Clasificación de eventos para Monitorización no Intrusiva de carga utilizando razonamiento basado en casos

José David Cortés Torres

Trabajo de Investigación para optar por el título de:

Magíster en Ingeniería Eléctrica

Director

Ph.D. en Electrical and Computer Engineering

César Antonio Duarte Gualdrón

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones E3T

Maestría en Ingeniería Eléctrica

Bucaramanga

2018

Agradecimientos

A mis Familiares

A mis Padres, han realizado mucho con tan poco, los admiro, gracias por su amor y dedicación. En sus manos está la posibilidad de desarrollar el más alto potencial y gracias a la educación temprana y un excelente ejemplo de vida.

A Dios

Por haberme bendecido con los dones que me ha otorgado, por permitir llegar al camino de la ciencia y la educación nuevamente. Por haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

A GISEL

Agradezco a todos mis compañeros y profesores que creyeron en mi capacidad, capacidad que tenemos todos, es grato contar con un grupo de trabajo fuerte y disciplinado.

A la UIS

A la Universidad Industrial de Santander y a cada profesor que entendió mis orígenes y competencias.

Tabla de Contenido

Introducción 14
1 Objetivos 17
1.1 Objetivo General 17
1.2 Objetivos Específicos 17
2 Monitorización no intrusiva de cargas eléctricas 17
3 Razonamiento basado en casos 22
3.1 Desarrollo Histórico 22
3.2 Fases del RBC 24
3.2.1 Representación de un caso en RBC 25
3.2.2 Indexación de la base de casos. 27
3.2.3 Fase de Recuperación. 28
3.2.4 Fase de Reuso 31
3.2.5 Fase de Revisión 32
3.2.6 Fase de Retención. 33
4 Diseño de los modelos de RBC para clasificación de eventos en sistemas de monitorización no-intrusiva de cargas eléctricas. 35
4.1 Preprocesamiento 36
4.1.1 Mediciones. 36
4.1.2 Construcción de la base de casos. 38
4.1.3 Extracción de Características 41
4.1.3.1 Indexación del Conocimiento. 42
4.1.4 Definición de los Descriptores 46
4.1.5 Selección de Características. 51
4.2 Procesamiento 51
4.2.1 Algoritmos de Recuperación. 53
4.2.2 Algoritmos de Reutilización. 55
4.2.3 Algoritmos de Revisión. 57
4.2.4 Algoritmos de Retención. 60
4.3 Modelo RBC 64

4.4 Combinaciones. 67
5 Análisis del desempeño de los sistemas diseñados para la clasificación de eventos.....68
5.1 Análisis de resultados obtenidos para cada modelo propuesto de RBC 68
5.2 Análisis de resultados obtenidos por bases de conocimiento. 70
5.3 Análisis de resultados obtenidos por régimen. 71
5.4 Análisis de resultados obtenidos por conexión de aparatos. 72
5.5 Sumario de desempeños de los diferentes clasificadores.....74
5.6 Comparación de los sistemas de identificación 75
5.7 Análisis de Varianza usando el Método Kruskal-Wallis para Datos No paramétricos ..76
6 Conclusiones y Trabajo Futuros 77
6.1 Conclusiones.....77
6.2 Recomendaciones y Trabajos Futuros 78
6.3 Divulgación de Trabajos 79
Apéndice A. Clases consideradas para cada una de las bases de casos 80
Apéndice B. Descriptores considerados para cada una de las bases de casos.....87
Apéndice C. Histogramas para los modelos de RBC considerados para cada una de las bases de casos en los diferentes regímenes. 92
Apéndice D. Datos obtenidos en cada partición para comparación.....96
Apéndice E. Resultados de las pruebas no paramétricas usando el método de Kruskal-Wallis .100
Referencias Bibliográficas.....103

Lista de Tablas

Tabla 1. Características extraídas para regímenes estacionario y transitorio de propuestas en la literatura.....21

Tabla 2. *Listado de aparatos eléctricos de uso común en un ambiente residencial se consideran en este trabajo de investigación. Se presenta la potencia nominal que entrega el fabricante, los estados están definidos por la operación del aparato* 38

Tabla 3. *Bases de datos construidas usando diferentes métodos de extracción de características.*.....39

Tabla 5. *Descripción de los regímenes de operación de los aparatos.* 43

Tabla 4. *Estados de operación considerados para aparatos electrónicos, inductivos y resistivos. Se mencionan algunos de los estados de operación de algunos aparatos.*.....43

Tabla 7. *Características de las combinaciones de dos aparatos eléctricos que fueron utilizadas en las mediciones. Se presenta el grupo de dos aparatos y los estados de funcionamiento definidos por la naturaleza de los aparatos y los transitorios que fueron considerados. Para combinaciones de 2 aparatos, solamente se consideran transitorios de encendido o de apagado para uno solo de los dos aparatos que participan en la combinación.* 44

Tabla 6. *Características de los aparatos eléctricos utilizado en la medición individual. Se presentan los estados de funcionamiento, los transitorios de encendido y apagado y la potencia dada por el fabricante.* 44

Tabla 8. *Características de las combinaciones de tres aparatos eléctricos que fueron utilizadas en las mediciones. Se presenta el grupo de tres aparatos y los estados de funcionamiento definidos por la naturaleza de cada aparato y los transitorios que fueron considerados. Para combinaciones de 3 aparatos, solamente se consideran transitorio de encendido o de apagado para uno solo de los 3 aparatos que participan en la combinación, los restantes están en régimen estacionario.* .45

Tabla 9. *Descriptores usados para la conformación de la base de casos llamada ST para estado transitorio.* 48

Tabla 10. *Descriptores usados para la conformación de la base de casos llamada ST para régimen transitorio.* 48

Tabla 11. *Clases definidas para los aparatos individuales y combinaciones de 2 y 3 aparatos para la construcción de la base de casos Wavelet, Fractional y S en régimen transitorio. TSX significa que es en estado transitorio. Para mayor información de las etiquetas se puede consultar en la Tabla 21.* 52

Tabla 12. *En la columna 3 se presentan las componentes principales extraídas que recoge el 95% de la información de todos los descriptores para cada una de las bases de casos consideradas. Para los casos en que no haya una cantidad de componentes principales que recoja el 95%, el entrenamiento se realiza con los descriptores originales.*.....53

Tabla 13. *Exactitudes obtenidas para los modelos de RBC del 1 al 12.* 74

Tabla 14. *Exactitudes obtenidas para los modelos de RBC del 13 al 24.* 74

Tabla 15. *Clases definidas para los aparatos sometidos a prueba para la construcción de la base de casos FVs en régimen estacionario* 80

Tabla 16. *Clases definidas para los aparatos sometidos a prueba para la construcción de la base de casos FVTime en régimen transitorio. TS2 significa que el transitorio es apagado-encendido.* 80

Tabla 17. *Clases definidas para los aparatos sometidos a prueba para la construcción de la base de casos Wavelet, Fractional y S en régimen estacionario. SS1 significa que es en estado estacionario de los aparatos individual.* 81

Tabla 18. *Clases definidas para las combinaciones de 2 aparatos para la construcción de la base de casos Wavelet, Fractional y S en régimen estacionario. SS1 significa que es en estado estacionario.* 81

Tabla 19. *Clases definidas para las combinaciones de 3 aparatos para la construcción de la base de Casos Wavelet, Fractional y S en régimen estacionario. SS1 significa que es en estado estacionario.* 82

Tabla 20. *Clases de aparatos tenidos en cuenta para transitorios* 83

Tabla 21. *Descriptores usados para la conformación de la base de casos llamada FV para estado transitorio. Los descriptores desde 1-13 y desde el 22 al 36 son tomados de la IEEE 1459 Standart. Los descriptores desde el 14 hasta el 21 son tomados de (IEEE, 2010).* 87

Tabla 22. *Descriptores usados para la conformación de la base de casos llamada FVTime para estado transitorio.* 88

Tabla 23. *Forma de cálculo de los descriptores usados en la construcción de la base de casos Wavelet estacionario y transitorio.* 89

Tabla 24. *Forma de cálculo de los descriptores usados en la construcción de la base de casos S-Transform estacionario y transitorio.* 89

Tabla 25. *Forma de cálculo de los descriptores usados en la construcción de la base de casos FRTF estacionario y transitorio.* 90

Tabla 26. *Descriptores usados para la conformación de la base de casos llamada IEEE1459 para estado estacionario. Los descriptores son tomados de la IEEE 1459 Standart.* 90

Tabla 27. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, Wavelet Estacionario.* 96

Tabla 28. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, Wavelet Transitorio.* 96

Tabla 29. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, FRFT Estacionario.* 97

Tabla 30. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, FRFT Transitorio.* 98

Tabla 31. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, S-Transform Estacionario.* 98

Tabla 32. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, S-Transform Transitorio.* 99

Lista de Figuras

Figura 1. Fases que comprenden el ciclo de razonamiento basado en casos. 26

Figura 2. Algunas indexaciones y modo de organización del conocimiento en RBC. Tomado de (Bichindaritz, 2008).....29

Figura 4. Formas de conexión para diferentes aparatos. Los aparatos pueden estar en modo individual o en combinaciones de 2 o 3 aparatos. 37

Figura 3. Módulos de conexión utilizados para la adquisición de las señales 37

Figura 5. Ontología de la base de casos para los modelos de RBC propuestos. 41

Figura 6. Procedimiento para la adaptación usando desviación estándar de los casos promisorios seleccionados en la fase de reutilización.....59

Figura 7. Procedimiento para el análisis de las clases candidatas a ser solución del caso consulta.....61

Figura 8. Algoritmo de aprendizaje por casos.....63

Figura 9. Algoritmo para la reducción de instancias ruidosas 64

Figura 10. Procedimiento de implementación para algoritmo de reducción de instancias ruidosas.....65

Figura 11. Mapa general de modelo de RBC para clasificación de eventos en monitorización no intrusiva de carga. 66

Figura 12. Combinaciones de los diferentes métodos para cada una de las fases del RBC. Se encuentran agrupados de a tres para mostrar que varios modelos tienen fases en común. Por ejemplo, los modelos 1-2-3 tienen en común el algoritmo de distancia Euclidiana, el algoritmo por regla de mayoría y la adaptación por reinstaciación, sólo difieren del modelo de retención usado. 67

Figura 13. Desempeños de los 24 modelos de RBC propuestos para todas las bases de casos consideradas y divididas por tipo de régimen.....69

Figura 14. Resultados de desempeño para los 24 modelos de RBC propuestos que se encuentran arriba del 70%. Para estos valores seleccionados en verde, se calcularon el promedio, la mediana, la media armónica y geométrica. 70

Figura 15. Resultados de desempeño para las bases de casos propuestas que se encuentran arriba del 70%. Para estos valores seleccionados en verde, se calcularon el promedio, la mediana, la media armónica y geométrica. 71

Figura 16. Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, S-Transform Estacionario y Transitorio. 73

Figura 17. Resumen donde se muestran los modelos de RBC que presentan diferencias estadísticamente significativas respecto al modelo de MSV considerado como referente. Se presentan para las bases de casos Wavelet, S-Transform y FRFT tanto en régimen estacionario como en régimen transitorio. 76

Figura 18. Resultados de desempeños para la base de casos FV en régimen estacionario. 92

Figura 19. Resultados de desempeños para la base de casos FVTime en régimen transitorio.92

Figura 20. Resultados de desempeños para la base de casos ST2 en régimen transitorio. 92

Figura 21. Resultados de desempeños para la base de casos ST3 en régimen transitorio. 92

Figura 22. Resultados de desempeños para la base de casos ST4 en régimen transitorio. 93

Figura 23. Resultados de desempeños para la base de casos ST5 en régimen transitorio. 93

Figura 24. Resultados de desempeños para la base de casos ST5 en régimen transitorio. 93

Figura 25. Resultados de desempeños para la base de casos ST7 en régimen transitorio. 93

Figura 26. Resultados de desempeños para la base de casos Wavelet en régimen estacionario. 94

Figura 27. Resultados de desempeños para la base de casos Wavelet en régimen transitorio.94

Figura 28. Resultados de desempeños para la base de casos S-Transformen régimen estacionario. 94

Figura 29. Resultados de desempeños para la base de casos S-Transformen régimen transitorio. 94

Figura 30. Resultados de desempeños para la base de casos FRFT en régimen estacionario..94

Figura 31. Resultados de desempeños para la base de casos FRFT en régimen transitorio. ...95

Figura 32. Resultados de desempeños para la base de casos FRFT en régimen transitorio. ...95

Figura 33. Resultados de desempeños para la base de casos FRFT en régimen transitorio. ...95

Figura 34. Prueba de significancia usando el método de Kruskal-Wallis para las bases de casos con información extraída de la transformada Wavelet. Las líneas con círculos rojos indica que grupos son los que presentan diferencias estadísticamente significativas. La línea con el círculo azul indica el modelo MSV 100

Figura 35. Prueba de significancia usando el método de Kruskal-Wallis para las bases de casos con información extraída de la transformada S. Las líneas con círculos rojos indica que grupos son los que presentan diferencias estadísticamente significativas. La línea con el círculo azul indica el modelo MSV.....101

Figura 36. Prueba de significancia usando el método de Kruskal-Wallis para las bases de casos con información extraída de la Transformada Fraccionaria de Fourier. Las líneas con círculos rojos indica que grupos son los que presentan diferencias estadísticamente significativas. La línea con el círculo azul indica el modelo MSV.....102

Lista de Apéndices

Apéndice A. Clases consideradas para cada una de las bases de casos.....80

Apéndice B. Descriptores considerados para cada una de las bases de casos 87

Apéndice C. Histogramas para los modelos de RBC considerados para cada una de las bases de casos en los diferentes regímenes. 92

Apéndice D. Datos obtenidos en cada partición para comparación 96

Apéndice E. Resultados de las pruebas no paramétricas usando el método de Kruskal-Wallis 100

Resumen en español

TITULO: CLASIFICACIÓN DE EVENTOS PARA MONITORIZACIÓN NO INTRUSIVA DE CARGAS ELÉCTRICAS UTILIZANDO RAZONAMIENTO BASADO EN CASOS[1]

AUTOR: JOSÉ DAVID CORTÉS TORRES[2]

PALABRAS CLAVE: NILM, Razonamiento basado en casos.

El consumo eléctrico en hogares y sitios de trabajo se ha incrementado. Esto puede ser evidenciado en el crecimiento de ciudades inteligentes, el uso de energías renovables y dispositivos tecnológicos de última generación. Panorama que se torna necesario para el diseño de estrategias referentes a la gestión de la demanda. La monitorización intrusiva de cargas (NILM), permite medir el consumo eléctrico de forma centralizada para estimar la potencia individual de cada aparato conectado en un sistema eléctrico usando algoritmos matemáticos. En la actualidad, la metodología de razonamiento basado en casos (RBC) ha mostrado una solución apropiada para tareas de clasificación. Como el RBC fracciona el conocimiento de acuerdo al contexto, representa una ventaja respecto con los métodos tradicionales. El presente trabajo de investigación describe el modelo para clasificación de cargas de tipo residencial en eventos de conexión y desconexión usando la metodología de RBC dividida en en tres etapas. La etapa de extracción de descriptores fue implementada usando transformadas *tiempo-frecuencia* y de las definiciones de la *norma IEEE1459* de 22 aparatos eléctricos de uso residencial. La etapa de clasificación está compuesta por 4 fases que dispone de 2 algoritmos por fase. La etapa de evaluación fue implementada usando la métrica de casos clasificados correctamente sobre total de casos. Los resultados fueron comparados con el método de clasificación basado en máquinas de soporte vectorial (SVM). Los hallazgos más importantes muestran que los 24 modelos implementados con RBC presentaron un 39.66% en promedio. Estos datos revelan que los modelos de RBC cuentan con un desempeño comparable con algunos métodos considerados en NILM, lo que se constituye en una metodología para la clasificación de aparatos en un sistema residencial.

[1] Trabajo de Investigación

[2] Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones. Director, César Antonio Duarte Gualdrón

Resumen en inglés

TITULO: CLASIFICACIÓN DE EVENTOS PARA MONITORIZACIÓN NO INTRUSIVA DE CARGAS ELÉCTRICAS UTILIZANDO RAZONAMIENTO BASADO EN CASOS[3]

AUTOR: JOSÉ DAVID CORTÉS TORRES[4]

PALABRAS CLAVE: NILM, Case Based Reasoning.

Electricity consumption in homes and workplaces has increased. This can be evidenced in the growth of smart cities, the use of renewable energy and cutting-edge technological dis-positives. Panorama that becomes necessary for the design of strategies related to demand management. The non-intrusive load monitoring (NILM), allows to measure the centralized power consumption to estimate the individual power of each connected device in an electrical system using mathematical algorithms. Currently, case-based reasoning (CBR) methodology has shown an appropriate solution for tasks classification. As the RBC divides the knowledge according to the context, it represents an advantage respect to the traditional methods. The present research describes CBR methodology model to classify connection and disconnection events. It was divided into three stages. The feature extraction stage was implemented using time-frequency transformations and IEEE1459 definition standard for 22 home electrical appliances. The classification stage is composed by 4 phases. Each phase has 2 algorithms. The evaluation stage was implemented using the metric of correctly classified cases by total cases. The results were compared with the classification method based on support vector machines (SVM). The results showed the 24 models implemented with CBR presented an average of 39.66%. These data reveal that some CBR models have a comparable performance with some methods considered in NILM, which constitutes a methodology for the classification of appliances in a residential system.

[3] Trabajo de Investigación

[4] Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones. Director, César Antonio Duarte Gualdrón

Introducción

Los países miembros de la Unión Europea establecieron las metas que a nivel ambiental y energético se deberían lograr para el año 2030. Una de ellas es la reducción de los niveles de polución y mejoramiento de la eficiencia energética (De Baets, Ruyssinck, Develder, Dhaene y Deschrijver, 2017). Esto puede ser un aval para que las energías renovables se incrementen en los próximos años, sin embargo, para la puesta en marcha de nuevas formas de producción de energía limpia implica superar ciertos desafíos, uno de ellos, la dispersión geográfica de las fuentes de producción de energía eléctrica disponibles. Otra de las opciones que está jugando un rol importante en la actualidad es la monitorización y gestión del consumo eléctrico. Esto se ha evidenciado en las redes inteligentes, donde la gestión del consumo y la energía generadas, se realiza de forma automática, además de la variedad de dispositivos que han sido desarrollados para el análisis del consumo eléctrico en los últimos años. El registro en tiempo real es información altamente valorada, dado que pueden ser extraídos indicadores para gestión del consumo. Un dispositivo de medida puede obtener datos del consumo de una residencia, analizarlo y detectar patrones de comportamiento para identificar cuáles aparatos se encuentran conectados. Este enfoque de análisis de patrones de consumo, está basado en la metodología conocida como Monitorización No intrusiva de Carga (NILM). Esta metodología usa un dispositivo de registro de variables primarias que, mediante algoritmos matemáticos para identificación de aparatos eléctricos, estiman los perfiles de potencia eléctrica en un ambiente residencial.

Por otro lado, el razonamiento basado en casos (RBC) ha sido reconocido como una metodología de aprendizaje que es capaz de generalizar el conocimiento en dominios que aún no han sido realmente entendidos (Aamodt y Plaza, 1994). Está compuesto por 4 fases que son: Fase de recuperación, fase de reuso, fase de revisión y fase de retención. Se basa en la forma de razonamiento

de los seres humanos y en cómo resuelven un problema. Tiene como premisa que la solución a un nuevo caso tiene similares soluciones, por lo que, los nuevos problemas son resueltos recuperando información contenida en casos similares con soluciones que son ya conocidas. Usualmente, los casos son representados como vectores que contienen valores de descriptores previamente seleccionados. Usando una métrica de similaridad entre el llamado caso consulta (nuevo caso) y los casos que se encuentran almacenados en la base de casos, se encuentra la solución a este nuevo problema. Parte del éxito de recuperar la solución apropiada al caso consulta se encuentra en seleccionar los apropiados descriptores, que capturen la mayor información posible de los casos recuperados y que pueda ser de utilidad en la clasificación del caso consulta.

Las técnicas para extracción de características basadas en transformadas tiempo-frecuencia como la transformada discreta de Fourier, la transformada S, la transformada Wavelet son las más implementadas en monitorización no intrusiva de carga (Chan y Fu, 1999; Cunningham, 2009; Duarte, Delmar, Goossen, Barner y Gomez-Luna, 2012; Jimenez, Duarte, Petit y Carrillo, 2014; Nilsson y Funk, 2004; Olsson, Funk y Xiong, 2004; Tzanetakis, Essl y Cook, 2001). Sin embargo, estás técnicas son adecuadas para señales que tienen una dinámica simple y predecible, pero que fallan en identificar patrones de mayor complejidad, como señales con ruido o aparatos combinados (Daw, Finney y Tracy, 2003). Otro problema en la extracción de características mediante métodos tiempo-frecuencia, es la cantidad de coeficientes o escalas que deben ser considerados, y como los tiempos de muestreo son considerablemente altos, generan vectores de alta dimensionalidad que tienen alta complejidad computacional.

Este trabajo de investigación propone un enfoque distinto para abordar el problema de clasificación de aparatos eléctricos, mediante la metodología de razonamiento basado en casos, donde

las señales registradas mediante sensores de corriente y tensión eléctricas, proveerán la información necesaria para extraer los descriptores que servirán para clasificar los distintos aparatos. Adicionalmente, se usan métricas de similaridad numérica y técnicas de adaptación del conocimiento que permitirán interpretar el concepto de clasificación en monitorización no intrusiva de cargas eléctricas de un modo novedoso.

El contenido de la tesis es desarrollado de la siguiente manera, en la Sección 2 se describe la monitorización no intrusiva de carga y algunas propuestas para el diseño de sistemas de identificación de cargas con algoritmos de reconocimiento de patrones. En la Sección 3 3, se presenta el desarrollo histórico, los algoritmos usados en cada una de las fases del razonamiento basado en casos y cómo se estructura la metodología. En la Sección 4 se presenta la propuesta de diseño usando la metodología de razonamiento basado en casos, en el que se han seleccionado dos métodos en cada una de las fases. Además, se explica la naturaleza del conocimiento extraído, los descriptores usados y todo lo relacionado con la etapa de preprocesamiento. En la Sección 5 se presentan los resultados obtenidos para los modelos propuestos de acuerdo a las fuentes de conocimiento que fueron usadas para la elaboración de los algoritmos basados en la metodología de estudio. Finalmente se presentan las conclusiones y trabajos futuros.

1 Objetivos

1.1 Objetivo General

- Diseñar las etapas de un sistema de identificación de eventos basado en la metodología razonamiento basado en casos para la monitorización no intrusiva de cargas eléctricas de tipo residenciales

1.2 Objetivos Específicos

- Recopilar los algoritmos recomendados en la literatura para cada etapa del proceso de CBR.
- Ajustar un sistema usando razonamiento basado en casos para la identificación de estado de cargas eléctricas residenciales monofásicas.
- Evaluar el desempeño del sistema propuesto con respecto a otros sistemas basados en el uso de técnicas de inteligencia artificial como, por ejemplo, máquinas de soporte vectorial

2 Monitorización no intrusiva de cargas eléctricas

La demanda mundial de energía eléctrica se proyecta con una cifra que alcanzará los 33,300 TWh hacia el año 2030. Esta cifra es un referente que exige que los operadores de red y plantas de generación de energía eléctrica se encuentren preparados ante una eventual circunstancia, sin embargo, tal parece que no se podrá satisfacer la demanda de energía eléctrica, si su ritmo de crecimiento sigue siendo más alto respecto a la producción de energía eléctrica (Lee, Jung, Kim, Lee y Kim, 2010).

Recientes estudios han mostrado que el ahorro de energía eléctrica es una solución para evitar el colapso que sucedería en caso que la producción de energía eléctrica sea insuficiente ante la creciente demanda (Paetz, Dütschke y Fichtner, 2012). De acuerdo a la Agencia Internacional de la Energía, "el uso racional de la energía tanto a nivel domiciliario como a nivel industrial implicaría un ahorro en el consumo del 15 al 20%. Este ahorro evitaría la extinción de recursos no renovables utilizados en la generación de electricidad" (M. Berges, Soibelman y Matthews, 2010). Otros estudios han mostrado que un consumidor que disponga de información relacionada con su consumo de energía eléctrica, tendrá una tendencia al ahorro de energía eléctrica (Mario Berges, Goldman, Matthews y Soibelman, 2008). Dada esta situación, la monitorización de consumo surge como propuesta que permite obtener información del comportamiento del consumo y a su vez, desarrollar estrategias para la gestión de la demanda.

Ahora bien, en la actualidad el consumo se registra mediante los contadores de energía eléctrica que son una de las opciones que se encuentran disponibles en ambientes residenciales. Permite a los operadores de red (OR) registrar la cantidad de kWh en un determinado tiempo. Sin embargo, no recolectan información relacionada con la actividad individual de cada elemento de circuito conectado a la red. Solo presentan valores globales de energía eléctrica, sin información detallada sobre el consumo eléctrico. Otra opción son los contadores inteligentes, que además de registrar el consumo eléctrico en tiempo real, que son capaces de ajustar la potencia instalada y verificar las lecturas de consumo de forma remota. Al realizar un registro en tiempo real, nuevos servicios tanto para el consumidor final como para los operadores de red pueden ser ofrecidos. Algunos ejemplos como la detección de consumo de aparatos que se encuentran conectados, las comparaciones de consumo entre poblaciones o regiones y la optimización de recursos para la mejora de eficiencia

energética. Básicamente, estos contadores desarrollan una labor de telegestión. Y finalmente, aunque esta modalidad no está relacionada directamente con los parámetros eléctricos de cada aparato, existen aparatos con servicio de conexión a internet, bases de datos en la nube y conexiones inter-dispositivos que permiten a los fabricantes hacer proyecciones de nuevos modelos más eficientes a nivel eléctrico, lo que en la actualidad es llamado el internet de las cosas. Todos los dispositivos de medición de consumo disponibles son herramientas para la implementación de monitorización de carga.

La monitorización no intrusiva de carga busca obtener información sobre el perfil de consumo eléctrico medido de forma centralizada y desagregarlo de forma individual para cada aparato conectado a la red. Para tal propósito se sitúa un dispositivo en el tablero principal para registrar en un solo punto de medición todos los datos de corriente y tensión eléctricas. La potencia eléctrica instantánea se calcula como el producto de la corriente y tensión eléctricas instantáneas, la cual representa la sumatoria de las potencias individuales de todos los aparatos o cargas eléctricas conectadas al sistema eléctrico residencial. La potencia total calculada es el punto de partida para conocer el consumo individual de cada aparato conectado en un instante determinado. De las mediciones registradas se pueden inferir variables secundarias que son construidas mediante algoritmos matemáticos para obtener firmas de carga. Una firma de carga está compuesta por parámetros eléctricos que brindan información acerca de la naturaleza y el estado de operación de cada aparato de forma individual. Para identificar a qué aparato o aparatos pertenecen los perfiles de carga obtenidos de la conexión o desconexión de aparatos individuales o combinados se utiliza la metodología de reconocimiento de patrones.

Formalmente, se define P_T como la potencia total registrada como suma de potencias individuales agregadas en un tiempo t mostrada en la Ecuación 1:

$$P_T(t) = P_1(t) + P_2(t) + \ldots + \mathrm{P}_j(t) + \ldots P_N(t)$$

Ecuación 1

donde N es el número de aparatos que se encuentran conectados al sistema eléctrico analizado. Como $P_T(t)$ es la medida total agregada, solo resta estimar qué aporte tendrían los N aparatos conectados a la red en el instante de tiempo de análisis. El problema para determinar el valor de los aportes individuales son las infinitas posibles soluciones, aumentando la complejidad computacional. Un evento se define como el cambio de estado de un aparato, es decir, de encendido a apagado o viceversa, en un determinado tiempo. Estos cambios pueden ser detectados y cuantificados usando extracción de características. La literatura ha propuesto extracción de características en detección de eventos implementadas en una tarea de clasificación (Mario Berges, Goldman, Matthews, Soibelman y Anderson, 2011; Jin, Tebekaemi, Berges y Soibelman, 2011; Leeb, Shaw y Kirtley, 1995; Norford y Leeb, 1996). Las firmas de carga parametrizan la solución de tal forma que las características extraídas acotan el grupo de las soluciones que pueden ser obtenidas (Zoha, Gluhak, Imran y Rajasegarar, 2012). Las firmas de carga pueden ser extraídas de dos tipos de regímenes de funcionamiento de cargas eléctricas. Las firmas en régimen transitorio, que se presentan justo después de la conexión o desconexión de un aparato a la tensión eléctrica y las firmas en régimen estacionario que ocurren cuando el aparato presenta un consumo constante que permanece invariante en el tiempo. La mayor parte de los enfoques abordados en Monitorización No intrusiva de carga se encuentran en el enfoque de aprendizaje supervisado, esto es, del fenómeno a estudiar se conoce a qué aparato o grupo de aparatos y a qué condiciones pertenecen las firmas de carga que fueron extraídas.

Se define un operador de extracción de características $V(\bullet)$ aplicado a una señal $x[k]$ del cual se obtiene un vector de características $C[j]^n = \langle C_1, C_2, ...C_D \rangle$ donde n es la clase a la que pertenece al experimento *j-ésimo* que permite extraer el *j-ésimo* vector característico. Se realiza una asignación de clase al vector característico extraído para representar la firma de carga, donde la clase asignada es el aparato que fue sometido a prueba de conexión y desconexión. Después de obtener los vectores de características, sólo resta realizar un proceso de entrenamiento. Para el

Tabla 1. *Características extraídas para regímenes estacionario y transitorio de propuestas en la literatura.*

Régimen	Método	Características
Estacionario	**Cambio en la potencia**	Variación de la potencia activa y reactiva en estado estable
	Características tomadas del dominio del Tiempo y Frecuencia	Corriente RMS, Corriente Promedio, Corriente Pico, Voltaje RMS y factor de potencia
	Trayectorias V-I	Característica de Forma V-I, asimetría, área, curvatura, pendiente, etc.
	Ruido	Ruido de alta frecuencia generada por interferencia electromagnética
Transitorio	**Potencia**	Perfiles de potencia, envolventes espectrales
	Corriente transitoria inicial	Picos de corriente, tamaño, duración, forma del transitorio, tiempo de respuesta
	Ruido	Componentes de la transformada de Fourier para ruido de alta frecuencia

enfoque de entrenamiento fuera de línea, se adquieren las mediciones durante un periodo de tiempo, se extraen las características y se les asigna una clase. Para el entrenamiento se utiliza una base de ejemplos ya existente. Un resultado exitoso de un sistema de identificación y clasificación de aparatos eléctricos depende de muchos factores, tales como, los descriptores que son extraídos de las señales registradas, del tipo de clasificador que se use, de cuántos descriptores son usados, de la naturaleza del aparato y su funcionamiento, de la tensión y corriente aplicadas, etc. Mencionando algunos trabajos desarrollados en Monitorización No intrusiva de carga, parten desde el enfoque por un simple agrupamiento propuesto por (Hart, 1992) en donde los aparatos eran clasificados usando potencias activa y reactivas en régimen estacionario y se aplicaban métricas de

distancia para clasificar un nuevo vector en un grupo de vectores conocidos y pertenecientes a una clase definida hasta los nuevos enfoques en los que se incluye suavizado de señales para el tratamiento de variaciones de potencia (Farinaccio y Zmeureanu, 1999), el uso de redes bayesianas (Marchiori, Hakkarinen, Han y Earle, 2011), Redes Neuronales Artificiales (Ruzzelli, Nicolas, Schoofs y O'Hare, 2010), modelo basados en cadenas ocultas de Markov (Zia, Bruckner y Zaidi, 2011),modelos con Máquinas de Soporte Vectorial (Lai, Lai, Huang y Chao, 2013) y enfoques de Redes Neuronales con aprendizaje profundo (Nalmpantis y Vrakas, 2018).

3 Razonamiento basado en casos

3.1 Desarrollo Histórico

El razonamiento basado en casos (RBC) es una metodología que busca emular mediante una inteligencia artificial la forma de razonamiento del ser humano en la resolución de un problema.

Aprovechando la capacidad de razonamiento del ser humano para resolver un problema se han diseñado modelos conceptuales de razonamiento, dado que la resolución de problemas ha jugado un papel importante en la historia de la humanidad, desde el ámbito educacional hasta el gubernamental.

Las investigaciones en el área de la ciencia cognitiva, filosofía y la psicología sobre modelos conceptuales de razonamiento humano, resolución de problemas y aprendizaje (Ludwig, Anscombe y Rhees, 1954; Smith y Medin, 1981), llevan a formular el primer modelo basado en los paquetes de memoria que simulan la forma en que el cerebro humano almacena la información y categoriza los conceptos abstractos del mundo (Bartlett y Burt, 1933). Ampliando esta idea, (Ludwig et al., 1954) observa que los conceptos que componen el mundo natural se explican mejor usando un conjunto de casos o instancias. Una primera aproximación en el área de la Inteligencia

Artificial (IA) es la teoría basada en modelos semánticos de memoria (Collins y Quillian, 1969; Tulving y others, 1972), que son representados mediante hechos del mundo observable que tienen la propiedad de invariancia en el tiempo, sin embargo, tanto la idea de Wittgenstein como la Teoría Semántica no explican cómo se constituye un caso (Voskoglou y Salem, 2014). Tulving combina la teoría semántica y la representación de los paquetes de memoria usando "episodios". Cada episodio se almacena con una etiqueta temporal que luego es usada para su recuperación. En este momento, se pueden vislumbrar las primeras formas de representar un caso en el área de la IA (Tulving y others, 1972).

A la par, en la década de los 70s, la IA observa un apogeo importante en sistemas de razonamiento basados en reglas para la resolución de problemas. Sin embargo, muchos de estos problemas no pueden ser solucionados ya que tienen algunas limitaciones: Las reglas son atemporales, invariantes y obligaban al problema a adoptar comportamientos distintos de su propia naturaleza, es decir, no se comprende el problema sino se le establecen órdenes. Siguiendo con los modelos conceptuales de memoria, Roger Schank planteó una estrategia para solución de problemas mediante una metodología de contextualización (Tulving y others, 1972). Esto permitió observar un problema usando memoria en episodios, pero en función del contexto, es decir, el conocimiento previo y la comprensión sobre el problema basándose en su contexto permiten concluir que la solución más apropiada se logra usando casos en vez de usar reglas. Entonces, Schank define los paquetes organizados en memoria (Memory Organised Packet MOP) representándolos mediante patrones o características que se observan en las situaciones de estudio. A partir de este momento, estás representaciones de memoria dinámica se transformaron en "casos". Ahora surge la pregunta, *¿cuáles son las características que explican mejor los conceptos y que puede dar consistencia a un caso?*

Para ello, Schank, diseñó una serie de marcos con situaciones conceptuales variadas que fueron sometidas a condiciones de estudio repetitivas. Como resultado, se comprueba que un concepto puede ser descrito mediante una característica general y características específicas. En otras palabras, las situaciones pueden entenderse con características homogéneas que establecen semejanzas y con características heterogéneas que establecen diferencias (Schank, 1975; Voskoglou y Salem, 2014). Así pues, los casos específicos con similares propiedades se pueden organizar bajo una estructura llamada episodio generalizado. Si un nuevo caso problema requiere una solución, esta es encontrada cuando ciertas características son equivalentes en un caso previo. Hasta aquí se pueden considerar los inicios del razonamiento basado en casos, desde su punto de partida como un modelo conceptual, basado en la forma de razonamiento humano para la solución de problemas, hasta la definición inicial de un caso de donde deriva su nombre.

3.2 Fases del RBC

El razonamiento basado en casos es una metodología propuesta por Schank y fue Janet Kolodner quien le dio la forma que metodología con la que se conoce en la actualidad (Janet Kolodner, 2014). Esta metodología se basa en la forma en que un ser humano razona sobre los problemas, adoptando como cierta la suposición: “casos similares tienen soluciones similares”, es decir, la solución de nuevos problemas se sustenta en las soluciones de problemas similares de experiencias pasadas (Janet L Kolodner, 1992).

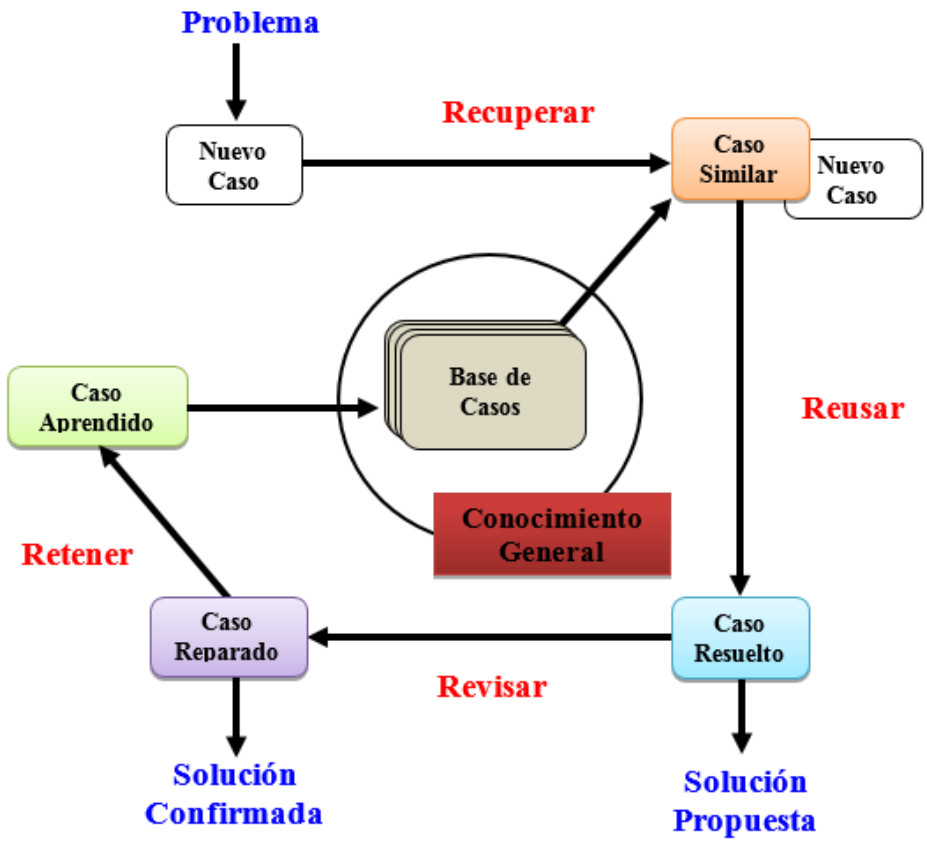
Como se mencionó en la sección anterior, la metodología RBC se basa en cuatro fases que son (Aamodt y Plaza, 1994):

- **RECUPERAR** (Retrieve) el caso o casos más similares. Retoma la experiencia de un problema anterior, el cual es similar a uno nuevo.

- **REUTILIZAR** (Reuse) la información y conocimiento del caso o casos recuperados para resolver un nuevo problema. Copia o integra la solución del caso o casos recuperados.
- **REVISAR** (Revise) la solución que se está proponiendo.
- **RETENER** (Retain) o guardar una nueva solución que ha sido confirmada o validada, para que sea útil en problemas futuros.

En la actualidad la fase de retención se ha dividido en dos más, llamadas "Review" (Repasar) y "Restore" (Restaurar) (Roth-Berghofer y Iglezakis, 2001), pero estas tienen más relación con la forma de organizar el conocimiento e indexarlo para aprovechar los casos que son guardados y que no son suficientemente entendidos para ser almacenados. En la Figura 1 se ilustran las fases de un sistema de razonamiento basado en casos.

3.2.1 Representación de un caso en RBC. Un caso es la representación del conocimiento como la mínima unidad básica de abstracción de un concepto dentro de un contexto. Además, puede ser representado usando diferentes formas y tamaños, con intervalos de tiempo cortos o largos, asociando soluciones con problemas, o salidas con situaciones o ambas. Un caso almacena la experiencia como una enseñanza útil, dado que tiene el potencial de colaborar en la construcción de un razonamiento para lograr un conjunto de objetivos o advertir de posibles fallas para una situación aún no prevista (Janet L Kolodner, 1992).



El caso debe tener los siguientes componentes:

- Una **situación** que describa el estado del problema y sus restricciones.
- Una **respuesta** que contenga la acción tomada para corregir el problema.
- Una **salida** que contenga la información que describe el éxito o fracaso cuando se aplicó la respuesta al problema.

La definición simbólica de un caso debe contener los componentes que se mencionaron en la lista anterior. En este trabajo de investigación se toma la representación de un caso como un conjunto de características o atributos asociados a una solución, definido en la literatura de Monitorización No intrusiva de Carga como **ejemplo**. Para mayor ampliación sobre la representación de un caso, se puede consultar en (El-Sappagh y Elmogy, 2015).

Formalmente, un *j-ésimo* ejemplo $E^{(j)}$, conformado por un vector numérico con d características y que pertenece a una *k-ésima* clase definida $C^{(k)}$ se conoce como un caso. La representación simbólica se puede observar en la Ecuación 2.

$$E^{(j)} = \left[f_1 \;\; f_2 \;\; f_3 \;\; \ldots \;\; f_{(n-1)} \;\; f_{(n)} \;\; \ldots\ldots \;\; f_d \;\; C^{(k)} \right]$$

Ecuación 2

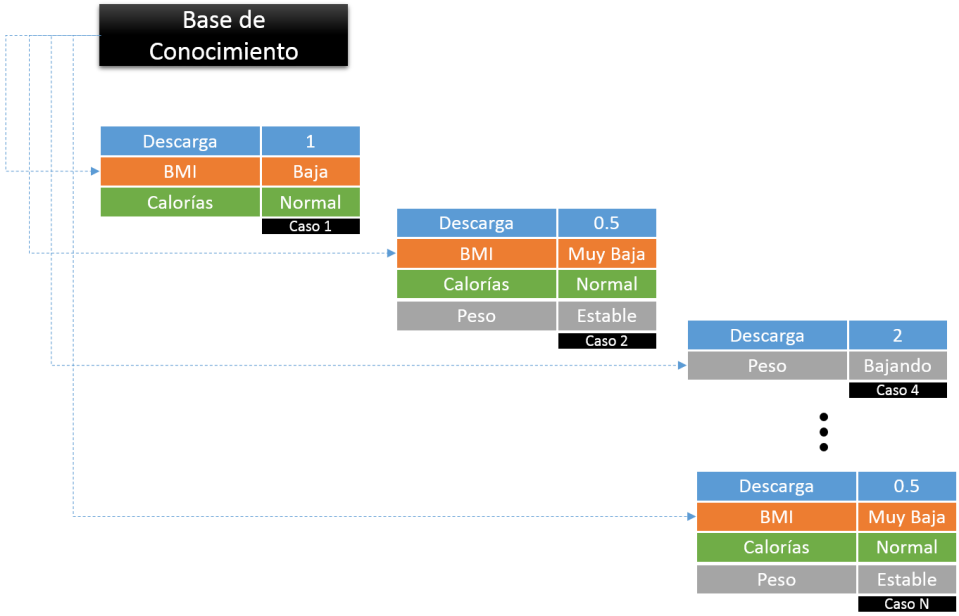
donde el *j-ésimo* ejemplo tiene *d* descriptores simbolizados por f .

3.2.2 Indexación de la base de casos. Al realizar el proceso de construcción y organización del conocimiento en una base de casos, la indexación de los casos hace referencia a la asignación de índices para una futura labor de comparación y extracción de casos útiles. La selección de índices adecuados para organizar los casos almacenados y garantizar la mayor extracción de conocimiento al realizar un proceso de recuperación determina la calidad de la información recuperada. Los índices deben ser de fácil uso, deben reflejar la importancia de los descriptores y la influencia en la solución del caso consulta. No deben ser abstractos y deben permitir la recuperación de los casos en todas las posibles circunstancias. En la Figura 2 se presentan algunos ejemplos de organización de bases de casos. Para mayor información sobre las propiedades que deben presentar los índices, puede ser consultada en (Wortmann, 1997). En cuanto a criterios para selección y construcción de índices se puede en (Birnbaum y Collins, 1989). Algunas técnicas propuestas para la elaboración automática de indexación son: Bonzano, Cunningham y Smyth (1997) usa técnicas inductivas para el aprendizaje de los niveles de importancia de cada descriptor usando casos similares de la base de casos. Brüninghaus y Ashley (1997) emplean la indexación utilizando el conocimiento relacionado en la base de casos, mediante reglas en la clasificación de texto. Sarkheyli y Söffker (2015) utiliza un Operador de Modelado de Situación (Situation Operator Model SOM)

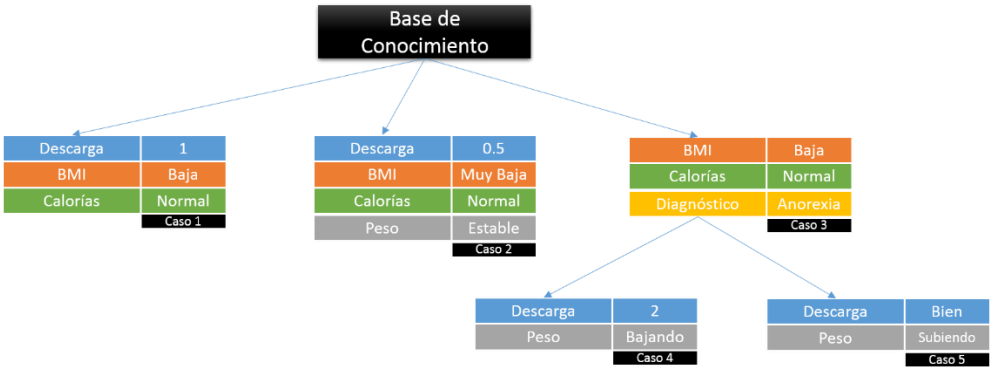
para representar el conocimiento de eventos, acciones relacionadas y los efectos en los casos. Vehviläinen, Alm y Hyvönen (2006) utiliza indexación semántica para crear situaciones de pregunta-respuesta en servicios para web server. Stéphane, Hector y others (2010) plantea una indexación por vecindades esféricas combinada con árboles de decisión. Referente a la temática de monitorización no intrusiva de carga la mejor aproximación referente al tipo de conocimiento que se desea abstraer se encuentra en el Capítulo 19 de (Richter y Weber, 2016), en donde para una aplicación de RBC para el reconocimiento de lenguaje hablado mediante señales temporales en el que un caso es un vector característico denotado por f_v, que pertenece a un conjunto fijo de índices I. En el capítulo 4 se presenta el modo de indexación considerado para este trabajo de investigación.

3.2.3 Fase de Recuperación. El proceso de recuperación es un proceso para encontrar los casos que son lo más cercanos al caso llamado el *caso consulta* definido como C_q. Para llevar a cabo esta labor, es necesario seleccionar un mecanismo que permita realizar la búsqueda en la base de casos y establecer un criterio de recuperación, con el fin de determinar cuáles son las posibles clases a las que puede pertenecer el *caso consulta*. Los métodos para recuperación empleados en investigación son extremadamente diversos, desde el más simple, como los vecinos más cercanos, hasta lo más complejos que incluyen el uso de agentes inteligentes. Independientemente de la técnica, se mide la similaridad de los casos almacenados con el caso consulta para seleccionar los *casos promisorios*. En la Sección 4 se presentan los algoritmos de recuperación considerados para este trabajo de investigación.

(a) Base de casos que tiene una indexación con ordenamiento jerárquico y en función del contexto.



(b) Base de casos que tiene una indexación plana. No existen jerarquías entre los casos.



(c) Base de casos que tiene una indexación semi-estructurada, donde algunos casos tienen asociados algunos sub-casos.

Figura 2. Algunas indexaciones y modo de organización del conocimiento en RBC. Tomado de (Bichindaritz, 2008)

La

medida de similaridad está definida para un par de casos, donde C_m es el caso tomado de la base de casos que se compara con el caso consulta C_q y donde se define la relación $Sim(C_m, C_q)$. La medida de similaridad debe cumplir las propiedades de distancia, tales como la propiedad de identidad, donde $Sim(X, X) = Sim(Y, Y)$, debe ser monótona, independiente y definida en el intervalo cerrado $[0,1]$.

La literatura acerca de medida de similaridad y distancias es muy amplia, puesto que es una técnica muy usada en la disciplina de inteligencia artificial. Uno de los trabajos clave es el desarrollado por (Tversky, 1977), donde se menciona que el proceso de similaridad que consiste en descripciones de equivalencias entre descriptores y valores de distancia. Cunningham organiza las distintas medidas de similaridad y establece una taxonomía, en los que se destacan, los métodos directos, basados en transformaciones, basados en información teórica y los métodos emergentes (Cunningham, 2009). Otro documentos recomendados donde mencionan las diferentes clases de similaridad aplicadas a RBC son: Wu y Palmer (1994) con medidas entre características taxonómicas; Stanfill y Waltz (1986) menciona las métricas solapadas y las basadas en diferencias de valor, que son extendidas más adelante en Wilson y Martinez (1996) y en Cheng, Li, Kwok y Li (2004); Boriah, Chandola y Kumar (2008) define medidas para atributos categóricos; Gabel (2005) define valores ordenados de distancia basados en taxonomías, Gamero, Meléndez y Colomer (2011) y Montani, Bottrighi, Leonardi, Portinale y Terenziani (2009) utilizan medidas de similaridad para implementarlas en series de tiempo, Armengol y Plaza (2012) y Ontañón y Plaza (2007) implementan medidas de similitud en características mediante razonamientos simbólicos y sistemas multiagente.

3.2.4 Fase de Reuso. Cuando uno o varios casos similares han sido recuperados, la solución que está contenida en esos casos es reusada para resolver el caso consulta. Copiar una solución puede ser la forma más simple para dar una respuesta al caso consulta, que usualmente se usa para tareas de clasificación con bases de conocimiento sencillas y con un número de casos limitado. Esto ocurre, porque usualmente la solución al caso consulta se encuentra en la base de conocimiento y no es necesario realizar procesos de adaptación de la solución. Sin embargo, para tareas de planeación o configuración, los procesos de adaptación de la solución son necesarios, ya que probablemente las soluciones no se encuentran disponibles en la base de conocimiento. Algunas técnicas para adaptar las soluciones han sido propuestas en Aha, Marling y Watson (2005) en donde se mencionan algunas como métodos de adaptación por transformación o métodos de adaptación generativa. Las transformaciones ocurren mediante reglas definidas u operadores que modifican las soluciones recuperadas. Se toman la solución o soluciones recuperadas se analizan y se realizan las modificaciones necesarias para proponer una solución. En la Sección 4 se presentan los algoritmos de reutilización considerados para este trabajo de investigación.

En Ji, Park y Lee (2011), realiza adaptaciones usando datos estadísticos de 129 proyectos ejecutados en Corea y probados luego con 13 casos. En Corchado, Lees, Fyfe, Rees y Aiken (1998) utilizan un modelo de adaptación basado en redes neuronales con una función de base radial y que obtiene la solución más representativa. En Bregón y Alfonso (2005) utilizan un método de adaptación basado en el algoritmo de *k-vecinos más cercanos*, donde la solución seleccionada es recuperada mediante votación del más cercano. En Pous, Colomer, Meléndez y de la Rosa (2004) utilizan un método de adaptación por transformación en el que definen jerarquías que llevan a la solución propuesta. En Grech y Main (2005) utilizan algoritmos genéticos para minimizar los pesos de los casos recuperados y así seleccionar la clase que da soluciona al caso consulta. En

Khussainova (2016) se realiza un proceso de adaptación guiada, en donde toma todas las soluciones de los casos recuperados y los usa para calcular puntuaciones y así obtener la solución al caso consulta. En general, la mayoría de los algoritmos implementados que reporta la literatura son basados en la adquisición del conocimiento de dominios expertos para construir una adaptación de los casos recuperados. Esto significa que una persona experta observa de los casos recuperados un modo de adaptar la solución al caso consulta. La solución adaptada es copiada a la base de conocimiento y es donde se produce el aprendizaje.

Por otro lado, los métodos generadores de adaptaciones consisten en algoritmos capaces de resolver problemas usando el conocimiento general, mediante la generación de cierto número de salidas en función de un número de restricciones. Sin embargo, estos métodos presentan complejidad computacional alta y mucho tiempo de cómputo para tener una adaptación apropiada para el caso consulta que requiere ser solucionado (Bergmann, Althoff, Minor, Reichle y Bach, 2009).

3.2.5 Fase de Revisión. En esta fase, al caso consulta ya se la ha asignado una solución que será ahora revisada. Para este proceso, el caso que debe ser revisado pasa por una serie de chequeos para confirmar que la solución copiada en la fase de reuso es la más apropiada y la correcta. La fase de revisión verifica que la solución dada al caso consulta desde el punto de vista de la realidad, sea la solución apropiada para ser aprendida. Este proceso de aprendizaje es básicamente la inclusión del caso consulta en la base de conocimiento. Algunos parámetros para revisión son ejecutados solamente cuando algún error es reportado. En el capítulo 4 se presentan los algoritmos de revisión considerados para este trabajo de investigación.

En Fdez-Riverola, Díaz, Borrajo, Yáñez y Corchado (2005) emplean en la fase de revisión un conjunto Tageno Sugeno Kang (TSK) con el fin de validar la solución inicial propuesta por el sistema. En esta fase se definen dos límites: límites de aceptación y límites de rechazo calculando

como umbral un número obtenido del dominio del problema. En Supic y Ribaric (2005) utilizan reglas de para creación de nuevas situaciones en dominios estructurados con el fin de reparar el caso que no haya sido correctamente adaptado en la fase de reuso. En Weber, Rinderle, Wild y Reichert (2005) utilizan como criterio un resultado negativo para decidir que el caso revisado puede pasar a la fase de retención o no. Por ejemplo, en Louvieris (2008), Baccigalupo y Plaza (2007), Jorro-Aragoneses, Díaz-Agudo y Recio-García (2015) y en Malherbe, Iwaszko y Aufaure (2015) la fase de revisión está diseñada para que un ser humano decida si la solución sugerida es correcta o debe ser adaptada aún. En Jayanthi, Chakraborti y Massie (2010) utilizan una matriz de similaridad con potenciales casos que pueden generar confusión y aplican reglas para definir umbrales de revisión. En Gómez y Plaza (2012) utilizan un algoritmo llamado Generación de Esquemas de Programación Seriales o SSGS (Serial Schedule Generation Scheme) que generaliza listas de programación para contrastar si el caso adaptado tiene una lista válida en función de las restricciones del proyecto. En Nemecs y Buga (2017) la fase de revisión la utilizan como adaptación de características en caso que no se cumplan los umbrales establecidos para los atributos. Y en Schnell, Couffignal, Lieber, Saleh y Jay (2017) la fase de revisión simplemente es la copia completa de todos los casos que han sido adaptados en la fase de reuso.

3.2.6 Fase de Retención. Usualmente en la fase de reuso ya se tiene la solución al caso consulta, sin embargo, cuando se requiere que la información nueva obtenida del caso consulta sea almacenada en la base de conocimiento se realiza un proceso de revisión. Ahora bien, si se decide en la fase de revisión que el caso consulta aporta información relevante a la base de conocimiento entonces será almacenada en la fase de retención. Algunos criterios de revisión simplemente se centran en tomar todas las soluciones de los casos consultados para ser procesados en la fase de retención. Otros criterios se centran en crear reglas en la fase de revisión para luego ser procesadas

en la fase de retención. A este proceso de almacenamiento de los casos revisados en la base de conocimiento se le conoce como fase de retención.

La fase de retención es también conocida como la fase de aprendizaje de un sistema RBC. Se considera aprendizaje cuando un caso revisado es agregado a la base de conocimiento, dado que el nuevo caso estará disponible para ser recuperado y reutilizado en experiencias futuras. A pesar de que podría representar una ventaja, el continuo incremento de la base de conocimiento degrada la eficiencia en la recuperación, además que haría el procesamiento de la solución más demorado. Es por eso, que algunas propuestas de RBC proponen en la fase de retención la eliminación de la información redundante, ruidosa y que no aporta significado a la base de conocimiento. A esto se le conoce como mantenimiento de la base de conocimiento y es una de las propuestas donde adicionan dos fases más en la metodología RBC, llamadas Repasar (Review) y Restaurar (Restore) (Roth-Berghofer y Iglezakis, 2001). En (Smyth y McKenna, 2001) desarrollan modelos de competencia explícita para activar la selección de los casos que van a ser agregados a la base de conocimiento. En Armengol y Plaza (2003) describen mediante aprendizaje inductivo como enfoque *lazy* como puede ser usado en la fase de retención del RBC. Hanney y Keane (1996) menciona ejemplos de cómo aprender mediante reglas de adaptación. Gupta, Aha y Moore (2004) desarrolla un enfoque de aprendizaje mediante taxonomías para la indexación de casos mediante un RBC conversacional. En Zhang, Zhang y Leake (2017) utilizan la técnica de k-medianas para hacer particiones y decidir qué casos almacenar o no.

Los primeros enfoques de mantenimiento de base de casos se encuentran en Leake y Wilson (1998), Smyth (1998) y en Smyth y Cunningham (1996) en donde se aborda la problemática de eliminar casos redundantes para evitar el crecimiento de la base de ejemplos. Reinartz, Iglezakis y Roth--Berghofer (2001) describe las fases de repaso y de restauración. En Iglezakis, Reinartz y

Roth-Berghofer (2004) se menciona una metodología para mantenimiento de RBC usando memorias. En Craw, Massie y Wiratunga (2007) se menciona una metodología de mantenimiento basado en el enfoque introspectivo mediante parámetros que ajusta un usuario y por último, Murdock y Goel (2001) y Mehta, Ontañón y Ram (2009) usan una perspectiva de meta-razonamiento para desarrollar un control y monitorización de la base de conocimiento. Otros enfoques de mantenimiento se basan en la eliminación de los casos que pueden generar ruido o que no aportan a la clasificación. En Ruiz Ordóñez y others (2008) se citan dos métodos para mantenimiento de bases de casos que son los métodos IB3 (Instance Based Learning) y DROP (Decremental Reduction Optimization Procedure). El IB3 elimina casos mediante un criterio estadístico basado en la distribución *t-student*, y utiliza intervalos de decisión, entre lo que se encuentra, rechazo, aceptación y una región de indecisión, mientras que el DROP elimina los casos de acuerdo a reglas de clasificación definiendo los k-vecinos más cercanos llamados *asociados*. A partir de los asociados, los casos son retenidos o eliminados de la base de casos general. Estos dos últimos algoritmos se consideran en el capítulo 4 como algoritmos de retención para este trabajo de investigación.

4 Diseño de los modelos de RBC para clasificación de eventos en sistemas de monitorización no-intrusiva de cargas eléctricas.

A continuación, se explican las consideraciones referentes a las etapas del modelo de razonamiento basado en casos y qué criterios se tuvieron en cuenta para la selección de cada uno de los algoritmos. Como el modelo RBC está compuesto por dos etapas, en la primera se hablará del preprocesamiento, todo lo referente a la adquisición de medidas, la construcción de bases de casos y la extracción y selección de características útiles para el modelo RBC. En la segunda etapa ya se explican el modelo RBC y los diferentes algoritmos en cada una de las fases que lo componen.

4.1 Preprocesamiento

Para acondicionar las señales para que el modelo RBC realice la clasificación, es preciso realizar una serie de actividades previas con el fin de entregar un vector apropiado para el modelo RBC.

4.1.1 Mediciones. Para la adquisición de los datos se realizan varios registros de medidas en el Laboratorio de Integración Energética del Grupo de Investigación en Sistemas de Energía Eléctrica (GISEL), adscrito a la Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones (E3T), de la Universidad Industrial de Santander (UIS), ubicado en Guatiguará- Piedecuesta (Lab 308). Para esto se utiliza una fuente Chroma Modelo 61511/61512 y tres módulos de adquisición de datos National Instrument con referencia NI9225, para medida de tensión entre -300V y 300V, con referencia NI9239 para medida de tensión entre -10V y 10V y con referencia NI9227 para medida de corriente de -5A a 5A. Para los casos en que la corriente del aparato excede la capacidad del módulo de corriente, se utiliza una pinza de corriente Fluke *i30s* que convierte la corriente en un nivel de tensión que se ajusta con el módulo NI9239. Para la obtención de los eventos se realiza mediante la conexión y desconexión de los diferentes aparatos eléctricos de uso común en un ambiente residencial. El objetivo es la obtención de señales punto a punto, ajustado a una tasa de muestreo de 50 000 muestras por segundo. Cuando se realiza la conexión y desconexión del aparato sometido a prueba, el módulo de adquisición cDAQ 9174 envía los datos a la computadora y los almacena en formato de matriz para ser manipulado usando el software MATLAB®. Se utiliza una interfaz de usuario usando el software LabView® para definir los tiempos de duración de la prueba, el lugar de almacenamiento y las etiquetas de los datos crudos. En la Figura 3 se muestran los módulos de corriente y tensión utilizados para las pruebas.

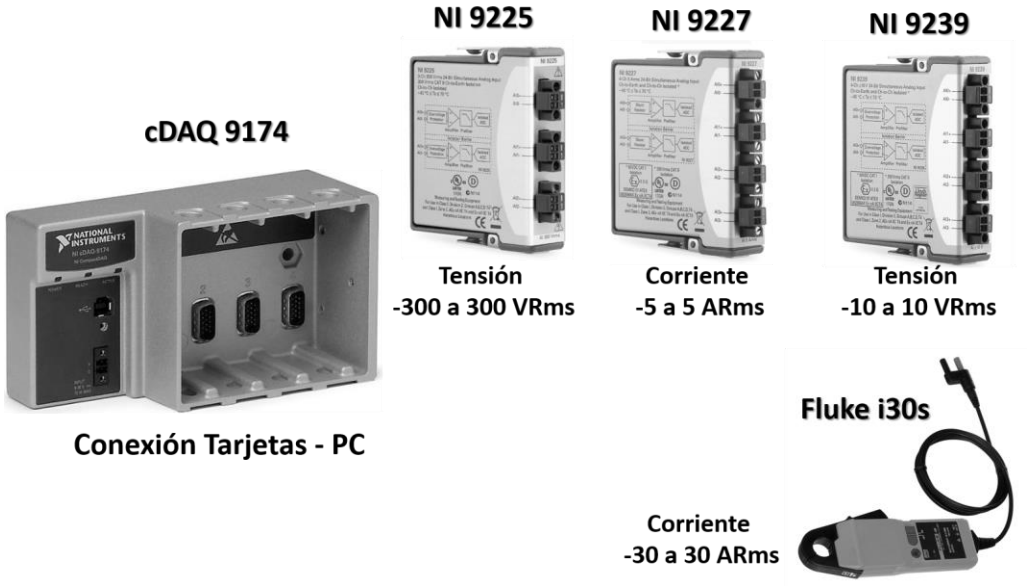


Figura 4. Módulos de conexión utilizados para la adquisición de las señales

Por otro lado, mencionando el modo en que fueron conectados los diferentes aparatos y con el objetivo de simular la operación natural del aparato, fueron alimentados directamente con la fuente

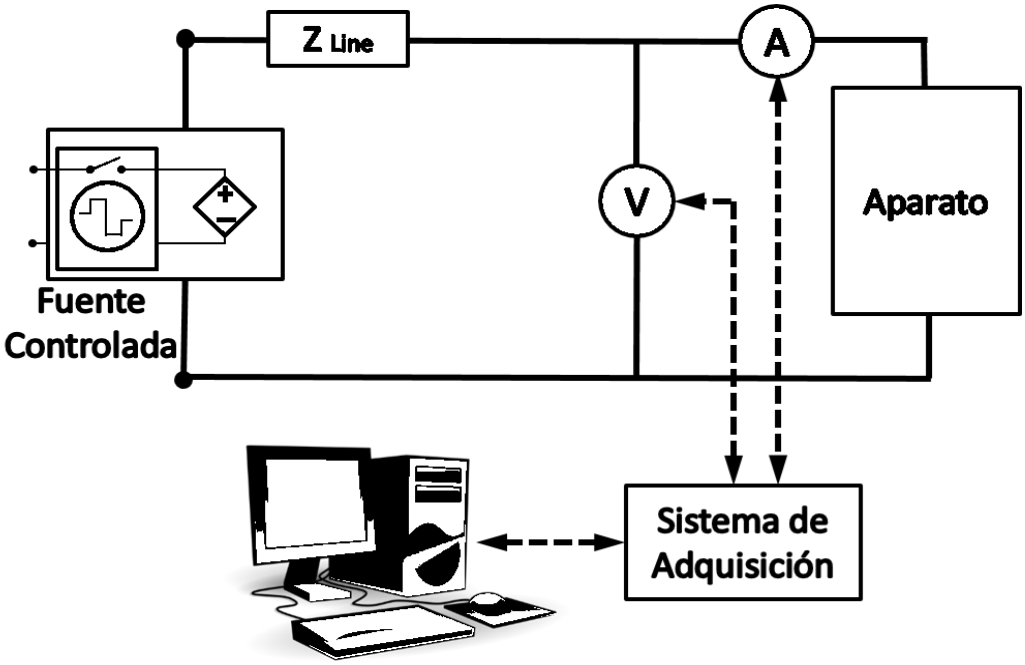


Figura 3. Formas de conexión para diferentes aparatos. Los aparatos pueden estar en modo individual o en combinaciones de 2 o 3 aparatos.

Chroma y mediante un interruptor se implementa la de conexión y desconexión del aparato. Referente a la señal de alimentación se utiliza una señal de 120 V de amplitud con 60Hz de frecuencia. Se midieron de forma individual y combinaciones entre ellos. En la Figura 4 se pueden observar el modelo circuital para un sólo aparato y el modelo circuital cuando se hace la conexión de varios aparatos. Finalmente, los aparatos que se consideran en este trabajo de investigación son seleccionados usando el criterio de frecuencia Rashid y Singh (2017), es decir, los aparatos de uso común y extendido en ambientes residenciales. En la Tabla 2 se muestra el listado de aparatos que se consideran en este trabajo de investigación donde algunos de ellos tienen dos estados de operación o más.

4.1.2 Construcción de la base de casos. La construcción de la base de casos tiene los siguientes componentes:

Tabla 2. *Listado de aparatos eléctricos de uso común en un ambiente residencial se consideran en este trabajo de investigación. Se presenta la potencia nominal que entrega el fabricante, los estados están definidos por la operación del aparato*

ID	APARATO	MARCA	POTENCIA [W]	ESTADOS	NATURALEZA
A1	Licuadora	Oster	450	S1 -S2	Inductivo
A2	Bombillo Incandescente	Sylvania	75	S1	Resistivo
A3	Bombillo Fluorescente	Daiku	20	S1	Electrónico
A4	Bombillo Fluorescente	Interlight	9	S1	Electrónico
A5	Cargador de Celular	Motorola	20	S1	Electrónico
A6	Computador de Escritorio	Dell	216	S1-S3	Electrónico
A7	Ventilador Silver Cool	Silver Cool	48	S1	Inductivo
A8	Nevera	Haceb	1150	S1-S2	Inductivo
A9	Secador de Cabello	Conair	1875	S1-S2	Resistivo - Inductivo
A10	Bombillo Halógeno	Opalux	50	S1	Resistivo
A11	Bombillo Halógeno	Opalux	70	S1	Resistivo
A12	Plancha	Recco	1200	S1	Resistivo
A13	Bombillo LED	Light	7	S1	Electrónico
A14	Cargador Portátil	Samsung	120	S1	Electrónico
A15	Cargador Portátil	Samsung	132	S1	Electrónico
A16	Cargador Portátil	Samsung	40	S1	Electrónico
A17	Cargador Portátil	Lenovo	65	S1	Electrónico
A18	Microondas	Samsung	1200	S1	Electrónico
A19	Sanduchera	Simply	750	S1	Resistivo
A20	Televisión TRC	Phillips	84	S1	Electrónico
A21	Televisión TRC	Sony	90	S1	Electrónico
A22	Cargador Portátil	Asus	65	S1	Electrónico

1. Tipo de conocimiento que se requiere organizar

2. Modo de organizar el conocimiento.

Si se refiere al tipo de conocimiento, la información se ha dispuesto usando la naturaleza de las señales obtenidas, esto significa que las mediciones se clasifican régimen transitorio y régimen estacionario tanto para aparatos individuales como para combinaciones de ellos. Las bases de conocimiento son función de las características extraídas de las mediciones previas. Para este caso se puede decir que el conocimiento extraído es obtenido usando transformaciones de tiempo-frecuencia y usando características temporales. En la Tabla 3 se muestran las bases de casos construidas en función de las características extraídas.

Tabla 3. *Bases de datos construidas usando diferentes métodos de extracción de características.*

ITEM	NOMBRE	CARACTERÍSTICAS EXTRAÍDAS	NATURALEZA	No CLASES	No EVENTOS
1	FVs	Normatividad IEEE 1459 (IEEE, 2010)	Estado Estacionario	18	1790
2	FVTime	Características tomadas de Jiménez (2018)	Estado Transitorio	23	1891
3	ST1	Descriptores estadísticos para cada columna de la matriz S completa (Jiménez, 2018)	Estado Transitorio	19	1891
4	ST2	Agregación a lo largo de perfiles de las frecuencias armónicas Jiménez (2018)	Estado Transitorio	19	1891
5	ST3	Agregación de frecuencias en potencias de 2 Jiménez (2018)	Estado Transitorio	19	1891
6	ST4	Agregación de frecuencias en potencias de 10 Jiménez (2018)	Estado Transitorio	19	1891
7	ST5	Proyección PCA Jiménez (2018)	Estado Transitorio	19	1891
8	ST6	Proyección PCA Jiménez (2018)	Estado Transitorio	19	1891
9	ST7	Descriptores estadísticos para cada fila de la matriz S completa Jiménez (2018)	Estado Transitorio	19	1891
10	Wavelet	Coeficientes de la transformada Wavelet	Estado Estacionario	62	12648
11	Wavelet	Coeficientes de la transformada Wavelet	Estado Transitorio	44	6875
12	FRFT	Órdenes de extracción de la FRFT (Cala, 2018)	Estado Estacionario	62	12648
13	FRFT	Órdenes de extracción de la FRFT(Cala, 2018)	Estado Transitorio	44	6875
14	S-Transform	Órdenes de extracción de la FRFT (Cala, 2018)	Estado Estacionario	62	12648
15	S-Transform	Órdenes de extracción de la FRFT (Cala, 2018)	Estado Transitorio	44	6875
16	IEEE1459	Normatividad IEEE 1459 (IEEE, 2010)	Estado Estacionario	13	1290
17	IEEE1459	Normatividad IEEE 1459 (IEEE, 2010)	Estado Transitorio	36	2190

Referente al modo de organización, las bases de casos son organizadas en matrices donde cada fila corresponde a la información de un evento y las columnas contienen la información de N características de acuerdo al respectivo método de extracción. Las bases de casos se dividen en *Casos Estacionarios* y *Casos Transitorios*. En los casos de régimen estacionario se encuentran las extraídas las bases de casos obtenidas a partir del estándar IEEE 1459, que son las propuestas en este trabajo de investigación, de la transformada de Wavelet, de la transformada S y de la de transformada Fraccional de Fourier, que son tomadas del trabajo de investigación (Cala, 2018). En los casos de régimen transitorio se encuentran las bases de casos obtenidas a partir del estándar IEEE 1459, propuesta en este trabajo de investigación, la transformada Wavelet y la Transformada Fraccional de Fourier, propuestas del trabajo de investigación desarrollado en (Cala, 2018) y la Transformada S, propuestas en (Cala, 2018) y (Jimenez, 2018). Además, se usaron, para la fase de reutilización y de razonamiento una base de casos llamada *Casos Representantes* que fue obtenida mediante la media aritmética de cada clase para todos los casos almacenados de cada clase y para reducción de dimensionalidad una base de datos llamada *Casos Iniciales* que aporta información del 95% de las muestras para tener un criterio inicial acerca de la clase a la que pertenecería un caso consulta y tener un criterio de decisión en la fase de razonamiento tanto en el método de adaptación por desviación estándar como el método de ajuste por centroide. En la Figura 5 se muestra la ontología de la base de casos que permite extraer conocimiento de acuerdo a los razonamientos tenidos en cuenta en la fase de revisión del modelo de razonamiento basado en casos. Cada caso se almacena mediante un vector con una dimensión igual al número de características extraídas en función del método y adicionando una etiqueta que define su clase, dado que este modelo considerado usa aprendizaje supervisado.

4.1.3 Extracción de Características. Una señal registrada por la tarjeta de adquisición tiene como característica principal su tiempo de duración. Esta puede ser dividida en dos regímenes. El régimen de estado transitorio, que ocurre justo después de la acción de encendido o apagado del aparato y el régimen de estado estacionario, que ocurre cuando la señal se estabiliza después del estado transitorio. Posterior a ello, se procede a realizar la extracción de información útil mediante algoritmos matemáticos. Finalmente, la información extraída es almacenada en un vector de características.

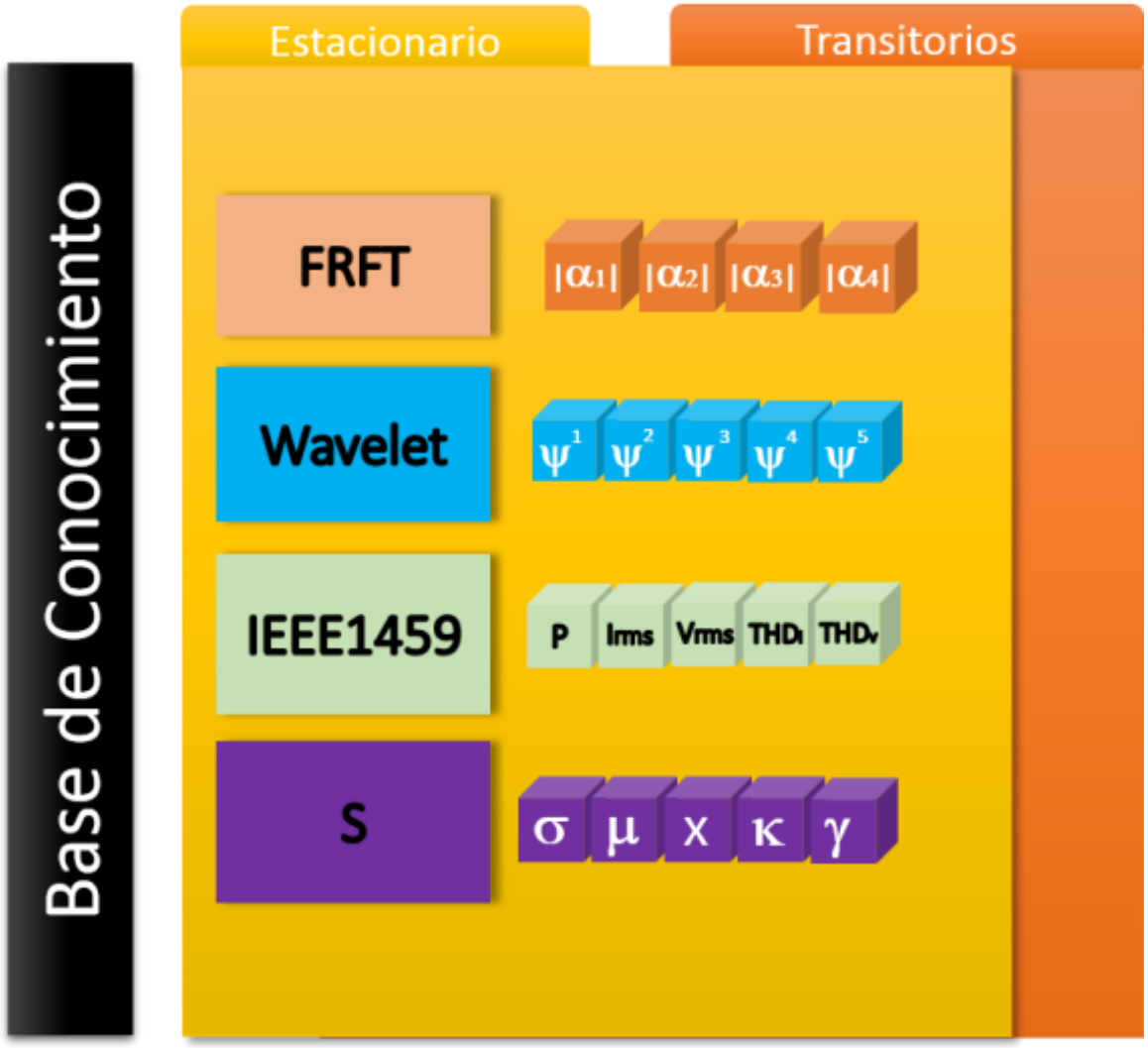


Figura 5. Ontología de la base de casos para los modelos de RBC propuestos.

Mediante un algoritmo implementado en MATLAB® se selecciona de la señal cruda cada uno de los regímenes, tanto de estado transitorio como de estado estacionario. La selección de los puntos se realiza visualmente y de forma manual, seleccionando los puntos de inicio y final tanto del régimen transitorio como del régimen estacionario. Luego, para cada régimen se extraen las características con base en el estado del arte de Monitorización No intrusiva de Carga. Usualmente, las propuestas para extracción de características en Monitorización No Intrusiva de Carga para señales en estado transitorio como en estado estable, obtienen información netamente numérica, por lo que la indexación de la base de casos se toma como un vector de características de valor numérico. Finalmente, con las características extraídas en cada uno de los regímenes se construyen las 17 bases de casos que sirven de conocimiento para someter a prueba los 24 modelos de RBC que se diseñaron. Para mayor detalle, en el apéndice C, para cada una de las bases de casos seleccionadas se muestran cuáles fueron específicamente las características extraídas y de qué literatura fueron consultadas.

4.1.3.1 Indexación del Conocimiento. En esta sección se presentan las notaciones que seguirán las clases, los aparatos eléctricos, sus estados y los tipos de señales de cada uno de los regímenes considerados en este trabajo de investigación. Estas notaciones se presentan en la Tabla 4 y en la Tabla 5, explicando con detalle el significado de cada una de ellas.

Con respecto al modo operativo de cada aparato, se consideran los estados de operación de cada uno de los aparatos según su naturaleza. Para este caso, se clasifican en Electrónicos, Inductivos y Resistivos. De acuerdo a la naturaleza se tipificaron los posibles estados. La notación SX , donde S significa 'ESTADO' y X es el número del estado que va desde 1 a 4 indicando el estado de operación del aparato. Por ejemplo, la Nevera se considera con dos estados de operación, Estado 1 (S1) se refiere al Ciclo de trabajo de la nevera y el Estado 2 (S2) se refiere a la apertura de la

Tabla 5. *Estados de operación considerados para aparatos electrónicos, inductivos y resistivos. Se mencionan algunos de los estados de operación de algunos aparatos.*

		NATURALEZA		
		ELECTRÓNICOS	**INDUCTIVOS**	**RESISTIVOS**
ESTADOS	**S1**	X	X	X
	S2	X	X	N/A
	S3	X	N/A	N/A
	S4	X	N/A	N/A

puerta. En la Tabla 4 pueden apreciarse los estados de operación tipificados. En cuanto al tipo de señal obtenida, se consideran los regímenes transitorios de encendido-apagado, con notación (TS1), transitorio de apagado-encendido, con notación (TS2) y estacionario, con notación (SS1), tanto para aparatos conectados de forma individual como en combinaciones de grupos de 2 y de 3, respectivamente. Para las mediciones registradas en las combinaciones, se plantean dos posibles

Tabla 4. *Descripción de los regímenes de operación de los aparatos.*

	RÉGIMEN
SS1	Estado Estacionario
TS1	Transitorio ON-OFF
TS2	Transitorio OFF-ON
TS3	Transitorio Intermedios

situaciones: o todos se encuentran en estado estacionario o alguno de los aparatos de la combinación se encuentran en estado transitorio, ya sea transitorio de encendido-apagado o transitorio de apagado-encendido, y los restantes en estado estacionario. En la Tabla 5, se muestran los tipos de regímenes que se extrajeron como información para clasificación de los aparatos. Finalmente, las mediciones registradas se clasifican, usando la cantidad de aparatos. En la Tabla 6, Tabla 7 y Tabla 8, se muestran los aparatos como clase en forma individual y combinaciones en grupos de 2 y 3 que se consideran en este trabajo de investigación.

Tabla 6. *Características de las combinaciones de dos aparatos eléctricos que fueron utilizadas en las mediciones. Se presenta el grupo de dos aparatos y los estados de funcionamiento definidos por la naturaleza de los aparatos y los transitorios que fueron considerados. Para combinaciones de 2 aparatos, solamente se consideran transitorios de encendido o de apagado para uno solo de los dos aparatos que participan en la combinación.*

ID	APARATOS	POTENCIA [W]	ESTADOS	TRANSITORIO
A1A4	Licuadora Bombillo Fluorescente	459	S1: Velocidad 2 S1: Encendido - Apagado	**TS1-TS2** **TS1-TS2**
A1A5	Licuadora Cargador de Celular	470	S1: Velocidad 2 S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A1A7	Licuadora Ventilador	498	S1: Velocidad 2 S1: Velocidad 1	**TS1-TS2** **TS1-TS2**
A1A8	Licuadora Nevera	1600	S1: Velocidad 2 S1: Apertura Puerta	**TS1-TS2** **TS1-TS2**
A1A11	Licuadora Bombillo Halógeno	520	S2: Velocidad 2 S1: Encendido - Apagado	**TS1-TS2** **TS1-TS2**
A1A12	Licuadora Plancha	1650	S2: Velocidad 2 S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A1A21	Licuadora Televisor TRC	540	S2: Velocidad 2 S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A4A5	Bombillo Fluorescente Cargador de Celular	29	S1: Encendido - Apagado S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A4A7	Bombillo Fluorescente Ventilador	57	S1: Encendido - Apagado S1: Velocidad 1	**TS1-TS2** **TS1-TS2**
A4A8	Bombillo Fluorescente Nevera	1159	S1: Encendido - Apagado S1: Apertura Puerta	**TS1-TS2** **TS1-TS2**
A4A11	Bombillo Fluorescente Bombillo Halógeno	79	S1: Encendido - Apagado S1: Encendido - Apagado	**TS1-TS2** **TS1-TS2**
A4A12	Bombillo Fluorescente Plancha	1209	S1: Encendido - Apagado S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A4A21	Bombillo Fluorescente Televisor TRC	99	S1: Encendido - Apagado S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A5A7	Cargador de Celular Ventilador	68	S1: Conexión - Desconexión S1: Velocidad 1	**TS1-TS2** **TS1-TS2**
A5A8	Cargador de Celular Nevera	1170	S1: Conexión - Desconexión S1: Apertura Puerta	**TS1-TS2** **TS1-TS2**
A5A11	Cargador de Celular Bombillo Halógeno	90	S1: Conexión - Desconexión S1: Encendido - Apagado	**TS1-TS2** **TS1-TS2**
A5A12	Cargador de Celular Plancha	1220	S1: Conexión - Desconexión S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A5A21	Cargador de Celular Televisor TRC	10	S1: Conexión - Desconexión S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A7A8	Ventilador Nevera	1198	S1: Velocidad 1 S1: Apertura Puerta	**TS1-TS2** **TS1-TS2**
A7A11	Ventilador Bombillo Halógeno	118	S1: Velocidad 1 S1: Encendido - Apagado	**TS1-TS2** **TS1-TS2**
A7A12	Ventilador Plancha	1248	S1: Velocidad 1 S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A7A21	Ventilador Televisor TRC	138	S1: Velocidad 1 S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A8A11	Nevera Bombillo Halógeno	1220	S1: Apertura Puerta S1: Encendido - Apagado	**TS1-TS2** **TS1-TS2**
A8A12	Nevera Plancha	2350	S1: Apertura Puerta S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A8A21	Nevera Televisor TRC	1240	S1: Apertura Puerta S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A11A12	Bombillo Halógeno Plancha	1270	S1: Encendido - Apagado S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A11A21	Bombillo Halógeno Televisor TRC	160	S1: Encendido - Apagado S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**
A12A21	Plancha Televisor TRC	1290	S1: Conexión - Desconexión S1: Conexión - Desconexión	**TS1-TS2** **TS1-TS2**

Tabla 8. *Características de las combinaciones de tres aparatos eléctricos que fueron utilizadas en las mediciones. Se presenta el grupo de tres aparatos y los estados de funcionamiento definidos por la naturaleza de cada aparato y los transitorios que fueron considerados. Para combinaciones de 3 aparatos, solamente se consideran transitorio de encendido o de apagado para uno solo de los 3 aparatos que participan en la combinación, los restantes están en régimen estacionario.*

ID	APARATOS	POTENCIA [W]	ESTADOS	TRANSITORIO
A1A2A4	Licuadora Bombillo Incandescente Bombillo Fluorescente	534	S1: Velocidad 1 S1: Encendido - Apagado S1: Encendido - Apagado	TS1 TS1 TS1
A1A3A12	Licuadora Bombillo Fluorescente Plancha	1670	S1: Velocidad 1 S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A1A3A22	Licuadora Bombillo Fluorescente Cargador de Portátil	535	S1: Velocidad 1 S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A1A19A21	Licuadora Sanduchera Televisor TRC	1290	S1: Velocidad 1 S1: Conexión - Desconexión S1: Conexión - Desconexión	TS1 TS1 TS1
A2A3A22	Bombillo Incandescente Bombillo Fluorescente Cargador de Portátil	160	S1: Encendido - Apagado S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A2A4A11	Bombillo Incandescente Bombillo Fluorescente Bombillo Halógeno	154	S1: Encendido - Apagado S1: Encendido - Apagado S1: Encendido - Apagado	TS1 TS1 TS1
A2A4A19	Bombillo Incandescente Bombillo Fluorescente Sanduchera	834	S1: Encendido - Apagado S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A2A19A21	Bombillo Incandescente Sanduchera Televisor TRC	915	S1: Encendido - Apagado S1: Conexión - Desconexión S1: Conexión - Desconexión	TS1 TS1 TS1
A3A11A13	Bombillo Fluorescente Bombillo Halógeno Bombillo LED	97	S1: Encendido - Apagado S1: Encendido - Apagado S1: Encendido - Apagado	TS1 TS1 TS1
A3A12A13	Bombillo Fluorescente Plancha Bombillo LED	1227	S1: Encendido - Apagado S1: Conexión - Desconexión S1: Encendido - Apagado	TS1 TS1 TS1
A7A11A13	Ventilador Bombillo Halógeno Bombillo LED	125	S1: Velocidad 1 S1: Encendido - Apagado S1: Encendido - Apagado	TS1 TS1 TS1
A7A11A19	Ventilador Bombillo Halógeno Sanduchera	868	S1: Velocidad 1 S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A7A11A21	Ventilador Bombillo Halógeno Televisor TRC	208	S1: Velocidad 1 S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A8A11A22	Nevera Bombillo Halógeno Cargador de Portátil	1285	S1: Apertura Puerta S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1
A11A13A21	Bombillo Halógeno Bombillo LED Televisor TRC	167	S1: Encendido - Apagado S1: Encendido - Apagado S1: Conexión - Desconexión	TS1 TS1 TS1

4.1.4 Definición de los Descriptores. Como se ha mencionado, se dispone de 17 bases de conocimiento que permitirán tomar decisiones sobre el desempeño del modelo de razonamiento basado en casos. Estás bases se encuentran conformadas por descriptores calculados usando las definiciones de mediciones de potencia eléctrica para una señal de referencia sinusoidal de acuerdo a la norma IEEE1459, descriptores calculados usando la Transformada S, descriptores calculados usando la transformada Wavelet y descriptores calculados usando la transformada fraccionaria de Fourier. A continuación de explicarán los descriptores que fueron considerados y las clases definidas en cada base de casos.

4.1.4.1 ***Base de Casos FVs construida de Mediciones en Régimen Estacionario.*** Para la base de casos llamada **FVs**, los descriptores desde el 1 al 21, fueron tomados de (Jimenez, 2018). Los descriptores desde el 22 hasta el 37 fueron tomados de IEEE 1459 Standart de calidad de potencia (IEEE, 2010). Tiene un total de 1790 ejemplos, dividido en 18 clases representadas por los eventos estacionarios de los aparatos de forma individual. La base de conocimiento está conformada por una matriz numérica de 1790 filas por 37 columnas. Para más detalles, en la Tabla 21 se muestran la definición de cada descriptor considerado, mientras que en la Tabla 15 se muestran los aparatos individuales y sus estados seleccionados.

4.1.4.2 ***Base de Casos FVTime construida de Mediciones en Régimen Transitorio***. Para la base de casos llamada **FVTime**, los descriptores fueron tomados de (Jimenez et al., 2015), en donde se realiza el análisis usando medidas estadísticas (Promedio, Kurtosis, Máximo, Factor de Cresta, Desviación Estándar, Entropía entre otras) para señales en régimen transitorio. Tiene un total de 1891 ejemplos, divididos en 19 clases representadas por los eventos transitorios de encendido (TS1 OFF-ON) de los aparatos en forma individual.

Para mayor detalle, en la Tabla 22 se muestran los descriptores estadísticos y la definición matemática para su obtención, mientras que en la Tabla 19 se muestran los aparatos individuales y sus estados seleccionados.

4.1.4.3 *Base de Casos ST1 a ST4 y ST7 - construida usando Métodos de Agregación de Mediciones en Régimen Transitorio.*

Para las bases de casos llamadas ST1, ST2, ST3, ST4 y ST7, Los descriptores fueron tomados de (Jimenez, 2018) y calculados usando medidas estadísticas de la matriz de tiempo-frecuencia obtenida de la transformada S. Se usaron dos tipos de métodos para reducción de dimensionalidad: *agregación* y *proyección*. Los métodos de agregación se enfocan en describir el contenido de las matrices de tiempo-frecuencia usando medidas estadísticas, así como algunos estadísticos usados para la construcción de la base de casos FVTime para eventos transitorios. Algunos trabajos (Altrabalsi, Stankovic, Liao y Stankovic, 2016), (Bhotto, Makonin y Bajić, 2017) han empleado estos métodos de agregación. Para estas bases de casos se usan los métodos de agregación y en la Tabla 9 se da una breve explicación sobre estos descriptores. Sin embargo, para mayor detalle sobre cómo fueron calculados los descriptores usando los métodos de agregación, puede ser consultado en (Jimenez, 2018). Tiene un total de 1891 ejemplos, divididos en 19 clases representadas por los eventos transitorios de encendido (TS1 - OFF-ON) de los aparatos. En la Tabla 19 se muestran los aparatos individuales y sus estados seleccionados.

4.1.4.4 *Base de Casos ST5-ST6 construida usando Métodos de Proyección de Mediciones en Régimen Transitorio.*

Los descriptores fueron tomados de (Jimenez, 2018). Las bases de casos están compuestas por 1891 ejemplos y por 19 clases. En la Tabla 16 se muestran los aparatos y los estados que fueron tenidos en cuenta. Para la construcción de las bases de casos se usaron dos métodos para reducción de dimensionalidad mediante proyección.

Los métodos de proyección representan la información de la matriz original en un vector. Los trabajos (Bhotto et al., 2017) y (Initiative y others, 2016) han empleado estos métodos de proyección mediante Análisis de Componentes Principales (PCA) o Análisis de Discriminante Lineal (LDA). La información de los métodos de proyección y cómo fueron calculados los descriptores se puede ampliar en (Jimenez, 2018).

Tabla 9. *Descriptores usados para la conformación de la base de casos llamada ST para estado transitorio.*

BASE DE CASOS	DESCRIPCIÓN	DIMENSIÓN
ST1	Agregación a lo largo de las columnas de la matriz de la transformada S completa.	1891 ejemplos con 24 descriptores
ST2	Agregación a lo largo de los perfiles de las frecuencias armónicas.	1891 ejemplos con 280 descriptores
ST3	Agregación a lo largo de frecuencias en escala de potencias de dos.	1891 ejemplos con 63 descriptores
ST4	Agregación a lo largo de frecuencias en escala de potencias de diez.	1891 ejemplos con 21 descriptores
ST7	Agregación a la largo de las filas de la matriz de la transformada S completa	1891 ejemplos con 21 descriptores

4.1.4.5 *Base de Casos WAVELET, S-TRANSFORM Y FRTF construidas usando Transformada Wavelet, S y Fourier Fraccionaria en Mediciones de Regímenes Estacionario y Transitorio*

. Los descriptores de las bases de casos construidas con el método de extracción Wavelet se proponen en (Duarte et al., 2012) y se encuentran en la Tabla 3 como los ítemes 10 y 11. Este método calcula la contribución de energía de un conjunto de escalas de acuerdo a unos parámetros de diseño. Sus parámetros de diseño son: Frecuencia central de la función Wavelet madre, Fc, ancho de banda de 3 dB de la Wavelet madre, F 3dB y el número de escalas seleccionados con

Tabla 10. *Descriptores usados para la conformación de la base de casos llamada ST para régimen transitorio.*

BASE DE CASOS	DESCRIPCIÓN	DIMENSIÓN
ST5	Proyección mediante Análisis de Componentes Principales.	1891 ejemplos con 24 descriptores
ST6	Proyección mediante Análisis de Discriminante Lineal.	1891 ejemplos con 2502 descriptores

base en las propiedades de los eventos estacionarios y transitorios. Para los eventos estacionarios se calculan como descriptores 40 escalas dando como resultado una base de casos de 12648 casos con 40 descriptores conformada por 62 clases, mientras que para los eventos transitorios se calculan 75 escalas, dando como resultado una base de casos de 6875 casos con 75 descriptores conformada por 44 clases.

Los descriptores usados para régimen estacionario y transitorio son mostrados en la Tabla 23, mientras que las clases son mostradas en la para aparatos individuales y en la Tabla 18 para aparatos en combinados de a 2 y Tabla 19 para aparatos combinados de a 3 en régimen estacionario, y Tabla 20 para individuales y combinados para régimen transitorio. En la Tabla 11 se hace un sumario de las etiquetas consideradas.

Para los descriptores usados en las bases de casos donde utiliza la técnica de extracción de características propuesta en Jimenez et al. (2014) (referenciadas en la Tabla 3 como S-Transform ítemes 14 y 15), en las que se obtienen parámetros estadísticos de tiempo-frecuencia de transformada S y se usa el algoritmo propuesto en (Stockwell, Mansinha y Lowe, 1996) con frecuencias de análisis seleccionadas con base en las propiedades de los eventos transitorios y estacionarios. De las 7 características propuestas en (Jimenez et al., 2014) se utilizan las 6 primeras características para el sistema de régimen estacionario dando como resultado una base de casos de 12648 casos con 6 descriptores conformadas por 62 clases. El rango de frecuencias en la transformada S es de 0[Hz] hasta 10[kHz] para 100 frecuencias equidistantes. Para el régimen transitorio se utilizan las 7 características propuestas en (Jimenez et al., 2015) pero con un rango de frecuencias de 0[Hz] hasta 25[kHz] para 100 frecuencias equidistantes, dando como resultado una base de casos de 6875 casos con 6 descriptores conformada por 44 clases. Los descriptores usados para régimen estacionario y transitorio son mostrados en la Tabla 10, mientras que las clases son mostradas en la para

aparatos individuales, en la Tabla 18 para aparatos en combinados de a 2 y Tabla 19 para aparatos combinados de a 3 en régimen estacionario y en la Tabla 20 para individuales y combinados para régimen transitorio. En la Tabla 11 se hace un sumario de las etiquetas consideradas.

Por último, para los descriptores usados en las bases de casos donde se utiliza la técnica de Transformada Fraccional de Fourier propuesta en (Cala, 2018) (referenciadas en la Tabla 3 como FRTF ítemes 12 y 13), se obtienen los órdenes fraccionarios óptimos utilizando la técnica de extracción de Wigner-Ville y el centroide espectral fraccionario. Para el régimen estacionario se extraen 20 descriptores dando como resultado una base de casos de 12648 casos con 20 descriptores conformada por 62 clases, mientras que para régimen transitorio se extrajeron 17 descriptores, dando como resultados una base de casos de 6875 casos con 17 descriptores conformadas por 44 clases. Los descriptores usados para régimen estacionario y transitorio son mostrados en la Tabla 25, mientras que las clases son mostradas en la para aparatos individuales, en la Tabla 18 para aparatos en combinados de a 2 y en la Tabla 19 para aparatos combinados de a 3 en régimen estacionario y en la Tabla 20 para individuales y combinados para régimen transitorio. En la Tabla 11 se hace un sumario de las etiquetas consideradas en la Tabla 20.

4.1.4.6 *Base de casos IEEE 1459- construida usando las definiciones de la Norma IEEE1459 para Medidas en Régimen Estacionario y Transitorio*

. Los descriptores de las bases de casos llamadas IEEE1459 que se calculan con las definiciones de la norma IEEE 1459 (referenciadas de la Tabla 3 como IEEE1459 ítemes 16 y 17), se calculan usando las definiciones de mediciones de potencia eléctrica para una señal de referencia sinusoidal. Para los eventos estacionarios se calculan 15 descriptores, dando como resultado una base de casos de 2190 casos con 15 descriptores conformada por 36 clases, mientras que para los eventos transitorios se calculan 7

descriptores dando como resultados una base de casos de 1290 casos con 7 descriptores conformada por 13 clases. Los descriptores usados para régimen estacionario y transitorio son mostrados en la Tabla 26, mientras que las clases son mostradas en la Tabla 21. y Tabla 22.

4.1.5 Selección de Características. Luego de la extracción de las diferentes características, es necesario decidir cuáles características son las que aportan la mayor información y cuáles tienen redundancias con el fin de reducir la dimensionalidad del vector de características y por lo tanto reducir la complejidad computacional. Para lograr este propósito, se utiliza de Análisis de Componentes Principales (ACP) para la reducción de dimensionalidad. En la Tabla 12 se presentan los resultados de las componentes principales extraídas en cada una de las bases de casos que garantizan el 95% de la varianza, siguiendo la regla establecida en (Jackson, 1993).

4.2 Procesamiento

En el preprocesamiento se analizan los datos registrados con las tarjetas de adquisición NI9127 y NI9137 y se ajustan a las entradas del modelo de razonamiento basado en casos. El modelo de razonamiento basado en casos está compuesto por 4 fases. La fase de recuperación, la fase de reuso, la fase de revisión y la fase de retención. Para cada fase se implementaron dos métodos. En este trabajo de investigación se estudia el desempeño que se obtiene al realizar la combinación de los métodos usados en cada fase y se identifica qué modelo de razonamiento basado en casos es el más adecuado para realizar la clasificación de eventos en Monitorización no intrusiva de carga.

Tabla 11. *Clases definidas para los aparatos individuales y combinaciones de 2 y 3 aparatos para la construcción de la base de casos Wavelet, Fractional y S en régimen transitorio. TSX significa que es en estado transitorio. Para mayor información de las etiquetas se puede consultar en la*
Tabla 20.

CLASE	ETIQUETA	EVENTOS POR CLASE
1	TS1A1-S1	180
2	TS1A1-S2	150
3	TS1A10-S1	99
4	TS1A11-S1	314
5	TS1A12-S1	200
6	TS1A13-S1	190
7	TS1A19-S1	180
8	TS1A2-S1	197
9	TS1A21-S1	140
10	TS1A22-S1	60
11	TS1A3-S1	201
12	TS1A4-S1	228
13	TS1A5-S1	170
14	TS1A6-S1	100
15	TS1A6-S2	100
16	TS1A6-S3	100
17	TS1A7-S1	224
18	TS1A8-S1	170
19	TS1A8-S2	100
20	TS1A9-S1	100
21	TS1A9-S2	100
22	TS2A1-S1	180
23	TS2A1-S2	170
24	TS2A10-S1	99
25	TS2A11-S1	314
26	TS2A12-S1	200
27	TS2A13-S1	190
28	TS2A19-S1	180
29	TS2A2-S1	197
30	TS2A21-S1	150
31	TS2A22-S1	60
32	TS2A3-S1	201
33	TS2A4-S1	238
34	TS2A5-S1	170
35	TS2A6-S1	100
36	TS2A6-S2	100
37	TS2A6-S3	100
38	TS2A7-S1	234
39	TS2A8-S1	189
40	TS2A8-S2	100
41	TS2A9-S1	100
42	TS2A9-S2	100
43	TS3A6-S2	100
44	TS4A6-S2	100

Tabla 12. *En la columna 3 se presentan las componentes principales extraídas que recoge el 95% de la información de todos los descriptores para cada una de las bases de casos consideradas. Para los casos en que no haya una cantidad de componentes principales que recoja el 95%, el entrenamiento se realiza con los descriptores originales.*

ÍTEM	BASES DE CASOS	COMPONENTES PRINCIPALES EXTRAÍDAS
1	FVs	Las primeras 37 componentes
2	FVTime	Las primeras 9 componentes
3	ST1	Ninguna componente
4	ST2	Las primeras 278 componentes
5	ST3	Las primeras 63 componentes
6	ST4	Las primeras 21 componentes
7	ST5	Las primeras 1885 componentes
8	ST6	Las primeras 1860 componentes
9	ST7	Ninguna componente
10	Wavelet	Las primeras 39 componentes
11	Wavelet	Las primeras 5 componentes
12	FRFT	las primeras 15 componentes
13	FRFT	las primeras 12 componentes
14	S-Transform	Las primeras 2 componentes
15	S-Transform	Las primeras 73 componentes
16	IEEE1459	Las primeras 12 componentes
17	IEEE1459	Las primeras 10 componentes

4.2.1 Algoritmos de Recuperación. De acuerdo con la literatura revisada, para la implementación de la fase de recuperación se utiliza el método de los k-vecinos cercanos mediante la métrica Euclidiana y la métrica Minkowsky. Más formalmente, estas operaciones de similaridad pueden ser expresadas de la siguiente manera. Sea $CB = e_1,\ e_2, ..., e_N$ una librería que tiene a su disposición N casos. Como cada caso está identificado por una indexación que corresponde a sus características, se usa una colección de características $f_j (j = 1,\ 2, ..., d)$ y un operador $V(Class)$ que asigna una clase al caso en mención. El i-ésimo caso puede ser representado como un vector con dimensión $(d+1)$, esto es, $e_i = \left(f_1^{(i)}, f_2^{(i)}, f_3^{(i)}....., f_j^{(i)}.....f_d^{(i)}, C_i \right)$, donde $f_j^{(i)}$ corresponde al valor de la característica $f_j (1 < j < d)$ y C_i a la clase correspondiente al valor del operador $V_{\square}(Class)\ (1, 2, ..., N)$.

4.2.1.1 ***Distancia Euclidiana.*** Con los datos previamente normalizados, el cálculo de la distancia se hace punto a punto del caso consulta con respecto a cada uno de los casos que se encuentran almacenados en la base de casos. Definiendo formalmente, sea $CB = e_1, e_2, ..., e_N$ una librería de conocimiento disponible y sea e_q un nuevo caso que debe ser analizado. Dados un operador $\mathcal{F}(Ind)$ que asigna del conjunto de características $f_j (j = 1, 2, ..., d)$ y un operador $\mathcal{V}$ que asigna una clase C_k del conjunto $(Class)(1, 2, ..., N)$, se establece para el caso *i-ésimo* almacenado en la base de casos como $e_i = \left(f_1^{(i)}, f_2^{(i)}, f_3^{(i)}....., f_j^{(i)}.....f_d^{(i)}, C_k \right)$. Siguiendo las reglas de indexado establecidas por la base de conocimiento, se define el caso consulta como $e_q = \left(f_1^{(i)}, f_2^{(i)}, f_3^{(i)}....., f_j^{(i)}.....f_d^{(i)}, S \right)$ donde S es la clase solución que debe ser encontrada con $S \in V_k(Class)$. La distancia Euclidiana entre un caso consulta y un caso i-ésimo almacenado en la librería de conocimiento se define como lo mostrado en la Ecuación 3.

$$D_{Euclidiana}(e_q, e_i) = \sqrt{\sum_{j=1}^{N} (eq_j - ei_j)^2}$$

Ecuación 3

4.2.1.2 ***Distancia Minkowski.*** Para esta distancia los datos se normalizan y el cálculo de la distancia se hace punto a punto respecto a todas las características seleccionadas. Siguiendo un razonamiento similar a la distancia Euclidiana, la distancia de Minkowski de parámetro $p = 2$, entre el caso consulta y un caso almacenado en la librería de conocimiento se define como en la Ecuación 4.

$$D_{Minkoswki}(e_q, e_j) = \left(\sum_{j=1}^{N} \left| eq_j - ei_j \right|^p \right)^{\frac{1}{p}}$$

Ecuación 4

El algoritmo de recuperación calcula la distancia entre el caso consulta y cada uno de los casos almacenados en la base de casos. Siguiendo la premisa del razonamiento basado en casos que menciona *La solución de un nuevo caso es similar a los casos almacenados*, definiendo la medida de similitud entre cada uno de los casos de la librería de casos y el caso consulta son ordenados de acuerdo al criterio de la mínima distancia. Posteriormente, se selecciona el 10 % de los casos almacenados. Este subconjunto de casos recuperado mediante el algoritmo de recuperación se conoce como *casos promisorios*. Son los casos promisorios los que pasan a la siguiente fase del RBC para ser utilizados.

4.2.2 Algoritmos de Reutilización. La fase de reutilización se centra en cómo definir cuál es la solución que se debería reutilizar. Esto significa que al definir el conjunto de *casos promisorios* y puesto que se trabaja con aprendizaje supervisado, los casos promisorios tienen una etiqueta asociada que daría la solución al caso consulta. Esto es, de la librería de casos $CB = e_1, e_2, ..., e_N$ debe ser asignada del conjunto de las K-clases $C_j (j = 1, 2, ...m, ..., K)$ una etiqueta. Para esto el operador $\mathcal{V}(Class)$ asigna una clase a cada caso de la librería de casos. Dentro de este conjunto de *casos promisorios* se plantean los siguientes escenarios:

1. La totalidad de los casos promisorios pertenecen a una única etiqueta y esta es la solución correcta del caso consulta.
2. La totalidad de los casos promisorios pertenecen a una única etiqueta, pero esta es una solución errada del caso consulta.

3. Los casos promisorios pertenecen a diferentes clases y exista una clase que tenga mayor frecuencia absoluta que las demás.

4. Los casos promisorios pertenecen a diferentes clases y dos clases o más tienen la misma frecuencia absoluta.

Estos escenarios son tratados en la fase de revisión, puesto que a partir de ellos viene el razonamiento. Regresando a la fase de reutilización, dos algoritmos se consideran para reutilizar las etiquetas de los casos promisorios y decidir una solución parcial para el caso consulta.

4.2.2.1 ***Regla de Votación por Mayoría.*** Para este caso, sea el conjunto de casos promisorios como $CP = p_1, p_2, ..., p_S$ y sea $\mathcal{V}(Class)$ el operador que asigna una clase a cada caso de la librería de casos. Las etiquetas de clase de los *S* casos promisorios son agrupadas en F conjuntos utilizando el operador $G_f(CP)(f = 1, 2, ..., f, ...S)$ que asigna cada caso promisorio a un conjunto G. Se calcula el cardinal para cada conjunto G_f, definido como $n(G_f)$. Inicialmente, la solución copiada para el caso consulta se asigna mediante el operador $\mathcal{V}(Class)$ del conjunto G que presenta la máxima cardinalidad de los F conjuntos, asumiendo que la cardinalidad de los F conjuntos es diferente. Si existen conjuntos que tienen la misma cardinalidad, la solución copiada al caso consulta es seleccionada de forma aleatoria. Formalmente, el operador $V(e_q)$ asignará una clase al caso consulta como lo mostrado en la Ecuación 5.

$$V(e_q) = \begin{cases} V(\max(n(\mathrm{G}_1), f_a(\mathrm{G}_2), ..., n(\mathrm{G}_s))) & \text{si } n(\mathrm{G}_1) \neq n(\mathrm{G}_2)...n(\mathrm{G}_f) \\ V(\mathrm{rand}(F)) & \text{si } n(\mathrm{G}1) = n(\mathrm{G}2)...n(\mathrm{G}_f) \end{cases}$$

Ecuación 5

4.2.2.2 ***Regla basada en Clase.*** Para este caso, sea el conjunto de casos representantes como $CR = r_1, r_2, ..r_l., r_K$ y donde el l-ésimo caso representante está definido como el promedio aritmético de todos los casos que pertenecen a la clase l. La solución $\mathcal{V}(e_q)$ es la clase que tenga la menor distancia desde los casos representantes hasta el caso consulta. Formalmente, el operador $V(e_q)$ asignará una clase al caso consulta como lo mostrado en la Ecuación 6.

$$V(e_q) = V(\min(d(r_1, e_q), d(r_2, e_q), ..., d(r_l, e_q), ..., d(r_k, e_q)))$$

Ecuación 6

Donde $d(r_l, e_q)$ es la distancia entre del caso promedio con clase l y el caso consulta.

4.2.3 Algoritmos de Revisión. La fase de revisión permite asegurar que la solución reutilizada (copiada) sea la correcta. Esta fase de revisión es también conocida como adaptación y permite ajustar un caso promisorio cuando la solución reutilizada sea difusa y no pueda establecerse una respuesta definitiva. Para este modelo de razonamiento basado en casos se utilizaron dos técnicas de adaptación. La primera realiza la adaptación de características ruidosas usando la desviación estándar. La segunda es un algoritmo de razonamiento utilizando el centroide de las tres clases de mayor frecuencia.

4.2.3.1 ***Adaptación por desviación estándar.*** Esta metodología se basa en adaptar parte de los descriptores que se encuentran fuera de una desviación estándar y cuyos límites son definidos con la información recolectada de la librería de casos. Está propuesta es tomada del método de *Adaptación por reinstanciación* (Balázs, Imre y Lajos, 1999). Esto es, sea e_m el caso m-ésimo de clase C_j y sea $f_i^{(m)} \in e_m$ donde m es el caso m-ésimo e i es el descriptor i-ésimo. para todo $f_i^{(m)}$ que

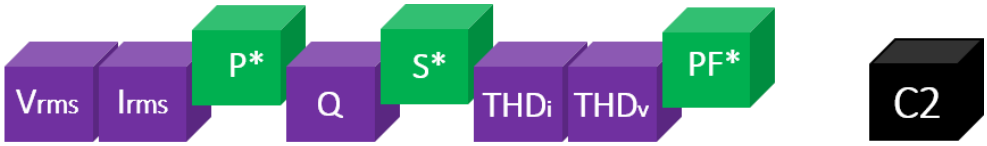
no pertenece al conjunto con límites $[\mu_i - \sigma_i, \mu_i + \sigma_i]$ se genera un número aleatorio con distribución normal que se encuentre dentro de los límites de conjunto $[\mu_i - \sigma_i, \mu_i + \sigma_i]$. A esta acción se le conoce como adaptación de los descriptores y permite adaptar los *casos promisorios* cuando las clases asignadas a dicho conjunto tiene datos multimodales. En la Figura 6 se ilustra el proceso de adaptación por desviación estándar.

4.2.3.2 ***Adaptación por Centroide***. Después de la fase de recuperación y luego de utilizar las dos métricas de similaridad, la distancia Euclidiana y la distancia Minkowski, se genera la base de casos llamada *casos promisorios*. Cada uno de los *casos promisorios* tienen una clase asociada desde el momento de su inclusión en la base de conocimiento dado el operador $\mathcal{V}$. Cuando se listan los casos promisorios, pueden ocurrir las condiciones consideradas en la fase de reutilización. En esos escenarios donde el caso consulta puede pertenecer a dos o más clases, se plantea la adaptación por centroide. De las 3 clases que tienen la máxima cardinalidad se extraen todos los casos que se encuentren en el círculo de radio σ y centro μ. Por ejemplo, para la clase 1, con la primera máxima cardinalidad, se calcula la media μ_1 y su desviación estándar σ_1 de todos los casos almacenados en la base de casos.

Luego, de todos los d descriptores de cada caso, se seleccionan dos descriptores aleatorios, llamados j,k, y se traza un círculo con radio $\sigma_1(j,k)$ y centro $\mu_1(j,k)$. Dentro del círculo se encuentra el 95% de los casos para la clase 1. El proceso se repite para las clases 2 y 3 que son las siguientes clases que tienen la máxima cardinalidad. La situación mencionada es ilustrada en la Figura 7(a). Este procedimiento se repite por un total veces como tantas combinaciones en grupos de 2 descriptores se calculen.

(a) Se tiene un caso promisorio con los descriptores mostrados.

(b) Se ha calculado previamente la desviación estándar por descriptor y se han seleccionado los descriptores que se encuentran por fuera a una desviación estándar. Por ejemplo, en este caso, los descriptores P, S y PF deben ser adaptados.



(c) Los descriptores son adaptados usando un valor aleatorio con distribución uniforme que pertenezca al conjunto $(\mu_i - \sigma_i, \mu_i + \sigma_i)$ del descriptor analizado.

Figura 6. Procedimiento para la adaptación usando desviación estándar de los casos promisorios seleccionados en la fase de reutilización.

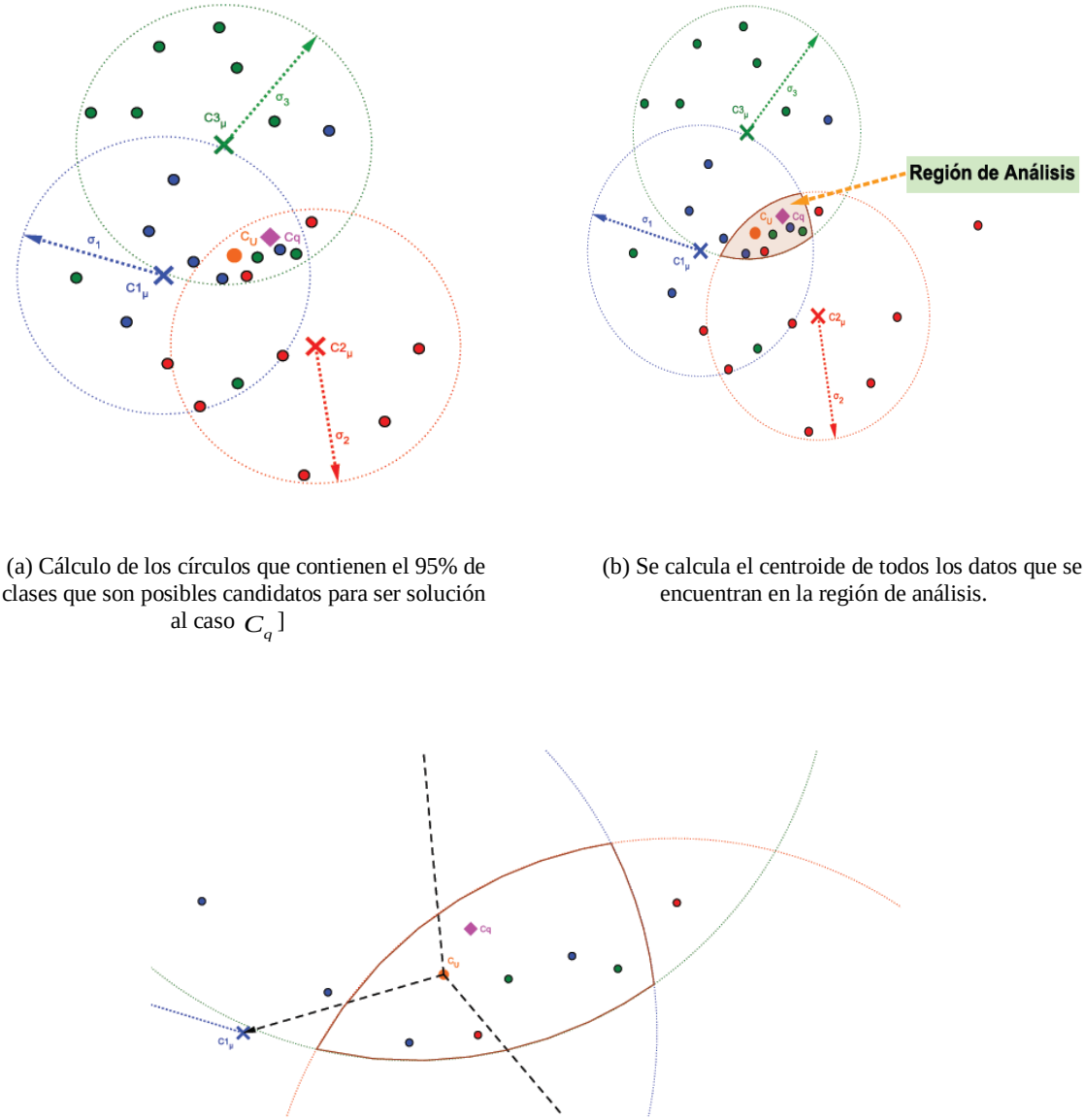
Para cada combinación Y, el razonamiento puede presentar los posibles escenarios:

- **Escenario 1:** Los descriptores seleccionados en la combinación Y del caso consulta resulten en la intersección de 2 o 3 clases.
- **Escenario 2**: Los descriptores seleccionados en la combinación Y del caso consulta resulten en una y solo una clase.
- **Escenario 3:** Los descriptores seleccionados en la combinación Y del caso consulta no pertenezcan a ninguna clase.

Para el escenario 1, el razonamiento usando el método del centroide se ilustra en la Figura 7. En la Figura 7(a) se encuentra Cq (caso consulta) que se encuentra en la intersección de las clases que pertenecen los descriptores del caso consulta. En la Figura 7(b) se muestra la región de análisis donde se encuentra C_u que es el centroide. En la Figura 7(c) se calcula las distancia euclidiana desde el centroide a cada una de las medias de las clases seleccionadas. La menor distancia es la clase que se selecciona. Para el escenario 2, la clase seleccionada es la clase en la que se encuentren los descriptores del caso consulta para la combinación Y. Para el escenario 3, se calcula la distancia euclidiana desde los descriptores del caso consulta hasta las medias de cada clase. La menor distancia es la clase seleccionada. Después de realizar las Y combinaciones, se realiza una votación simple. La clase que tenga más votos, es la clase que se selecciona como solución del caso consulta. En caso que el empate entre 2 o 3 clases persista, la selección se realiza de forma aleatoria.

4.2.4 Algoritmos de Retención. Una de las premisas del RBC es el almacenamiento de la experiencia revisada. Sin importar si los casos tienen solución correcta o incorrecta, este modelo aprovecha también ese conocimiento para mejorar sus razonamientos. En este modelo se implementaron dos algoritmos para retención de casos. El primero es un algoritmo de aprendizaje basado en los casos conocido como Instance Based Learning (IB). El segundo es un algoritmo de reducción de ruido conocido como Decremental Reduction Optimization Procedure (DROP).

4.2.4.1 *Instance Based-Learning (IB3).* Para este modelo de razonamiento basado en casos se implementó el algoritmo de aprendizaje por instancias que emplea un filtro selectivo para decidir cuáles casos serán guardados y luego ser usados en procesos de clasificación. Esto representa ventajas en la exactitud de la clasificación donde el número de casos correctamente clasificados



(a) Cálculo de los círculos que contienen el 95% de las clases que son posibles candidatos para ser solución al caso C_q]

(b) Se calcula el centroide de todos los datos que se encuentran en la región de análisis.

(c) La distancia más corta desde el punto del centroide a la media de cada clase es la solución del caso consulta.

Figura 7. Procedimiento para el análisis de las clases candidatas a ser solución del caso consulta.

sobre los intentos de clasificación se mantiene, además que emplea una prueba de significancia para decidir cuáles son los casos que pueden considerarse ruidosos.

Al estar basado en la prueba de significancia, el criterio decide un caso como aceptable si la exactitud de su clasificación local es más grande que la clasificación global, donde el intervalo de confianza es definido por el desempeño actual (número de clasificaciones correctas a nivel global)

y la exactitud que da su clase (número de casos clasificados correctamente de la clase analizada durante el proceso de clasificación). En la Ecuación 7 se puede observar cómo se define el intervalo de confianza para decidir si un caso es aceptable o no aceptable. En la Figura 8 se muestra el pseudocódigo del algoritmo IB3.

$$S = \frac{b + \frac{z^2}{2h} \pm z\sqrt{b\frac{1-b}{h} + \frac{z^2}{4h^2}}}{1 + \frac{z^2}{h}}$$

Ecuación 7

donde

- b = Número de Casos correctamente Clasificados / Número de Intentos.
- z =Son los límites del intervalo 0.9 para aceptación y 0.7 para no aceptación.
- h = Número de casos almacenados en la base de casos.

Para mayor amplitud del tema se puede dirigir (Aha, Kibler y Albert, 1991).

4.2.4.2 ***Decremental Reduction Optimization Procedure (DROP).*** Corresponde a una familia de algoritmos que son usados para remover casos ruidosos. Parten de la premisa: *si un caso P es ruidoso se debe a que tiene diferente clase respecto de la clase a la que fue asociado*. DROP define inicialmente un conjunto de casos denominado S, donde inicialmente S es igual al total de la base de casos.

```
function instance set SI = IB3(Training set TI)
    SI = ø
    for x = 1 to number of instances in TI
        for y = 1 to number of instances in SI
            sim[y] = similarity(x,y)
            if ∃y ∈ SI | acceptable(y)
                ymax = max(sim[y])
        else
            ymax = randomInstance(SI)
        endif
        if class(x) == class(y)
            classification = correct
        else
            classification = incorrect
            SI = SI ∪ x
        endif
        for y = 1 to number of instances in SI
            if (sim[y] ≥ sim[ymax])
                update classi_cation record of y
                if record classification of y is significantly poor
                    SI = remove(y from SI)
                endif
            endfor
        endfor
endfunction
```

Figura 8. Algoritmo de aprendizaje por casos

Para decidir si el caso P debe ser removido, se verifica el efecto en la exactitud global de remover el caso P de los asociados de P. Los asociados de un caso P se definen como los k-vecinos más cercanos al caso P. Esto es, cada caso en la base de casos, es definido como el caso P, y a su vez se deben calcular los k-asociados establecidos por el método de los k-vecinos más cercanos. Formalmente, sea P el caso a analizar, y sea $A_1, A_2, ...A_i, ...A_k$ los k-asociados de P. Se define el asociado i-ésimo de P como $P.A_i$. Si remover P causa que cada asociado $P.A_i$ use el k+1 vecino más cercano en lugar de P, entonces P podría ser un caso ruidoso. Se presentan entonces dos posibilidades: 1) Si P tiene la misma clase que $P.A_i$ y $P.A_i.N_{k+1}$ tiene diferente clase que $P.A_i$, esto debilita la clasificación y causa que $P.A_i$ no clasifique correctamente a sus vecinos más cercanos y 2) si P tiene diferente clase que $P.A_i$ y $P.A_i.N_{k+1}$ tiene diferente clase que $P.A_i$, remover

P causa que los casos clasificados incorrectamente sean ahora clasificados correctamente. Esencialmente, el algoritmo DROP prueba como removiendo el caso P se degrada la exactitud general cuando se realiza una validación cruzada dejando a P por fuera del proceso (Wilson y Martinez, 2000). En la Figura 9 se muestra el pseudocódigo de este método, mientras que en la Figura 10 se muestra un diagrama de procesos del método.

```
function instance set SI = DROP1(Training set TI)
        SI = TI
        for si = 1 to number of instances in SI
                find(k+1 nearest neighbors)
                add(si at each neighbors' associates list)
        endfor
        for si = 1 to number of instances in SI
                with = number of associates correctly classified
                        with si
        without = number of associates correctly classified
                                without si
if without  with ≥ with
                remove(si from SI)
                for a = 1 to number of associates of si
                        remove(si from nearest neighbors of a)
                        find(new nearest neighbor of a)
                        add(new neighbor to associates list of a)
                endfor
                for n = 1 to number of neighbors of si
                        remove(si from associates list of n)
                endfor
        endif
    endfor
endfunction
```

Figura 9. Algoritmo para la reducción de instancias ruidosas

4.3 Modelo RBC

El modelo de razonamiento basado en casos está compuesto por dos etapas generales, la primera de preprocesamiento, que a su vez se compone de los métodos de extracción y selección de características y de la ontología de los datos extraídos. La segunda etapa es el procesamiento, donde la información acondicionada y preparada, es sometida a prueba con el modelo de razonamiento basado en casos. Las fases de recuperación, reutilización, revisión y retención son implementadas de modo que finalmente el modelo entregue una clasificación sobre un caso consulta. En la Figura 11

se puede observar las etapas y sus respectivas fases para el modelo de razonamiento basado en casos.

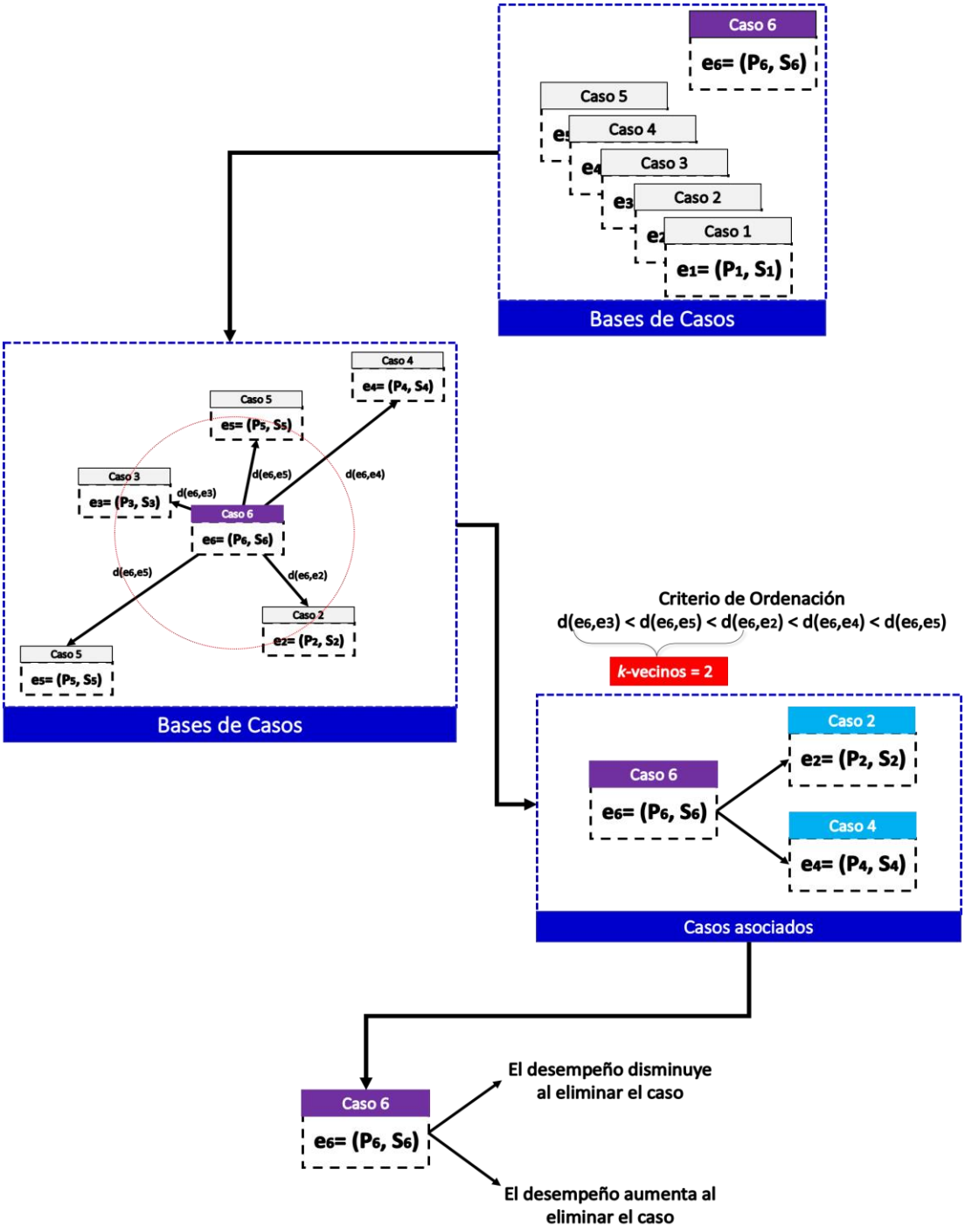


Figura 10. Procedimiento de implementación para algoritmo de reducción de instancias ruidosas

En resumen, para la etapa de preprocesamiento, en la extracción de características, se implementaron 17 métodos diferentes, algunos basados en el tiempo y otras basados en transformadas tiempo-frecuencia. Después de la extracción se utilizó Análisis de Componentes Principales para selección de características con el mayor aporte de la información. En la fase de procesamiento se implementaron 4 etapas. Para la fase de recuperación se utilizaron los algoritmos de distancia Euclidiana y la distancia de Minkowski. Para la fase de reutilización se implementaron los algoritmos de votación por regla de la mayoría y la regla basada en clase. Para la fase de revisión se utilizaron los algoritmos de reinstanciación o llamado adaptación por desviación estándar y el razonamiento por centroide. Para la fase de retención se utilizaron los algoritmos de aprendizajes por instancias y el algoritmo decremental de ruido. Para la etapa de validación y desempeño se utilizaron validación cruzada en grupos de a 10 utilizando la métrica de casos correctamente clasificados sobre el total de casos.

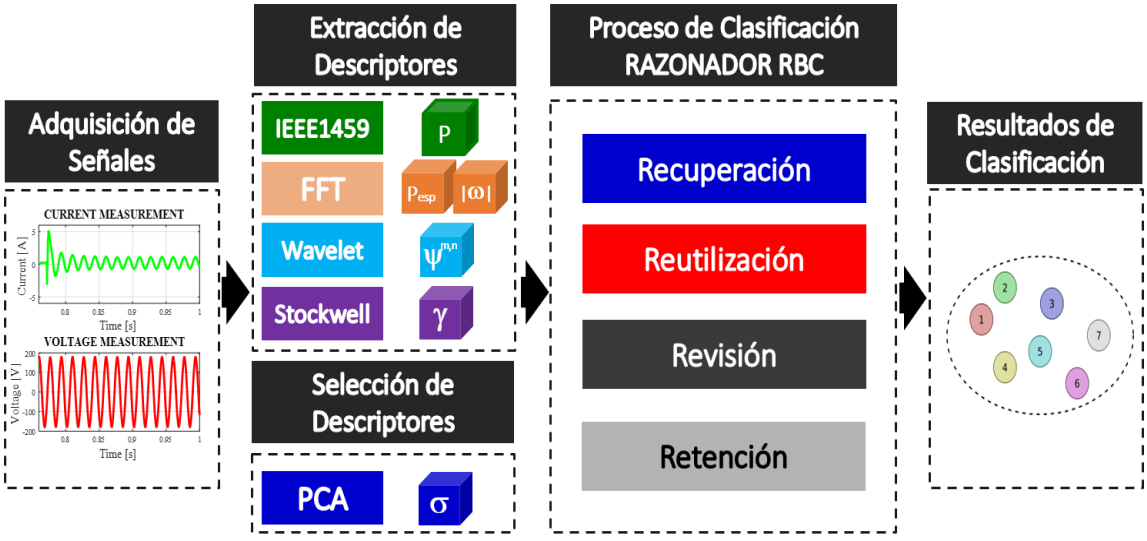


Figura 11. Mapa general de modelo de RBC para clasificación de eventos en monitorización no intrusiva de carga.

4.4 Combinaciones.

Para cuantificar el desempeño de la metodología del razonamiento basado en casos propuesta, dos algoritmos de cada fase fueron implementados y sometidos a prueba. Como son 4 fases se tendría un total de 16 combinaciones. Adicionalmente, en la fase de preprocesamiento, la selección de características fue realizada utilizando Análisis de Componentes Principales, por lo que se darían más posibilidades de modelos de razonadores basados en casos. Adicionalmente, como propuesta base, se estudió la influencia del agregar a la base de casos los casos que fueron aprendidos. Esto daría un total de modelos de razonamiento basado en casos de 24 con un total de 17 bases de casos. A continuación, la Figura 12 muestra de qué están compuestos los razonadores basado en casos.

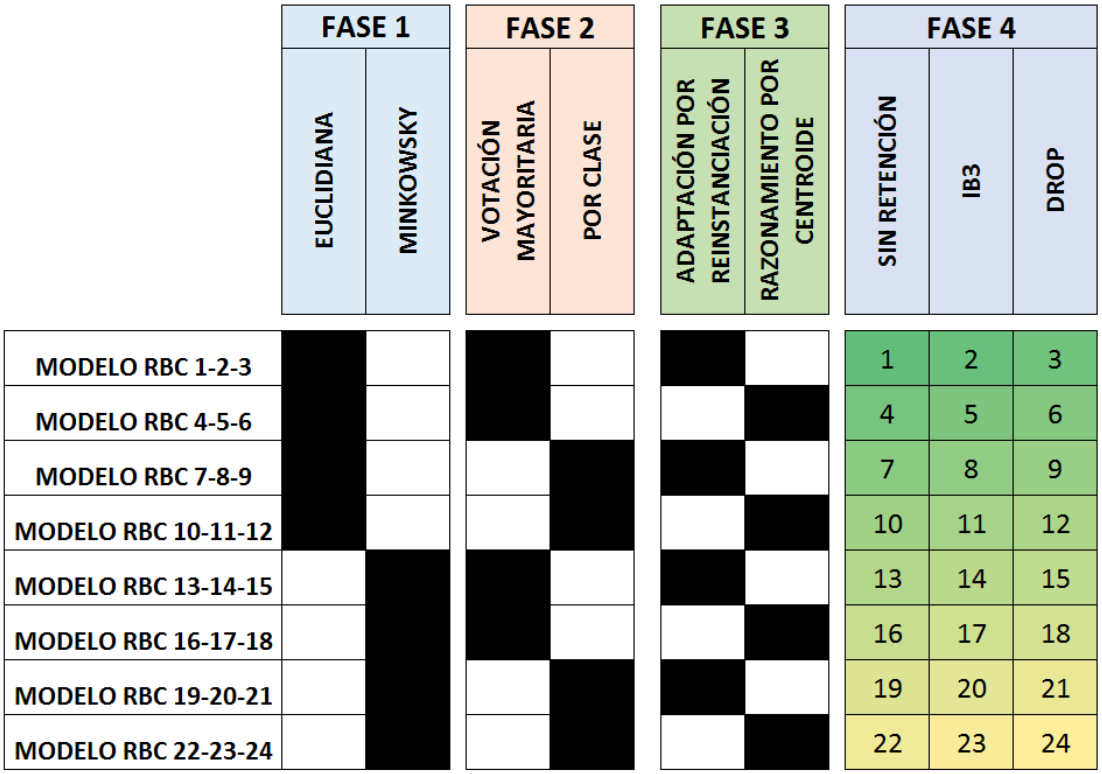


Figura 12. Combinaciones de los diferentes métodos para cada una de las fases del RBC. Se encuentran agrupados de a tres para mostrar que varios modelos tienen fases en común. Por ejemplo, los modelos 1-2-3 tienen en común el algoritmo de distancia Euclidiana, el algoritmo por regla de mayoría y la adaptación por reinstanciación, sólo difieren del modelo de retención usado.

5 Análisis del desempeño de los sistemas diseñados para la clasificación de eventos

En esta sección, se presentan los diferentes análisis de los resultados obtenidos en cada uno de los modelos de razonamiento basado en casos utilizando la información contenida en las bases de casos con los descriptores tomados del estándar IEEE1459, los calculados usando la Transformada Wavelet, la transformada S y la transformada Fraccionaria de Fourier, tanto para régimen estacionario como para régimen transitorio. Primero, se realiza el análisis de los resultados encontrados de cada modelo de RBC propuesto independientemente de la base de conocimiento considerado, esto es el desempeño promedio respecto a las 17 bases de conocimiento consideradas. Segundo, se realiza el análisis de los resultados encontrados de acuerdo al conocimiento extraído independientemente del modelo de RBC considerado, para este caso, de las diferentes formas de extracción de características consideradas. Tercero, se realiza el análisis de los resultados encontrados por el tipo de régimen considerado, para este caso, régimen estacionario y régimen transitorio. Cuarto, se realiza el análisis por la conexión de aparatos realizada, para este caso, conexión en modo individual y conexión de a 2 y 3 aparatos. Por último, se realiza un análisis de varianza mediante la técnica de Kruskal–Wallis para datos no-paramétricos con el fin de determinar una relación estadística del desempeño de los diferentes modelos de clasificación de RBC, usando métodos de extracción basado en FRFT, Wavelet y Transformada Stockwell, respecto al desempeño generado por un modelo basado en máquinas de soporte vectorial

5.1 Análisis de resultados obtenidos para cada modelo propuesto de RBC

A continuación, se presentan los resultados obtenidos para cada una de las bases de casos puestas a prueba mediante cada uno de los modelos de RBC propuestos. En la Figura 13 se pueden

observar los resultados obtenidos en porcentaje de clasificación para cada modelo de RBC respecto a cada base de casos considerada. Para cada porcentaje se presenta una escala de color que registra a 0% como el valor mínimo, representado por el color rojo, y 100% el valor máximo, representado por el color verde. Para la Figura 13 se realizan dos análisis: 1) ¿Cuál es el desempeño de los modelos independientemente de la base de casos tomada? y 2) ¿Cuál es la base de casos que tiene mejor desempeño independientemente del modelo de RBC considerado?

MODELOS RBC

BASES DE CONOCIMIENTO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
FVs	38	26	39	36	31	39	38	26	39	28	20	30	38	26	39	36	31	39	38	26	39	28	21	30
WSS	23	18	23	26	21	27	23	18	23	23	18	22	22	18	23	26	21	27	23	18	23	23	18	22
S-SS	77	41	77	70	38	72	76	40	77	70	36	69	75	41	76	69	36	69	76	40	76	70	36	69
FRFT SS	77	29	78	43	18	46	77	29	78	55	20	55	77	29	78	43	15	46	78	29	78	55	20	55
IEEE1459 SS	73	18	75	34	16	38	73	18	75	27	12	28	73	17	74	31	14	36	73	17	74	27	12	28
FVTime	97	54	97	95	53	95	97	54	97	70	32	69	97	54	97	95	53	95	97	54	97	70	32	69
ST1	89	47	89	68	46	69	88	47	89	79	38	79	88	47	89	68	46	69	88	47	89	79	38	79
ST2	53	45	54	47	43	47	53	46	55	38	38	38	52	47	54	46	44	45	52	48	55	38	38	38
ST3	68	30	72	32	23	32	67	30	72	45	22	44	67	28	72	28	12	28	67	28	72	45	22	44
ST4	43	14	46	31	17	34	43	14	46	31	13	33	43	14	46	32	19	35	43	13	46	31	13	33
ST5	5	7	5	5	8	5	5	8	5	5	5	5	5	8	5	5	6	5	5	8	5	5	5	5
ST6	36	30	40	33	34	35	36	30	40	29	22	31	36	30	40	33	33	35	36	30	40	29	22	31
ST7	29	25	32	35	30	37	29	25	32	33	25	33	29	25	32	35	30	37	29	25	32	33	25	33
WTS	78	61	79	70	56	70	78	61	79	68	52	68	78	54	79	71	43	73	78	54	80	67	52	68
S-TS	5	7	5	5	6	5	5	8	5	5	5	5	5	8	5	5	6	5	5	7	5	5	5	5
FRFT TS	43	15	45	30	16	31	43	15	45	27	11	27	43	15	45	31	16	33	43	15	45	27	11	27
IEEE1459 TS	47	29	48	26	20	29	47	29	48	18	16	18	47	29	48	26	20	29	47	29	48	18	16	18

Figura 13. Desempeños de los 24 modelos de RBC propuestos para todas las bases de casos consideradas y divididas por tipo de régimen.

De la Figura 14 se presenta el análisis de los modelos de RBC que tienen un desempeño promedio general de clasificación superior al 70%, indistintamente de la base de casos. Siguiendo la información de las filas inferiores, los promedios mostrados revelan que los modelos 1,3,7,9,13,15,19 y 21 presentan un desempeño superior al 70%. Referente a estos modelos, se observa que las fases que establecieron diferencias se encuentran en las fases de revisión y de retención. Para la fase de revisión son los modelos que utilizan el método de adaptación por desviación estándar. Para la

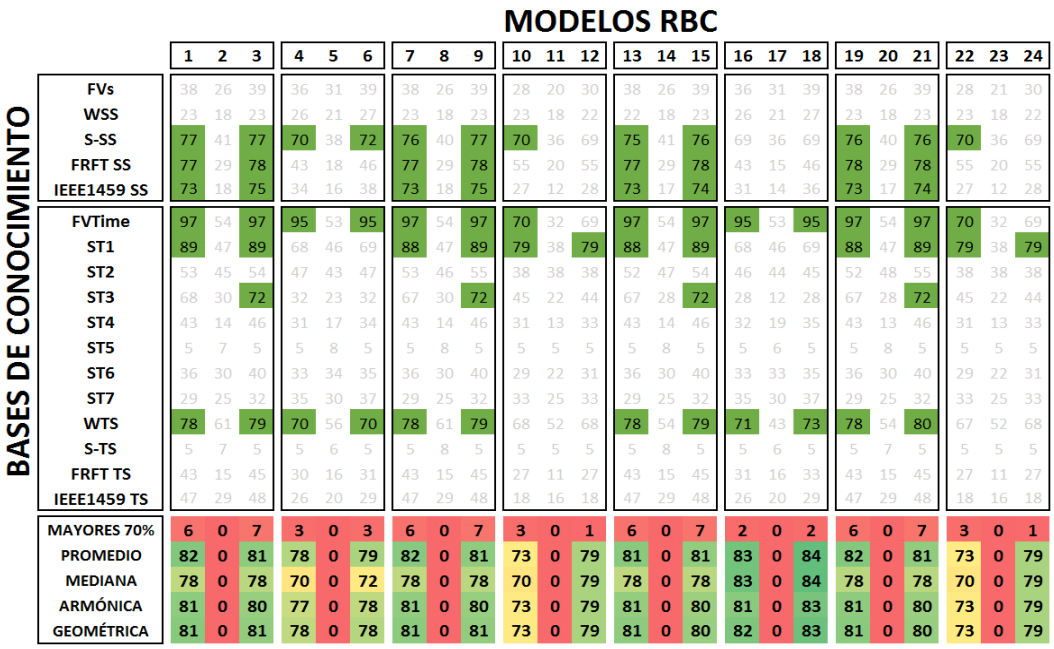


Figura 14. Resultados de desempeño para los 24 modelos de RBC propuestos que se encuentran arriba del 70%. Para estos valores seleccionados en verde, se calcularon el promedio, la mediana, la media armónica y geométrica.

fase de retención son los modelos considerados sin un método de retención y que utilizan el algoritmo de reducción de ruido. Sin embargo, en los modelos que consideran el algoritmo decremental de ruido siempre respecto a los modelos sin retención la cantidad de base de casos es siempre superior en todas los modelos de RBC que analizados. A partir de estas observaciones, se puede concluir que no importa la métrica de similaridad usada, ni el método de copiado de solución almacenada, podría representar una ventaja si se utiliza en la fase de revisión un modelo basado en el método de adaptación por desviación estándar y en la fase de retención un modelo basado en los algoritmos decrementales de ruido.

5.2 Análisis de resultados obtenidos por bases de conocimiento.

A continuación, se presentan los resultados obtenidos para cada una de bases de casos y el análisis de su desempeño a lo largo de los diferentes modelos de RBC. En la Figura 15 se pueden observar los resultados obtenidos en porcentaje de clasificación para cada base de casos. También, en la Figura 15 se presenta el análisis de las bases de casos que tienen en promedio un desempeño

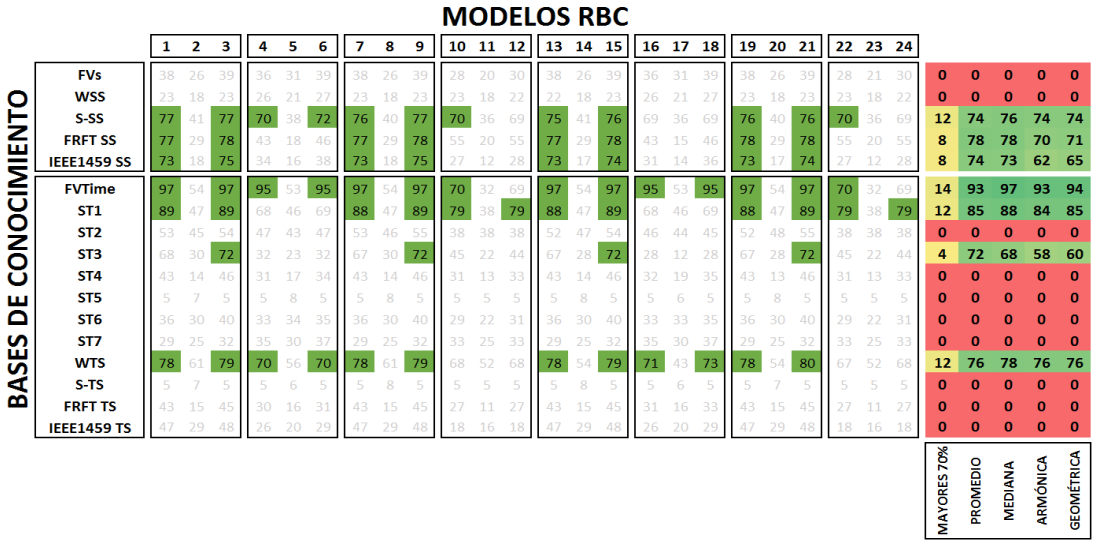


Figura 15. Resultados de desempeño para las bases de casos propuestas que se encuentran arriba del 70%. Para estos valores seleccionados en verde, se calcularon el promedio, la mediana, la media armónica y geométrica.

de clasificación superior al 70%, indistintamente del modelo de RBC considerado. Para cada porcentaje se presenta una escala de color que registra a 0% como el valor mínimo, representado por el color rojo, y 100% el valor máximo, representado por el color verde. Siguiendo la información de las columnas anexas, se puede observar que para las bases de casos usando S-Transform Estacionaria, ST1, FVTime y Wavelet Transitoria presentan exactitudes por encima del 70%. Sin embargo, la base de casos FVTime respecto del resto, presenta la mayor cantidad de modelos de RBC que tienen exactitudes superiores al 70%, con un total de 14 de 24, equivalente al 58%. La base de casos FVTime también presenta un desempeño promedio de 93% seguido de la base de casos ST1 con un 85%. Como consecuencia de estos resultados, se puede observar que la mejor base de casos para clasificar eventos de conexión y desconexión es la base de casos FVTime.

5.3 Análisis de resultados obtenidos por régimen.

A continuación, se presentan los resultados obtenidos para cada una de bases de casos y clasificados por el tipo de régimen del cual fueron extraídos. En la Figura 15 se pueden observar los resultados obtenidos en porcentaje de clasificación para cada base de casos. Para cada porcentaje

se presenta una escala de color que registra a 0% como el valor mínimo, representado por el color rojo, y 100% el valor máximo, representado por el color verde. Adicionalmente las bases de casos fueron ordenadas por régimen estacionario y régimen transitorio. Las bases de casos construidas con información de régimen estacionario corresponden a las primeras cinco. Las bases de casos construidas con información de régimen transitorio corresponden a las restantes. Siguiendo la información de las columnas anexas, se puede observar que para las bases de casos con información extraída del régimen estacionario y que presentan exactitudes por encima del 70% son las S-Transform, FRFT y IEEE1459. Tomando el promedio global a lo largo de los 24 modelos considerados, la base de casos FRFT presenta un valor promedio de 78% respecto a los otras bases de casos que cumplen la condición. Sin embargo, la base de casos S-Transform con un promedio global de 74%, el 50% de los modelos presentan exactitudes superiores al 70% en comparación con el 33% de modelos de la base de casos FRFT.

Siguiendo un análisis similar, las bases de casos que presentan los mejores desempeños globales son las bases de casos FVTime, ST1, ST3 y Wavelet. Tomando el promedio global a lo largo de los 24 modelos considerados, la base de casos FVTime presenta un 93% en comparación con el resto. Además, está base de casos tiene la mayor cantidad de modelos con un desempeño superior al 70%.

5.4 Análisis de resultados obtenidos por conexión de aparatos.

A continuación, se presentan los resultados obtenidos para cada una de bases de casos y divididos en conexión individual y conexión de a 2 y 3 aparatos. por el tipo de régimen del cual fueron extraídos. En la Figura 16 se pueden observar los resultados obtenidos en porcentaje de clasificación para cada base de casos en modo conexión individual. Para cada porcentaje se presenta una

escala de color que registra a 0% como el valor mínimo, representado por el color rojo, y 100% el valor máximo, representado por el color verde.

Si se analiza por desempeños por el modelo de RBC considerado, se puede observar que el modelo 21 presenta la mayor coincidencia independientemente de la fuente de conocimiento extraído y con un desempeño promedio de 81%, sin embargo, existen varios modelos que presentan el mayor desempeño global promedio de 85% independientemente de la base de casos considerada y son los modelos 1,3,6,7,12,15 y 16. Cabe notar que también hay cuatro modelos que no tuvieron coincidencia con ninguna base de casos y son los modelos 9,18,20 y 23. Si se analiza por desempeños de las bases de casos, se puede observar que la base de conocimiento IEEE1459 para estado estacionario, con un desempeño promedio global de 90% presenta un total de 18 de los 24 modelos con desempeño local superior al 70% y además es la más representativa. Si se analiza por tipo de régimen, la base de casos que presenta el mejor desempeño en estado estacionario es la llamada IEEE1459 que utiliza características tomadas de la definición de variables de calidad de la energía

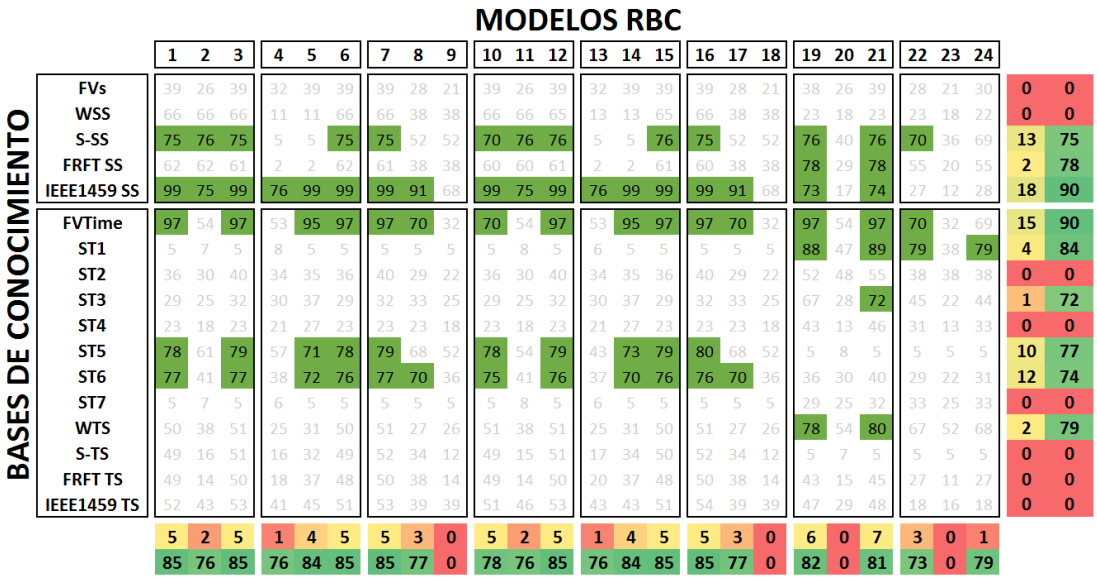


Figura 16. Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, S-Transform Estacionario y Transitorio.

en el estándar IEEE 1459, mientras que la que presenta el mejor desempeño en estado transitorio

es la llamada FVTime que utiliza parte de definiciones de variables de calidad de energía del estándar IEEE1459 y otras variables tomadas

5.5 Sumario de desempeños de los diferentes clasificadores

A continuación, en la Tabla 13 y en la Tabla 14 se pueden observar el resumen de todos los desempeños obtenidos por los diferentes modelos de razonamiento basado en casos para cada una de las bases de casos que fueron analizadas.

Tabla 13. *Exactitudes obtenidas para los modelos de RBC del 1 al 12.*

	MODELOS DE RBC											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
FVs	38,39	26,17	39,08	36,17	31,36	39,03	38,44	26,07	38,87	28,03	20,31	29,51
FVTime	96,59	53,85	96,59	95,42	52,91	95,42	96,54	53,85	96,70	69,94	32,01	69,50
ST1	88,54	46,53	88,81	68,40	46,44	69,18	88,26	46,53	88,72	79,04	37,72	79,41
ST2	53,49	45,35	54,19	47,29	43,49	47,05	52,56	45,74	54,65	38,14	38,22	37,60
ST3	67,51	29,87	71,51	31,80	23,48	32,38	67,40	29,63	71,65	45,47	21,66	44,06
ST4	43,45	13,64	45,61	30,58	17,08	34,34	42,60	13,53	45,79	31,04	13,13	32,71
ST5	5,24	6,77	5,24	5,24	7,83	5,24	5,24	8,30	5,24	5,24	5,24	5,24
ST6	36,01	29,56	39,66	32,68	33,63	34,59	35,96	29,67	39,61	29,09	22,16	30,51
ST7	29,30	25,01	31,78	35,06	29,93	36,54	29,40	24,91	31,78	32,95	24,64	32,68
Wavelet SS	22,58	17,56	22,79	26,12	21,20	26,76	22,69	17,56	22,85	23,27	17,82	21,68
Wavelet TS	78,11	61,18	78,53	70,23	56,48	70,49	77,95	60,97	79,27	67,69	51,56	67,79
S Transform SS	76,78	40,67	77,10	70,02	37,65	71,55	76,31	39,61	77,21	69,96	35,96	69,22
S-Transform TS	5,24	6,93	5,24	5,24	6,29	5,24	5,24	8,25	5,24	5,24	5,24	5,24
FRFT SS	77,48	29,44	77,93	42,94	17,62	45,67	77,42	29,29	78,00	54,63	19,97	55,46
FRFT TS	42,84	15,23	45,11	29,78	15,67	30,62	42,73	15,17	45,19	27,01	10,71	26,82
IEEE1459 SS	73,21	18,13	74,73	34,47	16,26	37,59	73,35	18,18	74,59	27,40	12,25	28,19
IEEE1459 TS	46,62	29,11	47,53	26,18	20,16	29,13	46,68	29,24	47,59	17,98	16,41	18,18

Tabla 14. *Exactitudes obtenidas para los modelos de RBC del 13 al 24.*

	13	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
FVs	38,34	25,96	39,24	36,12	31,36	39,03	38,50	26,02	39,13	28,03	20,83	29,51
FVTime	96,59	53,85	96,65	95,42	52,91	95,42	96,65	53,85	96,65	70,00	32,01	69,50
ST1	88,31	46,53	88,72	68,45	46,44	69,18	88,45	46,53	88,72	79,09	37,72	79,41
ST2	52,40	47,05	53,64	45,81	44,26	44,81	52,25	47,60	55,04	38,14	38,22	37,60
ST3	67,35	28,33	71,52	27,70	11,59	27,56	67,08	28,47	71,52	45,08	21,66	44,06
ST4	43,39	13,60	45,77	32,06	18,52	35,17	42,95	13,21	45,99	31,03	13,08	32,71
ST5	5,24	8,35	5,24	5,24	6,29	5,24	5,24	8,20	5,24	5,24	5,24	5,24
ST6	35,59	29,61	39,66	32,73	33,37	34,59	36,07	29,67	39,77	29,03	22,16	30,51
ST7	29,35	24,96	31,89	35,17	29,88	36,54	29,35	24,91	31,84	32,95	24,64	32,68
Wavelet SS	22,48	17,77	22,74	26,02	21,20	26,76	22,58	17,93	23,00	23,11	17,82	21,68
Wavelet TS	77,53	54,26	79,16	71,07	42,52	72,55	78,37	53,99	79,53	67,48	51,56	67,79
S Transform SS	74,93	41,04	76,04	68,69	36,49	69,33	75,83	40,40	75,94	70,23	35,96	69,22
S-Transform TS	5,24	8,41	5,24	5,24	6,24	5,24	5,24	7,19	5,24	5,24	5,24	5,24
FRFT SS	77,47	29,14	77,99	43,07	15,35	45,66	77,59	29,14	78,08	55,02	19,97	55,46
FRFT TS	42,73	15,11	45,12	31,08	16,19	32,66	42,94	15,27	45,08	27,03	10,71	26,82
IEEE1459 SS	73,36	16,71	74,38	30,77	13,69	35,63	73,38	16,67	74,34	27,06	12,25	28,19
IEEE1459 TS	46,78	29,11	47,59	25,57	20,20	28,81	46,84	29,06	47,58	17,91	16,41	18,18

5.6 Comparación de los sistemas de identificación

Para determinar en qué rango se encuentra el sistema de clasificación usando razonamiento basado en casos, se utilizó el clasificador propuesto en (Cala, 2018) basados en Máquinas de Soporte Vectorial, que fueron previamente entrenados y de los cuáles se obtuvieron sus medidas de desempeño. Se presentan los análisis obtenidos de las bases de casos Wavelet Estacionaria y Transitoria, FRFT Estacionaria y Transitoria y S-Transform Estacionaria y Transitoria. En la Tabla 27 y Tabla 31 se muestran las medidas de desempeño de cada una de las bases de casos mencionadas y las medidas de desempeño de los respectivos modelos de RBC. Las filas de cada tabla corresponden a las corridas que se realizaron, en este caso son 10 corridas. Las columnas muestran los desempeños de cada uno de los modelos de RBC propuesto y el modelo de MSV que se considera como referente.

Realizando un análisis superficial, se puede observar que para las bases de casos que tienen la información en régimen estacionario, la medida de desempeño del modelo de MSV respecto a los otros modelos de RBC, tiene una diferencia positiva de 25%, lo que indica que, para bases de conocimiento en régimen estacionario, el modelo de MSV tiene mejor desempeño. Sin embargo, para las bases de casos que tienen la información en régimen transitorio, la medida de desempeño del modelo de MSV respecto a los otros modelos de RBC, presentan valores cercanos incluso inferiores a algunos modelos de RBC propuestos. Esto indica que los modelos de RBC propuestos son comparables con el modelo MSV que se usa como referente.

5.7 Análisis de Varianza usando el Método Kruskal-Wallis para Datos No paramétricos

Para determinar si existen diferencias estadísticamente significativas entre el modelo de MSV referente y los diferentes modelos de RBC propuestos, se utiliza el método no paramétrico de Kruskal-Wallis (McKight y Najab, 2010), dado que después de realizar una prueba de normalidad, usando el método de Kolgomorov-Smirnov para pruebas en datos con media y varianza desconocida (Lilliefors, 1967), a los datos presentados en la Tabla 27 y Tabla 31, estos no presentan distribución normal, por lo que no es posible aplicar una ANOVA Simple. La prueba de comparar las varianzas de los datos en cada una de las bases de casos con el modelo referente basado en MSV. Para lograr este propósito, se utiliza el Toolbox de MATLAB para calcular los estadísticos de cada uno de desempeño de clasificación y se muestran las gráficas que permiten tomar la decisión sobre cuál combinación de RBC puede ser un clasificador prometedor en futuras tareas. En la Figura 17 se muestra el resumen de los modelos de RBC para cada una de los métodos de extracción de conocimiento considerados y que fueron comparados con el método de extracción basado en MSV. En la Figura 34, Figura 35 y Figura 36, ubicadas en el Apéndice E, se muestran los resultados obtenidos después de realizar la prueba de significancia para las bases de casos consideradas de los 24 modelos de RBC propuesto con el modelo MSV referente.

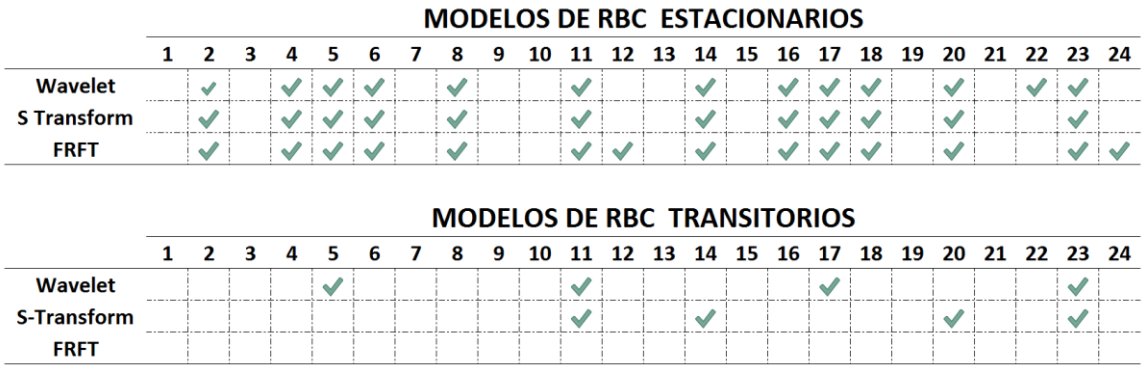


Figura 17. Resumen donde se muestran los modelos de RBC que presentan diferencias estadísticamente significativas respecto al modelo de MSV considerado como referente. Se presentan para las bases de casos Wavelet, S-Transform y FRFT tanto en régimen estacionario como en régimen transitorio.

6 Conclusiones y Trabajo Futuros

En este capítulo final se resumen los resultados obtenidos en este trabajo de investigación. Se presentan también algunas recomendaciones y trabajos futuros y por último se muestran los trabajos divulgados con los resultados de este trabajo de investigación.

6.1 Conclusiones

En este trabajo de investigación se diseñaron 24 modelos de RBC para la clasificación de aparatos eléctricos residenciales en condiciones de laboratorio y en conexiones individuales y combinaciones de 2 y 3 aparatos respectivamente. Se comparó con el modelo basado en máquinas de soporte vectorial teniendo diferentes fuentes de conocimiento. Los modelos están compuestos por dos instancias iniciales, como son el *preprocesamiento,* en donde se usan los algoritmos de extracción de características, en función de la base de conocimiento y un algoritmo de selección de características basado en análisis de componentes principales, y el *procesamiento* que considera los algoritmos propuestos para los 24 modelos de razonamiento basado en casos.

A nivel general, los resultados de clasificación muestran en promedio, considerando todos los modelos de RBC, presentan una exactitud de 39,66% con una desviación estándar 20,38. Este valor de desviación estándar es causado por las bajas tasas de exactitud en las bases de datos transitorias ST1-ST4 y la ST7. En promedio, las 5 primeras bases de casos utilizadas que presentan las mejores exactitudes son las FVs, IEEE1459, ST5, ST6, S-Transform con valores de 75,78%, 68,96%, 67,33%, 61,50% y 50,4%.

A nivel particular, en cuanto a los resultados de exactitud, clasificados por el tipo de régimen, las bases de conocimiento en régimen estacionario presentan un porcentaje de 37,49% con una desviación estándar de 14,91. En régimen transitorio presentan un porcentaje de 40,57% con una

desviación estándar de 22,81. Así entonces, las bases de conocimiento con información de régimen transitorio tienen mejores exactitudes pero información muy dispersa respecto a la información en régimen estacionario. Esto ocurre, debido a las bases de casos ST1 y ST7 que tienen bajas exactitudes y altas dispersiones.

Ahora, con respecto de las exactitudes clasificadas por el tipo de conexión, los aparatos individuales a lo largo de todas las bases de conocimiento presentan un porcentaje de exactitud de 39,51% con una desviación estándar de 9,40. Para aparatos combinados de 2 y 3 presentan un porcentaje de exactitud de 40,4% con una desviación estándar de 23,07. Se puede observar que la base de conocimiento para aparatos individuales es menos dispersa que la base de conocimiento de los aparatos combinados, a pesar que las exactitudes son muy parecidas. La dispersión puede ser explicada debido a la baja cantidad de ejemplos que se tienen de las diferentes combinaciones que fueron consideradas en contraste con la alta información obtenida de señales individuales.

6.2 Recomendaciones y Trabajos Futuros

El problema de investigación abordado en el presente trabajo, abre caminos en una serie de actividades en monitorización no intrusiva de carga para ambientes residenciales, las cuales son:

A nivel de grupo de investigación.

- Implementación de sistemas de monitorización no intrusiva de carga de bajo costo utilizando tarjetas de desarrollo.
- Consideración de bases de datos públicas con información relacionada con patrones de consumo y huellas digitales de aparatos que ya han sido caracterizados
- Diseño y prototipado de un sistema de monitorización no intrusiva de carga con procesamiento en la nube.

A nivel de trabajo de investigación.

- Mejoramiento en la indexación de las bases de casos e inclusión de índices lingüísticos.
- Estudio de las bases de datos dinámicas para implementar la fase de retención.
- Considerar otro tipo de algoritmos adaptativos en la fase de reuso y revisión para mejorar los procesos de razonamiento en situaciones de aparatos combinados de 2 y 3 cargas eléctricas.
- Considerar otro tipo de descriptores en situaciones donde se conecten más de 2 aparatos.

6.3 Divulgación de Trabajos

Algunos de los resultados de este trabajo de investigación se divulgaron en el Simposio Internacional de Tratamiento de señales, imágenes y visión artificial STSIVA 2016 en el mes de agosto-septiembre en la ciudad de Bucaramanga en el que se presentó el artículo de investigación titulado **Reasoner Design based on HYPO for classification of lighting loads**. Este fue publicado en la base de datos del IEEE.

Anexos

Apéndice A. Clases consideradas para cada una de las bases de casos

Tabla 15. *Clases definidas para los aparatos sometidos a prueba para la construcción de la base de casos FVs en régimen estacionario*

CLASE	ETIQUETA	APARATOS	POTENCIA	ESTADOS	EVENTOS
1	SS1A1	Licuadora	450	S1: Velocidad 1	100
2	SS1A1	Licuadora	450	S2: Velocidad 2	95
3	SS1A2	Bombillo Incandescente	75	S1: Encendido - Apagado	100
4	SS1A3	Bombillo Fluorescente	20	S1: Encendido - Apagado	100
5	SS1A4	Bombillo Fluorescente	9	S1: Encendido - Apagado	100
6	SS1A6	Computador de Escritorio	216	S1: Secuencia 1	99
7	SS1A7	Ventilador	48	S1: Velocidad 1	100
8	SS1A8	Nevera	1150	S1: Apertura Puerta	100
9	SS1A8	Nevera	1150	S2: Ciclo 1	100
10	SS1A9	Secador de Cabello	875	S1: Temperatura 1	100
11	SS1A9	Secador de Cabello	875	S2: Temperatura 2	100
12	SS1A10	Bombillo Halógeno	50	S1: Encendido - Apagado	99
13	SS1A11	Bombillo Halógeno	70	S1: Encendido - Apagado	98
14	SS1A12	Plancha	1200	S1: Conexión - Desconexión	100
15	SS1A13	Bombillo LED	7	S1: Encendido - Apagado	98
16	SS1A14	Cargador de Portátil	120	S1: Conexión - Desconexión	100
17	SS1A19	Sanduchera	750	S1: Conexión - Desconexión	101
18	SS1A21	Televisor TRC	90	S1: Conexión - Desconexión	100

Tabla 16. *Clases definidas para los aparatos sometidos a prueba para la construcción de la base de casos FVTime en régimen transitorio. TS2 significa que el transitorio es apagado-encendido.*

CLASE	ETIQUETA	APARATOS	POTENCIA	ESTADOS	EVENTOS
1	TS2A1	Licuadora	450	S1: Velocidad 1	100
2	TS2A1	Licuadora	450	S2: Velocidad 2	95
3	TS2A2	Bombillo Incandescente	75	S1: Encendido - Apagado	100
4	TS2A3	Bombillo Fluorescente	20	S1: Encendido - Apagado	100
5	TS2A4	Bombillo Fluorescente	9	S1: Encendido - Apagado	100
6	TS2A5	Cargador de Celular	20	S1: Conexión - Desconexión	100
7	TS2A6	Computador de Escritorio	216	S2: Secuencia 2	100
8	TS2A7	Ventilador	48	S1: Velocidad 1	100
9	TS2A8	Nevera	1150	S1: Apertura Puerta	100
10	TS2A8	Nevera	1150	S2: Ciclo 1	100
11	TS2A9	Secador de Cabello	875	S1: Temperatura 1	100
12	TS2A9	Secador de Cabello	875	S2: Temperatura 2	100
13	TS2A10	Bombillo Halógeno	50	S1: Encendido - Apagado	99
14	TS2A11	Bombillo Halógeno	70	S1: Encendido - Apagado	98
15	TS2A12	Plancha	1200	S1: Conexión - Desconexión	100
16	TS2A13	Bombillo LED	7	S1: Encendido - Apagado	98
17	TS2A14	Cargador de Portátil	120	S1: Conexión - Desconexión	100
18	TS2A19	Sanduchera	750	S1: Conexión - Desconexión	101
19	TS2A21	Televisor TRC	90	S1: Conexión - Desconexión	100

Tabla 17. *Clases definidas para los aparatos sometidos a prueba para la construcción de la base de casos Wavelet, Fractional y S en régimen estacionario. SS1 significa que es en estado estacionario de los aparatos individual.*

CLASE	ETIQUETA	APARATOS	POTENCIA	ESTADOS	EVENTOS
1	SS1A1	Licuadora	450	S1: Velocidad 1	100
2	SS1A1	Licuadora	450	S2: Velocidad 2	100
3	SS1A2	Bombillo Incandescente	75	S1: Encendido - Apagado	97
4	SS1A3	Bombillo Fluorescente	20	S1: Encendido - Apagado	101
5	SS1A4	Bombillo Fluorescente	9	S1: Encendido - Apagado	108
6	SS1A5	Cargador de Celular	20	S1: Conexión - Desconexión	100
7	SS1A6	Computador de Escritorio	216	S1: Secuencia 1	100
8	SS1A6	Computador de Escritorio	216	S2: Secuencia 2	400
9	SS1A6	Computador de Escritorio	216	S3: Secuencia 3	100
10	SS1A7	Ventilador	48	S1: Velocidad 1	104
11	SS1A8	Nevera	1150	S1: Apertura Puerta	100
12	SS1A8	Nevera	1150	S2: Ciclo 1	100
13	SS1A9	Secador de Cabello	875	S1: Temperatura 1	100
14	SS1A9	Secador de Cabello	875	S2: Temperatura 2	100
15	SS1A10	Bombillo Halógeno	50	S1: Encendido - Apagado	99
16	SS1A11	Bombillo Halógeno	70	S1: Encendido - Apagado	104
17	SS1A12	Plancha	1200	S1: Conexión - Desconexión	100
18	SS1A13	Bombillo LED	7	S1: Encendido - Apagado	110
19	SS1A19	Sanduchera	750	S1: Conexión - Desconexión	100

Tabla 18. *Clases definidas para las combinaciones de 2 aparatos para la construcción de la base de casos Wavelet, Fractional y S en régimen estacionario. SS1 significa que es en estado estacionario.*

CLASE	ETIQUETA	ESTADOS	APARATOS	POTENCIA	ESTADOS	EVENTOS
20	SS1A1A4	S2	Licuadora	450	S2: Velocidad 2	100
		S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	
21	SS1A1A5	S2	Licuadora	450	S2: Velocidad 2	100
		S1	Cargador de Celular	20	S1: Conexión - Desconexión	
22	SS1A1A7	S2	Licuadora	450	S2: Velocidad 2	200
		S1	Ventilador	48	S1: Velocidad 1	
23	SS1A1A8	S2	Licuadora	450	S2: Velocidad 2	200
		S1	Nevera	1150	S1: Apertura Puerta	
24	SS1A1A11	S2	Licuadora	450	S2: Velocidad 2	200
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
25	SS1A1A12	S2	Licuadora	450	S2: Velocidad 2	200
		S1	Plancha	1200	S1: Conexión - Desconexión	
26	SS1A1A21	S2	Licuadora	450	S2: Velocidad 2	200
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
27	SS1A4A5	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	100
		S1	Cargador de Celular	20	S1: Conexión - Desconexión	
28	SS1A4A7	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	200
		S1	Ventilador	48	S1: Velocidad 1	
29	SS1A4A8	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	100
		S1	Nevera	1150	S1: Apertura Puerta	
30	SS1A4A11	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	200
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
31	SS1A4A12	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	200
		S1	Plancha	1200	S1: Conexión - Desconexión	
32	SS1A4A21	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	200
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
33	SS1A5A7	S1	Cargador de Celular	20	S1: Conexión - Desconexión	100
		S1	Ventilador	48	S1: Velocidad 1	
34	SS1A5A8	S1	Cargador de Celular	20	S1: Conexión - Desconexión	100
		S1	Nevera	1150	S1: Apertura Puerta	
35	SS1A5A11	S1	Cargador de Celular	20	S1: Conexión - Desconexión	200

CLASE	ETIQUETA	ESTADOS	APARATOS	POTENCIA	ESTADOS	EVENTOS
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
36	SS1A5A12	S1	Cargador de Celular	20	S1: Conexión - Desconexión	100
		S1	Plancha	1200	S1: Conexión - Desconexión	
37	SS1A5A21	S1	Cargador de Celular	20	S1: Conexión - Desconexión	100
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
38	SS1A7A8	S1	Ventilador	48	S1: Velocidad 1	200
		S1	Nevera	1150	S1: Apertura Puerta	
39	SS1A7A11	S1	Ventilador	48	S1: Velocidad 1	200
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
40	SS1A7A12	S1	Ventilador	48	S1: Velocidad 1	200
		S1	Plancha	1200	S1: Conexión - Desconexión	
41	SS1A7A21	S1	Ventilador	48	S1: Velocidad 1	200
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
42	SS1A8A11	S1	Nevera	1150	S1: Apertura Puerta	200
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
43	SS1A8A12	S1	Nevera	1150	S1: Apertura Puerta	200
		S1	Plancha	1200	S1: Conexión - Desconexión	
44	SS1A8A21	S1	Nevera	1150	S1: Apertura Puerta	200
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
45	SS1A11A12	S1	Bombillo Halógeno	70	S1: Encendido - Apagado	200
		S1	Plancha	1200	S1: Conexión - Desconexión	
46	SS1A11A21	S1	Bombillo Halógeno	70	S1: Encendido - Apagado	200
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
47	SS1A12A21	S1	Plancha	1200	S1: Conexión - Desconexión	200
		S1	Televisor TRC	90	S1: Conexión - Desconexión	

Tabla 19. *Clases definidas para las combinaciones de 3 aparatos para la construcción de la base de Casos Wavelet, Fractional y S en régimen estacionario. SS1 significa que es en estado estacionario.*

CLASE	ETIQUETA	ESTADOS	APARATOS	POTENCIA	ESTADOS	EVENTOS
		S1	Licuadora	450	S1: Velocidad 1	
48	SS1A1A2A4	S1	Bombillo Incandescente	75	S1: Encendido - Apagado	375
		S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	
		S1	Licuadora	450	S1: Velocidad 1	
49	SS1A1A3A12	S1	Bombillo Fluorescente	20	S1: Encendido - Apagado	375
		S1	Plancha	1200	S1: Conexión - Desconexión	
		S1	Licuadora	450	S1: Velocidad 1	
50	SS1A1A3A22	S1	Bombillo Fluorescente	20	S1: Encendido - Apagado	375
		S1	Cargador de Portátil	65	S1: Conexión - Desconexión	
		S1	Licuadora	450	S1: Velocidad 1	
51	SS1A1A19A21	S1	Sanduchera	750	S1: Conexión - Desconexión	375
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
		S1	Bombillo Incandescente	75	S1: Encendido - Apagado	
52	SS1A2A3A22	S1	Bombillo Fluorescente	20	S1: Encendido - Apagado	375
		S1	Cargador de Portátil	65	S1: Conexión - Desconexión	
		S1	Bombillo Incandescente	75	S1: Encendido - Apagado	
53	SS1A2A4A11	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	375
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
		S1	Bombillo Incandescente	75	S1: Encendido - Apagado	
54	SS1A2A4A19	S1	Bombillo Fluorescente	9	S1: Encendido - Apagado	375
		S1	Sanduchera	750	S1: Conexión - Desconexión	
		S1	Bombillo Incandescente	75	S1: Encendido - Apagado	
55	SS1A2A19A21	S1	Sanduchera	750	S1: Conexión - Desconexión	375
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
		S1	Bombillo Fluorescente	20	S1: Encendido - Apagado	
56	SS1A3A11A13	S1	Bombillo Halógeno	70	S1: Encendido - Apagado	375
		S1	Bombillo LED	7	S1: Encendido - Apagado	
		S1	Bombillo Fluorescente	20	S1: Encendido - Apagado	
57	SS1A3A12A13	S1	Plancha	1200	S1: Conexión - Desconexión	375
		S1	Bombillo LED	7	S1: Encendido - Apagado	
58	SS1A7A11A13	S1	Ventilador	48	S1: Velocidad 1	375

CLASE	ETIQUETA	ESTADOS	APARATOS	POTENCIA	ESTADOS	EVENTOS
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
		S1	Bombillo LED	7	S1: Encendido - Apagado	
		S1	Ventilador	48	S1: Velocidad 1	
59	SS1A7A11A19	S1	Bombillo Halógeno	70	S1: Encendido - Apagado	375
		S1	Sanduchera	750	S1: Conexión - Desconexión	
		S1	Ventilador	48	S1: Velocidad 1	
60	SS1A7A11A21	S1	Bombillo Halógeno	70	S1: Encendido - Apagado	375
		S1	Televisor TRC	90	S1: Conexión - Desconexión	
		S1	Nevera	1150	S1: Apertura Puerta	
61	SS1A8A11A22	S1	Bombillo Halógeno	70	S1: Encendido - Apagado	375
		S1	Cargador de Portátil	65	S1: Conexión - Desconexión	
		S1	Bombillo Halógeno	70	S1: Encendido - Apagado	
62	SS1A11A13A21	S1	Bombillo LED	7	S1: Encendido - Apagado	375
		S1	Televisor TRC	90	S1: Conexión - Desconexión	

Tabla 20. *Clases de aparatos tenidos en cuenta para transitorios*

CLASE	ETIQUETA	APARATO	GRUPO APARATOS	ESTADO	EVENTOS POR GRUPO	EVENTOS POR CLASE
			A1		100	
			A1A2A4		20	
1	TS1A1-S1	Licuadora	A1A3A12	S1: Velocidad 1	20	180
			A1A3A22		20	
			A1A19A21		20	
			A1		100	
			A1A7		10	
			A1A8		10	
2	TS1A1-S2	Licuadora	A1A11	S2: Velocidad 2	10	150
			A1A12		10	
			A1A21		10	
3	TS1A10-S1	Bombillo Halógeno	A10	S1: Encendido - Apagado	99	99
			A11		104	
			A1A11		10	
			A4A11		10	
			A5A11		10	
			A7A11		10	
			A8A11		10	
			A11A12		10	
4	TS1A11-S1	Bombillo Halógeno	A11A21	S1: Encendido - Apagado	10	314
			A2A4A11		20	
			A3A11A13		20	
			A7A11A13		20	
			A7A11A19		20	
			A7A11A21		20	
			A8A11A22		20	
			A11A13A21		20	
			A12		100	
			A1A12		10	
			A4A12		10	
			A7A12		10	
5	TS1A12-S1	Plancha	A8A12	S1: Conexión - Desconexión	10	200
			A11A12		10	
			A12A21		10	
			A1A3A12		20	
			A3A12A13		20	
			A13		110	
			A3A11A13		20	
6	TS1A13-S1	Bombillo LED	A3A12A13	S1: Encendido - Apagado	20	190
			A7A11A13		20	
			A11A13A21		20	
			A19		100	
			A1A19A21		20	
7	TS1A19-S1	Sanduchera	A2A4A19	S1: Conexión - Desconexión	20	180
			A2A19A21		20	
			A7A11A19		20	
			A2		97	
8	TS1A2-S1	Bombillo Incandescente	A1A2A4	S1: Encendido - Apagado	20	197
			A2A3A22		20	

			A2A4A11		20	
			A2A4A19		20	
			A2A19A21		20	
9	TS1A21-S1	Televisor TRC	A1A21	S1: Conexión - Desconexión	10	140
			A4A21		10	
			A7A21		10	
			A8A21		10	
			A11A21		10	
			A12A21		10	
			A1A19A21		20	
			A2A19A21		20	
			A7A11A21		20	
			A11A13A21		20	
10	TS1A22-S1	Cargador de Portátil	A1A3A22	S1: Conexión - Desconexión	20	60
			A2A3A22		20	
			A8A11A22		20	
11	TS1A3-S1	Bombillo Fluorescente	A3	S1: Encendido - Apagado	101	201
			A1A3A12		20	
			A1A3A22		20	
			A2A3A22		20	
			A3A11A13		20	
			A3A12A13		20	
12	TS1A4-S1	Bombillo Fluorescente	A4	S1: Encendido - Apagado	108	228
			A1A4		10	
			A4A7		10	
			A4A8		10	
			A4A11		10	
			A4A12		10	
			A4A21		10	
			A1A2A4		20	
			A2A4A11		20	
			A2A4A19		20	
13	TS1A5-S1	Cargador de Celular	A5	S1: Conexión - Desconexión	100	170
			A1A5		10	
			A4A5		10	
			A5A7		10	
			A5A8		10	
			A5A11		10	
			A5A12		10	
			A5A21		10	
14	TS1A6-S1	Computador de Escritorio	A6	S1: Secuencia 1	100	100
15	TS1A6-S2	Computador de Escritorio	A6	S2: Secuencia 2	100	100
16	TS1A6-S3	Computador de Escritorio	A6	S3: Secuencia 3	100	100
17	TS1A7-S1	Ventilador	A7	S1: Velocidad 1	104	224
			A1A7		10	
			A4A7		10	
			A7A8		10	
			A7A11		10	
			A7A12		10	
			A7A21		10	
			A7A11A13		20	
			A7A11A19		20	
			A7A11A21		20	
18	TS1A8-S1	Nevera	A8	S1: Apertura Puerta	100	170
			A1A8		10	
			A7A8		10	
			A8A11		10	
			A8A12		10	
			A8A21		10	
			A8A11A22		20	
19	TS1A8-S2	Nevera	A8	S2: Ciclo 1	100	100
20	TS1A9-S1	Secador de Cabello	A9	S1: Temperatura 1	100	100
21	TS1A9-S2	Secador de Cabello	A9	S2: Temperatura 2	100	100
22	TS2A1-S1	Licuadora	A1	S1: Velocidad 1	100	180
			A1A2A4		20	
			A1A3A12		20	
			A1A3A22		20	
			A1A19A21		20	
23	TS2A1-S2	Licuadora	A1	S2: Velocidad 2	100	170
			A1A4		10	
			A1A5		10	
			A1A7		10	

			A1A8		10	
			A1A11		10	
			A1A12		10	
			A1A21		10	
24	TS2A10-S1	Bombillo Halógeno	A10	S1: Encendido - Apagado	99	99
			A11		104	
			A1A11		10	
			A4A11		10	
			A5A11		10	
			A7A11		10	
			A8A11		10	
			A11A12		10	
25	TS2A11-S1	Bombillo Halógeno	A11A21	S1: Encendido - Apagado	10	314
			A2A4A11		20	
			A3A11A13		20	
			A7A11A13		20	
			A7A11A19		20	
			A7A11A21		20	
			A8A11A22		20	
			A11A13A21		20	
			A12		100	
			A1A12		10	
			A4A12		10	
			A7A12		10	
26	TS2A12-S1	Plancha	A8A12	S1: Conexión - Desconexión	10	200
			A11A12		10	
			A12A21		10	
			A1A3A12		20	
			A3A12A13		20	
			A13		110	
			A3A11A13		20	
27	TS2A13-S1	Bombillo LED	A3A12A13	S1: Encendido - Apagado	20	190
			A7A11A13		20	
			A11A13A21		20	
			A19		100	
			A1A19A21		20	
28	TS2A19-S1	Sanduchera	A2A4A19	S1: Conexión - Desconexión	20	180
			A2A19A21		20	
			A7A11A19		20	
			A2		97	
			A1A2A4		20	
29	TS2A2-S1	Bombillo Incandescente	A2A3A22	S1: Encendido - Apagado	20	197
			A2A4A11		20	
			A2A4A19		20	
			A2A19A21		20	
			A1A21		10	
			A4A21		10	
			A5A21		10	
			A7A21		10	
			A8A21		10	
30	TS2A21-S1	Televisor TRC	A11A21	S1: Conexión - Desconexión	10	150
			A12A21		10	
			A1A19A21		20	
			A2A19A21		20	
			A7A11A21		20	
			A11A13A21		20	
31	TS2A22-S1	Cargador de Portátil	A1A3A22	S1: Encendido - Apagado	20	40
			A2A3A22		20	
			A3		101	
			A1A3A12		20	
32	TS2A3-S1	Bombillo Fluorescente	A1A3A22	S1: Encendido - Apagado	20	201
			A2A3A22		20	
			A3A11A13		20	
			A3A12A13		20	
			A4		108	
			A1A4		10	
			A4A5		10	
33	TS2A4-S1	Bombillo Fluorescente	A4A7	S1: Encendido - Apagado	10	238
			A4A8		10	
			A4A11		10	
			A4A12		10	

			A4A21		10	
			A1A2A4		20	
			A2A4A11		20	
			A2A4A19		20	
34	TS2A5-S1	Cargador de Celular	A5	S1: Conexión - Desconexión	100	170
			A1A5		10	
			A4A5		10	
			A5A7		10	
			A5A8		10	
			A5A11		10	
			A5A12		10	
			A5A21		10	
35	TS2A6-S1	Computador de Escritorio	A6	S1: Secuencia 1	100	100
36	TS2A6-S2	Computador de Escritorio	A6	S2: Secuencia 2	100	100
37	TS2A6-S3	Computador de Escritorio	A6	S3: Secuencia 3	100	100
38	TS2A7-S1	Ventilador	A7	S1: Velocidad 1	104	234
			A1A7		10	
			A4A7		10	
			A5A7		10	
			A7A8		10	
			A7A11		10	
			A7A12		10	
			A7A21		10	
			A7A11A13		20	
			A7A11A19		20	
			A7A11A21		20	
39	TS2A8-S1	Nevera	A8	S1: Apertura Puerta	100	189
			A1A8		10	
			A4A8		9	
			A5A8		10	
			A7A8		10	
			A8A11		10	
			A8A12		10	
			A8A21		10	
			A8A11A22		20	
40	TS2A8-S2	Nevera	A8	S2: Ciclo 1	100	100
41	TS2A9-S1	Secador de Cabello	A9	S1: Temperatura 1	100	100
42	TS2A9-S2	Secador de Cabello	A9	S2: Temperatura 2	100	100
43	TS3A6-S2	Computador de Escritorio	A6	S2: Secuencia 2	100	100
44	TS4A6-S2	Computador de Escritorio	A6	S2: Secuencia 2	100	100

Apéndice B. Descriptores considerados para cada una de las bases de casos

Tabla 21. *Descriptores usados para la conformación de la base de casos llamada FV para estado transitorio. Los descriptores desde 1-13 y desde el 22 al 36 son tomados de la IEEE 1459 Standart. Los descriptores desde el 14 hasta el 21 son tomados de (IEEE, 2010).*

DESCRIPTOR	DESCRIPCIÓN	NOTACIÓN
1-13	Coeficientes Armónicos de 1 al 13	H_i con $i = 1, 2, ...13$
14	Máximo del Estado Estable: El máximo valor que se encuentra en el intervalo definido para estado estable.	max
15	Factor de cresta de la señal estacionaria.	$FC = \frac{V_{max}}{V_{rms}}$
16	Medida de la entropía de la señal estacionaria.	$H = -\sum_{i=1}^{k} p_i log_2(p_i)$
17	Coeficiente de Asimetría de la señal transitoria.	$\gamma = \frac{\kappa_3}{\kappa^{\frac{3}{2}}}$
18	Medida de la Kurtosis de la señal transitoria.	$\kappa = \frac{\mu_4}{\sigma^4}$
19	Pendiente 1	m_1
20	Pendiente 2	m_2
21	Área: Área bajo la curva de la señal estacionaria.	$A = \int_{t_0}^{t} i(t)dt$
22	Voltaje Medio Cuadrático: Es la medida del voltaje medio cuadrático.	$V_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} V_i^2}$
23	Corriente Restante para Señales no Sinusoidales: Es el valor contenido de calcular la resta de la corriente instantánea con la corriente armónica de la fundamental.	$I_H = I_0 + \sqrt{2}\sum_{h\neq 1} I_h sin(h\omega t - \beta_h)$
24	Voltaje restante para señales no sinusoidales: Es el valor obtenido de calcular la resta del voltaje instantáneo con voltaje armónico fundamental.	$V_H = V_0 + \sqrt{2}\sum_{h\neq 1} V_h sin(h\omega t - \alpha_h)$
25	Potencia Activa: Es la medida la potencia activa extraída se la señal estacionaria.	$P = V_{rms} I_{rms}$
26	Potencia Armónica restante para señales no sinusoidales: Es el producto del valor de voltaje armónico con la corriente armónica.	$P_H = V_H I_H$
27	Potencia Reactiva: Es la medida la potencia reactiva extraída se la señal estacionaria.	$Q = \sqrt{S^2 - P^2}$
28	Potencia Aparente: Es la medida la potencia aparente calculada como producto del valor eficaz de corriente y de voltaje.	$S = V_{rms} I_{rms}$
29	Potencia Aparente Fundamental: Es la medida la potencia aparente calculada como producto del valor armónico fundamental de corriente y armónico fundamental de voltaje.	$S_1 = V_1 I_1$ donde $v = V_1 + V_H = I_1 + I_H$

DESCRIPTOR	DESCRIPCIÓN	NOTACIÓN
30	Potencia Aparente restante para señales no sinusoidales: Es la medida la potencia aparente calculada como producto del valor armónico fundamental de corriente y armónico fundamental de voltaje.	$S_H = V_H I_H$
31	Potencia Aparente no fundamental: Es la potencia aparente no fundamental (pendiente) .	$S_N = \sqrt{S^2 - S_1^2}$
32	Potencia de Distorsión de Corriente: Es el valor del producto del voltaje fundamental RMS y la corriente armónica.	$D_i = V_1 I_H = S_1 THD_i$
33	Potencia de Distorsión de Voltaje: Es el valor del producto del voltaje armónica por la corriente fundamental.	$D_v = V_H I_1 = S_1 THD_v$
34	Potencia de Distorsión Armónica.	$D_H = \sqrt{S_H^2 - P_H^2}$
35	Factor de Potencia Fundamental: Es el valor del factor de potencia resultante del producto.	$FP_1 = cos(\theta) = \frac{P_1}{S_1}$
36	Distorsión Armónica Total de Corriente: Es el valor de la sumatorio Pendiente.	$THD_i = \frac{I_H}{I_1}$
37	Distorsión Armónica Total de Voltaje: Es el valor de la sumatorio Pendiente.	$THD_v = \frac{V_H}{V_1}$

Tabla 22. *Descriptores usados para la conformación de la base de casos llamada FVTime para estado transitorio.*

DESCRIPTOR	DESCRIPCIÓN	NOTACIÓN
1	Punto inicial del transitorio: Este descriptor mide en qué punto respecto al punto cero de la señal.	$P_i \in T$ donde T es el régimen transitorio
2	Longitud del transitorio: Se mide la distancia temporal entre el punto de inicio y el punto final del transitorio.	$L = t_n - t_{n-1} = \Delta T$
3	Promedio del transitorio: Se toma la media aritmética de los datos registrados en la duración del transitorio.	$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$
4	Máximo valor del transitorio.	max
5	Factor de cresta de la señal transitoria.	$FC = \frac{V_{max}}{V_{rms}}$
6	Desviación estándar de la señal transitoria.	$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_i - \hat{x})^2}$
7	Medida de la entropía de la señal transitoria.	$H = -\sum_{i=1}^{k} p_i log_2(p_i)$
8	Coeficiente de Asimetría de la señal transitoria.	$\gamma = \frac{\kappa_3}{\kappa^{\frac{3}{2}}}$
9	Medida de la Kurtosis de la señal transitoria.	$\kappa = \frac{\mu_4}{\sigma^4}$

Tabla 23. *Forma de cálculo de los descriptores usados en la construcción de la base de casos Wavelet estacionario y transitorio.*

BASES DE CASOS	DESCRIPTORES	DESCRIPCIÓN	NOTACIÓN
Wavelet Estacionario	1-40	Escalas de la Transformada Wavelet: Este descriptor es la contribución de energía obtenida por cada representada por las escalas definida por una escala inicial a con una frecuencia central de 3dB de la transformada Wavelet de Morlet. $\hat{\Psi} = e^{-\pi^2 (f-F_c)^2 . \Delta F}$ donde ΔD está definido por: $\Delta F_{3dB} = \frac{2}{\pi}\sqrt{\frac{ln\sqrt{2}}{\Delta F}}$ con a_0 es la escala inicial definida por: $a_0 = \frac{2F_c}{F_s}$ y las escalas siguientes definidas por a_j sería: $a_j = \left(\frac{2F_c + 3\Delta F_{3dB}}{2F_c - 3\Delta F_{3dB}}\right)^j a_0$	$W_C = a_j$
Wavelet Transitorio	1-75		$W_C = a_j$

Tabla 24. *Forma de cálculo de los descriptores usados en la construcción de la base de casos S-Transform estacionario y transitorio.*

BASES DE CASOS	DESCRIPTORES	DESCRIPCIÓN	NOTACIÓN
S-Transform Transitorio	Desviación Estándar	Se toma de la matriz S y desde la primera columna A_1 hasta la l-ésima columna A_l se calcula la desviación estándar del conjunto de datos correspondiente al máximo de cada columna $\sigma_1((max(AC_l) \mid l = 0,1,...,N-1))$	$F_1 = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_i - \hat{x})^2}$
	Desviación Estándar	Se toma de la matriz S y desde la primera fila A_1 hasta la n-ésima fila A_n se calcula la desviación estándar del conjunto de datos correspondiente al máximo de cada fila $\sigma((max(AR_n) \mid l = 0,1,...,N-1))$	$F_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_i - \hat{x})^2}$
	Potencia	La potencia espectral de los datos correspondientes al máximo de cada columna	$F_3 = \frac{1}{N}\sum_{l=0}^{N-1}\left\|max\left(Ac_l\right)\right\|^2$

BASES DE CASOS	DESCRIPTORES	DESCRIPCIÓN	NOTACIÓN
	Desviación Estándar	Se toma de la matriz S Desviación estándar del conjunto de datos correspondiente al fase de contorno máximo de cada columna de la matriz S.	$F_4 = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_i - \hat{x})^2}$
	Kurtosis	La Kurtosis del conjunto de datos correspondiente a las magnitudes máximas de cada columna de la matriz S $\kappa = ((max(AC_l) \mid l = 0,1,...,N-1))$	$F_5 = \frac{\mu_4}{\sigma^4}$
	Factor de Cresta	Factor de cresta es la relación entre el amplitud pico y el valor medio cuadrático	$F_6 = \frac{V_{max}}{V_{rms}}$
	Duración del Transitorio	Es la característica calculada como la duración del transitorio	$F_7 = longitud(X[k])$
S-Transform Estacionario	$F_1 - F_6$	Usando los descriptores desde $F_1 - F_6$ de las características transitorias pero aplicadas a las señales extraídas del régimen estacionario.	$F_1 - F_6$

Tabla 25. *Forma de cálculo de los descriptores usados en la construcción de la base de casos FRTF estacionario y transitorio.*

BASES DE CASOS	DESCRIPTORES	DESCRIPCIÓN	NOTACIÓN
FRTF Estacionario	1-40	Órdenes Fraccionarios: Este descriptor es obtenido de la distribución de energía deWigner-Viles definida como:	$W_c = a_j$
FRTF Transitorio	1-75	$W[x](t,\mu) = \int_{-\infty}^{+\infty} x(t+y/2)x^*(y/2)e^{-2i\pi\mu y}dy$	$W_c = a_j$

Tabla 26. *Descriptores usados para la conformación de la base de casos llamada IEEE1459 para estado estacionario. Los descriptores son tomados de la IEEE 1459 Standart.*

DESCRIPTOR	DESCRIPCIÓN	NOTACIÓN
1	*Voltaje Medio Cuadrático*: Es la medida del voltaje medio cuadrático.	$V_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}V_i^2}$
2	*Corriente Media Cuadrático*: Es la medida de la corriente media cuadrático.	$I_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}I_i^2}$
3	*Voltaje Medio*: Es la medida del voltaje medio de la señal.	$V_{DC} = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} v(t)dt$
4	*Corriente Media*: Es la medida de la corriente media cuadrático.	$I_{DC} = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} i(t)dt$
5	Potencia Activa: Es la medida la potencia activa extraída se la señal estacionaria.	$P = V_{rms}I_{rms}$
6	*Potencia Aparente*: Es la medida la potencia aparente calculada como producto del valor eficaz de corriente y de voltaje.	$S = V_{rms}I_{rms}$
7	*Potencia Reactiva*: Es la medida la potencia reactiva extraída se la señal estacionaria.	$Q = \sqrt{S^2 - P^2}$

DESCRIPTOR	DESCRIPCIÓN	NOTACIÓN
8	*Factor de Potencia Fundamental*: Es el valor del factor de potencia resultante del producto.	$FP_1 = \cos(\theta) = \frac{P_1}{S_1}$
9	*Corriente Armónica Fundamental*: Es la valor contenido de calcular la resta de la corriente instantánea con la corriente armónica de la fundamental.	$I_H = I_0 + \sqrt{2}\sum_{h\neq 1} I_h \sin(h\omega t - \beta_h)$
10	Distorsión Armónica Total de Corriente: Es el valor de la sumatorio Pendiente.	$THD_i = \frac{I_H}{I_1}$
11	*Voltaje Armónico Fundamental*: Es la valor contenido de calcular la resta de la corriente instantánea con la corriente armónica de la fundamental.	$I_H = I_0 + \sqrt{2}\sum_{h\neq 1} I_h \sin(h\omega t - \beta_h)$
12	Distorsión Armónica Total de Voltaje: Es el valor de la sumatorio Pendiente.	$THD_v = \frac{V_H}{V_1}$
13	PCA_1	H_i con $i = 1, 2, \ldots 13$
14	PCA_2 de 1 al 13	H_i con $i = 1, 2, \ldots 13$
15	T^2	H_i con $i = 1, 2, \ldots 13$

Apéndice C. Histogramas para los modelos de RBC considerados para cada una de las bases de casos en los diferentes regímenes.

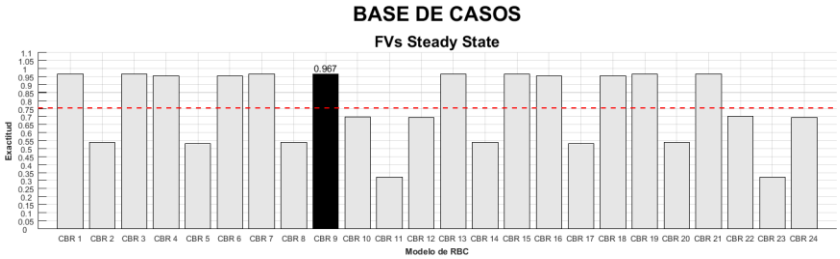


Figura 18. Resultados de desempeños para la base de casos FV en régimen estacionario.

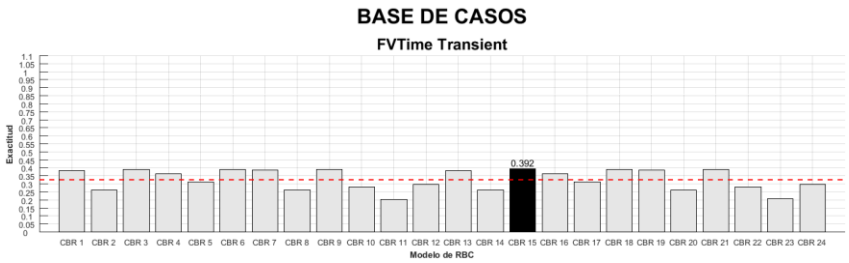


Figura 19. Resultados de desempeños para la base de casos FVTime en régimen transitorio.

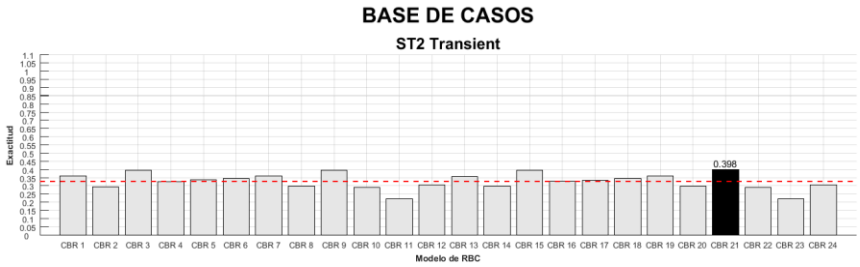


Figura 20. Resultados de desempeños para la base de casos ST2 en régimen transitorio.

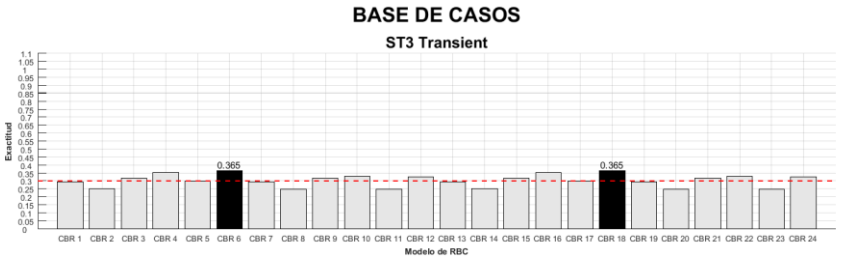


Figura 21. Resultados de desempeños para la base de casos ST3 en régimen transitorio.

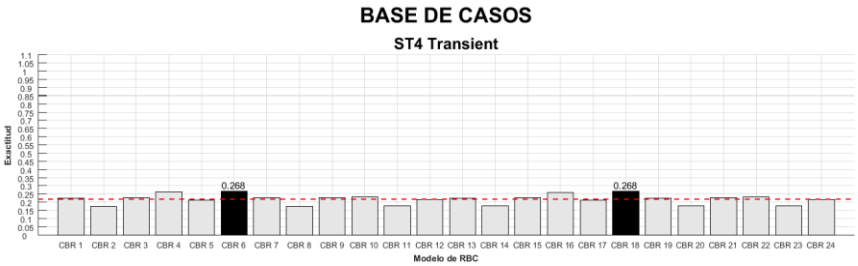


Figura 22. Resultados de desempeños para la base de casos ST4 en régimen transitorio.

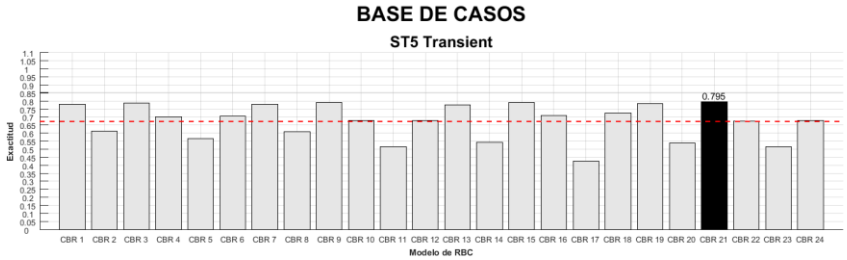


Figura 23. Resultados de desempeños para la base de casos ST5 en régimen transitorio.

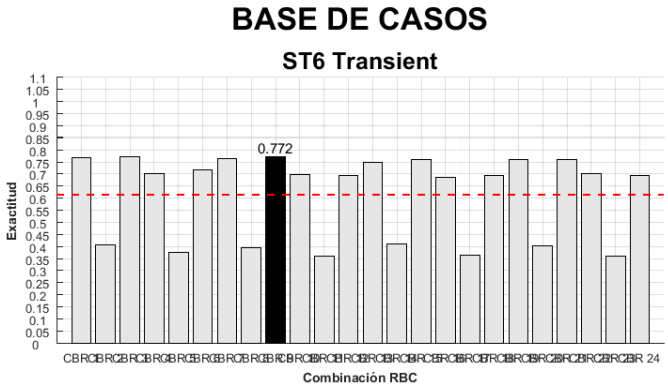


Figura 24. Resultados de desempeños para la base de casos ST5 en régimen transitorio.

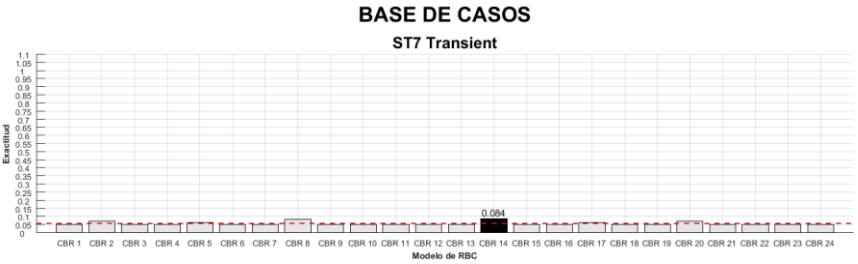


Figura 25. Resultados de desempeños para la base de casos ST7 en régimen transitorio.

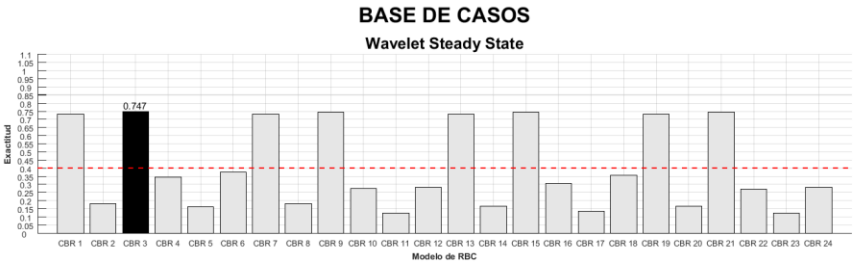


Figura 26. Resultados de desempeños para la base de casos Wavelet en régimen estacionario.

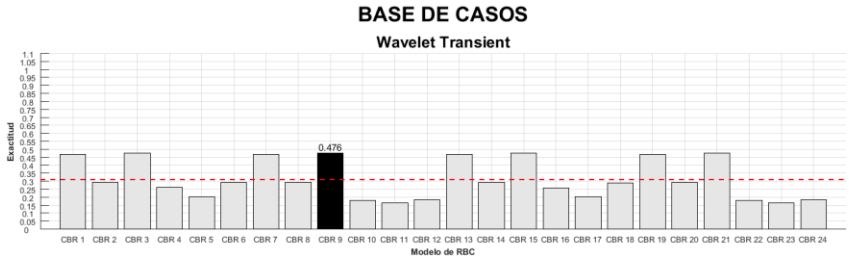


Figura 27. Resultados de desempeños para la base de casos Wavelet en régimen transitorio.

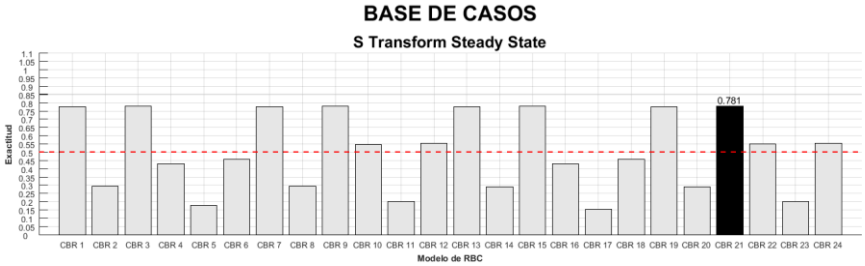


Figura 28. Resultados de desempeños para la base de casos S-Transformen régimen estacionario.

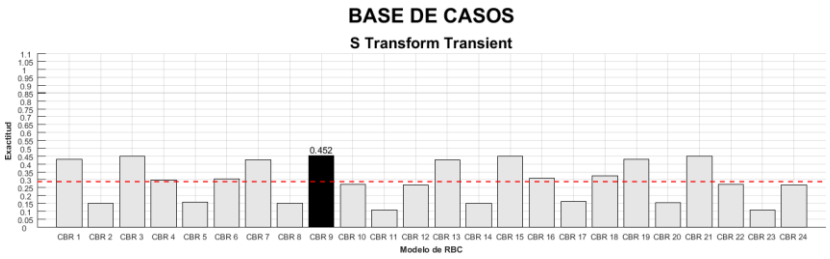


Figura 29. Resultados de desempeños para la base de casos S-Transformen régimen transitorio.

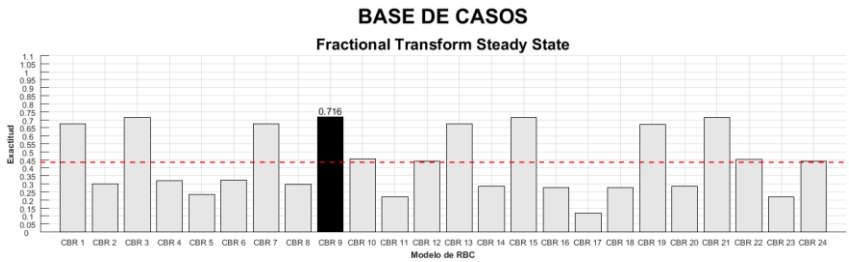


Figura 30. Resultados de desempeños para la base de casos FRFT en régimen estacionario.

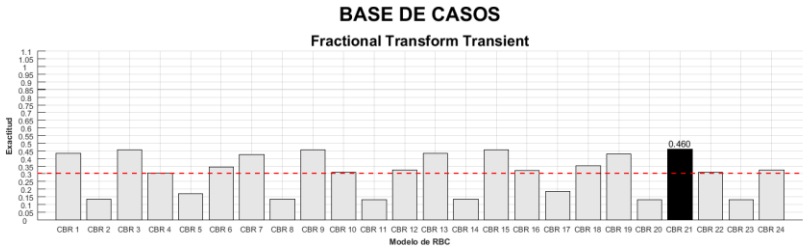


Figura 31. Resultados de desempeños para la base de casos FRFT en régimen transitorio.

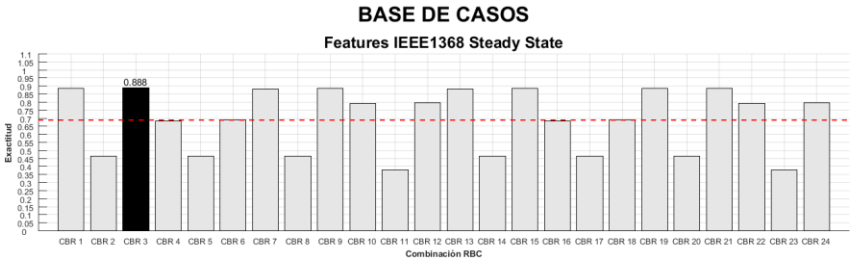


Figura 32. Resultados de desempeños para la base de casos FRFT en régimen transitorio.

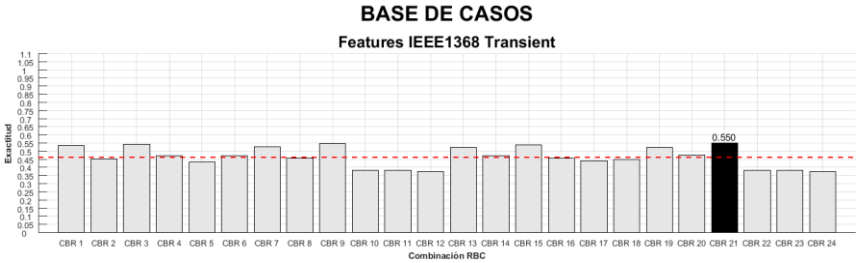


Figura 33. Resultados de desempeños para la base de casos FRFT en régimen transitorio.

Apéndice D. Datos obtenidos en cada partición para comparación

Tabla 27. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, Wavelet Estacionario.*

COMPARACIÓN BASE DE CASOS WAVELET RÉGIMEN ESTACIONARIO

	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
MSV	0,99	0,99	0,99	1	0,99	0,99	0,99	0,99	0,99	0,99
RBC 1	0,73	0,7	0,72	0,72	0,71	0,7	0,71	0,7	0,69	0,68
RBC 2	0,29	0,28	0,29	0,28	0,27	0,28	0,27	0,26	0,26	0,26
RBC 3	0,73	0,72	0,73	0,74	0,71	0,71	0,72	0,71	0,71	0,68
RBC 4	0,23	0,23	0,24	0,23	0,23	0,23	0,23	0,23	0,24	0,23
RBC 5	0,12	0,12	0,12	0,13	0,12	0,13	0,13	0,12	0,13	0,1
RBC 6	0,25	0,26	0,25	0,26	0,24	0,26	0,26	0,25	0,25	0,25
RBC 7	0,72	0,7	0,73	0,72	0,7	0,71	0,71	0,7	0,69	0,69
RBC 8	0,29	0,28	0,3	0,28	0,28	0,29	0,27	0,27	0,27	0,26
RBC 9	0,73	0,71	0,73	0,73	0,71	0,71	0,72	0,71	0,71	0,69
RBC 10	0,29	0,29	0,3	0,3	0,26	0,27	0,28	0,29	0,28	0,27
RBC 11	0,18	0,18	0,18	0,18	0,17	0,17	0,17	0,18	0,18	0,18
RBC 12	0,72	0,7	0,74	0,72	0,7	0,7	0,72	0,7	0,69	0,68
RBC 13	0,28	0,26	0,28	0,27	0,26	0,27	0,25	0,25	0,25	0,25
RBC 14	0,73	0,71	0,74	0,73	0,7	0,71	0,71	0,71	0,7	0,69
RBC 15	0,2	0,2	0,21	0,21	0,21	0,19	0,21	0,21	0,2	0,2
RBC 16	0,1	0,1	0,1	0,12	0,11	0,1	0,11	0,1	0,11	0,1
RBC 17	0,23	0,24	0,23	0,24	0,24	0,22	0,25	0,24	0,24	0,24
RBC 18	0,72	0,7	0,73	0,73	0,7	0,7	0,72	0,71	0,69	0,68
RBC 19	0,28	0,26	0,28	0,27	0,26	0,27	0,25	0,24	0,25	0,25
RBC 20	0,73	0,71	0,73	0,74	0,71	0,7	0,72	0,71	0,71	0,69
RBC 21	0,29	0,29	0,3	0,3	0,25	0,27	0,28	0,28	0,28	0,27
RBC 22	0,18	0,18	0,18	0,18	0,17	0,17	0,17	0,18	0,18	0,18
RBC 23	0,29	0,3	0,3	0,31	0,26	0,27	0,29	0,29	0,29	0,28
RBC 24	0,99	0,99	0,99	1	0,99	0,99	0,99	0,99	0,99	0,99

Tabla 28. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, Wavelet Transitorio.*

COMPARACIÓN BASE DE CASOS WAVELET RÉGIMEN TRANSITORIO

	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
MSV	0,34	0,36	0,34	0,34	0,35	0,37	0,38	0,34	0,36	0,35
RBC 1	0,51	0,48	0,5	0,46	0,49	0,49	0,5	0,51	0,48	0,5
RBC 2	0,36	0,34	0,36	0,34	0,34	0,34	0,37	0,34	0,33	0,37
RBC 3	0,5	0,49	0,5	0,48	0,51	0,51	0,5	0,52	0,48	0,51
RBC 4	0,28	0,28	0,27	0,27	0,26	0,26	0,28	0,29	0,27	0,26
RBC 5	0,24	0,22	0,23	0,2	0,21	0,22	0,24	0,22	0,23	0,23
RBC 6	0,29	0,29	0,29	0,27	0,29	0,29	0,29	0,3	0,29	0,28
RBC 7	0,51	0,48	0,5	0,46	0,49	0,49	0,49	0,52	0,47	0,5
RBC 8	0,36	0,33	0,36	0,34	0,35	0,34	0,37	0,35	0,33	0,37
RBC 9	0,5	0,49	0,51	0,49	0,51	0,5	0,5	0,53	0,48	0,51
RBC 10	0,26	0,25	0,24	0,23	0,25	0,24	0,27	0,26	0,24	0,25
RBC 11	0,24	0,24	0,24	0,22	0,24	0,22	0,25	0,23	0,23	0,24

	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RBC 12	0,26	0,25	0,25	0,24	0,25	0,23	0,27	0,26	0,25	0,25
RBC 13	0,51	0,48	0,5	0,46	0,5	0,49	0,5	0,51	0,48	0,5
RBC 14	0,36	0,33	0,37	0,34	0,34	0,34	0,37	0,34	0,33	0,37
RBC 15	0,5	0,49	0,51	0,48	0,51	0,51	0,5	0,52	0,48	0,52
RBC 16	0,27	0,28	0,27	0,26	0,26	0,26	0,28	0,28	0,27	0,26
RBC 17	0,24	0,22	0,23	0,21	0,21	0,22	0,24	0,22	0,23	0,23
RBC 18	0,29	0,28	0,29	0,27	0,28	0,28	0,3	0,3	0,28	0,27
RBC 19	0,51	0,48	0,5	0,47	0,5	0,49	0,5	0,5	0,48	0,49
RBC 20	0,36	0,33	0,35	0,34	0,34	0,34	0,37	0,35	0,33	0,38
RBC 21	0,51	0,49	0,51	0,48	0,51	0,5	0,5	0,53	0,48	0,52
RBC 22	0,26	0,25	0,25	0,23	0,25	0,24	0,27	0,25	0,24	0,25
RBC 23	0,24	0,24	0,24	0,22	0,24	0,22	0,25	0,23	0,23	0,24
RBC 24	0,26	0,25	0,25	0,24	0,25	0,23	0,27	0,26	0,25	0,25

Tabla 29. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, FRFT Estacionario.*

COMPARACIÓN BASE DE CASOS FRFT RÉGIMEN ESTACIONARIO

	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
SVM	0,93	0,92	0,92	0,92	0,92	0,92	0,92	0,91	0,93	0,93
RBC 1	0,65	0,64	0,63	0,66	0,64	0,64	0,61	0,65	0,65	0,62
RBC 2	0,42	0,4	0,39	0,43	0,41	0,41	0,39	0,41	0,41	0,4
RBC 3	0,67	0,64	0,65	0,69	0,67	0,66	0,65	0,66	0,67	0,63
RBC 4	0,2	0,19	0,21	0,22	0,2	0,2	0,21	0,21	0,2	0,21
RBC 5	0,17	0,16	0,16	0,17	0,16	0,16	0,16	0,17	0,16	0,16
RBC 6	0,21	0,21	0,21	0,22	0,2	0,2	0,21	0,22	0,21	0,2
RBC 7	0,65	0,63	0,63	0,66	0,64	0,63	0,62	0,65	0,64	0,62
RBC 8	0,4	0,39	0,39	0,42	0,41	0,41	0,4	0,41	0,4	0,4
RBC 9	0,67	0,66	0,65	0,68	0,67	0,66	0,67	0,66	0,67	0,64
RBC 10	0,44	0,42	0,41	0,45	0,41	0,41	0,42	0,41	0,42	0,4
RBC 11	0,31	0,29	0,28	0,32	0,29	0,29	0,29	0,28	0,29	0,28
RBC 12	0,44	0,41	0,4	0,43	0,41	0,41	0,41	0,4	0,42	0,39
RBC 13	0,65	0,63	0,62	0,64	0,66	0,64	0,63	0,64	0,64	0,62
RBC 14	0,4	0,39	0,38	0,41	0,4	0,4	0,38	0,41	0,4	0,38
RBC 15	0,66	0,65	0,64	0,67	0,66	0,67	0,66	0,67	0,67	0,66
RBC 16	0,16	0,18	0,18	0,18	0,17	0,17	0,18	0,17	0,18	0,17
RBC 17	0,1	0,1	0,1	0,11	0,09	0,09	0,1	0,09	0,1	0,1
RBC 18	0,17	0,18	0,18	0,19	0,17	0,18	0,18	0,17	0,18	0,16
RBC 19	0,63	0,64	0,62	0,65	0,65	0,63	0,63	0,64	0,65	0,61
RBC 20	0,39	0,39	0,39	0,41	0,41	0,4	0,38	0,42	0,39	0,39
RBC 21	0,67	0,65	0,65	0,68	0,68	0,64	0,65	0,66	0,68	0,63
RBC 22	0,45	0,42	0,41	0,45	0,41	0,41	0,42	0,4	0,42	0,4
RBC 23	0,31	0,29	0,28	0,32	0,29	0,29	0,29	0,28	0,29	0,28
RBC 24	0,44	0,41	0,4	0,43	0,41	0,41	0,41	0,4	0,42	0,39

Tabla 30. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, FRFT Transitorio.*

COMPARACIÓN BASE DE CASOS FRFT RÉGIMEN TRANSITORIO

	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
MSV	0,34	0,36	0,36	0,38	0,37	0,35	0,35	0,3	0,32	0,36
RBC 1	0,49	0,46	0,45	0,43	0,45	0,46	0,44	0,46	0,48	0,43
RBC 2	0,15	0,15	0,13	0,13	0,12	0,15	0,13	0,13	0,15	0,14
RBC 3	0,49	0,49	0,48	0,48	0,49	0,47	0,48	0,48	0,49	0,45
RBC 4	0,31	0,29	0,28	0,31	0,28	0,31	0,31	0,32	0,33	0,29
RBC 5	0,17	0,17	0,16	0,17	0,16	0,17	0,16	0,18	0,16	0,17
RBC 6	0,34	0,33	0,32	0,33	0,31	0,35	0,35	0,34	0,36	0,33
RBC 7	0,47	0,45	0,45	0,43	0,43	0,45	0,45	0,45	0,48	0,43
RBC 8	0,14	0,15	0,12	0,12	0,12	0,14	0,14	0,14	0,14	0,15
RBC 9	0,5	0,48	0,49	0,48	0,49	0,47	0,47	0,47	0,5	0,45
RBC 10	0,38	0,36	0,35	0,34	0,35	0,37	0,38	0,35	0,36	0,35
RBC 11	0,16	0,15	0,14	0,13	0,12	0,17	0,13	0,15	0,14	0,13
RBC 12	0,39	0,37	0,36	0,35	0,35	0,38	0,4	0,37	0,38	0,35
RBC 13	0,49	0,45	0,45	0,43	0,46	0,46	0,46	0,45	0,48	0,44
RBC 14	0,15	0,15	0,13	0,12	0,12	0,15	0,14	0,13	0,14	0,13
RBC 15	0,5	0,48	0,48	0,47	0,49	0,48	0,49	0,48	0,5	0,45
RBC 16	0,33	0,31	0,3	0,33	0,3	0,32	0,34	0,32	0,34	0,3
RBC 17	0,18	0,19	0,18	0,19	0,17	0,18	0,18	0,17	0,19	0,17
RBC 18	0,35	0,34	0,33	0,34	0,31	0,35	0,34	0,34	0,35	0,32
RBC 19	0,5	0,45	0,43	0,44	0,43	0,47	0,45	0,47	0,48	0,4
RBC 20	0,15	0,15	0,12	0,11	0,12	0,13	0,14	0,13	0,14	0,14
RBC 21	0,5	0,48	0,49	0,48	0,48	0,49	0,47	0,48	0,5	0,46
RBC 22	0,38	0,36	0,34	0,33	0,36	0,38	0,38	0,35	0,37	0,34
RBC 23	0,16	0,15	0,14	0,13	0,12	0,17	0,13	0,15	0,14	0,13
RBC 24	0,39	0,37	0,36	0,35	0,35	0,38	0,4	0,37	0,38	0,35

Tabla 31. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, S-Transform Estacionario.*

COMPARACIÓN BASE DE CASOS S-TRANSFORM RÉGIMEN ESTACIONARIO

	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
MVS	0,93	0,92	0,92	0,92	0,92	0,92	0,92	0,91	0,93	0,93
RBC 1	0,74	0,74	0,74	0,74	0,72	0,71	0,71	0,73	0,73	0,75
RBC 2	0,43	0,42	0,42	0,42	0,41	0,42	0,4	0,42	0,42	0,45
RBC 3	0,75	0,73	0,74	0,75	0,72	0,71	0,72	0,74	0,73	0,76
RBC 4	0,27	0,26	0,26	0,26	0,26	0,26	0,26	0,25	0,27	0,26
RBC 5	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,11
RBC 6	0,28	0,28	0,28	0,27	0,28	0,28	0,28	0,27	0,29	0,28
RBC 7	0,74	0,73	0,74	0,74	0,73	0,71	0,72	0,74	0,74	0,75
RBC 8	0,43	0,42	0,42	0,42	0,41	0,42	0,4	0,42	0,42	0,45
RBC 9	0,75	0,74	0,74	0,75	0,72	0,72	0,72	0,74	0,74	0,75
RBC 10	0,5	0,54	0,51	0,52	0,51	0,5	0,52	0,53	0,52	0,53
RBC 11	0,27	0,31	0,29	0,29	0,29	0,28	0,29	0,3	0,29	0,3

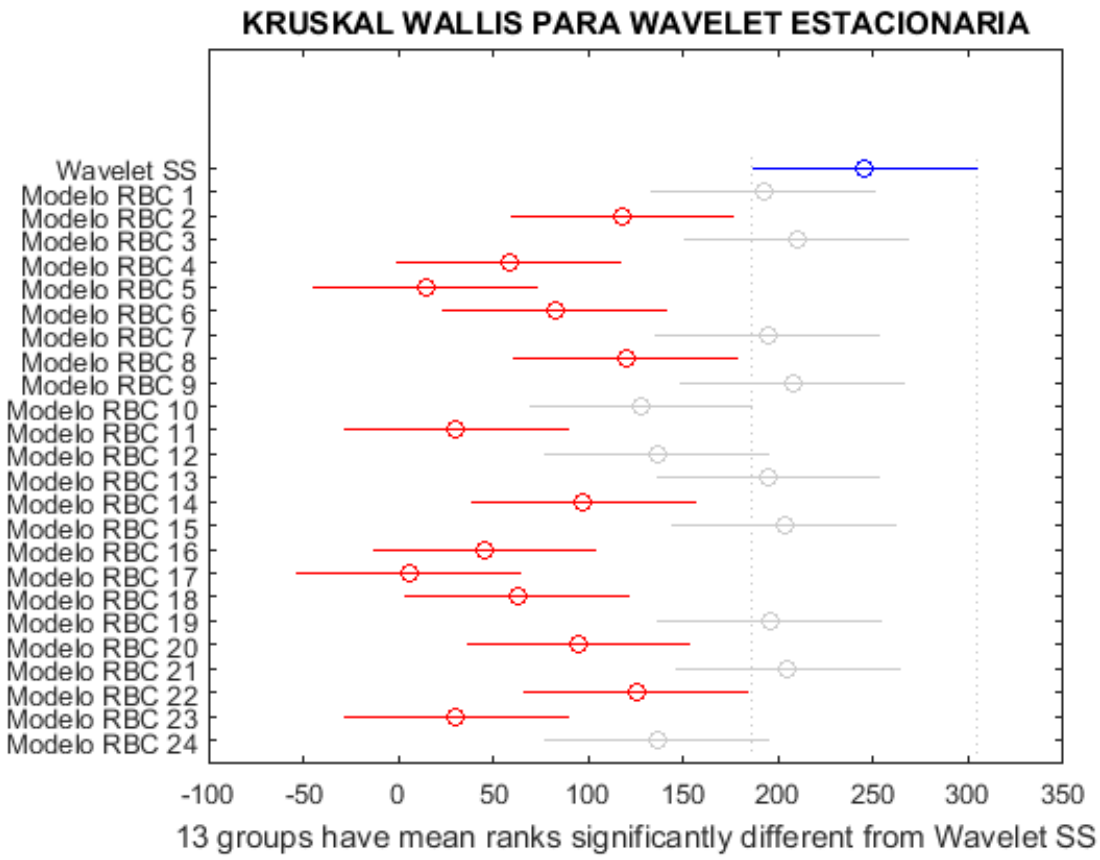
	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RBC 12	0,49	0,54	0,52	0,53	0,52	0,5	0,52	0,53	0,52	0,54
RBC 13	0,74	0,73	0,74	0,74	0,73	0,71	0,71	0,73	0,74	0,75
RBC 14	0,44	0,42	0,42	0,43	0,42	0,42	0,4	0,42	0,42	0,44
RBC 15	0,75	0,74	0,74	0,75	0,72	0,72	0,72	0,74	0,74	0,76
RBC 16	0,27	0,25	0,26	0,25	0,26	0,27	0,27	0,26	0,27	0,26
RBC 17	0,1	0,1	0,1	0,09	0,1	0,1	0,1	0,09	0,09	0,1
RBC 18	0,28	0,28	0,28	0,27	0,27	0,28	0,29	0,28	0,28	0,28
RBC 19	0,75	0,73	0,74	0,74	0,73	0,72	0,71	0,73	0,74	0,76
RBC 20	0,43	0,42	0,42	0,42	0,41	0,42	0,41	0,42	0,42	0,44
RBC 21	0,75	0,75	0,74	0,74	0,73	0,72	0,72	0,73	0,74	0,76
RBC 22	0,49	0,54	0,52	0,52	0,51	0,5	0,52	0,53	0,53	0,53
RBC 23	0,27	0,31	0,29	0,29	0,29	0,28	0,29	0,3	0,29	0,3
RBC 24	0,49	0,54	0,52	0,53	0,52	0,5	0,52	0,53	0,52	0,54

Tabla 32. *Desempeños de los 24 modelos de RBC propuestos y el modelo MSV referente para las bases de casos, S-Transform Transitorio.*

COMPARACIÓN BASE DE CASOS S-TRANSFORM RÉGIMEN TRANSITORIO

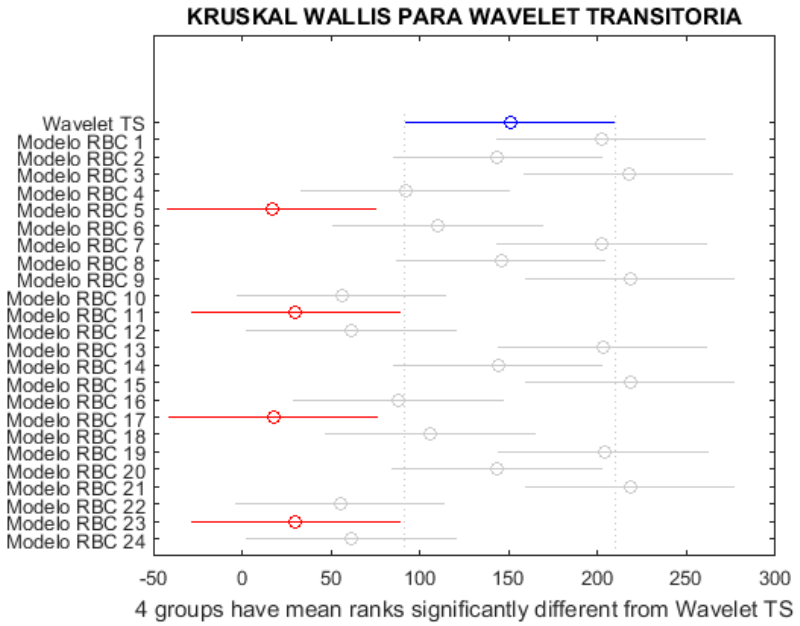
	PARTICIONES									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
MSV	0,33	0,36	0,38	0,39	0,38	0,36	0,37	0,36	0,39	0,37
RBC 1	0,49	0,52	0,45	0,47	0,46	0,47	0,47	0,48	0,46	0,46
RBC 2	0,15	0,18	0,13	0,13	0,13	0,15	0,15	0,15	0,14	0,15
RBC 3	0,51	0,52	0,49	0,48	0,49	0,5	0,5	0,52	0,47	0,48
RBC 4	0,29	0,29	0,29	0,27	0,27	0,29	0,33	0,31	0,31	0,29
RBC 5	0,14	0,15	0,15	0,14	0,15	0,14	0,15	0,15	0,13	0,15
RBC 6	0,31	0,3	0,3	0,27	0,3	0,31	0,33	0,3	0,31	0,31
RBC 7	0,49	0,5	0,45	0,47	0,47	0,46	0,47	0,47	0,47	0,47
RBC 8	0,15	0,17	0,14	0,13	0,14	0,14	0,16	0,15	0,13	0,16
RBC 9	0,51	0,52	0,48	0,48	0,49	0,5	0,5	0,51	0,48	0,47
RBC 10	0,34	0,34	0,32	0,32	0,33	0,32	0,34	0,33	0,32	0,33
RBC 11	0,13	0,11	0,12	0,14	0,1	0,12	0,14	0,11	0,14	0,12
RBC 12	0,35	0,33	0,32	0,32	0,33	0,32	0,33	0,32	0,32	0,32
RBC 13	0,48	0,51	0,45	0,48	0,46	0,47	0,47	0,48	0,46	0,46
RBC 14	0,14	0,17	0,13	0,13	0,13	0,14	0,13	0,15	0,14	0,14
RBC 15	0,51	0,51	0,49	0,48	0,49	0,5	0,5	0,51	0,48	0,48
RBC 16	0,3	0,32	0,31	0,28	0,28	0,31	0,33	0,33	0,32	0,32
RBC 17	0,15	0,15	0,15	0,14	0,17	0,15	0,15	0,16	0,15	0,16
RBC 18	0,33	0,33	0,32	0,29	0,31	0,33	0,34	0,33	0,34	0,34
RBC 19	0,48	0,51	0,46	0,48	0,46	0,47	0,46	0,49	0,47	0,46
RBC 20	0,14	0,17	0,14	0,13	0,14	0,14	0,14	0,15	0,13	0,14
RBC 21	0,52	0,52	0,49	0,48	0,48	0,5	0,5	0,51	0,47	0,48
RBC 22	0,34	0,34	0,32	0,33	0,32	0,32	0,34	0,33	0,32	0,34
RBC 23	0,13	0,11	0,12	0,14	0,1	0,12	0,14	0,11	0,14	0,12
RBC 24	0,35	0,33	0,32	0,32	0,33	0,32	0,33	0,32	0,32	0,32

Apéndice E. Resultados de las pruebas no paramétricas usando el método de Kruskal-Wallis

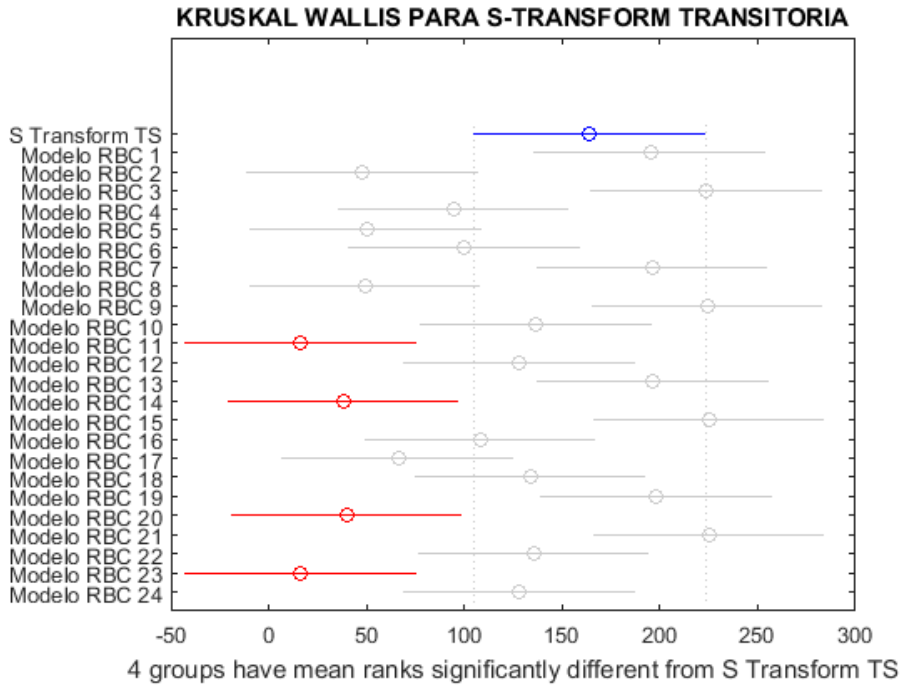


(a) En la base de casos Wavelet en régimen estacionario.

Figura 34. Prueba de significancia usando el método de Kruskal-Wallis para las bases de casos con información extraída de la transformada Wavelet. Las líneas con círculos rojos indica que grupos son los que presentan diferencias estadísticamente significativas. La línea con el círculo azul indica el modelo MSV

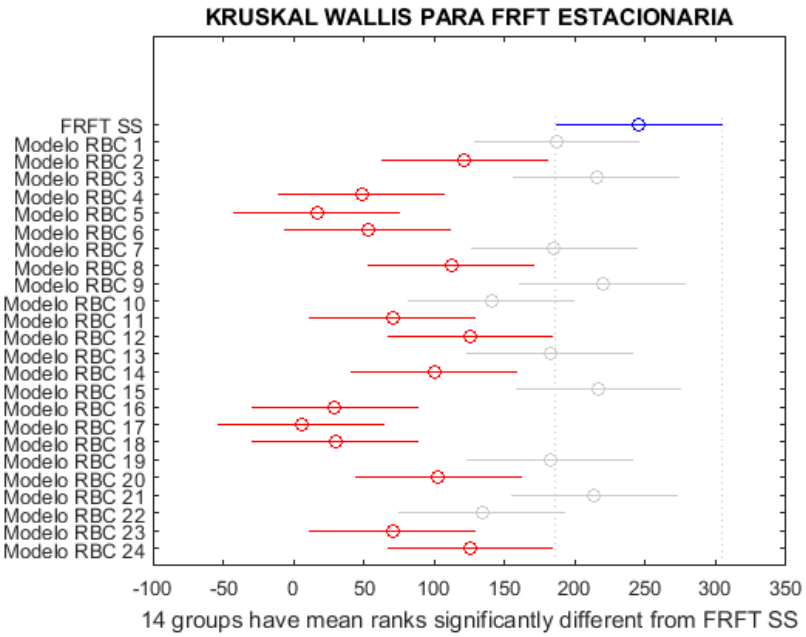


(a) En la base de casos S-Transformen régimen estacionario.

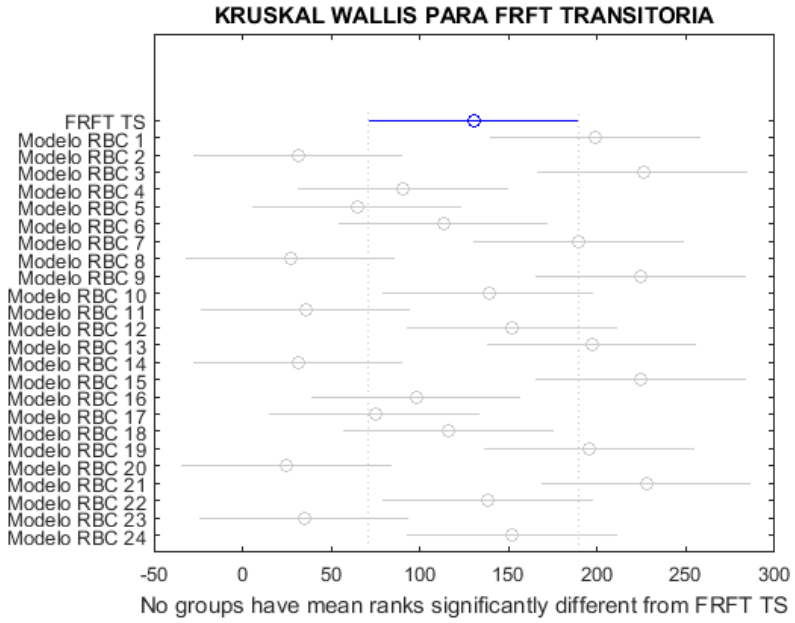


(b) En la base de casos S-Transformen régimen Transitorio.

Figura 35. Prueba de significancia usando el método de Kruskal-Wallis para las bases de casos con información extraída de la transformada S. Las líneas con círculos rojos indica que grupos son los que presentan diferencias estadísticamente significativas. La línea con el círculo azul indica el modelo MSV



(a) En la base de casos FRFT en régimen estacionario.



(b) En la base de casos FRFT en régimen Transitorio.

Figura 36. Prueba de significancia usando el método de Kruskal-Wallis para las bases de casos con información extraída de la Transformada Fraccionaria de Fourier. Las líneas con círculos rojos indica que grupos son los que presentan diferencias estadísticamente significativas. La línea con el círculo azul indica el modelo MSV

Referencias Bibliográficas

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications, 7*(1), 39-59.

Aha, D. W., Kibler, D., & Albert, M. K. (1991). Instance-based learning algorithms. *Machine learning, 6*(1), 37-66.

Aha, D. W., Marling, C., & Watson, I. (2005). Case-based reasoning commentaries: introduction. *The Knowledge Engineering Review, 20*(3), 201-202.

Altrabalsi, H., Stankovic, V., Liao, J., & Stankovic, L. (2016). Low-complexity energy disaggregation using appliance load modelling. *AIMS Energy, 4*(1), 884-905.

Armengol, E., & Plaza, E. (2003, 2003). *Discovery of toxicological patterns with lazy learning*.

Armengol, E., & Plaza, E. (2012). Symbolic explanation of similarities in case-based reasoning. *Computing and informatics, 25*(2-3), 153-171.

Baccigalupo, C., & Plaza, E. (2007, 2007). *A case-based song scheduler for group customised radio*.

Balázs, M., Imre, S., & Lajos, K. (1999). Adaptation Methods in Case-Based Reasoning. *microCAD'99*.

Bartlett, F. C., & Burt, C. (1933). Remembering: A study in experimental and social psychology. *British Journal of Educational Psychology, 3*(2), 187-192.

Berges, M., Goldman, E., Matthews, H. S., & Soibelman, L. (2008). Training load monitoring algorithms on highly sub-metered home electricity consumption data. *Tsinghua Science & Technology, 13*, 406-411.

Berges, M., Goldman, E., Matthews, H. S., Soibelman, L., & Anderson, K. (2011). User-centered nonintrusive electricity load monitoring for residential buildings. *Journal of computing in civil engineering, 25*(6), 471-480.

Berges, M., Soibelman, L., & Matthews, H. S. (2010, 2010). *Building commissioning as an opportunity for training non-intrusive load monitoring algorithms*.

Bergmann, R., Althoff, K.-D., Minor, M., Reichle, M., & Bach, K. (2009). Case-Based Reasoning. *KI, 23*(1), 5-11.

Bhotto, M. Z. A., Makonin, S., & Bajić, I. V. (2017). Load disaggregation based on aided linear integer programming. *IEEE Transactions on Circuits and Systems II: Express Briefs, 64*(7), 792-796.

Bichindaritz, I. (2008). Memory structures and organization in case-based reasoning. In (pp. 175-194): Springer.

Birnbaum, L., & Collins, G. (1989, 1989). *Remindings and engineering design themes: A case study in indexing vocabulary*.

Bonzano, A., Cunningham, P., & Smyth, B. (1997, 1997). *Using introspective learning to improve retrieval in CBR: A case study in air traffic control*.

Boriah, S., Chandola, V., & Kumar, V. (2008, 2008). *Similarity measures for categorical data: A comparative evaluation*.

Bregón, A., & Alfonso, A. (2005). Un sistema de razonamiento basado en casos para la clasificación de fallos en sistemas dinámicos. *Universidad de Valladolid*.

Brüninghaus, S., & Ashley, K. D. (1997, 1997). *Using machine learning for assigning indices to textual cases*.

Cala, H. M. (2018). Identificación de cargas eléctricas residenciales utilizando características basadas en la transformada fraccionaria de Fourier. Tesis Inédita. In.

Chan, K.-P., & Fu, W.-C. (1999, 1999). *Efficient time series matching by wavelets*.

Cheng, V., Li, C.-H., Kwok, J. T., & Li, C.-K. (2004). Dissimilarity learning for nominal data. *Pattern Recognition, 37*(7), 1471-1477.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior, 8*(2), 240-247.

Corchado, J. M., Lees, B., Fyfe, C., Rees, N., & Aiken, J. (1998). Neuro-adaptation method for a case-based reasoning system. *Computing and Information Systems, 5*, 15-20.

Craw, S., Massie, S., & Wiratunga, N. (2007, 2007). *Informed case base maintenance: A complexity profiling approach*.

Cunningham, P. (2009). A taxonomy of similarity mechanisms for case-based reasoning. *IEEE Transactions on Knowledge and Data Engineering, 21*(11), 1532-1543.

Daw, C. S., Finney, C. E. A., & Tracy, E. R. (2003). A review of symbolic analysis of experimental data. *Review of Scientific instruments, 74*(2), 915-930.

De Baets, L., Ruyssinck, J., Develder, C., Dhaene, T., & Deschrijver, D. (2017). On the Bayesian optimization and robustness of event detection methods in NILM. *Energy and Buildings, 145*, 57-66.

Duarte, C., Delmar, P., Goossen, K. W., Barner, K., & Gomez-Luna, E. (2012, 2012). *Non-intrusive load monitoring based on switching voltage transients and wavelet transforms*.

El-Sappagh, S. H., & Elmogy, M. (2015). Case based reasoning: case representation methodologies. *International Journal of Advanced Computer Science & Applications, 1*(6), 192-208.

Farinaccio, L., & Zmeureanu, R. (1999). Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings, 30*(3), 245-259.

Fdez-Riverola, F., Díaz, F., Borrajo, M. L., Yáñez, J. C., & Corchado, J. M. (2005, 2005). *Improving gene selection in microarray data analysis using fuzzy patterns inside a cbr system*.

Gabel, T. (2005, 2005). *On the use of vocabulary knowledge for learning similarity measures*.

Gamero, F. I., Meléndez, J., & Colomer, J. (2011). Qssi: A new similarity index for qualitative time series. Application to classify voltage sags. *Applied Artificial Intelligence, 25*(2), 141-162.

Gómez, M., & Plaza, E. (2012, 2012). *Case-Based Project Scheduling*.

Grech, A., & Main, J. (2005, 2005). *A case-based reasoning approach to formulating university timetables using genetic algorithms*.

Gupta, K. M., Aha, D. W., & Moore, P. (2004, 2004). *Learning feature taxonomies for case indexing*.

Hanney, K., & Keane, M. T. (1996, 1996). *Learning adaptation rules from a case-base*.

Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE, 80*(12), 1870-1891.

IEEE. (2010). Standard Definitions for the Measurement of Electric Power Quantities under Sinusoidal, Nonsinusoidal, Balanced, or Unbalanced Conditions - Redline. *IEEE Std 1459-2010 (Revision of IEEE Std 1459-2000) - Redline*, 1-52. doi:10.1109/IEEESTD.2010.5953405

Iglezakis, I., Reinartz, T., & Roth-Berghofer, T. R. (2004, 2004). *Maintenance memories: beyond concepts and techniques for case base maintenance*.

Initiative, M. I. T. E., & others. (2016). Utility of the future: An MIT Energy Initiative response to an industry in transition. *available at https://energy. mit. edu/wp-content/uploads/2016/12/Utility-of-the-Future-Full-Report. pdf, accessed on December, 20*, 2017-2017.

Jackson, D. A. (1993). Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology, 74*(8), 2204-2214.

Jayanthi, K., Chakraborti, S., & Massie, S. (2010, 2010). *Introspective knowledge revision in textual case-based reasoning*.

Ji, S.-H., Park, M., & Lee, H.-S. (2011). Case adaptation method of case-based reasoning for construction cost estimation in Korea. *Journal of Construction Engineering and Management, 138*(1), 43-52.

Jimenez, Y. (2018). *Automatic Disaggregation of Residential Electrical Consumption with Non-Intrusive Methods.*

Jimenez, Y., Duarte, C., Petit, J., & Carrillo, G. (2014, 2014). *Feature extraction for nonintrusive load monitoring based on S-Transform*.

Jimenez, Y., Duarte, C., Petit, J., Meyer, J., Schegner, P., & Carrillo, G. (2015). Characterization of current switching transients for appliance identification.

Jin, Y., Tebekaemi, E., Berges, M., & Soibelman, L. (2011, 2011). *Robust adaptive event detection in non-intrusive load monitoring for energy aware smart facilities*.

Jorro-Aragoneses, J. L., Díaz-Agudo, B., & Recio-García, J. A. (2015, 2015). *Addressing the Cold-Start Problem in Facial Expression Recognition*.

Khussainova, G. (2016, 2014). *Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning*.

Kolodner, J. (2014). *Case-based reasoning*: Morgan Kaufmann.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review, 6*(1), 3-34.

Lai, Y.-X., Lai, C.-F., Huang, Y.-M., & Chao, H.-C. (2013). Multi-appliance recognition system with hybrid SVM/GMM classifier in ubiquitous smart home. *Information Sciences, 230*, 39-55.

Leake, D. B., & Wilson, D. C. (1998, 1998). *Categorizing case-base maintenance: Dimensions and directions*.

Lee, J., Jung, D.-K., Kim, Y., Lee, Y.-W., & Kim, Y.-M. (2010, 2010). *Smart grid solutions, services, and business models focused on telco*.

Leeb, S. B., Shaw, S. R., & Kirtley, J. L. (1995). Transient event detection in spectral envelope estimates for nonintrusive load monitoring. *IEEE Transactions on Power Delivery, 10*(3), 1200-1210.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association, 62*(318), 399-402. doi:10.1080/01621459.1967.10482916

Louvieris, P. (2008). Case-Based Reasoning Adaptation for High Dimensional Solution Space.

Ludwig, W., Anscombe, G. E. M., & Rhees, R. (1954). Philosophical investigations.

Malherbe, E., Iwaszko, T., & Aufaure, M.-A. (2015, 2015). *A case-based approach for easing schema semantic mapping*.

Marchiori, A., Hakkarinen, D., Han, Q., & Earle, L. (2011). Circuit-level load monitoring for household energy management. *IEEE Pervasive Computing, 10*(1), 40-48.

McKight, P. E., & Najab, J. (2010). Kruskal-Wallis Test. *The corsini encyclopedia of psychology*, 1-1.

Mehta, M., Ontañón, S., & Ram, A. (2009, 2009). *Using meta-reasoning to improve the performance of case-based planning*.

Montani, S., Bottrighi, A., Leonardi, G., Portinale, L., & Terenziani, P. (2009, 2009). *Multi-level abstractions and multi-dimensional retrieval of cases with time series features*.

Murdock, J. W., & Goel, A. K. (2001, 2001). *Meta-case-based reasoning: Using functional models to adapt case-based agents*.

Nalmpantis, C., & Vrakas, D. (2018). Machine learning approaches for non-intrusive load monitoring: from qualitative to quantitative comparation. *Artificial Intelligence Review*, 1-27.

Nemecs, S. T., & Buga, A. (2017, 2017). *Towards a case-based reasoning approach to dynamic adaptation for large-scale distributed systems*.

Nilsson, M., & Funk, P. (2004, 2004). *A case-based classification of respiratory sinus arrhythmia*.

Norford, L. K., & Leeb, S. B. (1996). Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. *Energy and Buildings, 24*(1), 51-64.

Olsson, E., Funk, P., & Xiong, N. (2004). Fault diagnosis in industry using sensor readings and case-based reasoning. *Journal of Intelligent & Fuzzy Systems, 15*(1), 41-46.

Ontañón, S., & Plaza, E. (2007, 2007). *An argumentation-based framework for deliberation in multi-agent systems*.

Paetz, A.-G., Dütschke, E., & Fichtner, W. (2012). Smart homes as a means to sustainable energy consumption: A study of consumer perceptions. *Journal of consumer policy, 35*(1), 23-41.

Pous, C., Colomer, J., Meléndez, J., & de la Rosa, J. L. (2004). Case based reasoning as an extension of fault dictionary methods for linear electronic analog circuits diagnosis. *Recent Advances In Artificial Intelligence Research And Development*, 391-391.

Rashid, H., & Singh, P. (2017, 2017). *Energy disaggregation for identifying anomalous appliance*.

Reinartz, T., Iglezakis, I., & Roth--Berghofer, T. (2001). Review and Restore for Case-Base Maintenance. *Computational Intelligence, 17*(2), 214-234.

Richter, M. M., & Weber, R. O. (2016). *Case-based reasoning*: Springer.

Roth-Berghofer, T., & Iglezakis, I. (2001, 2001). *Six Steps in Case-Based Reasoning: Towards a maintenance methodology for case-based reasoning systems*.

Ruiz Ordóñez, M. L., & others. (2008). *Multivariate statistical process control and case-based reasoning for situation assessment of sequencing batch reactors*: Universitat de Girona.

Ruzzelli, A. G., Nicolas, C., Schoofs, A., & O'Hare, G. M. P. (2010, 2010). *Real-time recognition and profiling of appliances through a single electricity sensor*.

Sarkheyli, A., & Söffker, D. (2015). Case indexing in Case-Based Reasoning by applying Situation Operator Model as knowledge representation model. *IFAC-PapersOnLine, 48*(1), 81-86.

Schank, R. C. (1975). The structure of episodes in memory. In (pp. 237-272): Elsevier.

Schnell, M., Couffignal, S., Lieber, J., Saleh, S., & Jay, N. (2017, 2017). *Case-Based Interpretation of Best Medical Coding Practices—Application to Data Collection for Cancer Registries*.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts* (Vol. 9): Harvard University Press Cambridge, MA.

Smyth, B. (1998, 1998). *Case-base maintenance*.

Smyth, B., & Cunningham, P. (1996, 1996). *The utility problem analysed*.

Smyth, B., & McKenna, E. (2001). Competence models and the maintenance problem. *Computational Intelligence, 17*(2), 235-249.

Stanfill, C., & Waltz, D. (1986). Toward memory-based reasoning. *Communications of the ACM, 29*(12), 1213-1228.

Stéphane, N., Hector, R., & others. (2010). Effective retrieval and new indexing method for case based reasoning: application in chemical process design. *Engineering Applications of Artificial Intelligence, 23*(6), 880-894.

Stockwell, R. G., Mansinha, L., & Lowe, R. P. (1996). Localization of the complex spectrum: the S transform. *IEEE transactions on signal processing, 44*(4), 998-1001.

Supic, H., & Ribaric, S. (2005, 2005). *Autonomous creation of new situation cases in structured continuous domains*.

Tulving, E., & others. (1972). Episodic and semantic memory. *Organization of memory, 1*, 381-403.

Tversky, A. (1977). Features of similarity. *Psychological review, 84*(4), 327-327.

Tzanetakis, G., Essl, G., & Cook, P. (2001, 2001). *Audio analysis using the discrete wavelet transform*.

Vehviläinen, A., Alm, O., & Hyvönen, E. (2006, 2006). *Combining case-based reasoning and semantic indexing in a question-answer service*.

Voskoglou, M. G., & Salem, A.-B. M. (2014). Analogy-based and case-based reasoning: Two sides of the same coin. *arXiv preprint arXiv:1405.7567*.

Weber, B., Rinderle, S., Wild, W., & Reichert, M. (2005, 2005). *CCBR-driven business process evolution*.

Wilson, D. R., & Martinez, T. R. (1996, 1996). *Value difference metrics for continuously valued attributes*.

Wilson, D. R., & Martinez, T. R. (2000). Reduction techniques for instance-based learning algorithms. *Machine learning*, *38*(3), 257-286.

Wortmann, K. (1997). On case representation and indexing in a case-based reasoning system for waste management.

Wu, Z., & Palmer, M. (1994, 1994). *Verbs semantics and lexical selection*.

Zhang, Y., Zhang, S., & Leake, D. (2017, 2017). *Maintenance for Case Streams: A Streaming Approach to Competence-Based Deletion*.

Zia, T., Bruckner, D., & Zaidi, A. (2011, 2011). *A hidden markov model based procedure for identifying household electric loads*.

Zoha, A., Gluhak, A., Imran, M. A., & Rajasegarar, S. (2012). Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, *12*(12), 16838-16866.