



UNIVERSIDAD  
DE CIENFUEGOS

Universidad de Cienfuegos

“Carlos Rafael Rodríguez”

Facultad de Ciencias Económicas y Empresariales

Tesis en opción al grado científico  
de Master en Ingeniería Industrial.

## **Pronóstico de la demanda de energía eléctrica del sector residencial en el municipio de Cienfuegos.**

Autor: MSc. Lázaro Daniel Alvarez Basanta

Tutor: DrC.T. Jenny Correa Soto

MSc. Miguel Santana Justiz

5 de abril de 2022

---

---

# RESUMEN.

---

La energía es considerada uno de los motores impulsores del desarrollo social, desarrollo sustentado en fuentes no renovables y contaminantes del medioambiente; de ahí la necesidad de su gestión en todos los ámbitos incluyendo los municipios. El análisis del “marco teórico-referencial” sobre la gestión energética local permitió identificar como problema científico de la investigación la necesidad de incorporar pronósticos confiables de la demanda energética en el sector residencial. En consecuencia, se plantea como objetivo general, aplicar un paquete de técnicas estadísticas que permitan pronosticar la demanda para este sector. Los principales resultados obtenidos en esta investigación se resumen en el desarrollo de un paquete de técnicas estadísticas soportadas en una herramienta computacional de código abierto. Los valores teórico, metodológico, práctico, social, económico y medioambiental de esta investigación están en presentar: un instrumento metodológico basado en la actualización, reconceptualización, integración y contribución a mejorar el desempeño municipal. La validación del instrumento metodológico se realizó mediante su aplicación en el gobierno local del municipio de Cienfuegos.

---

---

# Índice general

---

<b>RESUMEN.</b>	<b>II</b>
<b>ÍNDICE GENERAL.</b>	<b>III</b>
<b>INTRODUCCIÓN.</b>	<b>1</b>
<b>1. CONSIDERACIONES SOBRE DEMANDA ELÉCTRICA</b>	<b>8</b>
1.1. Gestión de la energía . . . . .	8
1.1.1. ISO 50001 . . . . .	9
1.1.2. Gestión Energética Local . . . . .	10
1.1.3. Gestión Energética Local en Cuba . . . . .	13
1.2. Planificación de la energía . . . . .	14
1.3. Técnicas de pronóstico de demanda eléctrica. . . . .	18
1.3.1. Regresión Lineal . . . . .	21
1.3.2. Regresión Splines adaptativa multivariada (MARS) . . . . .	22
1.3.3. Redes neuronales artificiales . . . . .	25
1.4. Herramientas computacionales de apoyo a la planificación energética. . . . .	27
1.5. Conclusiones parciales: . . . . .	32
<b>2. ANÁLISIS DE LA GESTIÓN ENERGÉTICA LOCAL DE CIENFUEGOS</b>	<b>33</b>
2.1. Caracterización del municipio de Cienfuegos. . . . .	33
2.1.1. Situación energética municipal . . . . .	37
2.2. Antecedentes de investigación. . . . .	44
2.3. Críticas . . . . .	48
2.4. Conclusiones parciales: . . . . .	58

<b>3. APLICACIÓN DE LAS TÉCNICAS ANALIZADAS EN EL PRONÓSTICO DE LA ENERGÍA ELÉCTRICA EN EL SECTOR RESIDENCIAL</b>	<b>59</b>
3.1. Análisis municipal. . . . .	59
3.2. Consejos Populares representativos. . . . .	69
3.2.1. Rancho Luna . . . . .	69
3.2.2. Gloria . . . . .	72
3.2.3. Caunao . . . . .	75
3.2.4. Tulipán . . . . .	78
3.2.5. Resumen Procesamiento de CP . . . . .	80
3.3. Conclusiones parciales: . . . . .	87
<b>Conclusiones Generales.</b>	<b>88</b>
<b>Recomendaciones.</b>	<b>89</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.</b>	<b>90</b>

---

# INTRODUCCIÓN.

---

La gestión energética (GE) debe formar parte del sistema de gestión de cualquier organización, dedicada a desarrollar e implementar su política energética (Borroto Nordelo & Monteagudo Yanes, 2006). En el año 2011 la Organización Internacional de Normalización (ISO) aprueba la norma ISO 50001: 2011 (Correa Soto *et al.*, 2014), actualizada en el 2018 (ISO, 2018) que comprende el Sistema de Gestión de la Energía, así como, requisitos con orientación para su uso orientados a reducir los costos de energía y a la mejora de su imagen corporativa; por este motivo para muchas organizaciones la GE se ha convertido en una prioridad (Jovanović & Filipović, 2016).

Sin embargo, la GE no es solo exclusiva de las organizaciones de producción o de servicios, sino que abarca a toda la sociedad, siendo una necesidad que los gobiernos locales la integren a su gestión pública. La importancia de esta integración radica en que los gobiernos locales fomenten la eficiencia energética y el uso de la energía limpia, debido a su influencia sobre los sectores de la sociedad, la promoción de políticas y programas para el uso de la energía, de ahí el término gestión energética local (GEL) (Correa Soto, González Pérez, & Hernández Alonso, 2017; Correa Soto *et al.*, 2018; Correa Soto, Salmerón, Eras, Rivera, & Viñales, 2021).

La GEL es la planeación estratégica de las necesidades y usos de energía en la localidad a corto, mediano y largo plazo, de manera que resulten en la implementación de un sistema energético eficiente, económico y amigable con el medio ambiente (Denis & Parker, 2009); entre los elementos que integran la GEL se encuentra la planificación, el uso de recursos y su incidencia en la sociedad (Correa Soto *et al.*, 2018, 2021).

La planificación de la energía, generalmente, se realiza dentro de las organizaciones gubernamentales pero puede ser llevada a cabo por grandes empresas de energía con la colaboración de las diferentes partes interesadas procedentes de organismos gubernamentales, empresas de servicios públicos, locales, instituciones académicas y otros grupos de interés. En ella se debe reflejar los resultados de crecimiento de la poblacional; una nueva tendencia se conoce como Planificación de la Energía Sostenible, que ofrece un enfoque más integral al problema de la planificación de las necesidades futuras de energía (Ferrari, Bruni, Bramonti, *et al.*, 2020).

Existen muchas bases de conocimientos en la temática de la planificación de la energía en su fase inicial, pero poca voluntad de participar en un proceso intermedio que resultaría en la

transferencia de conocimientos y la creación de consenso debido a la incertidumbre sobre cómo se realizarán estas visiones en el futuro. Los municipios no tienen el poder para garantizar las ambiciones energéticas establecidas en la fase de planificación. Los planificadores se muestran escépticos sobre los proyectos que promueven el desarrollo de la eficiencia energética, ya que no pueden crear vínculos legales para los objetivos energéticos, lo que puede llevar a que las partes no cumplan lo acordado (Lindkvist *et al.*, 2019).

Los planes se basan en la trayectoria que han tenido los consumos y las ofertas energéticas, apoyándose en la situación actual y en las perspectivas, para proyectarse hacia el futuro. Esto significa que deben estar incluidos dentro de uno o varios planes económicos a mediano plazo, permitiendo mantener líneas de actuaciones claras y concretas, sobre las inversiones necesarias realizar. De ahí la necesidad de pronóstico, este servirá de guía a la planificación de la producción, comercio o distribución, operacionabilidad, etc. (Khuntia, Rueda, & Meijden, 2016; Hong & Fan, 2016; Ravadanegh, Jahanyari, Amini, & Taghizadeghan, 2016).

La demanda de electricidad depende de diversas variables: las climáticas (temperatura, velocidad del viento, humedad, etc.), demográficas (la distribución de la población, viviendas, industria, etc.) y estilos de vida vistos en intensidad de las actividades comerciales y cotidianas (horas pico frente a horas no pico, días laborables frente a fines de semana, feriados, vacaciones, etc.) (Black & Henson, 2013).

Las técnicas para realizar pronósticos se pueden clasificar en dos grupos principales: estadísticas y de inteligencia artificial, aunque los límites entre estos grupos son bastante ambiguos como resultado de la colaboración multidisciplinaria de la comunidad científica (Weron, 2014; Liu, Nowotarski, Hong, & Weron, 2015; Khuntia *et al.*, 2016). La literatura hace referencia a técnicas estadísticas, basadas en regresión o series cronológicas; y de inteligencia artificial (IA) (Xie, Hong, & Stroud, 2015; Hong & Fan, 2016; Bezerra, Veiga, Barroso, & Pereira, 2016).

En Cuba se ha demostrado la necesidad de incorporar la gestión de la energía a la gestión de gobierno local, a partir de la necesidad de impulsar un cambio de la matriz energética cubana hacia la gestión local, dado a que esta se sustenta en el 95,7% de combustibles fósiles; requiriendo un cambio en su estructura con una mayor penetración de las FER. El contexto cubano favorece esta transición dado a que el país con la proyección de la actualización del Modelo Económico y Social en el 2011 se inicia un cambio de enfoque hacia la energía sostenible, en el 2014 con la aprobación la Política para el desarrollo prospectivo de las fuentes renovables y el uso eficiente de la energía, haciendo énfasis en elevar la eficiencia energética, un cambio de la estructura de la matriz energética actual y su relación con la competitividad de la economía nacional; disminuyendo la dependencia de estos combustibles importados, sus costos energéticos y el

impacto medioambiental (Correa Soto *et al.*, 2017, 2018; Rodríguez Figueredo, 2019; Correa Soto *et al.*, 2021).

En el año 2017 se aprueban las bases del Plan Nacional de Desarrollo Económico y Social (PNDES) hasta el 2030 (PCC, 2017) relacionado con los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas para ese periodo (ONU, 2015). Con la declaración de la protección de los recursos y el medioambiente como dimensiones del desarrollo sostenible y ejes estratégicos en el PNDES hasta el 2030, así como, la actualización de los lineamientos de la Política Económica y Social referentes a los territorios, con el de la política energética (PCC, 2017; Correa Soto *et al.*, 2017), el Decreto-Ley No. 345/ 2017 “Del desarrollo de las fuentes renovables y el uso eficiente de la energía”, así como las instrucciones y resoluciones complementarias al respecto (Consejo de Estado, 2019); rectorado por Constitución de la República de Cuba que en su artículo 169 reconoce la autonomía del municipio y en el artículo 75, la responsabilidad de proteger el medio ambiente y los recursos naturales con estrecha vinculación al desarrollo sostenible de la economía y la sociedad.

El municipio de Cienfuegos ha ido incorporando la GEL a su gestión de gobierno, se han realizado una serie de investigaciones al respecto (Armas & Santana Cruz, 2017; Hurtado Espinosa, 2017; Díaz La Hoz, 2019), haciéndose un acercamiento al pronóstico de la demanda de energía eléctrica en el municipio. Sin embargo, un aspecto a señalar, en las investigaciones es que no se realizan las pruebas para analizar el cumplimiento de los supuestos de RLM, lo que puede conducir a la invalidez de los resultados al no cumplirse alguno de ellos (normalidad, homocedasticidad y la no autocorrelación de los residuos); mientras que la violación del resto de los supuestos solo afecta la veracidad y confiabilidad del modelo de pronóstico propuesto.

En los trabajos de Armas y Santana Cruz (2017), Hurtado Espinosa (2017) hace referencias a seleccionar el mejor modelo a partir de todas las combinaciones posibles, pero se presentan solamente cinco combinaciones para el pronóstico de demanda eléctrica, en cuanto a la ANN propuesta por Díaz La Hoz (2019) presenta una debilidad, y es que debido al uso de pesos aleatorios como parte de su función de transferencia, los resultados varían de en cada ejecución. Otra cuestión es que no ofrece un modelo funcional lineal, adicionalmente, se tiene su alto costo computacional, debido a que se utilizan herramientas auxiliares privativas (MatLab, Statgraphics) lo que conduciría al pago de licencias de uso así como sus actualizaciones en el paso del tiempo. Todo lo anterior planteado constituye la Situación Problemática.

Por tal razón, se plantea el siguiente:

## **Problema científico:**

¿Cómo contribuir a un pronóstico confiable de la demanda de energía eléctrica en el sector residencial en el municipio de Cienfuegos?

## **Hipótesis:**

La aplicación de un paquete de técnicas estadísticas soportadas en una herramienta computacional de código abierto, contribuye al pronóstico de la demanda de energía eléctrica en el sector residencial en el municipio de Cienfuegos.

**Variable independiente:** paquete de técnicas estadísticas.

**Variable dependiente:** pronóstico de la demanda de energía eléctrica en el sector residencial.

## **Objetivo general de la investigación:**

Aplicar un paquete de técnicas estadísticas soportadas en una herramienta computacional de código abierto que permitan pronosticar la demanda de energía eléctrica en el sector residencial en el municipio de Cienfuegos.

## **Objetivos específicos:**

- Construir un marco teórico-referencial de la investigación relacionado con la gestión de la energía, la gestión energética local, la planificación de la energía, técnicas de pronósticos de demanda eléctrica y herramientas computacionales de apoyo a la planificación energética.
- Realizar un análisis crítico de la GEL en el municipio de Cienfuegos.
- Pronosticar la demanda de energía eléctrica en el sector residencial del municipio de Cienfuegos mediante técnicas estadísticas soportadas en una herramienta computacional de código abierto.

## **Justificación de la investigación:**

La planificación de la demanda de energía eléctrica en Cuba para el sector residencial considera los consumos históricos, sin tener en cuenta otros elementos que inciden directamente en ellos, además esta planificación solo se realiza a nivel de municipio y circuitos eléctricos.

## **Metodología usada en la investigación:**

### **De los métodos teóricos:**

El **analítico–sintético**, para poder estructurar el fenómeno en sus múltiples relaciones y componentes y así facilitar su estudio.

El **Inductivo–deductivo**, para partir de hechos aislados y arribar a proposiciones generales.

El **Histórico–Lógico**, para resumir y sintetizar lo estudiado sobre la regresión múltiple, la selección de modelo y los métodos multicriterios.

La **Modelación**, específicamente el modelo teórico sobre el fenómeno de regresión y su asociación a técnicas computacionales.

El **tránsito**, de lo abstracto a lo concreto que permitió establecer relación y comparación entre la interpretación de los resultados del procesamiento y la realidad objetiva.

### **De los métodos empíricos:**

La **observación**, para la interpretación del sistema y sus partes.

El **análisis de documentos**, para la revisión de los modelos y métodos aplicables a la temática en cuestión.

### **De los métodos estadísticos matemáticos:**

El **empleo de técnicas estadísticas** como el método de los mínimos cuadrados ordinarios para la estimación de la regresión, las pruebas estadísticas para la validación de modelos y el cumplimiento de los supuestos.

Para el exitoso desarrollo de esta investigación se hizo necesario una amplia revisión bibliográfica que unida a técnicas de búsquedas, procesamientos y análisis de la información arrojó resultados valiosos. Entre los artículos consultados se exponen pruebas estadísticas, sus limitaciones, así como, alternativas robustas para la detección de problemas asociados al análisis de regresión múltiple.

La investigación forma parte del trabajo de tesis de doctoral “INSTRUMENTO METODOLÓGICO PARA LA GESTIÓN ENERGÉTICA EN LOS ÓRGANOS DE GOBIERNO LOCAL EN CUBA”, y responde a los siguientes proyectos:

- Apoyo a la Eficiencia energética y el uso sostenible de las Fuentes Renovables de Energía (FRE) en el municipio Cienfuegos. Programa PTCTI-02 Desarrollo Local y Territorio

(2020-2023).

- Modelo de Gestión Energética en los órganos cubanos de gobierno local. Fase II (2018-2021). Proyecto Nacional Asociado a Programa Ministerio de Energía y Minas, Ministerio de Ciencia Tecnología y Medio Ambiente, Oficina Nacional Uso Racional de la Energía, Gobierno de Municipio de Cienfuegos.
- Modelo de Gestión Energética en los órganos cubanos de gobierno local. Financiamiento desde (enero 2015- diciembre 2018). Proyecto Nacional Asociado a Programa Ministerio de Energía y Minas, Ministerio de Ciencia Tecnología y Medio Ambiente, Oficina Nacional Uso Racional de la Energía, Gobierno de Municipio de Cienfuegos.
- A Cuban network of cleaner production (CP) centres and strengthening education and research on CP. International Project. Katholieke Universiteit Leuven (KU Leuven). Belgium.
- Fortaleciendo capacidades en procesos de articulación y encadenamientos en la Ciudad Patrimonial para dinamizar el Desarrollo territorial Proyecto Internacional Plataforma para el desarrollo integral territorial (auspiciado por el Programa de Naciones Unidas para el Desarrollo, PNUD). Fase I y Fase II.

**El informe está estructurado en:**

- Resumen
- Introducción
- Tres Capítulos
- Conclusiones Generales
- Recomendaciones
- Referencias bibliográficas.

**Capítulo 1:** En este capítulo se realiza una revisión bibliográfica sobre la gestión de la energía, la gestión energética local, la planificación de la energía, técnicas de pronósticos de demanda eléctrica y herramientas computacionales de apoyo a la planificación energética.

**Capítulo 2:** Se realiza una caracterización del municipio de Cienfuegos, la situación energética actual, antecedentes y un análisis crítico de la GEL en el municipio.

**Capítulo 3:** Se realiza el pronóstico de la demanda de energía eléctrica en el sector residencial del municipio de Cienfuegos mediante la aplicación de las técnicas RLM, ANN, MARS, para las cuales se hace uso de varias librerías de Python: statsmodels, sklearn, pyearth.

**Conclusiones y recomendaciones** Las conclusiones brindan el análisis del alcance de los resultados de este trabajo y se compara con lo propuesto anteriormente, haciendo énfasis en el cumplimiento de los objetivos planteados.

Las recomendaciones se centran en futuras líneas de trabajo y resultados que pudieran aplicarse en otros campos de la ciencia.

---

# CONSIDERACIONES SOBRE DEMANDA ELÉCTRICA

---

En este capítulo se provee una visión general de las diferentes metodologías y técnicas utilizadas para el pronóstico de demanda eléctrica más utilizadas. Técnicas se refiere a un grupo de modelos tales como, regresión lineal múltiple (MLR) y redes de inteligencia artificial (ANN). Por otra parte, se utiliza metodologías para representar el esquema general de solución que puede ser implementado con múltiples técnicas.

## 1.1. Gestión de la energía

La gestión energética (GE) es parte del sistema de gestión de una organización dedicada a desarrollar e implementar su política energética. La GE o administración de la energía es un subsistema de la gestión organizacional que abarca las actividades de administración y aseguramiento que le confieren a la organización la aptitud para satisfacer de forma eficiente sus necesidades energéticas (Borroto Nordelo & Monteagudo Yanes, 2006; Ortiz Alvarado, 2015).

Con la aprobación de la norma ISO 50001: 2011 “Energy Management Systems - Requirements with guidance for use.” por la Organización Internacional de Normalización (ISO), como resultado de normas técnicas desarrolladas por países como Dinamarca en el año 2001, Suecia en el 2003, Estados Unidos e Irlanda en el 2005, España en el 2007 y la Unión Europea en el 2009 (Correa Soto *et al.*, 2018)); ha traído como consecuencia el aumento del interés internacional en la GE. Por este motivo para muchas organizaciones la GE se ha convertido en una prioridad por lo que se esfuerzan en reducir los costos de energía, ajustándose a los requisitos reglamentarios y por ende a mejorar su imagen corporativa (Jovanović & Filipović, 2016).

Para gestionar el análisis y la evaluación de la planificación, se han implementado numerosas herramientas cuantitativas. Se han propuesto varios métodos para clasificar las herramientas, particularmente para el modelado de sistemas energéticos. Estas se pueden clasificar como herramientas integrales y auxiliares. Las auxiliares son aquellas que generan información adicional que se necesita para que las herramientas integrales realicen análisis futuros (por ejemplo, basado en Delphi como un método auxiliar para la previsión). Las herramientas

integrales son aquellas que pueden evaluar todo el sistema energético desde el análisis primario de energía hasta la energía útil y los servicios para todos los portadores energéticos; evalúan los aspectos técnicos, económicos y ecológicos de los planes; y realizan un análisis intersectorial, ejemplos de estos se pueden apreciar en la tabla 1.1 (Mirakyan & De Guio, 2013).

Tabla 1.1: Herramientas computacionales. Fuente: (Mirakyan & De Guio, 2013)

Software	Foco de la aplicación	Método	Licencia	Enlace
Energy plan	Energía integrada Análisis ambiental	Simulación	Libre para uso académico	<a href="http://energy.plan.aau.dk">energy.plan.aau.dk</a>
etransport	Energía integrada Análisis ambiental	Optimización	Libre para uso académico	<a href="http://www.sintef.no">www.sintef.no</a>
ETEM ó MARKAL-Lite	Energía integrada Análisis ambiental	Optimización	Comercial	<a href="http://www.ordecys.com">www.ordecys.com</a>
GEMIS	Análisis de ciclo de vida	Simulación Conteo físico	Libre	<a href="http://www.gemis.de">www.gemis.de</a>
LEAP	Energía integrada Análisis ambiental	Simulación Conteo físico Optimización	Comercial y Libre	<a href="http://www.energycommunity.org">www.energycommunity.org</a>
(MESAP) PlaNet	Energía integrada Análisis ambiental	Simulación y conteo físico	Comercial	<a href="http://www.seven2one.de">www.seven2one.de</a>
MARKAL/TIMES	Energía integrada Análisis ambiental	Programación Lineal	Comercial	<a href="http://www.etsap.org">www.etsap.org</a>
Anylogic	Genérico	Simulación de eventos discretos Sistemas dinámicos Modelación basado en agentes	Comercial	<a href="http://www.xjtek.com">www.xjtek.com</a>
iThink	Genérico	Sistemas dinámicos	Comercial	<a href="http://www.iseesystem.com">www.iseesystem.com</a>
Vensim	Genérico	Sistemas dinámicos	Comercial	<a href="http://www.vensim.com">www.vensim.com</a>
Simapro	Análisis de ciclo de vida	Simulación y conteo físico	Comercial y Libre	<a href="http://www.simapro.com">www.simapro.com</a>
Umberto	Análisis de ciclo de vida	Simulación y conteo físico	Comercial	<a href="http://www.umberto.de">www.umberto.de</a>

### 1.1.1. ISO 50001

La solicitud para el desarrollo de la norma internacional ISO 50001: 2018 de gestión de la energía provino de la Oficina de las Naciones Unidas para el Desarrollo Industrial (ONUDI), quien reconoció que la industria necesitaba plantear una respuesta efectiva al cambio climático. Para la ISO, la gestión energética fue uno de los cinco campos principales dignos para el desarrollo de Normas Internacionales. La gestión eficaz de la energía es una prioridad, ya que cuenta con un potencial significativo en cuanto al ahorro de energía y la reducción de las emisiones de gases de efecto invernadero en todo el mundo (Silva Goncalves & Santos, 2019)(ISO, 2018).

El objetivo de este estándar internacional es permitir a las organizaciones establecer los sistemas y procesos necesarios para mejorar el rendimiento en el uso de la energía. El estándar lleva a reducciones de costo, emisiones de gases de efecto invernadero y otros impactos ambientales por medio de la gestión sistemática de la energía. Es aplicable a todo tipo de organizaciones independientemente de su ubicación geográfica, condiciones culturales o sociales. La implementación acertada depende del compromiso de todos los niveles y funciones de la organización y sobre todo de la dirección superior (Silva Goncalves & Santos, 2019)(ISO, 2018).

En la norma se definen los requisitos para un sistema de gestión energética (SGE), para desarrollar e implantar una política energética, establecer objetivos, metas y planes de acción, teniendo en cuenta los requisitos legales y la información pertinente al uso significativo de energía. Donde el SGE permite a una organización alcanzar sus compromisos de política, tomar las acciones que sean necesarias para mejorar su desempeño energético y demostrar la conformidad del sistema con los requisitos de esta Norma Internacional. La norma se basa en el ciclo de mejora continua Planear-Hacer-Verificar-Actuar e incorpora la gestión energética en las prácticas organizacionales diarias como se puede apreciar en la figura 1.1 (Correa Soto *et al.*, 2017, 2018).

La aplicación global de esta norma contribuye a lograr un uso más eficiente de las fuentes de energía disponibles, incrementar la competitividad y reducir el impacto ambiental asociado al uso de la energía, al establecer un marco internacional para la gestión de todos los aspectos relacionados con la energía, incluidos su uso y adquisición, por parte de las instalaciones industriales y comerciales, o de las compañías en su totalidad. La norma sugiere a las organizaciones las estrategias y herramientas de gestión, como los indicadores energéticos, para incrementar su eficiencia energética, reducir costos y mejorar su desempeño ambiental. Las bases de este enfoque se muestran a continuación en la figura (1.1) (Campillo Sabina, 2018).

Las ciudades juegan un papel importante en el cumplimiento de las ambiciones de reducir emisiones de carbono establecidas por muchos gobiernos internacionales. El Acuerdo de Paris (2015) destacó la urgencia de abordar el cambio climático a escala internacional.

### **1.1.2. Gestión Energética Local**

La GEL, en general, obtuvo gran parte de la atención formal seguida por la crisis del petróleo de la década de 1970. Los enfoques de gestión sólo enfatizaron entonces en la identificación de opciones de suministro eficientes. Esto siguió al desarrollo de enfoques de gestión muy reconocidos en los que es un proceso para explorar la solución de costo mínimo que

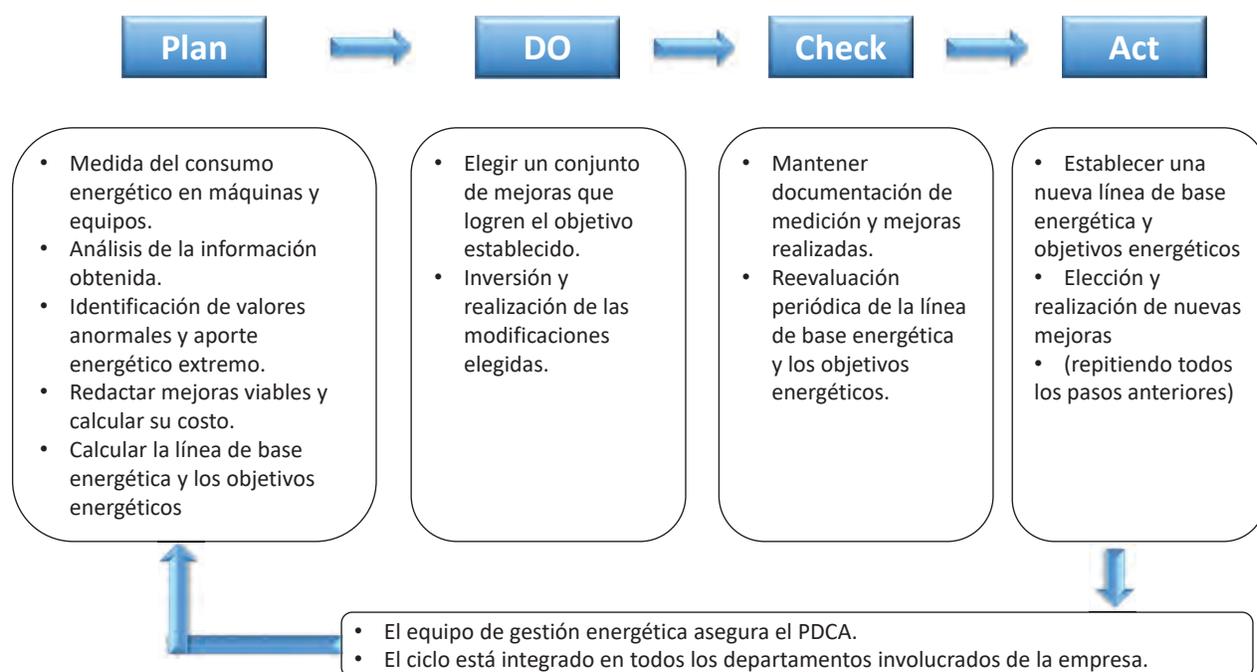


Figura 1.1: Modelo de sistema de gestión de la energía ISO 50001. Fuente: (ISO, 2018)

satisfaga las demandas de energía presentes y futuras. Otros criterios, como las preocupaciones medioambientales y la fiabilidad del suministro, reciben valores monetarios y se incluyen en los criterios de costos, por lo que se consideran únicamente como limitaciones (Mirjat *et al.*, 2017).

El incremento gradual de la demanda de energía y sus precios, especialmente en los países en desarrollo, plantea la necesidad de seguir mejorando tanto del lado de la generación como de su consumo. En el lado de la generación, la explotación del potencial de energía renovable puede considerarse un método factible y comercialmente aplicable, mientras que varias soluciones, como el uso de sistemas de almacenamiento conectados a la red, un uso más eficiente de la energía en los inmuebles y la implementación de programas de gestión de la demanda, generalmente se adoptan en el lado del consumo. En este sentido, incorporar todas estas tecnologías en un entorno común tiene importantes beneficios desde el punto de vista económico y técnico tanto para los consumidores como para las empresas. Las casas inteligentes con unidades de generación de energía residencial, que se han generalizado en la última década con la aparición de conceptos de redes inteligentes, pueden señalarse como uno de los mejores ejemplos de sistemas tan avanzados (Elma, Tascikaraoğlu, Ince, & Selamoğulları, 2017).

En la red eléctrica convencional, la electricidad generada por los generadores se transfieren a los consumidores en un sistema centralizado de generación de energía a través de sistemas de transmisión y distribución. En el sistema de generación centralizado, la energía es generada por

unas pocas unidades de generación a gran escala y se transfiere a una gama de consumidores domésticos, comerciales e industriales. Al aumentar el nivel de integración de los DER en el lado de los consumidores, se está formando una generación de energía distribuida, donde muchas unidades de generación a pequeña escala están conectadas a la red de distribución con capacidades que van desde unos pocos kilovatios hasta varios megavatios, lo que resulta en flujos de energía bidireccional. El creciente nivel de los DER ha cambiado el sistema de energía de una estructura centralizada a un modelo desregulado (Khorasany, Mishra, & Ledwich, 2018; Mengelkamp, Schönland, Huber, & Weinhardt, 2019; Sheikhahmadi *et al.*, 2020).

El despliegue masivo de DER plantea desafíos operativos y de mercado algunos de estos incluyen, entre otros, la falta de visibilidad de DER, problemas de congestión y voltaje. La gestión de energética local fomenta el manejo localizado de la distribución a partir de las demandas específicas de la localidad. Además, diferentes estudios muestran que la disposición a aceptar los recursos energéticos locales y compartir la inversión para el uso de DER se ha incrementado entre los consumidores en los últimos años. Como resultado, el diseño del nuevo mercado energético para el comercio de energía local en sistemas de generación distribuida se ha vuelto significativamente importante (Mirjat *et al.*, 2017; Celik, Roche, Suryanarayanan, Bouquain, & Miraoui, 2017; Yang *et al.*, 2017; Khorasany *et al.*, 2018; Mengelkamp *et al.*, 2019; Sheikhahmadi *et al.*, 2020).

En la actualidad, la GEL en los países más desarrollados incluye el uso de herramientas en línea, la planificación futura a corto, mediano y largo plazo mediante la modelación y los estudios de escenarios, la implantación de ideas innovadoras y su socialización (Correa Soto *et al.*, 2017, 2018).

Hay varios estudios sobre gestión energética que se encuentran en la literatura que han sido realizados tanto por gobiernos como por académicos a nivel mundial utilizando herramientas de modelado energético para respaldar la formulación de políticas energéticas. Estudios realizados en varios países, como China, Nigeria, Sudáfrica, Taiwán, Emiratos Árabes Unidos, España, Francia, Estados Unidos, Creta, Portugal, Irán, Panamá, Mongolia, Argentina, Turquía, Grecia, México, India, Indonesia, Brasil, Siria, Malasia, Tanzania, Irlanda, Letonia, Reino Unido, República de Macedonia, Pakistan, Suecia y Alemania, muestran la importancia de las solicitudes de herramientas de modelación para analizar las opciones de política energética a diferentes niveles (Marzband, Sumper, Domínguez-García, & Gumara-Ferret, 2013; Mirjat *et al.*, 2017; Celik *et al.*, 2017; Yang *et al.*, 2017; Mengelkamp *et al.*, 2019; Sheikhahmadi *et al.*, 2020).

Muchos de estos enfoques de GEL ahora se están trabajando bien utilizando varias herramientas de modelado de energía basadas en computadora. Asignación de mercado (MARKAL), Programa de evaluación de energía y potencia (ENPEP), Modelo para sistemas de suministro de energía y su impacto ambiental general (MENSAJE), Sistema de planificación de alternativas energéticas de largo alcance (LEAP) y EnergyPLAN son herramientas de modelado energético bien reconocidas utilizadas para la planificación energética en varios niveles. Estos modelos ayudan a organizar una gran cantidad de datos, reflejan el sistema energético completo de forma comprensible y proporcionan un marco coherente para probar hipótesis. Estas herramientas de modelado energético son útiles para diversas aplicaciones de gestión energética que son diversas en términos de las regiones que analizan, las tecnologías que considerar y objetivo que cumplen (Mirjat *et al.*, 2017; Biswas & Emadi, 2019; Nguyen *et al.*, 2020; Nieta *et al.*, 2020).

Con la creciente crisis energética mundial, las energías renovables y la tecnología de la información se han mejorado e integrado constantemente, lo que ha promovido el rápido desarrollo de la industria energética. Sin embargo, las fuentes de energía renovable (RES), como la energía eólica y la energía solar, con altos niveles de incertidumbre e intermitencia no pueden cumplir con la flexibilidad requerida para la programación de energía. Por otro lado, el acceso escalable a los sistemas de almacenamiento de energía (ESS) y los vehículos eléctricos enchufables (PEV) hace que sea más difícil controlar la energía en el sistema eléctrico. Sobre esta base, Jeremy Rifkin propuso el concepto de Internet de la energía (IE) que se basa en los conceptos e ideas de Internet. Como tendencia principal en la transición de la infraestructura energética existente, la IE desempeña un papel crucial en la promoción del desarrollo y la acomodación de todo tipo de opciones de generación y almacenamiento, especialmente, en permitir la participación activa del lado de la demanda del equilibrio generación/consumo. (Yang *et al.*, 2017)

### **1.1.3. Gestión Energética Local en Cuba**

Según plantea Correa Soto *et al.* (2017) en el estudio documental sobre la gestión energética local en Cuba se han identificado tres trabajos: la definición de indicadores sectoriales energéticos para el municipio Cienfuegos (Monteagudo *et al.*, 2013), la experiencia piloto en la utilización de las fuentes renovables de energía en el municipio de San José de las Lajas, Provincia de Mayabeque (Rojas, 2014), por otra parte el Centro de Gestión de la Información y Desarrollo de la Energía (CUBAENERGÍA) a través de la Red Nacional de Gestión del Conocimiento en Energía (REDENERG) logra la interrelación de actores vinculados directa o indirectamente al sector energético, con el propósito de acompañar y asesorar a los decisores en el proceso

de identificación de los problemas energéticos y en la determinación de soluciones (González et al., 2006), con la creación del Nodo Municipal de Energía (NOME) (González et al., 2013; Arencibia, 2014). Los estudios realizados por la REDENERG y CUBAENERGÍA se han basado en acciones puntuales de eficiencia energética y en el aprovechamiento de la informatización de la sociedad y no en la gestión energética municipal, reconociéndose como problema energético que los gobiernos municipales no han desarrollado los mecanismos para la gestión energética de subordinación territorial (Correa Soto *et al.*, 2017).

## 1.2. Planificación de la energía

La planificación de la energía a menudo se realiza dentro de las organizaciones gubernamentales, pero también puede ser llevada a cabo por grandes empresas de energía, tales como, centrales eléctricas, esta puede llevarse a cabo con la colaboración de las diferentes partes interesadas procedentes de organismos gubernamentales, empresas de servicios públicos, locales, instituciones académicas y otros grupos de interés (Ferrari *et al.*, 2020). Los enfoques integrados que tengan en cuenta tanto la provisión de los suministros de energía y el papel de la eficiencia energética en la reducción de las demandas son a menudo utilizados en la planeación de la energía. En esta debe reflejar siempre los resultados de crecimiento de la población, además tradicionalmente ha desempeñado un papel importante en establecer el marco de las regulaciones en el sector de la energía. Una nueva tendencia se conoce como Planificación de la Energía Sostenible, que tiene un enfoque más integral al problema de la planificación de las necesidades futuras de energía (Ferrari *et al.*, 2020).

La planificación de una ciudad sostenible conlleva cambios en las formas de la planificación tradicional. El proceso estándar de planificación urbana no prioriza la planificación energética, ya que, generalmente las características espaciales de un área son planificadas antes de su explotación mientras que la energética se lleva a cabo una vez ocurrida la anterior. La introducción del desempeño energético en la planificación de una localidad conducen a considerar varias variables, que incluyen transporte, uso de energía y suministro, clima, contaminación, ruido y emisiones, tratamiento de residuos comunes, entre otras (Resch & Andresen, 2017). Las tendencias generales en la evolución de los enfoques de planificación energética a nivel de ciudad son los siguientes:

- Creciente conciencia de la comunidad sobre los problemas ambientales.
- Creciente interés en el uso de tecnologías de generación distribuida basado en recursos

renovables y pequeños sistemas de cogeneración.

- Mayor número de decisores con diferentes intereses y preferencias que participan en el proceso de planificación.
- Promover el análisis intersectorial entre diferentes sectores, como industria, hogares y transporte.

En este sentido Mirakyan y De Guio (2013) plantea que “. . . la planificación energética integrada regional (subnacional) es un enfoque para encontrar soluciones respetuosas con el medio ambiente, institucionalmente sólidas, socialmente aceptables y rentables de la mejor combinación de opciones de oferta y demanda de energía para un área definida para apoyar el desarrollo regional sostenible a largo plazo . Es un proceso de planificación transparente y participativo, siendo una oportunidad para que los planificadores presenten problemas complejos e inciertos de manera estructurada, integrada y transparente, para que las partes interesadas revisen, comprendan y apoyen las decisiones de planificación . . . ”(Mirakyan & De Guio, 2013).

La planificación energética integrada requiere un intercambio de información entre los diferentes sistemas. Se requiere experiencia, conocimiento y la pericia de varias prácticas para lograr una mejor estrategia y transparencia. Estudios recientes subrayan la necesidad de un modelo integrador de una localidad sostenible, ya que, los aspectos energéticos y de planificación en la actualidad se abordan por separado. Por lo tanto, la creación de un modelo integral mediante la utilización de enfoques innovadores para abordar las preocupaciones energéticas mediante el uso de herramientas digitales basadas en sensores y Big Data que permita a los decisores trazar y planificar las estrategias y políticas energéticas es algo sumamente deseado. Sin embargo, esto implica que estos tengan objetivos y prioridades claramente definidos en el proceso de planificación urbana (Lindkvist *et al.*, 2019).

Al principio de las etapas de planificación, es importante lograr un compromiso entre las diferentes partes involucradas en los escenarios propuestos elaborados a través de técnicas de “retrospectivas” y de la participación comunitaria. Esto facilitan que las partes interesadas comprendan la naturaleza compartida de sus problemas y los aborden mediante soluciones prácticas. La intención es cambiar el proceso de planificación a través de la deliberación, que “da forma a los entendimientos, dando sentido a las acciones potenciales que a su vez motivan a los actores” (Innes & Booher, 2015). Entre estos se deben incluir los decisores, planificadores, inversores y promotores inmobiliarios, así como, ciudadanos y otros usuarios que tienen influencia directa e indirecta en el proceso de planificación.

Transcurre un largo periodo de tiempo desde la planificación energética integrada hasta su

implementación en las localidades esto significa que hay varias iteraciones en los planes como se reflejaba en la figura 1.1. Dentro de este periodo pueden existir fluctuaciones de los actores involucrados en el proyecto inicial dando paso a nuevos integrantes, es por esto que las ideas entendidas en una fase pueden adquirir otro significado en una nueva fase del mismo (Lindkvist *et al.*, 2019).

La planificación de localidades sostenibles recurre a múltiples actores que van desde contratistas, arquitectos, gobiernos municipales y otros que no han coordinado estructuralmente sus actividades y que tienen diversos intereses, metas y motivaciones dificultando la colaboración entre estos.

Para Lindkvist *et al.* (2019) los múltiples criterios que permiten vincular los diversos sistemas (por ejemplo, infraestructura y energía) están mal definidos, pero son aspectos más amplios cuando el uso de energía necesita consideración en una escala de vecindario. Tradicionalmente, los procesos participativos se centran en los aspectos sociales del desarrollo urbano, como la habitabilidad, la estética y la seguridad. El desarrollo de vecindarios energéticamente eficientes también debe abordar las necesidades y hábitos de los ciudadanos en términos de uso de energía. Los talleres de diseño son una herramienta participativa para facilitar la transferencia de conocimientos, pero puede ser performativo.

Los procesos participativos pueden ser intermediarios si se diseñan de manera que permitan a los residentes locales brindar información que sea directamente relevante para la eficiencia energética y la generación como un enfoque de la planificación del vecindario.

Los actores responsables en el proceso de planificación, que incluyen diferentes profesionales (por ejemplo, arquitectura y urbanismo), generan conocimiento temprano en el proceso, pero hay profesiones muy diferentes (por ejemplo, ingenieros de sistemas de energía y empresarios) que forman la base de conocimiento posterior para las decisiones de planificación. Estos dos grupos provienen de diferentes bases de conocimiento y hablan diferentes idiomas. Hay una falta de continuación de la mediación del conocimiento, ya que las diferentes bases de conocimiento pasan a primer plano en diferentes momentos del proceso de planificación, lo que hace que el conocimiento clave de los participantes ya no parezca relevante a medida que los nuevos participantes del proyecto ganan prominencia a lo largo del tiempo (Lindkvist *et al.*, 2019).

La planificación de los sistemas de energía es el dominio de los proveedores de energía y su participación es imprescindible para comprender los aspectos energéticos del área. Sin embargo, existe una renuencia a comprometerse entre sí. Los planificadores urbanos perciben que la participación temprana complican el proceso de planificación, mientras que los actores externos no ven una razón de por qué deberían participar al principio del proceso de planificación

(Lindkvist *et al.*, 2019).

Según Mirakyan y De Guio (2013) el proceso de planificación energética puede ser dividida en las cuatro fases siguientes:

- Preparación y orientación.
- Diseño de modelos y análisis detallado.
- Priorización y decisión
- Implementación y monitoreo.

Cada fase consta de varios subpasos interconectados debido a esto no se realizan necesariamente en una secuencia predeterminada. El procedimiento refleja los pasos de trabajo y el proceso de toma de decisiones. La selección de pasos puede variar de una situación a otra, dependiendo de la relación problema-objetivo, la cantidad de tiempo disponible, los recursos y los intereses de las partes interesadas.

Hay muchas bases de conocimientos en la fase de planificación inicial, pero poca voluntad de participar en un proceso intermedio que resultaría en la transferencia de conocimientos y la creación de consenso debido a la incertidumbre sobre cómo se realizarán estas visiones en el futuro. Los municipios no tienen el poder para garantizar las ambiciones energéticas establecidas en la fase de planificación. Los planificadores se muestran escépticos sobre los proyectos que promueven el desarrollo de la eficiencia energética, ya que no pueden crear vínculos legales para los objetivos energéticos, lo que puede llevar a que las partes no cumplan lo acordado (Lindkvist *et al.*, 2019).

Los planes se basan en la trayectoria que han tenido los consumos y las ofertas energéticas, apoyándose en la situación actual y en las perspectivas, para proyectarse hacia el futuro. Esto significa que deben estar incluidos dentro de uno o varios planes económicos a medio plazo, permitiendo mantener unas líneas de actuaciones claras y concretas, sobre las inversiones que es necesario realizar.

Según Barrachina Gomez *et al.* (1993) un plan energético debe considerar, durante su elaboración, cuáles son las directrices de política energética que se plantean en ese momento, este punto hace que la planificación energética esté directamente influida por las variaciones en el gobierno de la zona donde se quiera aplicar. Por tanto, está sujeta a continuas revisiones que se suman a aquellas que son necesarias en toda planificación. En el plan energético los factores más relevantes son el fomento del ahorro y eficiencia energéticos, la mejora de las

condiciones medioambientales y el cambio de tendencia hacia el aumento gradual de la tasa de autoabastecimiento mediante el uso de energías renovables.

Una buena previsión de la demanda de energía debe apoyarse en un modelo econométrico general que sintetice la evolución, a medio o largo plazo de la región considerada. Si no es posible la formulación de un modelo económico, o si las incertidumbres de evolución son muy grandes, la previsión de la demanda se planteará en términos más simples. En este caso se utilizarán hipótesis del crecimiento demográfico, estructura de población y nivel de ingreso (García Benedicto, 2004). Uno de los conceptos básicos que más importante en este momento es la denominada “intensidad energética”, este concepto representa la participación de la energía en la producción de un bien. La previsión de la demanda de energía debe matizarse precisando los suministros de carácter insustituible, de hecho, esta partida puede suponer un límite importante respecto de las actuaciones que se pretendan realizar.

El objeto final de la planificación energética es establecer los balances de energía en cada una de las situaciones futuras a que haga referencia el plan. Éstos se presentan en forma simplificada como el equilibrio entre la demanda prevista y la oferta obtenida a través de los distintos medios de abastecimiento elegidos en la planificación. Otro punto a considerar es la evolución del consumo de energía por sectores económicos, ello permitirá detectar cómo influyen en la demanda de energía primaria los cambios en los distintos sectores y actuar en consecuencia.

### **1.3. Técnicas de pronóstico de demanda eléctrica.**

Los pronósticos de demanda eléctrica juegan un papel crucial en los sistemas de producción de energía eléctrica en todo el mundo (Hong & Fan, 2016; Ravadanegh *et al.*, 2016). El pronóstico de demanda guía la planificación de la producción, comercio o distribución, operacionabilidad, etc. Sin embargo, la electricidad revela una estructura especial, que no puede ser almacenada, y debe proporcionar un equilibrio entre la producción y demanda (Khuntia *et al.*, 2016).

Al mismo tiempo, la demanda de electricidad depende del clima (temperatura, velocidad del viento, humedad, etc.), la distribución de la población (viviendas, industria, etc.) y estilos de vida vistos en intensidad de las actividades comerciales y cotidianas (horas pico frente a horas no pico, días laborables frente a fines de semana, feriados, vacaciones, etc.) (Black & Henson, 2013). Estas características únicas y específicas conducen a cambios de la de demanda eléctrica requiriéndose adaptación del suministro eléctrico, como se puede apreciar en la figura 1.2.

Por otro lado, estas características han conducido a una nueva área de investigación para

el desarrollo de técnicas de pronósticos más precisos y estables. Al mismo tiempo, un buen pronóstico de resultados ha permitido avanzar en las siguientes materias: variabilidad climática (calentamiento global), inclusión de vehículos eléctricos, generación eólica y solar, la eficiencia energética y respuesta a la demanda eléctrica. En general, no suele haber una solución exacta para la ecuación anterior debido a la variabilidad inherente a datos del mundo real, a la infinidad de factores que pueden influir en cada dato y a la incertidumbre de muchas mediciones (Berriel, Lopes, Rodrigues, Varejao, & Oliveira-Santos, 2017; Keshtegar, Mert, & Kisi, 2018).



Figura 1.2: Factores que influyen en el pronóstico de la demanda de energía eléctrica. Fuente: (Ortiz Alvarado, 2015)

Las técnicas de pronóstico se pueden clasificar en dos grupos principales: estadísticas y de inteligencia artificial aunque los límites entre estos grupos son bastante ambiguos como resultado de la colaboración multidisciplinaria de la comunidad científica (Weron, 2014; Khuntia *et al.*, 2016; Liu *et al.*, 2015). La literatura comúnmente revisada hace referencia a cuatro técnicas estadísticas, a saber, regresión lineal múltiple (MLR), modelos aditivos semiparamétricos, autorregresivos y de media móvil (ARIMA) y modelos de suavizado exponencial; y cuatro de inteligencia artificial (IA), a saber, ANN, modelos de regresión difusa, máquina de soporte vectorial (SVM) y, además, máquina de impulso de gradiente (Hong & Fan, 2016; Bezerra *et al.*, 2016; Xie *et al.*, 2015).

Regresión splines adaptativa multivariante (MARS) es una técnica no paramétrica y no lineal

de aprendizaje estadístico que se utiliza en modelación, regresión, identificación, predicción, pronóstico, etc. (Friedman, 1991). La metodología de red neuronal artificial (ANN) también es una técnica no lineal y de aprendizaje no paramétrico que se utiliza en estas áreas del saber. La regresión lineal (LR) es la forma más rápida de estimación por mínimos cuadrados en clasificación que tiene propiedades similares con ANN y MARS. ANN, MARS y LR proporcionan poderosos y muy exitosos métodos para construcción de pronósticos en grupos relacionados (Goude, Nedellec, & Kong, 2013). El pronóstico de demanda eléctrica puede ser clasificado según el período de tiempo abordado.

Aún no se ha determinado un estándar preciso para clasificar el rango de pronósticos de carga. Los procesos de pronóstico se pueden clasificar en cuatro categorías: pronóstico de carga a muy corto plazo (VSTLF), pronóstico de carga a corto plazo (STLF) (Saez-Gallego & Morales, 2017), pronóstico de carga a mediano plazo (MTLF) y pronóstico de carga a largo plazo (LTLF) (Hong & Fan, 2016; Berriel *et al.*, 2017). En esta clasificación, VSTLF aborda un período de hasta 1 día, STLF es un período que incluye 1 día-2 semanas, MTLF aborda un período de 2 semanas a 1 año, y LTLF se refiere a un período superior a 1 año. Según una clasificación aproximada, hay períodos STLF hasta 2 semanas, y período LTLF después de 2 semanas (Wang, Zhang, & Chen, 2016). LTLF es un tema importante en la planificación eficaz y eficiente del sistema de producción eléctrica (Khuntia *et al.*, 2016; Xie *et al.*, 2015).

La estimación sensible puede afectar en gran medida la hoja de ruta de las inversiones en sistemas eléctricos. La sobrestimación del futuro puede conducir a una pérdida económica en construir nuevas unidades de generación de energía para suministrar la demanda prevista. Subestimar la demanda futura puede causar problemas en el suministro (Hong & Fan, 2016). Por lo tanto, se necesita un método preciso para pronosticar cargas, ya que conduce a un modelo preciso que tiene en cuenta los factores que afectan el crecimiento de la carga durante varios años.

LTLF depende de varios factores como los hábitos humanos y las influencias ambientales. Estos factores se pueden clasificar en períodos de tiempo: horas del día (día / noche), día de la semana (día de la semana / fin de semana), época del año (temporada) y festivos: además pueden ser condiciones climáticas (temperatura, humedad y viento), tipos de clientes y distribución de la población, indicadores económicos y precio de la electricidad (Xiao, Shao, Liang, & Wang, 2016; Hong & Fan, 2016; Black & Henson, 2013; De Giorgi, Congedo, & Malvoni, 2014). Los parámetros meteorológicos medidos y los datos de carga son los parámetros más efectivos en términos de precisión de los métodos de pronóstico basados en datos históricos (Khuntia *et al.*, 2016).

### 1.3.1. Regresión Lineal

El análisis de regresión es un proceso estadístico para estimar la relación entre variables (Gujarati & Porter, 2013). Los modelos de regresión lineal múltiple han sido utilizados en tanto para STLF como LTLF. Consiste en encontrar la relación de una variable dependiente (de salida o endógena) con un conjunto de variables independientes (de entrada o exógenas). Formalmente, dado un conjunto de datos  $(x_i, y_i)$ , para  $i = 1, \dots, n$ , donde  $x_i \in \mathbb{R}$  e  $y_i$  es el valor de salida correspondiente al vector  $x_i$ , y dada una función  $f(x, A)$ , se quiere determinar el vector de parámetros  $A$ , tal que:

$$y_i = f(x_i, A), \text{ para } i = 1, \dots, n \quad (1.1)$$

La relación (1.1) se relaja definiendo una función de pérdida,  $L$ , que mide cómo los errores de predicción entre  $y_i$  y  $f(x_i, A)$  debieran penalizarse, con la idea de encontrar una solución lo más aproximadamente posible al cumplimiento de la igualdad. La elección habitual de función de pérdida es la norma  $L_p$ :

$$L_p(y - f(x, A), x) = |y - f(x, A)|^p \quad (1.2)$$

para algún número positivo  $p$ . La función de pérdida más usada es el *ajuste de mínimos cuadrados*, donde  $p = 2$ , que presenta ventajas analíticas, ya que  $L_1$  tiene el inconveniente de presentar discontinuidades en sus derivadas.

Es frecuente que  $f$  sea una función lineal. En caso contrario, por lo general, se puede linealizar el modelo mediante transformaciones de las variables.

El modelo clásico de regresión lineal, asume que se dispone de  $n$  observaciones independientes para las variables  $y, x_1, x_2, \dots, x_k$ . Se asume el modelo siguiente:

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \epsilon_i \quad (1.3)$$

donde los valores  $\beta_i$  son parámetros no conocidos;  $\epsilon$  término del error, y se requiere el cumplimiento de los siguientes supuestos: (Gujarati & Porter, 2013)

**Supuesto 1:** El modelo de regresión es lineal en los parámetros.

**Supuesto 2:** Los valores de las regresoras, las variables independientes ( $X$ ), son fijos, o los valores de  $X$  son independientes del término de error. Aquí, esto significa que se requiere covarianza cero entre  $\epsilon_i$  y cada variable  $X$ .

**Supuesto 3:** Para  $X$  dadas, el valor medio del término  $\epsilon_i$  es cero.

**Supuesto 4:** Para  $X$  dadas, la varianza de  $\epsilon_i$  es constante u homoscedástica.

**Supuesto 5:** Para  $X$  dadas, no hay autocorrelación, o correlación serial, entre las perturbaciones.

**Supuesto 6:** El número de observaciones  $n$  debe ser mayor que el número de parámetros por estimar.

**Supuesto 7:** Debe haber variación suficiente entre los valores de las variables  $X$ .

**Supuesto 8:** No hay colinealidad exacta entre las variables  $X$ .

**Supuesto 9:** El modelo está correctamente especificado, por lo que no hay sesgo de especificación.

**Supuesto 10:** El término estocástico (de perturbación)  $\epsilon_i$  está normalmente distribuido.

### 1.3.2. Regresión Splines adaptativa multivariada (MARS)

En la literatura, la regresión, es ampliamente usada para la predicción y el pronóstico, se basa principalmente en los métodos de estimación de mínimos cuadrados y máxima verosimilitud. Existen varios enfoques de regresión: modelos de regresión lineal, modelos no lineales, modelos lineales generalizados, modelos de regresión no paramétricos, modelos aditivos y modelos aditivos generalizados. MARS, un procedimiento de regresión adaptativo y no paramétrico propuesto por Jerome Friedman, se emplea particularmente para estimar funciones generales de argumentos multidimensionales (Friedman, 1991). Al mismo tiempo, MARS se puede definir como una generalización de la regresión lineal por el método paso a pasos o una modificación del algoritmo de árboles de regresión y clasificación (CART). No hay una suposición específica sobre la relación funcional subyacente entre las variables dependientes e independientes. Estima las contribuciones de las funciones base tanto para los efectos aditivos como interacciones de los predictores para determinar la variable de respuesta (Kuter, Weber, Özmen, & Akyürek, 2014; Özmen & Weber, 2014; Özmen, 2016; Vanegas & Vásquez, 2017; Cevik *et al.*, 2017; Kuter, Akyurek, & Weber, 2018). Construye e incluye expansiones en términos de funciones bases lineales por intervalos:

$$c^+(x, \tau) = [x - \tau]_+, c^-(x, \tau) = [-x + \tau]_+ \quad (1.4)$$

donde  $x, \tau \in \mathbb{R}$ . Estas dos funciones, como se muestra en la figura 1.3, se pueden nombrar

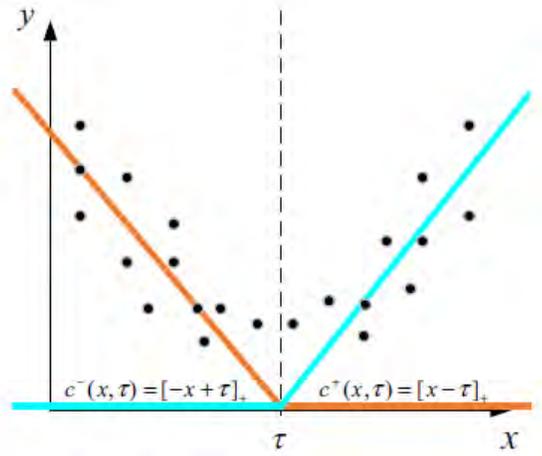


Figura 1.3: Detalles de funciones bases unidimensionales (basado en Friedman, 1991)

como un par reflejado. En el par, el símbolo (+) especifica solo las partes positivas utilizadas, y de lo contrario es cero. No es necesario centrar ni escalar, pero se sugieren. Los modelos MARS son resistentes a predictores altamente correlacionados de varianza cercana a cero. Pero esto puede conducir a una cantidad significativa de aleatoriedad durante el proceso de selección del predictor. La opción de dividir entre dos predictores altamente correlacionados se convierte en una elección al azar. Se considera la forma general del modelo, incluyendo variables y vectores aleatorios:

$$Y = f(X) + \epsilon, X = (X_1, X_2, \dots, X_p)^T \quad (1.5)$$

La meta es construir un conjunto de pares reflejados para cada variable de entrada  $x_j$  ( $j = 1, 2, \dots, p$ ):

$$\varphi := [x_j - \tau] + [\tau - x_j] + |\tau \in x_{1,j}, x_{2,j}, x_{3,j}, \dots, x_{N,j}, j \in 1, 2, \dots, p \quad (1.6)$$

Y puede representarse a partir de 1.5 como sigue:

$$Y = \theta_0 + \sum_{m=1}^M \theta_m T_m(X) + \epsilon \quad (1.7)$$

donde  $T_m$  son funciones base de  $\varphi$  o productos de dos o más de tales funciones. Las funciones de base de interacción se crean multiplicando una función base existente con una función lineal truncada que involucra una nueva variable. 0 y  $m$  son los coeficientes estimados minimizando la suma de cuadrados residual. Además,  $s$  puede ser un hilo para la selección del signo  $\pm 1$ . La  $v(j, m)$  etiqueta las variables predictoras y  $\tau_{jm}$  representa valores de los nudos correspondientes (Friedman, 1991). Según estos datos, las funciones de base multidimensionales tienen el siguiente

aspecto:

$$T_m(X) = \prod_{j=1}^{K_m} [S_{jm} \cdot (X_{v(j,m)} - \tau_{jm})] \quad (1.8)$$

El algoritmo MARS es la unión de dos subprocedimientos, denominados algoritmo Progresivo y

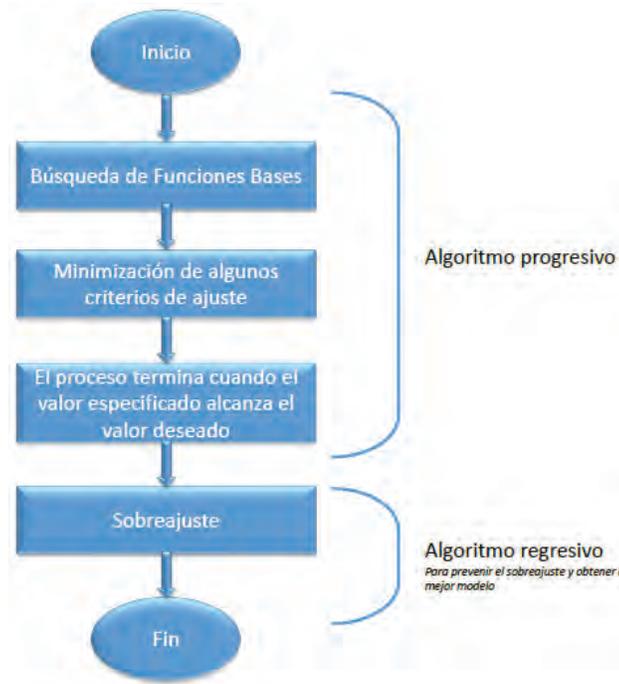


Figura 1.4: Diagrama de Flujo del algoritmo MARS. Fuente: Elaboración propia

algoritmo Regresivo, representados en la figura 1.4. Como se muestra en dicha figura, el algoritmo progresivo produce típicamente un ajuste excesivo de los datos: por tanto, posteriormente se aplica un procedimiento de borrado hacia atrás. El procedimiento de eliminación o algoritmo regresivo evita el sobreajuste al disminuir la complejidad del modelo sin degradar el ajuste a los datos. El procedimiento evalúa funciones base y resta al modelo tales BF que contribuyen al menor incremento en el error cuadrado residual en cada etapa, produciendo un modelo  $f_\mu$  óptimamente estimado con respecto a cada número de alguna complejidad en términos de estimación, llamado  $\mu$ . El valor óptimo de  $\mu$  podría calcularse con validación cruzada de acuerdo con el número de muestras  $N$ , pero el algoritmo MARS utiliza validación cruzada generalizada (GCV) para disminuir la carga computacional. GCV se puede definir de la siguiente manera, y también es conocido como falta de ajuste (LOF):

$$LOF(f_u) = GCV(\mu) = \sum_{i=1}^N \frac{(Y_i - f_\mu(x_i))^2}{(1 - M(\mu)/N)^2} \quad (1.9)$$

Aquí el denominador esta relacionado con la complejidad de la estimación. EL valor óptimo de  $M(\mu)$  puede ser calculado mediante:)

$$M(\mu) = u + d \cdot K \quad (1.10)$$

El algoritmo crea un modelo que contiene las funciones básicas vitales no repetitivas. Por otro lado, reduce la carga computacional y proporciona facilidad de procesamiento datos. Al mismo tiempo, el algoritmo es muy eficaz en aplicaciones de pronósticos.

### 1.3.3. Redes neuronales artificiales

Las redes neuronales son una rama del campo conocido como inteligencia artificial que también incluye razonamiento basado en casos, sistemas expertos y algoritmos genéticos (Barrow & Crone, 2016a, 2016b; Azad, Mekhilef, & Ganapathy, 2014). Una ANN se compone de una serie de elementos de procesamiento interconectados, cambiando su respuesta de estado dinámico a entradas externas. Dan un mejor rendimiento para hacer actividades humanoides en campos como procesamiento de voz, reconocimiento de imágenes, control robótico, pronóstico, etc.

En primer lugar, cuando se habla de una red neuronal en el ámbito de aprendizaje automático, se debe usar el término “red neuronal artificial” (Artificial Neural Network). Las redes neuronales biológicas son muchas más complicadas que los modelos matemáticos utilizados para las ANN, tanto en su tamaño (número de nodos) como en las conexiones de estos (Nalcaci, Özmen, & Weber, 2019).

No existe una definición universalmente aceptada de una ANN, pero una posible seria la siguiente: una ANN es una red constituida por muchas unidades de proceso sencillas (“neuronas”), cada una de ellas posiblemente con una pequeña cantidad de memoria local. Las neuronas están conectadas por canales de comunicación (conexiones”) que por lo general llevan asociados valores numéricos (como función peso  $w$ ), codificados por cualquiera de varios medios. Estos nodos operan solamente con sus datos locales y las entradas que reciben a través de las conexiones mediante una función llamada de activación (Barrow & Crone, 2016b).

La forma más simple de una neurona que contiene capas de entrada, ocultas y de salida se muestra en la figura 1.5. El número de neuronas puede variar en cada capa. La capa de entrada configura los datos de acuerdo a la capa de aprendizaje la cual a su vez es una capa oculta (puede contener varias), esta a su vez contiene las funciones de transferencia: sigmoide, umbral, lineal por partes y gaussiana, estas juegan un papel clave en el aprendizaje.

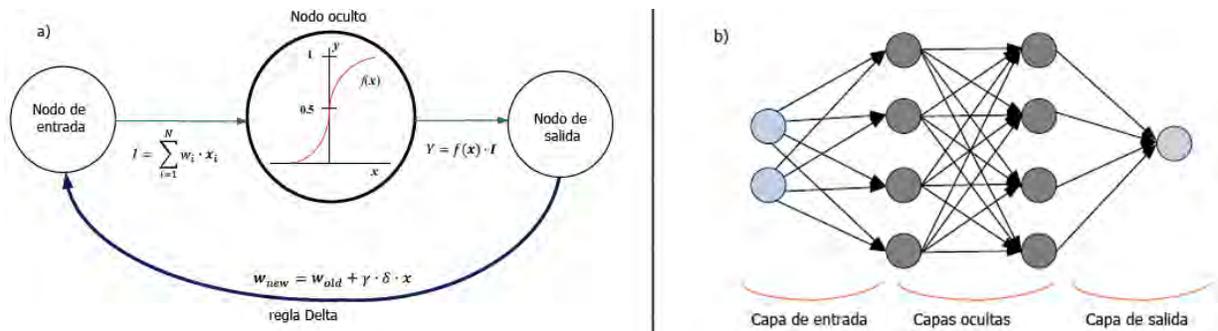


Figura 1.5: Esquema de un perceptrón. (a) simple, (b) multicapa. Fuente: Elaboración propia

Las redes neuronales tienen una alta tolerancia a los datos ruidosos unido a su capacidad de clasificar patrones sobre redes que no han sido entrenadas, lo cual es una ventaja significativa.

Finalmente, sin tener en cuenta si el tipo de problema es atacable mediante esta metodología, es necesario hacer un último inciso en un aspecto fundamental: es importante entender que no hay métodos para entrenar a las ANN para que puedan crear mágicamente información que no está contenida en conjunto de datos de entrenamiento. La ANN descubrirá el conocimiento si lo hay, previo ajuste correcto de su topología y sus hiperparámetros, pero eso es todo.

En general, las redes neuronales presentan ventajas respecto a otras técnicas porque:

- Se entrenan, auto organizan, aprenden y olvidan.
- Son robustas y tolerantes a fallas; la falla de una o varias neuronas no implica un fallo total en la red neuronal.
- Son flexibles, lo que les permite adaptarse fácilmente a nuevos ambientes, ya que pueden catalogarse como sistemas independientes.
- Se emplean en datos en los cuales el patrón es oscuro e imperceptible, que exhiben comportamiento impredecibles o no lineal, como en modelos tradicionales de series temporales y datos caóticos.
- La velocidad de respuesta es menor que la del cerebro humano.
- Son hábiles en el proceso de asociar, evaluar o reconocer patrones.

## 1.4. Herramientas computacionales de apoyo a la planificación energética.

En la actualidad, existe variedad de software que reúnen un sinnúmero de características y funcionalidades que brindan a los usuarios una mayor simplicidad y comodidad a la hora de realizar el cálculo y la estimación desde estadísticas básicas, hasta la aplicación más rigurosa de diferentes técnicas o métodos para la construcción de modelos econométricos.

Las herramientas que se encuentran disponibles en el mercado pueden ser clasificadas en dos grupos, atendiendo a su tipo de licencia: privativo o de pago; y, libre o de código abierto. Siguiendo esta línea se tiene a las herramientas computacionales más populares o conocidas: Análisis de Datos (complemento de Microsoft Excel), Minitab, SPSS, Statgraphic, SAS, Statistica, Matlab, RATS, Stata, Eviews.

De otro lado, se tiene, con licencia libre, herramientas más modestas como Tanagra, para fines académicos y de investigación, Italassi, enfocado en la interpretación de regresiones con dos variables independientes. Algunas herramientas más optimizadas, pero aún así con limitaciones son Octave, herramienta open-source con una variedad de funciones y paquetes, y JMatlab, un clon de Matlab.

Siguiendo con el análisis, las herramientas más potentes a tener en cuenta son SPSS, SAS, Matlab, Python con sus librerías y CRAN con su lenguaje R, de las cuales se valoran más detalle con el fin de tener una visión más amplia y a su vez hacer un paralelo entre cada una de ellas.

**SPSS:** proveer una amplia variedad de funcionalidades dirigido a una gran variedad de profesionales en diversos ámbitos con el fin de mejorar la eficiencia y minimizar el riesgo. Para ello, proporciona un catálogo de sus productos donde explican su contenido detalladamente (IBM.Inc, Consultado en:Feb, 2021; ?).

La suscripción base a este módulo cuesta 99\$ USD por usuario, con renovación mensual. Existen otras licencias más costosas, las cuales son la Base v25, Standard v25, Professional v25 y Premium v25. Estas suscripciones cuestan 1250\$ USD, 2770\$ USD, 5560\$ USD y 8290\$ USD respectivamente, de suscripción anual y por usuario con soporte técnico.

SPSS tiene un desarrollo potente ya que su interfaz gráfica tiene la posibilidad de mostrar diversos tipos de gráficos sin necesidad de programación, sino a través de su interfaz.

**SAS:** SAS BASE, hospedado en [https : //www.sas.com/en\\_us/software/base – sas.html](https://www.sas.com/en_us/software/base-sas.html)

donde se puede ver una descripción de sus funcionalidades, es una de las más grandes en el sector, con una gran capacidad para el análisis y con su propio lenguaje de programación, desarrollado por SAS Institute con un intérprete desarrollado por la empresa nombrada. También hay otro intérprete desarrollado por la empresa World Programming. Esta herramienta tiene un gran potencial en el tema de análisis de datos, ya que es capaz de realizar una amplia variedad de tareas estadísticas tanto básicas como avanzadas (SAS.Inc, Consultado en: Feb, 2021).

Es ampliamente utilizada en entornos empresariales, con un precio de suscripción elevado. La versión que se proporciona es la 9.4 y el precio para empezarla a usar es de 9,720\$ USD el primer año, después será entre el 25-30 % del precio dicho, aunque también existe una versión de prueba con el fin de probar sus ventajas. Esta licencia incluye SAS Base, SAS/STAT y SAS Graph. Existen más paquetes, tales como, SAS/ETS para series de tiempo, SAS/OR para optimización o SAS/IML para manipulación de matrices, por ejemplo. Cada una de ellas tiene un costo añadido de 3000\$ USD. Todo esto es un precio aproximado de inicio ya que SAS vende su producto dependiendo de los *cores o sockets* de los servidores que se vayan a usar. También existe una versión virtualizada de su herramienta orientada al entorno universitario, “University Edition”.

**MATLAB:** Abreviatura de MATrix LABoratory) orientado al campo de las ingenierías, contiene su propio lenguaje de programación, lenguaje M, involucrada en temas no solo de análisis de datos sino también en otros campos como la visión artificial, Deep Learning, Procesamiento de señales, robótica o sistemas de control (MathWorks.Inc, Consultado en: Feb, 2021).

MATLAB ofrece varios tipos de licencias. En primer lugar, está la licencia estándar, con un precio de 2000€ la licencia perpetua y 800€ la anual; la cual está orientada a un usuario final que operará, instalará y administrará el software.

Para el análisis de datos Matlab cuenta con un grupo de paquetes (Toolbox) que permiten extraer mayor partido a la herramienta, a decir: Global Optimization Toolbox, Parallel Computing Toolbox, Curve Fitting Toolbox, Neural Network Toolbox, Statistics and Machine Learning Toolbox, Optimization Toolbox, Database Toolbox, Text Analytics Toolbox, Symbolic Math Toolbox.

Por otra parte, Matlab también dispone de una comunidad con la que compartir problemas y conocimientos, un servicio web, Matlab online, un servicio de almacenamiento, Matlab Drive y un servicio móvil compatible con iPhone, iPad o un dispositivo android, Matlab mobile (Alvarez Basanta, 2020).

**Python:** Python es una herramienta que ha ido tomando popularidad en el mundo informático desde hace unos años, debido a su gran capacidad para tratar los datos y sus paquetes en diversos campos. En su página se puede encontrar como describen ellos en primera instancia su aplicación y su lenguaje. Este es de código abierto, bajo una licencia Open Source OSI-approved haciendo que sea libre para su distribución y uso (Python Core Team, Consultado en: Feb, 2021; Pajankar, 2021). Python provee una serie de paquetes/librerías para poder sacarle un buen partido al análisis de datos. Estas son las siguientes:

- **Pandas:** Librería diseñada para trabajar con datos para una rápida y fácil manipulación, adición y visualización de estos. Soporta dos tipos de estructuras, Series y Dataframes.
- **Statsmodels:** Librería que provee clases y funciones para la estimación de diferentes modelos estadísticos, la realización de pruebas estadísticas y exploración de datos.
- **Mlpy:** Librería que provee lo último en métodos de Machine Learning.
- **NumPy:** Librería fundamental para la computación científica, cual soporta matrices y arrays amplias y multidimensionales, además de una gama de funciones matemáticas para operaciones con arrays.
- **SciPy:** Librería usada para la computación científica y computación técnica, que provee módulos para optimización, algebra lineal, integración, interpolación, transformada de Fourier(FFT), procesamiento de señal e imagen, solución a ecuaciones diferenciales ordinarias, etc.
- **Scikit-Learn:** Librería que provee varios algoritmos de clasificación, regresión y cauterización. Este paquete ofrece la posibilidad de operar con las librerías NumPy and Scipy.
- **Theano:** Librería que permite definir, optimizar y evaluar expresiones matemáticas con matrices multidimensionales de forma eficiente.

Esta herramienta también tiene librerías desarrolladas para la visualización son:

- **Matplotlib:** Librería diseñada para la generación tanto de simples como de potentes gráficas con facilidad. Herramienta de bajo nivel en la que es necesario el uso de código para alcanzar niveles avanzados de visualizaciones, pero tiene una gran capacidad para realizar gráficos personalizados con cierto esfuerzo.
- **Seaborn:** Esta librería se centra en la visualización de modelos estadísticos. Sus estilos por defecto son más sofisticados que los de *matplotlib*. Esta librería de alto nivel, hace más fácil la creación de gráficas, tales como mapas de calor, series temporales o diagramas de violín.

- **Bokeh:** Librería con una gran visualización de gráficos. Su objetivo es la visualización interactiva e independiente de matplotlib como la anterior tratada. Esta muestra sus gráficas a través de un navegador usando JavaScript widgets. Una característica importante de esta librería es la capacidad que tiene mediante abstracciones de alto nivel permite crear gráficas complejas y a vista de bajo nivel dando flexibilidad a los desarrolladores.

Python dispone de una comunidad muy activa con la que compartir problemas y conocimientos. Gracias a su gran comunidad, es capaz de evolucionar y mejorar por su contribución a detectar bugs o errores, arreglarlos y documentarlos. También son estos seguidores los que desarrollan otras librerías. Como se ha comentado, su comunidad es su gran soporte haciendo de Python un lenguaje y herramienta a tener en cuenta.

**R-Project:** R, al igual que Python, es un lenguaje y entorno (sistema) con un reconocimiento en el mundo del análisis de datos, gracias a que esta fue diseñada con ese propósito, básicamente, es una herramienta estadística para estadísticos. R es de código abierto bajo las condiciones y términos de la Free Software Foundation's GNU General Public License (R Core Team, Consultado en: Feb, 2021).

Provee gran variedad de técnicas estadísticas y gráficas para la aplicación de modelos lineales o no lineales, la realización de pruebas estadísticas básicas, análisis de series de tiempo, entre muchos otros. Es altamente extensible, ofreciendo la posibilidad de incluir cuando sea necesario nuevos paquetes desarrollados por la comunidad que satisfagan la necesidad del usuario. Algunos de sus paquetes básicos son:

- |             |            |        |
|-------------|------------|--------|
| ▪ Stats     | ▪ utils    | ▪ base |
| ▪ Graphics  | ▪ datasets |        |
| ▪ grDevices | ▪ methods  |        |

Con estos paquetes se pueden realizar análisis decentes de manera eficiente. Pero se puede encontrar otros paquetes para análisis avanzados. Unos cuantos de ellos son:

- RMySQL, RPostgreSQL, RSQLite, RODBC, xlsx, XLConnect, RMongo, foreign. Estos paquetes sirven para tener acceso a diversos tipos de formatos, como Excel, .csv, SAS, SPSS, etc., o base de datos, ya sea MySQL, Mongo, etc.
- Dplyr. Paquete esencial para crear subconjuntos, resumir, reorganizar o unir conjuntos de datos, véase, datasets.
- Tidy. Paquete para cambiar el diseño de los datasets. Con esta se podrá dejar los en un formato ordenado.

- `Stringr`. Paquete con funciones de tratamiento de expresiones regulares, caracteres o Strings.
- `Data.table`. Paquete que sirve como una alternativa para organizar datasets operaciones muy rápidas, el cual es utilizado mayormente en Big Data.
- `Car`. Paquete muy útil para realizar tablas ANOVA de tipo II y III.
- `Caret`. Paquete para entrenar modelos de regresión y clasificación.
- `Mgcv`. Paquete que provee funciones para modelos aditivos generalizados.
- `Lme4`, `nlme`. Paquete para modelos de efectos mixtos lineales o no lineales.
- `Shiny`. Paquete que hace fácil construir aplicaciones web interactivas con R.
- `ggplot2`. Este paquete es el más conocido por los usuarios, por su capacidad de manipulación y representación gráfica.

Esta herramienta al igual que Python no tiene un servicio técnico en sí, pero al estar bien implementada y tener una gran comunidad se puede encontrar mucha información sobre dicha herramienta, tanto en tutoriales, libros, código e hilos de ayuda. Al igual que en Python, R ha ganado mucha fama y los desarrolladores, aparte de la propia empresa, han apostado por ella contribuyendo a su mejora. R consta de una gran variedad de paquetes con los que poder trabajar. Estos se pueden encontrar en CRAN, The Comprehensive R Archive Network (Alvarez Basanta, 2020).

En estos últimos años, la gran productividad de herramientas libres ha hecho que las de pago, hayan perdido seguidores a favor de aquellas que cuentan con las mismas capacidades y sobretodo con una mejor escalabilidad. Dando paso al uso de herramientas como Python o R, siendo actualmente líderes en el tema de análisis de datos. No obstante, Python está concebida con una visión más amplia, ya que no solo trabaja en el campo del “Data Science” (Ozgur, Colliau, Rogers, Hughes, *et al.*, 2017; Coleman, Lyon, Maliar, & Maliar, 2020; Ranjani, Sheela, & Meena, 2019; Raschka, Patterson, & Nolet, 2020; Suresh, 2021).

## 1.5. Conclusiones parciales:

- La gestión de la energía se aplica en cualquier tipo de organización incluyendo los gobiernos locales; en Cuba esta temática aún es incipiente conllevando a hacer énfasis en la planeación energética como uno de los elementos de la GEL.
- En el ciclo de la planeación energética pueden ser aplicadas numerosas técnicas y herramientas estadísticas, matemáticas y computacionales que permiten realizar el pronóstico de la futura demanda de energía eléctrica.
- Dentro de las técnicas estadístico-matemáticas destacan RLM, ANN y MARS, las últimas superan las limitaciones impuestas por los supuestos que debe cumplir RLM y consiguen mejorar la calidad del ajuste.
- Dentro de las herramientas computacionales destaca Python, un lenguaje de código abierto (libre) con abundantes librerías que ofrecen soporte para múltiples áreas del conocimiento, desde la realización de análisis matemático - estadístico hasta aplicaciones de automática e inteligencia artificial, y una comunidad de desarrollo activa.

# ANÁLISIS DE LA GESTIÓN ENERGÉTICA LOCAL DE CIENFUEGOS

---

En este capítulo expone una caracterización del municipio Cienfuegos, se exponen los CP que lo conforman, su distribución espacial, su consumo eléctrico. Se realiza en resumen de las principales investigaciones precedentes en el ámbito de la gestión energética local, sus aportes y deficiencias.

## **2.1. Caracterización del municipio de Cienfuegos.**

El municipio Cienfuegos es el municipio cabecera de la provincia de igual nombre este consta con una extensión superficial de 356,3 km<sup>2</sup>. Se encuentra situado en el centro-sur de la provincia a los 22° 7' y 30" de latitud Norte y 180 18' de longitud Oeste sobre la península de Majagua, limita al Norte con los municipios de Palmira y Rodas, por el Sur con el Mar Caribe, hacia el Este con el municipio de Cumanayagua y por el Oeste con el municipio de Abreus (ONEI Cienfuegos, 2019).

Sus características ambientales se reflejan en la tabla 2.1 y están determinadas por los indicadores de clima, por ejemplo, en el año 2016 se registró un total anual de 963,8 mm de lluvia, que abarcaron 121 días, una temperatura media anual 30,8°C para la máxima y 20,8°C para la mínima, dirección y rapidez de viento predominante 16 rumbos NE a 7,2 km/h, humedad relativa del 77 % y una nubosidad media de 3 octavos (ONEI Cienfuegos, 2019).

Tabla 2.1: Variables meteorológicas. Fuente: (ONEI Cienfuegos, 2019)

ESTACIONES Y AÑOS	Viento predominante		Humedad relativa	Nubosidad media
	Dirección en 16 rumbos <sup>(a)</sup>	Rapidez (km/h)	(%)	(En octavos)
<b>Cienfuegos</b>				
2013	NE	9,6	78	3
2014	NE	9,2	77	3
2015	ENE	10,7	79	3
2016	NE	7,2	77	3
2017	NE	6,4	77	3
2018	NE	6,4	80	3

<sup>(a)</sup>Se refiere a los puntos cardinales

Fuente: Delegación Provincial del Ministerio de Ciencia, Tecnología y Medio Ambiente

ESTACIONES Y AÑOS	Lluvia		Temperatura media	
	Total anual (mm)	Días con lluvias (U)	Máxima (°C)	Mínima (°C)
<b>Cienfuegos</b>				
2013	1 543,6	116	30,7	21,2
2014	1 067,0	104	31,1	20,9
2015	1 166,1	125	31,5	21,5
2016	963,8	121	30,8	20,8
2017	1 568,5	123	30,8	20,8
2018	1 437,6	117	30,4	20,8

La bahía de Cienfuegos se extiende en un área de 88.46 km<sup>2</sup> con un volumen de agua para el nivel medio del mar de 810 Km<sup>3</sup>. Su forma es ovalada y está orientada de NO a SE. Tiene 19 km de longitud máxima y 7.5 km en su parte más ancha, con una profundidad promedio de 9.5 m. De forma natural está dividida en dos lóbulos delimitados por el bajo “Las Cuevas” que tiene una profundidad promedio de 1.5 m. Presenta un estrecho y sinuoso canal de acceso con una longitud de 3 600 m y profundidades entre 30 y 50 m en el centro, limitado por dos puntas, al este por “Punta Los Colorados”. El territorio presenta un gran potencial natural, tanto para el desarrollo de la actividad humana: residencial, industrial, marítimo-portuaria, agropecuaria, forestal, minera, pesquera, turístico-recreativa y otros; así como, la conservación de ecosistemas de la flora y fauna, ejemplo de estos se pueden observar en el área protegida Guanaroca (Caravaca *et al.*, 2014; ONEI Cienfuegos, 2019).

Las características físico geográficas propician la vulnerabilidad del territorio ante la ocurrencia de fenómenos atmosféricos como las inundaciones por intensas lluvias, las penetraciones marinas y las afectaciones por fuertes vientos, dado por los ríos y arroyos, en el caso de la ciudad se incrementan las inundaciones a partir de los problemas de drenajes generados por la

urbanización. Las penetraciones marinas ponen en peligro a las costas bajas y acumulativas, manifestándose de manera diferente en el interior y exterior de la bahía. La exposición a los fuertes vientos se hace mayor en las áreas de llanuras al no contar con barreras naturales que las protejan frente a este peligro (Caravaca *et al.*, 2014).

El municipio cuenta con 19 Consejos Populares (CP) de ellos 11 urbanos y 8 mixtos que responden a las necesidades gubernamentales y político-administrativas siendo utilizados como base para el control territorial, a los cuales se refiere en la Tabla 2.2, se encuentran distribuidos espacialmente como se muestra en al figura 2.1.

Tabla 2.2: Consejos Populares. Fuente: Elaboración propia

<b>Municipio</b>	<b>Consejos Populares</b>
Cienfuegos	Reina, Centro Histórico, Pastorita, Junco Sur, La Juanita, Juanita II, Pueblo Griffo, Caunao, La Gloria, Tulipán, La Barrera, Buenavista, San Lázaro, Paraíso, Rancho Luna, Punta Gorda, Guaos, Pepito Tey, Castillo CEN

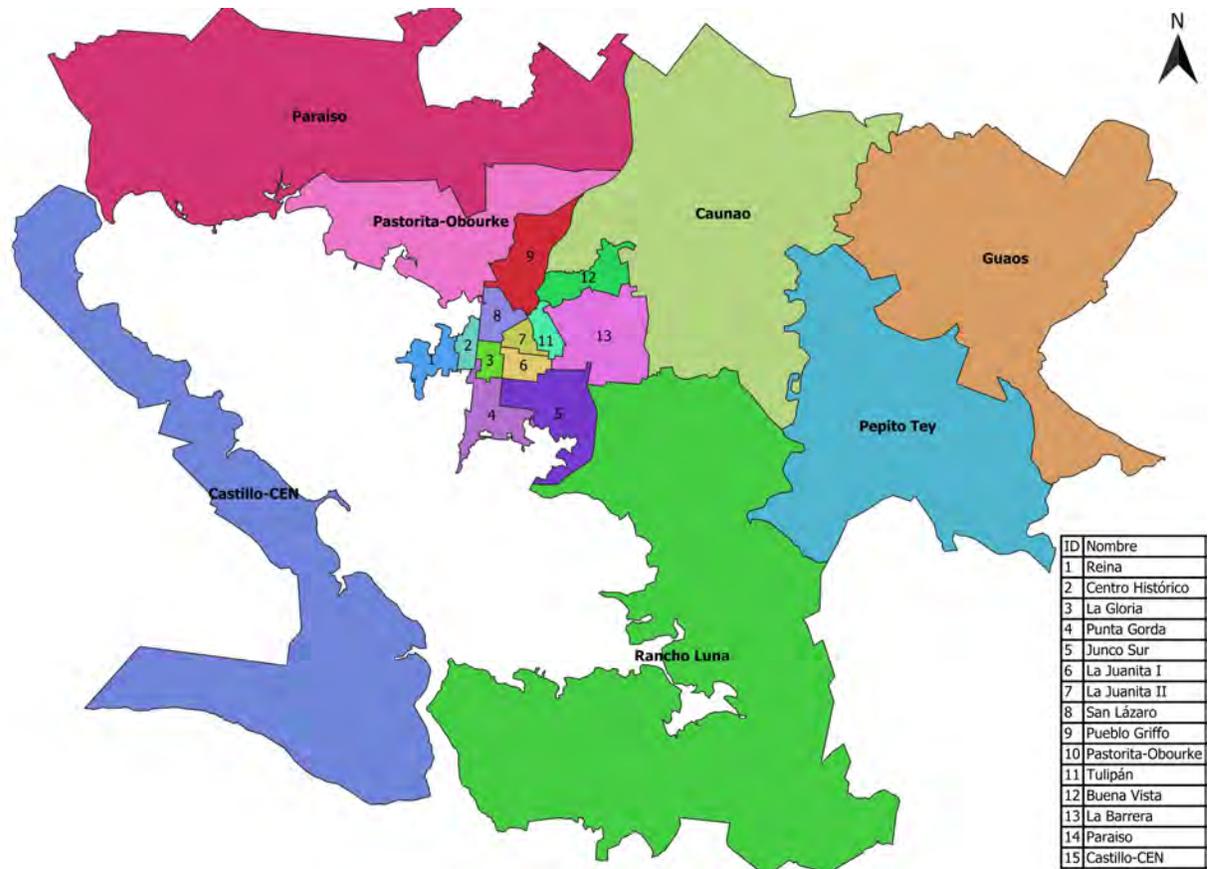


Figura 2.1: Distribución Espacial de los CP. Fuente: (IPF,2021)

El municipio tiene una población residente de 177617 habitantes y una densidad de población de 499,4 hab/km<sup>2</sup>, con 87599 mujeres y 90018 hombres, los menores de 15 años representan el 5% de la población, las edades entre 15 y 59 años el 64.92% y los mayores de 60 años son 34 086 representando el 19.19% de toda la población cienfueguera, el Índice de Rocet es de 17.5% por lo que se clasifica como una población muy envejecida y la esperanza de vida al nacer para los hombres es de 76 años y las mujeres 79.6 años. El municipio tiene una tasa anual de crecimiento de 2,8 y una relación de masculinidad 973 y un total de 56946 viviendas (ONEI Cienfuegos, 2019).

La base económica del municipio es fundamentalmente industrial y de servicios. El territorio cuenta con 3 zonas industriales y otra más pequeña en Guabairo con la Fábrica de Cemento como su principal representante, 3 zonas portuarias, una red de almacenes, talleres y pequeñas industrias dispersas dentro de la trama urbana. En la actividad agropecuaria se destacan la producción de alimentos como: cultivos varios, frutales y ganadería. Una actividad con futuro es el turismo, que cuenta con 9 hoteles, se desarrolla la actividad inmobiliaria en Punta Gorda y su ampliación en el Centro Histórico y proyecciones de desarrollo hasta el 2030, existe una

base de campismo y cabañas de recreación (Correa Soto *et al.*, 2018).

De los 115 470 habitantes del municipio en edad laboral 96 280 están empleados en el sector estatal con un salario promedio de 804 pesos. El sector estatal está conformado en el municipio por 121 organismos (58 empresas, 50 unidades presupuestas, 12 cooperativas y 1 empresas mixtas), estos organismos para el cumplimiento de su objeto social consumen energía que se desglosa en energía eléctrica, el gas, la gasolina motor, el combustible diésel, los aceites, grasas y lubricantes, petróleo crudo, petróleo combustible, donde los organismos mayores consumidores pertenecen al Ministerio de Energía y Minas, Ministerio de la Construcción y el MINAL (ONEI Cienfuegos, 2019).

Por otra parte, el sector residencial compuesto por las 56946 viviendas consume energía eléctrica, gas, donde el portador de mayor significancia es la energía eléctrica siendo el Consejo Popular Centro Histórico el de mayor consumo y Guaos el de menor. Así como la generación de energía eléctrica a través de la Termoeléctrica “Carlos Manuel de Céspedes” (ETE) y la inserción de fuentes renovables de energía (FRE), como parque fotovoltaico (Cantarrana), biodigestores, calentadores solares, paneles solares, arietes hidráulicos (Campillo Sabina, 2018).

### **2.1.1. Situación energética municipal**

La Oficina Municipal de Estadística e Información (OMEI) es la encargada de la recopilación estadística del consumo de energía eléctrica por organismos y actividades económicas, a través del Anuario Estadístico Municipal de los siguientes portadores energéticos:

- Energía eléctrica
- Diesel
- Gasolina
- Gas licuado
- Petróleo crudo y derivados
- Fuel oil
- Solventes
- Asfalto de petróleo

En esta no se registran los consumos de energía eléctrica residencial. En tal sentido, Campillo Sabina (2018) realizó un diagnóstico energético del municipio en cuestión donde se determinó el consumo residencial de los CP.

En la tabla 2.3 se pone en evidencia el crecimiento de la generación de energía eléctrica alternativa en el municipio, tomando como base el año 2016, de estas el mayor generador es el parque fotovoltaico ubicado en el consejo popular Paraíso el cual tiene un capacidad de

generación instalada de 2.6 MW (ONEI Cienfuegos, 2019).

Tabla 2.3: Dispositivos generadores de energía renovable y energía sustituida.  
Fuente: (ONEI Cienfuegos, 2019)

CONCEPTO	UM	2013	2014	2015	2016	2017	2018
<b>Dispositivos</b>	<b>U</b>	-	-	-	381	674	736
Molinos de viento		-	-	-	47	241	232
Digestores de biogás		-	-	-	178	381	456
Plantas de biogás		-	-	-	-	-	-
Sistema de paneles fotovoltaicos		-	-	-	156	52	48
Aerogeneradores		-	-	-	-	-	-
<b>Energía sustituida</b>	<b>tep</b>	-	-	-	754,8	1 350,0	663,5
Molinos de viento		-	-	-	70,5	357,0	352,5
Digestores de biogás		-	-	-	684,3	993,0	311,0
Plantas de biogás		-	-	-	-	-	-
Sistema de paneles fotovoltaicos		-	-	-	-	-	-
Aerogeneradores		-	-	-	-	-	-

Estos generadores alternativos se encuentran distribuidos espacialmente en diferentes CP del municipio, donde su mayor parte se encuentran en consejos populares mixtos como se aprecia en la tabla 2.4.

Tabla 2.4: Generación alternativa por CP. Fuente: (Rodríguez Figueredo, 2019)

Generador Alternativo	Consejo Popular	Cantidad
Solar fotovoltaica	La Gloria, Reina, Paraiso	3
Biogás	Paraiso(1), 29 viviendas en (Buena Vista, Caunao, Guaos, Pueblo Grifo, Paraíso, Pepito Tey, Punta Gorda y Tulipán)	30
Ariete Hidráulico	Pepito Tey	1

En el consumo de combustibles utilizados para la generación se observa una disminución de 69,1 toneladas de combustibles entre el 2013 y el 2017 como consecuencia del aumento de los generadores alternativos, como se muestra en la tabla 2.5.

Tabla 2.5: Consumo de portadores energéticos. Fuente: (ONEI Cienfuegos, 2019)

Primarios						
PRODUCTOS	UM	2013	2014	2015	2016	2017
Petróleo crudo	T	288,1	244,2	250,6	264,9	219,0
Secundarios						
CONCEPTO	UM	2013	2014	2015	2016	2017
Derivados del petróleo <sup>(a)</sup>		41 969,5	37 794,9	36 876,5	34 462,6	34 000,4

<sup>(a)</sup> Comprende el consumo de petróleo crudo utilizado directamente en sustitución de fuel oil

La energía producida por los diferentes medios de generación es resumida en la tabla 2.6 donde se observa un aumento de 1,8 veces entre el 2014 y el 2015, mientras que en el 2018 se aprecia una disminución del 28.31 % con respecto al año anterior.

Tabla 2.6: Generación bruta de electricidad por tipo de planta productora. Fuente: (ONEI Cienfuegos, 2019)

AÑOS	GigaWatt hora							
	Generación térmica			Grupos Electrógenos			Renovables	
	Total	Termo-eléctricas	Auto-productores	Turbinas de gas	Interconectados al sistema	Aislados	Hidro-eléctricas	Otras
2013	2 204,6	1 913,3	53,1	-	238,2	-	-	-
2014	1 883,4	1 633,2	55,9	-	194,3	-	-	-
2015	3 476,4	1 705,8	61,4	-	-	-	-	1 709,2
2016	3 481,7	1 703,2	67,8	-	-	-	-	1 710,8
2017	3 251,1	1 592,5	53,6	-	-	-	-	1 605,0
2018	2 330,6	1 130,7	49,4	-	-	-	-	1 150,5
AÑOS	Estructura del total				Tasas			
	Generación térmica	Turbinas de gas	Grupos Electrógenos	Renovables	Generación térmica	Turbinas de gas	Grupos Electrógenos	Renovables
2013	89,2	-	10,8	-	63,6	-	30,0	-
2014	89,7	-	10,3	-	-14,1	-	-18,4	-
2015	50,8	-	-	49,2	4,6	-	-	-
2016	50,9	-	-	49,1	0,2	-	-	0,1
2017	50,9	-	-	49,4	0,2	-	-	-6,2
2018	50,6	-	-	49,4	-28,3	-	-	-28,3

En la tabla 2.7 se puede apreciar un aumento del consumo en el sector residencial, así como, una fluctuación del promedio mensual consumido alrededor de una media de 312.1 kW.h.

Tabla 2.7: Consumo de energía residencial. Fuente: (ONEI Cienfuegos, 2019)

CONCEPTO	UM	2013	2014	2015	2016	2017	2018
Electricidad	GW.h	149,4	153,1	157,6	160,5	161,7	171,1
Electricidad		5,2	2,5	2,9	1,8	0,7	5,8
	Tasas(%)						
<b>Consumo promedio mensual</b>	<b>kW.h/cliente</b>	<b>322,1</b>	<b>303,2</b>	<b>328,1</b>	<b>302,9</b>	<b>308,8</b>	<b>307,6</b>
De ello: Residencial		216,3	210,3	218,4	210,3	215,7	271,4
Nivel de Electrificación	%	100,0	100,0	100,0	100,0	100,0	100,0

El gas licuado en el municipio de Cienfuegos hasta mayo de 2017 para el sector residencial era por asignación, solo el 29% de las viviendas contaban con este combustible.

A partir de junio de dicho año cambia la estructura de consumo a partir del inicio de su venta liberada, en la figura 2.2 se puede apreciar el incremento considerable tanto en clientes como del consumo promedio mensual de este portador posterior a su liberación tomado de (Campillo Sabina, 2018).



Figura 2.2: Consumo de GLP. Fuente: Elaboración propia

### Análisis de la Gestión Energética Local en el municipio de Cienfuegos

Con la investigación de Campillo Sabina (2018) se obtuvo que los consumos de energía eléctrica a nivel municipal se llevaban sin tener en cuenta las características de los CP que permitan tomar decisiones al gobierno local al respecto y fomentar la utilización de fuentes alternativas de energía. En este estudio inicial se considera la energía eléctrica demandada por el sector

estatal y residencial para un período de nueve años que comprende desde el año 2007 al año 2016, en el municipio de Cienfuegos.

Los datos provienen de la Organización Básica Eléctrica(OBE), donde el consumo de energía eléctrica se registra a través de las cinco sucursales en el municipio, siendo estas:

- Bahía
- Centro
- Gloria
- Caunao
- CEN

De forma general, el sector estatal en el periodo 2007-2016 tuvo la tendencia a disminuir su demanda; sin embargo, el sector privado presenta un comportamiento creciente e irregular debido a que en el mismo se encuentran las cooperativas y el sector residencial donde confluyen no solo los hogares sino una gran inserción del sector privado que demandan energía eléctrica esto se evidencia en la figura 2.3.

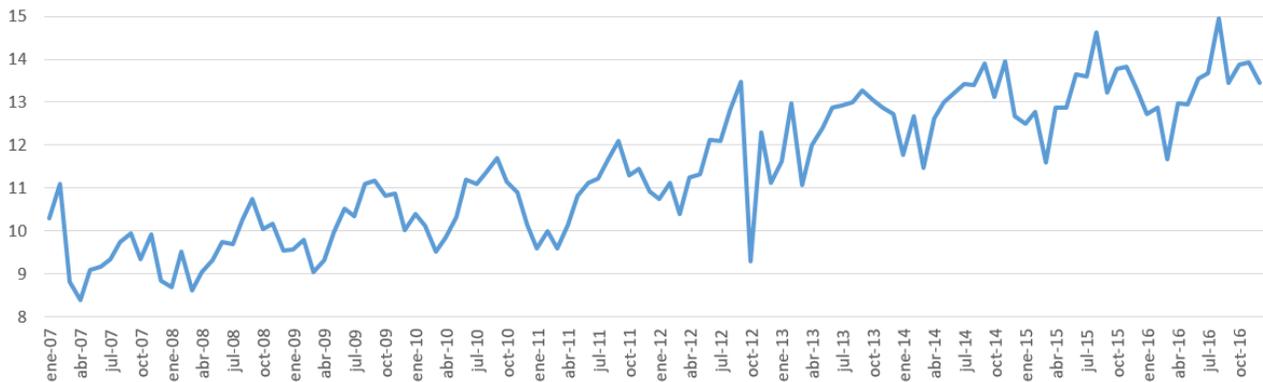


Figura 2.3: Consumo de energía eléctrica del municipio Cienfuegos en el periodo 2007-2016 para el sector privado en GW/h. Fuente: Elaboración propia

Un análisis de consumo de energía eléctrica en el sector privado es de vital importancia, con sus rutas correspondientes a cada Consejo Popular (CP); a partir del año 2015 la OBE comienza a registrarlo como información. Haciéndose necesario determinar los CP por cada Sucursal, para esto, Campillo Sabina (2018) realizó el análisis de las 243 rutas (trayectoria por la que se hacen las lecturas de los metrocontadores de los consumidores residenciales), en la tabla 2.8 se muestran los resultados.

Tabla 2.8: Consejos Populares por sucursal. Fuente: (Campillo Sabina, 2018)

Sucursal	Rutas	Consejos populares	Cantidad CP
Bahía	77	Buena Vista-Esperanza, Junco Sur, Paraíso, Pastorita-Obourke, Pueblo Griffó, Rancho Luna San Lázaro	7
Caunao	20	Caunao, Guaos y Pepito Tey	3
Centro	56	Centro Histórico, Junco Sur, Punta Gorda y Reina	4
Gloria	72	Juanita, Juanita II, Junco Sur, La Barrera, La Gloria, San Lázaro y Tulipán	7
CEN	18	Castillo CEN	1

En el estudio realizado por Campillo Sabina (2018) determinó que en todos los CP del municipio de Cienfuegos el mes de febrero constituye el de menor consumo de energía eléctrica y el mes de julio, el de mayor consumo en el año coincidiendo en el mes más frío del año y el inicio del verano respectivamente.

En la figura 2.4 se muestran el comportamiento del consumo de energía eléctrica por CP durante el año 2016, ordenados descendientemente, se puede apreciar su significación en el consumo municipal. Evidenciándose que los CP de mayor consumo de energía eléctrica son: Centro Histórico, Punta Gorda, Juanita y Tulipán, siendo los de menor consumo Rancho Luna, La Barrera y Guaos.

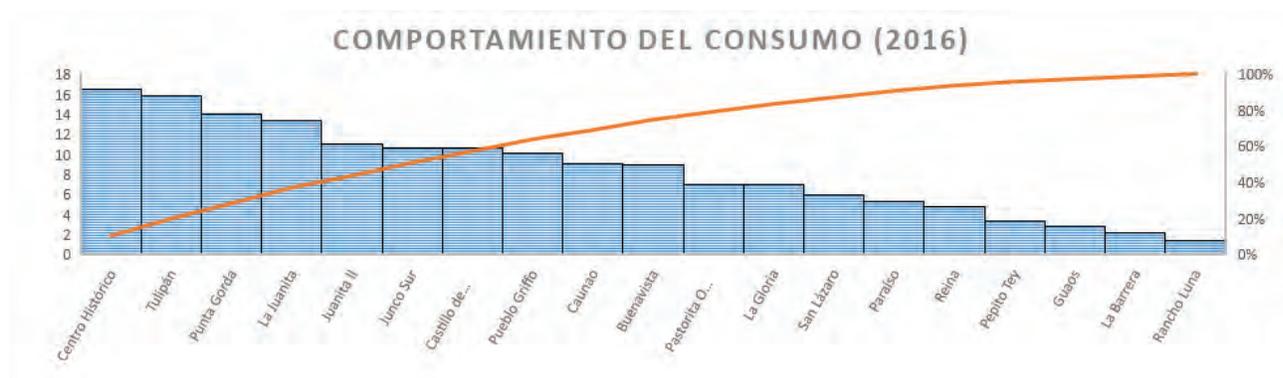


Figura 2.4: Diagrama Pareto sobre la demanda en los CP de Cienfuegos. Fuente: Elaboración propia

A partir de los datos obtenidos durante el periodo 2007-2016 se realiza un análisis de cluster con el objetivo de descubrir si los CP presentan un comportamiento homogéneo en cuanto al consumo de energía eléctrica.

En el dendograma (figura 2.5) se puede observar la formación de 4 grupos que poseen consumos similares. Estos se han enmarcado y enumerados para su fácil apreciación; el primer grupo, se corresponde con los CP que presentan un mayor consumo siendo el que posee una distancia significativa respecto a los demás, confirmando el comportamiento apreciado en el diagrama de Pareto anteriormente presentado. Si se ordenan los CP por población, consumo y hogares que cocinan con energía eléctrica no coinciden los de mayor consumo con alguna de estas dos variables. Cabe destacar que el consumo energético de los consejos populares no parece guardar relación con la población que poseen o la cantidad de hogares que cocinan con equipos eléctricos; a criterio del autor está más asociado a los hábitos de consumo de los pobladores en cada uno estos, siendo esta una variable que se tuvo en cuenta en varios estudios (Black & Henson, 2013; Mirjat *et al.*, 2017; Fallah, Deo, Shojafar, Conti, & Shamsirband, 2018; Karunathilake, Perera, Ruparathna, Hewage, & Sadiq, 2018; Nalcaci *et al.*, 2019; Zheng & Weng, 2019).

Este grupo representa el 21 % de los CP y su consumo equivale al 37 % del total del año 2016, si a este se le añaden los del segundo grupo representan el 52 % de los consejos populares de Cienfuegos y el 75 % del consumo de energía.

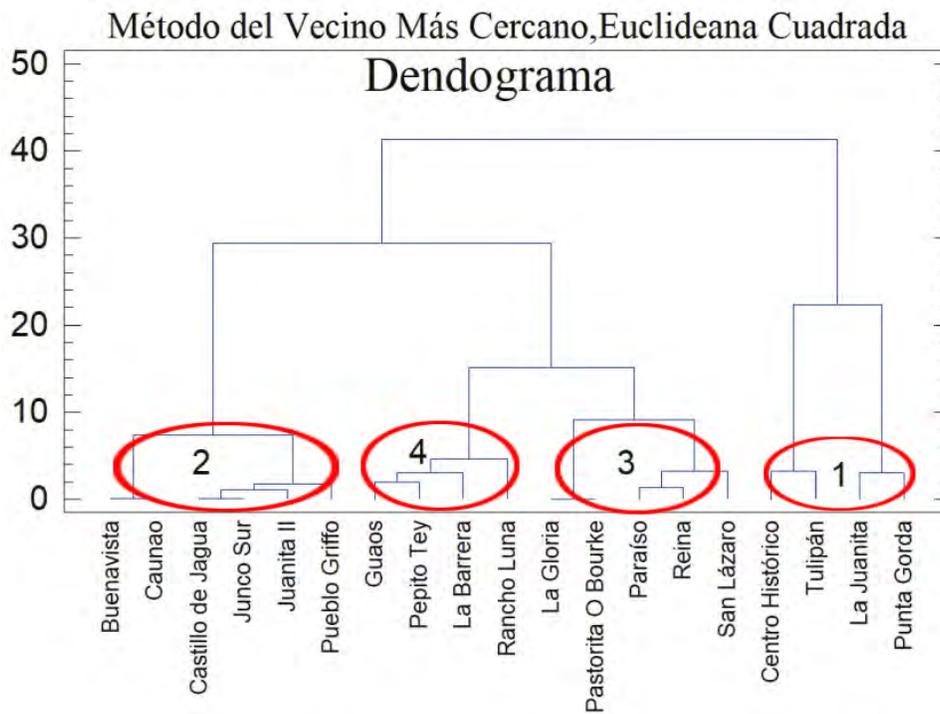


Figura 2.5: Clasificación de los CP por grupos de consumo. Fuente: Elaboración propia

En la figura 2.6 se puede observar la distribución espacial del consumo de energía mediante la degradación de colores en una escala de rojo que va desde color fuerte para aquellos CP con mayor consumo hasta una tonalidad más clara, para este análisis se tomó como base el consumo anual de cada CP en el 2016.

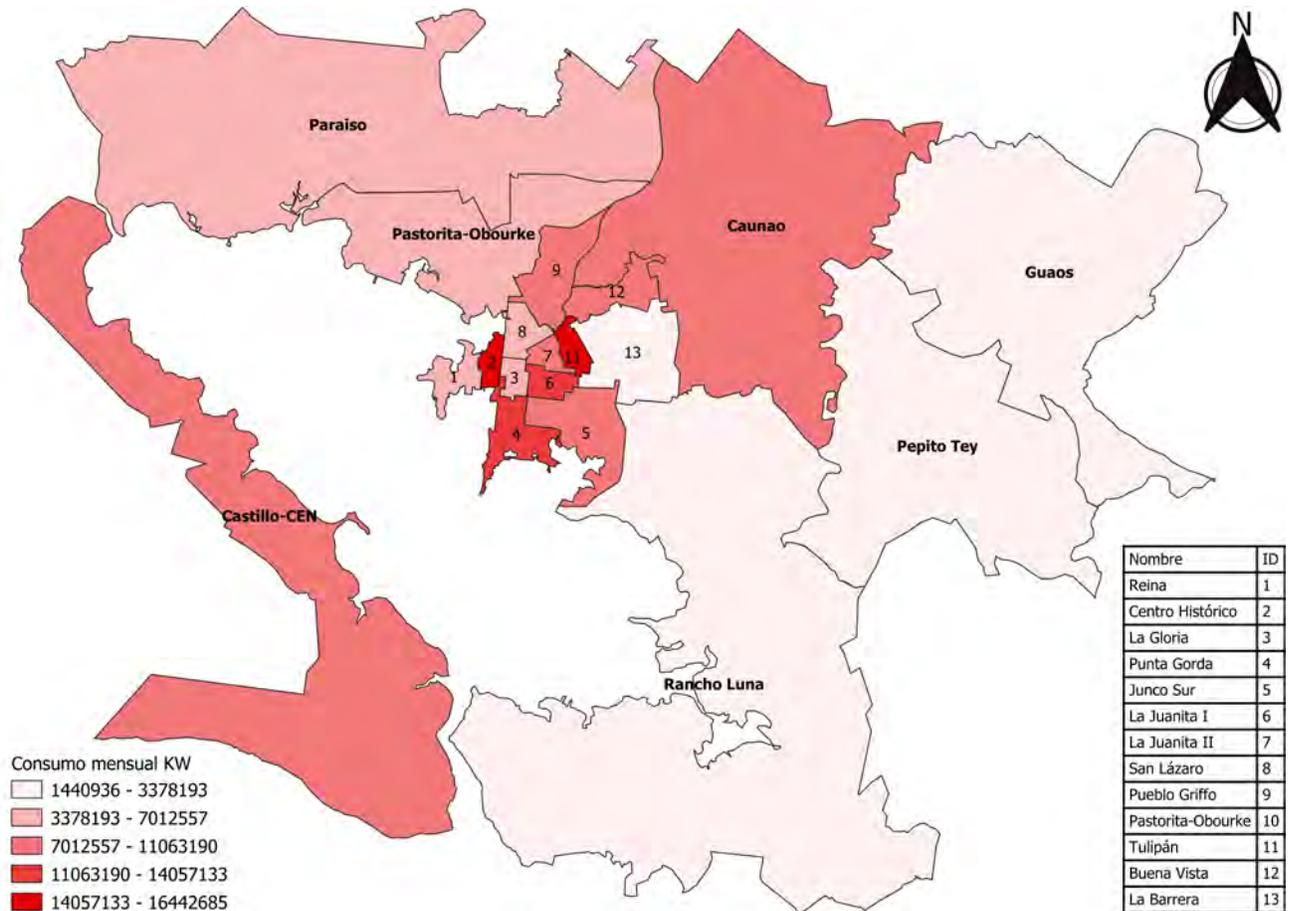


Figura 2.6: Consumo de energía eléctrica por CP. Fuente: Elaboración propia

## 2.2. Antecedentes de investigación.

La GEL se basa en la planificación de la energía, las matrices de generación y consumo energético incluyendo las fuentes renovables de energía (FRE) e indicadores energéticos que posibilitan la gestión de los gobiernos locales sobre los recursos energéticos territoriales (Correa Soto *et al.*, 2018).

En Cuba, el balance energético se realiza a nivel país mediante la metodología de OLADE, donde la matriz energética cubana se sustenta en el 95,7% de combustibles fósiles (Rodríguez *et al.*, 2018) requiriendo un cambio en su estructura con una mayor penetración de las FRE.

El contexto cubano favorece esta transición dado a que el país con la proyección de la actualización del Modelo Económico y Social en el 2011 se inicia un cambio de enfoque hacia la energía sostenible, en el 2014 con la aprobación la Política para el desarrollo prospectivo de las fuentes renovables y el uso eficiente de la energía, haciendo énfasis en elevar la eficiencia energética, un cambio de la estructura de la matriz energética actual y su relación con la competitividad de la economía nacional; disminuyendo la dependencia de estos combustibles importados, sus costos energéticos y el impacto medioambiental (Correa Soto *et al.*, 2017).

En el año 2017 se aprueban las bases del Plan Nacional de Desarrollo Económico y Social (PNDES) hasta el 2030 (PCC, 2017) relacionado con los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas para ese periodo (ONU, 2015). Con la declaración de la protección de los recursos y el medioambiente como dimensiones del desarrollo sostenible y ejes estratégicos en el PNDES hasta el 2030, así como, la actualización de los lineamientos de la Política Económica y Social referentes a los territorios, con el de la política energética (Correa Soto *et al.*, 2017) (PCC, 2017), el Decreto-Ley No. 345/ 2017 “Del desarrollo de las fuentes renovables y el uso eficiente de la energía”, si como la instrucción y resoluciones complementarias al respecto (Consejo de Estado, 2019); rectorado por Constitución de la República de Cuba que en su artículo 169 reconoce la autonomía del municipio y en el artículo 75, la responsabilidad de proteger el medio ambiente y los recursos naturales con estrecha vinculación al desarrollo sostenible de la economía y la sociedad.

Una particularidad de importancia es que en el sistema cubano la mayoría de las organizaciones productoras y de servicios son propiedad del Estado, lo que facilita el papel del gobierno local para mejorar la gestión, eficiencia y uso racional de la energía (Correa Soto *et al.*, 2017, 2018). Por lo que en el país desde el año 2001 con el objetivo de mejorar la eficiencia energética en las organizaciones se aplicó la Tecnología de Gestión Total Eficiente de la Energía (TGTEE) en más de 200 organizaciones de producción y de servicios, cuyos resultados responden a los objetivos trazados a nivel de sus organismos superiores (Ministerios, Uniones de empresa, Grupos empresariales) y al Programa de la Revolución Energética a partir del 2005.

A partir de la adopción de la NC ISO 50001 comenzaron a aplicar procedimientos y metodologías en consonancia con ella (Correa Soto *et al.*, 2014). Aunque es insuficiente en la actualidad en Cuba se han certificado por esta norma empresas del Ministerio de Energía y Minas (MINEM) y del Ministerio de la Construcción (MICONS), sin embargo, muchas otras han incorporado la gestión de la energía como una potencialidad en la mejora de su desempeño empresarial, resultando obligatoria, a partir del 2019 para las entidades grandes consumidoras de portadores energéticos, la certificación por la NC ISO 50001 vigente (MINEM, 2019).

A pesar del consumo significativo de energía en el país, la gestión energética no ha sido una prioridad para los gobiernos locales; y se circunscriben a tratar de cumplir lo establecido por el MINEM y el Ministerio de Economía y Planificación (MEP). Por otra parte, la gestión de los recursos energéticos en Cuba se ha realizado a escala nacional sin tener en cuenta las características de cada territorio que les permita conocer y gestionar sus potencialidades energéticas tanto de la oferta como de la demanda, lo que no ha incentivado a los órganos de gobierno locales a incorporar la gestión de la energía en la gestión pública, teniendo un papel pasivo en el tratamiento y uso de la información.

Lo anterior evidencia la necesidad de que los gobiernos locales se involucren en la gestión de la energía en Cuba; sin embargo, los gobiernos municipales no poseen herramientas que les permita gestionar los recursos energéticos presentes en el territorio y desconocen la contabilidad de todos los flujos energéticos y las relaciones entre entradas y salidas de energía y sus transformaciones.

La gestión energética en el municipio de Cienfuegos ha transitado en los últimos años por varias investigaciones que van desde el diseño de los indicadores energéticos para el sector residencial en el municipio de Cienfuegos realizado por Armas y Santana Cruz (2017) para los cuales se auxiliaron de varias técnicas, a saber, diagrama causa-efecto, 5Ws y 2Hs, diagrama de flujo, análisis de regresión múltiple, entre otras.

Al mismo tiempo, Hurtado Espinosa (2017) se encontraba desarrollando un producto informático para la gestión de la energía en el territorio a partir de una aplicación web llamada “GEM” con varios módulos de interacción que abarcan las aristas de Educación, Socialización y Operacional. Ambas investigaciones abordan el problema de pronosticar la demanda de energía eléctrica a través de regresión lineal múltiple considerando todas las combinaciones posibles de las variables independientes y comparando los modelos obtenidos mediante  $\bar{R}^2$ , el  $C_p$  de Mallows y el  $MSE$ , que permitieron la determinación de líneas bases energéticas para el sector residencial por CP.

Más adelante, cronológicamente hablando, Cabello Justafré (2018) crea un indicador que permite medir la eficiencia energética municipal. Se valida las líneas bases energéticas concluyéndose que los pronósticos mediante RLM son validos, sin embargo, no lo suficientemente precisos. A razón de los resultados obtenidos, se selecciona ANN para el pronóstico teniendo en cuenta las mismas variables utilizadas por (Armas & Santana Cruz, 2017; Hurtado Espinosa, 2017).

La ANN se modeló a través de la herramienta Neural Network Toolbox de Matlab, del tipo feedforward backpropagation, con una función de entrenamiento (TRAINLM), formada por 3 capas (dos ocultas y una de salida), 5 variables independientes de entrada y una variable de salida. Las capas están formadas por 10,5 y 1 neuronas respectivamente, todas del tipo tangencial

sigmoidal como se aprecia en la figura 2.7.

Para el entrenamiento se empleó el método de gradiente descendente con retro-propagación del error. Estableciendo una comparación entre el valor real y el pronosticado.

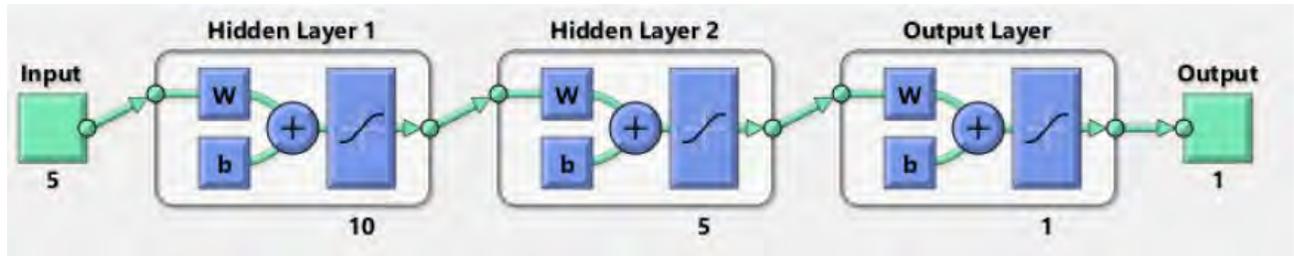


Figura 2.7: Modelo de ANN diseñada para cada CP. Fuente: (Cabello Justafré, 2018)

Chaviano Díaz y Puebla Alfonso (2018) realizan un análisis de calidad de la herramienta antes creada por Hurtado Espinosa (2017). El trabajo con expertos le permitió determinar las deficiencias existentes en el producto GEM. Se aplica el método general de solución de problemas para la definición y análisis de los problemas identificados, así como, la selección de las alternativas de solución y su implementación. Se utilizan herramientas como el análisis de modo fallo y errores (AMFE) para la priorización de los problemas existentes y el despliegue de la función calidad (QFD) para la mejora del producto GEM con el fin de permitir una mejor interacción con los diferentes actores de la GEL.

Campillo Sabina (2018) realiza un diagnóstico para determinar la situación energética del municipio de Cienfuegos permitiendo conocer la generación y consumo de energía en el territorio, la matriz de fuentes renovables de energía, así como, las potencialidades y barreras relacionadas con la sostenibilidad energética. Su propuesta se sustenta en el uso de técnicas y herramientas tales como: entrevista, revisión de documentos, trabajo con expertos, tormentas de ideas, 5Ws y 2Hs, diagrama pareto, mapeo de procesos, flujogramas, análisis de RLM, variabilidad, estabilidad, tendencia y pronósticos, histogramas, tabulación cruzada, prueba de ajuste e independencia, gráficos de control, gráficos de secuencias y de modelos ajustados, gráficos de radar/araña y MapInfo. Como conclusión fundamental se tiene que los gobiernos municipales, en Cuba, no poseen herramientas que les permitan gestionar los recursos energéticos presentes en el territorio.

Rodríguez Figueredo (2019) realizó una serie de propuestas con el objetivo de integrar las potencialidades energéticas locales a la Estrategia de Desarrollo Económico Social Municipal (EDSM); ofreciéndole al gobierno local una guía en la utilización de fuentes renovables de energía como recursos indispensables para un desarrollo próspero y sostenible.

Díaz La Hoz (2019) aplicando las consideraciones sugeridas por Chaviano Díaz y Puebla Alfonso

(2018) en su trabajo de diploma, perfecciona el producto informático GEM agregándole al apartado Educación nuevos contenidos que abarcan: artículos científicos, investigaciones, capacitación, Libros, Legislación y contenido multimedia e interactivo. En el apartado Operacional se trabaja con diferentes roles que permiten usar diferentes herramientas de acuerdo a los permisos asignados para cada cual con el fin de evitar una manipulación indebida de los contenidos y evitar el acceso a los datos utilizados para los pronósticos de demanda; en el de Socialización se pueden conocer noticias y eventos asociados a la gestión energética local.

## 2.3. Críticas

Para el diseño de indicadores energéticos en el sector residencial, Armas y Santana Cruz (2017), Hurtado Espinosa (2017) desarrollan el procedimiento de RLM y dejan propuestas las ANN para futuras investigaciones. Cabello Justafre (2018), Díaz La Hoz (2019), Rodríguez Figueredo (2019) retoman el procedimiento RLM y aplican ANN realizando una comparación del pronóstico obtenido entre ambas técnicas. Estos refieren que utilizan un procedimiento estadístico auxiliándose del Statgraphics, como una única acción; sin embargo, este software en su programación, tiene implementado la selección de modelos a partir de los criterios  $R^2$  Ajustado o  $C_p$  de Mallow indistintamente como una acción. Una vez seleccionado el “mejor modelo”, se procede al análisis de la regresión a partir de la ejecución de otro procedimiento (otra acción).

Se han definido múltiples criterios de selección de modelos en la literatura estadística. Según Gujarati y Porter (2013), algunos de ellos son muy empleados y los paquetes estadísticos incluyen, desde hace tiempo, información de este tipo, intercalada en sus rutinas de regresión. García (1996) afirma que cada criterio de selección posee condiciones ideales que se aspiran lograr para poder concluir cuál de los modelos es el mejor, por lo menos de forma individual, según dicho criterio (Gujarati & Porter, 2013; Olive, 2017). Algunos de estos se pueden apreciar en la tabla 2.9.

Castle, Doornik, y Hendry (2011, 2012) señalan que a menudo, la computadora (es decir, el modelador automático) puede encontrar un mejor modelo que el ser humano; asimismo puede ser capaz de encontrar varios modelos que son más o menos igual de buenos, proporcionando información adicional para tomar decisiones; entre estos métodos se tiene a los algoritmos de ampliación, reducción, métodos híbridos y métodos combinatorios.

Los algoritmos de ampliación, reducción y sus variantes tienen en común la estimación de las ecuaciones de regresión con un conjunto de variables y a continuación añadir o eliminar

Tabla 2.9: Criterios de selección. Fuente: Elaboración propia

Medidas de bondad del ajuste	$R^2$	$R^2 = 1 - \frac{SCE}{SCT}$
	$\bar{R}^2$	$\bar{R}^2 = 1 - \frac{SCE/(n-k)}{SCT/(n-1)}$
Criterios de Información	Cp	$Cp = \frac{SRC_k}{\hat{\sigma}^2} - (n - 2k)$
	AIC	$AIC = e^{2k/n} \frac{SCR}{n}$
	BIC	$BIC = n^{k/n} \frac{SCR}{n}$
Medidas de Error	MAD	$MAD = \frac{\sum_{t=1}^N  E_t }{N}$
	MAPE	$MAPE = \frac{\sum_{t=1}^N  \frac{E_t}{Y_t} }{N}$
	PMAD	$PMAD = \frac{\sum_{t=1}^N  E_t }{\sum_{t=1}^N  Y_t }$
	MSE	$MSE = \frac{\sum_{t=1}^N E_t^2}{N}$
	RMSE	$RMSE = \sqrt{\frac{\sum_{t=1}^N E_t^2}{N}}$

selectivamente variables hasta que se logre con alguna medida de criterio conjunta o regla de parada (Penelas & Gómez, 1999). En este caso ninguno de estos métodos asegura la satisfacción simultánea de los múltiples criterios de adecuación. También existe el riesgo de seleccionar modelos “aparentemente adecuados” cuando no lo son (por ejemplo modelos con significancia estadística pero no teórica). Douglas *et al.* (2002) afirman que los analistas sin experiencia pueden verse tentados a creer que como todos los procedimientos secuenciales terminan con una ecuación final, dicho modelo es óptimo en algún sentido, lo cual es incorrecto.

En cuanto a la selección del mejor modelo a partir de un único criterio, la literatura especializada en este tema propone el uso de más de un criterio de selección, no estando implementado en los softwares utilizados en las investigaciones analizadas.

Otro aspecto a señalar, en las investigaciones de Armas y Santana Cruz (2017), Hurtado Espinosa (2017), Campillo Sabina (2018), Díaz La Hoz (2019), Rodríguez Figueredo (2019) es que no se realizan las pruebas para analizar el cumplimiento de los supuestos de RLM mediante MCO enumerados en el epígrafe (1.3.1), lo que puede conducir a la invalidez de los resultados en caso de no cumplirse alguno de los siguientes: la normalidad, heterocedasticidad y autocorrelación; mientras que la violación del resto de los supuestos solo afecta la veracidad y confiabilidad del modelo. En resumen, los contrastes elaborados a partir de los estimadores MCO ya no son válidos

Ampliando la idea anterior, el MCO en sí para estimar el RLM, no requiere ningún supuesto

sobre la distribución de probabilidades de los residuos. Ahora bien, para hacer inferencia estadística sí se necesita de su cumplimiento o puede afectar las relaciones y la significación. La normalidad permite la estimación por intervalos de confianza para los coeficientes de regresión, para la predicción, permite el planteamiento de dójimas sobre los parámetros del modelo. Cuando los residuos no son normales, los intervalos y las pruebas de hipótesis no son exactas y pueden llegar a ser inválidas. En presencia de heterocedasticidad la estimación de la varianza estaría sesgada y sería no consistente; mientras la autocorrelación produce estimadores MCO ineficientes.

A criterio del autor cuando se realice la regresión mediante MCO se debería seguir el algoritmo que se muestra en la figura 2.8

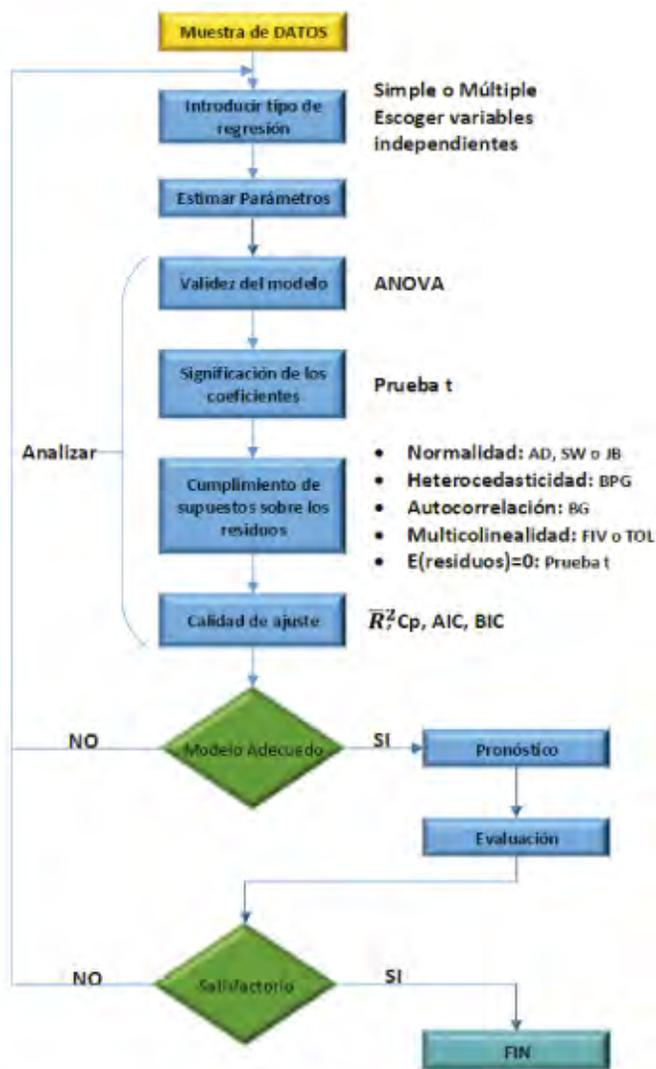


Figura 2.8: Algoritmo de regresión mediante MCO. Fuente: Elaboración propia

En los trabajos de Armas y Santana Cruz (2017), Hurtado Espinosa (2017) hace referencias

a seleccionar el mejor modelo a partir de todas las combinaciones posibles, pero se presentan solamente 5 combinaciones. Siguiendo el principio de la teoría combinatoria, se pueden obtener  $2^k - 1$  combinaciones siendo  $k$  el número de variables independientes, para este caso, 31. Haciendo impracticable mediante inspección visual la selección sobre la base de los dos criterios escogidos por los autores antes mencionados.

Al escoger el modelo de menor  $C_p$  no se tiene en cuenta las sugerencias de Gujarati y Porter (2013) sobre el mal uso de este criterio en cuanto a la selección de modelos a partir de su valor más pequeño; lo correcto sería seleccionar aquel modelo cuyo valor sea más cercano al número de variables independientes. Si se utiliza el algoritmo de selección de modelos propio de Statgraphic, el modelo seleccionado no concuerda con el planteado en estas investigaciones, a modo de comparación se expone en la figura 2.9.

**Tabla 3.5:** Criterio de evolución por número de variables independientes para el Consejo Popular Caonao. **Fuente:** Elaboración propia a partir de Kialashaki & Reisel, (2013).

Número de variables independientes	1	2	3	4	5
$R^2_{ajustada}$	69.9616	76.1035	77.9265	77.7836	77.6067
CME	2.4967E9	1.9862E9	1.83468E9	1.84656E9	78.5476
$C_p$	42.2854	10.8537	2.3434	4.09148	6.0
Variables incluidas	$H_1$	$X_3$ y $H_1$	$X_1$ , $X_3$ y $H_1$	$X_2$ , $X_3$ , $P_1$ y $H_1$	$X_1$ , $X_2$ , $X_3$ , $P_1$ y $H_1$

El modelo de regresión seleccionado para el CP Caonao responde a los criterios de decisión, la R-Cuadrada ajustada (mayor) y la estadística  $C_p$  (menor) quedando la recta resultante:

$$Y_1 = 33381.7 - 6823.75 \cdot X_1 + 11929.6 \cdot X_3 + 539.837 \cdot H_1$$

### Selección del Modelo de Regresión - y

Modelos con Mayor R-Cuadrada Ajustada				
		<i>R-Cuadrada</i>		<i>Variables</i>
<i>CME</i>	<i>R-Cuadrada</i>	<i>Ajustada</i>	<i>Cp</i>	<i>Incluidas</i>
1,75974E9	79,3619	78,8282	2,03483	BCE
1,77456E9	79,3675	78,6499	4,00383	BCDE
1,775E9	79,3624	78,6445	4,0322	ABCE
1,78102E9	79,1123	78,5721	3,41388	ACE
1,79006E9	79,3682	78,4633	6,0	ABCDE

Variable dependiente: y  
 Variables independientes:  
 A=x1  
 B=x3  
 C=x2  
 D=P1  
 E=H1

Figura 2.9: Comparación de modelos seleccionados. Fuente: Elaboración propia

De hecho, en todas estas investigaciones aparece la temperatura de bochorno en el modelo resultante siendo esta combinación de la humedad y la temperatura ambiente provocando la

existencia de multicolinealidad, otro supuesto de la regresión. Basta con analizar los supuestos ya sea gráficamente para tener una idea intuitiva del cumplimiento de estos o a través de dóxicas para su confirmación; esto se ejecuta con las limitaciones propias del software en cuanto a la robustez de las técnicas implementadas (Alvarez Basanta, 2020).

ANN constituyen una poderosa técnica para modelar fenómenos complejos donde la relación es aparentemente lineal en presencia de multicolinealidad, datos ruidos o contaminados, con distribuciones no normales de las variables. Esto se debe fundamentalmente a la capacidad de aprendizaje de los algoritmos que posee y de utilizar relaciones de transferencia no lineales. Para lo primero, se basa en particiones en los datos en tres conjuntos: de entrenamiento, de prueba, y un tercero de validación, que puede omitirse. En la figura 2.10 se muestra el proceso de aprendizaje automático, organizados en cuatro fases .

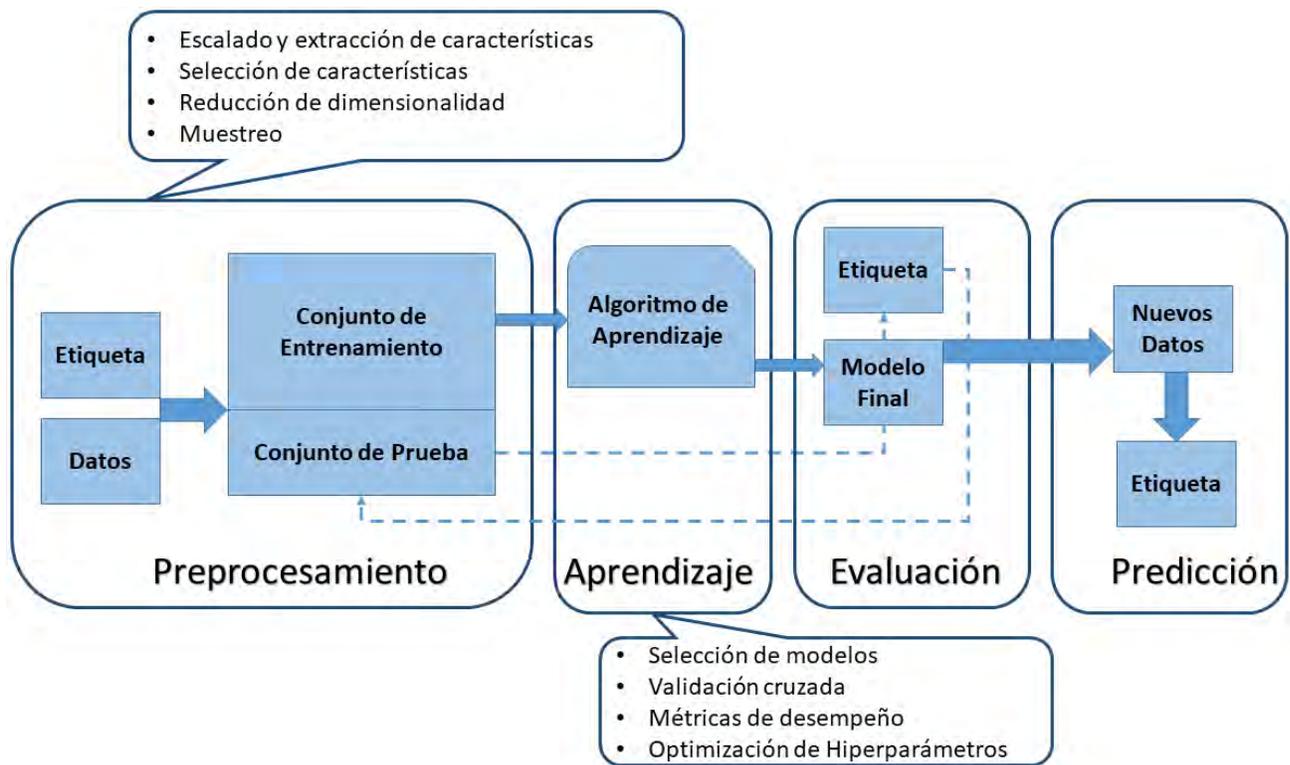


Figura 2.10: ANN. Aprendizaje automático. Fuente: Elaboración propia

Se enumeran algunas desventajas sobre las ANN referidas en (Mijwel, 2018; Tu, 1996):

- Dependencia del hardware: Las redes neuronales artificiales requieren procesadores con potencia de procesamiento en paralelo, de acuerdo con su estructura. Por este motivo, requiere una alta disponibilidad de recursos computacionales.
- Comportamiento inexplicable de la red: este es el problema más importante de la ANN.

Cuando produce una solución de sondeo, no da una pista de por qué es esa. y cómo se obtuvo. Esto reduce la confianza en la red.

- Determinación de la estructura de red adecuada: no existe una regla específica para determinar la estructura de las ANN. La estructura de red adecuada se logra mediante la experiencia, el ensayo y error.
- Dificultad para mostrar el problema a la red: las ANN trabajan con información numérica. Los problemas deben traducirse en valores numéricos antes de ser introducidos en la ANN. El mecanismo de visualización que se determine aquí influirá directamente en el desempeño de la red. Esto depende de las habilidades del usuario.
- Se desconoce la duración de la red: la red se reduce a un cierto valor del error en la muestra significa que la formación se ha completado. Este valor no nos da resultados óptimos.
- ANN funciona como una “caja negra” y tiene capacidad limitada para identificar explícitamente posibles relaciones causales.
- Tienen a sobreajustarse. El sobreajuste es una causa frecuente de un alto error de generalización. implica que el modelo ha sobreinterpretado patrones en los datos de entrenamiento y se ha ajustado demasiado bien incluyendo ruido.
- Su desarrollo es empírico y quedan muchos problemas metodológicos por resolver.
- Tienen problemas en la estimación de cálculos precisos. Funcionan bien con problemas complejos de difícil cálculo pero que no requieren respuestas exactas, sino solo respuestas rápidas y buenas.
- Se comportan bien en predicción a largo plazo con componentes no lineales; en cambio, no están claras las mejorías observadas en series cortas y estacionales.
- Tiempo de aprendizaje. esto depende de dos factores: el primero si se incrementa la cantidad de patrones a identificar o clasificar y segundo si se requiere mayor flexibilidad o capacidad de adaptación de la ANN para reconocer patrones que sean sumamente parecidos, se deberá invertir más tiempo en lograr que la red converja a valores de pesos que representen lo que se quiere enseñar.
- No permite interpretar lo que se ha aprendido, la red por si sola proporciona una salida, un número, que no puede ser interpretado por ella misma.

A criterio del autor, para el pronóstico de demanda eléctrica, ANN presenta una debilidad, y es que debido al uso de pesos aleatorios como parte de su función de transferencia, los resultados varían de en cada ejecución. Otra cuestión es que no ofrece un modelo funcional lineal, adicionalmente, se tiene su alto costo computacional. En la figura 2.11 se ofrece de manera resumida los elementos de esta técnica en comparación de la RLM y MARS.

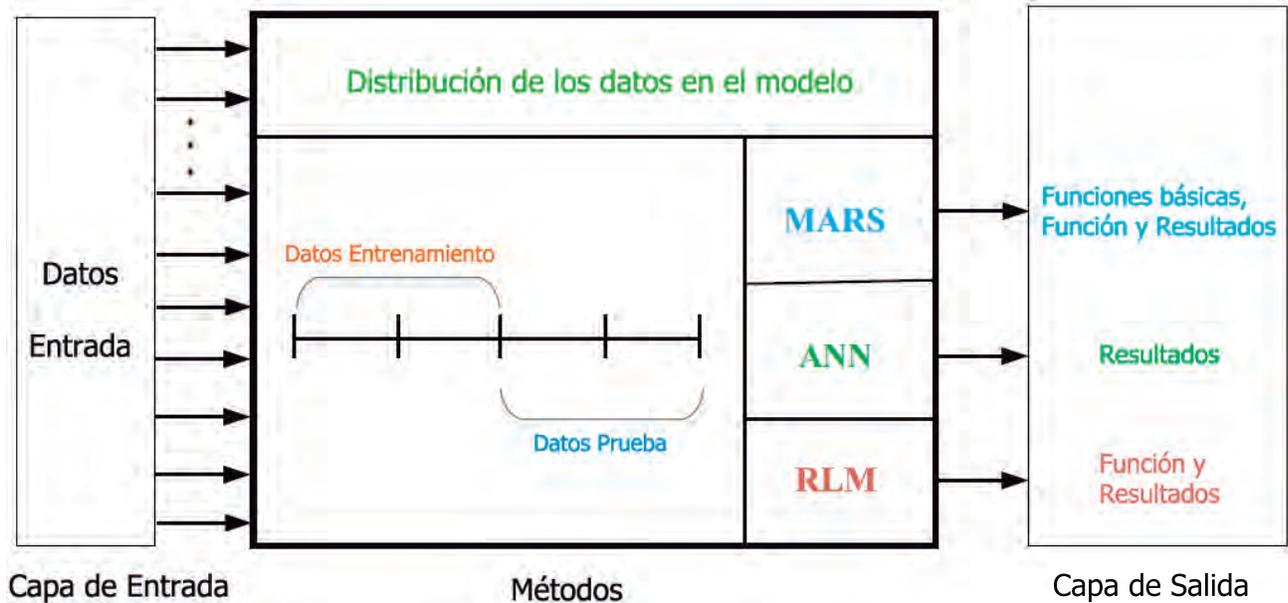


Figura 2.11: Resumen de las técnicas. Fuente: Elaboración propia

Se utilizan herramientas auxiliares privativas (MatLab, Statgraphics) lo que conduciría al pago de licencias de uso así como sus actualizaciones en el paso del tiempo como se refiere en epígrafe 1.4, en caso contrario se incurriría en una violación de licencia de uso.

En 2020, la revista IEEE spectrum, elabora anualmente un ranking de popularidad de los principales lenguajes de programación (ver figura 2.12), siguiendo un proceso que permite al usuario escoger los criterios para ordenar esta lista, además de establecer 4 categorías: web, móvil, empresarial y sistemas embebidos. El gran ganador en esta comparativa fue Python, siendo el primero en todas las categorías en las que está incluido.

Cuando se comenta sobre las características del sistema diseñado, tales como, flexibilidad etc, se menciona el costo computacional. El sistema web no presenta problemas al respecto, por su parte MatLab es un alto consumidor de recursos, este necesita un mínimo de 4 Gb de RAM siendo 8 Gb lo recomendado, mientras que para una instalación típica necesita de 5-8 Gb recomendando el uso de SSD (MathWorks.Inc, Consultado en: Feb, 2021).

Language Ranking: IEEE Spectrum							
Rank	Language	Type	Score	Rank	Language	Type	Score
1	Python	🌐 🖥️ ⚙️	100.0	13	SQL	🖥️	64.6
2	Java	🌐 📱 🖥️	95.3	14	PHP	🌐	63.8
3	C	📱 🖥️ ⚙️	94.6	15	Assembly	⚙️	63.7
4	C++	📱 🖥️ ⚙️	87.0	16	Scala	🌐 📱 🖥️	63.5
5	JavaScript	🌐	79.5	17	HTML	🌐	61.4
6	R	🖥️	78.6	18	Kotlin	🌐 📱	57.8
7	Arduino	⚙️	73.2	19	Julia	🖥️	56.0
8	Go	🌐 🖥️	73.1	20	Rust	🌐 🖥️ ⚙️	55.6
9	Swift	📱 🖥️	70.5	21	Shell	🖥️	52.0
10	Matlab	🖥️	68.4	22	Processing	🌐 🖥️	49.2
11	Ruby	🌐 🖥️	66.8	23	C#	🌐 📱 🖥️ ⚙️	48.1
12	Dart	🌐 📱	65.6	24	SAS	🖥️	45.2

Figura 2.12: Ranking de uso de lenguajes de programación 2020.

Fuente: (<https://spectrum.ieee.org/static/interactive-the-top-programming-languages-2020>)

Existen dos formas de ejecutar un código de MatLab desde una aplicación web, la primera sería convertir el código hacia Javascript mediante el toolbox MatLab Coder, este nuevo código se puede ejecutar de forma embebida desde cualquier navegador, en este caso ya no sería MatLab el software a utilizar sino Javascript. La segunda opción sería instalarlo en el servidor y ejecutar el *código* previamente guardado en un archivo de manera silente (sin interfaz gráfica), esta opción conllevaría a un alto consumo de recursos del lado del servidor.

Statgraphic, por su parte, debe estar instalado físicamente en la pc en cual se está haciendo uso de la herramienta GEM solo existiendo de este version para Windows, lo cual afectaría la portabilidad de dicha herramienta web. Además, el usuario debe interactuar con este software estadístico con el fin de realizar el procedimiento para el que debe poseer conocimientos medios sobre la herramienta y sus procedimientos, así como, del uso de técnicas estadísticas.

Díaz La Hoz (2019) plantea que la herramienta web está soportada sobre python 3.0, haciendo alusión a que pudiera ser la herramienta fundamental en el procesado de los datos; sin embargo, a pesar de esta reflexión, no se utiliza para este propósito.

Python es una herramienta de desarrollo verdaderamente maravillosa que no solo sirve como un lenguaje de programación de propósito general, sino que también se adapta a nichos específicos de cualquier proyecto o flujos de trabajo. Con múltiples bibliotecas y paquetes que amplían sus capacidades, lo convierten en una herramienta versátil de ajuste perfecto para cualquiera que busque desarrollar programas y algoritmos. Consta con bibliotecas modernas de aprendizaje, automático y profundo. Las más conocidas para el procesamiento estadístico fueron referidas en el epígrafe 1.4; para ANN las más conocidas son:

- TensorFlow
- Keras
- Scikit-learn
- Theano
- PyTorch
- NeuroLab
- ffnet
- MXNet
- Lasagne
- pyrenn
- Deepy
- Blocks
- Nolearn
- NeuPy
- Scikit-Neural Network

Como interfaz universal para gran variedad de plataformas de aprendizaje automático, Keras se utiliza actualmente en muchos proyectos del sector de la IA. Si la biblioteca ya registraba más de 250 000 usuarios individuales a mediados de 2018, esta cantidad ha aumentado significativamente, al menos desde que incorporó TensorFlow. La libre elección del framework subyacente, la licencia gratuita y el concepto multiplataforma hacen de Keras la solución perfecta para las aplicaciones de redes neuronales profesionales, tanto en la industria como en la investigación. Por ejemplo, empresas famosas como Netflix, Uber y Yelp, así como, grandes organizaciones como la NASA o la CERN, confían en Keras o en el paquete TensorFlow-Keras para llevar a cabo sus proyectos.

Una ventaja resulta en la flexibilidad de hacer uso de varias bibliotecas al mismo tiempo para el procesado de datos. De esta forma, los resultados obtenidos de un análisis pueden utilizarse como datos de entrada al usar cualquier otra librería como se puede apreciar en la figura 2.13.



Figura 2.13: Mejores bibliotecas de Python para Machine Learning y Deep Learning.

Fuente: (<https://towardsdatascience.com/best-python-libraries-for-machine-learning-and-deep-learning-b0bd40c7e8c>)

A pesar de todo lo antes mencionado en estas investigaciones se crea un precedente en cuanto al cambio de pensamiento sobre la necesidad de la gestión energética desde los consejos populares, la necesidad imperativa sobre el uso fuentes alternativas de energía, se crea una herramienta informática que permite la interacción con los diferentes actores que intervienen en la planeación energética y los dota de un mayor conocimiento y poder de decisión en cuanto a las políticas y objetivos a trazarse para lograr en un futuro constar con localidades sostenibles.

## 2.4. Conclusiones parciales:

- Se caracteriza el consumo de energía eléctrica residencial municipal, realizando una clasificación a nivel de CP aportándose una representación espacial del mismo.
- En las investigaciones precedentes en el municipio Cienfuegos no se consideran los supuestos de la regresión lo cual puede invalidar los resultados. No se ofrecen medidas de calidad y del error de las técnicas aplicadas.
- El producto GEM está implementado en *Python* y no utiliza las potencialidades del mismo en contraposición del uso de software privativo.

# APLICACIÓN DE LAS TÉCNICAS ANALIZADAS EN EL PRONÓSTICO DE LA ENERGÍA ELÉCTRICA EN EL SECTOR RESIDENCIAL

---

En este capítulo se tiene la implementación de las técnicas (*RLM*, *ANN*, *MARS*) para las cuales se hace uso de varias librerías de *Python* a saber: *statsmodels*, *sklearn*, *pyearth* las cuales se aplican a varios CP y se realizan comparaciones entre estas.

## 3.1. Análisis municipal.

En nuestro país el único proveedor de energía eléctrica es la Union Eléctrica (UNE) esta a partir de su redes de distribución llega al 100 % de la población y es la encargada de gestionar aquellas FER capaces de abastecer a varios clientes, ejemplo de esto en nuestro municipio se puede citar el parque fotovoltaico de Cantarranas.

Para un correcto entendimiento de la demanda eléctrica se debe tener claro que los datos procesados corresponden al consumo del mes anterior, ya que el sector residencial a diferencia de los demás clientes de la empresa eléctrica que deben pagar su demanda con anterioridad, estos pagan después de haberla consumido cobrándose mensual de ahí que solo se tengan datos de la demanda una vez realizada la lectura por parte de los lectores de la empresa eléctrica.

A partir de los datos obtenidos en las diferentes fuentes de difusión de estos se procede a realizarla un análisis de los mismo con el fin de detectar datos atípicos, falta de consistencia, etc.

Tabla 3.1: Definición de variables a utilizar

Variable	Descripción	Variable	Descripción
X1	Temperatura seca del aire	X4	Población
X2	Temperatura de bochorno	X5	Viviendas que usan cocción eléctrica
X3	Humedad	Y	Consumo del CP

Se cargan las librerías necesarias para el análisis y procesamiento mediante el siguiente código

```
#####
#                               GENERAL
#####
import pandas as pd # Import XLS File
import numpy as np
import scipy as sci
import matplotlib.pyplot as plt
import pyod
import statistics as stat
#####
#                               OLS
#####
import statsmodels.api as sm
import statsmodels.stats.api as sms
#####
#                               SKLEARN
#####
from sklearn.model_selection import train_test_split
from sklearn.neural_network import MLPRegressor
from sklearn.preprocessing import StandardScaler
from sklearn import metrics as skm
#####
#                               MARS
#####
from pyearth import Earth
#####
# Load Data
data = pd.read_excel(r 'BD.xls ', sheet_name='Tulipan ')
X = data [['X1', 'X2', 'X3', 'X4', 'X5']].values
y = data ['Y'].values
#####
```

A partir de esta implementación los datos pueden ser introducidos a un archivo de Excel(\*.xls) el cual puede contener los CP en hojas diferentes dentro de un mismo libro, esto facilita el trabajo de recopilación y organización de datos pudiéndose generar desde cualquier medio informático sin importar su entorno de trabajo, entiéndase sistema operativo, paquete de ofimática.

En el primer grupo de líneas se cargan aquellas librerías utilizadas para la manipulación, preprocesado de los datos y generación de gráficos. Posteriormente, le corresponden a aquellos

paquetes utilizados para efectuar la regresión mediante el método MCO. El tercer conjunto corresponden a las librerías *SKlearn* que permiten la modelación, entrenamiento y análisis mediante ANN y finalmente, se importa el módulo *pyearth* para *MARS*.

Por último se hace lectura del archivo de datos y se le asigna a las variables los valores correspondientes del mismo.

En la figura 3.1 se puede observar la existencia de un dato atípico correspondiente al mes de octubre de 2012. Este mismo fenómeno se aprecia en la base de datos de cada CP.

Para la detección de los datos atípicos se utiliza la librería *pyod* implementada por Zhao, Nasrullah, y Li (2019) que contiene más de 20 algoritmos los cuales cubren desde técnicas clásicas de detección hasta arquitectura de ANN. Se utiliza el método *kNN* para la detección y sustitución de los outliers por la media de los *k* vecinos próximos, en este caso  $k = 3$ .

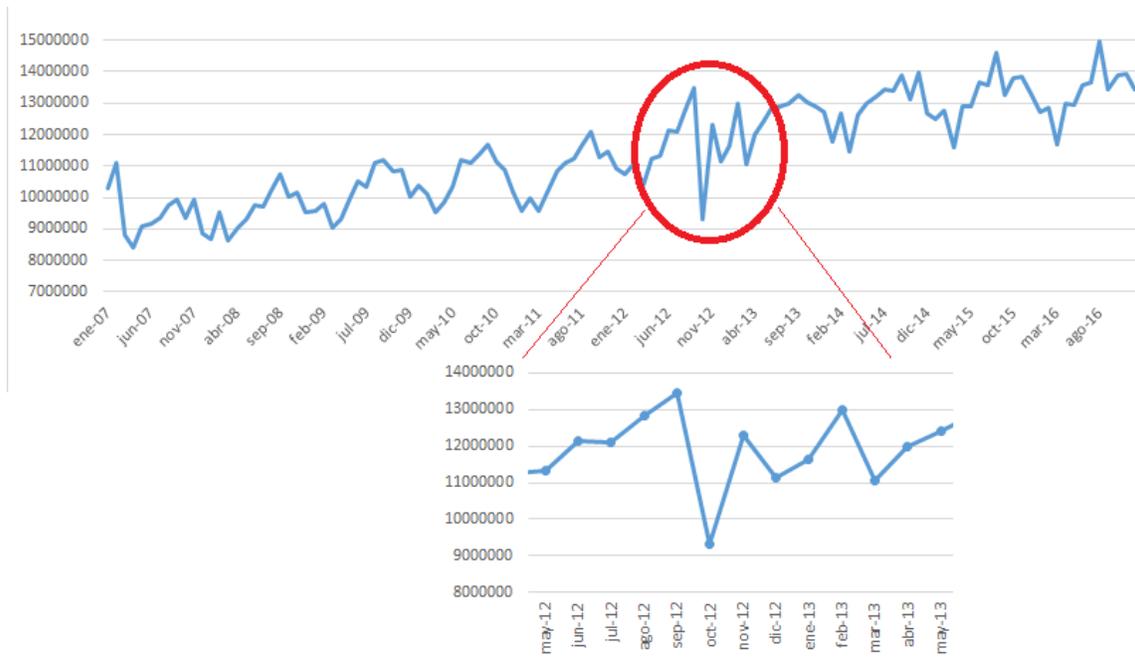


Figura 3.1: Datos atípicos municipio Cienfuegos. Fuente: Elaboración propia

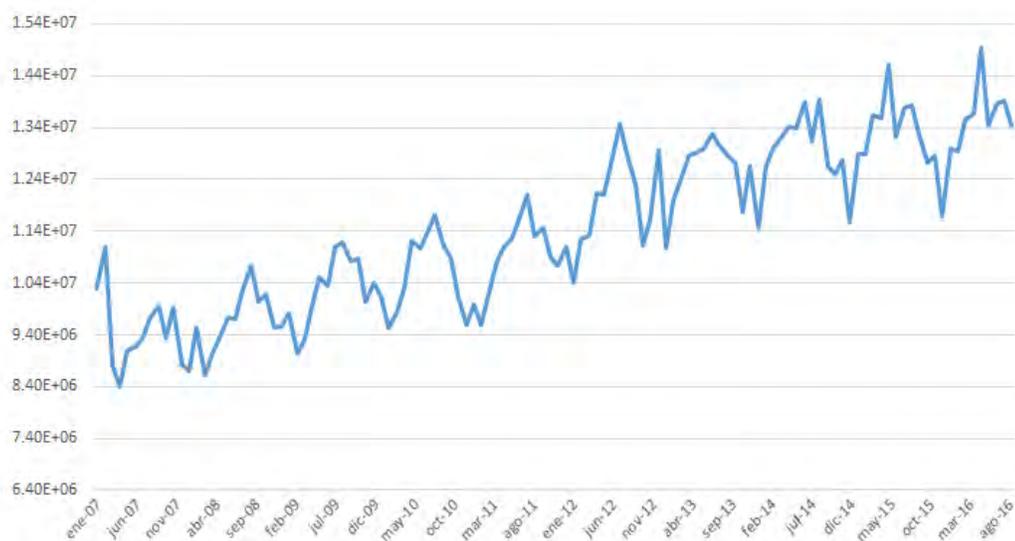


Figura 3.2: Datos procesados municipio Cienfuegos. Fuente: Elaboración propia

Como se puede apreciar en la figura 3.2 se obtiene una secuencia más suave, con una tendencia creciente, con cierta estacionalidad anual; una característica revelada en el gráfico es el aumento

de la demanda a partir del mes de marzo haciendo un pico en el tercer trimestre del año mientras que a partir del cuarto trimestre desciende el consumo presentando su mínimo en el mes de marzo. Repitiéndose el mismo patrón a lo largo de los 10 años analizados.

Ya preprocesados los datos están listos para proceder a la obtención del modelo por las técnicas analizadas. Tratando de extraer información de los datos se realiza un análisis de correlación de las variables (premisa fundamental para establecer un modelo).

La figura 3.3 muestra la matriz de correlación entre las variables analizadas, a su vez, como un mapa de colores en el cual se aprecia el coeficiente de *Pearson* una mayor intensidad de los colores equivale a una mayor relación entre las variables el azul representa la relación positiva, mientras que el rojo indica que esta es negativa.

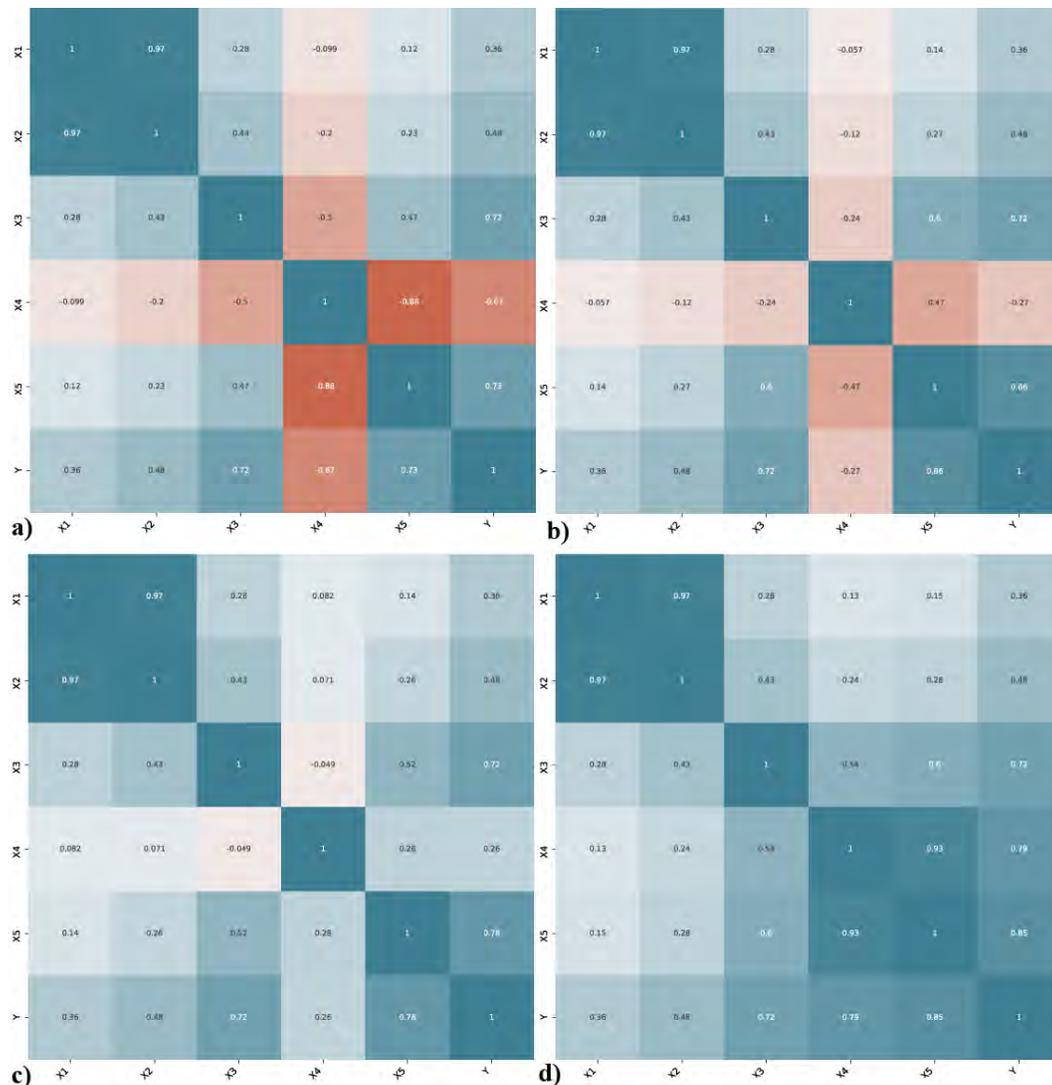


Figura 3.3: Covarianza entre variables de los CP. Fuente: Elaboración propia  
a)Tulipán, b)R Luna, c)La Gloria, d)Caunao

En el caso de Tulipán se tiene que las variables que presentan una correlación leve son  $X1$  con  $X4$  y  $X5$ , cabe resaltar que  $X4$  tiene un efecto negativo tanto con  $X1$  como con la variable de respuesta, por otro lado el par  $X4, X5$  presenta una fuerte relación entre ambas variables; este mismo fenómeno se aprecia en Rancho Luna pero de una menor intensidad. El en CP La gloria por otra parte solo presentan correlación negativa del par  $X3, X4$ .

De manera general, se puede afirmar que en todos los CP se observa el mismo fenómeno de correlación entre las variables, esto ofrece evidencia de que las variables independientes no lo son o comparten información. Todas las variables independientes muestran una correlación ya sea positiva o negativa respecto a la variable dependiente. Al existir algunas variables independientes cuya relación es leve se procederá al análisis de regresión lineal y comparación de los 3 modelos susceptibles de obtenerse a partir de estas variables ( $X1, X5$ ) para determinar los que posean una mayor explicación del fenómeno y no incumplan aquellos supuestos de la regresión que no pueden violarse.

Siguiendo el algoritmo propuesto en 2.8 se analizaron todos los posibles modelos (31 modelos) dejando solo los candidatos que cumplen las restricciones fundamentales (normalidad, homocedasticidad, no multicolinealidad). Estos candidatos se refinan escogiendo el que ofrece mejor ajuste de acuerdo a  $\bar{R}^2$  de este se muestran los resultados del ajuste de regresión lineal y se compara con ANN y MARS.

Con este fin se definen las variables dependientes e independientes y se procede a generar las  $2^k - 1$  combinaciones mediante la función potencia definida en el pseudocódigo; posteriormente se recorre mediante un ciclo todos los modelos generados y se les realiza la regresión guardando sus resultados en un diccionario para su posterior comparación.

```

### Generate models 2^k-1
vardep='Y~'
varind = ['X1', 'X2', 'X3', 'X4', 'X5']
model=[]
def potencia(c):
    if len(c) == 0:
        return [[]]
    a = potencia(c[:-1])
    return a+[s+[c[-1]] for s in a]
def imprime(c, prefix):
    for e in sorted(c, key=lambda s: (len(s), s)):
        model.append(prefix + "-" + ".join(e))
    return model
models =imprime( potencia(varind), vardep)
del models[0]
#####
#OLS
for i in range(0, len(models)):
    results = sm.ols(formula=models[i], data=data).fit()

```

En la tabla 3.2 se muestran los modelos obtenidos para el municipio de Cienfuegos. Se hace notar que los modelos expuestos son todos válidos, los coeficientes resultan significativos, los residuos resultan distribuidos normalmente, con media cero, homocedásticos y no existe presencia de multicolinealidad aparente. Todos los modelos presentan residuos autocorrelacionados.

También se puede apreciar como el modelo  $\hat{Y} = X_1 + X_3 + X_5$  el que mayor explicación ofrece de la variabilidad del consumo (81 %) siendo este el seleccionado para realizar la comparación.

Tabla 3.2: Modelos de RLM para el municipio de Cienfuegos

	$\bar{R}^2$	ANOVA	Prueba $t$	AD	$E \mu = 0$	BPG	BG	TOL
$\hat{Y} = X_5$	0.695463	SI	SI	SI	SI	SI	NO	SI
$\hat{Y} = X_1 + X_5$	0.753184	SI	SI	SI	SI	SI	NO	SI
$\hat{Y} = X_2 + X_5$	0.766977	SI	SI	SI	SI	SI	NO	SI
$\hat{Y} = X_1 + X_3 + X_5$	0.810078	SI	SI	SI	SI	SI	NO	SI

El modelo obtenido mediante MCO para el municipio de Cienfuegos se expresa matemáticamente como:

$$Y = -3,965 \cdot 10^6 + 1,288 \cdot 10^5 X_1 + 6,751 \cdot 10^4 X_3 + 157,1793 X_5.$$

Se aplica la técnica ANN y MARS, con el objetivo de mejorar la calidad de ajuste y mejorar el pronóstico sobre la variable consumo eléctrico.

Para configurar y el entrenamiento de ANN, la implementación se muestra a continuación. Se definen 2 capas ocultas con 10 y 5 neuronas respectivamente, con la función de activación *relu* y el solver *adam*, con una tasa de aprendizaje adaptativa.

Como medida de referencia para el error se utilizó el *MSE*, *RMSE* y *MAE*.

```
#####
# SKLEARN
#####
X = data[['X1', 'X2', 'X3', 'X4', 'X5']].values
regressor = MLPRegressor(hidden_layer_sizes=(10,5), activation='relu',
                           solver='adam', alpha=1e-7, batch_size='auto', learning_rate='adaptive',
                           learning_rate_init=0.001, power_t=0.9, max_iter=1000000, shuffle=True,
                           random_state=None, tol=1e-4, verbose=False, warm_start=False,
                           momentum=0.9, nesterovs_momentum=True, early_stopping=False,
                           validation_fraction=0.1, epsilon=1e-08, n_iter_no_change=10, max_fun=15000)
#####
X_train, X_test, Y_train, Y_test = train_test_split(X, y, test_size=0.2,
                                                    random_state=0)
regressor = regressor.fit(X_train, Y_train)
SK_y = regressor.predict(X)
```

```
#####
#           MARS
#####
mars = Earth() #max_degree=(1),penalty=(1.0), endspan=(5)
mars.fit(X,y)
print(mars.trace())
print(mars.summary())
mars_y = mars.predict(X)
mars_resid = y-mars_y
```

Basándose en el patrón gráfico (primera gráfica de abajo, de izquierda a derecha) se puede apreciar como MARS tiene un mejor ajuste que las otras 2 técnicas analizadas. Evidenciándose en la tabla 3.3 donde se aprecia que esta técnica ajusta al 88 % de la varianza y presenta valores de error inferiores a las otras.

En la figura 3.4 de muestra la gráfica de valores observados frente a pronosticados para Y.

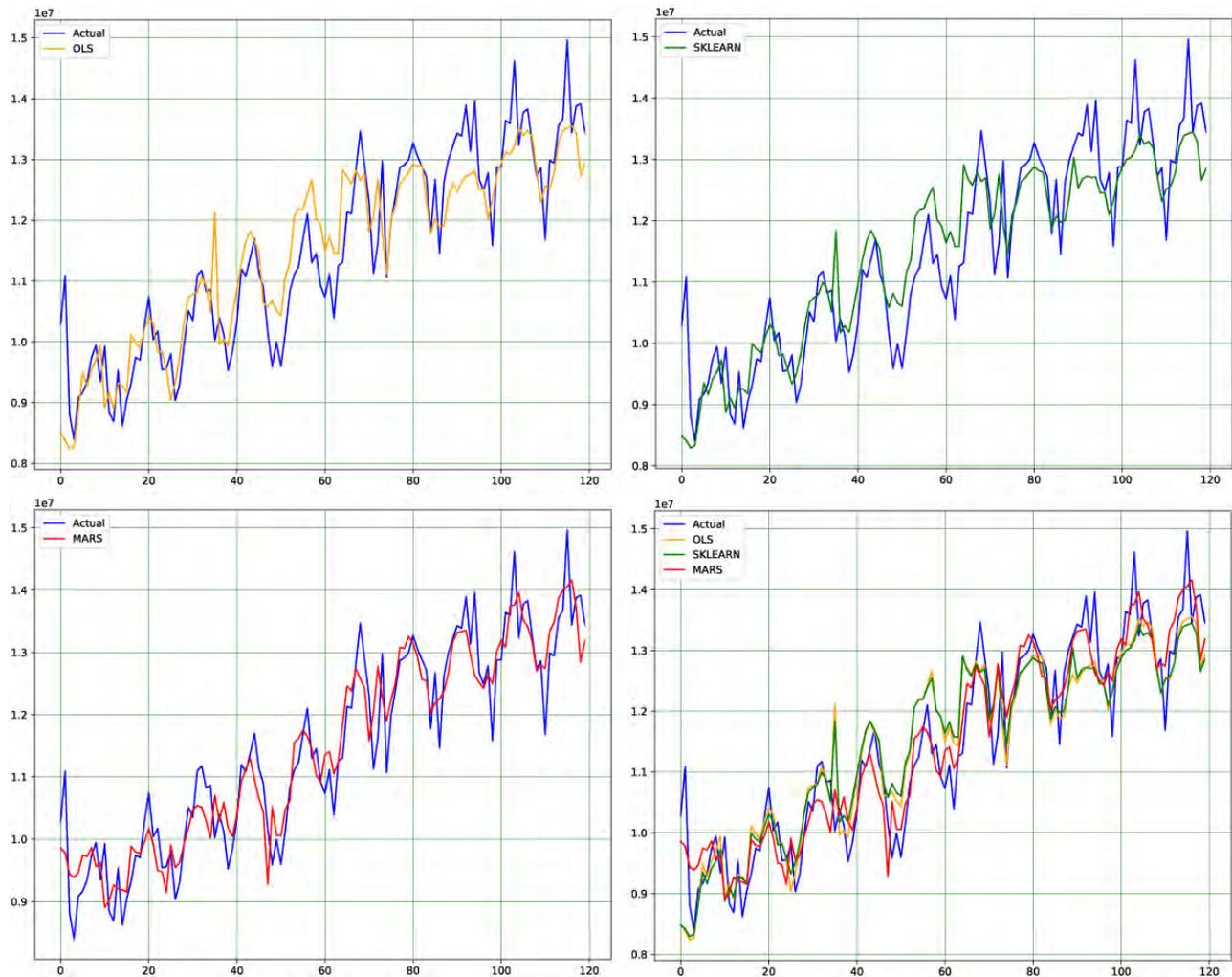


Figura 3.4: Observados vs Predichos municipio Cienfuegos. Fuente: Elaboración propia

En la tabla se compara el  $\bar{R}^2$  de cada modelo, así como, los errores. Se puede apreciar el ajuste de cada modelo y como las medidas del error son menores en el método *MARS*.

Tabla 3.3: Comparación de técnicas municipio Cienfuegos

	$\bar{R}^2$	MAE	MSE	RMSE	AD
MCO	0.810078	516850.683407	4.833628e+11	683557.409919	0.187731
ANN	0.826457	437050.968487	3.479232e+11	589850.186143	0.178678
MARS	0.877286	367273.057873	2.253185e+11	474677.297743	0.974001

En la figura 3.5 se aprecia que los residuos de MARS muestran mejor simetría respecto a 0, presentan una menor amplitud y curtosis.

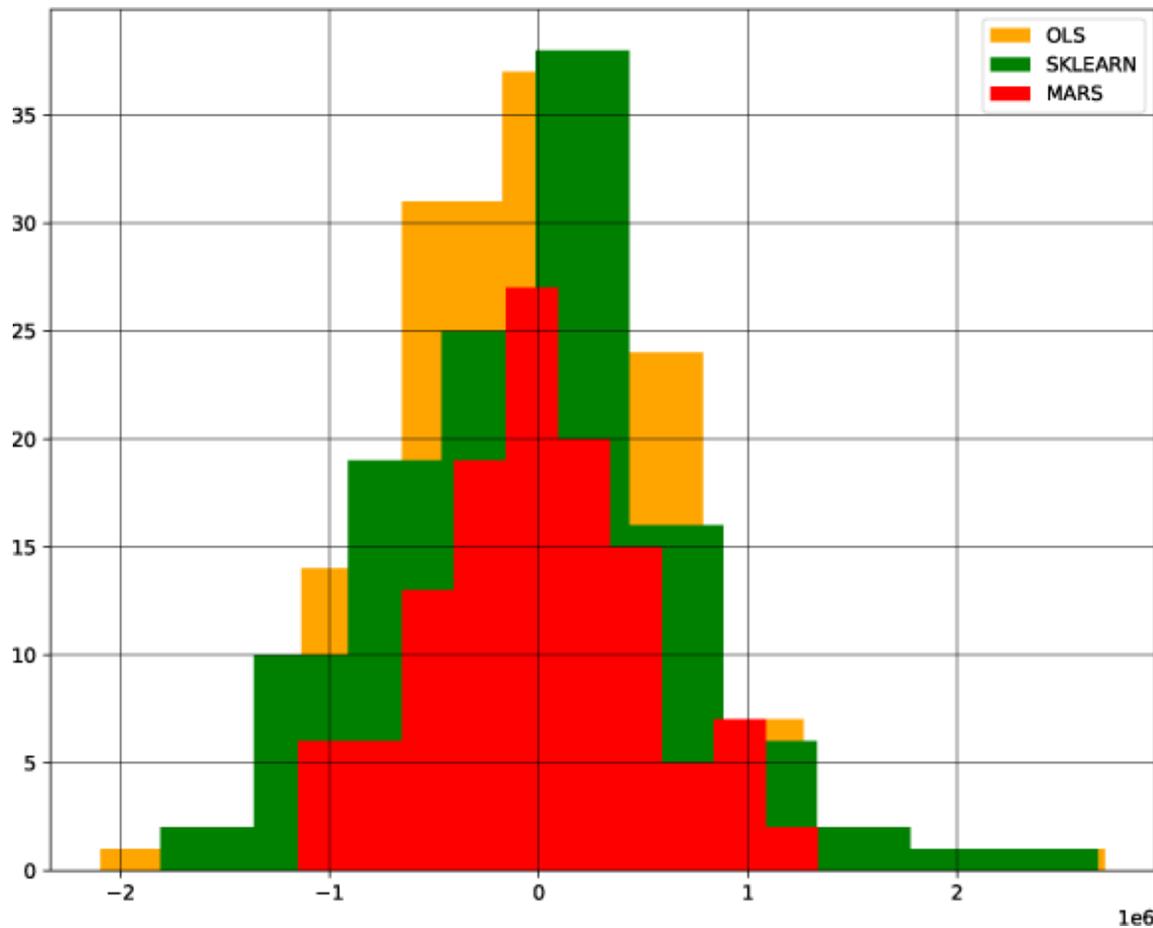


Figura 3.5: Histograma de los residuos, municipio Cienfuegos. Fuente: Elaboración propia

A pesar de no ser requisito de las últimas técnicas se verifica la normalidad de los residuos

para complementar el análisis de comportamiento de estas técnicas. En la tabla 3.3 se refleja el  $p$  - *value* de la prueba de normalidad AD.

La expresión matemática que caracteriza al modelo *MARS* queda definida como:

$$\begin{aligned}
 Y = & 1,20826 \cdot 10^{+7} + 1284,23 \cdot \max(0, X5 - 49836) - 87606 \cdot \max(0, 32,1333 - X2) + \dots \\
 & -154402 \cdot \max(0, X3 - 82,5484) - 65280,6 \cdot \max(0, 82,5484 - X3) + \dots \\
 & +4684,57 \cdot \max(0, 12701 - X4) - 869,598 \cdot \max(0, X5 - 50784) + \dots \\
 & -192,756 \cdot \max(0, 42742 - X5)
 \end{aligned}$$

A modo de ilustración de lo abordado en 2.3 respecto a las variaciones en la matriz de pesos que producen variaciones en la solución de ANN se muestran varias ejecuciones del código empleado en la figura 3.6, señalándose en rojo algunos comportamientos erráticos.

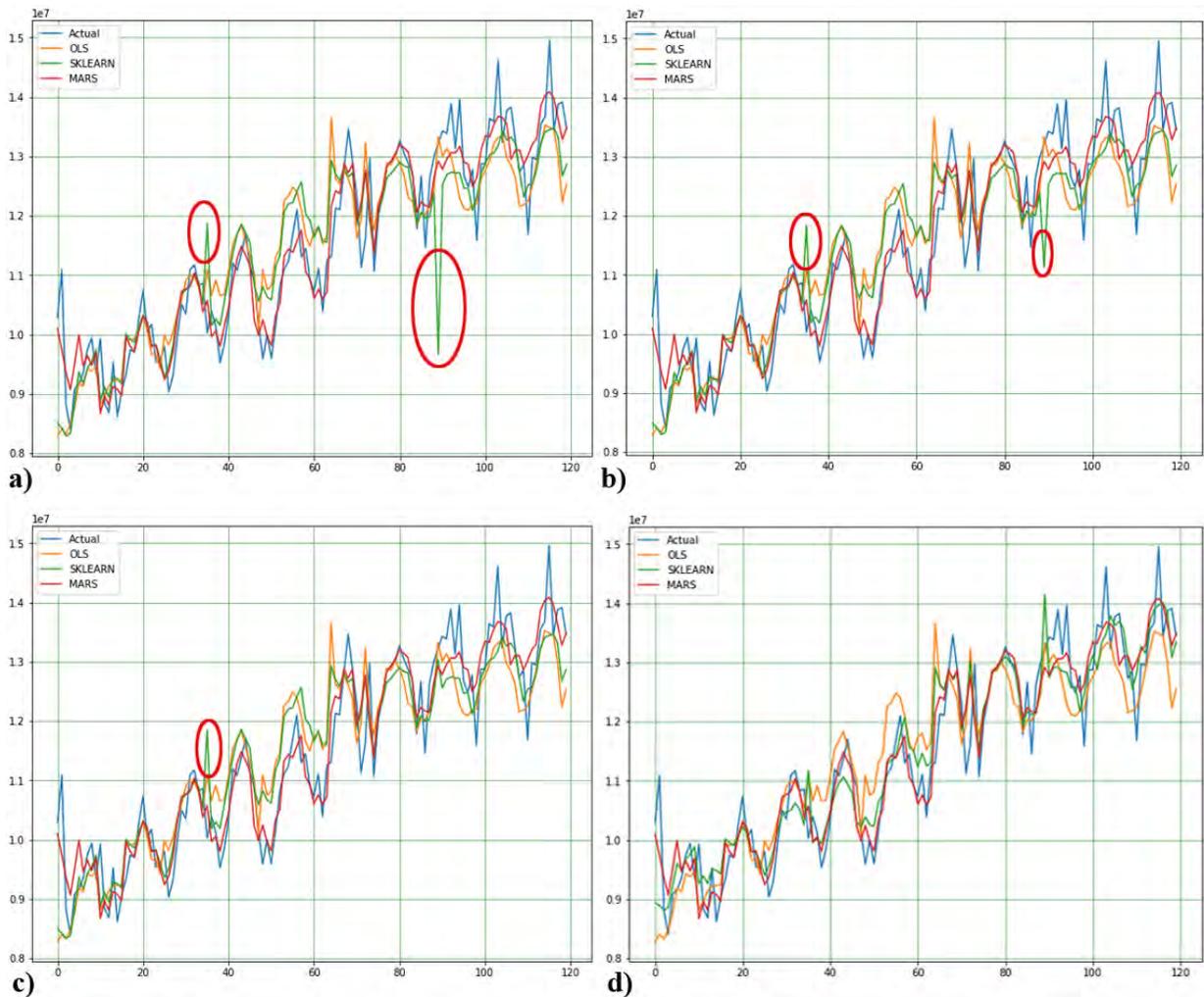


Figura 3.6: Aleatoriedad del pronóstico mediante ANN. Fuente: Elaboración propia

## 3.2. Consejos Populares representativos.

Para realizar el pronóstico de demanda eléctrica en el sector residencial se tomó de manera aleatoria un CP representativo de cada grupo formado a partir del análisis de cluster realizado anteriormente el cual se puede observar en la figura 2.5.

### 3.2.1. Rancho Luna

Se realiza el análisis de los modelos lineales en el CP, resultando aquel que relaciona las variables  $X_1, X_3, X_4, X_5$ , se hace notar que presenta autocorrelación en los residuos. se

Tabla 3.4: Modelos de RLM para el CP Rancho Luna

	$\bar{R}^2$	ANOVA	Prueba $t$	AD	E $\mu = 0$	BPG	BG	TOL
$Y \sim X_5$	0.734738	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_5$	0.790967	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_2 + X_5$	0.801195	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_3 + X_5$	0.830138	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_4 + X_5$	0.810378	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_2 + X_4 + X_5$	0.820791	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_3 + X_4 + X_5$	0.847047	SI	SI	SI	SI	SI	NO	SI

Del resumen de la tabla ANOVA se obtiene que este modelo es válido y sus coeficientes significativos individualmente, logrando explicar el 85 % de la variabilidad del consumo eléctrico con el modelo RLM.

Tabla 3.5: Coeficientes de la regresión CP

	Coef	std err	t	$P >  t $	[0.025	0.975]
const	-1.021e+05	2.23e+04	-4.583	0.000	-1.46e+05	-5.8e+04
X1	1201.3186	235.666	5.098	0.000	734.510	1668.128
X3	500.9021	93.323	5.367	0.000	316.048	685.756
X4	28.4627	7.655	3.718	0.000	13.299	43.626
X5	76.2052	4.998	15.247	0.000	66.305	86.105

La ecuación obtenida mediante MCO para el CP Rancho Luna se expresa matemáticamente como:

$$Y = -1,021 \cdot 10^{+5} + 1201,3186X_1 + 500,9021X_3 + 28,4627X_4 + 76,2052X_5.$$

*MARS* logra un ajuste de la variabilidad de la variable de respuesta de un 86% en la iteración 9 de su algoritmo y el menor error. El modelo explicativo se expone a continuación:

$$\begin{aligned} Y = & 64380,2 - 47,9656 \cdot \max(0, 858 - X_5) - 2886,29 \cdot \max(0, X_2 - 32,1333) + \dots \\ & - 978,602 \cdot \max(0, 32,1333 - X_2) - 399,341 \cdot \max(0, 82,5484 - X_3) + \dots \\ & + 209,496 \cdot \max(0, X_5 - 889) - 201,462 \cdot \max(0, X_5 - 969) + 16,9368 \cdot X_4 \end{aligned}$$

En la figura 3.7 de muestra la gráfica de valores observados frente a pronosticados para *Y*.

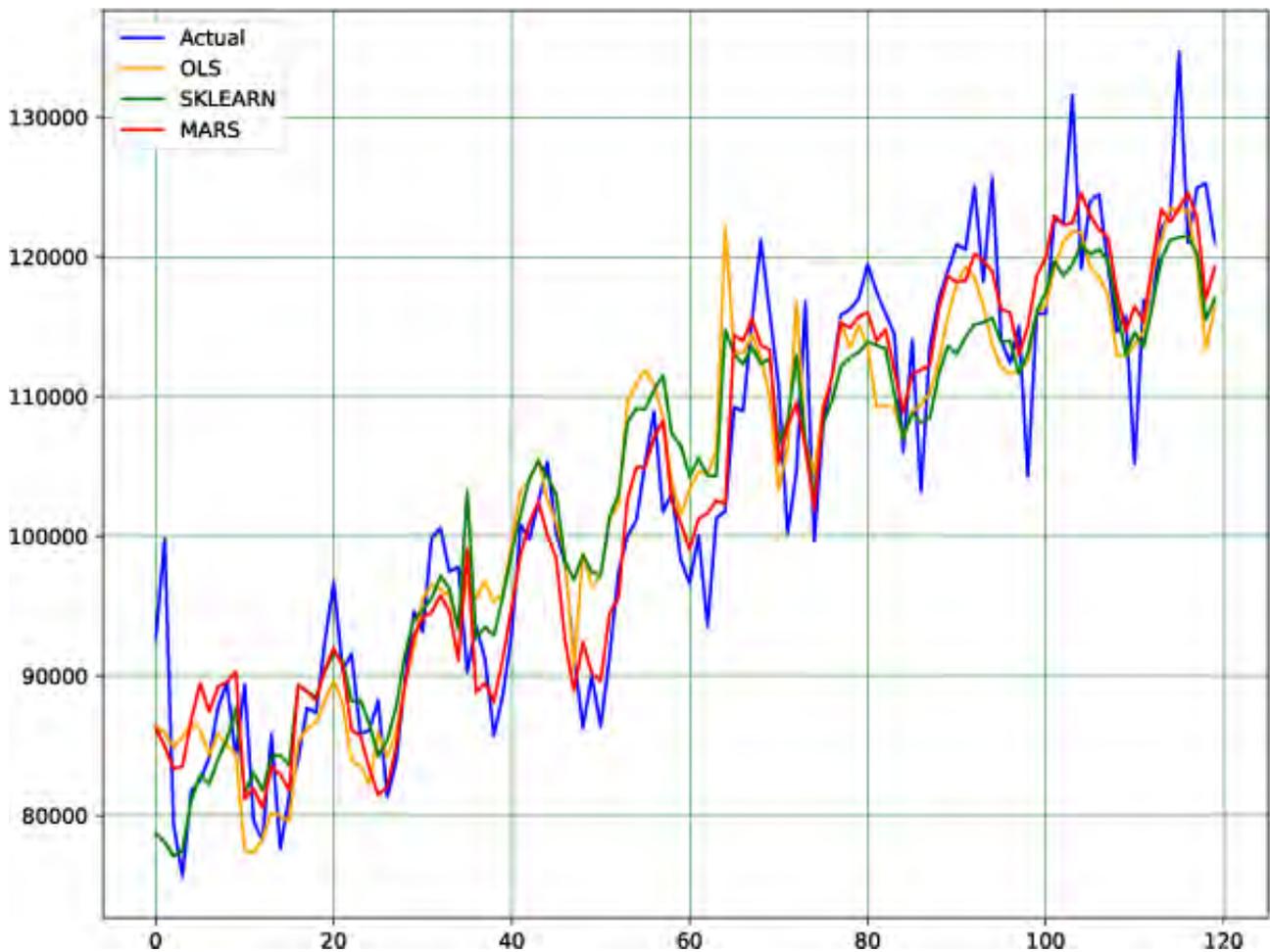


Figura 3.7: Observados vs Predichos CP Rancho Luna. Fuente: Elaboración propia

En la tabla 3.6 se realiza la comparación con las otras técnicas para mejorar el ajuste. En

esta ejecución ANN obtiene la menor explicación de la varianza y en adicción, un mayor error. Mientras que *MARS* posee una diferencia de  $10^7$  en la métrica *MSE*.

Tabla 3.6: Comparación de técnicas CP Rancho Luna

	$\bar{R}^2$	MAE	MSE	RMSE	AD
MCO	0.847047	4341.126579	3.153745e+07	5497.580379	0.673217
ANN	0.814893	4510.793140	3.382156e+07	5815.630319	0.526903
MARS	0.863803	3438.294005	2.026005e+07	4501.116014	0.268382

Se verifica la normalidad de los residuos de las técnicas cumpliéndose este supuesto para el RLM y siendo la misma distribución en los residuos de las restantes. En la gráfica 3.8 se aprecia como los residuos de *MARS* son los más simétricos y centrados.

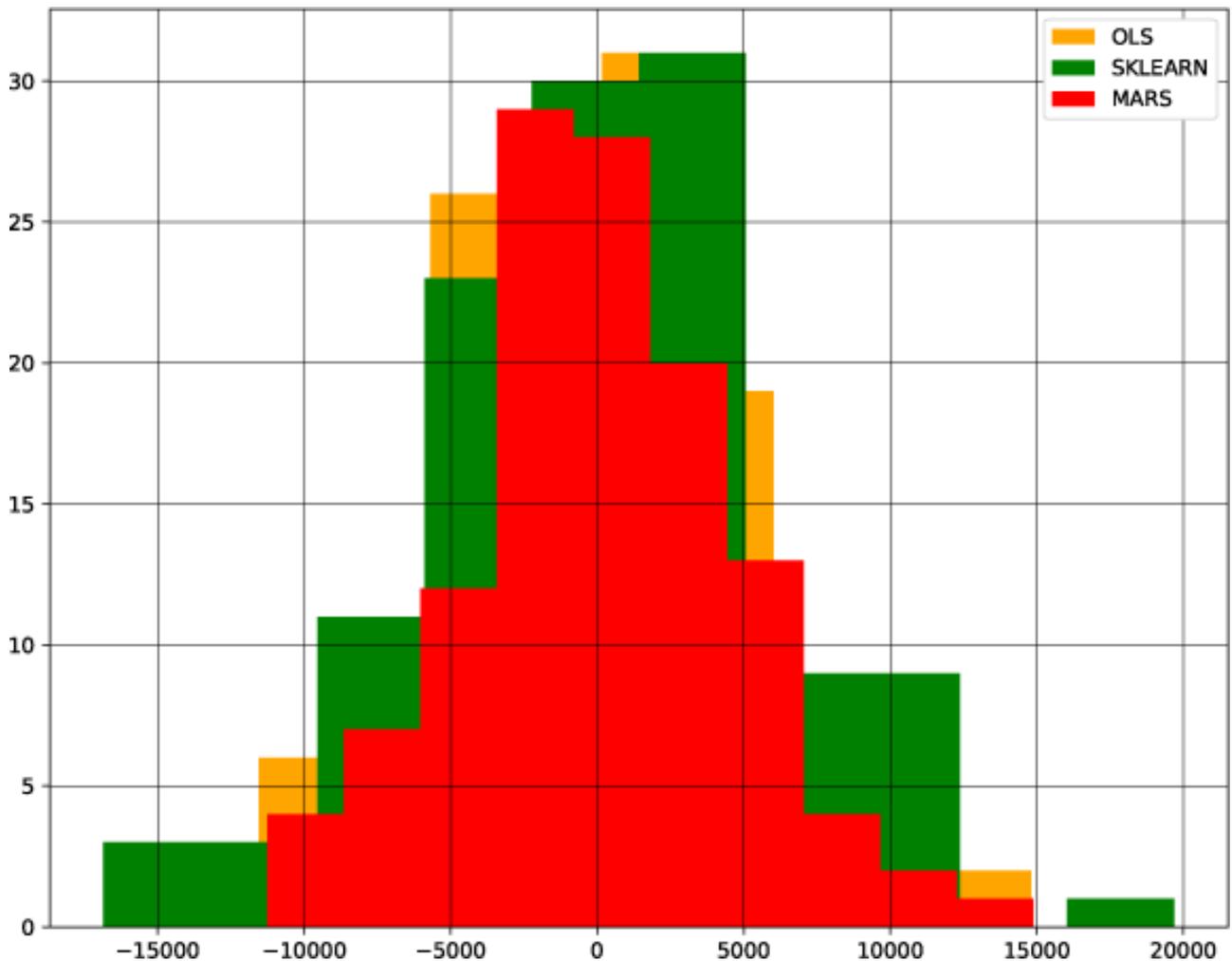


Figura 3.8: Histograma de los residuos, CP Rancho Luna. Fuente: Elaboración propia

### 3.2.2. Gloria

Se realiza el análisis de los modelos lineales en el CP, resultando aquel que relaciona las variables  $X1, X3, X5$ , se hace notar que presenta autocorrelación en los residuos.

Tabla 3.7: Modelos de RLM para el CP La Gloria

	$\bar{R}^2$	ANOVA	Prueba $t$	AD	$E \mu = 0$	BPG	BG	TOL
$Y \sim X5$	0.610027	SI	SI	SI	SI	SI	NO	SI
$Y \sim X1 + X5$	0.672401	SI	SI	SI	SI	SI	NO	SI
$Y \sim X2 + X5$	0.693348	SI	SI	SI	SI	SI	NO	SI
$Y \sim X1 + X3 + X4$	0.611911	SI	SI	SI	SI	SI	NO	SI
$Y \sim X1 + X3 + X5$	0.767025	SI	SI	SI	SI	SI	NO	SI

Del resumen de la tabla ANOVA se obtiene que este modelo es válido y sus coeficientes significativos individualmente, logrando explicar el 77 % de la variabilidad del consumo eléctrico con el modelo RLM.

Tabla 3.8: Coeficientes de la regresión CP Gloria

	Coef	std err	t	$P >  t $	[0.025	0.975]
const	-2.086e+05	4.24e+04	-4.920	0.000	-2.93e+05	-1.25e+05
X1	5401.2439	1411.471	3.827	0.000	2605.647	8196.841
X3	3649.8509	523.979	6.966	0.000	2612.044	4687.657
X5	115.3051	10.626	10.851	0.000	94.258	136.352

La ecuación obtenida mediante MCO para el CP se expresa matemáticamente como:

$$Y = -2,086 \cdot 10^{+5} + 5401,2439X1 + 3649,8509X3 + 115,3051X5.$$

*MARS* logra un ajuste de la variabilidad de la variable de respuesta de un 87 % en la iteración 11 de su algoritmo y el menor error. El modelo explicativo se expone a continuación:

$$\begin{aligned}
Y = & -3,69245 \cdot 10^{+7} + 189,795 \cdot \max(0, X5 - 2729) - 6567,99 \cdot X2 + \dots \\
& - 12271,4 \cdot \max(0, X2 - 32,1333) - 4205,8 \cdot \max(0, 32,1333 - X2) + \dots \\
& - 618,607 \cdot \max(0, 2953 - X5) - 494,58 \cdot \max(0, X5 - 2098) + 3703,8 \cdot X3 + \dots \\
& - 9492,86 \cdot \max(0, 82,5484 - X3) + 1153,7 \cdot \max(0, 10751 - X4) + \dots \\
& - 3226,06 \cdot \max(0, X4 - 10311) - 473,703 \cdot \max(0, X4 - 10885)
\end{aligned}$$

En la figura 3.9 de muestra la gráfica de valores observados frente a pronosticados para Y.

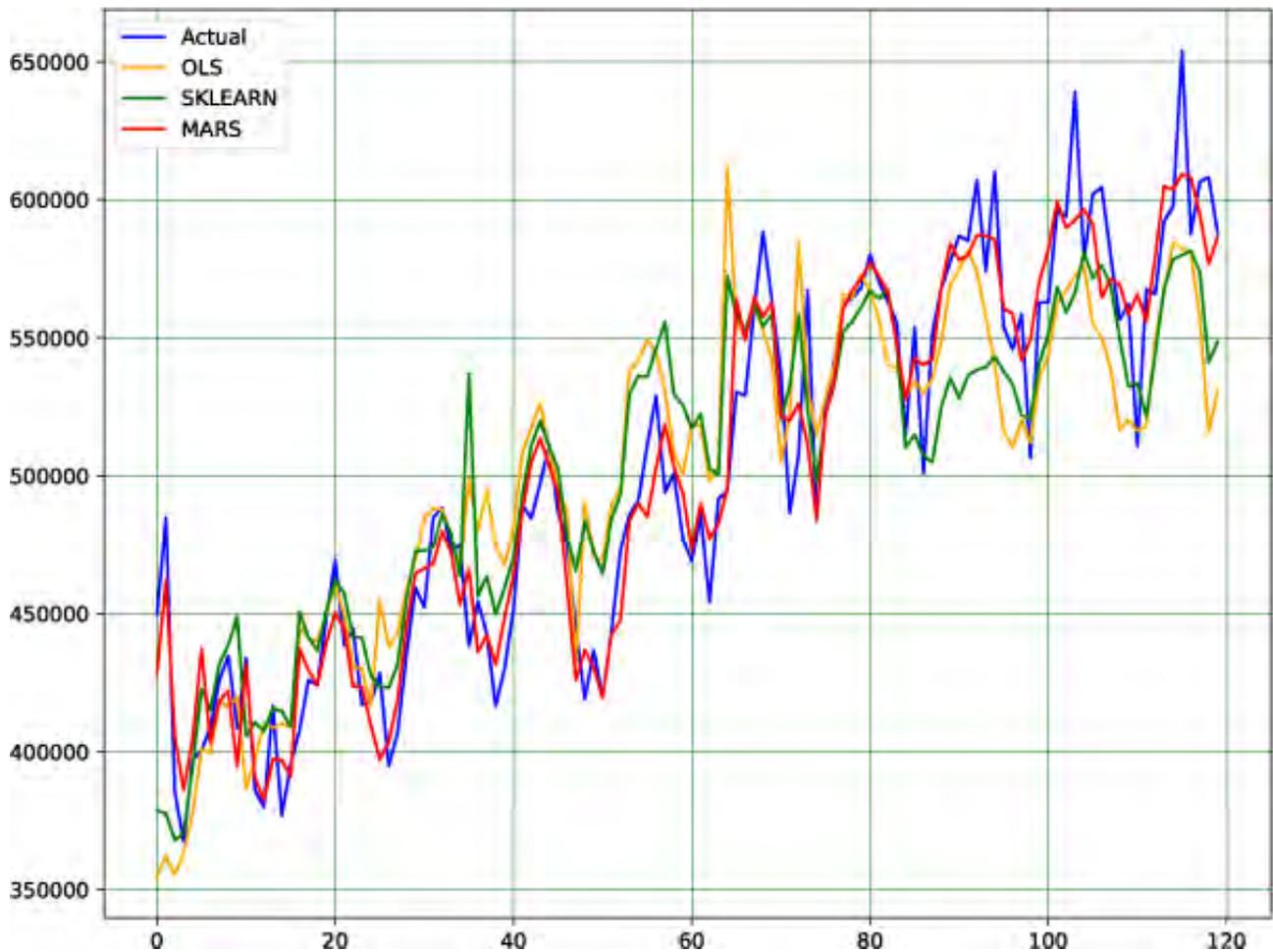


Figura 3.9: Observados vs Predichos CP La Gloria. Fuente: Elaboración propia

En la tabla 3.9 se realiza la comparación con las técnicas para mejorar el ajuste. En esta ejecución el pronóstico mediante ANN obtiene la menor explicación de la varianza, sin embargo, su  $MSE$  es inferior al MCO pero posee una diferencia de  $7,234718 \cdot 10^8$  con respecto a  $MARS$  evidenciando una vez más la robustez de esta última.

Tabla 3.9: Comparación de técnicas CP La Gloria

	$\bar{R}^2$	MAE	MSE	RMSE	AD
MCO	0.767025	24932.891141	1.132549e+09	33087.729306	0.137311
ANN	0.722907	25978.954158	1.062897e+09	32602.096372	0.861601
MARS	0.880908	14011.417620	3.394252e+08	18423.495430	0.398430

Se verifica la normalidad de los residuos de las técnicas cumpliéndose este supuesto para el RLM y siendo la misma distribución en los residuos de las restantes.

En la gráfica 3.10 se aprecia como los residuos de *MARS* son los más simétricos y centrados.

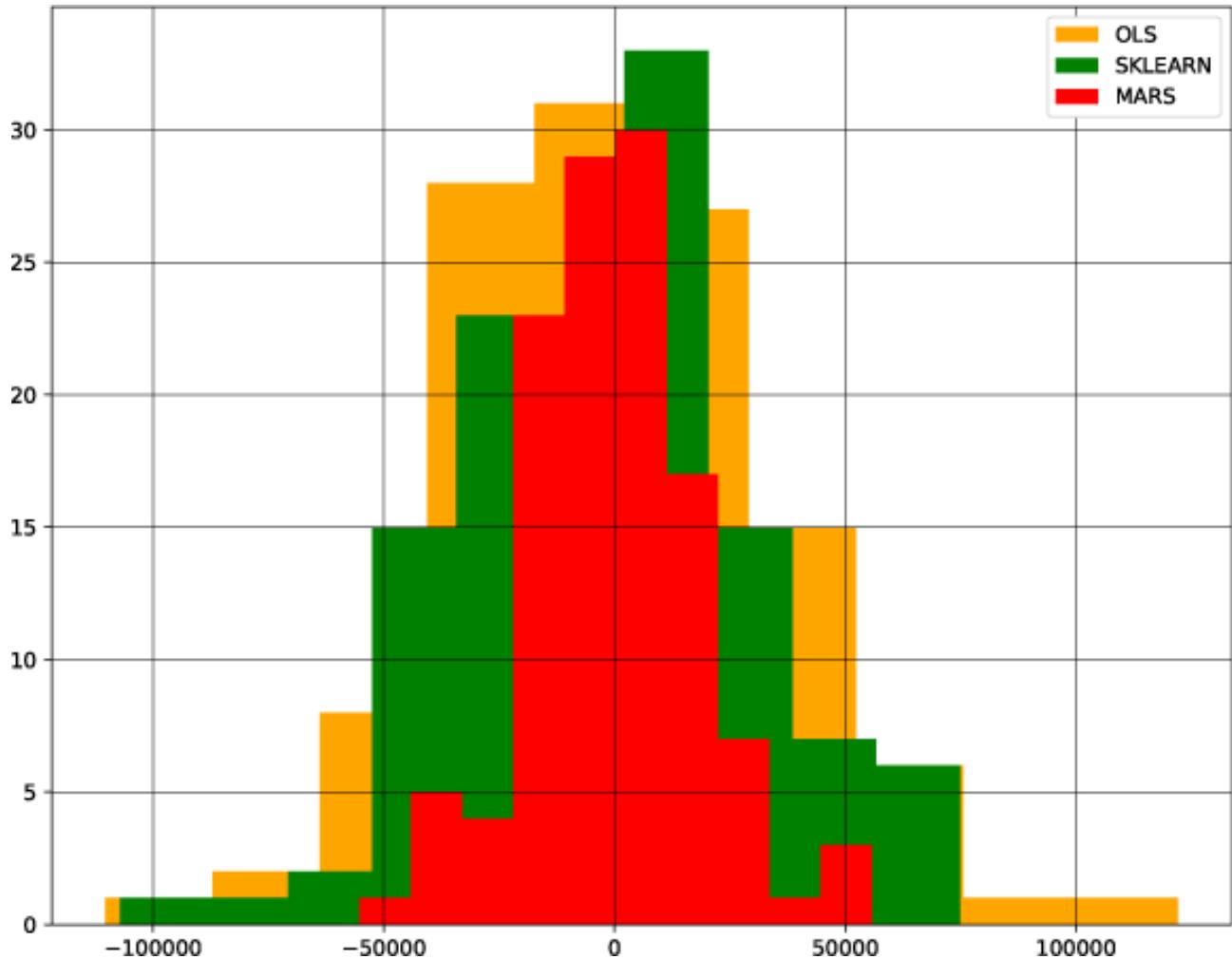


Figura 3.10: Histograma de los residuos, CP La Gloria. Fuente: Elaboración propia

### 3.2.3. Caunao

Se realiza el análisis de los modelos lineales en el CP, resultando aquel que relaciona las variables  $X_1, X_3, X_5$ , se hace notar que presenta autocorrelación en los residuos.

Tabla 3.10: Modelos de RLM para el CP Caunao

	$\bar{R}^2$	ANOVA	Prueba $t$	AD	E $\mu = 0$	BPG	BG	TOL
$Y \sim X_5$	0.725887	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_5$	0.778939	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_2 + X_4$	0.715478	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_2 + X_5$	0.790163	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_3 + X_4$	0.772439	SI	SI	SI	SI	SI	NO	SI
$Y \sim X_1 + X_3 + X_5$	0.822544	SI	SI	SI	SI	SI	NO	SI

Del resumen de la tabla ANOVA se obtiene que este modelo es válido y sus coeficientes significativos individualmente, logrando explicar el 83 % de la variabilidad del consumo eléctrico con el modelo RLM.

Tabla 3.11: Coeficientes de la regresión CP

	Coef	std err	t	$P >  t $	[0.025	0.975]
const	-1.941e+05	4.7e+04	-4.127	0.000	-2.87e+05	-1.01e+05
X1	7284.7342	1604.239	4.541	0.000	4107.336	1.05e+04
X3	3440.5614	630.800	5.454	0.000	2191.183	4689.939
X5	484.1069	35.041	13.816	0.000	414.705	553.509

La ecuación obtenida mediante MCO para el CP se expresa matemáticamente como:

$$Y = -1,941 \cdot 10^5 + 7284,7342X_1 + 3440,5614X_3 + 484,1069X_5.$$

*MARS* logra un ajuste de la variabilidad de la variable de respuesta de un 87 % en la iteración 17 de su algoritmo y los menores errores en las diferentes medidas evaluadas, destacándose que posee un *MSE* de un orden inferior en notación científica.

El modelo explicativo se expone a continuación:

$$Y = 1,22396 \cdot 10^{+6} + 1032,03 \cdot \max(0, X5 - 878) + \dots \\ - 20716 \cdot \max(0, 32,1333 - X2) - 237,913 \cdot \max(0, 980 - X5) - 19115,9X1$$

En la figura 3.11 se muestra la gráfica de valores observados frente a pronosticados para Y.

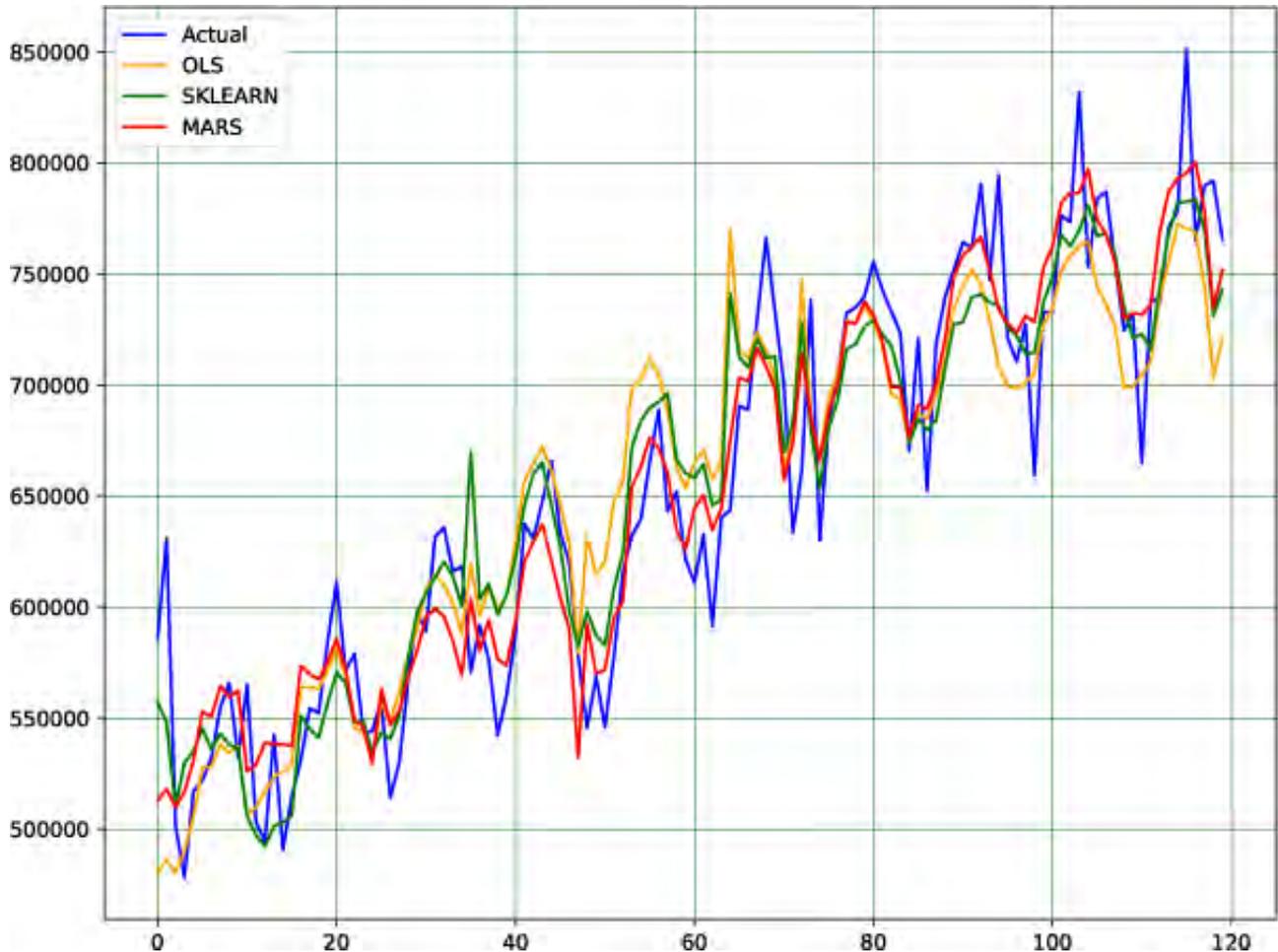


Figura 3.11: Observados vs Predichos CP Caunao. Fuente: Elaboración propia

En la tabla 3.12 se realiza la comparación con las técnicas para mejorar el ajuste. En esta ejecución el pronóstico mediante *MCO* obtiene la menor explicación de la varianza y en adicción, un mayor error. *ANN* no presenta diferencias significativas con respecto a este en cuanto al ajuste, sin embargo, sus errores son menores teniendo una diferencia de  $2,012097 \cdot 10^8$  con respecto a *MARS*.

Tabla 3.12: Comparación de técnicas CP Caunao

	$\bar{R}^2$	MAE	MSE	RMSE	AD
MCO	0.822544	28716.940763	1.462515e+09	37600.057328	0.208571
ANN	0.826057	26358.184394	1.124965e+09	33540.492100	0.286194
MARS	0.865283	23743.057694	9.237553e+08	30393.342830	0.349216

Se verifica la normalidad de los residuos de las técnicas cumpliéndose este supuesto para el RLM y siendo la misma distribución en los residuos de las restantes.

En la gráfica 3.12 se aprecia como los residuos de *MARS* son los más simétricos y centrados.

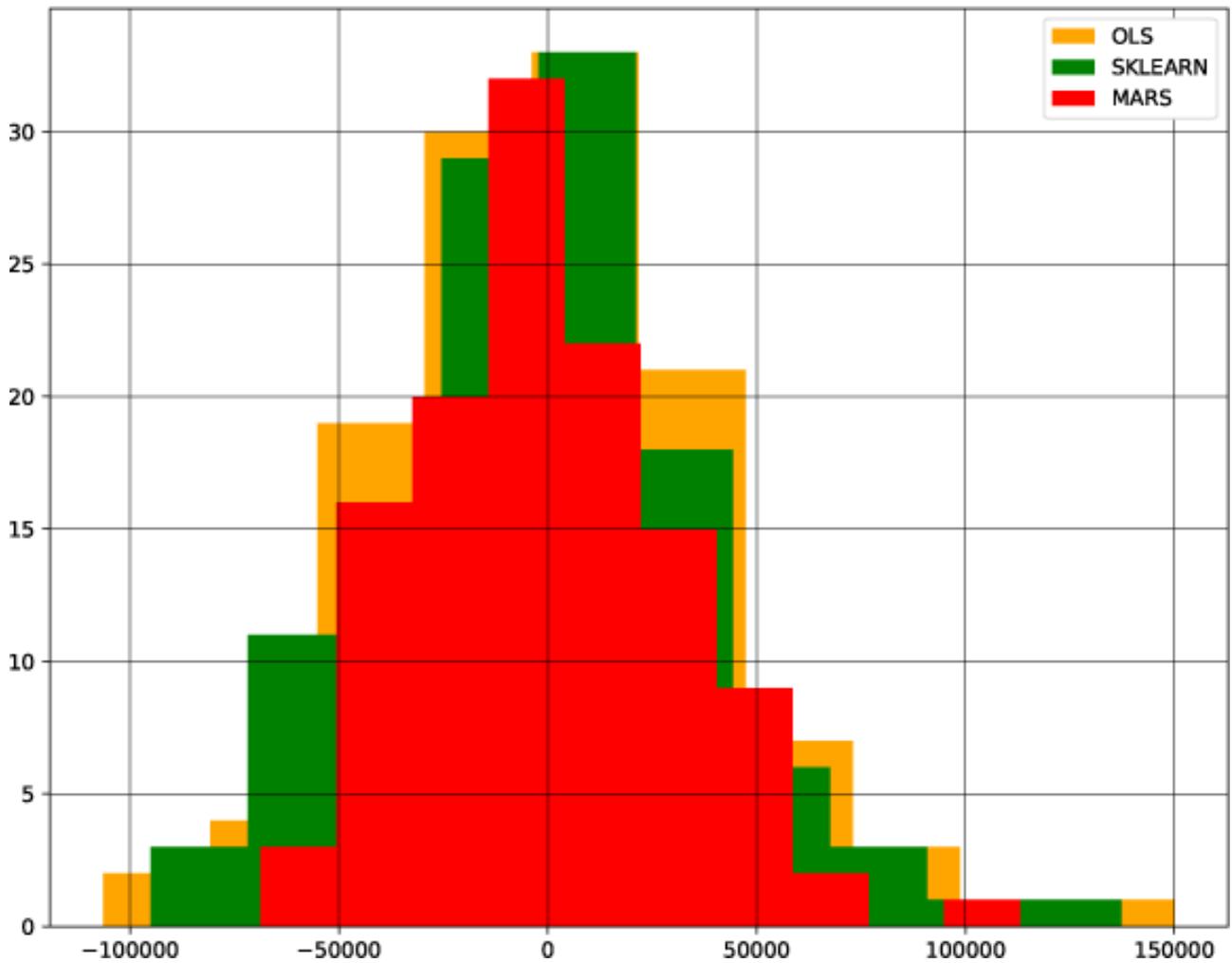


Figura 3.12: Histograma de los residuos, CP Caunao. Fuente: Elaboración propia

### 3.2.4. Tulipán

Se realiza el análisis de los modelos lineales en el CP, resultando aquel que relaciona las variables  $X1, X3, X4$  se hace notar que presenta autocorrelación en los residuos.

Como se puede apreciar en la tabla 3.13 este explica el 67% de la varianza, cumple los supuestos exceptuando la no autocorrelación de los residuos. Como se puede apreciar este valor de  $\bar{R}^2$  denota un ajuste moderado, lo que está asociado, lógicamente a valores altos de las medidas del error que se muestran en la tabla 3.15.

Tabla 3.13: Modelos de RLM para el CP Tulipán

	$\bar{R}^2$	ANOVA	Prueba $t$	AD	$E \mu = 0$	BPG	BG	TOL
$Y \sim X1 + X3 + X4$	0.671727	SI	SI	SI	SI	SI	NO	SI

Del resumen de la tabla ANOVA se obtiene que este modelo es válido y sus coeficientes significativos individualmente, logrando explicar el 67% de la variabilidad del consumo eléctrico con el modelo RLM.

Tabla 3.14: Coeficientes de la regresión CP Tulipán

	Coef	std err	t	$P >  t $	[0.025	0.975]
const	1.403e+06	2.55e+05	5.498	0.000	8.97e+05	1.91e+06
X1	1.31e+04	3784.736	3.461	0.001	5603.773	2.06e+04
X3	9973.2621	1385.533	7.198	0.000	7229.039	1.27e+04
X4	-148.7132	21.055	-7.063	0.000	-190.415	-107.011

En la tabla 3.15 se realiza la comparación con las técnicas para mejorar el ajuste. En esta ejecución el pronóstico mediante MCO obtiene la menor explicación de la varianza y en adicción, un mayor error. *ANN* logra explicar el 70% de la varianza; los errores de *MARS* con respecto a esta son menores en  $3,686397 \cdot 10^9$  siendo los más pequeños.

La ecuación obtenida mediante MCO para el CP se expresa matemáticamente como:

$$Y = 1,403 \cdot 10^6 + 1,31 \cdot 10^4 X1 + 9973,2621 X3 - 148,7132 X4.$$

En la figura 3.13 se muestra la gráfica de valores observados frente a pronosticados para  $Y$ . Se pueden observar algunos valores errados tanto para *MCO* como *ANN*. En todos los casos no se logran ajustar a los valores reales en los primeros 3 valores debido a su comportamiento brusco y

fuera del patrón lógico subsiguiente sin información de años precedentes no se puede determinar si este comportamiento es atípico o producido por error en la recolección de los datos.

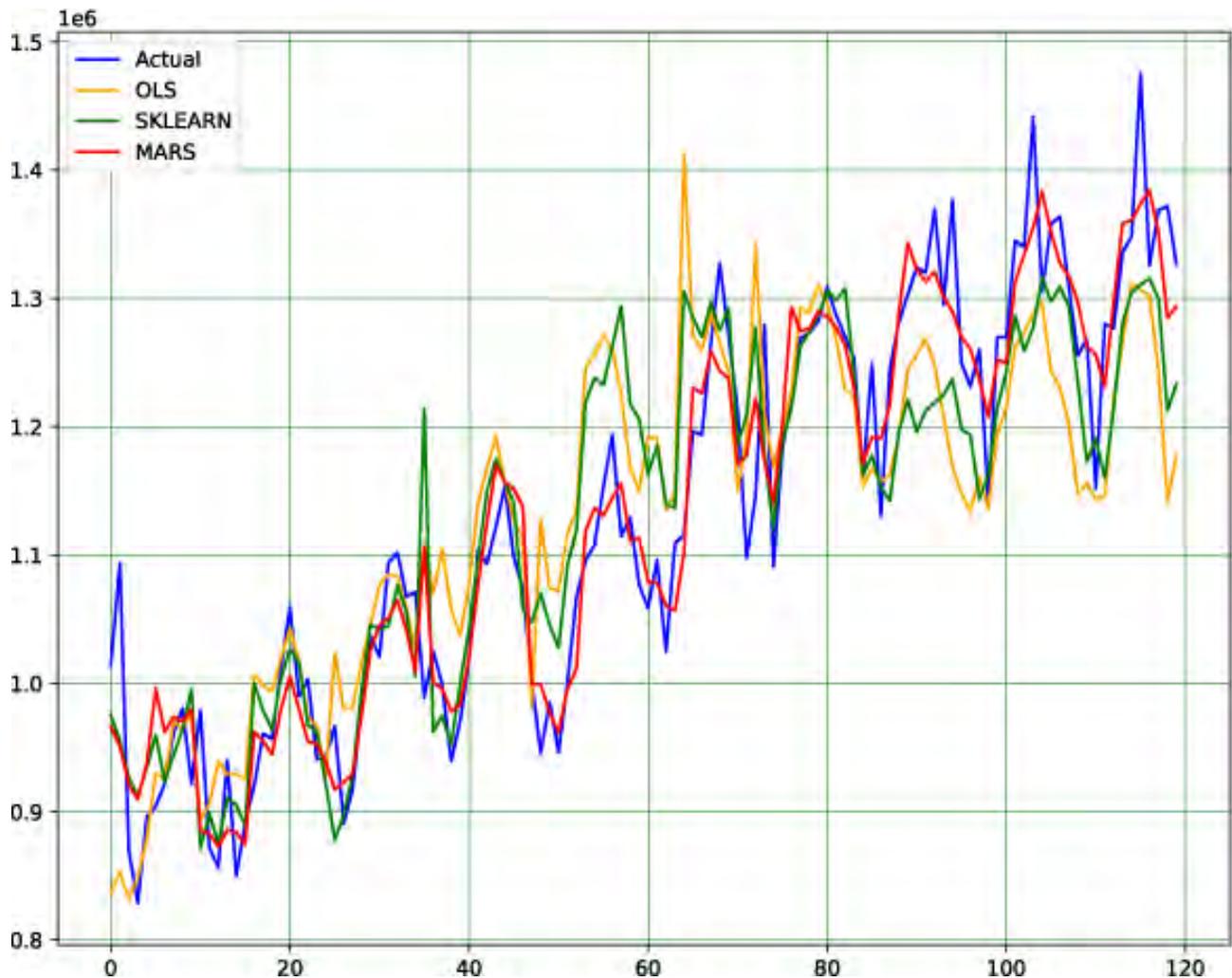


Figura 3.13: Observados vs Predichos CP Tulipán. Fuente: Elaboración propia

Tabla 3.15: Comparación de técnicas CP Tulipán

	$\bar{R}^2$	MAE	MSE	RMSE	AD
MCO	0.671727	72060.845778	8.124054e+09	88618.575993	0.290830
ANN	0.702858	59200.430337	5.799373e+09	76153.614792	0.731598
MARS	0.869073	35845.302570	2.112976e+09	45967.120938	0.699354

*MARS* logra un ajuste de la variabilidad de la variable de respuesta de un 87% en la iteración 10 de su algoritmo y el menor error.

El modelo explicativo se expone a continuación:

$$\begin{aligned}
 Y = & 950838 - 4083,12 \cdot \max(0, X4 - 8576) - 3905,96 \cdot \max(0, X4 - 8540) + \dots \\
 & - 381,327 \cdot \max(0, 8540 - X4) + 327,708 \cdot \max(0, X4 - 9246) + \dots \\
 & + 344,813 \cdot \max(0, X5 - 2531) + 16716 \cdot \max(0, X2 - 27) - 44161,1 \cdot \max(0, X2 - 32,1333)
 \end{aligned}$$

Se verifica la normalidad de los residuos de las técnicas cumpliéndose este supuesto para el RLM y siendo la misma distribución en los residuos de las restantes.

En la gráfica 3.14 se aprecia como los residuos de *MARS* son los más simétricos y centrados.

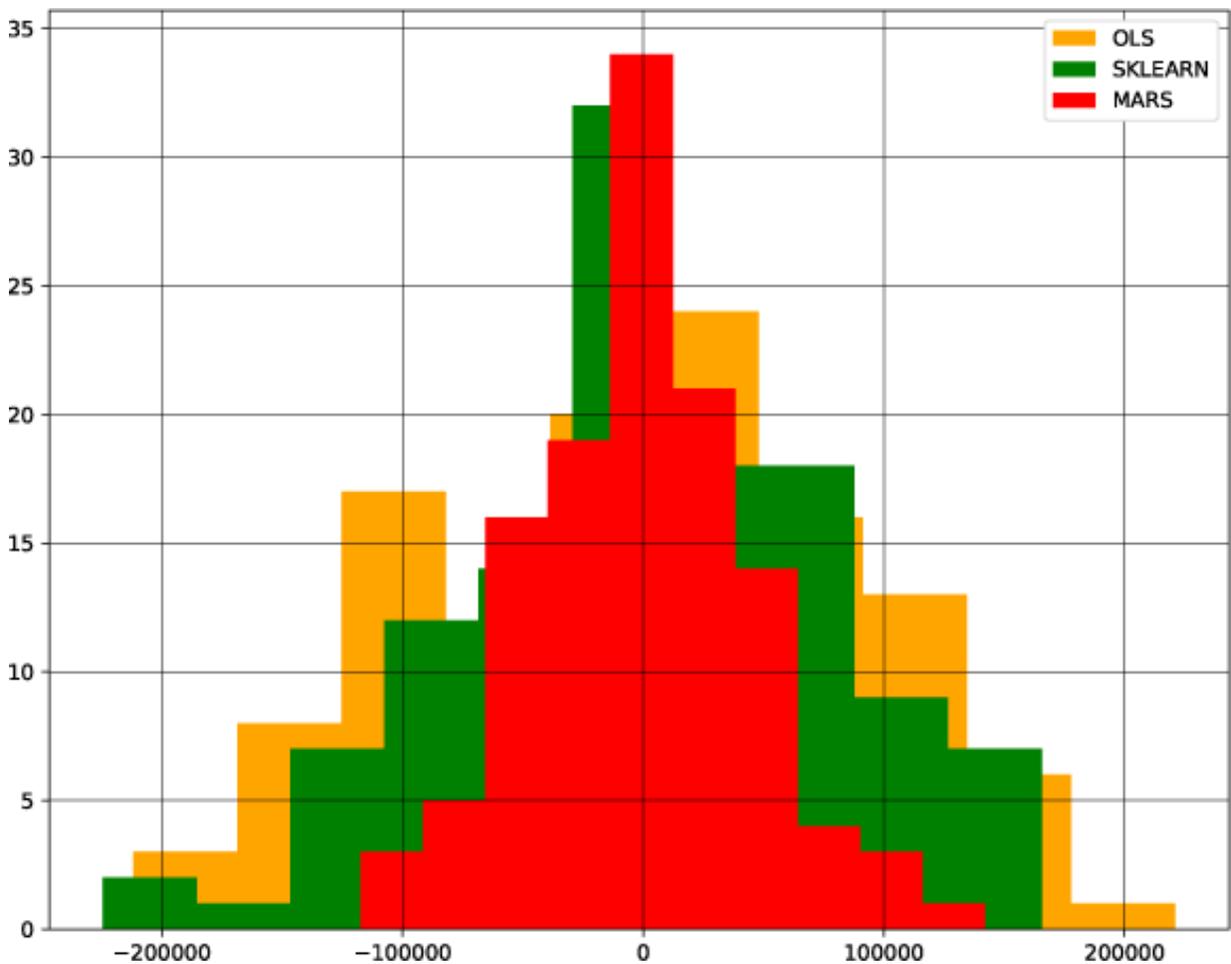


Figura 3.14: Histograma de los residuos, CP Tulipán. Fuente: Elaboración propia

### 3.2.5. Resumen Procesamiento de CP

A continuación se presenta un resumen que contiene los modelos de *MCO* y *MARS*, resultados del procesamiento de los restantes CP, así como, la varianza explicada por estos. Estos serán

ordenados atendiendo a su consumo de energía, de menor a mayor y al grupo al que pertenecen según el análisis de cluster realizado en el capítulo anterior.

■ **Grupo 4**

**La Barrera** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 78\%$

$$Y = -6,758 \cdot 10^{+4} + 1732,4765X1 + 1134,6059X3 + 39,1758X5.$$

MARS:  $\bar{R}^2 = 88\%$

$$\begin{aligned} Y = & 168532 - 102,773 \cdot \max(0, 8886 - X4) + 54,4779 \cdot \max(0, X5 - 2546) + \dots \\ & - 809,106 \cdot \max(0, 82,5484 - X3) - 36,4937 \cdot \max(0, 2843 - X5) + \dots \\ & + 130,427 \cdot \max(0, 1997 - X5) + 2437,05 \cdot \max(0, X2 - 27) + \dots \\ & - 6568,31 \cdot \max(0, X2 - 32,1333) \end{aligned}$$

**Guaos** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 83\%$

$$Y = -5,71 \cdot 10^{+4} + 2307,9953X1 + 1096,5400X3 + 154,6279X5.$$

MARS:  $\bar{R}^2 = 87\%$

$$\begin{aligned} Y = & 189855 - 1878,49 \cdot \max(0, 32,1333 - X2) - 247,193 \cdot \max(0, X5 - 795) + \dots \\ & + 576,839 \cdot \max(0, 596 - X5) - 854,389 \cdot \max(0, 82,5484 - X3) + \dots \\ & + 794,859 \cdot \max(0, X4 - 849) + 150,728 \cdot \max(0, X5 - 612) \end{aligned}$$

**Pepito Tey** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 86\%$

$$Y = 1,962 \cdot 10^{+5} + 2707,3405X1 + 964,2155X3 - 87,1156X4 + 293,5700X5.$$

MARS:  $\bar{R}^2 = 87\%$

$$\begin{aligned} Y = & -323382 + 1317,26 \cdot \max(0, 915 - X5) - 8372,47 \cdot \max(0, 32,1333 - X2) + \dots \\ & - 518,479 \cdot \max(0, X5 - 991) - 337,452 \cdot \max(0, X5 - 875) + \dots \\ & + 945,452 \cdot \max(0, X5 - 680) + 136,532X4 - 585,45 \cdot \max(0, 794 - X5) \dots \\ & - 8087,61X1 \end{aligned}$$

■ **Grupo 3**

**Reina** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 77\%$

$$Y = -1,321 \cdot 10^{+5} + 3750,0678X1 + 2401,0823X3 + 61,0416X5.$$

MARS:  $\bar{R}^2 = 87\%$

$$\begin{aligned}
Y &= 492801 + 91,4448 \cdot \max(0, X5 - 3457) + 171,464 \cdot \max(0, X4 - 10758) + \dots \\
&- 575,573 \cdot \max(0, 10758 - X4) - 5224,1 \cdot \max(0, X3 - 82,5484) + \dots \\
&- 1428,69 \cdot \max(0, 82,5484 - X3) - 8072,59 \cdot \max(0, 32,1333 - X2) + \dots \\
&+ 404,385 \cdot \max(0, 11052 - X4) - 7715,81X1
\end{aligned}$$

**Paraíso** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 77\%$

$$Y = -1,436 \cdot 10^{+5} + 5901,3991X1 + 3411,7575X3 + 22,8703X4.$$

MARS:  $\bar{R}^2 = 87\%$

$$\begin{aligned}
Y &= 312708 + 260,538 \cdot \max(0, X4 - 12203) + 4781,69X2 + \dots \\
&+ 92,3543 \cdot \max(0, X5 - 2069) + 261,597 \cdot \max(0, 2069 - X5) + \dots \\
&- 2295,75 \cdot \max(0, 82,5484 - X3) - 13343,3 \cdot \max(0, X2 - 32,1333)
\end{aligned}$$

**San Lázaro** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 81\%$

$$Y = -1,471 \cdot 10^{+5} + 4705,1884X1 + 2438,1859X3 + 100,0794X5.$$

MARS:  $\bar{R}^2 = 87\%$

$$\begin{aligned}
Y &= -117169 - 151,727 \cdot \max(0, 2934 - X5) - 3904,86 \cdot \max(0, 32,1333 - X2) + \dots \\
&- 1399,61 \cdot \max(0, 84,5484 - X3) + 266,37 \cdot \max(0, X5 - 2162) + \dots \\
&- 175,995 \cdot \max(0, 10786 - X4) + 352,496 \cdot \max(0, 11065 - X4) + \dots \\
&+ 249,363 \cdot \max(0, 9086 - X4) + 180,167 \cdot \max(0, X4 - 9186) + \dots \\
&- 1419,164 \cdot \max(0, 9186 - X4) - 242,542 \cdot \max(0, X4 - 11360)
\end{aligned}$$

**Pastorita-Obourque** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 77\%$

$$Y = -1,436 \cdot 10^{+5} + 5901,3991X1 + 3411,7575X3 + 22,8703X4.$$

MARS:  $\bar{R}^2 = 87\%$

$$\begin{aligned}
Y &= 312708 + 260,538 \cdot \max(0, X4 - 12203) + 4781,69X2 + \dots \\
&+ 92,3543 \cdot \max(0, X5 - 2069) + 261,597 \cdot \max(0, 2069 - X5) + \dots \\
&- 2295,75 \cdot \max(0, 82,5484 - X3) - 13343,3 \cdot \max(0, X2 - 32,133)
\end{aligned}$$

■ **Grupo 2**

**Buena Vista** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 76\%$

$$Y = -9462,8281 + 7061,7706X1 + 4891,8578X3 - 17,0944X4 + 109,8172X5.$$

$$\text{MARS: } \bar{R}^2 = 86 \%$$

$$\begin{aligned} Y = & -4,19563 \cdot 10^{+6} - 606,514 \cdot \max(0, X4 - 9632) - 3363,26 \cdot \max(0, 82,5484 - X3) + \dots \\ & + 527,65X4 + 2073,54 \cdot \max(0, X5 - 2724) - 181,12 \cdot \max(0, 2926 - X5) \dots \\ & - 2018,64 \cdot \max(0, X5 - 2668) - 5977,6 \cdot \max(0, 32,1333 - X2) \end{aligned}$$

**Caunao** Los modelos seleccionados para este CP se definen:

$$\text{MCO: } \bar{R}^2 = 82 \%$$

$$Y = -1,941 \cdot 10^{+5} + 7284,7342X1 + 3440,5614X3 + 484,1069X5.$$

$$\text{MARS: } \bar{R}^2 = 87 \%$$

$$\begin{aligned} Y = & 1,22396 \cdot 10^{+6} + 1032,03 \cdot \max(0, X5 - 878) - 20716 \cdot \max(0, 32,1333 - X2) + \dots \\ & - 237,913 \cdot \max(0, 980 - X5) - 19115,9X1 \end{aligned}$$

**Pueblo Grifo** Los modelos seleccionados para este CP se definen:

$$\text{MCO: } \bar{R}^2 = 82 \%$$

$$Y = 3,849 \cdot 10^{+5} + 8816,2738X1 + 2769,3455X3 - 29,0910X4 + 58,1394X5.$$

$$\text{MARS: } \bar{R}^2 = 86 \%$$

$$\begin{aligned} Y = & 768474 - 251,524 \cdot \max(0, X5 - 3529) - 6180,23 \cdot \max(0, 32,1333 - X2) + \dots \\ & + 235,506 \cdot \max(0, X5 - 2995) - 108,234 \cdot \max(0, 2995 - X5) + \dots \\ & - 3382,59 \cdot \max(0, 82,5484 - X3) - 252,823 \cdot \max(0, 8966 - X4) \end{aligned}$$

**Castillo** Los modelos seleccionados para este CP se definen:

MCO: Ningún modelo cumple aquellos supuestos que no deben violarse.

$$\text{MARS: } \bar{R}^2 = 87 \%$$

$$\begin{aligned} Y = & 808703 - 3465,98 \cdot \max(0, 82,5484 - X3) - 5815,22 \cdot \max(0, 32,1333 - X2) + \dots \\ & + 569,555 \cdot \max(0, X4 - 8913) - 70,4497 \cdot \max(0, 4427 - X5) \end{aligned}$$

**Junco Sur** Los modelos seleccionados para este CP se definen:

$$\text{MCO: } \bar{R}^2 = 80 \%$$

$$Y = -2,274 \cdot 10^{+5} + 8465,3682X1 + 4621,5179X3 + 121,4278X5.$$

$$\text{MARS: } \bar{R}^2 = 88 \%$$

$$\begin{aligned} Y = & 2,74058 \cdot 10^{+6} + 199,372 \cdot \max(0, X5 - 2772) + 323,412 \cdot \max(0, 2772 - X5) + \dots \\ & - 10855,7 \cdot \max(0, 82,5484 - X3) - 561,366 \cdot \max(0, X4 - 7024) + \dots \\ & - 7298,86 \cdot \max(0, X3 - 63,9333) + 280,493 \cdot \max(0, 10129 - X4) + \dots \\ & - 840,169 \cdot \max(0, 10386 - X4) \end{aligned}$$

**Juanita II** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 84\%$

$$Y = -2,292 \cdot 10^{+5} + 9001,1488X1 + 3687,3255X3 + 186,4267X5.$$

MARS:  $\bar{R}^2 = 88\%$

$$\begin{aligned} Y = & 950469 - 25814,2 \cdot \max(0, 32,1333 - X2) - 690,095 \cdot \max(0, X5 - 3196) + \dots \\ & + 821,788 \cdot \max(0, X5 - 2192) - 349,277 \cdot \max(0, X5 - 2435) + \dots \\ & - 25189,1X1 - 12350,9 \cdot \max(0, 63,9333 - X3) + 422,178 \cdot \max(0, 3256 - X5) \end{aligned}$$

■ **Grupo 1**

**Juanita I** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 84\%$

$$Y = -2,635 \cdot 10^{+5} + 1,076 \cdot 10^{+4}X1 + 4344,8258X3 + 213,2731X5.$$

MARS:  $\bar{R}^2 = 88\%$

$$\begin{aligned} Y = & 1,15849 \cdot 10^{+6} + 293,865 \cdot \max(0, 2976 - X5) - 8941,18 \cdot \max(0, 32,1333 - X2) + \dots \\ & + 1597,06 \cdot \max(0, 12568 - X4) - 4309,33 \cdot \max(0, 82,5484 - X3) + \dots \\ & - 453,562 \cdot \max(0, 3281 - X5) - 811,153 \cdot \max(0, 12649 - X4) \end{aligned}$$

**Punta Gorda** Los modelos seleccionados para este CP se definen:

MCO: Ningún modelo cumple aquellos supuestos que no deben violarse.

MARS:  $\bar{R}^2 = 83\%$

$$\begin{aligned} Y = & 1,57599 \cdot 10^{+6} - 700,289 \cdot \max(0, X5 - 2535) + 693,555 \cdot \max(0, X5 - 2241) + \dots \\ & - 28817 \cdot \max(0, 32,1333 - X2) + 1717,15 \cdot \max(0, X4 - 9030) - 26539,8X1 + \dots \\ & - 16379,9 \cdot \max(0, 63,9333 - X3) - 1714,01 \cdot \max(0, X4 - 9135) + \dots \\ & + 439,973 \cdot \max(0, 9135 - X4) \end{aligned}$$

**Centro Histórico** Los modelos seleccionados para este CP se definen:

MCO:  $\bar{R}^2 = 85\%$

$$Y = -6,883 \cdot 10^{+6} + 1,421 \cdot 10^{+4}X1 + 6259,3693X3 + 766,5879X4.$$

MARS:  $\bar{R}^2 = 86\%$

$$\begin{aligned} Y = & 1,14479 \cdot 10^{+6} - 5789,89 \cdot \max(0, 82,5484 - X3) + 215,274 \cdot \max(0, X5 - 3815) + \dots \\ & - 10359,6 \cdot \max(0, 32,1333 - X2) + 1231,77 \cdot \max(0, 9456 - X4) + \dots \\ & + 1049,19 \cdot \max(0, X4 - 9377) - 2036,84 \cdot \max(0, 9377 - X4) \end{aligned}$$

Para la planificación del consumo de la energía eléctrica en el sector residencial se realiza el pronóstico para el año 2022, se utilizan para ello los modelos obtenidos.

Para las variables  $X_2$  y  $X_3$  se realiza determinación del año meteorológico típico según Yang, Wan, Li, y Lam (2011), dando como resultado los meses meteorológicos típicos que lo integran que se muestran en la tabla 3.16. En cuanto a la variable  $X_4$  se considera la tasa anual de crecimiento poblacional del municipio que representa 2,8 para el 2018, según la OMEI (ONEI Cienfuegos, 2019), para la variable  $X_5$  se constante debido a que la adquisición de estos medios de cocción es en su mayoría de reposición.

Tabla 3.16: Meses meteorológicos típicos del año meteorológico típico para el municipio de Cienfuegos. Fuente: elaboración propia.

TMM	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Años	1992	1991	2009	2006	2015	2011	2000	1999	2008	2017	2019	1999
(1990-2020)												

En la tabla 3.17 se muestra el pronóstico para el municipio de Cienfuegos y sus 19 CP para el año 2022 constituyendo una herramienta a disposición del gobierno local para la planificación del consumo de energía eléctrica en el sector residencial.

La aplicación de los modelos pronósticos obtenidos para el municipio de Cienfuegos y sus 19 CP ha permitido al gobierno local tomar decisiones relacionadas con las dimensiones siguientes:

**Económicas:** conocer la demanda de energía eléctrica del sector residencial permitiendo incorporar al plan de la economía del municipio acorde con sus necesidades de funcionamiento y desarrollo. También estos resultados permiten una mejor distribución de este recurso en conciliación con el nivel central, establecer la prioridad por CP.

**Sociales:** el establecimiento de un equilibrio entre consumo de energía eléctrica y calidad de vida por CP y del municipio.

**Gubernamentales:** el uso de los resultados por parte del Consejo energético municipal, en cuanto a la toma de decisiones en las asignaciones de energía eléctrica en el sector residencial tanto desde la planificación como en contingencia energética, la priorización de sectores con una incidencia en el Desarrollo Local y la calidad de vida de la población, reflejado en el Plan de la Economía del municipio. Por otra parte, el perfeccionamiento de las estructuras de gobierno tales como: Consejo energético municipal, Creación del Polo

Científico de la Provincia de Cienfuegos con el frente “Desarrollo Energético Sostenible” y Capacitación de los delegados a la Asamblea Municipal del Poder Popular y los presidentes de los CP.

Tabla 3.17: Pronóstico del consumo de energía eléctrica para el sector residencial municipal de Cienfuegos, año 2022. Fuente: elaboración propia.

Consejo Popular	Mes											
	ene.	feb.	mar.	abr.	may.	jun.	jul.	ago.	sep.	oct.	nov.	dic.
Caonao	729,81	732,28	731,73	735,99	769,22	787,23	792,76	795,65	800,42	778,59	734,39	751,55
La Gloria	569,18	557,98	565,48	555,29	578,67	604,65	603,51	609,07	607,94	594,57	576,32	586,08
Tulipán	1296,89	1262,85	1256,27	1231,73	1297,31	1358,69	1360,92	1373,90	1384,07	1351,55	1284,78	1293,98
Rancho Luna	117,45	114,67	116,41	115,46	119,47	123,39	122,47	123,48	124,55	122,83	116,94	119,25
Buena Vista	728,01	705,95	717,09	705,24	733,59	767,54	777,49	781,51	781,36	766,54	726,18	739,22
Castillo-Cen	864,69	836,88	851,30	844,13	875,77	910,92	915,11	920,21	915,72	909,43	865,42	874,37
Centro Histórico	1333,99	1303,71	1327,48	1315,51	1373,30	1433,22	1445,22	1446,15	1440,28	1424,34	1341,05	1355,96
Guaos	233,74	228,00	231,85	229,80	237,33	245,55	248,24	249,16	249,29	244,93	233,18	237,51
Juanita II	886,34	891,02	887,39	890,19	929,54	951,77	957,34	960,26	965,91	941,17	888,45	908,06
Junco Sur	870,56	855,80	866,66	855,60	885,60	918,27	928,07	931,16	925,96	905,27	878,27	891,48
La Barrera	182,62	178,00	177,23	173,95	183,57	192,42	192,50	194,24	195,58	190,94	181,27	182,84
La Juanita	1086,35	1056,97	1073,78	1061,99	1101,91	1142,68	1155,16	1158,49	1158,49	1137,11	1080,23	1100,14
Paraiso	415,86	418,04	418,57	420,72	439,38	449,75	446,04	450,47	455,28	447,11	425,15	434,82
Pastorita-O´Bourque	571,53	557,79	568,96	578,00	590,64	600,07	599,66	605,38	607,27	595,91	567,01	580,72
Pepito Tey	272,38	274,43	268,60	268,53	279,50	290,32	287,89	294,27	298,22	295,66	276,41	278,47
Pueblo Griffó	823,30	800,60	811,68	801,12	830,25	861,07	870,18	872,74	872,66	857,82	816,87	830,50
Punta Gorda	1139,42	1144,23	1143,59	1141,42	1193,00	1218,07	1224,36	1226,90	1232,06	1201,60	1138,56	1162,50
Reina	393,17	390,19	383,84	374,26	396,16	416,74	418,11	414,78	414,19	404,46	393,53	399,81
San Lázaro	478,83	469,63	478,08	476,29	492,65	507,68	513,28	514,60	515,12	506,29	483,91	492,99
<b>Municipio Cienfuegos</b>	<b>13110,4</b>	<b>12868</b>	<b>13023,4</b>	<b>13207,8</b>	<b>13293,1</b>	<b>13863,4</b>	<b>14034</b>	<b>14088,6</b>	<b>13979,4</b>	<b>13639,6</b>	<b>13283,4</b>	<b>13484,8</b>

### 3.3. Conclusiones parciales:

- Se presenta una implementación en *python* de *RLM*, *ANN* y *MARS*.
- Se realiza una comparación entre los diferentes modelos obtenidos mediante RLM que cumplen los supuestos de la regresión. Para los CP Punta Gorda y Castillo no se presentan modelos que cumplen los supuestos.
- La técnica *MARS* ofrece los mejores resultados en cuanto a error, ajuste y proporciona un modelo que permite realizar pronósticos confiables.
- Se realiza el pronóstico de la demanda energética del sector residencial de los 19 CP y del municipio.

---

## Conclusiones generales:

---

- La planeación energética como elemento de la GEL requiere de pronósticos que pueden ser realizados mediante técnicas estadísticas a través de herramientas computacionales.
- En las investigaciones precedentes no se consideran el análisis de los supuestos y se utilizan software privativos; consecuencia de ello, se pueden invalidar los resultados obtenidos y el pago de licencias.
- Se ofrece una implementación en *Python* que facilita la integración con el producto GEM. Se realiza una comparación entre las técnicas utilizadas y se ofrece el modelo de mejor ajuste para realizar pronósticos de energía eléctrica.
- El sistema de toma de decisiones del gobierno local en cuanto a la demanda de energía eléctrica para el sector residencial, se basa en la utilización de técnicas estadístico-matemáticas como MLR mediante MCO, ANN y MARS soportadas en *Python* (lenguaje de código abierto).
- La creación de un modelo integral mediante la utilización de enfoques innovadores para la planificación de la energía mediante la captación y procesamiento de datos que permita a los decisores gubernamentales trazar y planificar las estrategias y políticas energéticas en función del desarrollo local con un menor impacto ambiental.

---

---

## Recomendaciones:

---

Las recomendaciones que se realizan en este trabajo están orientadas a propuestas de trabajo futuro.

1. Reducir la unidad de análisis.
2. Capacitación de los decisores en el uso del producto GEM
3. Continuar con la puesta en marcha, entrenamiento e inserción al portal del ciudadano del municipio de Cienfuegos del Producto GEM.
4. Continuar con la divulgación de los resultados obtenidos en esta investigación, a través de presentaciones en eventos científicos, artículos, libros, talleres, cursos de capacitación y/o formación, con el objetivo de extender estos resultados y aplicaciones en otros municipios del país.

---

# REFERENCIAS BIBLIOGRÁFICAS.

---

- Alvarez Basanta, L. D. (2020). *Procedimiento para selección de modelos de regresión múltiple basada en métodos multicriterios*. Tesis de maestría, Universidad de Cienfuegos.
- Armas, M. R. Blanco de, & Santana Cruz, F. E. (2017). *Diseño de indicadores energéticos para el sector residencial en el municipio de Cienfuegos*. Tesis de grado, Universidad de Cienfuegos.
- Azad, H. B., Mekhilef, S., & Ganapathy, V. G. (2014). Long-term wind speed forecasting and general pattern recognition using neural networks. *IEEE Transactions on Sustainable Energy*, 5(2), 546–553.
- Barrachina Gomez, M., Cerrolaza Asenjo, J., Garcia Alonso, J., Iranzo Martin, J., Lopez Perez, B., Minguez Perres, E., Minguez Torres, E., Pascualena Cambra, M., Poza Galiano, A., & Secades Ariz, I. (1993). 222 questions about the energy.
- Barrow, D. K., & Crone, S. F. (2016a). Cross-validation aggregation for combining autoregressive neural network forecasts. *International Journal of Forecasting*, 32(4), 1120–1137.
- Barrow, D. K., & Crone, S. F. (2016b). A comparison of adaboost algorithms for time series forecast combination. *International Journal of Forecasting*, 32(4), 1103–1119.
- Berriel, R. F., Lopes, A. T., Rodrigues, A., Varejao, F. M., & Oliveira-Santos, T. (2017). Monthly energy consumption forecast: A deep learning approach. En *2017 international joint conference on neural networks (ijcnn)* (pp. 4283–4290).
- Bezerra, B., Veiga, A., Barroso, L. A., & Pereira, M. (2016). Stochastic long-term hydrothermal scheduling with parameter uncertainty in autoregressive streamflow models. *IEEE Transactions on Power Systems*, 32(2), 999–1006.
- Biswas, A., & Emadi, A. (2019). Energy management systems for electrified powertrains: State-of-the-art review and future trends. *IEEE Transactions on Vehicular Technology*, 68(7), 6453–6467.

- Black, J. D., & Henson, W. L. (2013). Hierarchical load hindcasting using reanalysis weather. *IEEE Transactions on Smart Grid*, 5(1), 447–455.
- Borroto Nordelo, A., & Monteagudo Yanes, J. P. (2006). Gestión y economía energética. *Cienfuegos, Cuba: Editorial Universo Sur*, 98.
- Cabello Justafré, J. J. (2018). *Indicador para la eficiencia energética municipal en Cuba. Caso de estudio municipio de Cienfuegos*. Tesis de grado, Universidad de Cienfuegos.
- Campillo Sabina, E. (2018). *Diagnóstico Energético al Municipio de Cienfuegos*. Tesis de maestría, Universidad de Cienfuegos.
- Caravaca, A. M., Marrero, R. H. H., Fichez, R., Douillet, P., García, O. D., & Fernández, J. M. (2014). Influencia de las características hidrodinámicas y morfométricas en la distribución de <sup>210</sup>Pb en los sedimentos superficiales de la bahía de Cienfuegos, Cuba. *Revista de Investigaciones Marinas*, 31(1), 11–22.
- Castle, J. L., Doornik, J. A., & Hendry, D. F. (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics*, 3(1).
- Castle, J. L., Doornik, J. A., & Hendry, D. F. (2012). Model selection when there are multiple breaks. *Journal of Econometrics*, 169(2), 239–246.
- Celik, B., Roche, R., Suryanarayanan, S., Bouquain, D., & Miraoui, A. (2017). Electric energy management in residential areas through coordination of multiple smart homes. *Renewable and Sustainable Energy Reviews*, 80, 260–275.
- Cevik, A., Weber, G. W., Eyüboğlu, B. M., Karlı-Oğuz, K., The Alzheimer’s Disease Neuroimaging Initiative, et al.. (2017). Voxel-mars: a method for early detection of Alzheimer’s disease by classification of structural brain MRI. *Annals of Operations Research*, 258(1), 31–57.
- Chaviano Díaz, J., & Puebla Alfonso, J. C. (2018). *Despliegue de la Función Calidad para mejorar la Planificación Energética en el Municipio de Cienfuegos*. Tesis de grado, Universidad de Cienfuegos.
- Coleman, C., Lyon, S., Maliar, L., & Maliar, S. (2020). Matlab, python, julia: What to choose in economics? *Computational Economics*, 1–26.

- Consejo de Estado. (2019). Decreto-Ley No. 345 (GOC-2019-1063-O95) Del desarrollo de las fuentes renovables y el uso eficiente de la energía. *Gaceta Oficial de la República de Cuba*, Ministerio de Justicia. (95), 2123-2128.
- Correa Soto, J., Borroto Nordelo, A. E., González Álvarez, R., Curbelo Martínez, M., Díaz Rodríguez, A. M., *et al.*. (2014). Diseño y aplicación de un procedimiento para la planificación energética según la nc-iso 50001: 2011. *Ingeniería energética*, 35(1), 38–47.
- Correa Soto, J., Cabello Eras, J. J., Nogueira Rivera, D., Haeseldonckx, D., Sagastume Gutierrez, A., & Silva Oliveira, L. F. (2018). Municipal energy management model for cuban first level municipalities.
- Correa Soto, J., González Pérez, S., & Hernández Alonso, Á. (2017). La gestión energética local: elemento del desarrollo sostenible en cuba. *Revista Universidad y Sociedad*, 9(2), 59–67.
- Correa Soto, J., Salmerón, D. M. S., Eras, J. J. C., Rivera, D. N., & Viñales, Y. A. D. (2021). Balance energético como elemento de la gestión de gobierno local en cuba: caso estudio municipio de cienfuegos. *Universidad y Sociedad*, 13(1), 266–275.
- De Giorgi, M. G., Congedo, P. M., & Malvoni, M. (2014). Photovoltaic power forecasting using statistical methods: impact of weather data. *IET Science, Measurement & Technology*, 8(3), 90–97.
- Denis, G. S., & Parker, P. (2009). Community energy planning in canada: The role of renewable energy. *Renewable and Sustainable Energy Reviews*, 13(8), 2088–2095.
- Díaz La Hoz, M. (2019). *Mejora de la gestión energética municipal de cienfuegos*. Tesis de maestría, Universidad de Cienfuegos.
- Douglas, M., *et al.*. (2002). *Diseño y análisis de experimentos*.
- Elma, O., Tascikaraoğlu, A., Ince, A. T., & Selamoğulları, U. S. (2017). Implementation of a dynamic energy management system using real time pricing and local renewable energy generation forecasts. *Energy*, 134, 206–220.
- Fallah, S. N., Deo, R. C., Shojafar, M., Conti, M., & Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: state of the art, future challenges, and research directions. *Energies*, 11(3), 596.

- Ferrari, S., Bruni, E., Bramonti, L., *et al.*. (2020). Effective implementation of iso 50001: A case study on energy management for heating load reduction for a social building stock in northern italy. *Energy and Buildings*, *219*, 110029.
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 1–67.
- García Benedicto, L. (2004). *Desarrollo de un modelo multicriterio-multiobjetivo de oferta de energías renovables: aplicación a la comunidad de madrid*. Tesis de doctorado, Montes.
- Goude, Y., Nedellec, R., & Kong, N. (2013). Local short and middle term electricity load forecasting with semi-parametric additive models. *IEEE transactions on smart grid*, *5*(1), 440–446.
- Gujarati, D. N., & Porter, D. C. (2013). *Econometría* (5 ed.). Mexico: McGraw-Hill.
- Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), 914–938.
- Hurtado Espinosa, L. (2017). *Diseño de un producto informático para la gestión de la energía del gobierno municipal de Cienfuegos*. Tesis de grado, Universidad de Cienfuegos.
- IBM.Inc. (Consultado en:Feb, 2021). *Spss: Statistical package for the social sciences*.
- Innes, J. E., & Booher, D. E. (2015). A turning point for planning theory? overcoming dividing discourses. *Planning theory*, *14*(2), 195–213.
- Jovanović, B., & Filipović, J. (2016). Iso 50001 standard-based energy management maturity model–proposal and validation in industry. *Journal of Cleaner Production*, *112*, 2744–2755.
- Karunathilake, H., Perera, P., Ruparathna, R., Hewage, K., & Sadiq, R. (2018). Renewable energy integration into community energy systems: A case study of new urban residential development. *Journal of Cleaner Production*, *173*, 292–307.
- Keshtegar, B., Mert, C., & Kisi, O. (2018). Comparison of four heuristic regression techniques in solar radiation modeling: Kriging method vs rsm, mars and m5 model tree. *Renewable and sustainable energy reviews*, *81*, 330–341.
- Khorasany, M., Mishra, Y., & Ledwich, G. (2018). Market framework for local energy trading: a review of potential designs and market clearing approaches. *IET Generation, Transmission & Distribution*, *12*(22), 5899–5908.

- Khuntia, S. R., Rueda, J. L., & Meijden, M. A. van der. (2016). Forecasting the load of electrical power systems in mid-and long-term horizons: a review. *IET Generation, Transmission & Distribution*, 10(16), 3971–3977.
- Kuter, S., Akyurek, Z., & Weber, G.-W. (2018). Retrieval of fractional snow covered area from modis data by multivariate adaptive regression splines. *Remote sensing of environment*, 205, 236–252.
- Kuter, S., Weber, G. W., Özmen, A., & Akyürek, Z. (2014). Modern applied mathematics for alternative modeling of the atmospheric effects on satellite images. En *Modeling, dynamics, optimization and bioeconomics i* (pp. 469–485). Springer.
- Lindkvist, C., Juhasz-Nagy, E., Nielsen, B. F., Neumann, H.-M., Lobaccaro, G., & Wyckmans, A. (2019). Intermediaries for knowledge transfer in integrated energy planning of urban districts. *Technological Forecasting and Social Change*, 142, 354–363.
- Liu, B., Nowotarski, J., Hong, T., & Weron, R. (2015). Probabilistic load forecasting via quantile regression averaging on sister forecasts. *IEEE Transactions on Smart Grid*, 8(2), 730–737.
- Marzband, M., Sumper, A., Domínguez-García, J. L., & Gumara-Ferret, R. (2013). Experimental validation of a real time energy management system for microgrids in islanded mode using a local day-ahead electricity market and minlp. *Energy Conversion and Management*, 76, 314–322.
- MathWorks.Inc. (Consultado en: Feb, 2021). *Matlab: Matrix laboratory*.
- Mengelkamp, E., Schönland, T., Huber, J., & Weinhardt, C. (2019). The value of local electricity-a choice experiment among german residential customers. *Energy Policy*, 130, 294–303.
- Mijwel, M. M. (2018). Artificial neural networks advantages and disadvantages. *Retrieved from LinkedIn <https://www.linkedin.com/pulse/artificial-neuralnet Work>*.
- MINEM. (2019). Resolución 124 (GOC-2019-1067-O95) Regulaciones para elevar la gestión, eficiencia y conservación energética. Gaceta Oficial de la República de Cuba, Ministerio de Justicia. (95), 2123-2128.
- Mirakyan, A., & De Guio, R. (2013). Integrated energy planning in cities and territories: A review of methods and tools. *Renewable and sustainable energy reviews*, 22, 289–297.

- Mirjat, N. H., Uqaili, M. A., Harijan, K., Valasai, G. D., Shaikh, F., & Waris, M. (2017). A review of energy and power planning and policies of pakistan. *Renewable and Sustainable Energy Reviews*, 79, 110–127.
- Nalcaci, G., Özmen, A., & Weber, G. W. (2019). Long-term load forecasting: models based on mars, ann and lr methods. *Central European Journal of Operations Research*, 27(4), 1033–1049.
- Nguyen, T.-H., Nguyen, L. V., Jung, J. J., Agbehadji, I. E., Frimpong, S. O., & Millham, R. C. (2020). Bio-inspired approaches for smart energy management: State of the art and challenges. *Sustainability*, 12(20), 8495.
- Nieta, A. A. S. de la, Ilieva, I., Gibescu, M., Bremdal, B., Simonsen, S., & Gramme, E. (2020). Optimal midterm peak shaving cost in an electricity management system using behind customers smart meter configuration. *Applied Energy*, 116282.
- Olive, D. J. (2017). *Linear regression*. Springer.
- ONEI Cienfuegos. (2019). *Anuario Estadístico 2018 Cienfuegos* (Tech. Rep.). Tomado de: <http://www.onei.gob.cu>.
- ONU. (2015). Proyecto de resolución remitido a la cumbre de las Naciones Unidas para la aprobación de la agenda para el desarrollo después del 2015 por la Asamblea General en su lxxix periodo de sesiones. transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible. A/70/L.1. *ISSN:15-15900*.
- Ortiz Alvarado, M. E. (2015). *Pronóstico de la demanda eléctrica residencial basado en el modelo de regresión adaptativa multivariante spline (mars)*. Tesis de maestría.
- Ozgun, C., Colliau, T., Rogers, G., Hughes, Z., et al.. (2017). MatLab vs. Python vs. R. *Journal of Data Science*, 15(3), 355–371.
- Özmen, A. (2016). *Robust optimization of spline models and complex regulatory networks* (Vol. 139). Springer.
- Özmen, A., & Weber, G. W. (2014). Rmars: robustification of multivariate adaptive regression spline under polyhedral uncertainty. *Journal of Computational and Applied Mathematics*, 259, 914–924.
- Pajankar, A. (2021). Introduction to Python. En *Practical Python Data Visualization* (pp. 1–16). Springer.

- PCC. (2017). Documentos del 7mo. Congreso de Partido aprobados por el III Pleno del Comité Central del PCC el 18 de mayo de 2017 y respaldados por la Asamblea Nacional del Poder Popular el 1 de junio de 2017. Tabloides I y II. Cuba.
- Penelas, A., & Gómez, M. (1999). Aplicación del análisis cluster. *Hair, Joseph; Anderson, Rolph; Tatham, Ronald; Black, William. Análisis multivariante*, 533–536.
- Python Core Team. (Consultado en: Feb, 2021). *Python: A dynamic, open source programming language*.
- R Core Team. (Consultado en: Feb, 2021). *R: A language and environment for statistical computing*. Vienna, Austria.
- Ranjani, J., Sheela, A., & Meena, K. P. (2019). Combination of numpy, scipy and matplotlib/pylab-a good alternative methodology to matlab-a comparative analysis. En *2019 1st international conference on innovations in information and communication technology (iciict)* (pp. 1–5).
- Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information*, 11(4), 193.
- Ravadanegh, S. N., Jahanyari, N., Amini, A., & Taghizadeghan, N. (2016). Smart distribution grid multistage expansion planning under load forecasting uncertainty. *IET Generation, Transmission & Distribution*, 10(5), 1136–1144.
- Resch, E., & Andresen, I. (2017). Current challenges of urban energy planning in a norwegian municipality. En *Proceedings of world sustainable built environment conference 2017 hong kong. transforming our built environment through innovation and integration: Putting ideas into action. wsbe17*.
- Rodríguez, R. P., *et al.* (2018). Combustibles alternativos de segunda y tercera generación para motores de combustión interna. *Anales de la Academia de Ciencias de Cuba*, 8(1).
- Rodríguez Figueredo, S. (2019). *Integración de las potencialidades energéticas al desarrollo local del municipio de Cienfuegos*. Tesis de maestría, Universidad de Cienfuegos.
- Saez-Gallego, J., & Morales, J. M. (2017). Short-term forecasting of price-responsive loads using inverse optimization. *IEEE Transactions on Smart Grid*, 9(5), 4805–4814.
- SAS.Inc. (Consultado en: Feb, 2021). *Sas: Analytics software & solutions*.

- Sheikhahmadi, P., Bahramara, S., Shahrokhi, S., Chicco, G., Mazza, A., & Catalão, J. P. (2020). Modeling local energy market for energy management of multi-microgrids. En *2020 55th international universities power engineering conference (upec)* (pp. 1–6).
- Silva Goncalves, V. A. da, & Santos, F. J. M.-H. dos. (2019). Energy management system iso 50001: 2011 and energy management for sustainable development. *Energy Policy*, *133*, 110868.
- Suresh, K. (2021). *Design Optimization using MATLAB and SOLIDWORKS*. Cambridge University Press.
- Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225–1231.
- Vanegas, J., & Vásquez, F. (2017). Multivariate adaptative regression splines (mars), una alternativa para el análisis de series de tiempo. *Gaceta sanitaria*, *31*, 235–237.
- Wang, L., Zhang, Z., & Chen, J. (2016). Short-term electricity price forecasting with stacked denoising autoencoders. *IEEE Transactions on Power Systems*, *32*(4), 2673–2681.
- Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, *30*(4), 1030–1081.
- Xiao, L., Shao, W., Liang, T., & Wang, C. (2016). A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting. *Applied energy*, *167*, 135–153.
- Xie, J., Hong, T., & Stroud, J. (2015). Long-term retail energy forecasting with consideration of residential customer attrition. *IEEE transactions on smart grid*, *6*(5), 2245–2252.
- Yang, L., Wan, K. K., Li, D. H., & Lam, J. C. (2011). A new method to develop typical weather years in different climates for building energy use studies. *Energy*, *36*(10), 6121–6129.
- Yang, X., Zhang, Y., Zhao, B., Huang, F., Chen, Y., & Ren, S. (2017). Optimal energy flow control strategy for a residential energy local network combined with demand-side management and real-time pricing. *Energy and Buildings*, *150*, 177–188.
- Zhao, Y., Nasrullah, Z., & Li, Z. (2019). Pyod: A python toolbox for scalable outlier detection. *arXiv preprint arXiv:1901.01588*.

Zheng, Y., & Weng, Q. (2019). Modeling the effect of climate change on building energy demand in los angeles county by using a gis-based high spatial and temporal resolution approach. *Energy*, 176, 641–655.