquia*(46), 110-118.
- Teknomu, K. (Julio de 2007). *Revoledu - Revolution Education*. Obtenido de <http://www.revoledu.org/>
- The Pennsylvania State University, Department of Statistics. (2013). En *Peen State STAT 510 Applied Time Series Analysis* (págs. Lesson 3, 1:20). Pennsylvania : The Pennsylvania State University.
- Unidad de Planeación Minero Energética, Subdirección de Planeación Energética - Grupo de Demanda Energética. (2013). *Proyección de Demanda de Energía Eléctrica en Colombia*. Bogotá.
- Universidad de Nariño. (26 de Enero de 2010). Acuerdo 005. *Por el cual se deroga el Acuerdo No. 043 del 30 de abril de 2002 y se adopta la nueva reglamentación del trabajo de grado para los programas de Ingeniería Civil, Ingeniería de Sistemas, Ingeniería Electrónica y los demás que administre la Facultad*. San Juan de Pasto, Nariño, Colombia.
- Uriel Jiménez, E. (1992). *Análisis de series temporales: modelos ARIMA*. Valencia: Paraninfo.
- Valencia, A. L., Lozano, C. A., & Moreno, C. (2007). Modelo de promedios móviles para el pronóstico horario de potencia y energía eléctrica. *El Hombre y la Máquina*(29), 96-105.
- Viloria de la Hoz, J. (2007). Economía del departamento de Nariño: Ruralidad y aislamiento geográfico. *Documentos de trabajo sobre economía regional*, 5-76.

ANEXO 1.
**Características de los municipios agrupados en 14 clusters con el método de
mixtura de Gaussianas**

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Natalidad | Pob_total | Consumo Residen | Consumo Comercial | Consumo Industrial | Subregión | PersXviv | ConsXpers | CLUSTER |
|----------------------|----------------|-----------|---------|------|-----------|-----------|-----------------|-------------------|--------------------|----------------|----------|-----------|---------|
| ARBOLEDA | 2200 | 40.2 | 3.84 | 4.4 | 11.85 | 7513 | 1050554 | 6476.06 | 1254.75 | Juanambú | 4.07 | 139.83 | |
| BUESACO | 1950 | 48.35 | 3.34 | 4.07 | 11.64 | 23887 | 3801098 | 6138.21 | 4342 | Juanambú | 4.35 | 159.13 | |
| CONTADERO | 2500 | 38.86 | 3.27 | 4.18 | 11.7 | 6839 | 1128962 | 3851.86 | | Exprovincia de | 4.17 | 165.08 | |
| CORDOBA | 2867 | 48.51 | 3.5 | 4.28 | 15.31 | 13844 | 2193281 | 4945.77 | 1700.75 | Exprovincia de | 3.31 | 158.43 | |
| CUASPUD | 3050 | 39.96 | 3.09 | 4.04 | 12.39 | 8395 | 1264091 | 6243.27 | 14642 | Exprovincia de | 4.46 | 150.58 | |
| EL_PENOL | 2200 | 46.43 | 3.12 | 4.00 | 7.08 | 6642 | 1034510 | 6356.00 | | Guambuyaco | 3.68 | 155.75 | |
| GUAITARILLA | 2650 | 41.12 | 3.06 | 3.99 | 9.16 | 12661 | 1924369 | 9178.36 | 2378 | Sabana | 3.74 | 151.99 | |
| ILES | 2950 | 42.38 | 3.29 | 4.1 | 9.35 | 8342 | 1145079 | 4200.06 | 4212.75 | Exprovincia de | 4.28 | 137.27 | 1 |
| LA_FLORIDA | 2180 | 43.21 | 2.92 | 3.83 | 9.98 | 10116 | 1634925 | 7271.19 | | Centro | 3.03 | 161.62 | |
| LINARES | 1500 | 43.08 | 3.05 | 3.99 | 10.6 | 10755 | 1711920 | 4919.04 | 9246 | Occidente | 3.52 | 159.17 | |
| POTOSI | 2750 | 50.4 | 2.56 | 3.75 | 11.34 | 12527 | 1885140 | 7394.06 | | Exprovincia de | 4.27 | 150.49 | |
| SAN_PABLO | 1750 | 38.00 | 2.19 | 3.51 | 9.17 | 17778 | 2638035 | 7505.58 | | Río Mayo | 4.73 | 148.39 | |
| TAMINANGO | 1500 | 44.37 | 3.12 | 4.01 | 10.57 | 19112 | 3298302 | 10187.62 | 22628.11 | Cordillera | 4.23 | 172.58 | |
| TANGUA | 2400 | 47.36 | 3.1 | 4.00 | 8.98 | 10130 | 1685174 | 6050.69 | 1628 | Centro | 4.07 | 166.35 | |
| YACUANQUER | 2700 | 42.22 | 3.24 | 4.15 | 11.43 | 10582 | 1708598 | 5406.13 | | Centro | 3.94 | 161.46 | |
| COLON_GENOVA | 1750 | 38.24 | 3.06 | 4.05 | 10.24 | 9959 | 1186005 | 788.50 | | Río Mayo | 4.89 | 119.09 | |
| CUMBAL | 3100 | 51.51 | 2.55 | 3.7 | 11.65 | 34858 | 4403992 | 4814.28 | 19814.15 | Exprovincia de | 5.65 | 126.34 | 2 |
| FUNES | 2300 | 55.46 | 3.39 | 4.17 | 10.32 | 6685 | 909421 | 12742.87 | 2778 | Exprovincia de | 4.38 | 136.04 | |
| LA_LLANADA | 2300 | 52.61 | 2.52 | 3.61 | 10.77 | 6127 | 755036 | 37030.25 | 17971.56 | Guambuyaco | 5.87 | 123.23 | |
| OSPINA | 2850 | 40.52 | 3.15 | 4.08 | 6.55 | 8547 | 1060157 | 266.67 | 1561.22 | Sabana | 4.86 | 124.04 | |
| IPIALES | 2892 | 45.08 | 2.33 | 3.42 | 12.21 | 126335 | 31411264 | 5474.57 | 8775.21 | Exprovincia de | 4.95 | 248.63 | 3 |
| EL_ROSARIO | 1600 | 53.88 | 3.93 | 4.39 | 8.6 | 10701 | 2865055 | 17006.63 | | Cordillera | 5.09 | 267.74 | |
| LEIVA | 2000 | 49.3 | 3.47 | 4.16 | 7.06 | 13039 | 3473079 | 11053.20 | | Cordillera | 7.19 | 266.36 | 4 |
| SAPUYES | 2900 | 45.63 | 2.4 | 3.59 | 8.83 | 6797 | 1776246 | 8711.89 | 7501.65 | Sabana | 3.11 | 261.33 | |
| ALDANA | 3050 | 39.56 | 2.18 | 3.61 | 10.63 | 6397 | 1420885 | 9383.34 | 1881.67 | Exprovincia de | 3.58 | 222.12 | |
| ANCUYA | 1358 | 40.09 | 2.71 | 3.83 | 10.03 | 7776 | 1466379 | 2009.11 | 5685.83 | Occidente | 3.55 | 188.58 | |
| BELEN | 2450 | 38.51 | 1.69 | 3.34 | 9.81 | 7133 | 1402939 | 3200.39 | 5812.04 | Río Mayo | 4.79 | 196.68 | |
| CONSACA | 1640 | 43.16 | 3.09 | 4.05 | 11.68 | 9759 | 2275858 | 3911.21 | 8671.14 | Occidente | 3.17 | 233.21 | |
| EL_TABLON_DE_GOMEZ | 1650 | 48.16 | 2.96 | 3.92 | 16.00 | 13247 | 2523433 | 642.50 | | Río Mayo | 3.98 | 190.49 | |
| EL_TAMBO | 2250 | 46.33 | 2.94 | 3.81 | 13.07 | 13006 | 2517800 | 9046.18 | 6641.16 | Guambuyaco | 3.27 | 193.59 | |
| GUACHUCAL | 3285 | 41.83 | 2.43 | 3.69 | 10.34 | 16152 | 3421172 | 5640.86 | 52407.12 | Exprovincia de | 3.86 | 211.81 | |
| GUALMATAN | 2900 | 37.39 | 2.51 | 3.86 | 14.29 | 5738 | 1221775 | 7071.81 | 1846.5 | Exprovincia de | 3.46 | 212.93 | 5 |
| IMUES | 2500 | 42.57 | 2.98 | 4.02 | 8.18 | 6721 | 1537659 | 5668.62 | 4923.83 | Sabana | 2.83 | 228.78 | |
| LA_CRUZ | 2370 | 44.11 | 3.66 | 4.27 | 9.7 | 18049 | 3583750 | 5853.32 | 941.14 | Río Mayo | 4.27 | 198.56 | |
| LA_UNION | 1727 | 36.66 | 2.63 | 3.64 | 14.02 | 26819 | 5568184 | 5338.65 | 854.93 | Juanambú | 3.59 | 207.62 | |
| MALLAMA | 1809 | 56.54 | 2.77 | 3.83 | 9.71 | 8338 | 1603978 | 3664.56 | 1291 | Pie de Monte | 4.00 | 192.37 | |
| PUERRES | 2800 | 52.69 | 2.68 | 3.83 | 12.51 | 8635 | 1732624 | 4664.87 | 1347 | Exprovincia de | 3.55 | 200.65 | |
| PUPIALES | 3014 | 39.64 | 2.54 | 3.71 | 10.83 | 19029 | 4077257 | 5148.24 | 58451.35 | Exprovincia de | 3.93 | 214.27 | |
| SANDONA | 1871 | 35.04 | 2.26 | 3.75 | 8.1 | 25547 | 4660424 | 3484.92 | 10187.46 | Occidente | 4.69 | 182.43 | |
| TUQUERRES | 3070 | 37.17 | 2.83 | 3.72 | 13.28 | 40966 | 9605326 | 6456.14 | 1757.82 | Sabana | 3.93 | 234.47 | |
| CUMBITARA | 1600 | 50.9 | 4.61 | 4.61 | 7.43 | 13590 | 1408626 | 8714.99 | | Cordillera | 8.59 | 103.65 | 6 |
| ALBAN | 1935 | 29.89 | 2.92 | 3.82 | 5.39 | 20980 | 1776298 | 3857.25 | | Río Mayo | 6.92 | 84.67 | |
| LOS_ANDES | 1588 | 55.93 | 3.47 | 4.23 | 5.97 | 18804 | 1602493 | 1633.40 | 8595 | Guambuyaco | 6.90 | 85.22 | |
| RICAURTE | 1181 | 57.8 | 3.7 | 4.19 | 14.45 | 17090 | 1462803 | 6716.55 | 3388.75 | Pie de Monte | 5.19 | 85.59 | |
| SAMANIEGO | 1750 | 43.69 | 3.14 | 4.06 | 7.13 | 49906 | 4320418 | 6031.35 | 4844.18 | Abades | 7.16 | 86.57 | 7 |
| SAN_BERNARDO | 2100 | 35.38 | 3.3 | 4.05 | 4.85 | 17116 | 1378722 | 1351.33 | | Río Mayo | 8.63 | 80.55 | |
| SAN_LORENZO | 2150 | 44.21 | 3.22 | 4.05 | 10.02 | 19261 | 1762450 | 6583.25 | 152.16 | Juanambú | 4.39 | 91.50 | |
| SAN_PEDRO_DE_CARTAGO | 1750 | 39.95 | 3.14 | 4.04 | 11.02 | 7348 | 658683 | 419.21 | 2960 | Juanambú | 5.55 | 89.64 | |
| BARBACOAS | 160 | 59.55 | 3.56 | 4.3 | 11.81 | 34627 | 8495130 | 6911.45 | 4868.5 | Telebí | 5.37 | 245.33 | 8 |
| SAN_ANDRES_DE_TUMACC | 3 | 29.75 | 2.77 | 3.89 | 16.51 | 183006 | 52137667 | 16946.38 | 9721.99 | Pacífico Sur | 6.48 | 284.90 | 9 |
| CHCHAGUI | 1950 | 42.98 | 2.83 | 3.85 | 11.56 | 13408 | 3869544 | 12605.76 | | Centro | 2.86 | 288.60 | 10 |
| POLICARPA | 1000 | 50.98 | 2.86 | 3.84 | 10.78 | 15583 | 4471695 | 3900.92 | 0 | Cordillera | 4.66 | 286.96 | |
| MAGUI_PAYAN | 40 | 60.83 | 3.46 | 4.41 | 5.2 | 19822 | 1357591 | | | Telebí | 8.68 | 68.49 | 11 |
| ROBERTO_PAYAN | 24 | 59.2 | 3.65 | 4.29 | 7.05 | 20132 | 1230507 | | | Telebí | 6.35 | 61.12 | |
| PROVIDENCIA | 2300 | 34.65 | 3.48 | 4.17 | 4.98 | 12644 | 677131 | 5651.00 | | Abades | 5.71 | 53.55 | 12 |
| SANTA_CRUZ | 2400 | 49.79 | 3.72 | 4.22 | 3.54 | 24886 | 847780 | 2413.60 | 390 | Abades | 9.24 | 34.07 | 13 |
| PASTO | 2527 | 32.07 | 1.28 | 2.79 | 12.05 | 417484 | 135670821 | 5579.78 | 28049.63 | Centro | 4.41 | 324.97 | 14 |

ANEXO 2.

Características de los municipios agrupados en 8 clusters con el método k-means

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Natalidad | Pob_total | Consumo Residen | Consumo Comercial | Consumo Industrial | Subregión | PersXviv | ConsXpers | CLUSTER |
|----------------------|----------------|-----------|---------|------|-----------|-----------|-----------------|-------------------|--------------------|----------------|----------|-----------|---------|
| ALDANA | 3050 | 39.56 | 2.18 | 3.61 | 10.63 | 6397 | 1420885 | 9383.34 | 1881.67 | Exprovincia de | 3.58 | 222.12 | 1 |
| CONSACA | 1640 | 43.16 | 3.09 | 4.05 | 11.68 | 9759 | 2275858 | 3911.21 | 8671.14 | Occidente | 3.17 | 233.21 | |
| GUACHUCAL | 3285 | 41.83 | 2.43 | 3.69 | 10.34 | 16152 | 3421172 | 5640.86 | 52407.12 | Exprovincia de | 3.86 | 211.81 | |
| GUALMATAN | 2900 | 37.39 | 2.51 | 3.86 | 14.29 | 5738 | 1221775 | 7071.81 | 1846.5 | Exprovincia de | 3.46 | 212.93 | |
| IMUES | 2500 | 42.57 | 2.98 | 4.02 | 8.18 | 6721 | 1537659 | 5668.62 | 4923.83 | Sabana | 2.83 | 228.78 | |
| LA_UNION | 1727 | 36.66 | 2.63 | 3.64 | 14.02 | 26819 | 5568184 | 5338.65 | 854.93 | Juanambú | 3.59 | 207.62 | |
| PUPIALES | 3014 | 39.64 | 2.54 | 3.71 | 10.83 | 19029 | 4077257 | 5148.24 | 58451.35 | Exprovincia de | 3.93 | 214.27 | |
| TUQUERRES | 3070 | 37.17 | 2.83 | 3.72 | 13.28 | 40966 | 9605326 | 6456.14 | 1757.82 | Sabana | 3.93 | 234.47 | |
| ANCUYA | 1358 | 40.09 | 2.71 | 3.83 | 10.03 | 7776 | 1466379 | 2009.11 | 5685.83 | Occidente | 3.55 | 188.58 | |
| BELEN | 2450 | 38.51 | 1.69 | 3.34 | 9.81 | 7133 | 1402939 | 3200.39 | 5812.04 | Río Mayo | 4.79 | 196.68 | |
| EL_TABLON_DE_GOMEZ | 1650 | 48.16 | 2.96 | 3.92 | 16.00 | 13247 | 2523433 | 642.50 | | Río Mayo | 3.98 | 190.49 | |
| EL_TAMBO | 2250 | 46.33 | 2.94 | 3.81 | 13.07 | 13006 | 2517800 | 9046.18 | 6641.16 | Guambuyaco | 3.27 | 193.59 | |
| LA_CRUZ | 2370 | 44.11 | 3.66 | 4.27 | 9.7 | 18049 | 3583750 | 5853.32 | 941.14 | Río Mayo | 4.27 | 198.56 | |
| MALLAMA | 1809 | 56.54 | 2.77 | 3.83 | 9.71 | 8338 | 1603978 | 3664.56 | 1291 | Pie de Monte | 4.00 | 192.37 | |
| PUERRES | 2800 | 52.69 | 2.68 | 3.83 | 12.51 | 8635 | 1732624 | 4664.87 | 1347 | Exprovincia de | 3.55 | 200.65 | |
| SANDONA | 1871 | 35.04 | 2.26 | 3.75 | 8.1 | 25547 | 4660424 | 3484.92 | 10187.46 | Occidente | 4.69 | 182.43 | |
| CHCHAGUI | 1950 | 42.98 | 2.83 | 3.85 | 11.56 | 13408 | 3869544 | 12605.76 | | Centro | 2.86 | 288.60 | |
| PASTO | 2527 | 32.07 | 1.28 | 2.79 | 12.05 | 417484 | 135670821 | 5579.78 | 28049.63 | Centro | 4.41 | 324.97 | |
| POLICARPA | 1000 | 50.98 | 2.86 | 3.84 | 10.78 | 15583 | 4471695 | 3900.92 | 0 | Cordillera | 4.66 | 286.96 | |
| SAN_ANDRES_DE_TUMACC | 3 | 29.75 | 2.77 | 3.89 | 16.51 | 183006 | 52137667 | 16946.38 | 9721.99 | Pacífico Sur | 6.48 | 284.90 | |
| BUESACO | 1950 | 48.35 | 3.34 | 4.07 | 11.64 | 23887 | 3801098 | 6138.21 | 4342 | Juanambú | 4.35 | 159.13 | |
| CONTADERO | 2500 | 38.86 | 3.27 | 4.18 | 11.7 | 6839 | 1128962 | 3851.86 | | Exprovincia de | 4.17 | 165.08 | |
| CORDOBA | 2867 | 48.51 | 3.5 | 4.28 | 15.31 | 13844 | 2193281 | 4945.77 | 1700.75 | Exprovincia de | 3.31 | 158.43 | |
| CUASPUD | 3050 | 39.96 | 3.09 | 4.04 | 12.39 | 8395 | 1264091 | 6243.27 | 14642 | Exprovincia de | 4.46 | 150.58 | |
| EL_PENOL | 2200 | 46.43 | 3.12 | 4.00 | 7.08 | 6642 | 1034510 | 6356.00 | 8232 | Guambuyaco | 3.68 | 155.75 | |
| GUAITARILLA | 2650 | 41.12 | 3.06 | 3.99 | 9.16 | 12661 | 1924369 | 9178.36 | 2378 | Sabana | 3.74 | 151.99 | |
| LA_FLORIDA | 2180 | 43.21 | 2.92 | 3.83 | 9.98 | 10116 | 1634925 | 7271.19 | | Centro | 3.03 | 161.62 | |
| LINARES | 1500 | 43.08 | 3.05 | 3.99 | 10.6 | 10755 | 1711920 | 4919.04 | 9246 | Occidente | 3.52 | 159.17 | |
| POTOSI | 2750 | 50.4 | 2.56 | 3.75 | 11.34 | 12527 | 1885140 | 7394.06 | | Exprovincia de | 4.27 | 150.49 | |
| SAN_PABLO | 1750 | 38.00 | 2.19 | 3.51 | 9.17 | 17778 | 2638035 | 7505.58 | | Río Mayo | 4.73 | 148.39 | |
| TAMINANGO | 1500 | 44.37 | 3.12 | 4.01 | 10.57 | 19112 | 3298302 | 10187.62 | 22628.11 | Cordillera | 4.23 | 172.58 | |
| TANGUA | 2400 | 47.36 | 3.1 | 4.00 | 8.98 | 10130 | 1685174 | 6050.69 | 1628 | Centro | 4.07 | 166.35 | |
| YACUANQUER | 2700 | 42.22 | 3.24 | 4.15 | 11.43 | 10582 | 1708598 | 5406.13 | | Centro | 3.94 | 161.46 | |
| BARBACOAS | 160 | 59.55 | 3.56 | 4.3 | 11.81 | 34627 | 8495130 | 6911.45 | 4868.5 | Telebí | 5.37 | 245.33 | |
| EL_ROSARIO | 1600 | 53.88 | 3.93 | 4.39 | 8.6 | 10701 | 2865055 | 17006.63 | | Cordillera | 5.09 | 267.74 | |
| IPIALES | 2892 | 45.08 | 2.33 | 3.42 | 12.21 | 126335 | 31411264 | 5474.57 | 8775.21 | Exprovincia de | 4.95 | 248.63 | |
| LEIVA | 2000 | 49.3 | 3.47 | 4.16 | 7.06 | 13039 | 3473079 | 11053.20 | | Cordillera | 7.19 | 266.36 | |
| SAPUYES | 2900 | 45.63 | 2.4 | 3.59 | 8.83 | 6797 | 1776246 | 8711.89 | 7501.65 | Sabana | 3.11 | 261.33 | |
| ARBOLEDA | 2200 | 40.2 | 3.84 | 4.4 | 11.85 | 7513 | 1050554 | 6476.06 | 1254.75 | Juanambú | 4.07 | 139.83 | |
| COLON_GENOVA | 1750 | 38.24 | 3.06 | 4.05 | 10.24 | 9959 | 1186005 | 788.50 | | Río Mayo | 4.89 | 119.09 | |
| CUMBAL | 3100 | 51.51 | 2.55 | 3.7 | 11.65 | 34858 | 4403992 | 4814.28 | 19814.15 | Exprovincia de | 5.65 | 126.34 | |
| FUNES | 2300 | 55.46 | 3.39 | 4.17 | 10.32 | 6685 | 909421 | 12742.87 | 2778 | Exprovincia de | 4.38 | 136.04 | |
| ILES | 2950 | 42.38 | 3.29 | 4.1 | 9.35 | 8342 | 1145079 | 4200.06 | 4212.75 | Exprovincia de | 4.28 | 137.27 | |
| LA_LLANADA | 2300 | 52.61 | 2.52 | 3.61 | 10.77 | 6127 | 755036 | 37030.25 | 17971.56 | Guambuyaco | 5.87 | 123.23 | |
| OSPINA | 2850 | 40.52 | 3.15 | 4.08 | 6.55 | 8547 | 1060157 | 266.67 | 1561.22 | Sabana | 4.86 | 124.04 | |
| MAGUI_PAYAN | 40 | 60.83 | 3.46 | 4.41 | 5.2 | 19822 | 1357591 | | | Telebí | 8.68 | 68.49 | |
| PROVIDENCIA | 2300 | 34.65 | 3.48 | 4.17 | 4.98 | 12644 | 677131 | 5651.00 | | Abades | 5.71 | 53.55 | |
| ROBERTO_PAYAN | 24 | 59.2 | 3.65 | 4.29 | 7.05 | 20132 | 1230507 | | | Telebí | 6.35 | 61.12 | |
| SANTA_CRUZ | 2400 | 49.79 | 3.72 | 4.22 | 3.54 | 24886 | 847780 | 2413.60 | 390 | Abades | 9.24 | 34.07 | |
| ALBAN | 1935 | 29.89 | 2.92 | 3.82 | 5.39 | 20980 | 1776298 | 3857.25 | | Río Mayo | 6.92 | 84.67 | |
| CUMBITARA | 1600 | 50.9 | 4.61 | 4.61 | 7.43 | 13590 | 1408626 | 8714.99 | | Cordillera | 8.59 | 103.65 | |
| LOS_ANDES | 1588 | 55.93 | 3.47 | 4.23 | 5.97 | 18804 | 1602493 | 1633.40 | 8595 | Guambuyaco | 6.90 | 85.22 | |
| RICAURTE | 1181 | 57.8 | 3.7 | 4.19 | 14.45 | 17090 | 1462803 | 6716.55 | 3388.75 | Pie de Monte | 5.19 | 85.59 | |
| SAMANIEGO | 1750 | 43.69 | 3.14 | 4.06 | 7.13 | 49906 | 4320418 | 6031.35 | 4844.18 | Abades | 7.16 | 86.57 | |
| SAN_BERNARDO | 2100 | 35.38 | 3.3 | 4.05 | 4.85 | 17116 | 1378722 | 1351.33 | | Río Mayo | 8.63 | 80.55 | |
| SAN_LORENZO | 2150 | 44.21 | 3.22 | 4.05 | 10.02 | 19261 | 1762450 | 6583.25 | 152.16 | Juanambú | 4.39 | 91.50 | |
| SAN_PEDRO_DE_CARTAGO | 1750 | 39.95 | 3.14 | 4.04 | 11.02 | 7348 | 658683 | 419.21 | 2960 | Juanambú | 5.55 | 89.64 | |

ANEXO 3.

Características de los municipios agrupados en 9 clusters con el método de RNA

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Natalidad | Pob_total | Consumo Residen | Consumo Comercial | Consumo Industrial | Subregión | PersXviv | ConsXpers | CLUSTER |
|----------------------|----------------|-----------|---------|------|-----------|-----------|-----------------|-------------------|--------------------|----------------|----------|-----------|---------|
| ANCUYA | 1358 | 40.09 | 2.71 | 3.83 | 10.03 | 7776 | 1466379 | 2009.11 | 5685.83 | Occidente | 3.55 | 188.58 | 1 |
| BELEN | 2450 | 38.51 | 1.69 | 3.34 | 9.81 | 7133 | 1402939 | 3200.39 | 5812.04 | Río Mayo | 4.79 | 196.68 | |
| EL_TABLON_DE_GOMEZ | 1650 | 48.16 | 2.96 | 3.92 | 16.00 | 13247 | 2523433 | 642.50 | | Río Mayo | 3.98 | 190.49 | |
| EL_TAMBO | 2250 | 46.33 | 2.94 | 3.81 | 13.07 | 13006 | 2517800 | 9046.18 | 6641.16 | Guambuyaco | 3.27 | 193.59 | |
| LA_CRUZ | 2370 | 44.11 | 3.66 | 4.27 | 9.7 | 18049 | 3583750 | 5853.32 | 941.14 | Río Mayo | 4.27 | 198.56 | |
| MALLAMA | 1809 | 56.54 | 2.77 | 3.83 | 9.71 | 8338 | 1603978 | 3664.56 | 1291 | Pie de Monte | 4.00 | 192.37 | |
| PUERRES | 2800 | 52.69 | 2.68 | 3.83 | 12.51 | 8635 | 1732624 | 4664.87 | 1347 | Exprovincia de | 3.55 | 200.65 | |
| SANDONA | 1871 | 35.04 | 2.26 | 3.75 | 8.1 | 25547 | 4660424 | 3484.92 | 10187.46 | Occidente | 4.69 | 182.43 | |
| BUESACO | 1950 | 48.35 | 3.34 | 4.07 | 11.64 | 23887 | 3801098 | 6138.21 | 4342 | Juanambú | 4.35 | 159.13 | |
| CONTADERO | 2500 | 38.86 | 3.27 | 4.18 | 11.7 | 6839 | 1128962 | 3851.86 | | Exprovincia de | 4.17 | 165.08 | |
| CORDOBA | 2867 | 48.51 | 3.5 | 4.28 | 15.31 | 13844 | 2193281 | 4945.77 | 1700.75 | Exprovincia de | 3.31 | 158.43 | |
| EL_PENOL | 2200 | 46.43 | 3.12 | 4.00 | 7.08 | 6642 | 1034510 | 6356.00 | 8232 | Guambuyaco | 3.68 | 155.75 | 2 |
| LA_FLORIDA | 2180 | 43.21 | 2.92 | 3.83 | 9.98 | 10116 | 1634925 | 7271.19 | | Centro | 3.03 | 161.62 | |
| LINARES | 1500 | 43.08 | 3.05 | 3.99 | 10.6 | 10755 | 1711920 | 4919.04 | 9246 | Occidente | 3.52 | 159.17 | |
| TAMINANGO | 1500 | 44.37 | 3.12 | 4.01 | 10.57 | 19112 | 3298302 | 10187.62 | 22628.11 | Cordillera | 4.23 | 172.58 | |
| TANGUA | 2400 | 47.36 | 3.1 | 4.00 | 8.98 | 10130 | 1685174 | 6050.69 | 1628 | Centro | 4.07 | 166.35 | |
| YACUANQUER | 2700 | 42.22 | 3.24 | 4.15 | 11.43 | 10582 | 1708598 | 5406.13 | | Centro | 3.94 | 161.46 | |
| ALBAN | 1935 | 29.89 | 2.92 | 3.82 | 5.39 | 20980 | 1776298 | 3857.25 | | Río Mayo | 6.92 | 84.67 | 3 |
| CUMBITARA | 1600 | 50.9 | 4.61 | 4.61 | 7.43 | 13590 | 1408626 | 8714.99 | | Cordillera | 8.59 | 103.65 | |
| LOS_ANDES | 1588 | 55.93 | 3.47 | 4.23 | 5.97 | 18804 | 1602493 | 1633.40 | 8595 | Guambuyaco | 6.90 | 85.22 | |
| RICARTE | 1181 | 57.8 | 3.7 | 4.19 | 14.45 | 17090 | 1462803 | 6716.55 | 3388.75 | Pie de Monte | 5.19 | 85.59 | |
| SAMANIEGO | 1750 | 43.69 | 3.14 | 4.06 | 7.13 | 49906 | 4320418 | 6031.35 | 4844.18 | Abades | 7.16 | 86.57 | |
| SAN_BERNARDO | 2100 | 35.38 | 3.3 | 4.05 | 4.85 | 17116 | 1378722 | 1351.33 | | Río Mayo | 8.63 | 80.55 | |
| SAN_LORENZO | 2150 | 44.21 | 3.22 | 4.05 | 10.02 | 19261 | 1762450 | 6583.25 | 152.16 | Juanambú | 4.39 | 91.50 | |
| SAN_PEDRO_DE_CARTAGO | 1750 | 39.95 | 3.14 | 4.04 | 11.02 | 7348 | 658683 | 419.21 | 2960 | Juanambú | 5.55 | 89.64 | |
| ALDANA | 3050 | 39.56 | 2.18 | 3.61 | 10.63 | 6397 | 1420885 | 9383.34 | 1881.67 | Exprovincia de | 3.58 | 222.12 | |
| CONSACA | 1640 | 43.16 | 3.09 | 4.05 | 11.68 | 9759 | 2275858 | 3911.21 | 8671.14 | Occidente | 3.17 | 233.21 | |
| GUACHUCAL | 3285 | 41.83 | 2.43 | 3.69 | 10.34 | 16152 | 3421172 | 5640.86 | 52407.12 | Exprovincia de | 3.86 | 211.81 | 4 |
| GUALMATAN | 2900 | 37.39 | 2.51 | 3.86 | 14.29 | 5738 | 1221775 | 7071.81 | 1846.5 | Exprovincia de | 3.46 | 212.93 | |
| IMUES | 2500 | 42.57 | 2.98 | 4.02 | 8.18 | 6721 | 1537659 | 5668.62 | 4923.83 | Sabana | 2.83 | 228.78 | |
| LA_UNION | 1727 | 36.66 | 2.63 | 3.64 | 14.02 | 26819 | 5568184 | 5338.65 | 854.93 | Juanambú | 3.59 | 207.62 | |
| PUPIALES | 3014 | 39.64 | 2.54 | 3.71 | 10.83 | 19029 | 4077257 | 5148.24 | 58451.35 | Exprovincia de | 3.93 | 214.27 | |
| TUQUERRES | 3070 | 37.17 | 2.83 | 3.72 | 13.28 | 40966 | 9605326 | 6456.14 | 1757.82 | Sabana | 3.93 | 234.47 | |
| ARBOLEDA | 2200 | 40.2 | 3.84 | 4.4 | 11.85 | 7513 | 1050554 | 6476.06 | 1254.75 | Juanambú | 4.07 | 139.83 | 5 |
| CUASPUD | 3050 | 39.96 | 3.09 | 4.04 | 12.39 | 8395 | 1264091 | 6243.27 | 14642 | Exprovincia de | 4.46 | 150.58 | |
| GUAITARILLA | 2650 | 41.12 | 3.06 | 3.99 | 9.16 | 12661 | 1924369 | 9178.36 | 2378 | Sabana | 3.74 | 151.99 | |
| ILES | 2950 | 42.38 | 3.29 | 4.1 | 9.35 | 8342 | 1145079 | 4200.06 | 4212.75 | Exprovincia de | 4.28 | 137.27 | |
| POTOSI | 2750 | 50.4 | 2.56 | 3.75 | 11.34 | 12527 | 1885140 | 7394.06 | | Exprovincia de | 4.27 | 150.49 | |
| SAN_PABLO | 1750 | 38.00 | 2.19 | 3.51 | 9.17 | 17778 | 2638035 | 7505.58 | | Río Mayo | 4.73 | 148.39 | |
| MAGUI_PAYAN | 40 | 60.83 | 3.46 | 4.41 | 5.2 | 19822 | 1357591 | | | Telembí | 8.68 | 68.49 | 6 |
| PROVIDENCIA | 2300 | 34.65 | 3.48 | 4.17 | 4.98 | 12644 | 677131 | 5651.00 | | Abades | 5.71 | 53.55 | |
| ROBERTO_PAYAN | 24 | 59.2 | 3.65 | 4.29 | 7.05 | 20132 | 1230507 | | | Telembí | 6.35 | 61.12 | |
| SANTA_CRUZ | 2400 | 49.79 | 3.72 | 4.22 | 3.54 | 24886 | 847780 | 2413.60 | 390 | Abades | 9.24 | 34.07 | |
| CHCHAGUI | 1950 | 42.98 | 2.83 | 3.85 | 11.56 | 13408 | 3869544 | 12605.76 | | Centro | 2.86 | 288.60 | |
| PASTO | 2527 | 32.07 | 1.28 | 2.79 | 12.05 | 417484 | 135670821 | 5579.78 | 28049.63 | Centro | 4.41 | 324.97 | |
| POLICARPA | 1000 | 50.98 | 2.86 | 3.84 | 10.78 | 15583 | 4471695 | 3900.92 | 0 | Cordillera | 4.66 | 286.96 | |
| SAN_ANDRES_DE_TUMACC | 3 | 29.75 | 2.77 | 3.89 | 16.51 | 183006 | 52137667 | 16946.38 | 9721.99 | Pacífico Sur | 6.48 | 284.90 | |
| BARBACOAS | 160 | 59.55 | 3.56 | 4.3 | 11.81 | 34627 | 8495130 | 6911.45 | 4868.5 | Telembí | 5.37 | 245.33 | 8 |
| EL_ROSARIO | 1600 | 53.88 | 3.93 | 4.39 | 8.6 | 10701 | 2865055 | 17006.63 | | Cordillera | 5.09 | 267.74 | |
| IPIALES | 2892 | 45.08 | 2.33 | 3.42 | 12.21 | 126335 | 31411264 | 5474.57 | 8775.21 | Exprovincia de | 4.95 | 248.63 | |
| LEIVA | 2000 | 49.3 | 3.47 | 4.16 | 7.06 | 13039 | 3473079 | 11053.20 | | Cordillera | 7.19 | 266.36 | |
| SAPUYES | 2900 | 45.63 | 2.4 | 3.59 | 8.83 | 6797 | 1776246 | 8711.89 | 7501.65 | Sabana | 3.11 | 261.33 | |
| COLON_GENOVA | 1750 | 38.24 | 3.06 | 4.05 | 10.24 | 9959 | 1186005 | 788.50 | | Río Mayo | 4.89 | 119.09 | |
| CUMBAL | 3100 | 51.51 | 2.55 | 3.7 | 11.65 | 34858 | 4403992 | 4814.28 | 19814.15 | Exprovincia de | 5.65 | 126.34 | |
| FUNES | 2300 | 55.46 | 3.39 | 4.17 | 10.32 | 6685 | 909421 | 12742.87 | 2778 | Exprovincia de | 4.38 | 136.04 | |
| LA_LLANADA | 2300 | 52.61 | 2.52 | 3.61 | 10.77 | 6127 | 755036 | 37030.25 | 17971.56 | Guambuyaco | 5.87 | 123.23 | |
| OSPINA | 2850 | 40.52 | 3.15 | 4.08 | 6.55 | 8547 | 1060157 | 266.67 | 1561.22 | Sabana | 4.86 | 124.04 | |

ANEXO 4.

Características de los municipios de la Región 1 de la clasificación con el método de RNA

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Prod_Leche 2011 [t] | Prod_Significativa leche | Prod_Significativa Carne | Prod_Significativa pesca | Pob_total | Suscriptores Res. | Consumo Est1 kWh | Consumo Est2 [kWh] | Consumo Est3 [KWh] | Consumo Est4 [kWh] | Consumo Resid. | Consumo Comerc. | Consumo Indust. | Subregión | Personas X vivienda | Cons. Per Cápita | Cluster final |
|--------------------|----------------|-----------|---------|------|---------------------|--------------------------|--------------------------|--------------------------|-----------|-------------------|------------------|--------------------|--------------------|--------------------|----------------|-----------------|-----------------|--------------|---------------------|------------------|---------------|
| CUMBITARA | 1600 | 50.9 | 4.61 | 4.61 | 1250 | 0 | 1 | 1 | 13590 | 1583 | 770.27 | 1815.06 | 3532.00 | 0.00 | 1408626 | 8714.99 | 0 | Cordillera | 8.59 | 103.65 | 1 |
| SANTA_CRUZ | 2400 | 49.79 | 3.72 | 4.22 | 1624 | 0 | 0 | 0 | 24886 | 2694 | 402.61 | 670.26 | 0.00 | 0.00 | 847780 | 2413.60 | 390 | Abades | 9.24 | 34.07 | 1 |
| ALBAN | 1935 | 29.89 | 2.92 | 3.82 | 227.5 | 0 | 1 | 0 | 20980 | 3030 | 787.77 | 1056.92 | 2288.00 | 0.00 | 1776298 | 3857.25 | 0 | Río Mayo | 6.92 | 84.67 | 2 |
| SAMANIEGO | 1750 | 43.69 | 3.14 | 4.06 | 1200 | 0 | 0 | 1 | 49906 | 6973 | 445.63 | 831.62 | 1309.23 | 706.3333333 | 4320418 | 6031.35 | 4844.18 | Abades | 7.16 | 86.57 | 2 |
| SAN_BERNARDO | 2100 | 35.38 | 3.3 | 4.05 | 2800 | 0 | 0 | 0 | 17116 | 1984 | 849.33 | 1069.95 | 0.00 | 0.00 | 1378722 | 1351.33 | 0 | Río Mayo | 8.63 | 80.55 | 2 |
| LA_LLANADA | 2300 | 52.61 | 2.52 | 3.61 | 2000 | 0 | 0 | 0 | 6127 | 1044 | 725.25 | 706.94 | 0.00 | 0.00 | 755036 | 37030.25 | 17971.56 | Guambuyaco | 5.87 | 123.23 | 3 |
| POLICARPA | 1000 | 50.98 | 2.86 | 3.84 | 1174.2 | 0 | 1 | 0 | 15583 | 3341 | 1315.94 | 2239.53 | 6785.00 | 0.00 | 4471695 | 3900.92 | 0 | Cordillera | 4.66 | 286.96 | 3 |
| EL_ROSARIO | 1600 | 53.88 | 3.93 | 4.39 | 720 | 0 | 1 | 1 | 10701 | 2101 | 1253.73 | 3014.77 | 0.00 | 0.00 | 2865055 | 17006.63 | 0 | Cordillera | 5.09 | 267.74 | 4 |
| LEIVA | 2000 | 49.3 | 3.47 | 4.16 | 2765 | 0 | 1 | 0 | 13039 | 1812 | 1895.23 | 2827.00 | 0.00 | 0.00 | 3473079 | 11053.20 | 0 | Cordillera | 7.19 | 266.36 | 4 |
| LOS_ANDES | 1588 | 55.93 | 3.47 | 4.23 | 3200 | 0 | 0 | 0 | 18804 | 2726 | 511.48 | 986.16 | 1120.33 | 0.00 | 1602493 | 1633.40 | 8595 | Guambuyaco | 6.90 | 85.22 | 4 |
| RICAUARTE | 1181 | 57.8 | 3.7 | 4.19 | 1640 | 0 | 0 | 1 | 17090 | 3295 | 804.90 | 1361.78 | 1796.33 | 0.00 | 1462803 | 6716.55 | 3388.75 | Pie de Monte | 5.19 | 85.59 | 4 |
| LA_UNION | 1727 | 36.66 | 2.63 | 3.64 | 400 | 0 | 0 | 0 | 26819 | 7466 | 543.61 | 1079.44 | 1512.16 | 0.00 | 5568184 | 5338.65 | 854.93 | Juanambú | 3.59 | 207.62 | 5 |
| SAN_PABLO | 1750 | 38.00 | 2.19 | 3.51 | 1470 | 0 | 1 | 0 | 17778 | 3758 | 618.52 | 1414.61 | 2685.71 | 0.00 | 2638035 | 7505.58 | 0 | Río Mayo | 4.73 | 148.39 | 5 |
| SANDONA | 1871 | 35.04 | 2.26 | 3.75 | 4800 | 0 | 0 | 0 | 25547 | 5453 | 708.43 | 977.68 | 1347.53 | 1546.66667 | 4660424 | 3484.92 | 10187.46 | Occidente | 4.69 | 182.43 | 5 |
| ANCUYA | 1358 | 40.09 | 2.71 | 3.83 | 846 | 0 | 0 | 0 | 7776 | 2192 | 610.31 | 756.62 | 1048.86 | 0.00 | 1466379 | 2009.11 | 5685.83 | Occidente | 3.55 | 188.58 | 6 |
| ARBOLEDA | 2200 | 40.2 | 3.84 | 4.4 | 562.5 | 0 | 0 | 0 | 7513 | 1848 | 556.72 | 661.06 | 1406.50 | 0.00 | 1050554 | 6476.06 | 1254.75 | Juanambú | 4.07 | 139.83 | 6 |
| CHCHAGUI | 1950 | 42.98 | 2.83 | 3.85 | 3156 | 0 | 1 | 0 | 13408 | 4682 | 712.76 | 928.97 | 866.97 | 1003.79113 | 3869544 | 12605.76 | 0 | Centro | 2.86 | 288.60 | 6 |
| CONSACA | 1640 | 43.16 | 3.09 | 4.05 | 4860 | 0 | 1 | 0 | 9759 | 3078 | 691.83 | 798.84 | 1172.96 | 545 | 2275858 | 3911.21 | 8671.14 | Occidente | 3.17 | 233.21 | 6 |
| EL_PENOL | 2200 | 46.43 | 3.12 | 4.00 | 2640 | 0 | 1 | 0 | 6642 | 1803 | 533.55 | 684.43 | 0.00 | 0.00 | 1034510 | 6356.00 | 8232 | Guambuyaco | 3.68 | 155.75 | 6 |
| EL_TABLON_DE_GOMEZ | 1650 | 48.16 | 2.96 | 3.92 | 3650 | 0 | 0 | 0 | 13247 | 3328 | 750.52 | 868.85 | 0.00 | 0.00 | 2523433 | 642.50 | 0 | Río Mayo | 3.98 | 190.49 | 6 |
| EL_TAMBO | 2250 | 46.33 | 2.94 | 3.81 | 4550 | 0 | 0 | 0 | 13006 | 3979 | 444.43 | 874.54 | 1568.07 | 0.00 | 2517800 | 9046.18 | 6641.16 | Guambuyaco | 3.27 | 193.59 | 6 |
| LA_FLORIDA | 2180 | 43.21 | 2.92 | 3.83 | 16660 | 1 | 0 | 1 | 10116 | 3337 | 406.27 | 544.51 | 638.63 | 0.00 | 1634925 | 7271.19 | 0 | Centro | 3.03 | 161.62 | 6 |
| LINARES | 1500 | 43.08 | 3.05 | 3.99 | 2704 | 0 | 0 | 0 | 10755 | 3054 | 487.99 | 635.11 | 673.00 | 0.00 | 1711920 | 4919.04 | 9246 | Occidente | 3.52 | 159.17 | 6 |
| SAN_LORENZO | 2150 | 44.21 | 3.22 | 4.05 | 3000 | 0 | 1 | 0 | 19261 | 4386 | 355.22 | 687.26 | 1522.00 | 0.00 | 1762450 | 6583.25 | 152.16 | Juanambú | 4.39 | 91.50 | 6 |
| TAMINANGO | 1500 | 44.37 | 3.12 | 4.01 | 2100 | 0 | 1 | 1 | 19112 | 4521 | 588.18 | 1060.40 | 1202.51 | 2341.1 | 3298302 | 10187.62 | 22628.11 | Cordillera | 4.23 | 172.58 | 6 |

ANEXO 5.

Características de los municipios de la Región 2 de la clasificación con el método de RNA

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Prod_Leche 2011 [lt] | Prod_Significativa leche | Prod_Significativa Carne | Prod_Significativa pesca | Pob_total | Suscriptores Res. | Consumo Est1 kWh | Consumo Est2 [kWh] | Consumo Est3 [KWh] | Consumo Est4 [kWh] | Consumo Resid. | Consumo Comerc. | Consumo Indust. | Subregión | Personas X vivienda | Cons. Per Cápita | Clasificación primaria | Cluster final |
|---------------------|----------------|-----------|---------|------|----------------------|--------------------------|--------------------------|--------------------------|-----------|-------------------|------------------|--------------------|--------------------|--------------------|----------------|-----------------|-----------------|--------------|---------------------|------------------|------------------------|---------------|
| BARBACOAS | 160 | 59.55 | 3.56 | 4.3 | 0 | 0 | 1 | 1 | 34627 | 6448 | 2875.10 | 4721.18 | 5155.14 | 0.00 | 8495130 | 6911.45 | 4868.5 | Telembí | 5.37 | 245.33 | 2 | 1 |
| BUESACO | 1950 | 48.35 | 3.34 | 4.07 | 8315.3 | 0 | 1 | 1 | 23887 | 5493 | 630.43 | 796.79 | 1102.85 | 1630 | 3801098 | 6138.21 | 4342 | Juanambú | 4.35 | 159.13 | 2 | 1 |
| MAGUI_PAYAN | 40 | 60.83 | 3.46 | 4.41 | 280 | 0 | 0 | 0 | 19822 | 2284 | 2359.30 | 0.00 | 0.00 | 0.00 | 1357591 | | | Telembí | 8.68 | 68.49 | 2 | 2 |
| ROBERTO_PAYAN | 24 | 59.2 | 3.65 | 4.29 | 0 | 0 | 1 | 0 | 20132 | 3168 | 2282.22 | 0.00 | 0.00 | 0.00 | 1230507 | | | Telembí | 6.35 | 61.12 | 2 | 2 |
| SAN_ANDRES_DE_TUMAI | 3 | 29.75 | 2.77 | 3.89 | 1580 | 0 | 1 | 1 | 183006 | 28262 | 1807.19 | 2060.91 | 2038.03 | 2383.96702 | 52137667 | 16946.38 | 9721.99 | Pacífico Sur | 6.48 | 284.90 | 2 | 3 |

ANEXO 6.

Características de los municipios de la Región 3 de la clasificación con el método de RNA

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Prod_Leche 2011 [lt] | Prod_Significativa leche | Prod_Significativa Carne | Prod_Significativa pesca | Pob_total | Suscriptores Res. | Consumo Est1 kWh | Consumo Est2 [kWh] | Consumo Est3 [KWh] | Consumo Est4 [kWh] | Consumo Resid. | Consumo Comerc. | Consumo Indust. | Subregión | Personas X vivienda | Cons. Per Cápita | Clasificación primaria | Cluster final |
|---------------------|----------------|-----------|---------|------|----------------------|--------------------------|--------------------------|--------------------------|-----------|-------------------|------------------|--------------------|--------------------|--------------------|----------------|-----------------|-----------------|-----------------------|---------------------|------------------|------------------------|---------------|
| COLON_GENOVA | 1750 | 38.24 | 3.06 | 4.05 | 1576 | 0 | 1 | 0 | 9959 | 2036 | 541.55 | 1024.51 | 0.00 | 0.00 | 1186005 | 788.50 | | Río Mayo | 4.89 | 119.09 | 3 | 1 |
| OSPINA | 2850 | 40.52 | 3.15 | 4.08 | 19696 | 1 | 0 | 0 | 8547 | 1759 | 542.70 | 673.35 | 776.07 | 0.00 | 1060157 | 266.67 | 1561.22 | Sabana | 4.86 | 124.04 | 3 | 1 |
| PROVIDENCIA | 2300 | 34.65 | 3.48 | 4.17 | 1800 | 0 | 0 | 0 | 12644 | 2215 | 367.84 | 631.23 | 0.00 | 0.00 | 677131 | 5651.00 | | Abades | 5.71 | 53.55 | 3 | 1 |
| SAN_PEDRO_DE_CARTAC | 1750 | 39.95 | 3.14 | 4.04 | 2400 | 0 | 0 | 0 | 7348 | 1323 | 468.43 | 770.00 | 0.00 | 0.00 | 658683 | 419.21 | 2960 | Juanambú | 5.55 | 89.64 | 3 | 1 |
| GUAITARILLA | 2650 | 41.12 | 3.06 | 3.99 | 5337 | 0 | 0 | 0 | 12661 | 3382 | 565.50 | 580.88 | 39.00 | 0.00 | 1924369 | 9178.36 | 2378 | Sabana | 3.74 | 151.99 | 3 | 2 |
| ILES | 2950 | 42.38 | 3.29 | 4.1 | 14902.8 | 0 | 0 | 0 | 8342 | 1951 | 495.88 | 644.47 | 1494.94 | 1227 | 1145079 | 4200.06 | 4212.75 | Exprovincia de Obando | 4.28 | 137.27 | 3 | 2 |
| IMUES | 2500 | 42.57 | 2.98 | 4.02 | 2748.6 | 0 | 0 | 0 | 6721 | 2373 | 597.65 | 691.27 | 994.59 | 0.00 | 1537659 | 5668.62 | 4923.83 | Sabana | 2.83 | 228.78 | 3 | 2 |
| LA_CRUZ | 2370 | 44.11 | 3.66 | 4.27 | 5300 | 0 | 0 | 0 | 18049 | 4230 | 697.30 | 1351.97 | 1687.22 | 0.00 | 3583750 | 5853.32 | 941.14 | Río Mayo | 4.27 | 198.56 | 3 | 2 |
| TANGUA | 2400 | 47.36 | 3.1 | 4.00 | 3965 | 0 | 0 | 0 | 10130 | 2486 | 643.97 | 700.22 | 1366.68 | 0.00 | 1685174 | 6050.69 | 1628 | Centro | 4.07 | 166.35 | 3 | 2 |
| YACUANQUER | 2700 | 42.22 | 3.24 | 4.15 | 9729 | 0 | 0 | 0 | 10582 | 2685 | 607.40 | 695.39 | 1239.86 | 2117 | 1708598 | 5406.13 | | Centro | 3.94 | 161.46 | 3 | 2 |
| CUMBAL | 3100 | 51.51 | 2.55 | 3.7 | 102296 | 1 | 0 | 1 | 34858 | 6174 | 658.22 | 846.30 | 1345.66 | 860 | 4403992 | 4814.28 | 19814.15 | Exprovincia de Obando | 5.65 | 126.34 | 3 | 3 |
| FUNES | 2300 | 55.46 | 3.39 | 4.17 | 1755 | 0 | 0 | 0 | 6685 | 1527 | 569.36 | 620.05 | 1378.52 | 0.00 | 909421 | 12742.87 | 2778 | Exprovincia de Obando | 4.38 | 136.04 | 3 | 3 |
| MALLAMA | 1809 | 56.54 | 2.77 | 3.83 | 2795 | 0 | 1 | 1 | 8338 | 2086 | 772.06 | 759.25 | 158.00 | 0.00 | 1603978 | 3664.56 | 1291 | Pie de Monte | 4.00 | 192.37 | 3 | 3 |
| ALDANA | 3050 | 39.56 | 2.18 | 3.61 | 21630 | 1 | 0 | 0 | 6397 | 1788 | 802.77 | 764.65 | 1272.10 | 3201.5 | 1420885 | 9383.34 | 1881.67 | Exprovincia de Obando | 3.58 | 222.12 | 3 | 4 |
| BELEN | 2450 | 38.51 | 1.69 | 3.34 | 1605 | 0 | 0 | 0 | 7133 | 1488 | 818.05 | 1467.58 | 1661.81 | 19 | 1402939 | 3200.39 | 5812.04 | Río Mayo | 4.79 | 196.68 | 3 | 4 |
| GUALMATAN | 2900 | 37.39 | 2.51 | 3.86 | 9900 | 0 | 0 | 0 | 5738 | 1660 | 679.98 | 758.57 | 1177.73 | 0.00 | 1221775 | 7071.81 | 1846.5 | Exprovincia de Obando | 3.46 | 212.93 | 3 | 4 |
| SAPUYES | 2900 | 45.63 | 2.4 | 3.59 | 44100 | 1 | 0 | 0 | 6797 | 2182 | 746.18 | 862.95 | 1673.09 | 0.00 | 1776246 | 8711.89 | 7501.65 | Sabana | 3.11 | 261.33 | 3 | 4 |

ANEXO 7.

Características de los municipios de la Región 4 de la clasificación con el método de RNA

| Municipio | Altitud (msnm) | Ruralidad | Miseria | NBI | Prod_Leche 2011 [lt] | Prod_Significativa leche | Prod_Significativa Carne | Prod_Significativa pesca | Pob_total | Suscriptores Res. | Consumo Est1 kWh | Consumo Est2 [kWh] | Consumo Est3 [KWh] | Consumo Est4 [kWh] | Consumo Resid. | Consumo Comerc. | Consumo Indust. | Subregión | Personas X vivienda | Cons. Per Cápita | Clasificación primaria | Cluster final |
|-----------|----------------|-----------|---------|------|----------------------|--------------------------|--------------------------|--------------------------|-----------|-------------------|------------------|--------------------|--------------------|--------------------|----------------|-----------------|-----------------|-----------------------|---------------------|------------------|------------------------|---------------|
| PASTO | 2527 | 32.07 | 1.28 | 2.79 | 73800 | 1 | 0 | 1 | 417484 | 94564 | 1217.48 | 1355.86 | 1606.56 | 1746.61704 | 135670821 | 5579.78 | 28049.63 | Centro | 4.41 | 324.97 | 4 | 1 |
| GUACHUCAL | 3285 | 41.83 | 2.43 | 3.69 | 126256 | 1 | 0 | 0 | 16152 | 4181 | 740.04 | 946.27 | 2498.38 | 0.00 | 3421172 | 5640.86 | 52407.12 | Exprovincia de Obando | 3.86 | 211.81 | 4 | 2 |
| IPIALES | 2892 | 45.08 | 2.33 | 3.42 | 73707 | 1 | 0 | 1 | 126335 | 25501 | 902.54 | 1277.05 | 1471.05 | 1639.49804 | 31411264 | 5474.57 | 8775.21 | Exprovincia de Obando | 4.95 | 248.63 | 4 | 2 |
| PUPIALES | 3014 | 39.64 | 2.54 | 3.71 | 78116.4 | 1 | 0 | 0 | 19029 | 4838 | 747.98 | 921.23 | 1170.19 | 0.00 | 4077257 | 5148.24 | 58451.35 | Exprovincia de Obando | 3.93 | 214.27 | 4 | 2 |
| TUQUERRES | 3070 | 37.17 | 2.83 | 3.72 | 63240 | 1 | 0 | 0 | 40966 | 10433 | 775.07 | 1108.27 | 1572.29 | 1996.1 | 9605326 | 6456.14 | 1757.82 | Sabana | 3.93 | 234.47 | 4 | 2 |
| POTOSI | 2750 | 50.4 | 2.56 | 3.75 | 18264.8 | 1 | 0 | 0 | 12527 | 2932 | 627.21 | 660.10 | 669.84 | 0.00 | 1885140 | 7394.06 | | Exprovincia de Obando | 4.27 | 150.49 | 4 | 3 |
| PUERRES | 2800 | 52.69 | 2.68 | 3.83 | 9350 | 0 | 0 | 0 | 8635 | 2434 | 638.27 | 811.23 | 1145.21 | 0.00 | 1732624 | 4664.87 | 1347 | Exprovincia de Obando | 3.55 | 200.65 | 4 | 3 |
| CORDOBA | 2867 | 48.51 | 3.5 | 4.28 | 7164 | 0 | 0 | 0 | 13844 | 4184 | 465.81 | 608.64 | 1792.33 | 0.00 | 2193281 | 4945.77 | 1700.75 | Exprovincia de Obando | 3.31 | 158.43 | 4 | 4 |
| CONTADERO | 2500 | 38.86 | 3.27 | 4.18 | 15240 | 1 | 0 | 0 | 6839 | 1639 | 659.60 | 714.53 | 1392.00 | 0.00 | 1128962 | 3851.86 | | Exprovincia de Obando | 4.17 | 165.08 | 4 | 5 |
| CUASPUD | 3050 | 39.96 | 3.09 | 4.04 | 25848 | 1 | 0 | 0 | 8395 | 1882 | 632.07 | 705.44 | 999.33 | 0.00 | 1264091 | 6243.27 | 14642 | Exprovincia de Obando | 4.46 | 150.58 | 4 | 5 |



Universidad de **Nariño**

ANEXO 8.
MANUAL DE USUARIO
CLASIFICACIÓN Y PREDICCIÓN DEL CONSUMO DE ENERGÍA ELÉCTRICA EN EL DEPARTAMENTO DE
NARIÑO
INTERFAZ GRÁFICA DE USUARIO

Camilo Ernesto Pazmiño
Gabriel Esteban Narvaez Morales
AUXILIARES DE INVESTIGACIÓN

COORDINADOR
John Barco Jiménez
Ing. Electrónico, Magíster en Ingeniería Electrónica

ANALISIS DE OPORTUNIDADES ENERGETICAS CON FUENTE ALTERNATIVAS EN EL
DEPARTAMENTO DE NARIÑO

UNIVERSIDAD DE NARIÑO
San Juan de Pasto – Nariño
Colombia
2015

TABLA DE CONTENIDO

INTRODUCCIÓN 80

| | | |
|--------|--|----|
| 1. | Módulo de Clasificación 81 | |
| 1.1. | Barra de Menú | 82 |
| 1.1.1. | Módulo de Clasificación | 82 |
| 1.1.2. | Módulo de Predicción | 83 |
| 1.1.3. | Análisis de Registro Histórico..... | 83 |
| 1.2. | Lista de Selección de Municipio..... | 84 |
| 1.3. | Total Habitantes..... | 85 |
| 1.4. | Total Viviendas..... | 85 |
| 1.5. | Producción Agrícola | 85 |
| 1.6. | Producción Pecuaria | 86 |
| 1.7. | Cuadro de Mapas | 86 |
| 1.8. | Botón Calcular..... | 87 |
| 1.9. | Clasificación Exitosa | 88 |
| 1.10. | Botón Limpiar..... | 89 |
| 2. | Módulo de Predicción 90 | |
| 3. | Módulo Análisis de Registro Histórico 94 | |
| 3.1. | Tratamiento de la información | 94 |

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1. Módulo de Clasificación | 81 |
| Figura 2. Barra de Menú | 82 |
| Figura 3. Opciones del Módulo de Clasificación en el Módulo de Clasificación. | 82 |
| Figura 4. Opciones del Módulo de Predicción en el Módulo de Clasificación. | 83 |
| Figura 5. Error al abrir el Módulo de Predicción | 83 |
| Figura 6. Opciones del Módulo de Análisis de Registro Histórico en el Módulo de Clasificación. | 83 |
| Figura 7. Lista desplegable para selección de municipio. | 84 |
| Figura 8. Total Habitantes..... | 85 |
| Figura 9. Total Viviendas..... | 85 |
| Figura 10. Sección de Selección de Productos Agrícolas. | 86 |
| Figura 11. Sección de selección de productos de explotación animal. | 86 |
| Figura 12. Cuadro de mapa con su título. | 87 |
| Figura 13. Botón Calcular..... | 87 |
| Figura 14. Mensaje de Error. | 88 |
| Figura 15. Mensaje de advertencia de selección de características..... | 88 |
| Figura 16. Mensaje de finalización del proceso de clasificación. | 89 |
| Figura 17. Botón Limpiar..... | 89 |
| Figura 18. Advertencia al seleccionar Limpiar | 89 |
| Figura 19. Módulo de Predicción | 90 |
| Figura 20. Elegir modelo ARIMA | 91 |
| Figura 21. Elección de parámetros | 91 |
| Figura 22. Mensaje de error, los valores deben ser enteros positivos..... | 91 |
| Figura 23. Consumo y Predicción..... | 92 |
| Figura 24. Parámetros modelo ARIMA | 92 |
| Figura 25. Prueba AIC/BIC..... | 93 |
| Figura 26. Mensaje de error, calcular los tres modelos..... | 93 |
| Figura 27. Mapa del <i>cluster</i> | 93 |
| Figura 28. Módulo Análisis de Registro Histórico | 95 |
| Figura 29. Selección de los pesos del filtro | 95 |
| Figura 30. Parámetros PSD | 96 |
| Figura 31. Periodicidad | 96 |
| Figura 32. Cuadros de visualización de las series | 97 |

INTRODUCCIÓN

Este documento presenta el manual explicativo de la interfaz gráfica de usuario desarrollada para realizar la clasificación, análisis y predicción de la demanda de energía eléctrica en los municipios de Nariño. Este software fue desarrollado por auxiliares de investigación adscritos al Grupo de Investigación en Ingeniería Eléctrica y Electrónica (GIIEE) en el marco del proyecto de Análisis de Oportunidades Energéticas para el Departamento de Nariño (ALTERNAR).

La aplicación fue desarrollada en el entorno Matlab®. Utiliza diferentes tipos de funciones para los distintos procesos que se llevan a cabo. El algoritmo de clasificación con que trabaja es el de Redes Neuronales Artificiales (RNA). El modelamiento y la predicción se realiza mediante métodos univariados ARIMA y un método desarrollado por los autores que utiliza la información de suscriptores y habitantes de las poblaciones objeto de estudio.

1. Módulo de Clasificación

Este es el módulo inicial que se abre cuando se ejecuta la aplicación.

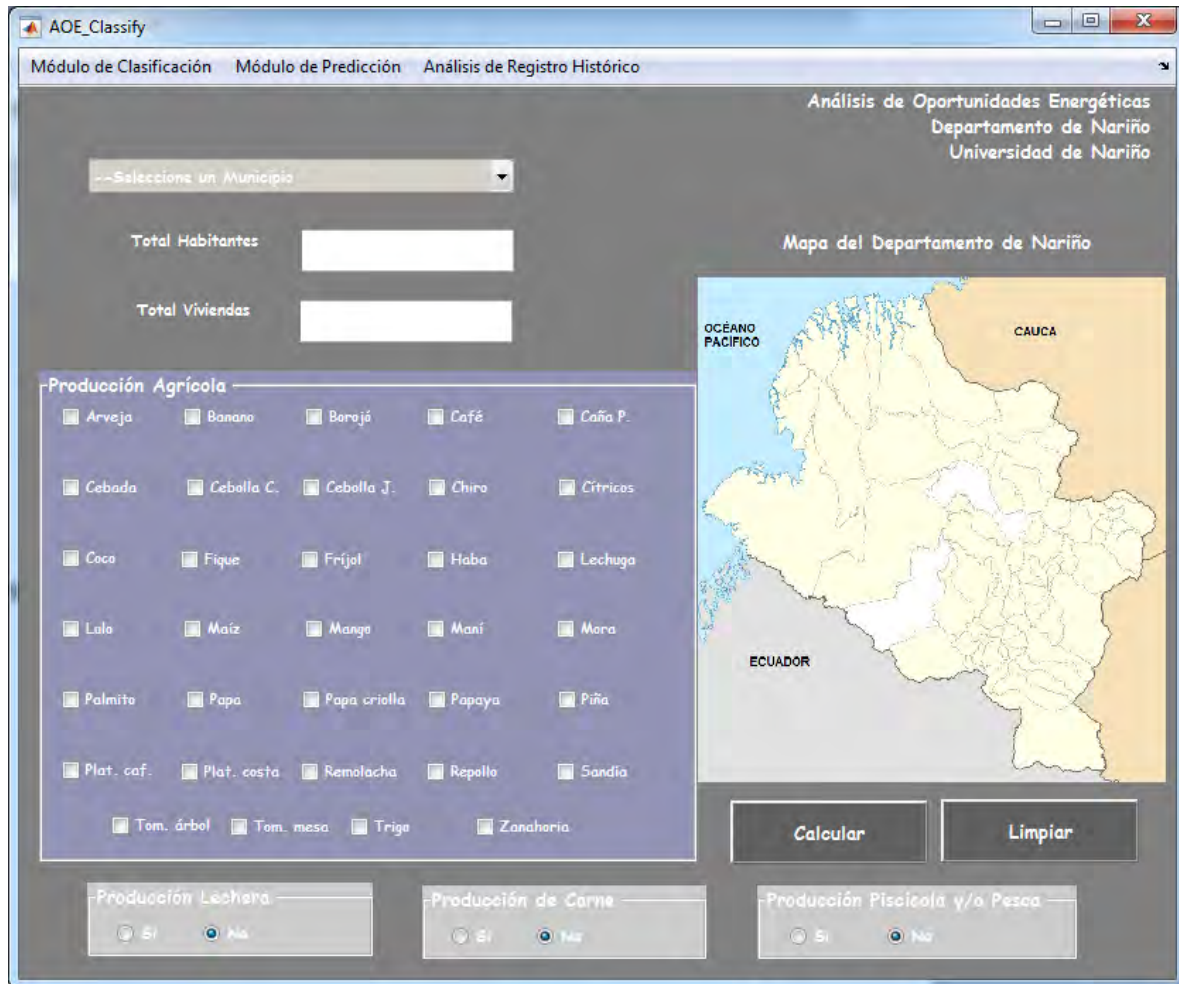


Figura 27. Módulo de Clasificación

El Módulo de Clasificación cuenta con 4 secciones con diferente tipo de ingreso de datos para realizar el proceso de clasificación. La primera de éstas es una lista desplegable en la cual puede seleccionarse el municipio al que pertenece la población que se quiere clasificar. Bajo la lista desplegable se encuentra dos cuadros para ingresar los valores correspondientes al total de habitantes de la población y el total de viviendas respectivamente. Más abajo se encuentra una gran sección con 34 casillas de selección que corresponden a productos agrícolas. En la parte inferior de la ventana se encuentran tres cuadros de selección única, en éstos puede seleccionarse la variable producción

pecuaria. Para la selección de las acciones que se pueden realizar aquí, existen dos secciones: la primera es la barra de menú en la parte superior; la segunda corresponde a los botones ubicados bajo el cuadro donde se grafica el mapa del Departamento. Por último, puede verse un cuadro donde se muestra el mapa del Departamento de Nariño.

1.1. Barra de Menú

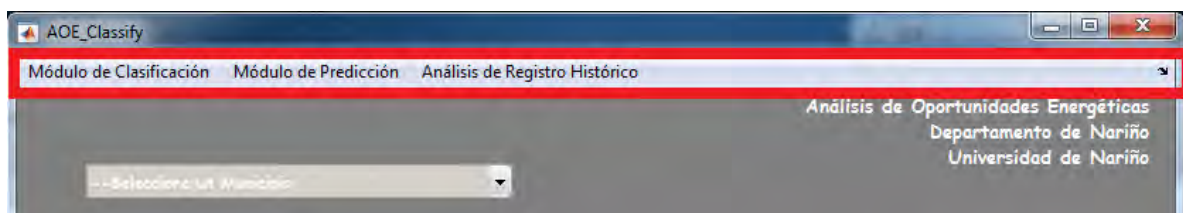


Figura 28. Barra de Menú

La Barra de Menú cuenta con tres opciones de selección: Módulo de Clasificación, Módulo de Predicción y Análisis de Registro Histórico.

1.1.1. Módulo de Clasificación

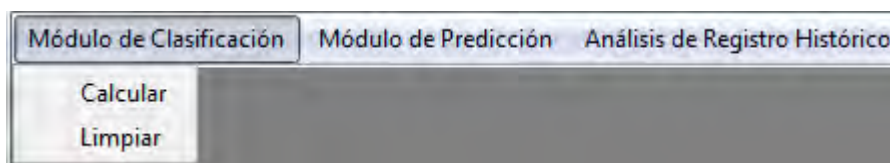


Figura 29. Opciones del Módulo de Clasificación en el Módulo de Clasificación.

Al seleccionar la opción del Módulo de Clasificación se despliega una lista con dos opciones: Calcular y Limpiar. Estas dos opciones realizan las mismas acciones que los botones que tienen los mismos nombres. Dichas acciones se explican más adelante.

1.1.2. Módulo de Predicción

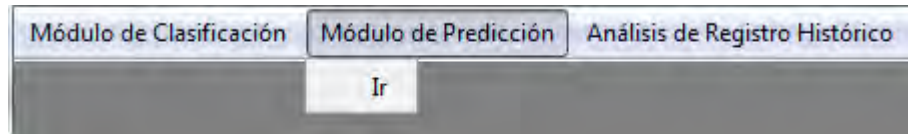


Figura 30. Opciones del Módulo de Predicción en el Módulo de Clasificación.

La selección del Módulo de Predicción despliega la opción Ir. La selección de ésta opción, permite acceder al siguiente módulo de la aplicación, donde pueden realizarse pruebas de predicción después de haber efectuado la clasificación de una población ingresando los datos solicitados en el Módulo de Clasificación. Si se selecciona la opción *Ir*, sin haber realizado el proceso de clasificación, aparece un cuadro de error (Figura 5), señalando que primero debe completarse el proceso actual.

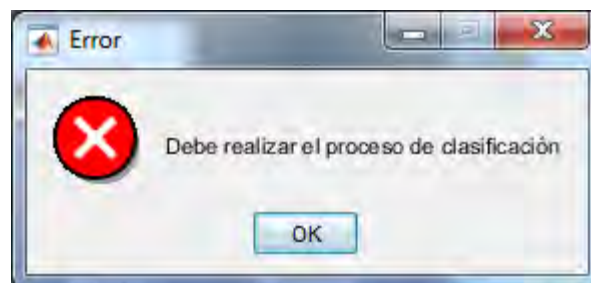


Figura 31. Error al abrir el Módulo de Predicción

1.1.3. Análisis de Registro Histórico

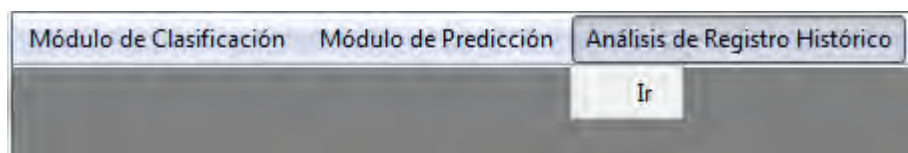


Figura 32. Opciones del Módulo de Análisis de Registro Histórico en el Módulo de Clasificación.

La selección del Análisis de Registro Histórico despliega la opción Ir. La selección de ésta opción, permite acceder a dicho módulo. Allí pueden realizarse diferentes pruebas a los datos de consumo organizados en series de tiempo (Por defecto se carga los datos

correspondientes al consumo de energía eléctrica reportados al Sistema Único de Información SUI en el periodo 2003-2014).

1.2. Lista de Selección de Municipio

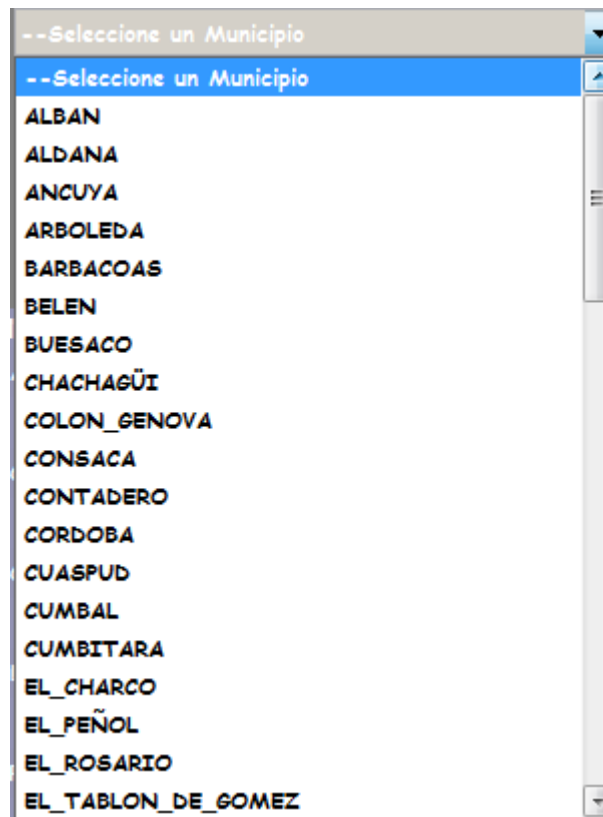


Figura 33. Lista desplegable para selección de municipio.

La selección de la lista desplegable permite escoger uno de los sesenta y tres municipios que conforman el Departamento de Nariño.

1.3. Total Habitantes



Total Habitantes

Figura 34. Total Habitantes.

Total Habitantes es un cuadro de captura de texto donde se debe introducir la cantidad de personas que habitan en la población objeto de estudio.

1.4. Total Viviendas



Total Viviendas

Figura 35. Total Viviendas.

Total Viviendas es un cuadro de captura de texto donde se debe introducir la cantidad de unidades residenciales existentes en la población objeto de estudio.

1.5. Producción Agrícola



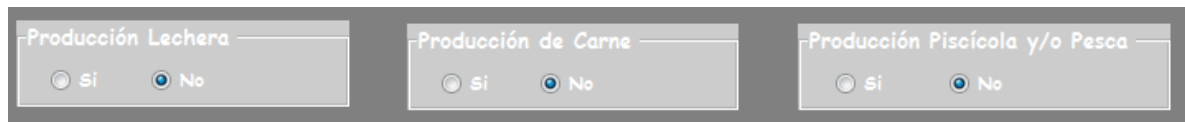
Producción Agrícola

| | | | | |
|-------------------------------------|--------------------------------------|---------------------------------------|------------------------------------|-----------------------------------|
| <input type="checkbox"/> Arveja | <input type="checkbox"/> Banano | <input type="checkbox"/> Borojó | <input type="checkbox"/> Café | <input type="checkbox"/> Caña P. |
| <input type="checkbox"/> Cebada | <input type="checkbox"/> Cebolla C. | <input type="checkbox"/> Cebolla J. | <input type="checkbox"/> Chiro | <input type="checkbox"/> Cítricos |
| <input type="checkbox"/> Coco | <input type="checkbox"/> Figue | <input type="checkbox"/> Fríjol | <input type="checkbox"/> Haba | <input type="checkbox"/> Lechuga |
| <input type="checkbox"/> Lulo | <input type="checkbox"/> Maíz | <input type="checkbox"/> Mango | <input type="checkbox"/> Maní | <input type="checkbox"/> Mora |
| <input type="checkbox"/> Palmito | <input type="checkbox"/> Papa | <input type="checkbox"/> Papa criolla | <input type="checkbox"/> Papaya | <input type="checkbox"/> Piña |
| <input type="checkbox"/> Plat. caf. | <input type="checkbox"/> Plat. costa | <input type="checkbox"/> Remolacha | <input type="checkbox"/> Repollo | <input type="checkbox"/> Sandía |
| <input type="checkbox"/> Tom. árbol | <input type="checkbox"/> Tom. mesa | <input type="checkbox"/> Trigo | <input type="checkbox"/> Zanahoria | |

Figura 36. Sección de Selección de Productos Agrícolas.

En esta sección se puede escoger los productos agrícolas cuya producción sea significativa en la economía de la población objeto de estudio. Un producto agrícola se considera significativo si la producción de éste es mucho mayor que la demanda de ese producto en la población objeto de estudio. Para la clasificación de una población es necesario que se seleccione por lo menos un producto. El modelo en el cual se basa la aplicación utilizó entre uno y cinco productos por municipio, por lo tanto se recomienda que la selección de productos agrícolas se encuentre en ese intervalo.

1.6. Producción Pecuaria



The image shows a horizontal interface with three selection boxes. The first box is titled 'Producción Lechera' and has radio buttons for 'Si' and 'No', with 'No' selected. The second box is titled 'Producción de Carne' and has radio buttons for 'Si' and 'No', with 'No' selected. The third box is titled 'Producción Piscícola y/o Pesca' and has radio buttons for 'Si' and 'No', with 'No' selected.

Figura 37. Sección de selección de productos de explotación animal.

La selección de productos de explotación pecuaria, al igual que en el caso de la que se realiza con los productos agrícolas, depende de lo significativo de la explotación de cada producto. En este caso, sí el municipio produce leche, carne y/o pescado en cantidades que superen ampliamente la demanda interna, se considera significativo. La selección se realiza escogiendo la opción *Si* en caso de que la obtención de alguno de estos productos sea significativa. A diferencia de la elección de productos agrícolas, el caso de la explotación animal puede tener todas sus selecciones en *No*.

1.7. Cuadro de Mapas

En la zona derecha de la interfaz se encuentra un cuadro donde se grafica la ubicación del municipio que se selecciona en la Lista Desplegable en el mapa del Departamento de Nariño. Cuando no se ha seleccionado ningún municipio o al momento de inicializar la aplicación, el mapa que se carga en esta sección corresponde al mapa del Departamento de Nariño.

En la parte superior del mapa se ubica el nombre del municipio que se encuentra seleccionado en el momento.



Figura 38. Cuadro de mapa con su título.

1.8. Botón Calcular

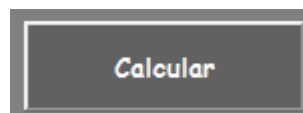


Figura 39. Botón Calcular

Al oprimir este botón se lleva a cabo el proceso de clasificación de la población objeto de estudio. Para que éste se realice de forma correcta, se debe haber seleccionado un municipio de la lista y se debe haber introducido los valores: Total Habitantes y Total Viviendas. En el caso de no haber efectuado alguno de estos pasos, aparecerá el mensaje de error que se muestra en la Figura 13 indicando cuál es el error. Sí la población objeto de estudio no cuenta con explotación significativa de productos

agrícolas, un cuadro de advertencia informa que no se ha realizado la selección de ninguno de los productos y pregunta si se desea continuar con el proceso. Para el caso de los productos obtenidos de la explotación animal, el proceso es el mismo que en el caso de la producción agrícola, el no seleccionar ningún tipo de explotación pecuaria no detiene el proceso de clasificación pero el usuario es advertido de la situación y la continuación del proceso se condiciona al criterio del usuario.



Figura 40. Mensaje de Error.

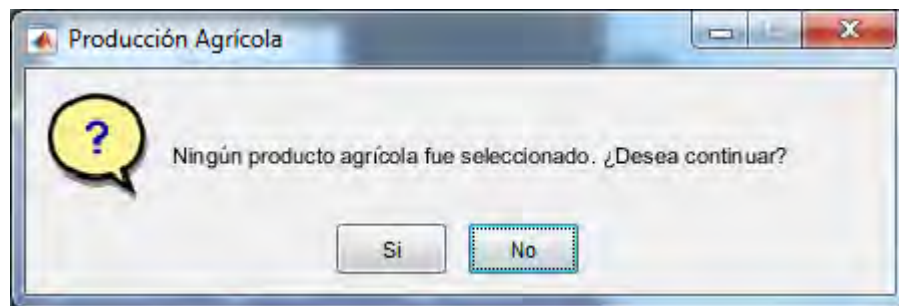


Figura 41. Mensaje de advertencia de selección de características.

1.9. Clasificación Exitosa

Al introducir todos los parámetros de clasificación y oprimir el botón calcular, el sistema entrega un mensaje donde muestra en qué región ha sido ubicada la población objeto de estudio y a que cluster pertenece. Los valores de clasificación que se muestran en el mensaje son utilizados en el módulo de predicción para las funciones de dicho módulo.

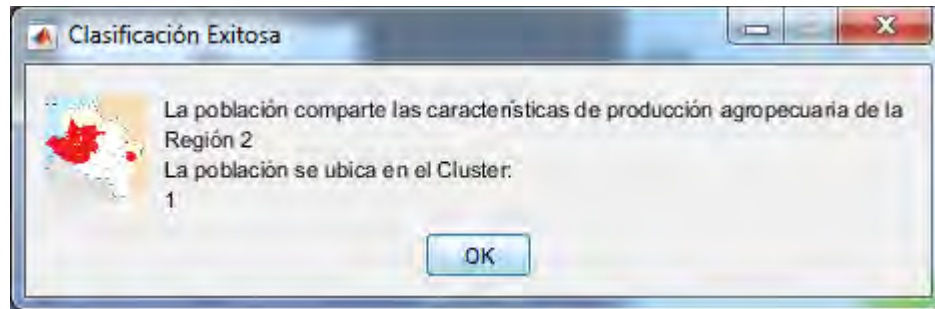


Figura 42. Mensaje de finalización del proceso de clasificación.

1.10. Botón Limpiar

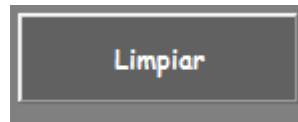


Figura 43. Botón Limpiar

El Botón Limpiar permite borrar todos los datos que hasta el momento se hayan seleccionado y reiniciar el sistema para una nueva clasificación. Tras ejecutar esta función no se podrá pasar al Módulo de Predicción, pues éste borra todos los datos de la clasificación como lo advierte el cuadro de mensaje que aparece al oprimir este botón.

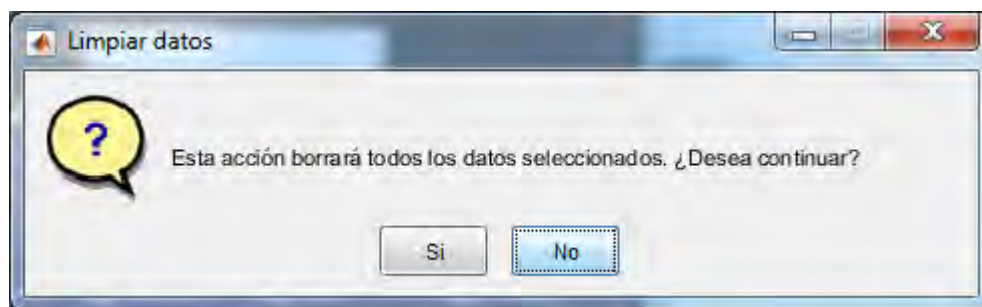


Figura 44. Advertencia al seleccionar Limpiar

2. Módulo de Predicción

Este módulo permite realizar el análisis de los modelos ARIMA, con el fin de seleccionar el que mejor se ajuste a la serie de consumo del *cluster* que se haya seleccionado según el Módulo de Clasificación. El primer paso es seleccionar uno de los tres modelos ARIMA del cuadro de botones llamado Predicción (Figura 20), al seleccionar uno de los modelos, se debe elegir los parámetros del mismo, por defecto aparecen unos valores, por ejemplo para el segundo modelo de la lista se debe elegir los parámetros P, D Q (Figura 21), éstos por defecto son $P=2$, $D=1$ y $Q=1$, también se debe seleccionar el número de meses al que se quiere realizar el pronóstico. También se debe tener en cuenta que los valores de los parámetros y el número de meses deben ser enteros positivos, de no serlo, un mensaje de error lo advierte (Figura 22).

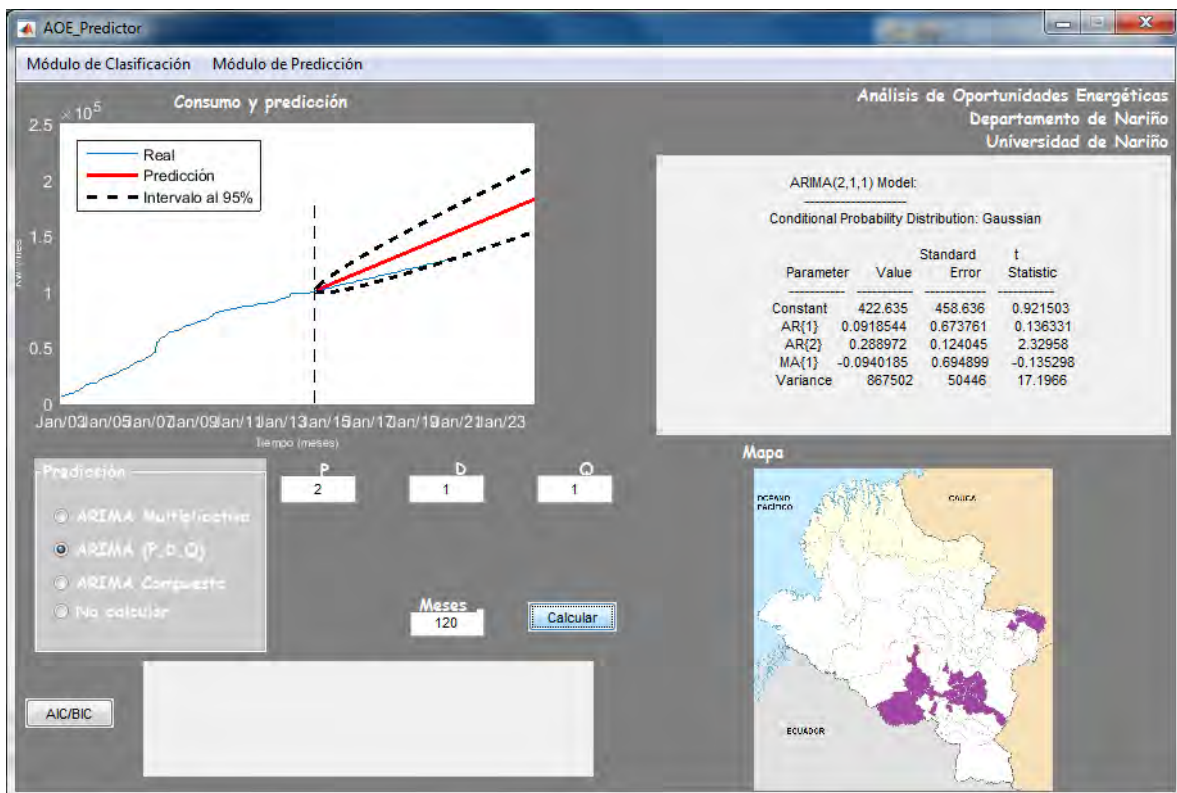


Figura 45. Módulo de Predicción

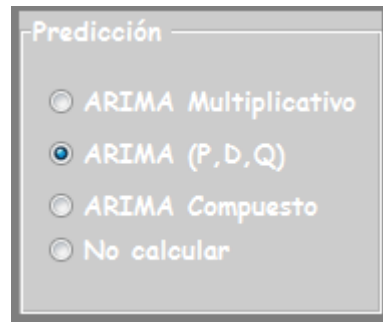


Figura 46. Elegir modelo ARIMA

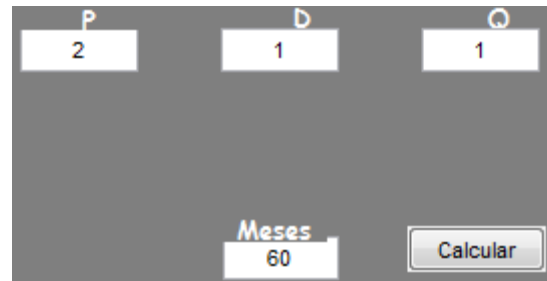


Figura 47. Elección de parámetros

Al presionar el botón Calcular, se grafica el comportamiento del consumo, la predicción realizada con la población y los suscriptores y el pronóstico realizado con el modelo. Estas gráficas se muestran en el cuadro llamado Consumo y Predicción (Figura 23). Una línea punteada paralela al eje “y”, divide el consumo en enero de 2014, que es el momento en que empieza la predicción, ya que la serie de consumo comprende el período entre enero de 2003 a diciembre de 2013. Se observa por tanto a partir de ese momento la predicción a máximo 84 meses, que se obtiene con la proyección realizada con la población y suscriptores, y la predicción realizada con el modelo ARIMA, según el número de meses seleccionado. En la figura 24 se puede observar los parámetros del modelo seleccionado.

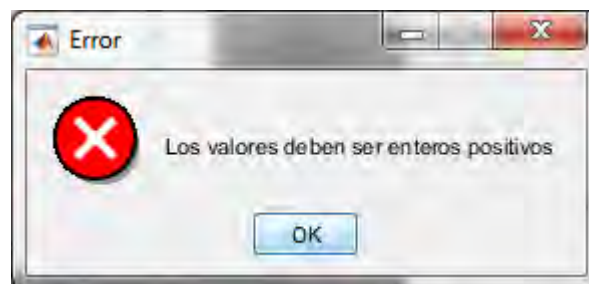


Figura 48. Mensaje de error, los valores deben ser enteros positivos

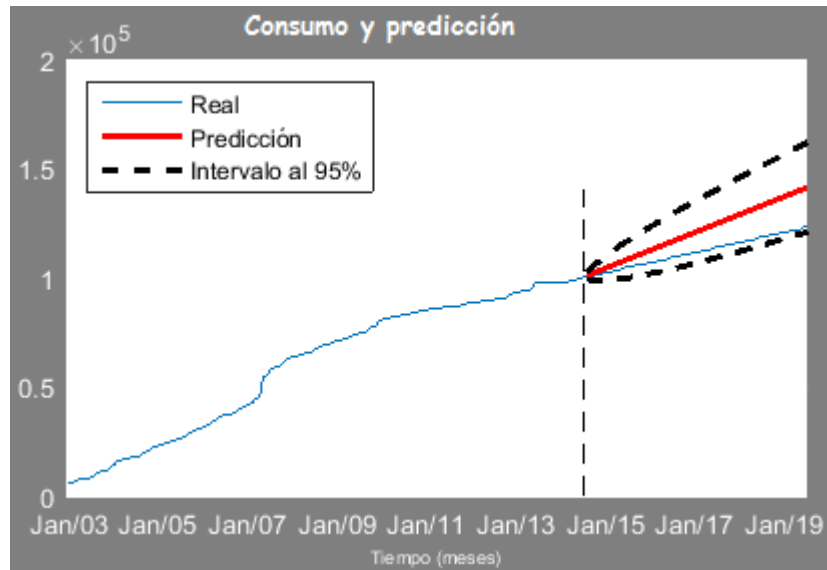


Figura 49. Consumo y Predicción

ARIMA(2,1,1) Model:

Conditional Probability Distribution: Gaussian

| Parameter | Value | Standard Error | t Statistic |
|-----------|------------|----------------|-------------|
| Constant | 422.635 | 458.636 | 0.921503 |
| AR{1} | 0.0918544 | 0.673761 | 0.136331 |
| AR{2} | 0.288972 | 0.124045 | 2.32958 |
| MA{1} | -0.0940185 | 0.694899 | -0.135298 |
| Variance | 867502 | 50446 | 17.1966 |

Figura 50. Parámetros modelo ARIMA

Una vez hecho el análisis con los tres modelos, se puede realizar una prueba AIC/BIC (Figura 25), para seleccionar el modelo más robusto, según el tipo de serie. Sí no se han ejecutado los tres modelos no se puede realizar la prueba y aparece un mensaje de error (Figura 26).

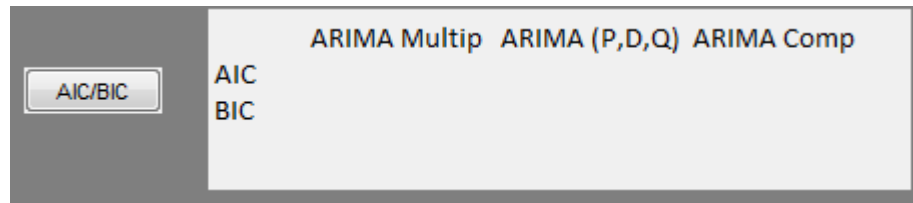


Figura 51. Prueba AIC/BIC

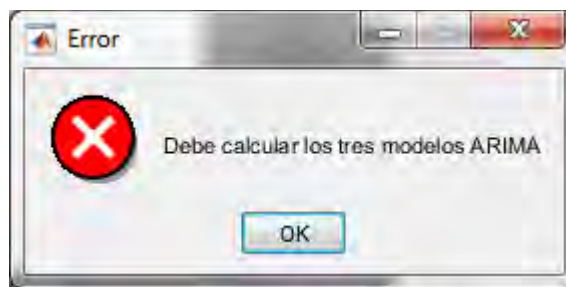


Figura 52. Mensaje de error, calcular los tres modelos

En la sección mapa, se muestra el *cluster* primario al que pertenece la población según el análisis realizado en el Modulo Clasificación (Figura 27).

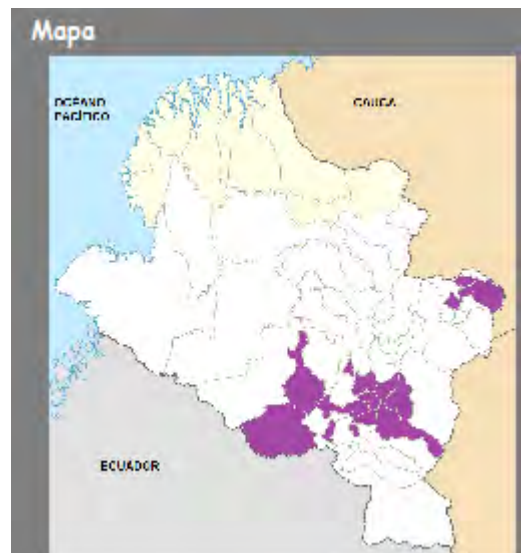


Figura 53. Mapa del *cluster*.

3. Módulo Análisis de Registro Histórico

3.1. *Tratamiento de la información*

La información de consumo de energía eléctrica pertenece a series de tiempo, por ende, se realizaron pruebas de filtrado y descomposición de las series, buscando eliminar información errónea de dichas series.

El módulo permite visualizar la información reportada al Sistema Único de Información de Servicios Públicos SUI por parte de las empresas prestadoras del servicio, esta información corresponde a los 58 municipios de Nariño que pertenecen a la zona interconectada (ZI), los cuales se muestran en la tabla 1.

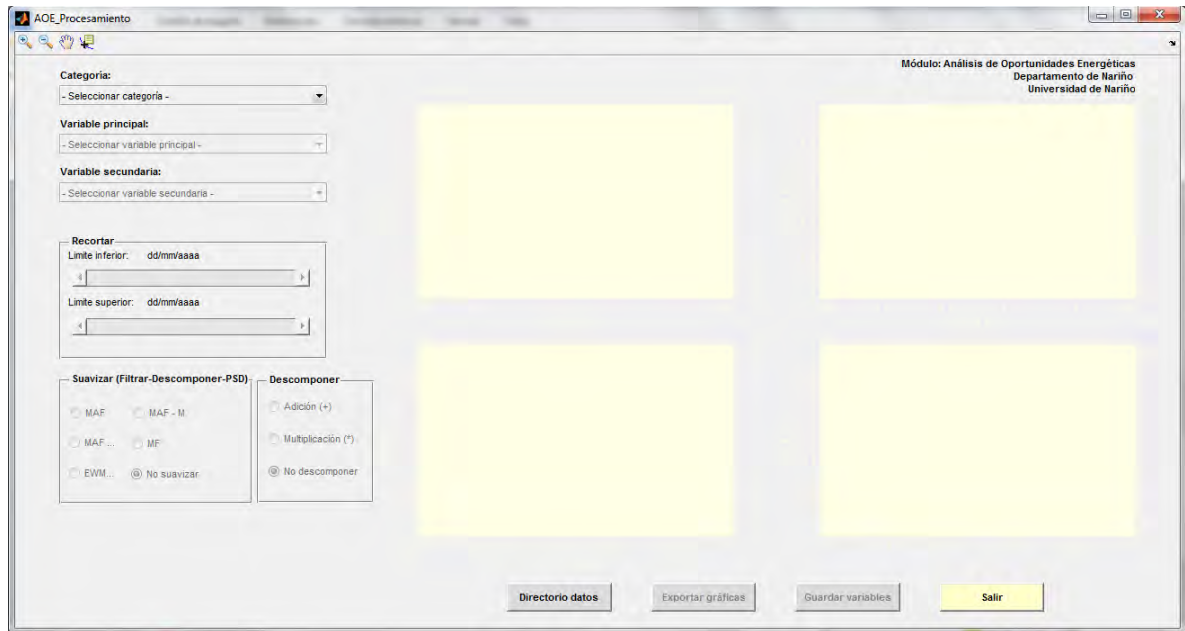


Figura 54. Módulo Análisis de Registro Histórico

La primera lista desplegable (Categoría) permite seleccionar la categoría en la que se agrupa la información a visualizar, las opciones son: Municipios, Subregión. La segunda lista (Variable Principal) permite la elección del municipio o subregión según se escogió en la primera lista, posteriormente se debe escoger la variable secundaria que se quiere visualizar en la tercera lista. El cuadro “Recortar” permite reducir la serie mes a mes entre los límites inferior y superior de la serie cargada. Al seleccionar una de las opciones dentro del cuadro “Suavizar” aparece una ventana (figura 29) en la cual se deben introducir los pesos del filtro, hay valores asignados por defecto que en la mayoría de casos funcionan de manera apropiada.

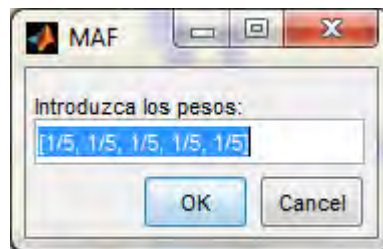


Figura 55. Selección de los pesos del filtro

El filtro MAF es un promedio móvil con promedio de ventana variable que puede estar o no relacionada con un patrón de repetición o estacionalidad. El MAF-M es un promedio

móvil ajustado a la repetición anual de las observaciones. El filtro EWMAF da más peso a los valores más recientes del conjunto de datos utilizados para realizar la suavización alrededor de cada observación de la serie.

El cuadro “Descomponer” permite seleccionar entre dos tipos de descomposición; aditiva y multiplicativa según el tipo de periodicidad de la serie, al igual que la opción anterior este despliega una ventana (figura 31) que permite seleccionar el valor en meses de repetición en la serie, al seleccionar el valor se despliega otra ventana (figura 30) para el cálculo de la densidad espectral de potencia PSD (por sus siglas en inglés), el primer valor realiza el cambio de la tasa de muestreo de meses a segundos y el segundo valor deja escoger el valor del orden del modelo auto-regresivo AR, estos valores también están ajustados por defecto. Los valores que se han cargado por defecto en las anteriores opciones fueron seleccionados basados en el curso *Applied Time Series Analysis* (Penn State), el cual explica paso a paso los parámetros que se deben tener en cuenta al descomponer una serie de tiempo para caracterizarla.

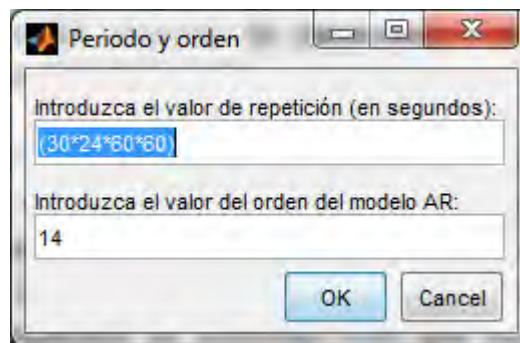


Figura 56. Parámetros PSD

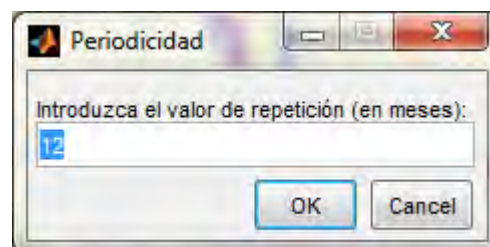


Figura 57. Periodicidad

Una vez se han seleccionado todas las opciones la interfaz muestra las gráficas de la variable secundaria seleccionada, en el primer cuadro de visualización. Junto a esta en el segundo cuadro aparece la gráfica de la tendencia que se obtiene al filtrar la señal quitando la información de estacionalidad. El tercer cuadro muestra la estacionalidad dependiendo de la selección que se haya hecho en la sección de descomposición, por último se presenta la imagen correspondiente a la PSD de la variable escogida (Figura 32).

En la parte inferior de la interfaz se encuentran cuatro botones; el primero de ellos permite seleccionar un directorio de datos diferente a la ubicación de los archivos de la interfaz, el segundo accede a una ventana donde se elige la ubicación y la extensión con que se quiere guardar las gráficas, en el tercero se puede guardar los datos de las gráficas que se presentan, finalmente el botón que permite salir de la aplicación.

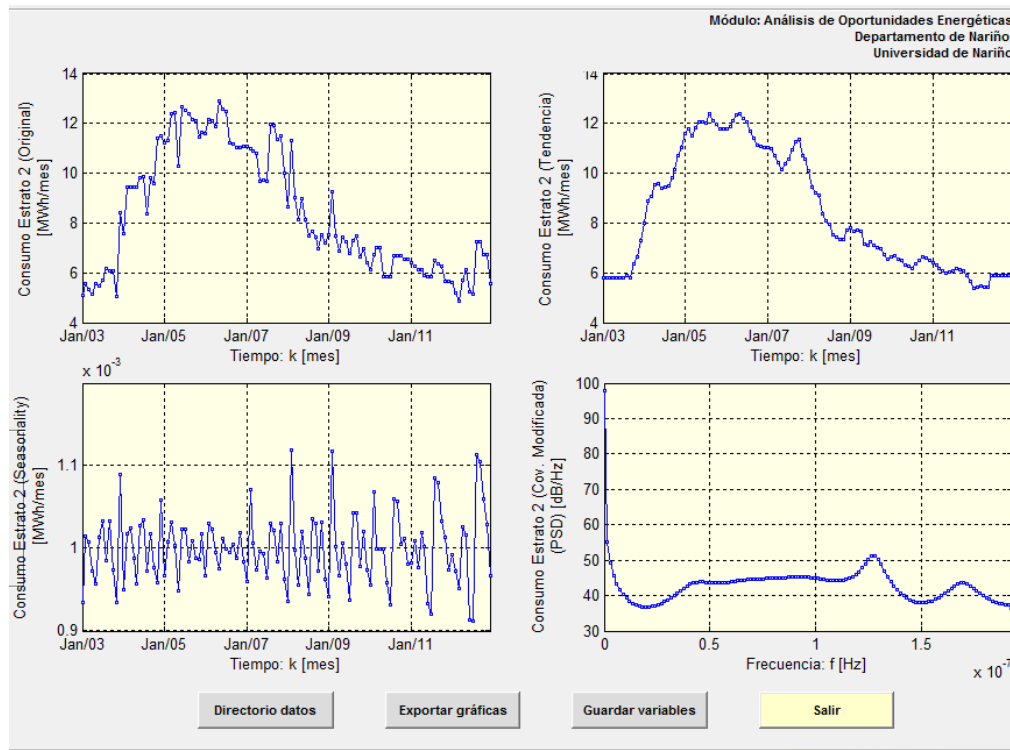


Figura 58. Cuadros de visualización de las series

4. Bibliografía

- Achicanoy Martínez, W. O. (2014). *Modelo de demanda de energía eléctrica para El Departamento de Nariño (Enfoque univariable clásico)*. San Juan de Pasto: Plan de Energización Rural Sostenible para el Departamento de Nariño PERS-Nariño.
- Akaike, H. (Diciembre de 1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, AC-19(6), 716-723.
- Alpaydin, E. (2010). *Introduction to Machine Learning*. London: The MIT Press.
- Arriaza, H. (2005). *Metodología para la inclusión de variables sociales en la formulación, ejecución y administración de proyectos de energía rural*. Quito: OLADE.
- Benavente, J. M., Galetovic, A., Sanhueza, R., & Serra, P. (2005). ESTIMANDO LA DEMANDA RESIDENCIALPOR ELECTRICIDAD EN CHILE: EL CONSUMO ES SENSIBLE AL PRECIO. *Cuadernos de Economía*, 42, 31-61.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Birmingham: Clarendon press - Oxford.
- Castillo Burbano, Á. M., & Jurado, J. A. (2014). *Caracterización social y económica del Departamento de Nariño*. San Juan de Pasto: Plan de Energización Rural Sostenible del Departamento de Nariño PERSN.
- Centro de Estudios de Desarrollo Regional y Empresarial CEDRE. (2012). *Indicadores socio-económicos Departamento de Nariño*. Pasto.
- de la Fuente Fernández, S. (2011). *SERIES TEMPORALES: MODELO ARIMA*. Madrid: Universidad Autónoma de Madrid.
- Departamento Administrativo Nacional de Estadística DANE. (7 de Abril de 2015). *Estadísticas Sociales: Necesidades Básicas Insatisfechas*. Obtenido de www.dane.gov.co
- Egelioglu, F., Mohamad, A., & Guven, H. (2001). Economic Variables and Electricity Consumption in Northern Cyprus. *Energy*, 26(4), 355-362.
- Fernández Sampedro, J. Á. (2006). *Descomposición tradicional de series temporales*. Madrid: Universidad Autónoma de Madrid.
- Garre, M., Cuadrado, J. J., & Sicilia, M. Á. (2005). *Comparación de diferentes algoritmos de clustering en la estimación de coste en el desarrollo de software*. Madrid: Universidad de Alcalá.
- Garrido Puerto, A. (2011). *MODELO DE PREVISIÓN GEOGRÁFICO Y TEMPORAL DE LA DEMANDA ELÉCTRICA*. Madrid.
- Gobernación de Nariño. (2004). *Informe de Gestión*. San Juan de Pasto.

- Gurrutxaga Goikoetxea, I. (2010). *Aportaciones a la clasificación no supervisada y a su validación. Aplicación a la seguridad informática*. Donostia.
- Llanos, D., Azuaje, A., & Romero, I. M. (4 de Junio de 2013). *Redes Neuronales*. Obtenido de <http://redesneuronal.blogspot.com/>
- MathWorks Inc. (2015). *MathWorks*. Obtenido de <http://www.mathworks.com>
- Medina Hurtado, S., & García Aguado, J. (2005). Protección de la demanda de energía en Colombia mediante un sistema de inferencia difuso neuronal. *Energética*(33), 15-24.
- Ministerio de Comercio, Industria y Turismo. (2013). *Departamento de Nariño*. Bogotá D.C.: Oficina de estudios económicos.
- Oropeza Rodríguez, J. L. (2006). Algoritmos y métodos para el reconocimiento de voz en español mediante sílabas. *Computación y Sistemas*, 270-286.
- Penn State. (s.f.). *psu*. (Eberly College of Science) Recuperado el Septiembre de 2014, de <https://onlinecourses.science.psu.edu/stat510/>
- Quispe Llanos, R. (1998). *Técnicas Estadísticas de Predicción*. Lima: Universidad Inca Garcilaso de la Vega.
- Rojas Zerpa, J. C. (2012). *Planificación del suministro eléctrico en áreas rurales de los países en vías de desarrollo: un marco de referencia para la toma de decisiones*. Zaragoza: Universidad de Zaragoza.
- Rueda, M., Velásquez Henao, J., & Franco Cardona, C. (2011). Avances recientes en la predicción de la demanda de electricidad usando modelos no lineales. *Dyna*(167), 36-43.
- Secretaría de Agricultura y Medio Ambiente - Gobernación de Nariño. (2014). *TRAZAS Sistema de Trazabilidad Agrícola*. Recuperado el Diciembre de 2014, de <http://trazasnarino.org/>
- Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel Methods For Pattern Analysis*. Cambridge: Cambridge University Press.
- Superintendencia de Servicios Públicos Domiciliarios. (2010). *Sistema Unico de Información de Sistemas Públicos*. Recuperado el Junio de 2014, de <http://www.sui.gov.co/>
- Tabares, H., & Hernández, J. (2008). Mapeo curvas típicas demanda de energía eléctrica del sector residencial, comercial e industrial de la ciudad de Medellín, usando redes neuronales artificiales y algoritmos de interpolación. *Revista Facultad de Ingeniería Universidad de Antioquia*(46), 110-118.
- Teknomo, K. (Julio de 2007). *Revoledu - Revolution Education*. Obtenido de <http://www.revoledu.org/>

The Pennsylvania State University, Department of Statistics. (2013). En *Peen State STAT 510 Applied Time Series Analysis* (págs. Lesson 3, 1:20). Pennsylvania : The Pennsylvania State University.

Unidad de Planeación Minero Energética, Subdirección de Planeación Energética - Grupo de Demanda Energética. (2013). *Proyección de Demanda de Energía Eléctrica en Colombia*. Bogotá.

Uriel Jiménez, E. (1992). *Análisis de series temporales: modelos ARIMA*. Valencia: Paraninfo.

Valencia, A. L., Lozano, C. A., & Moreno, C. (2007). Modelo de promedios móviles para el pronóstico horario de potencia y energía eléctrica. *El Hombre y la Máquina*(29), 96-105.

Viloria de la Hoz, J. (2007). Economía del departamento de Nariño: Ruralidad y aislamiento geográfico. *Documentos de trabajo sobre economía regional*, 5-76.