

**MONITORIZACIÓN NO INTRUSIVA DE CARGA: RECOLECCION DE  
DATOS Y CLASIFICACIÓN EN EL TIEMPO**

**JEISSON DAVID BELLO PEÑA**

**CARLOS ERIXÓN BELLO PEÑA**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO – MECÁNICAS  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA  
Y DE TELECOMUNICACIONES  
BUCARAMANGA**

**2014**

**MONITORIZACIÓN NO INTRUSIVA DE CARGA: RECOLECCION DE DATOS Y  
CLASIFICACIÓN EN EL TIEMPO**

**JEISSON DAVID BELLO PEÑA**

**Trabajo de grado para optar por el título de**

**Ingeniero electricista**

**CARLOS ERIXÓN BELLO PEÑA**

**Trabajo de grado para optar por el título de**

**Ingeniero electrónico**

**DIRECTOR**

**Ph.D(c) YULIETH JIMENEZ MANJARRES**

**CODIRECTOR**

**Ph.D GABRIEL ORDÓÑEZ PLATA**

**UNIVERSIDAD INDUSTRIAL DE SANTADER**

**FACULTAD DE INGENIERÍAS FÍSICO – MECÁNICAS**

**ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA**

**Y DE TELECOMUNICACIONES**

**BUCARAMANGA**

**2014**

## TABLA DE CONTENIDO

<b>INTRODUCCIÓN .....</b>	<b>14</b>
<b>1. SELECCIÓN DE EQUIPOS .....</b>	<b>18</b>
<b>2. RECOLECCIÓN DE DATOS .....</b>	<b>20</b>
<b>3. PROCESAMIENTO Y ANÁLISIS DE DATOS. ....</b>	<b>24</b>
3.1 Pre-procesamiento de las señales.....	24
3.2 Procesamiento de señales.....	27
3.3 Análisis de gráficas. ....	29
3.3.1 Análisis <i>clustering</i> directo.....	41
3.3.2 Análisis de Componentes Principales (PCA). ....	43
3.3.3 Análisis de Discriminación Lineal (LDA).....	45
3.4 Comparación métodos de análisis. ....	46
<b>4. CONCLUSIONES. ....</b>	<b>47</b>
<b>Referencias bibliográficas: .....</b>	<b>49</b>
<b>Bibliografía: .....</b>	<b>51</b>

## LISTA DE FIGURAS

Figura 1 Esquema de procesos seguidos por NILM [4].	16
Figura 2 Analizador de Redes Power Visa DRANETZ – BMI 1000	20
Figura 3 Pinzas de Corriente TR2510 Dranetz BMI	20
Figura 4 Diagrama unifilar distribución de circuitos ramales laboratorio máquinas eléctricas.	21
Figura 5 Plano eléctrico multitoma.	22
Figura 6 Multitoma físico	22
Figura 7 Puertos de conexiones del multitoma.	23
Figura 8 Diagrama conexiones del multitoma y medidor.	23
Figura 9 Esquema general del tratamiento de datos	24
Figura 10 Transitorio de tensión en encendido del ventilador en velocidad uno	26
Figura 11 Transitorio de corriente en encendido del ventilador en velocidad uno.	26
Figura 12 Factor de Cresta Vs. Desviación – encendido.	30
Figura 13 Factor de Cresta Vs. Desviación – apagados.	31
Figura 14 Entropía Vs. Desviación – encendidos.	31
Figura 15 Entropía Vs. Desviación – apagados.	32
Figura 16 Curtosis Vs. Desviación – encendidos.	32
Figura 17 Curtosis Vs. Desviación – apagados.	33
Figura 18 Potencia Vs. Desviación – encendidos	33
Figura 19 Potencia Vs. Desviación – apagados.	34
Figura 20 Entropía Vs. Factor de Cresta – encendidos.	34
Figura 21 Entropía Vs. Factor de Cresta – apagados.	35
Figura 22 Curtosis Vs. Factor de Cresta – encendidos.	35
Figura 23 Curtosis Vs. Factor de Cresta – apagados.	36

Figura 24 Potencia Vs. Factor de Cresta – encendidos.....	36
Figura 25 Potencia Vs. Factor de Cresta – pagados. ....	37
Figura 26 Curtosis Vs. Entropía – Encendidos. ....	37
Figura 27 Curtosis Vs. Entropía – Apagados.....	38
Figura 28 Potencia Vs. Entropía – Encendidos. ....	38
Figura 29 Potencia Vs. Entropía – Apagados.....	39
Figura 30 Potencia Vs. Curtosis – Encendidos.....	39
Figura 31 Potencia Vs. Curtosis – Apagados. ....	40
Figura 32 Frecuencias por grupo con clustering directo. ....	42
Figura 33 Frecuencias por grupo con PCA y posterior clustering.....	44
Figura 34 Frecuencias por grupo con LDA y posterior clustering. ....	46
Figura 35 Resultados finales de los tres análisis. ....	47

## LISTA DE TABLAS

Tabla 1 Características posibles medidores .....	19
Tabla 2 Conteo de encendidos y apagados por elemento.....	25
Tabla 3 Código de colores utilizado en las figuras 12-31.....	30
Tabla 4 Resultados del clustering.....	42
Tabla 5 Resultados del clustering después de PCA.....	44
Tabla 6 Resultados del clustering después de LDA.....	46

## RESUMEN

Título: Monitorización no intrusiva de carga: Recolección de datos y clasificación en el tiempo.<sup>1</sup>

Autores: Jeisson David Bello Peña y Carlos Erixón Bello Peña.<sup>2</sup>

Palabras claves: Monitorización No Intrusiva de Carga (NILM), desagregación de cargas, Analisis de Discriminación Lineal (LDA), Análisis de Componentes Principales (PCA), *clustering*, varianza, gestión de energía.

Los estudios relacionados con la medición de energía eléctrica siempre han sido un tema de gran importancia para el mundo actual. Una de las tecnologías que cada día toma más fuerza en esta área son las tecnologías Monitorización No Intrusiva de Carga (NILM). Una de las limitaciones más grandes que se encuentra al tratar de implementar tecnologías NILM es la desagregación de los datos para cada elemento de consumo. En este artículo se presenta un conjunto de datos de las señales de tensión y corriente muestreadas a frecuencia de 15,4 kHz bajo un ambiente controlado para cada electrodoméstico trabajado (plancha, licuadora, ventilador, lámpara incandescente y lámpara fluorescente), con el fin de presentar y autorizar el uso público dando una base para el análisis y prueba de algoritmos de desagregación, base fundamental de implementaciones NILM. Además de los equipos y procedimientos seguidos para la adquisición de datos, se muestra un primer análisis de las cargas, enfocado a la desagregación de cargas.

Al momento de tener métodos de mediciones energéticas eficientes, se puede realizar una mejor gestión de energía, lo cual podría llegar a ser uno de los grandes pilares para la disminución de la emisión de gases de efecto invernadero. Pensar en mediciones energéticas eficientes es pensar en grandes avances tecnológicos, avances como los que podría alcanzar la monitorización no intrusiva de carga, por sus características de bajo costo y, dependiendo del algoritmo de desagregación utilizado, la amplia gama de información que se podría obtener con este.

---

<sup>1</sup> Trabajo de grado

<sup>2</sup> Facultad de ingenierías Fisicomecánicas, Escuela de Ingeniería eléctrica, electrónica y de telecomunicaciones, Director: Yulieth Jimenez Manjarres Codirector: Gabriel Ordóñez Plata.

## ABSTRACT

Title: non-intrusive load monitoring: Data collection and classification in time.<sup>3</sup>

Authors: Jeisson David Bello Peña y Carlos Erixón Bello Peña.<sup>4</sup>

Keywords: non-intrusive load monitoring (NILM), Disaggregation of loads, linear discriminant analysis (LDA), principal component analysis (PCA), *clustering*, variance, power management.

Studies related to the measurement of electrical energy have always been a major issue in today's world. One of the technologies that every day takes more strength in this area is the technology of Non-Intrusive Load Monitoring (NILM). One of the biggest limitations find it when we want to implement technologies NILM is the disaggregation of data of consumption for each appliance. In this paper is presented a data set of voltage and current signals sampled at a frequency of 15.4 kHz where the environment of operation in each appliances is controlled (iron, blender, fan, incandescent and fluorescent lamp). This data set is built in order to present and authorize its use in public domain giving a dataset for the analysis and test of disaggregation algorithms, foundation in NILM implementations. In addition the equipment and procedures used in the data acquisition, and an initial analysis of loads, focused on the breakdown of charges is shown.

When having measurements of energy efficient methods, you can make a better power management, which could become one of the major pillars for reducing the emission of greenhouse gases. Think of efficient energy measurements is thinking in huge technological advances, advances like that could achieve non-intrusive load monitoring due to its characteristics of low cost and, depending on the algorithm used disaggregation, the wide range of information that could be obtained with this

---

<sup>3</sup> Bachelor thesis

<sup>4</sup> Faculty of Physicomechanics. School of Electric, Electronic and the Telecommunications Engineering. Director: Yulieth Jimenez Manjarres Codirector: Gabriel Ordóñez Plata.

## **Monitorización no intrusiva de carga: Recolección de datos y clasificación en el tiempo.**

**Jeisson David Bello Peña.**

**Ingeniero Electricista UIS.**

**Jeisson.bello@correo.uis.edu.co**

**Carlos Erixón Bello Peña.**

**Ingeniero Electrónico UIS.**

**Carlos.bello@correo.uis.edu.co**

### **Resumen**

Los estudios relacionados con la medición de energía eléctrica siempre han sido un tema de gran importancia para el mundo actual. Una de las tecnologías que cada día toma más fuerza en esta área son las tecnologías Monitorización No Intrusiva de Carga (NILM). Una de las limitaciones más grandes que se encuentra al tratar de implementar tecnologías NILM es la desagregación de los datos para cada elemento de consumo. En este artículo se presenta un conjunto de datos de las señales de tensión y corriente muestreadas a frecuencia de 15,4 kHz bajo un ambiente controlado para cada electrodoméstico trabajado (plancha, licuadora, ventilador, lámpara incandescente y lámpara fluorescente), con el fin de presentar y autorizar el uso público dando una base para el análisis y prueba de algoritmos de desagregación, base fundamental de implementaciones NILM. Además de los equipos y procedimientos seguidos para la adquisición de datos, se muestra un primer análisis de las cargas, enfocado a la desagregación de cargas.

Al momento de tener métodos de mediciones energéticas eficientes, se puede realizar una mejor gestión de energía, lo cual podría llegar a ser uno de los grandes pilares para la disminución de la emisión de gases de efecto invernadero. Pensar en mediciones energéticas eficientes es pensar en grandes avances tecnológicos, avances como los que podría alcanzar la monitorización no intrusiva de carga, por sus características de bajo costo y, dependiendo del algoritmo de

desagregación utilizado, la amplia gama de información que se podría obtener con este.

### **Palabras claves**

Monitorización No Intrusiva de Carga (NILM), desagregación de cargas, Analisis de Discriminación Lineal (LDA), Análisis de Componentes Principales (PCA), *clustering*, varianza, gestión de energía.

### **INTRODUCCIÓN**

Es un hecho que el clima global está siendo alterado como resultado del aumento de concentraciones de gases que favorecen el efecto invernadero. Este cambio acarrea problemáticas en los ecosistemas terrestres, problemáticas como el aumento en cantidad e intensidad de tormentas, sequias, migraciones de animales, deshielos, entre otros [1]. Gran parte de estos gases son generados por el uso irresponsable de la energía eléctrica, que desde su generación, pasando por el transporte y llegando a su consumo final, está continuamente aportando a la contaminación ambiental.

Se considera que una adecuada gestión de energía logra disminuir desde un 4 hasta un 15% el consumo total de energía [2], lo que ayudará a la mitigación de los efectos de tipo invernadero. Además de la importancia de los impactos ambientales, en la actualidad se cuenta con un mercado energético en desarrollo, donde cada día se necesita tener mayor información del consumo no solamente a nivel industrial sino también residencial y comercial.

Para seguir el ritmo acelerado de los avances tecnológicos en procesos como robótica, electrónica, automatización, etc, del momento, se necesita información veraz, completa y en tiempo real, del consumo energético. Esta información es de gran importancia tanto para consumidores como para comercializadores de energía, de modo que deben ser suplidos de manera eficaz, lo cual ha llevado a una continua búsqueda de tecnologías con la capacidad de suplir estas necesidades.

Colombia es un país que en este momento realiza enormes trabajos hacia la búsqueda de nuevas soluciones energéticas y de un mejor manejo en la medición de los consumos energéticos. Una gran prueba de estos esfuerzos es la creación del Instituto Nacional de Metrología Colombia (3 de noviembre del 2011 por medio del decreto 4175 del 2011), donde uno de sus principales enfoques es el manejo energético. Como este, hay grandes grupos empresariales que se encuentran impulsando trabajos dirigidos a manejos energéticos [3]. Lo que se espera con la creación de estas entidades es regular la facturación energética y crear conciencia tanto en los usuarios como en los comercializadores acerca de sus pérdidas energéticas, las cuales se podrían disminuir notoriamente con una correcta gestión energética.

Una de las soluciones que mejores resultados ha brindado hasta el momento, para la gestión de energía, es la Monitorización Intrusiva de Carga (ILM del inglés *Intrusive Load Monitoring*), la cual tiene la capacidad de identificar la operación de los electrodomésticos y su consumo individual en el hogar, y realizarle seguimiento. La falencia que conlleva un sistema ILM es que se debe instalar un sensor, o más, por cada elemento a monitorizar dentro del hogar, hecho que acarrea una gran inversión tanto en los sensores como en la compleja instalación necesaria de los sistemas de comunicación para recolectar los datos de ellos.

Una tecnología que brinda una solución a los problemas de sobrecostos acarreados en la ILM es la Monitorización No Intrusiva de Carga (NILM del inglés *Non Intrusive Load Monitoring*), la cual, al igual que la ILM, pretende monitorizar todos los elementos del hogar, con la diferencia que realiza una única medición en la entrada principal de energía de las viviendas. Se han venido realizando investigaciones en sistemas NILM y cada vez toma más fuerza como verdadera solución tanto por sus cómodos costos, como por su gran flexibilidad, dado que cada día los medidores brindan más información y esto aporta a que se puedan obtener sistemas NILM con mayor precisión [4].

El concepto de monitorización no intrusiva fue propuesto inicialmente por el MIT en los ochentas. Ese trabajo propone dividir los electrodomésticos caseros en cuatro categorías [5]:

1. Dispositivos de consumo permanente: Permanecerán encendidos durante 24 horas al día, 7 días a la semana, con consumo aproximadamente constante de energía activa y reactiva. Ejemplos de dispositivos de esta categoría incluyen detectores de humo cableados y cámaras de seguridad.
2. Máquinas de estados finitos: Pasan a través de varios estados de conmutación, por lo que el ciclo de conmutación completo está repetido con frecuencia en el ciclo diario o semanal de eventos. Ejemplos de máquinas de estados finitos son una lavadora o secadora de ropa.
3. Dispositivos de consumo variable: No tienen patrón periódico de cambio en los estados o de potencia. Ejemplos de tales aparatos son luces regulables y taladros, martillos y peines eléctricos.
4. Dispositivos de encendido y apagado: Aquellos que funcionan a voluntad del usuario y mantienen un consumo aproximadamente permanente, entre esta categoría se encuentran la mayoría de los aparatos domésticos, tales como una tostadora, bombilla y la bomba de agua.

En la Figura 1 se observa un esquema general de un sistema NILM [4]:

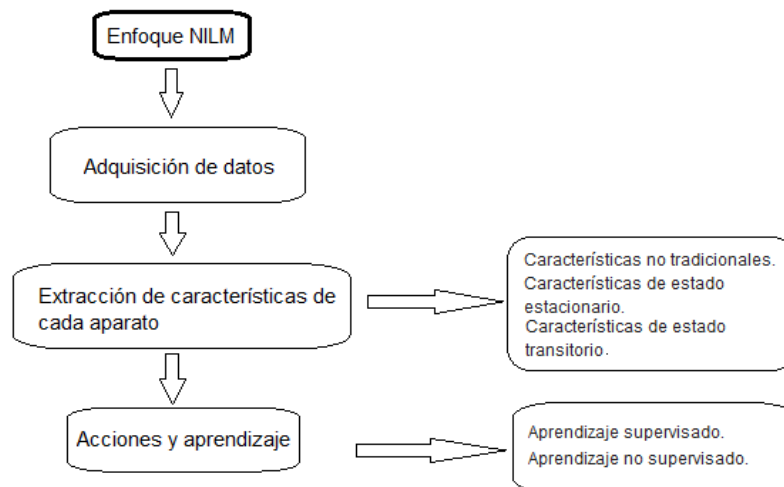


Figura 1 Esquema de procesos seguidos por NILM [4].

Este proyecto realiza la adquisición de datos, la correspondiente extracción de características del estado transitorio, y finalmente se analiza la estrategia de aprendizaje supervisado no supervisado.

A partir de estos estudios se han creado diferentes modelos pero siguiendo siempre una filosofía similar a la planteada por el MIT. Entre algunos de los más destacados se encuentran:

- El método MIT con filtros armónicos y formas de transitorios [6].
- *Cole and Albicki method* [7] [8].
- *Concordia University (CU) data disaggregation* [9].
- *Baranski's method* [10] [11] [12].

El objetivo final de la implementación de las tecnologías NILM es desagregar la carga mediante la identificación de eventos de encendido y apagado de los elementos que encuentre bajo su supervisión. Debe entenderse como evento no solo la conexión y desconexión, sino también cualquier cambio de estado que sufra determinado aparato durante su funcionamiento.

Para el desarrollo de algoritmos de desagregación de cargas para las tecnologías NILM se requiere un conjunto de datos de señales eléctricas correspondientes a la operación de los equipos. Existen bases de datos públicamente disponibles tales como, *REDD: Reference Energy Disaggregation Data Set* [13] y la *BLUED: Building-Level fully-labeled data set for Electricity Disaggregation* [14]. No obstante presentan limitaciones como son la baja frecuencia de muestreo de la REDD y la falta de etiquetas de encendido de apagado de la BLUED. Con el fin de mejorar estas limitaciones, se construyó el conjunto de datos CDHB: Conjunto de Datos Hermanos Bello.

En este documento se presentará el conjunto de datos del funcionamiento de electrodomésticos básicos de uso residencial, incluyendo la información de tensión, corriente y consumo de potencia en los distintos estados. Además están

presentes datos no solo de operación individual sino también simultánea, con sus respectivas etiquetas de encendido y apagado.

Este conjunto de datos se hará público a la comunidad investigativa con el fin de aportar y motivar el desarrollo de nuevos algoritmos NILM. Además se entrega un análisis con técnicas estadísticas para la representación y discriminación de las características transitorias como sobrepicos y energía del transitorio.

Cabe resaltar que este conjunto de datos, a pesar de estar enfocado a estudios en tecnologías NILM, puede ser útil en estudios de patrones de funcionamiento de electrodomésticos, diagnóstico de fallas, detección de anomalías, uso seguro de electrodomésticos, entre otros. Así que varios interesados, como consumidores, empresas de servicios y distribución, fabricantes de electrodomésticos, etc, podrían hacer uso de esta información.

El resto del presente documento se conforma de la siguiente manera. La sección 2 enseña la selección de los equipos necesarios para la implementación de la toma de datos. La sección 3 muestra el montaje del hardware utilizado para la recolección de los datos. La sección 4 presenta el análisis del comportamiento de los electrodomésticos utilizados. Finalmente en la sección 5 se discuten las conclusiones de este trabajo.

## **1. SELECCIÓN DE EQUIPOS**

Entre los equipos a utilizar durante la recopilación de datos, sin duda, el más importante fue el equipo de medida. La Tabla 1 muestra un cuadro comparativo donde se plasmaron las características más relevantes para cada equipo (analizador de redes) disponibles para GISEL<sup>5</sup>, con el fin de realizar la selección.

---

<sup>5</sup> Grupo de Investigación en Sistemas de Energía Eléctrica.

EQUIPO	Frecuencia de muestreo [kHz]	PRECISIÓN		RANGO		Canales de corriente	Canales de tensión	Almacenamiento de datos
		Tensión [V]	Corriente [A]	Tensión [Vrms]	Corriente [A]			
Power Visa	15,4	±0,1% ±0,05	±1% ±0,01	10-600	0,1-10	4	4	Memoria
Fluke 434-II	15,4	±0,5% de la tensión nominal	±0,5%±5 recuento	1-1000	0,5-600	4	4	Memoria
P-Cube	15,4	±0.05%rdg ±0.05% FS typical	±0.2%rdg ±0.2% FS typical <sup>6</sup>	0-360	0-14	4	4	Memoria

Tabla 1 Características posibles medidores

Con el análisis de las características de los equipos, se concluyó que el equipo que mejor se ajustaba a las pruebas a realizar fue el analizador de redes Power Visa DRANETZ – BMI 1000 (Figura 2), con el cual se logra una frecuencia de muestreo de alrededor de 15,4 kHz tanto para la tensión como para la corriente simultáneamente. Además se incluye el software Dran – View 6.3 que permite exportar los datos de potencias, energías, análisis de armónicos, caídas, sobrepicos e interrupciones, entre otros, de acuerdo a la norma IEEE 1159. Otra de las grandes ventajas que presentaba el Power Visa, fue que incluía una reciente calibración, de manera que se podía obtener un mayor grado de confianza en los datos obtenidos con él.

Luego de tener seleccionado el equipo de medición a utilizar, se llevó a cabo la búsqueda de las pinzas de corriente que mejor se adecuaban a la adquisición de muestras, para ello se tomaron las pinzas de corriente TR2510 Dranetz BMI (Figura 3), las cuales aportan la exactitud y el rango mostrado en la Tabla 1.

<sup>6</sup> El rango, la precisión y la cantidad de canales, en este caso, están dados por el módulo de corriente externo (CT4-50A-00) que se debe instalar al analizador de red.



Figura 2 Analizador de Redes Power Visa DRANETZ – BMI 1000



Figura 3 Pinzas de Corriente TR2510 Dranetz BMI

## 2. RECOLECCIÓN DE DATOS

El objetivo que se planteó inicialmente fue obtener datos locales de tensión, corriente y las correspondientes etiquetas de encendido y apagado para electrodomésticos de uso común. Basados en el estudio realizado por la Unidad de Planeación Minero Energética (UPME) [15] de los elementos de mayor uso doméstico de electrodomésticos, y teniendo en cuenta la disponibilidad, se seleccionaron los siguientes para la realización del conjunto de datos: plancha, licuadora, bombillo incandescente, lámpara fluorescente y ventilador. Se puede

encontrar especificaciones de cada uno de los electrodomésticos a probar en el Anexo 1.

Se selecciona el laboratorio de máquinas eléctricas para realizar las mediciones porque se tiene pleno conocimiento de la distribución de los circuitos ramales dentro del área del laboratorio (ver Figura 4.). Adicionalmente, este es uno de los lugares donde la acometida principal llega lo más directamente posible desde el transformador hasta el tablero principal.

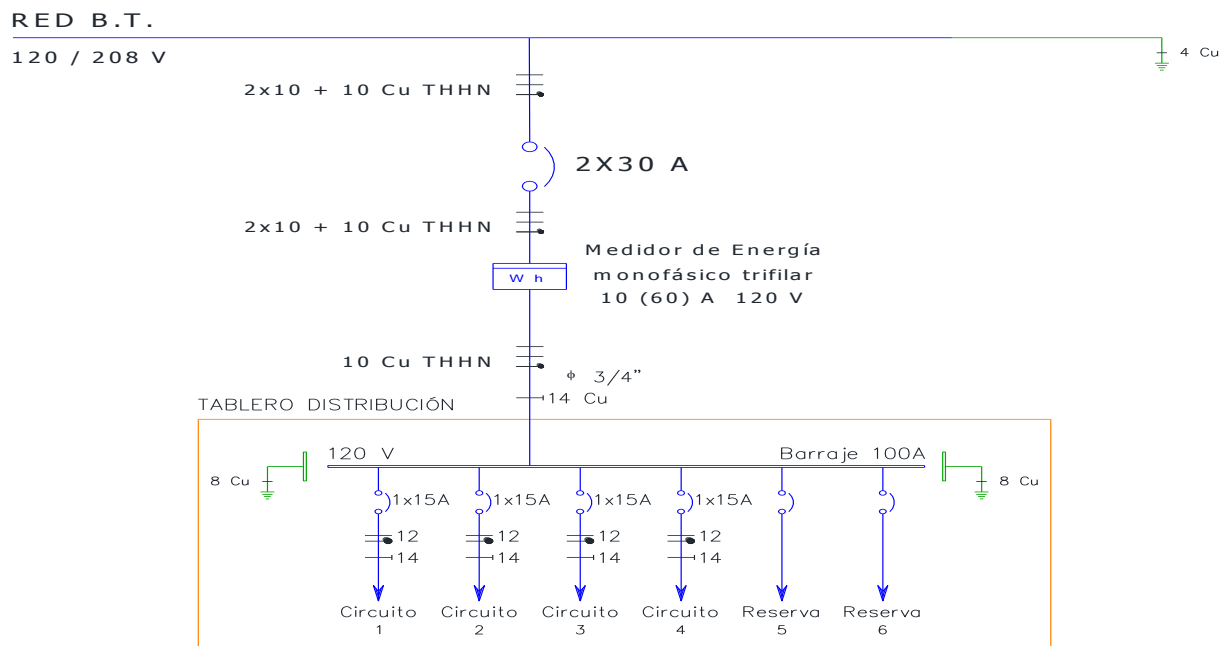


Figura 4 Diagrama unifilar distribución de circuitos ramales laboratorio máquinas eléctricas.

Conocer la distribución exacta de los circuitos ramales en el laboratorio permitiría conectar los aparatos a medir siempre en un circuito lo más independiente posible al circuito que energizaba el medidor, para evitar posibles interferencias.

Para mayor practicidad en la toma de datos, fue necesario el diseño de un multitoma que se acoplara correctamente a las distancias del cableado disponible, no comprometiera directamente las instalaciones del laboratorio, ya que en varias ocasiones es necesario tener acceso al cableado de energiza los

electrodomésticos, y que además brindara seguridad a los elementos de medida como a los electrodomésticos disponibles. En la Figura 5 se tiene un plano eléctrico de cómo está conformado el multitoma.

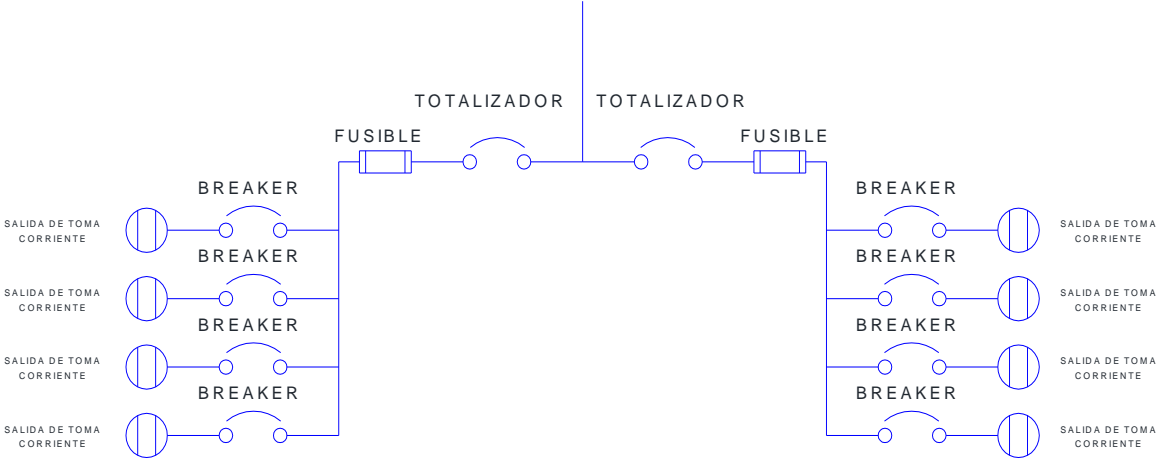


Figura 5 Plano eléctrico multitoma.

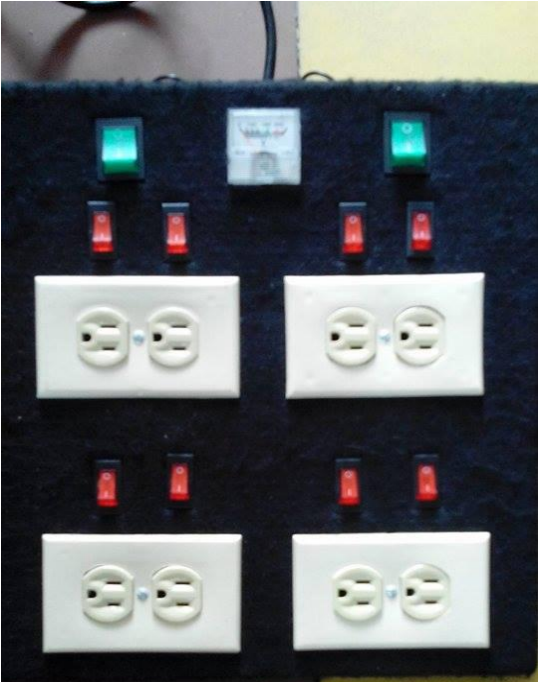


Figura 6 Multitoma físico



Figura 7 Puertos de conexiones del multitoma.

Con todos los elementos disponibles para realizar el montaje del multitoma y el medidor, que como ya se mencionó anteriormente, cada uno de ellos conectados a circuitos diferentes del tablero de distribución. El diagrama de la Figura 6 representada gráficamente la conexión.

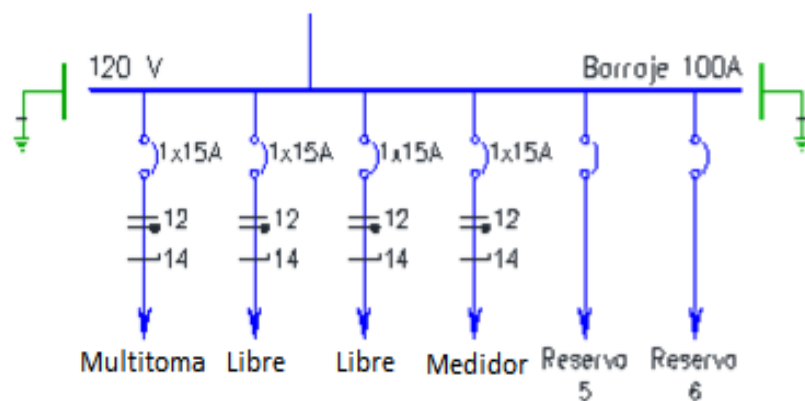


Figura 8 Diagrama conexiones del multitoma y medidor.

Contando con el montaje disponible para iniciar las mediciones, se prosigió a realizar las conexiones totales entre el multitoma, el analizador de redes y cada uno de los elementos a medir, donde se toma la señal de tensión de la entrada principal del multitoma y la señal de corriente desde el puerto totalizador ubicado

al respaldo del multitoma (ver Figura 7), logrando así obtener la intensidad de corriente total en pruebas con múltiples electrodomésticos.

Además de las descripciones aquí presentadas, el Anexo 2 muestra una guía rápida de cómo utilizar el conjunto de datos y algunos elementos relevantes de la toma de datos.

### 3. PROCESAMIENTO Y ANÁLISIS DE DATOS.

Para el análisis final de los datos fueron necesarias tres etapas posteriores a la recolección de datos, las cuales se representan en la Figura 9.

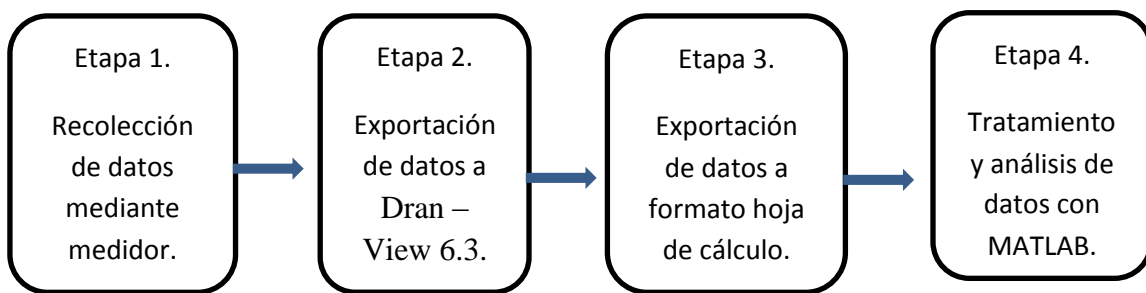


Figura 9 Esquema general del tratamiento de datos

#### 3.1 Pre-procesamiento de las señales

En la etapa tres (datos en hoja de cálculo) se tenía el conjunto de señales completas, transitorios junto con estado estable, con todos los encendidos y apagados juntos, de modo que para el análisis que se deseaba realizar era necesario tener solamente los transitorios de cada encendido y apagado, además que separables unos de los otros. Para tal fin se diseñó un código en MATLAB (etapa 4 de la Figura 9) que lograra, no solamente la identificación y segmentación de los transitorios, sino también el filtrado de las señales obtenidas, dado que se deseaba analizar el transitorio puro, de modo que el filtro eliminaba los aportes de la frecuencia fundamental y todos sus múltiplos, hasta el armónico 80. Finalmente se realizó un corte más fino del transitorio, el cual se almacenaba continuamente para su posterior estudio.

En la Tabla 2 se muestra el total de transitorios de los encendidos y apagados, plenamente identificados, con los que se contó (en total 847) para el análisis de las diferentes propiedades de cada elemento incluido en el conjunto de datos.

	<b>Datos on</b>	<b>Datos off</b>
bombillo licuadora	84	84
bombillo ventilador	41	41
fluorescente	35	35
plancha	22	22
ventilador bombillo	40	40
ventilador licuadora	41	40
ventilador vel1	40	40
ventilador vel3	40	40
bombillo	41	41
licuadora	40	40
<b>Total</b>	424	423

Tabla 2 Conteo de encendidos y apagados por elemento.

Inicialmente se planteó trabajar el análisis para las señales de tensión y de corriente, pero al segmentar las señales transitorias de tensión no se visualizaba cambio alguno al encenderse o apagarse los equipos, dado que la frecuencia de muestreo so es suficientemente alta (ver Figura 10), de este modo, que al no encontrar información relevante en dicha señal, se centró todo análisis en los transitorios de las señales de corriente, en los cuales se observó un notorio cambio (ver Figura 11), lo cual facilitaba la extracción de características para su posterior análisis.

Transitorio de Tensión

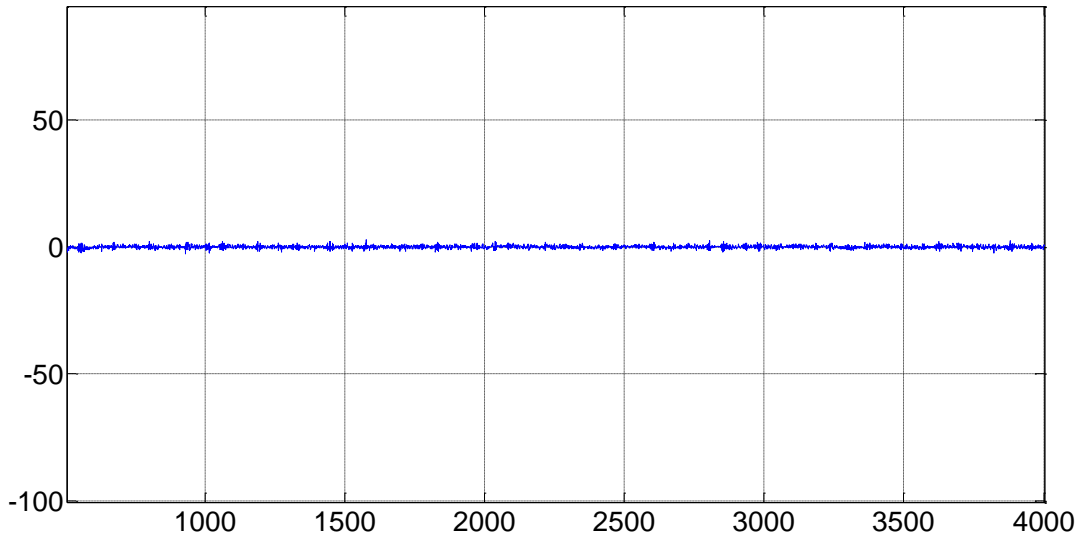


Figura 10 Transitorio de tensión en encendido del ventilador en velocidad uno

Transitorio de Corriente

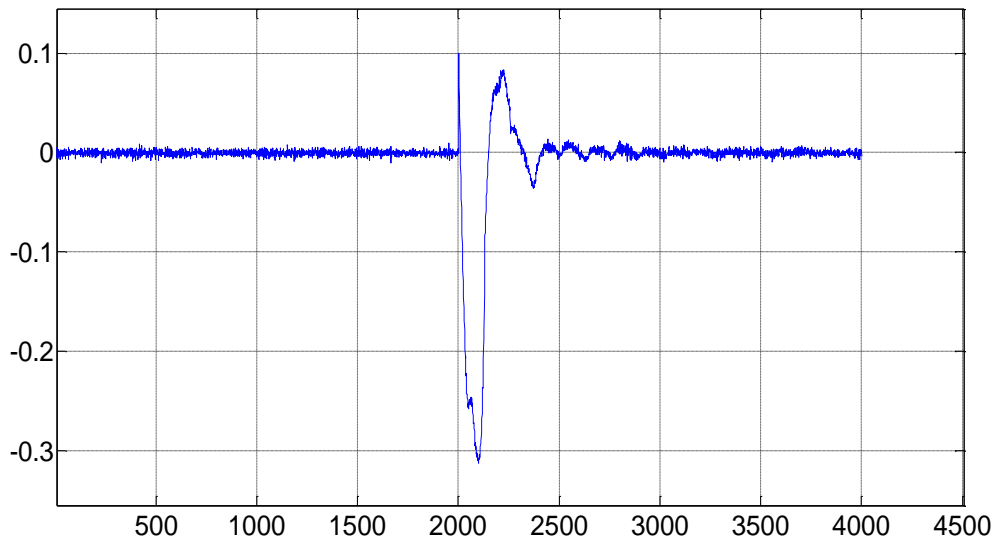


Figura 11 Transitorio de corriente en encendido del ventilador en velocidad uno.

### 3.2 Procesamiento de señales.

Para iniciar el análisis se extrajeron las siguientes características, en el dominio del tiempo, para cada uno de los transitorios de corriente:

- Desviación estándar:

Se operó mediante el comando “std” de la herramienta MATLAB, la cual define esta como [16]:

$$s = \left( \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{\frac{1}{2}}$$

Dónde:

- s = desviación estándar.
  - n = número de datos.
  - $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .
- Factor de cresta:

Se halló mediante los siguientes comandos de MATLAB:  $\max(\text{abs}(x))/((\text{sum}(x.^2)/\text{length}(x))^{\wedge}0.5)$ , lo que se puede interpretar como valor pico, dividido en el valor RMS.

- Entropía:
- Curtosis:

Se trabajó con el comando “kurtosis” de la herramienta MATLAB, este la calcula mediante la función [16]:

$$k_i = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})^4}{\left( \frac{1}{n} \sum_i (x_i - \bar{x})^2 \right)^2}$$

- $k_i$  = curtosis.
- n = número de datos.
- $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .

- Potencia

Se operó mediante los comandos de MATLAB:  $\text{sum}(x.^2)/\text{length}(x)$ , lo que se interpreta como la energía por unidad de tiempo.

El estudio de cada una de las características extraídas, se llevó a cabo mediante cuatro diferentes técnicas de análisis:

- Gráfica de pares de características:

Se realizó la combinación, en grupos de a dos, de las cinco características extraídas de los transitorios (Desviación estándar - Factor de cresta, Desviación estándar – Entropía, Desviación estándar – Curtosis, etc.) logrando un grupo total de 12 diferentes conjuntos de parejas de datos, los cuales fueron graficados con el fin de identificar gráficamente agrupaciones de cada uno de los electrodomésticos.

- *Clustering*:

Se basa en intentar responder cómo es que ciertos objetos (casos) pertenecen naturalmente a cierto número de clases o grupos, de tal manera que estos objetos comparten ciertas características.

Para el caso de clustering mediante la función “kmeans” de MATLAB, se encarga de encontrar medidas de disimilitud (similitud), por medio de distancias euclidianas, tomando como referencia el punto medio de cada “k” grupo [16].

- Análisis de Componentes Principales (PCA) por sus siglas en inglés, y posterior uso de *clustering*:

PCA es un método no paramétrico que utiliza medidas estadísticas simples (varianza y covarianza), junto con algoritmos de álgebra lineal, proporcionando una ruta para la representación de un conjunto de datos con una dimensión menor, pero que en dicha representación este presente la información de mayor relevancia [17].

- Análisis de Discriminación Lineal (LDA) por sus siglas en inglés, y posterior uso de *clustering*:

LDA es un método paramétrico que busca, al igual que PCA, disminuir la dimensionalidad de un determinado conjunto de datos, pero esta busca además conservar la mayor cantidad de información discriminante posible, para ello utiliza una proyección de orden  $C-1$  clases donde  $C$  son las clases totales, enfocándose en que estas proyecciones estén lo más separadas posibles, y que a su vez en ellas, los casos de una misma clase se encuentre muy cercanos uno del otro [18].

### **3.3 Análisis de gráficas.**

A continuación se presentan el conjunto de gráficas resultantes de las diferentes combinaciones de características extraídas a los transitorios (Gráficas 1-20). En la Tabla 3 se presenta el código de colores utilizado para representar cada elemento.

Las combinaciones de dos elementos se tomaron como un elemento individual y diferente, además las pruebas realizadas con dos elementos a la vez, consistieron en realizarle encendidos y apagados a solo uno de ellos, mientras el otro elemento permanecía en funcionamiento continuo. La nomenclatura utilizada ubica al elemento en funcionamiento continuo al inicio, y al que realiza la intermitencia al final. Siempre que se combinaba el ventilador con algún otro elemento, se mantenía el ventilador operando en la velocidad 1. Por sus características de funcionamiento, el ventilador se operó y supervisó, en dos de sus diferentes velocidades de funcionamiento.

Para cada una de las siguientes gráficas se utilizó la función “scatter” de la herramienta MATLAB, junto con el comando “hold on” para lograr ubicar en una sola gráfica el conjunto de datos totales. El fin de estas gráficas es realizar una evaluación visual de los diferentes electrodomésticos utilizados, y así tener un primer tamiz de las características que mejor podrían diferenciar un elemento del otro.

Elemento	Representación
bombillo licuadora	○ (rojo)
bombillo ventilador	○ (amarillo)
fluorescente	○ (magenta)
plancha	○ (cian)
ventilador bombillo	○ (verde)
ventilador licuadora	○ (azul)
ventilador vel1	○ (negro)
ventilador vel3	● (amarillo)
bombillo	● (rojo)
licuadora	● (azul)

Tabla 3 Código de colores utilizado en las figuras 12-31

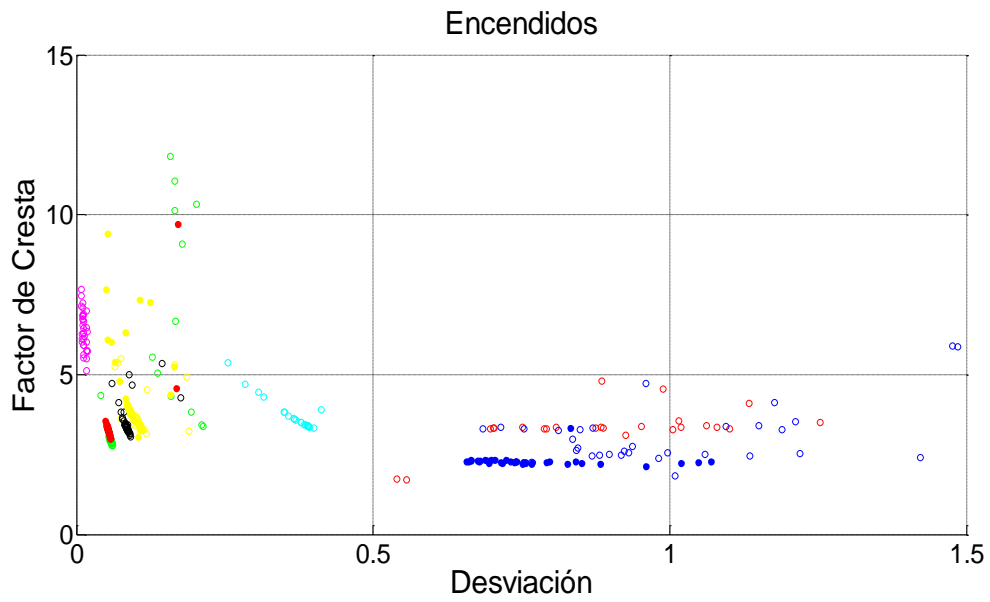


Figura 12 Factor de Cresta Vs. Desviación – encendido.

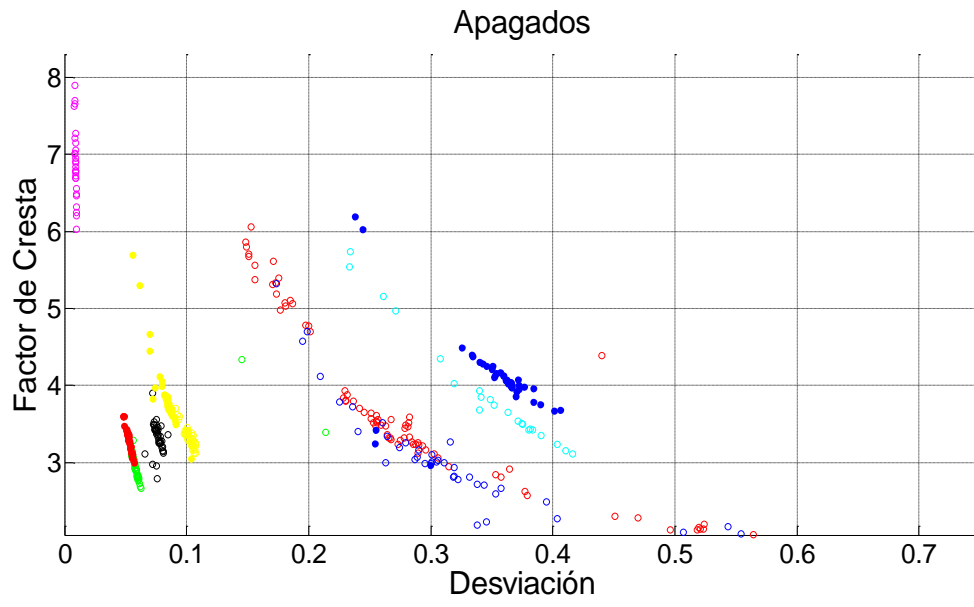


Figura 13 Factor de Cresta Vs. Desviación – apagados.

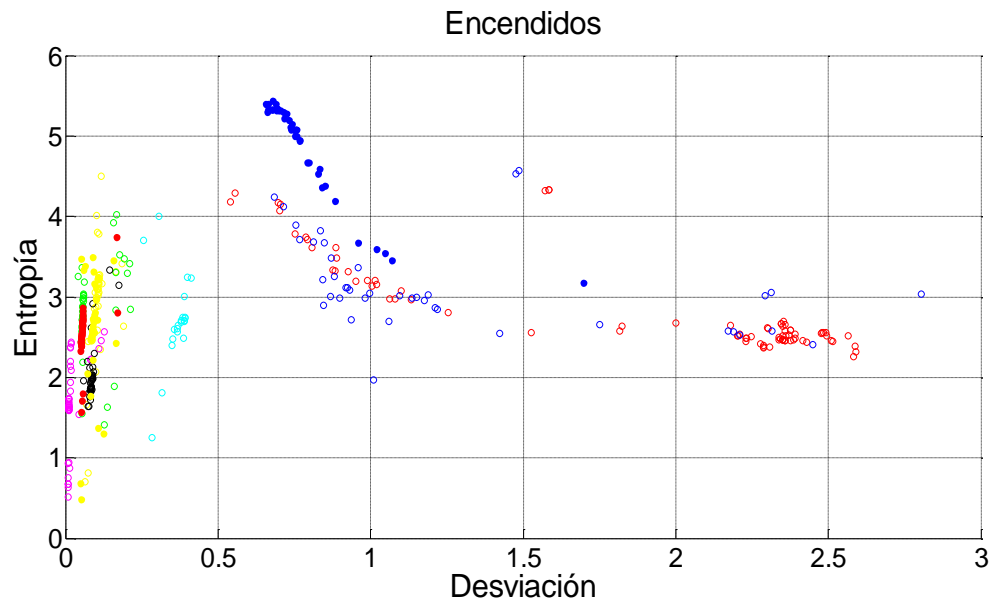


Figura 14 Entropía Vs. Desviación – encendidos.

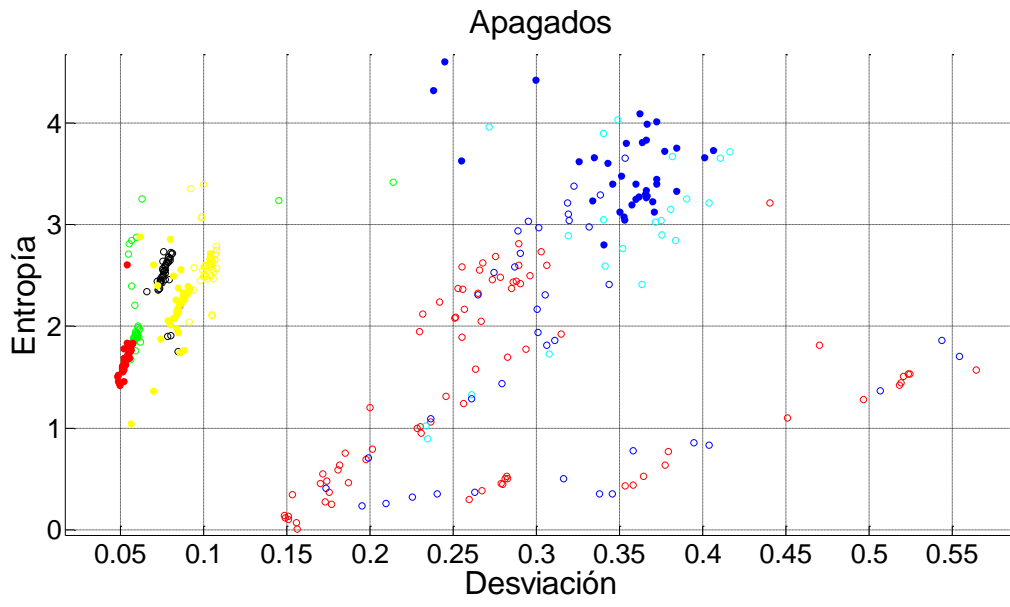


Figura 15 Entropía Vs. Desviación – apagados.

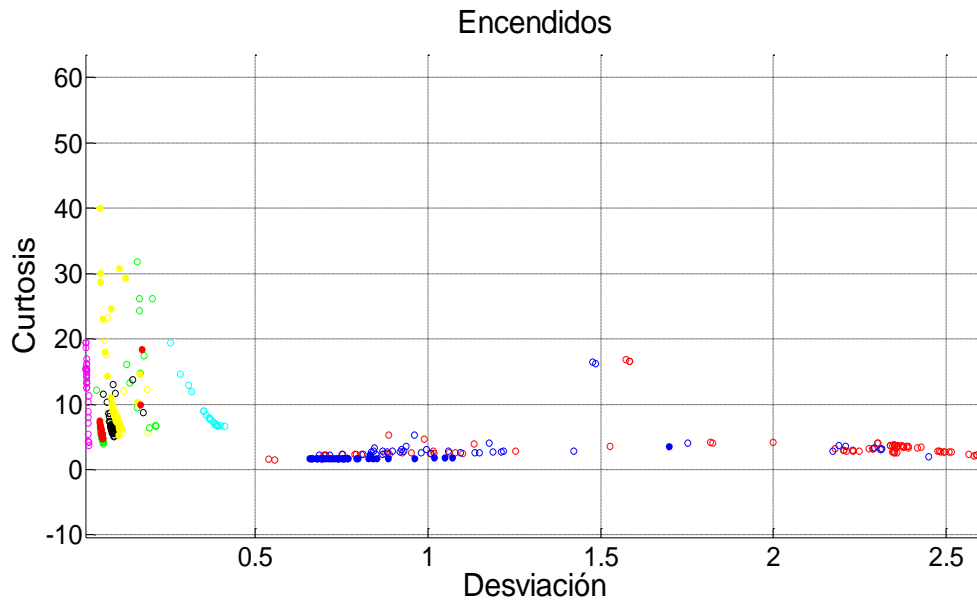


Figura 16 Curtosis Vs. Desviación – encendidos.

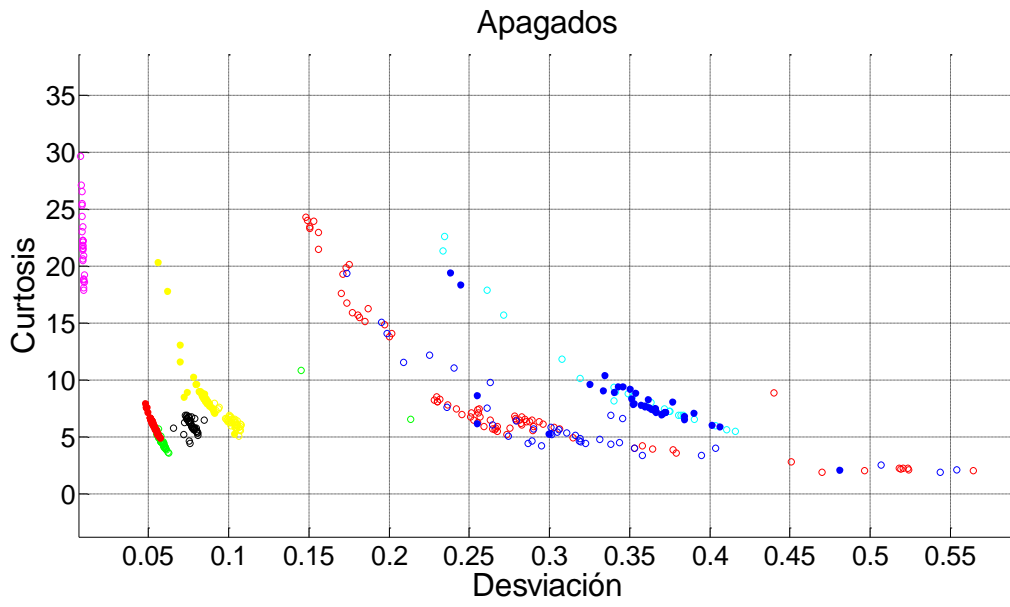


Figura 17 Kurtosis Vs. Desviación – apagados.

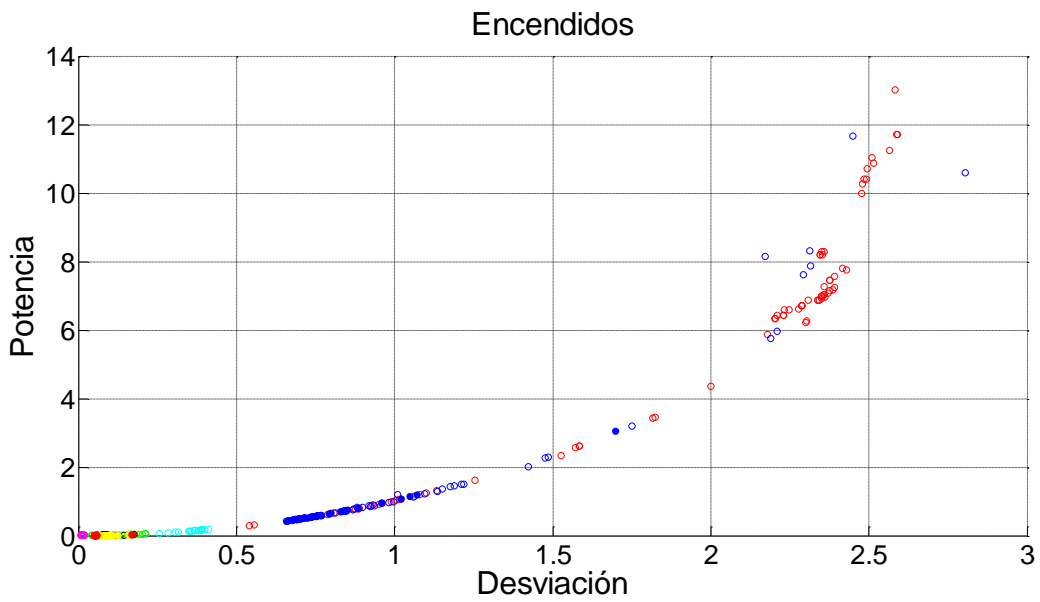


Figura 18 Potencia Vs. Desviación – encendidos

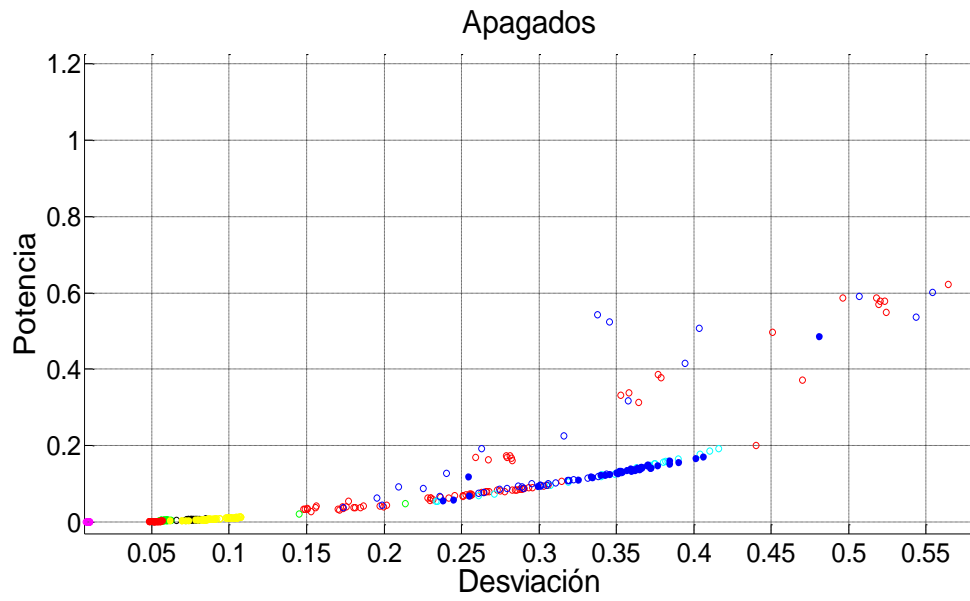


Figura 19 Potencia Vs. Desviación – apagados.

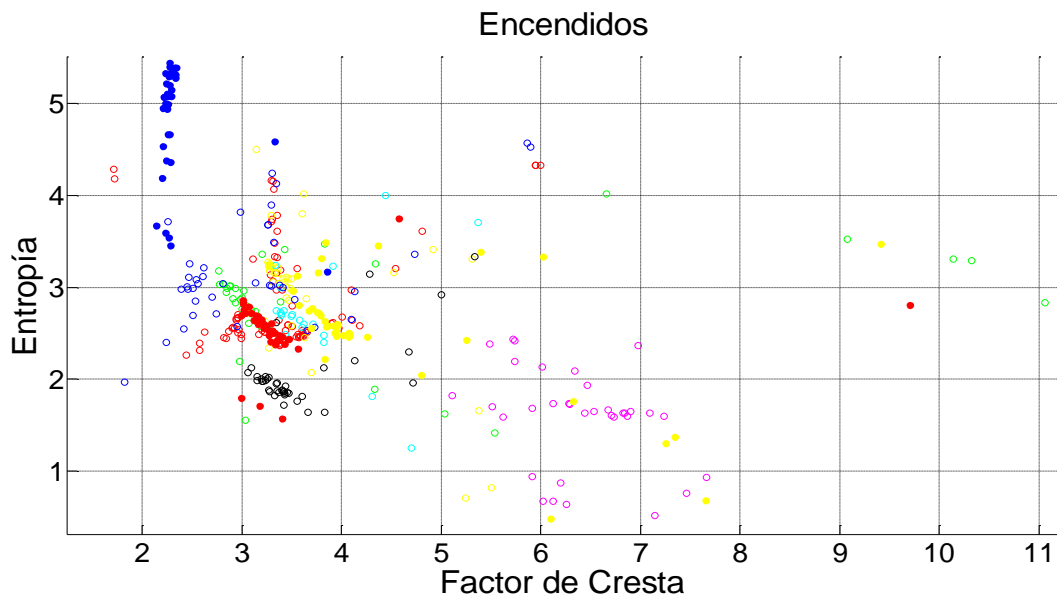


Figura 20 Entropía Vs. Factor de Cresta – encendidos.

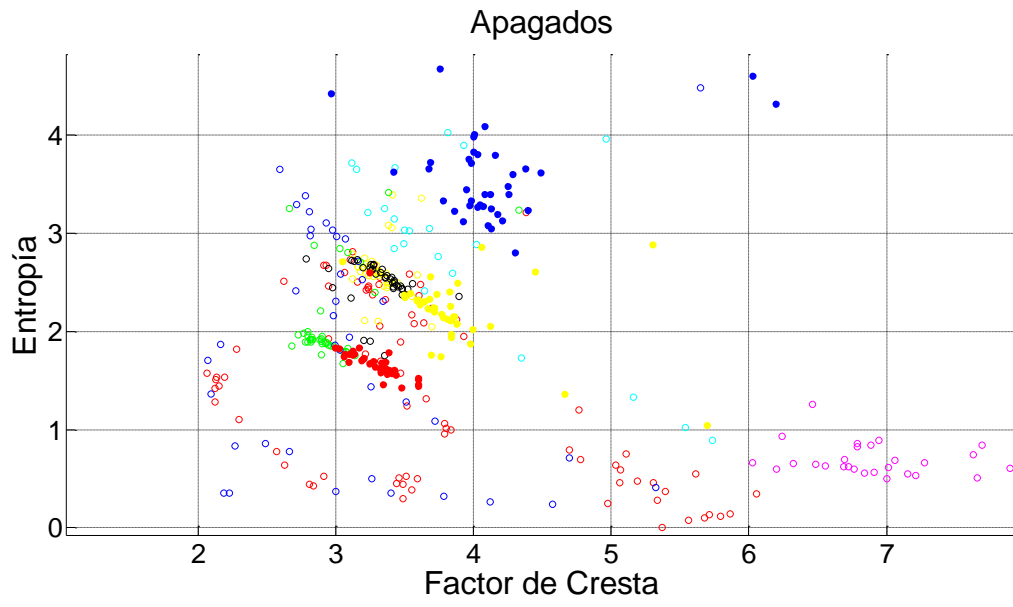


Figura 21 Entropía Vs. Factor de Cresta – apagados.

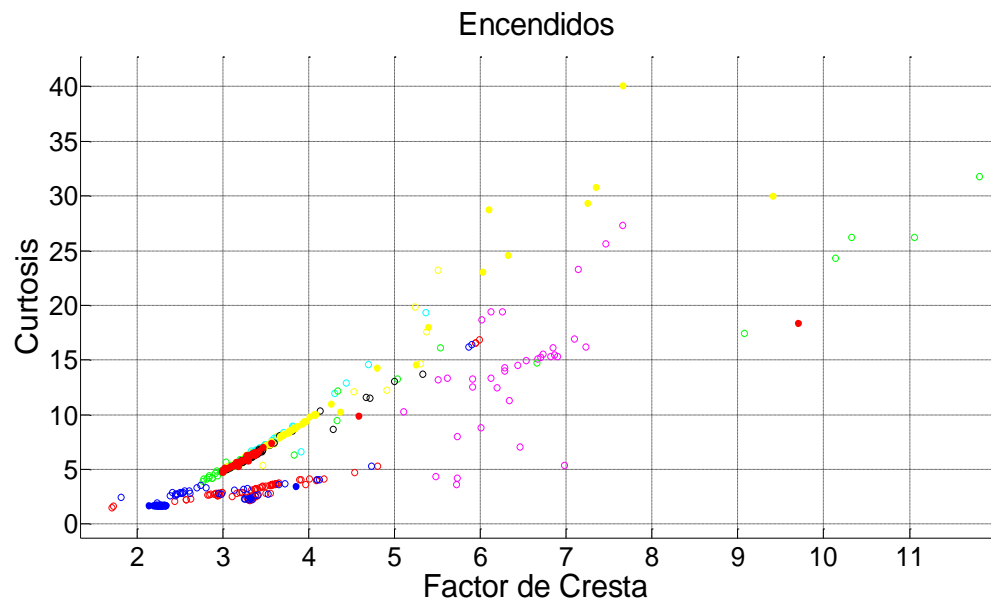


Figura 22 Curtosis Vs. Factor de Cresta – encendidos.

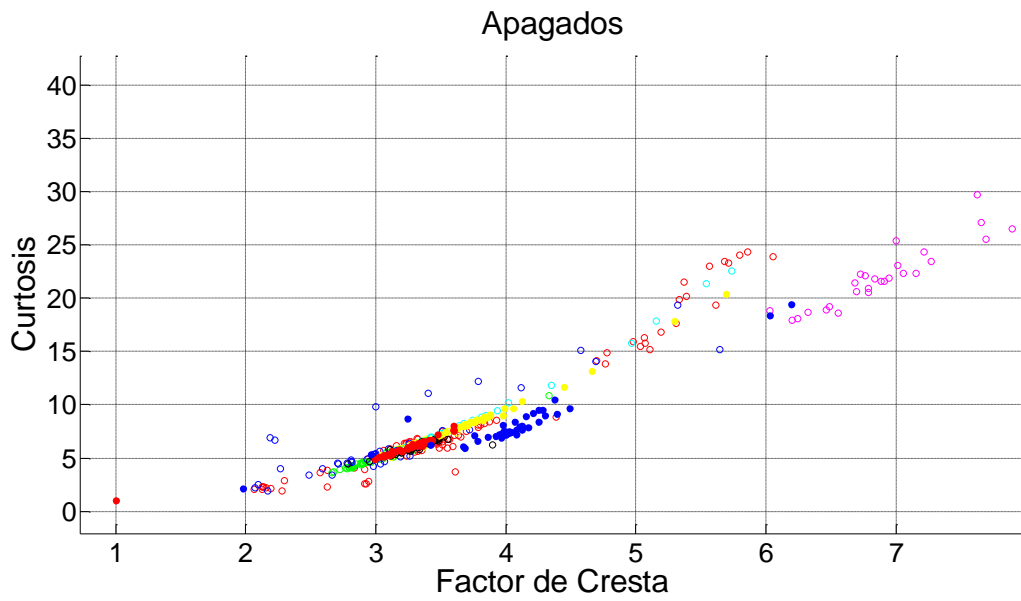


Figura 23 Kurtosis Vs. Factor de Cresta – apagados.

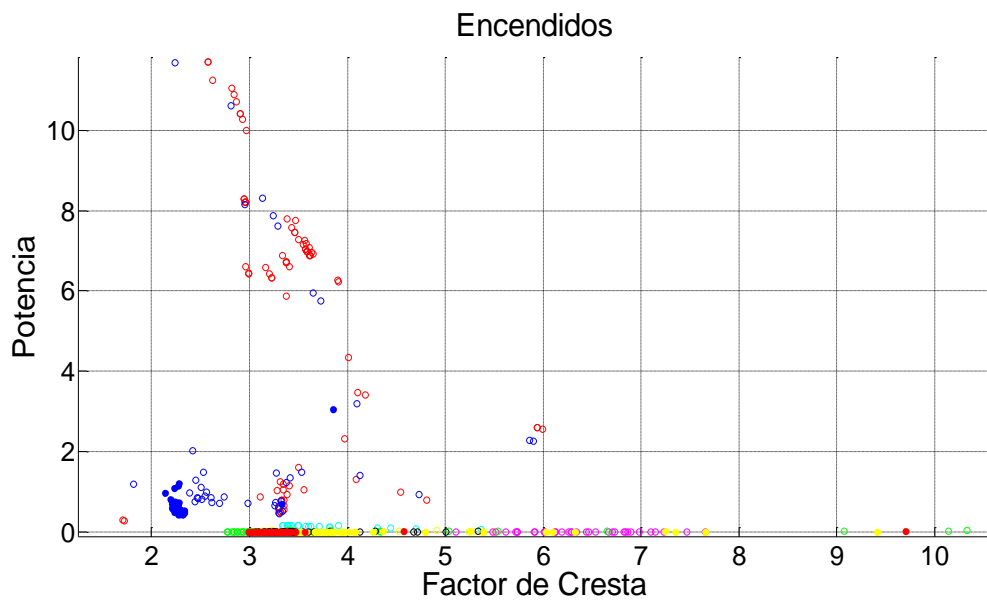


Figura 24 Potencia Vs. Factor de Cresta – encendidos.

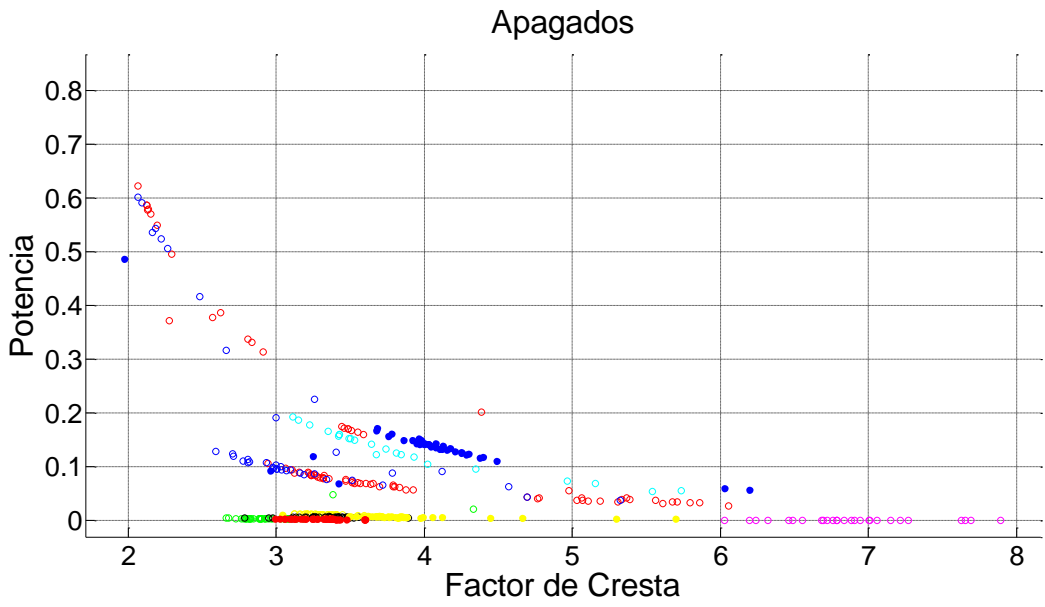


Figura 25 Potencia Vs. Factor de Cresta – pagados.

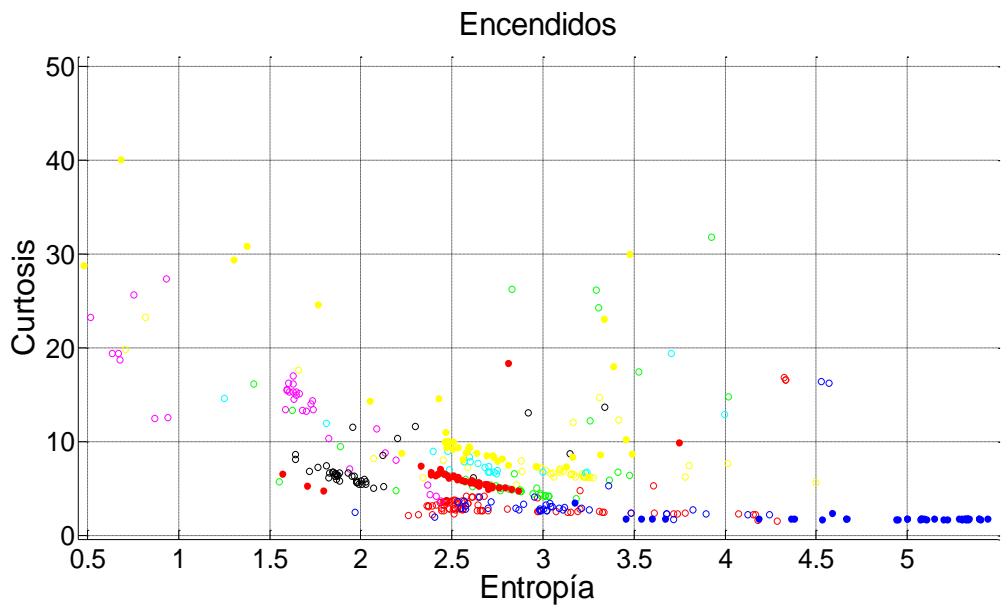


Figura 26 Curtosis Vs. Entropía – Encendidos.

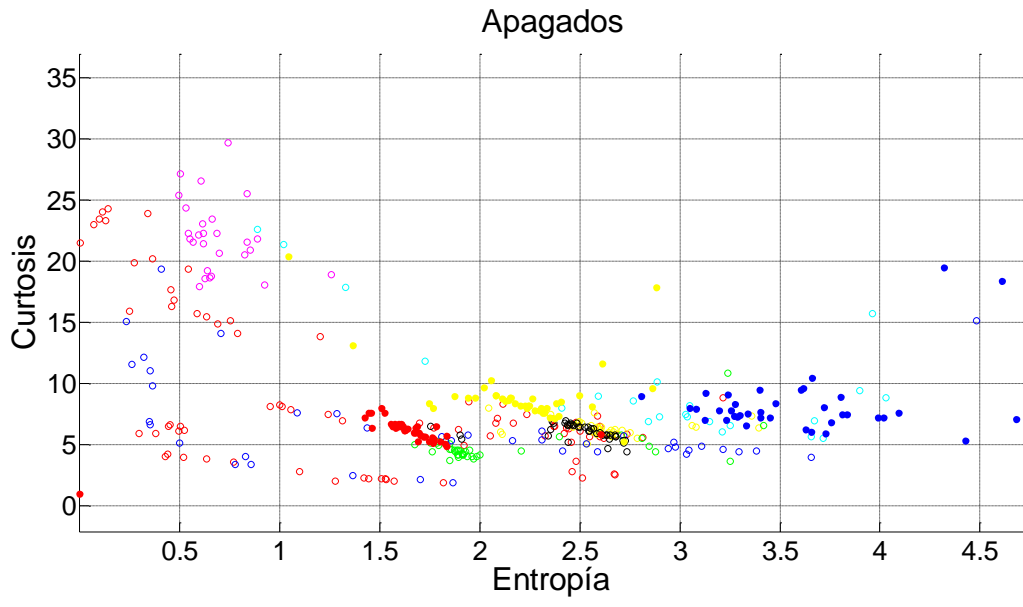


Figura 27 Kurtosis Vs. Entropía – Apagados.

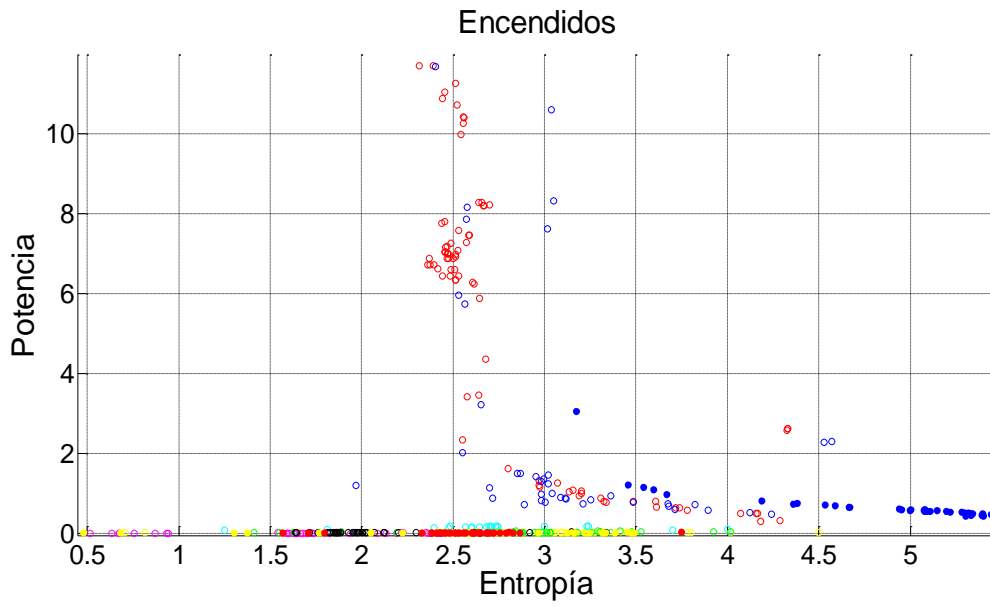


Figura 28 Potencia Vs. Entropía – Encendidos.

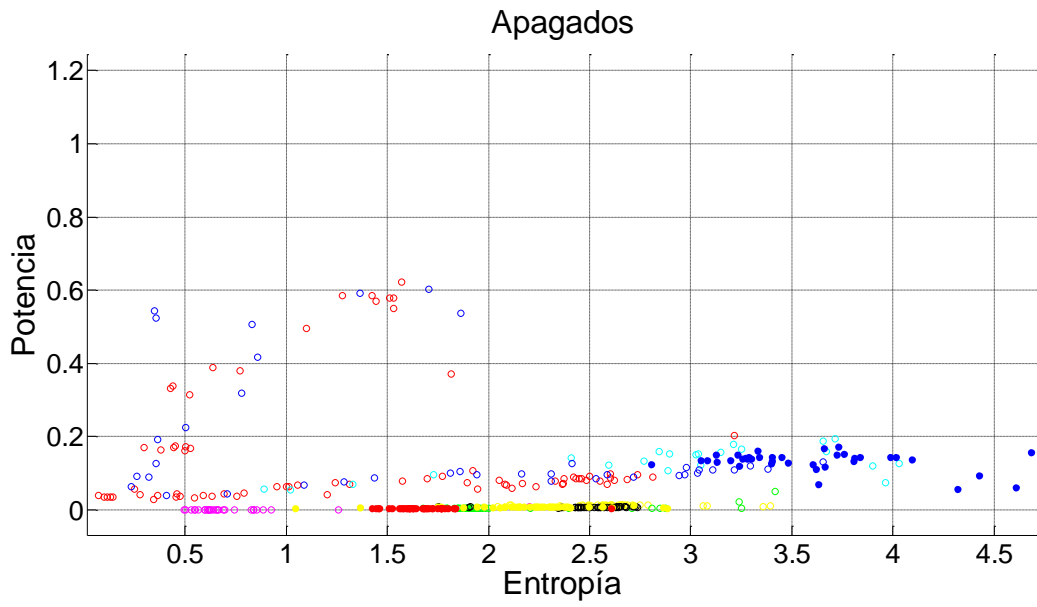


Figura 29 Potencia Vs. Entropía – Apagados.

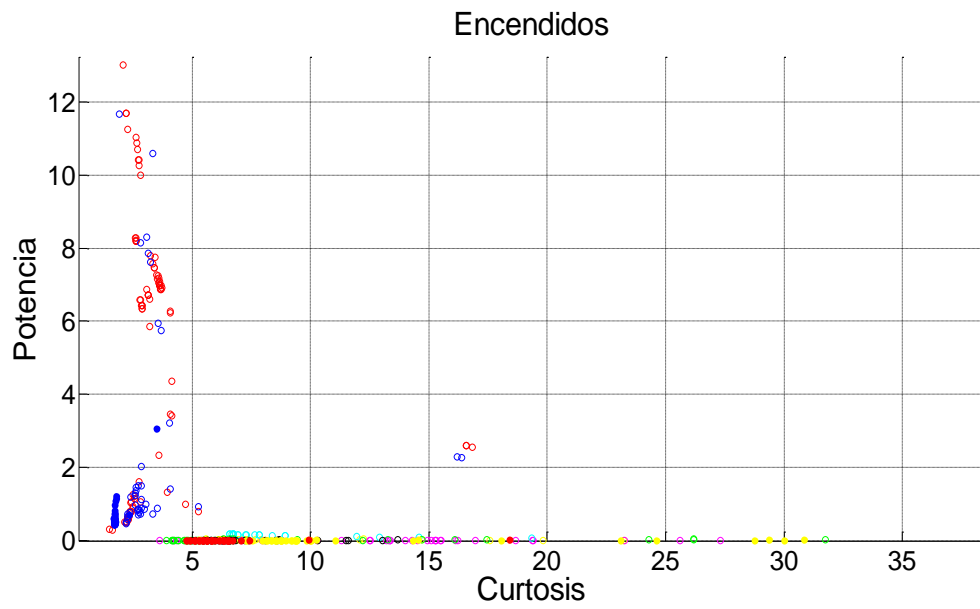


Figura 30 Potencia Vs. Curtosis – Encendidos.

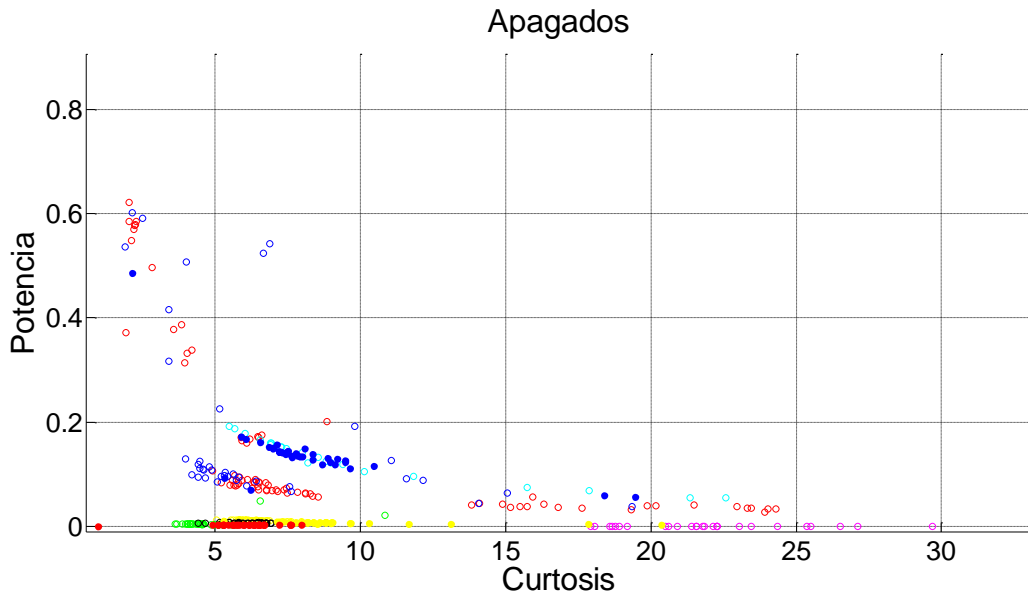


Figura 31 Potencia Vs. Curtosis – Apagados.

Aunque algunas de las combinaciones (Gráficas 12, 13, 16, 17, 24 y 25) brindan claras separaciones para algunos de los elementos (algunas más que otras), en general no son lo suficientemente diferenciables. Además que no es sencillo hacer un análisis del significado de cada gráfica, dado que las características con las se trabajó, las cuales se encuentran en los ejes de las gráficas, no tiene una manera clara de relacionarse, o como tal una representación espacial definida.

Como apoyo a estas gráficas se tienen los siguientes métodos de análisis, los cuales en comparación a estas, son medibles en su precisión y su eficiencia en desagregación, de modo tal que se pueda presentar resultados puntuales de la clasificación.

Para cada una de las técnicas de análisis a continuación presentadas, se extrajo la precisión con la que cada método agrupo cada electrodoméstico, en una determinada clase, mediante la siguiente ecuación:

$$Precisión = \frac{\text{número de eventos agrupados}}{\text{número de eventos totales}}$$

Además se obtuvo la tasa de desagregación total del método de análisis como:

$$Tasa\ de\ Desagregación = \frac{cantidad\ de\ grupos\ con\ un\ solo\ elemento}{cantidad\ de\ elementos\ totales}$$

Se entiende por elemento, cada uno de los electrodomésticos, y por evento cada uno de los datos transitorio de cada elemento, obtenidos del muestreo realizado y posterior segmentación de la señal.

### **3.3.1 Análisis *clustering* directo.**

El *clustering* se llevó a cabo con el función “kmeans” de MATLAB, la cual funciona como una clasificación no supervisada de datos, clasificando cada evento mediante iteraciones de distancias euclidianas a cada uno de los centroides de los grupos, y seleccionando la menor de estas [19].

En la Tabla 4 se observan las diferentes agrupaciones logradas por este algoritmo con sus respectivas precisiones. De la Tabla 4 se tiene que la mayor precisión se presenta para el encendido de la licuadora con una precisión del 85%, lo cual es una gran precisión para un algoritmo de selección, además se puede observar que es plenamente identificable entre las otras muestras, dado que este es el único evento agrupado en la clase uno. Por otra parte si se analiza el apagado de la licuadora, el cual también aporta una muy buena precisión al algoritmo (55%), es el único evento ubicado en el grupo 14, de modo que este método resulta bastante eficiente para la licuadora.

En general el método de *clustering*, con el conjunto de características totales, tiene un promedio de precisión de 49,07%, y como se puede apreciar en la Figura 32, una buena tasa de desagregación, dado que el 35% de los eventos fueron diferenciados completamente y la máxima repetición fue de tres elementos diferentes en un mismo grupo.

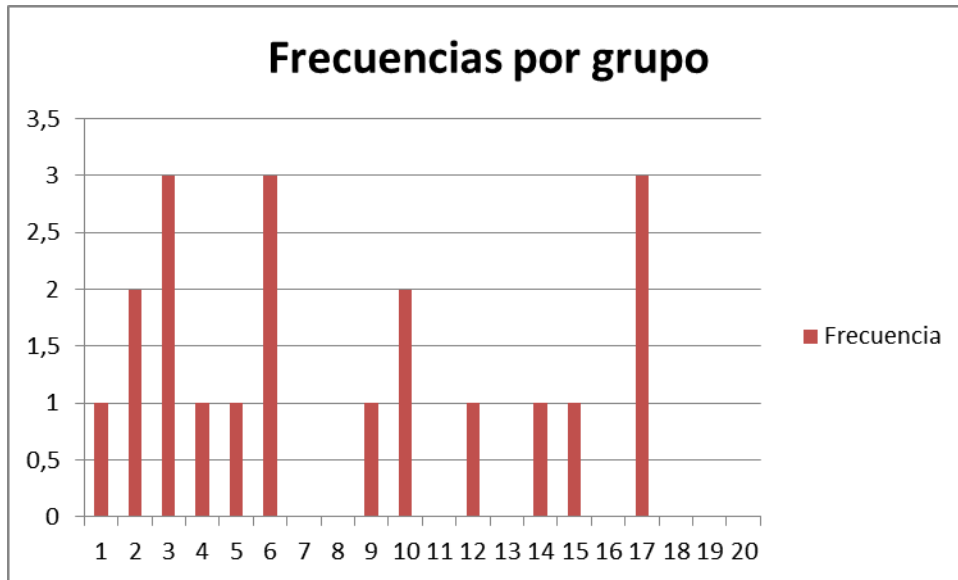


Figura 32 Frecuencias por grupo con clustering directo.

<i>clustering (kmeans)</i>			
		grupo	precision
on	bombillo licuadora	15	62,65%
off	bombillo licuadora	17	19,28%
on	bombillo ventilador	12	41,46%
off	bombillo ventilador	10	51,22%
on	fluorescente	5	27,50%
off	fluorescente	9	60,71%
on	plancha	6	57,14%
off	plancha	6	22,73%
on	ventilador bombillo	3	42,50%
off	ventilador bombillo	3	70,00%
on	ventilador licuadora	4	46,34%
off	ventilador licuadora	3	27,50%
on	ventilador vel1	17	40,00%
off	ventilador vel1	6	45,00%
on	ventilador vel3	2	50,00%
off	ventilador vel3	2	70,00%
on	bombillo	10	51,22%
off	bombillo	17	56,10%
on	licuadora	1	85,00%
off	licuadora	14	55,00%

Tabla 4 Resultados del clustering.

### 3.3.2 Análisis de Componentes Principales (PCA).

Para el PCA se operó el comando “princomp” de MATLAB, el cual realiza el análisis de componentes principales, mediante el uso de varianzas y covarianzas, encontrando las componentes principales de las etiquetas y calculando el peso que cada una de ellas representa [19].

Al examinar la Tabla 5 se encuentra que la mayor precisión se presenta nuevamente para el encendido de la licuadora, con una precisión del 95%, y al igual que el caso anterior es plenamente diferenciable de los demás, puesto que este es el único elemento agrupado en la clase 19. Uniendo esta información con el apagado de la licuadora, que aporta una precisión al algoritmo de 42,5% y es el único elemento ubicado en el grupo 18, se concluye que el algoritmo identifica claramente la licuadora, tanto en su encendido como en su apagado.

En general al aplicar PCA antes del *clustering*, se tiene un promedio de precisión de 36,42%, observable en la Figura 34, y una tasa de desagregación del 45%, 10% mejor que *clustering* directo, pero tiene la deficiencia que alcanzo a agrupar 7 elementos diferentes en una misma clase. Vale la pena aclarar que comparando la clasificación realizada utilizando el PCA, y realizando el *clustering* directamente, se tiene 12,64%, en promedio, mejor precisión para el caso en que se aplica el *clustering* directamente.

PCA y posterior clustering			
		grupo	precision
on	bombillo licuadora	17	32,53%
off	bombillo licuadora	4	15,66%
on	bombillo ventilador	4	29,27%
off	bombillo ventilador	20	39,02%
on	fluorescente	9	32,50%
off	fluorescente	3	53,57%
on	plancha	4	42,86%
off	plancha	4	18,18%
on	ventilador bombillo	16	20,00%
off	ventilador bombillo	7	32,50%
on	ventilador licuadora	12	46,34%
off	ventilador licuadora	16	22,50%
on	ventilador vel1	4	22,50%
off	ventilador vel1	4	37,50%
on	ventilador vel3	6	35,00%
off	ventilador vel3	6	52,50%
on	bombillo	5	31,71%
off	bombillo	4	26,83%
on	licuadora	19	95,00%
off	licuadora	18	42,50%

Tabla 5 Resultados del clustering después de PCA.



Figura 33 Frecuencias por grupo con PCA y posterior clustering.

### 3.3.3 Análisis de Discriminación Lineal (LDA).

Para el LDA se utilizó el código realizado por Deng Cai [20], el cual realiza el análisis de discriminación lineal, donde se ingresan las respectivas etiquetas de las clases, junto con la matriz cuyas columnas son las características extraídas, y las filas son el conjunto de puntos o eventos captados, para finalmente obtener como resultado las proyecciones de características más adecuadas para la diferenciación de datos.

Observando la Tabla 6 se tiene que la mayor precisión se presenta para el apagado del fluorescente, con una precisión del 96,43%, y la mínima se presentó con el apagado de la plancha (31,82%). Observando la discriminación entre grupos (Figura 34), se tiene una tasa de desagregación del 45%, lo cual es mejor que los dos casos anteriores. En total, al aplicar LDA a los datos y posteriormente *clustering* se tiene un promedio de precisión de 65,22%.

LDA y posterior <i>clustering</i>			
		grupo	precision
on	bombillo licuadora	6	59,04%
off	bombillo licuadora	7	39,76%
on	bombillo ventilador	9	63,41%
off	bombillo ventilador	17	78,05%
on	fluorescente	10	55,00%
off	fluorescente	14	96,43%
on	plancha	8	71,43%
off	plancha	3	31,82%
on	ventilador bombillo	17	57,50%
off	ventilador bombillo	18	80,00%
on	ventilador licuadora	19	39,02%
off	ventilador licuadora	11	30,00%
on	ventilador vel1	18	80,00%
off	ventilador vel1	17	87,50%
on	ventilador vel3	3	52,50%
off	ventilador vel3	8	65,00%
on	bombillo	17	87,80%
off	bombillo	18	95,12%
on	licuadora	13	82,50%
off	licuadora	4	52,50%

Tabla 6 Resultados del clustering después de LDA.

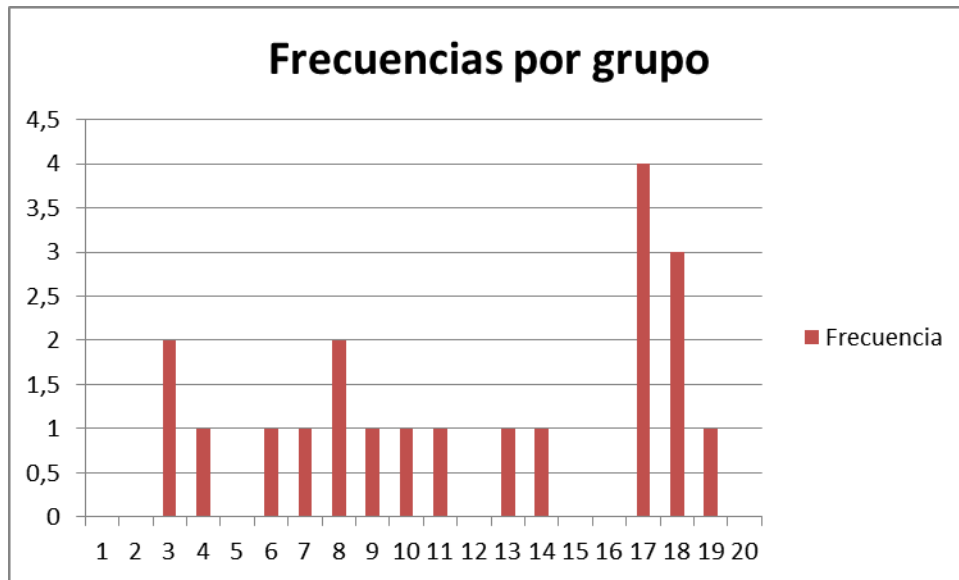


Figura 34 Frecuencias por grupo con LDA y posterior clustering.

### 3.4 Comparación métodos de análisis.

En la Figura 35 se encuentra representados los resultados finales de los tres métodos utilizados para el análisis de los transitorios. La Figura 35 muestra la gran ventaja, en cuanto a la precisión de selección, que obtuvo el método de *clustering* con LDA, sobrepasando al *clustering* con PCA y al *clustering* directo en 28,80% y 16,15% respectivamente. Además el *clustering* con LDA y PCA lograron una tasa de desagregación de 45%, lo cual es 10% mejor que el *clustering* directo..

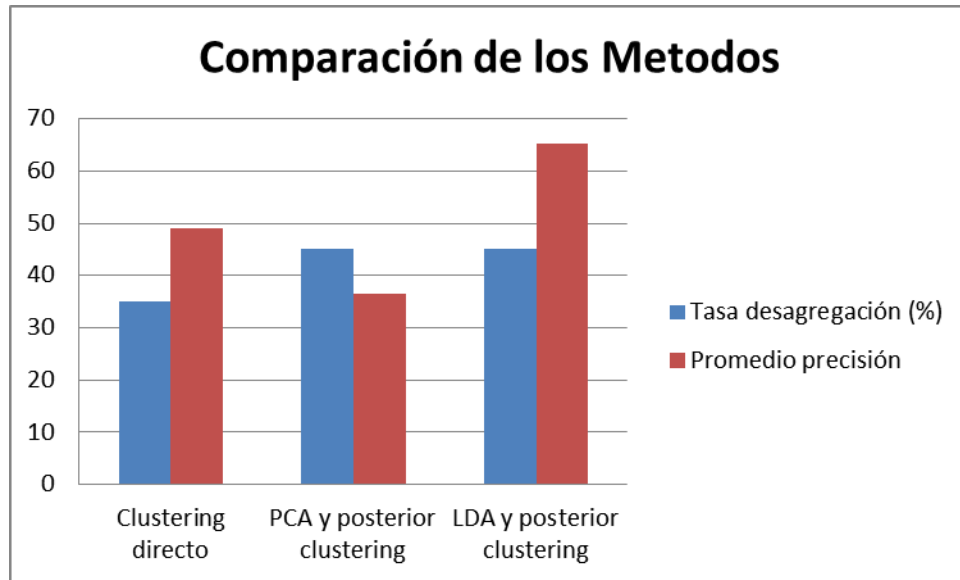


Figura 35 Resultados finales de los tres análisis.

#### 4. CONCLUSIONES.

- Se presentó un conjunto de datos totalmente etiquetados, especificando el tipo de evento sucedido para cada una de las etiquetas, con el fin de suministrar información para estudios NILM, específicamente en el área de prueba de algoritmos de desagregación.
- De los analizadores de redes disponibles para el grupo GISEL, el adecuado para llevar a cabo la recolección de datos fue el Power Visa DRANETZ – BMI 1000 con su frecuencia de muestreo de 15,4 kHz y los ocho canales disponibles, cuatro para tensión y cuatro para corriente.
- Antes de iniciar una toma de datos, es de gran importancia cubrir la mayor cantidad de imprevistos que puedan presentarse, con el fin de garantizar la confiabilidad de los datos a adquirir. Entre la gran cantidad de sucesos posibles, se pueden resaltar los siguientes:
  - Disponibilidad completa de los equipos a utilizar.
  - El buen estado de la totalidad de los equipos a trabajar.
  - Disponibilidad del sitio destinado para la toma de datos.
  - Continuidad y periodicidad en la toma de datos.

- Verificación y continua revisión, de la lista de elementos que intervienen directamente en la toma de datos.
- Es de gran importancia el continuo aporte a la comunidad de investigación, por parte de la universidad, y si estos son por medio de proyectos de grado, estos están representando un importante avance en la participación de la universidad en este ámbito. Para ello es vital que siempre se respalden correctamente los procedimientos realizados para dichos aportes, en nuestro caso, la toma del conjunto de datos, de manera que se contaba con una serie de equipos y un hardware en general, que se estaba manipulando constantemente, así que continuamente se debía revisar el buen estado de las conexiones utilizadas, la calibración de los equipos de medida y los equipos utilizados para la supervisión de sus consumos.
- La descripción de las cargas se enfocó hacia su comportamiento mediante un sistema NILM para la desagregación de cargas. Finalmente para el análisis gráfico, el caso que mejor logró diferenciar todos los grupos fue la gráfica de Desviación Estándar Vs. Factor de Cresta. El elemento que en la mayoría de los casos alcanzo un mayor grado de diferenciación fue la licuadora, lo cual concuerda con los resultados de las Tablas 4, 5 y 6.
- La característica que menos apporto a la desagregación de las cargas fue la entropía, hecho que se refleja en las gráficas 3,4,9,10 y 15-18, de modo que esta característica podría no ser tenida en cuenta para mejorar futuros trabajos de desagregación.
- El método que mejor logró desagregar las cargas fue la aplicación de LDA a las características extraídas de los transitorios, y el posterior uso de *clustering* al conjunto de datos modificado. Este método alcanzó un máximo, mínimo y promedio de precisión en las agrupaciones de 96,43%, 30% y 65,22%, respectivamente, una tasa de desagregación de las cargas de 45% y agrupo un máximo de cuatro cargas en un mismo grupo.

## Referencias bibliográficas:

- [1] (IPCC), Intergovernmental Panel on Climate Change, Climate Change, 2007.
- [2] J. Borstein y K. Blackmore, In-Home Display Units:, March, 2008.
- [3] Ministerio de comercio, Industria y Turismo, Decreto 4175, 2011.
- [4] A. Zoha, A. Gluhak, M. A. Imran y S. Rajasegarar, Non-Intrusive Load Monitoring Approaches for Disaggregated, 2012.
- [5] R. Michael y Z. Kurt, "Nonintrusive Appliance Load Monitoring: Review and Outlook" IEEE Transactions on Consumer Electronics,, Vol 57, No. 1, February 2011.
- [6] L. K. Norfold y S. B. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," Energy and Buildings, Vol. 24, pp. 51-64, 1996..
- [7] A. I. Albicki y A. Cole, "Data Extraction for Effective Non-Intrusive Identification of Residential Power Loads," IEEE Instrumentation and Measurement Technology Conference, pp. 812-815, 1998.
- [8] A. Albicki y A. Cole, "Algorithm for Non-Intrusive Identification of Residential Appliances," Circuits and Systems, Proceedings of the IEEE International Symposium on, Vol. III, pp. 338-341, 1998.
- [9] L. Farinaccio y R. Zmeureanu, "Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses," Energy and Buildings, vol. 30, pp. 245-259, 1999.
- [10] M. Baranski y J. Voss, "Non-Intrusive Appliance Load Monitoring Based on an Optical Sensor," EEE Power Tech Conference,, Bologna, 2003.
- [11] M. Baranski y J. Voss, "Genetic Algorithm for Pattern Detection in NIALM Systems," IEEE International Conference on Systems, Man and Cybernetics, pp. 3462-3468, 2004.
- [12] M. Baranski y J. Voss, "Detecting Patterns of Appliances from Total Load Data Using a Dynamic Programming Approach," Fourth IEEE International Conference on Data Mining (ICDM'04), 2004.

- [13] Z. Kolter y M. Johnson, REDD: A public data set for energy disaggregation research, In proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011.
- [14] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe y M. Berges, "BLUED:A FullyLabeledPublicDatasetforEvent-Based Non-Intrusive Load MonitoringResearch" in Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability, Beijing, China,, 2012.
- [15] Universidad Nacional de Colombia; UPME, Determinación del consumo final de energía en el sector residencial urbano y comercial y determinación de consumos para equipos domésticos de energía eléctrica y gas, Medellin, 2006.
- [16] L. Hubert, H.-F. Köhn y D. Steinley, «Cluster Analysis: A Toolbox for MATLAB,» 2009.
- [17] J. Shlens, TUTORIAL ON PRINCIPAL COMPONENT ANALYSIS, <http://www.sn1.salk.edu/~shlens/pub/notes/pca.pdf>, 2005.
- [18] S. Balakrishnama y A. Ganapathiraju, LINEAR DISCRIMINANT ANALYSIS - A BRIEF TUTORIAL, Mississippi, USA.
- [19] MathWorks.inc, <http://www.mathworks.com/help/stats.html>.
- [20] D. cai, LDA: Linear Discriminant Analysis; <http://www.cad.zju.edu.cn/home/dengcai/Data/code/LDA.m>.

## Bibliografía:

- A. Albicki y A. Cole, "Algorithm for Non-Intrusive Identification of Residential Appliances," Circuits and Systems, Proceedings of the IEEE International Symposium on, Vol. III, pp. 338-341, 1998.
- A. I. Albicki y A. Cole, "Data Extraction for Effective Non-Intrusive Identification of Residential Power Loads," IEEE Instrumentation and Measurement Technology Conference, pp. 812-815, 1998.
- A. Zoha, A. Gluhak, M. A. Imran y S. Rajasegarar, Non-Intrusive Load Monitoring Approaches for Disaggregated, 2012.
- D. cai, LDA: Linear Discriminant Analysis;  
<http://www.cad.zju.edu.cn/home/dengcai/Data/code/LDA.m>.
- IPCC, Intergovernmental Panel on Climate Change, Climate Change, 2007.
- J. Borstein y K. Blackmore, In-Home Display Units:, March, 2008.
- L. Farinaccio y R. Zmeureanu, "Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses," Energy and Buildings, vol. 30, pp. 245-259, 1999.
- L. Hubert, H.-F. Köhn y D. Steinley, «Cluster Analysis: A Toolbox for MATLAB,» 2009.
- MathWorks.inc, <http://www.mathworks.com/help/stats.html>.
- M. Baranski y J. Voss, "Detecting Patterns of Appliances from Total Load Data Using a Dynamic Programming Approach," Fourth IEEE International Conference on Data Mining (ICDM'04), 2004.
- M. Baranski y J. Voss, "Genetic Algorithm for Pattern Detection in NIALM Systems," IEEE International Conference on Systems, Man and Cybernetics, pp. 3462-3468, 2004.
- M. Baranski y J. Voss, "Non-Intrusive Appliance Load Monitoring Based on an Optical Sensor," IEEE Power Tech Conference,, Bologna, 2003.
- S. Balakrishnama y A. Ganapathiraju, LINEAR DISCRIMINANT ANALYSIS - A BRIEF TUTORIAL, Mississippi, USA.

Universidad Nacional de Colombia; UPME, Determinación del consumo final de energía en el sector residencial urbano y comercial y determinación de consumos para equipos domésticos de energía eléctrica y gas, Medellín, 2006.

Z. Kolter y M. Johnson, REDD: A public data set for energy disaggregation research, In proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011.