

**EVALUACIÓN DE ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN  
PARA EL DIAGNÓSTICO DE HUECOS DE TENSIÓN**

**RONALD AUGUSTO VELANDIA ORTEGA  
FREDY LEONARDO HERNÁNDEZ SUÁREZ**

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

**2010**

**EVALUACIÓN DE ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN  
PARA EL DIAGNÓSTICO DE HUECOS DE TENSIÓN**

**RONALD AUGUSTO VELANDIA ORTEGA  
FREDY LEONARDO HERNÁNDEZ SUÁREZ**

**Trabajo de grado presentado como requisito para optar al título de  
INGENIERO ELECTRICISTA**

**Director**

**Hermann Raúl Vargas Torres**

**Codirector**

**Víctor Augusto Barrera Núñez**

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

**2010**

## **AGRADECIMIENTOS**

Los autores expresan sus más sinceros agradecimientos a todos los profesores de la Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones de la Universidad Industrial de Santander por el conocimiento transmitido durante la carrera universitaria, en particular al profesor Hermann Raúl Vargas Torres por su compromiso y orientación durante el desarrollo de este proyecto y al profesor Víctor Augusto Barrera Núñez por todo el apoyo y orientación brindada para hacer posible la realización de este proyecto.

También especial agradecimiento a sus padres, amigos y compañeros, soporte invaluable durante el transcurso de este proceso.

## **DEDICATORIA**

A Dios Padre Todo Poderoso por ser mi guía en todo momento, a mis padres Hermes Alberto Velandia y Felsomina Ortega por toda la ayuda brindada y su entrega cada día, a mi hermana Laura Maritza, mi sobrino Camilo Alberto, su padre Edward, y a mi novia Nathalí Ramírez único gran amor incondicional en mi vida. Ellos son la hermosa familia y grandes regalos que Dios puso en mi camino.

Ronald Augusto Velandia Ortega.

A mis padres, y a mis hermanos que siempre han estado conmigo en cada instante de mí existir:

Miguel, Edy, Sandra y Mitu

Fredy Leonardo Hernández Suárez.

## TABLA DE CONTENIDO

	<b>Pág.</b>
INTRODUCCIÓN	24
1. GENERALIDADES	27
1.1 ANTECEDENTES	27
1.2 JUSTIFICACIÓN	28
1.3 PLANTEAMIENTO DEL PROBLEMA	29
1.4 OBJETIVOS DEL PROYECTO	30
2. REGLAS DE APRENDEZAJE	31
2.1 REGLAS DE ASOCIACIÓN Y DEPENDENCIA	31
2.2 SISTEMAS DE MINERÍA DE DATOS Y REGLAS DE DECISIÓN	32
2.2.1 Herramienta computacional Orange Canvas	33
2.2.2 Herramienta computacional WEKA	33
2.4 ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN	35
2.4.1 Algoritmo PRISM	35
2.4.2 Algoritmo 1R	36
2.4.3 Algoritmo AQ	36
2.4.4 Algoritmo CN2	37
2.4.5 Algoritmo RIPPER	38
2.4.6 Algoritmo APRIORI	39
2.4.7 Algoritmo CONJUNCTIVE RULE	40
2.4.8 Algoritmo RIDOR ( <i>Ripple Down Rule</i> )	40
3. DEFINICIÓN Y SELECCIÓN DE DESCRIPTORES PARA EL DIAGNOSTICO DE EVENTOS DE TENSIÓN	41
3.1 LOCALIZACIÓN RELATIVA U ORIGEN DE LOS EVENTOS DE TENSIÓN	41
3.1.1 Cambio de fase en la corriente de secuencia positiva PCSC	41
3.1.2 Relación entre la magnitud de la corriente de falla y pre-falla de secuencia positiva ( <i>I1ratio</i> )	42

3.1.3 Signo de la resistencia equivalente durante la perturbación ( $R_{ex}$ $R_{ey}$ $R_e$ )	42
3.1.4 Cambio de la magnitud y fase ( $Z_{ratio}$ y $AngleZ_{ag}$ ) de la impedancia vista desde el alimentador antes y durante el evento.	43
3.1.5 Componente Real de la Corriente (RCC)	44
3.2 DESCRIPTORES QUE CARACTERIZAN LA CAUSA DE LOS EVENTOS DE TENSIÓN	45
3.2.1 Relación entre la potencia aparente antes y después del evento ( $S_{ratio}$ )	45
3.2.2 Relación de tensión ( $V_{ratio}$ ) de corriente ( $I_{ratio}$ )	46
3.2.3 Máxima desviación estándar de los valores de tensión y corriente eficaz ( $\sigma_{maxV}$ y $\sigma_{maxI}$ )	48
3.2.4 Pendientes de tensión y de corriente durante el inicio del evento ( $PV_{max}$ y $PI_{max}$ )	49
3.2.5 Desviación estándar Pendientes de tensión y desviación estándar pendientes de corriente durante el inicio del evento ( $\sigma_{PV}$ y $\sigma_{PI}$ )	50
3.2.6 Comparación de Fases ( $C_{fasesV}$ , $C_{fasesI}$ )	51
3.2.7 Voltaje Complejo Remanente RCV y Factor Positivo - Negativo PNF	52
3.2.8 Máxima corriente y tensión de neutro ( $I_{nratio}$ y $V_{nratio}$ )	53
3.2.9 Ángulo de inserción de la falla (FIPA)	54
3.2.10 Cambio del ángulo de fase entre la tensión y la corriente ( $\theta_{VI}$ y $\theta_{VI\_cruce}$ )	54
3.2.11 Relación entre los factores de potencia ( $FPratio_{FFT}$ , $FPratio_{P\&S}$ , $FPratio_{cruce}$ )	55
3.2.12 Cambio en la potencia reactiva ( $\Delta Q$ )	57
3.2.13 Magnitud de la impedancia de falla ( $Z_{falla}$ )	58
3.2.14 Máxima tensión de arco ( $V_{arc_{max}}$ )	59
3.2.15 Frecuencia más alta de oscilación durante el evento ( $HighOscFrec$ )	60
3.2.16 Ganancia de tensión de secuencia cero ( $GV0$ ) y máxima tensión de secuencia cero ( $MaxV0$ )	61
3.2.17 Magnitud del segundo armónico de corriente $I_2$	62
3.2.18 Numero de segmentos de transición ( $Segmentos_{\delta=0.1, m=128}$ y $Segmentos_{\delta=8, m=64}$ )	62

3.2.19 <i>Coeficiente de Forma (CF)</i>	64
3.2.20 <i>Desviación de los Estados (<math>\sigma E</math>)</i>	65
3.2.21 <i>Relación Angular de las Longitudes (<math>\delta L</math>)</i>	67
3.2.22 <i>Desviación de las Longitudes Eléctricas (<math>\sigma \delta L</math>)</i>	69
3.2.23 <i>Delta de Longitudes (<math>\Delta LV</math> y <math>\Delta LI</math>)</i>	70
3.2.24 <i>Pérdida de Tensión (<math>L_v</math>, <math>L_{v3}</math>)</i>	71
4. ANALISIS DE LOS DESCRIPTORES MEDIANTE ESTADÍSTICA MULTIVARIANTE	73
4.1 <i>ANÁLISIS MULTIVARIANTE DE LA VARIANZA – MANOVA</i>	73
4.2 <i>IDENTIFICACIÓN DE LOS DESCRIPTORES RELEVANTES DE ACUERDO A LA CAUSA DE LOS EVENTOS DE TENSIÓN.</i>	76
4.3 <i>CLASIFICACIÓN DE LOS DESCRIPTORES SEGÚN EL ORIGEN DE LOS EVENTOS DE TENSIÓN</i>	85
5. EXTRACCIÓN DE REGLAS DE DECISIÓN	86
5.1 <i>ALGORITMOS DE LA HERRAMIENTA COMPUTACIONAL WEKA</i>	86
5.2 <i>ALGORITMOS DE LA HERRAMIENTA COMPUTACIONAL Orange Canvas</i>	89
5.3 <i>RESULTADOS DE LOS ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN</i>	90
5.4 <i>VALIDACIÓN DE ALGUNAS REGLAS QUE DIAGNOSTICAN LA CAUSA DE LOS EVENTOS DE TENSIÓN</i>	103
6. CONCLUSIONES	107
BIBLIOGRAFÍA	110
ANEXOS	113
A.2 Algoritmos de aprendizaje de extracción de reglas de decisión	113

## LISTA DE FIGURAS

	<b>Pág.</b>
Figura 1. Árbol de decisión para recomendar o no una cirugía ocular.....	34
Figura 2. Reglas de decisión que corresponden al árbol de la Figura2 .....	34
Figura 3. Sistema de potencia interconectado .....	43
Figura 4. Sistema de dos fuentes para el análisis de la localización de un hueco de tensión .....	44
Figura 5. Formas de onda ante el arranque de un motor de inducción.....	46
Figura 6. Formas de ondas debidas a la energización de un banco de capacitores a) Tensión RMS b) Corriente RMS .....	47
Figura 7. Desviación estándar entre las tensiones eficaces de un evento multiestado.....	48
Figura 8. a) Estimación de Pendientes de la señal de tensión RMS b) Estimación de pendientes de señal de corriente durante el primer estado transitorio en una falla monofásica .....	49
Figura 9. a) Hueco de tensión trifásico b) Estado transitorio durante la caída.....	50
Figura 10. Deltas de tensión para un hueco monofásico .....	51
Figura 11. Estimación del ángulo entre la tensión y la corriente ( $V_{Base} = 25Kv$ , $I_{Base} = 54A$ ) .....	55
Figura 12. Comportamiento de la potencia reactiva antes y después de la energización de bancos de capacitores .....	58
Figura 13. Tensión de arco en los estados de prefalla, falla y posfalla de un evento monofásico.....	59
Figura 14. Descomposición wavelet de la señal de voltaje en dos escalas de frecuencia .....	60
Figura 15. Tensión de secuencia cero en un evento de tensión multiestado.....	62
Figura 16. Segmentos de transición de un evento causado por la energización de bancos de capacitores. El evento contiene dos (2) segmentos de transición o un (1) estado transitorio .....	63

Figura 17. Segmentos de transición de un evento monoestado .....	64
Figura 18. Hueco de Tensión con la zona de operación del descriptor CF.....	65
Figura 19. Hueco de tensión Multiestado.....	66
Figura 20. Metodología para el cálculo $\delta L$ .....	67
Figura 21. Triangulo de longitudes eléctricas .....	68
Figura 22. Hueco de tensión originado por una falla de red trifásica .....	70
Figura 23. Hueco de tensión monofásico con oscilación de alta frecuencia durante el despeje de la falla .....	82
Figura 24. Entorno WEKA en la aplicación Explorer.....	87
Figura 25. Tres reglas de decisión arrojadas por el algoritmo JRip para la clase eventos desbalanceados y balanceados. ....	88
Figura 26. Matriz de confusión para las reglas del algoritmo JRip, evento desbalanceado y balanceado .....	89
Figura 27. Dos reglas de decisión arrojadas por el algoritmo CN2 para el escenario de eventos desbalanceados y balanceados. ....	90
Figura 28. Pasos propuestos para el diagnóstico de los eventos que involucran el grado de desbalance .....	97
Figura 29. Pasos propuestos para el diagnóstico de la causa para eventos multiestado y monoestado .....	99
Figura 30. Pasos propuestos para el diagnóstico de la localización relativa .....	102
Figura A.1. Algoritmo de aprendizaje de reglas por "Cobertura.....	113
Figura A.2. Algoritmo de PRISM.....	114
Figura A.3. Algoritmo de 1R.....	115
Figura A.4. Algoritmo de AQ .....	116
Figura A.5. Algoritmo de APRIORI.....	117
Figura A.6. Función Apriori Genera .....	118
Figura A.7. Función Reglas de Asociación .....	118
Figura A.8. Función Genera Reglas.....	118
Figura B.1. Hueco de tensión a un 50% de la amplitud de la señal.....	119
Figura B.2. Características de un hueco de tensión .....	119

Figura C.1.Sistema modelado en ATP.....	123
Figura C.2. a) Falla bifásica. b) Energización de banco de capacitores. c) Arranque de motor de inducción.....	127

## LISTA DE TABLAS

	<b>Pág.</b>
Tabla 1. Resultados MANOVA según el grado de desbalance del evento de tensión .....	78
Tabla 2. Resultados MANOVA para eventos desbalanceados .....	80
Tabla 3. Resultados MANOVA para eventos balanceados .....	83
Tabla 4. Resultados MANOVA para eventos monoestado y multiestado .....	84
Tabla 5. Resultado MANOVA según la localización relativa de los eventos de tensión .....	85
Tabla 6. Resultados de los algoritmos de extracción de reglas para la clase: eventos desbalanceados y balanceados .....	92
Tabla 7. Resultados de los algoritmos de extracción de reglas para la clase: eventos desbalanceados .....	93
Tabla 8. Resultados de los algoritmos de extracción de reglas para la clase: eventos balanceados .....	95
Tabla 9. Resultados de los algoritmos de extracción de reglas para la clase: eventos monoestado y multiestado .....	98
Tabla 10. Resultados de los algoritmos de extracción de reglas para la clase: localización aguas arriba y aguas abajo .....	100
Tabla 11. Resultados de la clasificación obtenida por algunas reglas que diagnostican eventos de tensión .....	103
Tabla 12. Resultados de la clasificación obtenida por algunas reglas que identifican eventos monoestado y multiestado .....	105
Tabla 13. Resultados de la clasificación obtenida por algunas reglas que diagnostican la localización relativa .....	106
Tabla B.1. Clasificación general de los huecos de tensión .....	121
Tabla C.1. Configuraciones de las líneas de transmisión eléctrica .....	123
Tabla C.2. Configuración de la línea de transmisión .....	124
Tabla C.3. Especificaciones de los transformadores .....	126

Tabla C.4. Cargas concentradas .....	126
Tabla C.5. Cargas distribuidas.....	126
Tabla C.6. Especificaciones de los capacitores .....	127

## LISTA DE ANEXOS

Pág.

A. ALGORITMOS DE APRENDIZAJE .....	113
A.1 Algoritmo de aprendizaje de reglas por cobertura .....	113
A.2.1 Algoritmo de aprendizaje de reglas PRIMS .....	113
A.2.2 Algoritmo de aprendizaje de reglas 1R.....	115
A.2.3 Algoritmo de aprendizaje de reglas AQ .....	116
A.2.4 Algoritmo de aprendizaje de reglas APRIORI.....	117
B. HUECOS DE TENSIÓN .....	119
B.1 Clasificación de los huecos de tensión.....	120
C. SIMULACIÓN DE HUECOS DE TENSIÓN PARA VALIDAR LA BASE DE DATOS SUMINISTRADA .....	122
C.1 Simulación en ATP del sistema de potencia IEEE de 34 nodos.....	122

## NOMECLATURA

<b>UIS</b>	Universidad Industrial de Santander
<b>UTP</b>	Universidad Tecnológica de Pereira
<b>UDG</b>	Universitat de Girona
<b>SPSS</b>	Statistical Package for the Social Sciences
<b>IEEE</b>	Institute for Electric and Electronic Engineer
<b>ORANGE</b>	Herramienta para la extracción de reglas de decisión
<b>CN2</b>	Algoritmo propuesto por Clark Niblett
<b>ATP</b>	Software para la simulación de transitorios electromagnéticos.
<b>ANOVA</b>	Analysis of Variance
<b>MANOVA</b>	Multivariate Analysis of Variance.
<b>RIPPER</b>	Repeated Incremental Pruning to Produce Error Reduction
<b>PQM</b>	Power Quality Monitor – Medidor de la calidad de potencia
<b>PCSC</b>	Cambio de fase en la corriente de secuencia positiva
<b><math>I_{1ratio}</math></b>	Relación entre la magnitud de la corriente de falla y prefalla de secuencia positiva
<b><math>Re</math></b>	Resistencia equivalente durante el evento
<b><math>Re_x</math></b>	Parte real de la impedancia equivalente durante el evento
<b><math>Re_y</math></b>	Parte imaginaria de la impedancia equivalente durante el evento
<b><math>Z_{ratio}</math></b>	Relación entre las magnitudes de la impedancia de prefalla y posfalla

<b><i>Angle<sub>zag</sub></i></b>	Ángulo de la impedancia durante la falla
<b><i>RCC</i></b>	Componente Real de la Corriente
<b><i>Sratio</i></b>	Relación entre la potencia aparente de posfalla y prefalla
<b><i>Vratio</i></b>	Relación entre la tensión de posfalla y prefalla
<b><i>Iratio</i></b>	Relación entre la corriente de posfalla y prefalla
<b><i><math>\sigma_{max}V</math></i></b>	Máxima desviación estándar de los valores de tensión
<b><i><math>\sigma_{max}I</math></i></b>	Máxima desviación estándar de los valores de corriente
<b><i>PVmax</i></b>	Máxima pendiente de tensión durante el inicio del evento
<b><i>PImax</i></b>	Máxima pendiente de corriente durante el inicio del evento
<b><i><math>\sigma PV</math></i></b>	Desviación estándar pendientes de tensión
<b><i><math>\sigma PI</math></i></b>	Desviación estándar pendientes de corriente
<b><i><math>C_{fases}V</math></i></b>	Comparación de fases de tensión
<b><i><math>C_{fases}I</math></i></b>	Comparación de fases de corriente
<b><i>RCV</i></b>	Voltaje complejo remanente
<b><i>PNF</i></b>	Factor positivo - negativo
<b><i>Vnratio</i></b>	Relación entre la tensión de neutro de falla y prefalla
<b><i>Inratio</i></b>	Relación entre la corriente de neutro de falla y prefalla
<b><i>FIPA</i></b>	Ángulo de inserción de la fase fallada
<b><i><math>\theta_{VI}</math></i></b>	Cambio en el ángulo de la tensión y la corriente
<b><i><math>\theta_{VI\_cruce}</math></i></b>	Cambio en el ángulo de la tensión y la corriente utilizando el cruce por cero de las señales

<b><math>FPratio_{FFT}</math></b>	Relación entre los factores de potencia antes y durante el evento utilizando la transformada rápida de Fourier
<b><math>FPratio_{P\&amp;S}</math></b>	Relación entre los factores de potencia antes y durante el evento utilizando las potencias activa y aparente
<b><math>FPratio_{cruce}</math></b>	Relación entre los factores de potencia antes y durante el evento utilizando el cruce por cero de las señales.
<b><math>\Delta Q</math></b>	Diferencia entre la potencia reactiva de prefalla y la potencia reactiva de posfalla
<b><math>Zfalla</math></b>	Magnitud de la impedancia de falla
<b><math>Varc_{max}</math></b>	Máximo arco de tensión
<b><math>HighOscFrec</math></b>	Frecuencia más alta de oscilación durante el evento
<b><math>GV_0</math></b>	Ganancia de tensión de secuencia cero
<b><math>MaxV_0</math></b>	Máxima tensión de secuencia cero
<b><math> I_2 </math></b>	Magnitud del segundo armónico de corriente
<b><math>CF</math></b>	Coeficiente de forma
<b><math>\sigma E</math></b>	Desviación de los estados
<b><math>\delta L</math></b>	Relación angular de las longitudes
<b><math>\sigma \delta L</math></b>	Desviación de las longitudes eléctricas
<b><math>\Delta LV</math></b>	Delta de longitudes de la señal de tensión
<b><math>\Delta LI</math></b>	Delta de longitudes de la señal de corriente
<b><math>Lv</math></b>	Pérdida de tensión basada en minRCV
<b><math>Lv3</math></b>	Pérdida de tensión trifásica

***minRCV***                      Mínima tensión compleja remanente

***minPNF***                      Mínimo factor positivo-negativo

***Segmentos* <sub>$\delta=0.1, m=128$</sub>**     Cantidad de segmentos transitorios con umbral  $\delta=0.1\%$   
y rata de submuestreo  $m=128$  muestras

***Segmentos* <sub>$\delta=8, m=64$</sub>**         Cantidad de segmentos transitorios con umbral  $\delta=8\%$  y  
rata de submuestreo  $m=64$  muestras

## RESUMEN

### TÍTULO

EVALUACIÓN DE ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN PARA EL DIAGNÓSTICO DE HUECOS DE TENSIÓN.\*

### AUTORES

Velandia Ortega Ronald Augusto, Hernández Suárez Fredy Leonardo.\*\*

### PALABRAS CLAVES

Análisis Multivariable, Calidad de Energía, Huecos de Tensión, Técnicas para la Extracción de Conocimiento.

### DESCRIPCIÓN

Este trabajo de grado ofrece nuevas herramientas para el análisis de fallas. Se describe una metodología que mediante el uso de varios algoritmos que emplean diferentes descriptores propuestos, permiten estimar la causa y el origen de huecos de tensión mediante la extracción de reglas de decisión. El análisis previo a la elaboración de la metodología consiste en una clasificación detallada de los registros suministrados por la empresa de energía eléctrica de Barcelona (ENDESA). A partir de los registros proporcionados por ENDESA, se realiza una formulación de nuevos descriptores planteados con el propósito de extraer una mayor información sobre los eventos de tensión. Este conjunto de descriptores se complementa con otros ya propuestos en la literatura y de esta forma se tiene un gran número de datos de entrada para el respectivo análisis de extracción de reglas de decisión, mostrando resultados con mayor precisión.

Se hace el análisis estadístico multivariante a los descriptores planteados para verificar la existencia de grupos o clases de acuerdo a cada una de las causas y el origen de los eventos de tensión. Con estos descriptores y usándolos en los algoritmos de extracción de reglas de conocimiento y se comparan los resultados para establecer falencias y debilidades en el diagnóstico de los huecos de tensión.

\* Proyecto de Grado: Trabajo de Investigación

\*\* Facultad de Ingenierías Físico-mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director Hermann Raúl Vargas Torres. Codirector Víctor Augusto Barrera Núñez.

## SUMMARY

### TITLE

EVALUATION OF ALGORITHMS OF EXTRACTION OF DECISION RULES FOR VOLTAGE SAGS DIAGNOSE.\*

### AUTHORS

Velandia Ortega, Ronald Augusto. Hernández Suárez, Fredy Leonardo.\*\*

### KEY WORDS

Multivariate Statistical Analysis, Power Quality, Voltage Sags (dips), Extraction Knowledge Techniques.

### DESCRIPTION

This document offers new tools for the faults analysis. A methodology is described using various algorithms to get different propose features, it allows to estimate of the cause and the origin of voltage events through the extraction of decision rules. The previous analysis to the elaboration of the methodology consists of a detailed classification of the registries provided by the distribution company of Barcelona (ENDESA). From the registries provided by ENDESA, a formulation of new description is made in order to extract the most important information of the voltage events. This set of description is complemented with other already proposed in literature and so have a big into number files for this analysis of extraction of decision rules, showing results with bigger precision than allow working in a more reliable methodology in order to estimate the location and cause in the voltage sags diagnose on the real power systems.

Doing multivariate statistical analysis is applied in order to verify the existence of groups or classes according to each one of the causes and the origin of associate voltage events. With this features and using their in the algorithms of extraction of decision rules and the results are compare to establish false results and weakness in the voltage sags diagnose.

\*Degree Project: Research Work.

\*\*Faculty of Physical-Mechanic Engineering. School of Electrical, Electronic and Telecommunications Engineering. Dir. Hermann Raúl Vargas Torres. Codir. Víctor Augusto Barrera Núñez.

## INTRODUCCIÓN

Es de gran importancia hoy en día el estudio de la calidad de la energía eléctrica, ya que de ella dependen los procesos industriales y el tiempo de vida útil de los dispositivos conectados a la red eléctrica.

La carga es sensible a cambios en la señal de tensión; de esta manera si ocurre un cambio en la amplitud de dicha señal, la calidad de la potencia es afectada y por lo tanto los equipos conectados al sistema de energía se encuentran sometidos a salidas de operación.

La aparición de cargas no lineales ha incrementado los niveles de distorsión armónica, los cuales pueden llevar al sistema a condiciones de resonancia, ocasionando un alto grado de sensibilidad ante la presencia de perturbaciones electromagnéticas a lo largo del sistema (huecos de tensión, subidas de tensión, picos de tensión, armónicos, etcétera).

La aparición de perturbaciones electromagnéticas produce gran impacto económico. Por esta razón las empresas realizan frecuentemente estudios para valorar los índices de calidad de energía, con la intención de reducir los efectos producidos por la aparición de tales perturbaciones en los sistemas; de allí la importancia del análisis y el diagnóstico de la calidad de la energía.

En los departamentos de operación de la red, dentro de las empresas de distribución de energía eléctrica, se realizan continuamente estudios de localización y causología.

Los estudios de localización tienen como objetivo determinar hacia donde está la fuente del evento (aguas arriba o aguas abajo) y su distancia con respecto al punto de medida (sólo para el caso aguas abajo).

Recibe el nombre de *localización relativa* cuando se intenta determinar la dirección de la fuente del evento (aguas arriba/abajo) y *localización precisa* cuando se intenta estimar la distancia con respecto al punto de registro.

Desde el punto de vista de estudios de localización de fallas, este trabajo se centra sólo en el análisis de la localización relativa del evento, debido a que se carecen de formas de onda cuya localización de la fuente sea conocida.

Los estudios de causología tienen como objetivo determinar cuál es la causa de la perturbación o conjunto de perturbaciones. Las causas de las perturbaciones eléctricas pueden ser internas o externas. Las causas internas son aquellas producidas por operación de elementos propios de la red, tales como, energización de transformadores, capacitores, arranque de motores, cambios de carga, etcétera; mientras que las causas externas están asociadas a perturbaciones generadas por agentes externos al sistema de energía, como lo son: descargas atmosféricas, contacto de las redes con animales y ramas de árboles, ruptura de empalmes, excavadoras, etcétera.

Para identificar la localización y la causa de la perturbación, se hace uso de seis algoritmos de extracción de reglas de decisión seleccionados, cuarenta y seis descriptores, algunos propuestos en este trabajo y otros extraídos de la literatura, son utilizados como elementos de entrada de los mencionados algoritmos de extracción de reglas. Estos descriptores son analizados (eléctricamente/estadísticamente) y validados.

Finalmente se demuestra que éstas pueden ser utilizadas en estudios que orientados al diagnóstico automático de la localización relativa de la fuente y de la causa de perturbaciones registradas en circuitos de distribución.

Consecuentemente, las empresas pueden tomar correctivos con el propósito de mejorar los índices de continuidad respecto a la duración y la frecuencia de las interrupciones, debido a que se identifican las fuentes de huecos de tensión, las cuales pueden ser intervenidas en el mantenimiento, evitando daños que conlleven a una eventual interrupción del servicio.

Se resalta como material valioso para el desarrollo del proyecto la información sobre huecos de tensión suministrados por la empresa de distribución de energía de Barcelona ENDESA, que forman un conjunto de perturbaciones registradas en 10 subestaciones de 25 kV.

Toda la información consignada en la base de datos de ENDESA fue procesada utilizando MATLAB, SPSS, WEKA y Orange Canvas. SPSS es un software para el análisis estadístico, mientras que WEKA y Orange Canvas son dos herramientas de libre distribución para el análisis de datos. A partir de éstos últimos y de su combinación con MATLAB se aplicaron los algoritmos de extracción de reglas.

## **1. GENERALIDADES**

### **1.1 ANTECEDENTES**

La calidad del suministro de energía eléctrica se ha convertido en un tema de sumo interés para empresas de energía eléctrica y usuarios.

Como consecuencia directa del nuevo mercado en energía eléctrica y del aumento en problemas relacionados con la calidad de la energía, las compañías de suministro y transmisión han puesto en marcha programas de monitorización de la red que permitirán capturar/registrar los fenómenos electromagnéticos que en ella ocurren.

El análisis de las perturbaciones electromagnéticas registradas, debe servir para determinar acciones que mejoren la calidad de energía en los sistemas de transmisión y distribución. Un resultado común de los programas de monitorización es agrupar disturbios en clases, a partir de los valores RMS mínimos o máximos de tensión y/o corriente, o simplemente comparando las formas de onda registradas con la ideal [1].

En importantes estudios acerca de la calidad de la energía, se ha tenido en cuenta el problema de los huecos de tensión, mediante el análisis de algunas causas y características más relevantes de dicho problema [2].

La gran cantidad de eventos registrados y almacenados por los sistemas de monitoreo (relés, protecciones, SCADA, etcétera) pueden ser utilizados para el diagnóstico de fallas.

Tales eventos contienen información útil sobre el comportamiento del sistema a lo largo de toda la perturbación, pero dada la complejidad de su análisis, generalmente sólo se extraen parámetros, llamados comúnmente descriptores. Los descriptores más comunes son, la magnitud y duración del hueco de tensión.

Existen metodologías útiles para encontrar la localización relativa y la causa del hueco de tensión basada en el uso de algunos descriptores que utilizan diferentes

formas de obtener características sobresalientes de las señales [3] y [4], así como estrategias utilizadas para la localización en sistemas de distribución [7]. Además de esto, estudios realizados acerca de desarrollos metodológicos para la detección de la causa del evento [23] y estrategias de segmentación para señales de huecos de tensión con componentes de alta frecuencia.

## **1.2 JUSTIFICACIÓN**

Este trabajo se encuentra justificado en la necesidad de determinar la localización relativa y causa de los eventos de tensión, ya que actualmente en Colombia las empresas distribuidoras de energía eléctrica están instalando equipos de monitorización de energía en cada uno de sus alimentadores. Tales equipos continuamente se encuentran registrando/capturando perturbaciones que se propagan a lo largo de la red eléctrica. Consecuentemente, la cantidad de eventos crece continuamente y las herramientas informáticas que permitan extraer información útil para la operación de la red a partir de los eventos registrados se convierten en un nuevo reto para la calidad de la energía eléctrica.

Igualmente, la utilización de equipos de monitorización ha propiciado la proliferación de las Redes Inteligentes o Smart Grids, en donde la cantidad de equipos de supervisión, control y almacenamiento juegan un papel importante, así como también, la extracción de información a partir de las señales medidas en tiempos cercanos al tiempo real.

De igual manera este trabajo contribuye al fortalecimiento del grupo de investigación (GISEL) en el estudio y análisis de huecos de tensión, donde no se han realizado hasta el momento otros trabajos basados en el análisis estadístico multivariante aplicado a este fenómeno.

Este trabajo sirve de apoyo al proyecto de investigación de doctorado “Monitorización de Huecos de Tensión en Sistemas de Distribución de Energía Eléctrica”, en el que se encuentran vinculadas entidades educativas como la

Universidad Industrial de Santander, la Universidad Tecnológica de Pereira y la Universitat de Girona en España.

Este proyecto en conjunto pretende promover el continuo desarrollo del tema de la calidad de la energía, en el que empresas del sector eléctrico como *ENDESA* Barcelona y *CODENSA* en Colombia se encuentran comprometidas.

### **1.3 PLANTEAMIENTO DEL PROBLEMA**

Las empresas del sector de transmisión y distribución de energía eléctrica se preocupan por cumplir ciertos criterios de disponibilidad y regulación, que dependen de la señal de tensión para asegurar la calidad y el suministro sin interrupciones y garantizar que la confiabilidad de los equipos conectados a la red eléctrica sea óptima en todo momento.

Durante el proceso de generación, transmisión y distribución la señal de tensión eventualmente se ve afectada por cambios en su amplitud y su frecuencia los cuales desvían de una onda senoidal pura de frecuencia 60Hz.

Por lo general en los sistemas de distribución, las fallas importantes siempre se conocen en el centro de control de la red, otras menos trascendentes son reportadas por los clientes cuyas quejas son remitidas a la empresa de energía eléctrica. Sin embargo, la localización relativa de una falla y su causa son desconocidas, por tanto, partiendo de las señales registradas por los equipos de monitorización y/o protección, es posible determinar la localización relativa y la causa del evento de tensión.

En este trabajo se proponen y analizan estadísticamente los descriptores obtenidos a partir de señales eléctricas reales (*ENDESA*) y simuladas (*ATP*). Luego, se demuestra que éstos descriptores pueden ser utilizados en herramientas informáticas para el diagnóstico (localización relativa y causa) de

eventos de tensión causados por energización de transformadores, bancos de capacitores, arranque de motores y/o fallas de la red.

#### **1.4 OBJETIVOS DEL PROYECTO**

Aplicar distintos algoritmos de extracción de reglas de decisión en el diagnóstico de huecos de tensión.

Para cumplir este objetivo se establecieron los siguientes objetivos específicos:

- a) Estudiar y seleccionar algoritmos de extracción de reglas encontradas en la literatura.
- b) Seleccionar descriptores que permitan caracterizar el origen y la causa de los huecos de tensión.
- c) Aplicar las reglas seleccionadas en la estimación del origen y causa de los huecos de tensión, utilizando los descriptores propuestos.
- d) Comparar las reglas de decisión obtenidas y establecer sus falencias y debilidades en diagnósticos de huecos de tensión.

## 2. REGLAS DE APRENDIZAJE

La necesidad de metodologías para el análisis inteligente de datos, que permiten descubrir un conocimiento útil a partir de una base de datos, es suministrado por las reglas de aprendizaje. Estas son definidas como un proceso no trivial que identifica patrones de un conjunto de datos, con las siguientes características: validez y utilidad. Este proceso es un conjunto de pasos iterativos entre los que se incluye el pre-procesamiento de los datos para corregir los posibles datos erróneos, incompletos e inconsistentes.

### 2.1 REGLAS DE ASOCIACIÓN Y DEPENDENCIA

Las reglas de asociación son la manera de expresar modelos específicos o patrones de datos de una base de datos cualquiera. “Estos patrones pueden servir para conocer el comportamiento general del problema que genera la base de datos, y de esta manera, se tenga más información que pueda asistir en la toma de decisiones” [8].

Básicamente el aprendizaje de las reglas de asociación se basa en su cobertura (soporte) que es el número de instancias que la regla predice correctamente y la confianza (precisión) que es la medida en porcentaje de veces que la regla cumple cuando se puede aplicar. En este caso los algoritmos de aprendizaje trabajan en la búsqueda de reglas que cumplan unos requisitos mínimos en estas medidas [8].

“Dado el gran volumen de datos de los problemas con el que los algoritmos de aprendizaje de reglas de asociación trabajan, la tarea de buscar patrones que cumplan estos requisitos pueden parecer muy tediosa, ya que los conjuntos de observaciones a ser analizados crece exponencialmente con respecto al número de variables de los datos” [8]. Sin embargo, en casos reales existen realmente pocos conjuntos frecuentes y los métodos que exigen una confianza y/o cobertura mínimas se benefician de este hecho [8].

Un reconocido algoritmo de aprendizaje de reglas de asociación es el algoritmo *Apriori*, el cuál basa su funcionamiento en la búsqueda de los conjuntos de

observaciones con una determinada cobertura. En primer lugar este algoritmo construye los conjuntos formados por una sola observación que supera la cobertura mínima. Este conjunto de conjuntos se usa para luego construir el conjunto de conjunto de dos observaciones y así sucesivamente hasta llegar a un tamaño en el cual no existan conjuntos de dichas observaciones con la cobertura requerida. Por tanto una vez se ha seleccionado los conjuntos de ítems que cumplen con la cobertura mínima, el algoritmo extrae las reglas de estos conjuntos de reglas que tengan un nivel de confianza mínimo [8].

Por tanto, el aprendizaje de reglas de asociación se divide en dos etapas: la extracción de los conjuntos de ítems (las observaciones mencionadas) que cumplan con la cobertura requerida desde los datos y la generación de reglas a partir de estos conjuntos de datos.

Un ejemplo de regla de asociación sería:

***SI Día nublado Y Vientos fuertes ENTONCES Lloverá***

El uso de reglas de asociación y dependencia representan una potente herramienta para la búsqueda de diferentes patrones en una base de datos. Dichas reglas son muy parecidas a las reglas de clasificación, con la diferencia que no hay una clase definida, ya que cualquier atributo puede estar en la parte derecha de una regla de asociación. Además este cambio no afectará al funcionamiento del algoritmo de aprendizaje de reglas de asociación, excepto en la eficiencia, ya que el número de observaciones a considerar crece, por lo que también sería válido para el aprendizaje de este tipo [8].

## **2.2 SISTEMAS DE MINERÍA DE DATOS Y REGLAS DE DECISIÓN**

Existen plataformas informáticas para la implementación de algoritmos de minería de datos. En concreto se utilizaron las herramientas *Orange Canvas* y *WEKA* para la extracción de reglas de decisión en el diagnóstico de huecos de eventos

de tensión. En ellas se encuentran implementados varios algoritmos para la extracción de reglas de decisión.

### **2.2.1 Herramienta computacional Orange Canvas**

Un recurso computacional importante para la extracción de reglas de decisión es el Orange Canvas. Es un programa informático para realizar minería de datos y análisis predictivo de la información. Consta de una serie de componentes desarrollados en C++ que implementan algoritmos de minería de datos, así como operaciones de pre-procesamiento y representación gráfica de datos.

Orange Canvas utiliza el algoritmo CN2, basado en la regla de aprendizaje CN2 para la extracción de reglas de conocimiento.

### **2.2.2 Herramienta computacional WEKA**

WEKA es un conjunto de librerías en lenguaje de programación Java que permite aplicar algoritmos para extraer conocimiento desde bases de datos.

Esta aplicación computacional posee varios algoritmos que extraen diferentes reglas de decisión de una base de datos. Algunos de estos algoritmos basados en reglas de aprendizaje son: JRip, Ridor, Conjunctive Rule, entre otros.

## **2.3 SISTEMAS DE APRENDIZAJE DE REGLAS POR COBERTURA**

Los sistemas de aprendizaje basados en árboles de decisión son unos de los métodos más fáciles de utilizar y entender.

“Un árbol de decisión es un conjunto de condiciones organizadas en una estructura jerárquica, de tal manera que la decisión final a tomar se puede determinar siguiendo las condiciones que se cumplen desde la raíz del árbol hasta alguna de sus hojas” [8].

Los árboles de decisión tienen ventaja de que las posibles opciones a partir de una determinada condición son excluyentes; permitiendo el análisis a una situación y siguiendo el árbol de decisión apropiadamente para llegar a una decisión o acción a tomar.

En la Figura 1 se observa el árbol de decisión para resolver el problema de realizar o no una cirugía ocular.

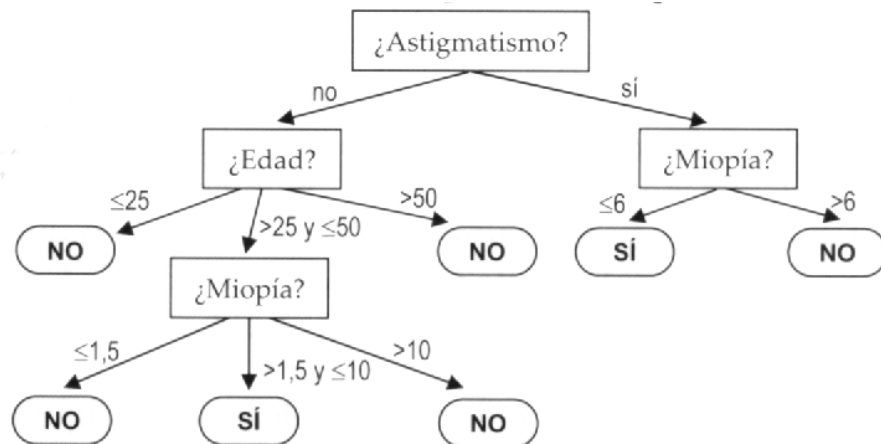


Figura 1. Árbol de decisión para recomendar o no una cirugía ocular

Los sistemas de reglas son una forma general de los árboles de decisión, un árbol de decisión se puede expresar como un conjunto de reglas.

A continuación en la Figura 2 se muestra el ejemplo de la cirugía ocular del árbol de decisión anterior expresado en forma de reglas.

**¿Operación?**

**SI Astig.=No Y 25<Edad≤50 Y 1.5<Miopía≤10 ENTONCES Sí**

**SI Astig.=Sí Y Miopía≤6 ENTONCES Sí**

**EN OTRO CASO NO**

Figura 2. Reglas de decisión que corresponden al árbol de la Figura2.

Este conjunto de reglas se derivan de particiones, en las cuales para cualquier condición siempre aparece además la o las condiciones complementarias. Existen en cambio muchos conjuntos de reglas que no cumplen estas condiciones y sin embargo son capaces de clasificar la evidencia de una manera conveniente.

Existen métodos que generan conjuntos de reglas que podrían ser contradictorias para algunos ejemplos en particular. La manera de resolver este problema es dar un orden a las reglas, denominado *lista de decisión*.

Estos métodos van añadiendo reglas, una detrás de otra, mientras vayan cubriendo ejemplos de una manera consistente, descartando los ejemplos ya cubiertos por las reglas ya obtenidas y con los ejemplos que quedan se empieza de nuevo [8]. Esto hace que puedan aparecer nuevas condiciones que solapen o no con las anteriores condiciones. Esta es la forma en la que actúan los métodos de cobertura (ver Anexo A.1).

## **2.4 ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN**

En la literatura actual se mencionan diferentes algoritmos de extracción de reglas de decisión usadas por diferentes programas de tipo estadístico y de extracción de conocimiento. A continuación se mencionan las diferentes estructuras que manejan distintos algoritmos de extracción de conocimiento, algunos de ellos son empleados por las herramientas computacionales que extraen las reglas que se usarán para diagnosticar el origen y las causas de los huecos de tensión.

A continuación se presentan algunos algoritmos usados por los entornos *WEKA* y *Orange Canvas* para la extracción de reglas de decisión.

### **2.4.1 Algoritmo PRISM**

Es un algoritmo básico de aprendizaje de reglas que asume que no hay ruido en los datos, es decir que toda la información no presenta distorsión en los datos clasificados [10]. Ver Anexo A.2.1.

Una de sus principales ventajas es la no consideración de ruido en los datos y crea reglas que cubre la mayor parte de las observaciones, separando las instancias para analizarlas por separado y cumplir su cometido.

### 2.4.2 Algoritmo 1R

El principio de este algoritmo es hacer reglas que prueban un solo par atributo-valor. Se prueban todos los pares atributo-valor y se selecciona el que ocasione el menor número de errores [10]. Ver Anexo A.2.2.

### 2.4.3 Algoritmo AQ

Uno de los primeros sistemas de reglas fue AQ, desarrollado originalmente por Michalski en 1979, ahora re-implementado y mejorado por otros autores, AQ11 y AQ15.

Su salida es un conjunto de reglas de clasificación del tipo “*if...then...*”. Esto es útil en aplicaciones de sistemas expertos basadas en reglas y normalmente son más fáciles de entender que los árboles de decisión.

El algoritmo de AQ, induce un conjunto de reglas de decisión, una para cada clase. Cada regla es de la forma:

$$if \langle cover \rangle then \langle clase \rangle , donde: \langle cover \rangle$$

En donde esta regla es una combinación Booleana de pruebas de descriptores.

Así los descriptores tienen su propio nombre:

*Selector*. Es la prueba básica de un atributo. Ejemplo:

$$(sangre = caliente)$$

*Complex*. Es el nombre que recibe una conjunción de *selectores*. Ejemplo:

$$(sangre = caliente) \& (pelo = si)$$

*Cover*. Es una disyunción de *complexes*. Ejemplo:

$$(sangre = caliente) \& (pelo = si) \text{ OR } (leche = si)$$

Un *complex* vacío cubre todos los ejemplos, mientras que un *cover* vacío no cubre ningún ejemplo. Un *cover* es almacenado junto con el valor de la clase asociada. Ejemplo:

Si (*sangre = caliente & pelo = si*) OR (*leche = si*) Entonces (*clase = mamífero*)

El algoritmo AQ selecciona un ejemplo de una clase (positivo) y toma *selectores* de éste, hasta que no se cubra ningún ejemplo de cualquier otra clase (negativo), para formar un *complex*. Cada *complex* encontrado es añadido al *cover* hasta que se cubran todos los ejemplos positivos [11]. Ver Anexo A.2.3.

#### **2.4.4 Algoritmo CN2**

Propuesto en 1988 por Clark Niblett , con la idea de analizar datos con ruido y evitar el sobre ajuste que se encontraba en sistemas como AQ [12].

Su contribución principal es eliminar la dependencia de un ejemplo específico durante su búsqueda y forma la base de muchos de los algoritmos de reglas actuales.

La heurística de búsqueda original que sigue es basada en entropía:

$$Entr = - \sum_i p_i \log_2 (p_i) \quad (2.1)$$

Donde  $p_i$  es la distribución de las clases que cubre cada regla.

Se selecciona la regla de menor entropía, es decir la regla que cubre muchos ejemplos de primera clase y pocos de las demás clases.

CN2 maneja también una medida de significancia para las reglas. El usuario proporciona un límite para la medida de significancia, abajo del cual las reglas son rechazadas.

La medida está basada en la razón de verosimilitud (*likelihood ratio statistic*) que mide una distancia entre dos distribuciones, definida por:

$$2 \sum_{i=1}^n f_i \log \left( \frac{f_i}{e_i} \right) \quad (2.2)$$

Donde:

- $F = (f_1, \dots, f_n)$  es la distribución de frecuencias observada de ejemplos dentro de las clases que satisfacen una regla dada (número de observaciones que satisfacen la regla entre el número total de ejemplos que satisface la regla).
- $E = (e_1, \dots, e_n)$  es la frecuencia esperada del mismo número de ejemplos bajo la suposición de que la regla selecciona ejemplos aleatoriamente (número de ejemplos cubiertos por la regla siguiendo la distribución de ejemplos del total de los ejemplos).
- Entre más baja es la medida es más probable que la aparente regularidad expresada en la regla sea por casualidad.

Estas medidas, entropía y significancia determinan que reglas son buenas (tienen alta precisión cuando predicen la clase mayoritaria cubierta) y confiables (la alta precisión sobre datos de entrenamiento no se debe solo a casualidad) [12].

#### **2.4.5 Algoritmo RIPPER**

RIPPER (Repeated Incremental Pruning to Produce Error Reduction) utiliza varias medidas e ideas al mismo tiempo. Construye un conjunto de reglas usando *covering*, las reduce usando la técnica heurística (Ecuación 2.1) con un conjunto de reglas de entrenamiento por separado, y luego “optimiza” al mismo tiempo ese conjunto de reglas.

RIPPER utiliza un conjunto de ejemplos separados para decidir podar reglas (eliminar descriptores de poco peso contenidos en la conjunción de cada regla), utiliza ganancia de información para crecer las reglas, la medida para podar reglas, esto es una medida basada en un conjunto global de reglas.

Una vez que construye un conjunto inicial de reglas, toma una regla  $R_i$  del conjunto total de reglas y la hace crecer (revisión) y también hace crecer una

nueva regla desde el principio. Al final se queda con la regla original o alguna de las otras dos (la que hizo crecer o construyo desde cero) pero tomando en cuenta el error sobre el conjunto total de las reglas [13] y [14].

#### **2.4.6 Algoritmo APRIORI**

Este algoritmo se basa en el conocimiento previo o “a priori” de los conjuntos de datos que aparecen con mayor frecuencia, esto sirve para reducir el espacio de búsqueda y aumentar la eficiencia.

El algoritmo Apriori se usa en minería de datos para encontrar reglas de asociación para encontrar correlaciones entre las observaciones de un conjunto de datos.

El proceso del algoritmo se resume en dos pasos:

1. Genera todos los conjuntos de observaciones con una única observación. Usa estos para generar los de dos observaciones, y así sucesivamente.
2. Se toman todos los posibles pares que cumplen con las medidas mínimas de soporte. Esto permite ir eliminando posibles combinaciones ya que no todas se tienen que considerar.
3. Genera las reglas revisando que cumplan con el criterio mínimo de confianza.

Si una conjunción de consecuentes de una regla cumple con los niveles mínimos de soporte y confianza, sus subconjuntos (consecuentes) también los cumplen.

Por el contrario, si alguna observación no los cumple, no tiene caso considerar sus super-conjuntos. Esto da una forma de ir construyendo reglas con un solo consecuente, y a partir de ellas construir de dos consecuentes y así sucesivamente. Este método hace una pasada por la base de datos para cada conjunto de observaciones de diferente tamaño.

El esfuerzo computacional depende principalmente de la cobertura mínima requerida, y se lleva prácticamente todo en el primer paso [15] y [16]. Ver Anexo A.2.4.

#### **2.4.7 Algoritmo CONJUNCTIVE RULE**

Este algoritmo implementa solo una regla de aprendizaje conjuntiva que funciona tanto para datos numéricos como nominales/categóricos.

Una regla consiste en varios antecedentes " Y" juntos y el valor de la clase para la clasificación. Lo siguiente que realiza el algoritmo es la distribución de las clases disponibles o el término medio para un valor numérico. Si la instancia experimental es poco notoria por esta regla, entonces es prevista utilizando una distribución predeterminada de la clase en los datos cubiertos por la regla de aprendizaje [15].

#### **2.4.8 Algoritmo RIDOR (Ripple Down Rule)**

Genera primero una regla por defecto (predeterminada) y luego toma las excepciones para la regla predeterminada con la mínima tasa de error. Entonces genera la mejor excepción para cada excepción iterando hasta lograr disminuir el error. Luego genera una expansión similar a un árbol de excepciones.

La excepción es un conjunto de reglas que predice clases. Este algoritmo es usado para generar dichas excepciones [17].

### **3. DEFINICIÓN Y SELECCIÓN DE DESCRIPTORES PARA EL DIAGNÓSTICO DE EVENTOS DE TENSIÓN**

La definición y selección de los descriptores es presentada en este capítulo. El análisis es realizado teniendo en la cuenta cada una de las distintas causas que generan eventos de tensión (energización de transformadores, bancos de capacitores, fallas balanceadas y desbalanceadas), así como también la localización relativa de la fuente del evento (aguas arriba y aguas abajo). Finalmente, son presentados los resultados del análisis estadístico, con el cual es posible evidenciar el grado de relevancia/importancia de cada descriptor respecto a cada causa y/o localización relativa del evento.

#### ***3.1 LOCALIZACIÓN RELATIVA U ORIGEN DE LOS EVENTOS DE TENSIÓN***

Los descriptores descritos a continuación son tomados de los algoritmos existentes para la estimación de la localización relativa de la fuente de los eventos de tensión. La localización relativa de un evento de tensión hace referencia a si la fuente del evento se encuentra en la dirección (aguas abajo) o en dirección contraria (aguas arriba) del flujo de la potencia con respecto al punto de medida.

##### ***3.1.1 Cambio de fase en la corriente de secuencia positiva PCSC***

Este descriptor corresponde a la diferencia del ángulo de fase entre las componentes de secuencia positiva de la corriente de falla y prefalla para estimar el origen del evento de tensión.

Los tres fasores de corriente de prefalla son tomados un periodo antes del punto de inserción de la falla, mientras que los tres fasores de corriente de falla son tomados un periodo después del instante de inserción de la falla. Luego, la diferencia angular entre la fase la corriente de secuencia positiva de prefalla y la componente de secuencia positiva de falla es calculada. Una explicación detallada de este descriptor y su respectivo algoritmo es dada en [18] y [26].

$$PCSC = \theta I_{falla(1)} - \theta I_{prefalla(1)} \quad (3.1)$$

$\theta I_{falla(1)}$  : Ángulo de la componente de secuencia positiva de la corriente de falla

$\theta I_{prefalla(1)}$ : Ángulo de la componente de secuencia positiva de la corriente de prefalla

### **3.1.2 Relación entre la magnitud de la corriente de falla y pre-falla de secuencia positiva ( $I_{1ratio}$ )**

La estimación de la relación entre las magnitudes de la corriente de secuencia positiva, de la falla y pre falla de un sistema trifásico depende del punto de inicio de la falla y consecuentemente corresponde al cambio en la dirección del flujo de potencia según donde se encuentre el equipo de medida que registra dicha perturbación. De acuerdo a los resultados presentados en [18], cuando se presenta un evento de tensión aguas abajo del punto de medida, la magnitud de la corriente de secuencia positiva sufre un ligero incremento inmediatamente después del instante de inserción de la falla en el sistema. Caso contrario ocurre cuando la ubicación de la falla se encuentra aguas arriba del punto de medida. Por tanto  $I_{1ratio}$  (Ecuación 3.2) toma valores mayores a la unidad para eventos aguas abajo y menores a la unidad para aguas arriba.

$$I_{1ratio} = \frac{I_{falla_{sec+}}}{I_{prefalla_{sec+}}} \quad (3.2)$$

### **3.1.3 Signo de la resistencia equivalente durante la perturbación**

#### **( $R_{ex}$ $R_{ey}$ $R_e$ )**

Estos descriptores se basan en la estimación de la parte real de la impedancia equivalente  $R_e$  durante el evento de tensión. El algoritmo propuesto en [19] utiliza las componentes de tensión y corriente de secuencia positiva para calcular los valores de  $R_e$ .  $R_e$  es estimada tomando  $n$  ciclos de las ondas de tensión y corriente incluyendo ciclos de pre-falla y falla. El número de ciclos es determinado por la reversión del flujo de potencia.  $R_{ex}$  corresponde a la impedancia estimada a

partir de la parte real de las componentes de secuencia positiva de la tensión, mientras que  $R_{ey}$  a la parte imaginaria, ecuaciones (3.3) y (3.4) [20] y [26]

$$\begin{bmatrix} R \\ X \\ E_{RE} \end{bmatrix} = \begin{bmatrix} I_{RE}(1) & I_{IM}(1) & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ I_{RE}(1) & I_{IM}(1) & 1 \end{bmatrix}^{-1} \begin{bmatrix} V_{RE}(1) \\ \cdot \\ \cdot \\ V_{RE}(n) \end{bmatrix} \quad (3.3)$$

$$\begin{bmatrix} R \\ X \\ E_{IM} \end{bmatrix} = \begin{bmatrix} I_{IM}(1) & I_{RE}(1) & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ I_{IM}(1) & I_{RE}(1) & 1 \end{bmatrix}^{-1} \begin{bmatrix} V_{IM}(1) \\ \cdot \\ \cdot \\ V_{IM}(n) \end{bmatrix} \quad (3.4)$$

Por lo tanto, si ambas resistencias ( $R_{ex}$  y  $R_{ey}$ ) tienen diferente signo el algoritmo no determina el origen del hueco de tensión. Signos positivos significa que la fuente del hueco está aguas arriba y signo negativo lo contrario.

### 3.1.4 Cambio de la magnitud y fase ( $Z_{ratio}$ y $Angle_{zag}$ ) de la impedancia vista desde el alimentador antes y durante el evento.

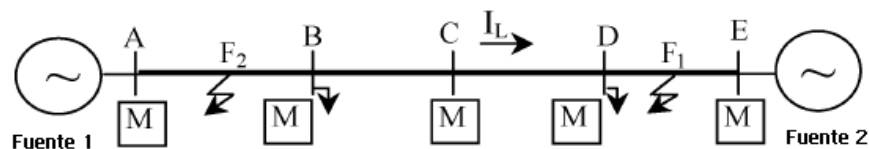


Figura 3. Sistema de potencia interconectado

El cálculo de estos descriptores se basa en el principio de la magnitud y ángulo de la impedancia antes y después del evento, indicando claramente la localización relativa del hueco de tensión con respecto al punto de medida (ver Figura 3). Para el caso de fallas aguas arriba, la dirección de la corriente se invierte y por tanto la impedancia vista por el medidor cambia su magnitud y ángulo de fase.

Para fallas aguas abajo, la magnitud de la impedancia vista durante la falla se reduce con respecto a la impedancia vista en estado estable, mientras que su ángulo de fase durante el evento se incrementa [20].

$Z_{ratio}$  es el cociente entre la impedancia de falla y la impedancia de prefalla capaz de identificar la localización relativa del hueco de tensión de acuerdo a la tendencia de este descriptor.

$$Z_{ratio} = \frac{Z_{falla}}{Z_{prefalla}} \quad (3.5)$$

### 3.1.5 Componente Real de la Corriente (RCC)

Para el análisis de este descriptor se considera un sistema con dos fuentes de energía y dos puntos de medición  $M_A$  y  $M_B$ . Se considera que la falla ocurrió en el punto X, ver Figura 4.

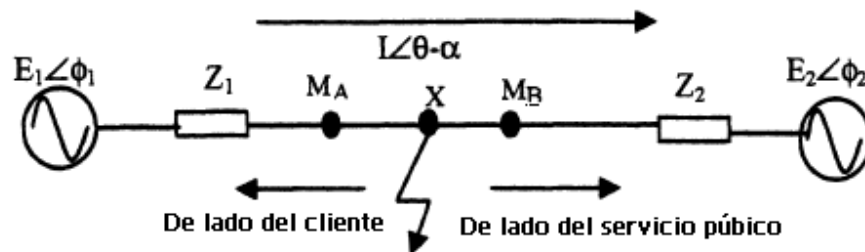


Figura 4. Sistema de dos fuentes para el análisis de la localización de un hueco de tensión

Para este caso (medidor  $M_A$ ) la corriente que fluye desde  $E_1$  a X, figura 4, está dada por  $I \cos(\theta - \alpha) > 0$ , (la  $RCC > 0$ ) y el hueco de tensión es visto desde el lado del cliente (aguas abajo). Esta corriente tiene la misma dirección que la corriente de pre-falla desde  $E_1$  a  $E_2$ . En el caso de monitorizar un punto después de la falla ( $M_B$ ) la corriente es  $I \cos(\theta - \alpha) < 0$  porque la corriente fluye a tierra en el punto de falla y esta corriente fluye desde  $E_2$  a X.

La falla del lado del servicio público y el origen del hueco de tensión es visto en este mismo lado (aguas arriba).

Cuando el hueco de tensión inicia, la corriente puede ser significativamente más alta que la corriente en estado estable [22].

### **3.2 DESCRIPTORES QUE CARACTERIZAN LA CAUSA DE LOS EVENTOS DE TENSIÓN**

A continuación se propone un conjunto de descriptores que se asocian a una o varias causas del evento de tensión. En el proceso de extracción de cada señal analizada se cuantifica la información relevante de cada perturbación y la formulación de los descriptores que permiten identificar la correcta asociación de un hueco de tensión con la causa que lo produjo. Se aclara que las señales extraídas de tensión y corriente eficaces empleadas en el análisis pertenecen a la componente fundamental.

#### **3.2.1 Relación entre la potencia aparente antes y después del evento (*Sratio*)**

La variación de la potencia aparente permite identificar los eventos de tensión provocados por el arranque de grandes motores de inducción. La relación *Sratio* se define como la relación entre la potencia aparente de posfalla y prefalla y está definida de la siguiente manera:

$$S_{prefalla} = \frac{1}{N} \sum_{n=1}^N S_{rms} \quad (3.6)$$

$$S_{posfalla} = \frac{1}{M} \sum_{m=1}^M S_{rms} \quad (3.7)$$

$$S_{ratio} = \frac{S_{posfalla}}{S_{prefalla}} \quad (3.8)$$

*N*: número de muestras en el estado de prefalla  
*M* número de muestras en el estado de posfalla

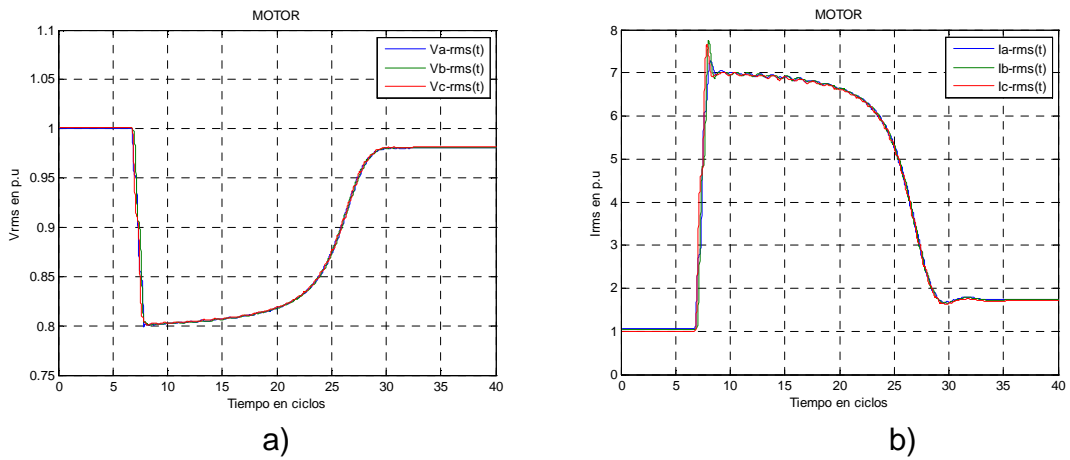


Figura 5. Formas de onda ante el arranque de un motor de inducción.

a) Tensión RMS    b) Corriente RMS

En la Figura 5 se presentan los valores eficaces de tensión y corriente de un evento de tensión causado por el arranque de un motor de inducción. Se observa un considerable incremento de la corriente después del evento; tal incremento ocasiona un crecimiento significativo de la potencia suministrada.

$Sratio$  tomará valores mayores a la unidad para eventos causados por arranque de grandes motores de inducción, debido al incremento en la potencia aparente después de la falla.

### 3.2.2 Relación de tensión ( $Vratio$ ) de corriente ( $Iratio$ )

Ambos descriptores sirven para estimar las variaciones de tensión y corriente respecto al estado de posfalla y prefalla.  $Vratio$  Se define como el cociente entre los respectivos estados de la señal de tensión, e  $Iratio$  como el cociente entre los respectivos estados de la señal de corriente. Estos descriptores son útiles para caracterizar la causa del evento, ya que  $Vratio$  es mayor a la unidad e  $Iratio$  es menor a la unidad cuando ocurre la energización de un banco de capacitores, reduciendo las pérdidas y mejorando la regulación.

$$V_{prefalla} = \frac{1}{N} \sum_{n=1}^N V_{rms} \quad (3.9)$$

$$V_{posfalla} = \frac{1}{M} \sum_{m=1}^M V_{rms} \quad (3.10)$$

$$V_{rpopre} = \frac{V_{posfalla}}{V_{prefalla}} \quad (3.11)$$

$$I_{prefalla} = \frac{1}{N} \sum_{n=1}^N I_{rms} \quad (3.12)$$

$$I_{posfalla} = \frac{1}{M} \sum_{m=1}^M I_{rms} \quad (3.13)$$

$$I_{rpopre} = \frac{I_{posfalla}}{I_{prefalla}} \quad (3.14)$$

$N$ : número de muestras en el estado de prefalla

$M$  número de muestras en el estado de posfalla

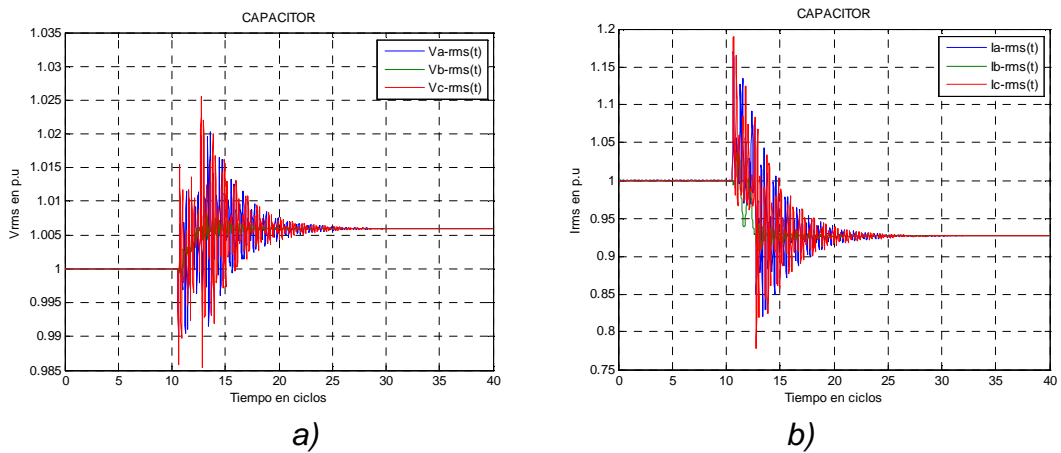


Figura 6. Formas de ondas debidas a la energización de un banco de capacitores

a) Tensión RMS b) Corriente RMS

La Figura 6 muestra los valores eficaces de tensión y de corriente registrados durante la energización de un banco de capacitores. En ella se nota que la tensión se incrementa alrededor de un 6%, mientras que la corriente se reduce alrededor de un 8%. Consecuentemente,  $V_{ratio}$  e  $I_{ratio}$  tienen valores mayores y menores que la unidad respectivamente.

### 3.2.3 Máxima desviación estándar de los valores de tensión y corriente eficaz ( $\sigma_{max}V$ y $\sigma_{max}I$ )

Estos descriptores permiten medir el grado de desbalance que existe tanto en las tres señales de tensión como en las tres señales de corriente. Se hace el cálculo tomando en cada muestra por ciclo las tres magnitudes eficaces tanto para la tensión como para la corriente, estimando para cada caso la desviación entre ellas. Es así como se obtiene un barrido de las desviaciones estándar entre las respectivas señales durante toda la evolución del hueco de tensión y se toma el mayor valor de estas desviaciones tanto para la tensión como para la corriente para la estimación de los respectivos descriptores (ver Figura 7).

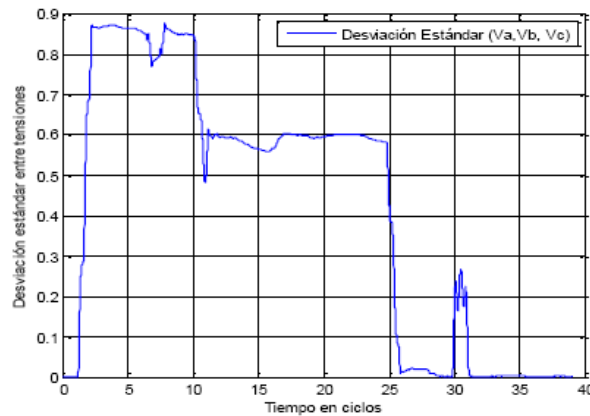


Figura 7. Desviación estándar entre las tensiones eficaces de un evento multiestado.

Ambos descriptores  $\sigma_{max}V$  y  $\sigma_{max}I$  toman valores altos ante fallas monofásicas, valores medios para fallas bifásicas, y los valores más bajos para fallas trifásicas.

$$\sigma_{Vrms}(t) = \sigma(V_{a\ rms}(t), V_{b\ rms}(t), V_{c\ rms}(t)) \quad (3.15)$$

$$\sigma_{max}Vrms = \max(\sigma_{Vrms}(t)) \quad (3.16)$$

$$\sigma_{Irms}(t) = \sigma(I_{a\ rms}(t), I_{b\ rms}(t), I_{c\ rms}(t)) \quad (3.17)$$

$$\sigma_{max}Irms = \max(\sigma_{Irms}(t)) \quad (3.18)$$

### 3.2.4 Pendientes de tensión y de corriente durante el inicio del evento ( $PV_{max}$ y $PI_{max}$ )

En el respectivo análisis y seguimiento de la señal de tensión y de corriente, se pudo notar que un factor relevante al presentarse un hueco de tensión es la rapidez con que cambian dichas señales, cuando se presenta el primer estado transitorio del evento.

Para el cálculo de las pendientes de cada fase ( $m_1$ ,  $m_2$ ,  $m_3$ ), ver figura 8, se hace un registro de las muestras en su primer estado transitorio, mediante el cálculo de regresión lineal se obtiene un valor de pendiente por cada fase. Los descriptores pendientes de tensión y de corriente denotados como  $PV_{max}$  y  $PI_{max}$  respectivamente, se obtienen al extraer el máximo valor de las magnitudes de las pendientes calculadas, tanto para las tres señales de tensión como para las tres señales de corriente. Aprovechando esta información extraída del primer estado transitorio, se proponen estos dos descriptores como indicadores que faciliten la estimación de la causa de la perturbación (ver Figura 8).

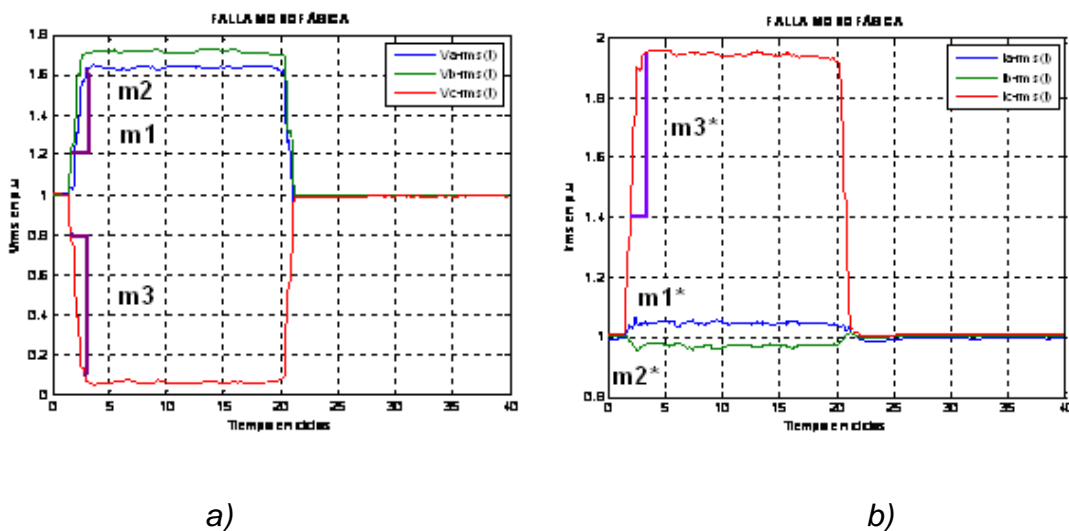


Figura 8. a) Estimación de pendientes de la señal de tensión RMS b) Estimación de pendientes de señal de corriente durante el primer estado transitorio en una falla monofásica

La expresión para obtener el descriptor  $PV_{max}$  se ilustra de la siguiente manera:

m1 : pendiente fase a  
 m2 : pendiente fase b  
 m3: pendiente fase c

$$PV_{max} = \max(|m1|, |m2|, |m3|) \quad (3.19)$$

### 3.2.5 Desviación estándar pendientes de tensión y desviación estándar pendientes de corriente durante el inicio del evento ( $\sigma_{PV}$ y $\sigma_{PI}$ )

El descriptor desviación estándar de las pendientes de las tres señales de tensión RMS  $\sigma_{PV}$  y el descriptor desviación estándar de las pendientes de las tres señales de corriente RMS  $\sigma_{PI}$  es el resultado de la desviación entre las pendientes calculadas en las respectivas señales eficaces durante la caída en el primer estado transitorio. Se emplean estos descriptores para estimar la causa del hueco de tensión. Ya que una desviación pequeña entre estas señales es un indicador de que ha ocurrido una falla trifásica debido a que las tres señales tienden a superponerse durante la evolución de la perturbación (ver Figura 9).

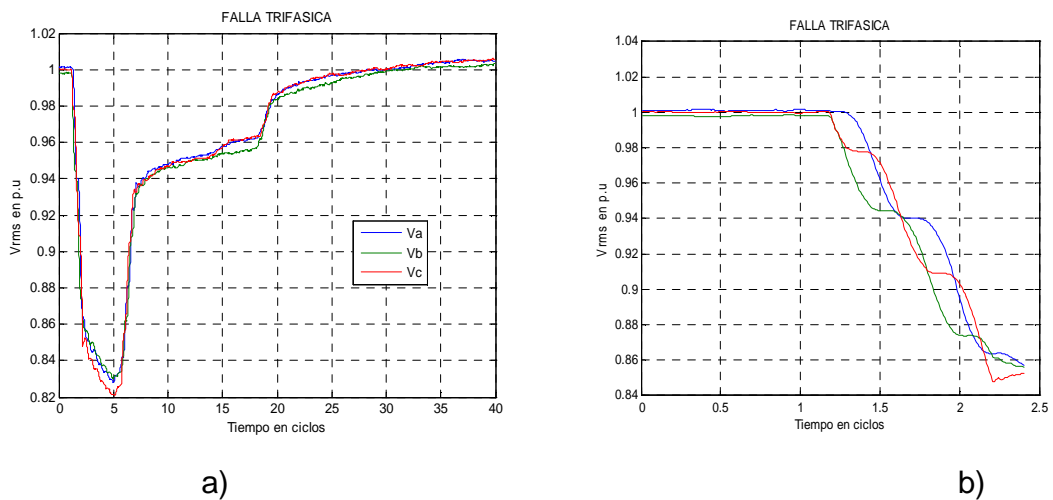


Figura 9. a) Hueco de tensión trifásico b) Estado transitorio durante la caída

La estimación de la desviación estándar entre las pendientes de la señal de tensión durante el primer estado transitorio en una falla de red se calcula:

m1 : pendiente fase c

m2 : pendiente fase a

m3: pendiente fase b

$$\sigma_{PV} = \delta(m1, m2, m3) \quad (3.20)$$

### 3.2.6 Comparación de fases ( $C_{fasesV}$ , $C_{fasesI}$ )

Este descriptor calcula los valores mínimos de cada señal de tensión para luego categorizarlos como máximo, medio y mínimo de acuerdo a los valores previamente obtenidos. Seguidamente se estiman las siguientes cantidades:

$$\Delta V1 = \text{Valor medio} - \text{Valor mínimo} \quad (3.21)$$

$$\Delta V2 = \text{Valor máximo} - \text{Valor mínimo} \quad (3.22)$$

$$C_{fasesV} = \min(\Delta V1, \Delta V2) \quad (3.23)$$

De acuerdo a los valores estimados por  $C_{fasesV}$  se puede identificar el tipo de falla ocurrida, es decir cuando ocurre una falla monofásica, este descriptor toma valores considerables como lo ilustra la Figura 10, en el caso de un evento bifásico se tendrán valores medios y en el caso trifásico este valor será muy pequeño.

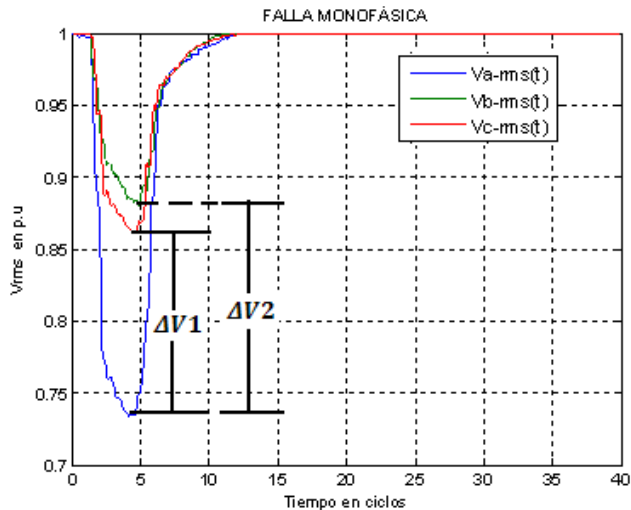


Figura 10. Deltas de tensión para un hueco monofásico.

### 3.2.7 Voltaje Complejo remanente RCV y factor positivo - negativo PNF

Estos descriptores contienen información del tipo y el grado de desbalance del evento de tensión respectivamente [24] y [25].

Comúnmente estos descriptores se han usado para caracterizar eventos de tensión originados por fallas de red, según resultados obtenidos en investigaciones [23].

Seis fases y componente simétrica son dos algoritmos usados para calcular el RCV y PNF [24], en conjunto con la caracterización de eventos desbalanceados basados en la clasificación de huecos tipo D (monofásico) y C (bifásicos) [23]. Una perturbación tipo  $D_a$  corresponde a un evento ocasionado por una falla monofásica en la fase 'a', en cuanto  $C_a$  hace referencia a una falla bifásica en las fases 'b' y 'c'. El algoritmo de las seis fases elimina la componente de secuencia cero para cada una de las tensiones de fase. Estas seis tensiones se calculan de la siguiente manera:

$$\begin{aligned} V_A &= rms \left\{ v_a - \frac{1}{3}(v_a + v_b + v_c) \right\} \\ V_B &= rms \left\{ v_b - \frac{1}{3}(v_a + v_b + v_c) \right\} \end{aligned} \quad (3.24)$$

$$\begin{aligned} V_C &= rms \left\{ v_c - \frac{1}{3}(v_a + v_b + v_c) \right\} \\ V_{AB} &= rms \left\{ \frac{v_a - v_b}{\sqrt{3}} \right\} \\ V_{BC} &= rms \left\{ \frac{v_b - v_c}{\sqrt{3}} \right\} \end{aligned} \quad (3.25)$$

$$V_{CA} = rms \left\{ \frac{v_c - v_a}{\sqrt{3}} \right\}$$

$$\begin{aligned} V_{rms}\{V_A(t)\} &\Rightarrow \text{hueco } D_a \\ V_{rms}\{V_B(t)\} &\Rightarrow \text{hueco } D_b \\ V_{rms}\{V_C(t)\} &\Rightarrow \text{hueco } D_c \end{aligned} \quad (3.26)$$

$$\begin{aligned} V_{rms}\{V_{AB}(t)\} &\Rightarrow \text{hueco } C_c \\ V_{rms}\{V_{BC}(t)\} &\Rightarrow \text{hueco } C_a \\ V_{rms}\{V_{CA}(t)\} &\Rightarrow \text{hueco } C_b \end{aligned} \quad (3.27)$$

Luego se calcula la secuencia de valores RMS para cada tensión de fase. Una vez identificado el tipo de hueco se halla el *RCV* y *PNF*.

Para un hueco tipo D:

$$\begin{aligned} RCV(t) &= V_1(t) + V_2(t) \\ PNF(t) &= V_1(t) - V_2(t) \end{aligned} \quad (3.28)$$

Para un hueco tipo C:

$$\begin{aligned} RCV(t) &= V_1(t) - V_2(t) \\ PNF(t) &= V_1(t) + V_2(t) \end{aligned} \quad (3.29)$$

*V1: Voltaje de secuencia positiva*

*V2: Voltaje de secuencia negativa*

Por tanto, según la literatura, para fallas monofásicas, los valores *RCV* y *PNF* tienden a la unidad y para fallas bifásicas y trifásicas sus magnitudes decaen significativamente, siendo mayor la caída para las trifásicas.

### **3.2.8 Máxima corriente y tensión de neutro (*Inratio* y *Vnratio*)**

*Inratio* se define como el cociente entre la corriente de neutro máxima durante el evento y la corriente de fase de prefalla. De manera similar fue calculado el descriptor *Vnratio*. Ambos descriptores fueron formulados con la intención de discriminar eventos ocasionados por energización de transformadores y arranque de motores. La hipótesis se basa en que *Inratio* y *Vnratio* tomarán valores mayores, cuanto mayor sea la diferencia de los niveles de saturación de los devanados del transformador.

Esto es válido solo en sistemas  $\Delta$ -Y, ya que de otra manera la corriente y la tensión de neutro no estarían disponibles [27]. La estimación de estos cocientes se realiza así:

$$V_{prefalla} = \frac{1}{N} \sum_{n=1}^N Vn_{rms} \quad (3.30)$$

$$I_{prefalla} = \frac{1}{N} \sum_{n=1}^N In_{inst} \quad (3.33)$$

$$V_{falla} = \frac{1}{M} \sum_{m=1}^M Vn_{rms} \quad (3.31)$$

$$I_{falla} = \frac{1}{M} \sum_{m=1}^M In_{inst} \quad (3.34)$$

$$Vn_{ratio} = \frac{vn_{falla}}{vn_{prefalla}} \quad (3.32)$$

$$In_{ratio} = \frac{in_{falla}}{in_{prefalla}} \quad (3.35)$$

*N: número de muestras en el estado de prefalla*

*M número de muestras en el estado de posfalla*

### **3.2.9 Ángulo de inserción de la falla (FIPA)**

Este descriptor es propuesto en [28] bajo la hipótesis de que las fallas debidas al contacto de la red eléctrica con animales y árboles se insertan alrededor del pico (ángulo de fase 90°) de la señal de tensión, debido a que en el pico es donde existe el gradiente máximo de tensión entre el animal/árbol y la red de distribución.

Según resultados presentados en [28] con señales reales, los eventos causados por contacto de la red con animales y árboles, tienen ángulos de fase de inserción de la falla alrededor de 60° y 120° (tanto en semiciclo positivo como negativo).

*FIPA* (Fault Insertion Phase Angle) es calculado como la diferencia entre la señal instantánea de tensión registrada y su correspondiente fundamental.

A lo largo del evento, el ángulo de inserción es obtenido como el punto donde se supera el umbral establecido para la diferencia entre la señal registrada y la componente fundamental [28].

### **3.2.10 Cambio del ángulo de fase entre la tensión y la corriente ( $\theta_{VI}$ y $\theta_{VI\_cruce}$ )**

Luego de la energización de capacitores, arranque de motores y energización/desenergización de grandes cargas se generan cambios en el ángulo de deslizamiento entre la tensión y la corriente. Tales cambios dependen de la

naturaleza inductiva o capacitiva de la carga, por ejemplo, ante la energización de un banco de capacitores se espera que el ángulo de deslizamiento se reduzca con respecto al deslizamiento en estado de prefault, caso contrario ocurriría con el arranque de un motor de inducción o una carga inductiva representativa.

Este descriptor se calcula empleando los valores instantáneos de la tensión y de la corriente antes y después de la perturbación, el deslizamiento del ángulo es calculado mediante la diferencia de muestras que existe entre la señal de tensión y de corriente durante el primer cruce por cero en el primer ciclo de prefault y en el último ciclo de posfalla tal como se ilustra en la Figura 11 [27].

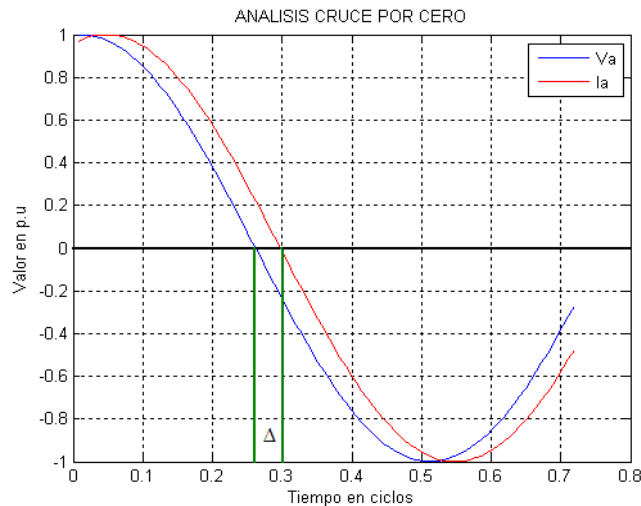


Figura 11. Estimación del ángulo entre la tensión y la corriente ( $V_{Base} = 25Kv$ ,  $I_{Base} = 54A$ )

### 3.2.11 Relación entre los factores de potencia ( $FPratio_{FFT}$ , $FPratio_{P\&S}$ , $FPratio_{cruce}$ )

Se propone el cociente entre el factor de potencia de posfalla y prefalla partiendo de la hipótesis expuesta en el descriptor anterior.

Se plantean tres maneras de calcular los ángulos de desfase entre tensión y corriente en prefalla y posfalla y el factor de potencia.

1. Utilizando la Transformada Rápida de Fourier (FFT): A partir de la estimación de los ángulos de los fasores de tensión y corriente en el primer y último ciclo de la señal.

$$FP = \cos(\theta_V - \theta_I) \quad (3.36)$$

$$FP_{prefalla} = \frac{1}{N} \sum_{n=1}^N FP_n \quad (3.37)$$

$$FP_{posfalla} = \frac{1}{M} \sum_{m=1}^M FP_m \quad (3.38)$$

$$FPratio_{FFT} = \frac{FP_{posfalla}}{FP_{prefalla}} \quad (3.39)$$

*N: número de muestras en el estado de prefalla*

*M número de muestras en el estado de posfalla*

2. Utilizando la potencia activa (P) y potencia aparente (S): El factor de potencia fue tomado como P/S en ambos estados. P y S fueron obtenidas a partir del producto de los fasores de tensión y corriente arrojados por la aplicación de la FFT con una ventana deslizante de longitud igual a un ciclo.

$$FP_{prefalla} = \frac{1}{N} \sum_{n=1}^N \frac{P_n}{S_n} \quad (3.40)$$

$$FP_{posfalla} = \frac{1}{M} \sum_{m=1}^M \frac{P_m}{S_m} \quad (3.41)$$

$$FPratio_{P\&S} = \frac{FP_{posfalla}}{FP_{prefalla}} \quad (3.42)$$

*N: número de muestras en el estado de prefalla*

*M número de muestras en el estado de posfalla*

3. Utilizando los cruces por cero de la señal de tensión y corriente: Los ángulos de fase de tensión y corriente fueron estimados utilizando los cruces por cero de la señal de tensión y de corriente instantánea.

$$FP_{prefalla} = \cos(\theta_{V_{pre}} - \theta_{I_{pre}}) \quad (3.43)$$

$$FP_{posfalla} = \cos(\theta_{V_{pos}} - \theta_{I_{pos}}) \quad (3.44)$$

$$FP_{ratio_{cruce}} = \frac{FP_{posfalla}}{FP_{prefalla}} \quad (3.45)$$

*N*: número de muestras en el estado de prefalla

*M* número de muestras en el estado de posfalla

La relación del factor de potencia fue calculada de tres maneras distintas con el objetivo de evaluar la capacidad de los tres distintos enfoques.

La relación entre el factor de potencia es superior a la unidad ante la energización de bancos de capacitores en diferentes puntos del sistema [23].

### **3.2.12 Cambio en la potencia reactiva ( $\Delta Q$ )**

En este descriptor las potencias reactivas de prefalla y posfalla son calculadas en el primer y último ciclo del evento respectivamente. Mediante esta diferencia entre potencia reactivas es posible identificar eventos causados por energización de bancos de capacitores, ya que al entrar al sistema generan cambios significativos en la potencia reactiva demandada por la carga (ver Figura 12).

$\Delta Q$  tiende a valores positivos debido al comportamiento natural de la potencia reactiva ante este hecho [27]. El cálculo de este descriptor se obtiene empleando los valores relativos de la potencia reactiva en la etapa de prefalla y posfalla.

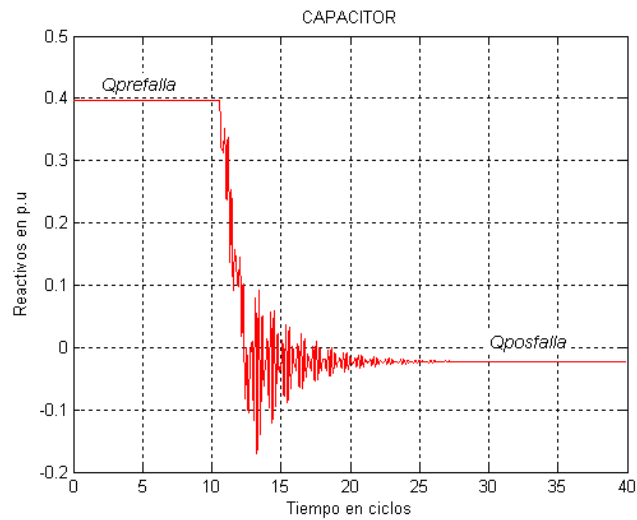


Figura 12. Comportamiento de la potencia reactiva antes y después de la energización de bancos de capacitores

$$\Delta Q = Q_{prefalla} - Q_{posfalla} \quad (3.46)$$

Donde:

$Q_{prefalla}$ : Potencia reactiva de prefalla

$Q_{posfalla}$ : Potencia reactiva de posfalla

### 3.2.13 Magnitud de la impedancia de falla ( $Z_{falla}$ )

La magnitud de la impedancia de falla es estimada mediante la tensión de la fase fallada y la corriente de neutro (3.47). Se utilizó la magnitud de la corriente de neutro porque depende de la impedancia de falla.

Este descriptor fue incluido en el análisis con la intención de discriminar las fallas de red de acuerdo a la magnitud de su impedancia [28].

$$Z_{falla} = \min \left| \frac{V_i < \theta_i}{I_{n_i} < \phi_i} \right| \quad (3.47)$$

Donde:

$Z_{falla}$  = Valor estimado de la impedancia de falla

$i =$  Número de ciclos de la corriente y la tensión durante el evento

$V_i =$  magnitud de la componente fundamental de la tensión instantánea de la fase fallada en el ciclo  $i$

$\theta_i =$  Ángulo de la tensión de la fase fallada

$I_{n_i} =$  magnitud de la componente fundamental de la corriente de neutro en la fase fallada en el ciclo  $i$

$\phi_i =$  Ángulo de fase de la corriente neutral en el ciclo  $i$

### 3.2.14 Máxima tensión de arco ( $V_{arc_{max}}$ )

Con este descriptor se busca medir la magnitud de la tensión del arco producido durante el contacto de animales y árboles con la red, así como también de rayos que impactan sobre la misma [28].

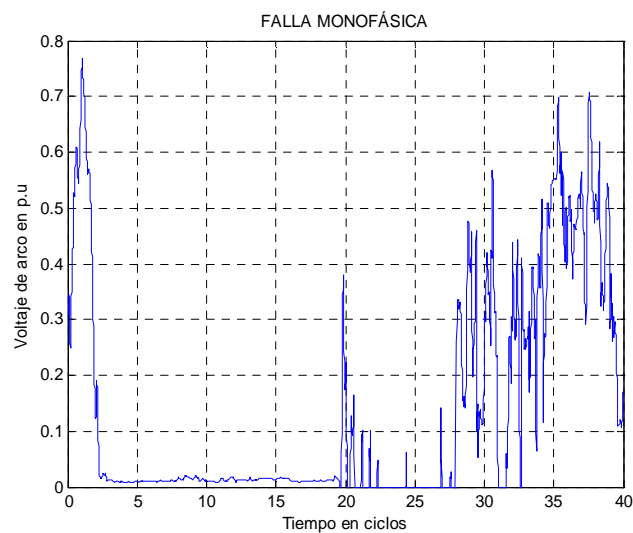


Figura 13. Tensión de arco en los estados de prefalla, falla y posfalla de un evento monofásico

Para fallas permanente  $Varc_{max}$  toma valores cercanos a cero, mientras que para fallas recurrentes  $Varc_{max}$  toma valores significativos. En este trabajo de grado  $Varc_{max}$  ha sido calculado como se describe en [28].

La tensión de arco sólo toma valores válidos durante el estado estable del evento (ciclos 3 a 18 aproximadamente) ver Figura 13.

### 3.2.15 Frecuencia más alta de oscilación durante el evento (HighOscFrec)

Este descriptor permite estimar la magnitud de las oscilaciones de alta frecuencia presentes en la señal de tensión mediante la utilización de la transformada Wavelet (ver Figura 14). Las señales de tensión de cada evento fueron descompuestas en dos escalas (1920-960Hz y 960-480Hz) en las cuales según los resultados consignados en [28] contienen información respecto a las componentes de alta frecuencia presentes durante eventos causados por animales, árboles y rayos.

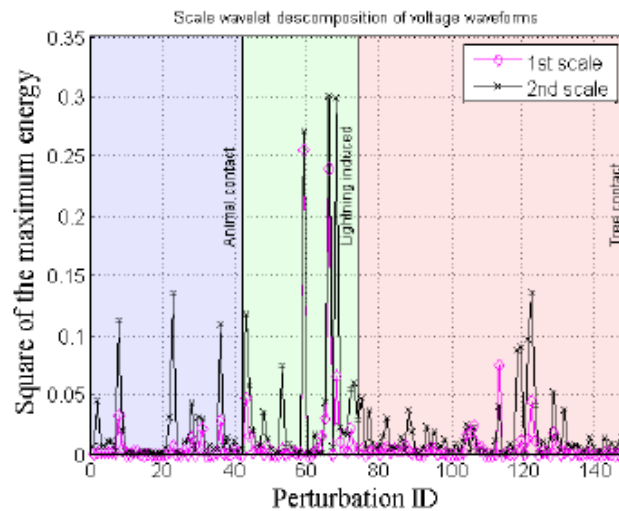


Figura 14. Descomposición wavelet de la señal de voltaje en dos escalas de frecuencia

Este descriptor fue utilizado con la intención de discriminar los eventos causados por la energización de bancos de capacitores, los cuales presentan oscilaciones

de alta frecuencia debidas al transitorio entre la impedancia capacitiva del banco y la inductiva de la red.

### **3.2.16 Ganancia de tensión de secuencia cero ( $GV_0$ ) y máxima tensión de secuencia cero ( $MaxV_0$ )**

El descriptor  $GV_0$  se utiliza para estimar el grado de desbalance de la señal de tensión.  $GV_0$  se define como la integral del valor eficaz de la tensión RMS de secuencia cero durante el evento de tensión, así:

$$GV_0 = \sum_{sag} \frac{V_0(t)_{rms}}{V_{prefalla}} \quad (3.48)$$

La estimación de  $GV_0$  se realiza durante la perturbación y consiste en la sumatoria de la magnitud de la tensión de secuencia cero en por unidad con respecto al valor de prefalla en cada instante de tiempo. Este resultado permite discriminar los tres tipos de fallas monofásicas, bifásicas y trifásicas, a manera de ejemplo se muestra el evento multiestado de la Figura 15, donde  $GV_0$  se espera que tome valores considerables para las fallas monofásicas, medios para bifásicas y bajos para fallas trifásicas [29].

El descriptor  $MaxV_0$  corresponde al valor máximo tomado por la tensión eficaz de secuencia cero que se presenta durante la falla.

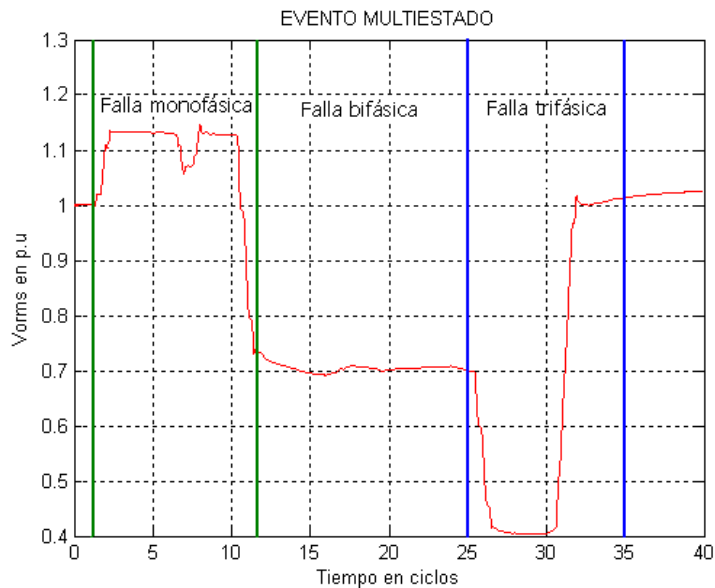


Figura 15. Tensión de secuencia cero en un evento de tensión multiestado

### 3.2.17 Magnitud del segundo armónico de corriente $|I_2|$

La magnitud de la segunda componente armónica de corriente se utiliza para identificar eventos de tensión causados por la energización de transformadores, ya que la energización de estos introducen componentes armónicas de segundo orden debido al flujo magnético remanente en los devanados del mismo.

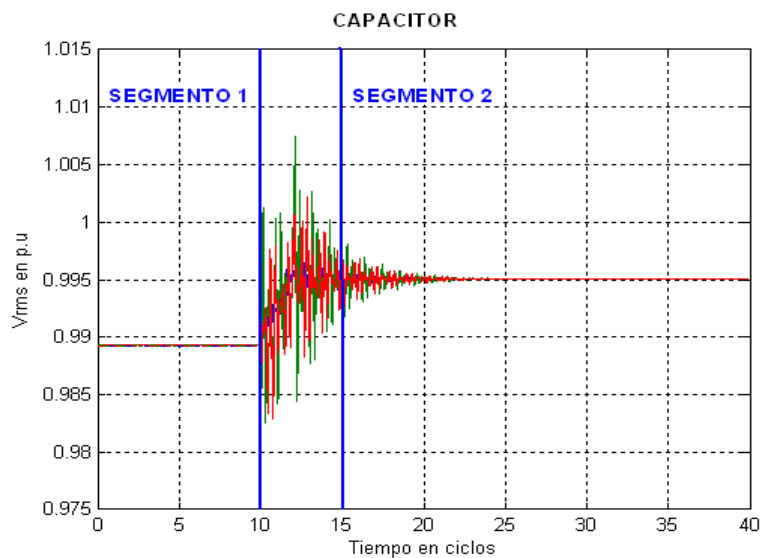
Este descriptor es calculado utilizando la transformada discreta de Fourier y una ventana deslizante de un ciclo de longitud [23].

### 3.2.18 Número de segmentos de transición (Segmentos $_{\delta=0.1, m=128}$ y Segmentos $_{\delta=8, m=64}$ )

Este descriptor identifica el número de segmentos de transición durante la evolución de una perturbación, por ejemplo en la Figura 16 se presenta un evento correspondiente a la energización de un banco de capacitores el cual presenta dos (2) segmentos de transición, o en otra palabras, un (1) estado transitorio.

La identificación de los segmentos de transición es realizada aplicando la segmentación basada en la primera derivada de la señal de tensión eficaz descrito en [27].

La segmentación para el descriptor  $Segmentos_{\delta=0.1, m=128}$  fue llevada a cabo con una tasa de submuestreo de 128 muestras y un valor umbral de 0.1%, son los parámetros de segmentación utilizados según los resultados consignados en [27] son ideales para la identificación de bancos de capacitores y cambios de carga.



*Figura 16. Segmentos de transición de un evento causado por la energización de bancos de capacitores. El evento contiene dos (2) segmentos de transición o un (1) estado transitorio*

Con la intención de distinguir entre eventos multiestado y monoestado se utilizó el descriptor  $Segmentos_{\delta=8, m=64}$  con una tasa de submuestreo de 64 muestras y un valor umbral de 8%. Tales parámetros son adecuados para segmentar eventos de tensión [27]. Por tanto este descriptor tiende a tomar valores superiores a cuatro para eventos multiestado e igual a cuatro para monoestado, ver Figura 17.

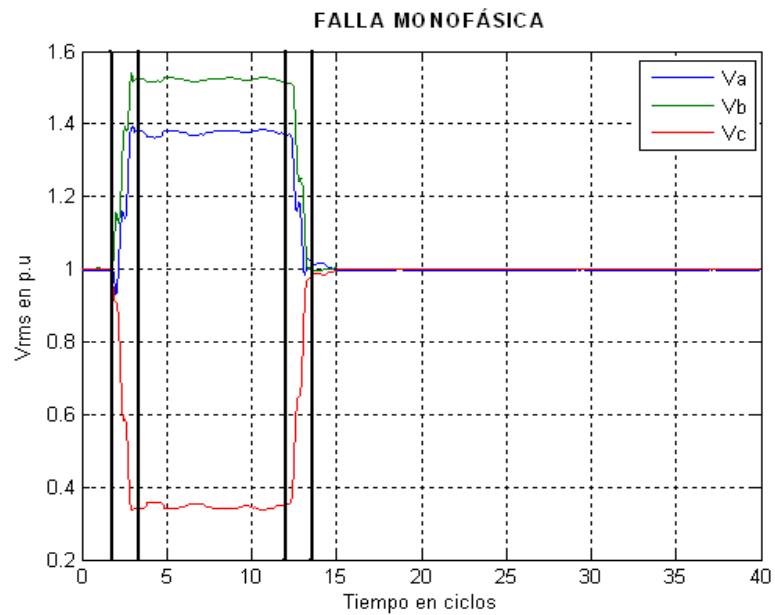


Figura 17. Segmentos de transición de un evento monoestado

### 3.2.19 Coeficiente de Forma (CF)

Observando el comportamiento de la secuencia de tensiones eficaces de cada uno de los huecos de tensión, se evidencia que algunos presentan características rectangulares, mientras otros características de tipo triangular (huecos de tensión causados por la energización de transformadores). El *Coeficiente de Forma CF* es propuesto para cuantificar si la secuencia de valores eficaces de la tensión durante la evolución del hueco sigue una tendencia triangular. Se espera que si se trata de una falla de red, *CF* obtendrá valores bajos debido a que la forma del hueco de tensión tiene una tendencia rectangular [23].

La manera de cuantificar el grado de tendencia triangular de la secuencia de valores eficaces consiste en comparar dicha secuencia con el triángulo comprendido entre los vértices denotados por el instante en que inicia la

perturbación, el instante en donde ocurre la máxima profundidad de tensión y el instante donde finaliza la perturbación, (figura 18) [23].

Como se observa en la Figura 18.b, un hueco de tensión causado por una falla de red no tiende a seguir los lados del triángulo, a diferencia de un hueco de tensión causado por la energización de un transformador en el cual su tensión eficaz está acotada por el triángulo y traerá consigo un alto valor en  $CF$ , como se muestra en la Figura 18.a [23].

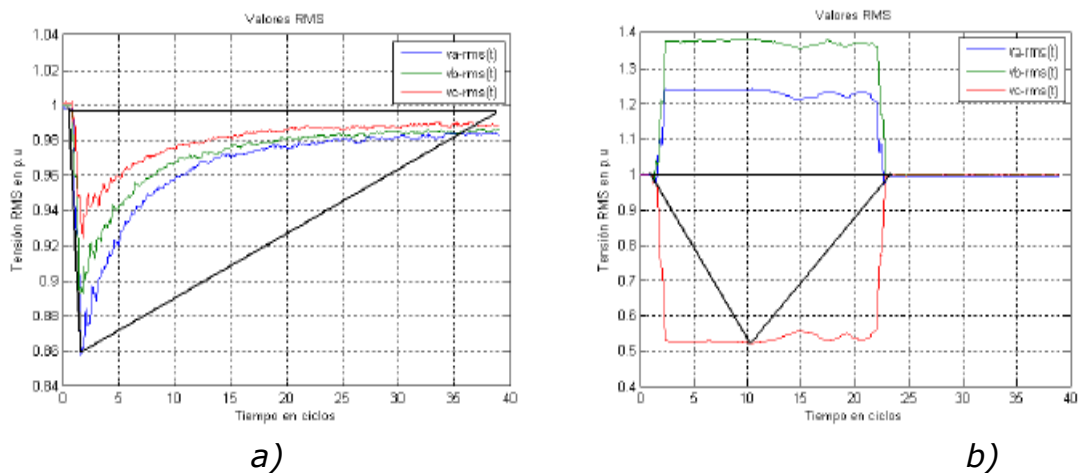


Figura 18. Hueco de Tensión con la zona de operación del descriptor  $CF$   
 a) Energización de un Transformador b) Falla de Red Monofásica

Este descriptor es calculado partir de tres índices relacionados con el grado de adecuación de la secuencia de valores eficaces con el lado de pendiente negativa del triángulo, uno con respecto al lado de pendiente positiva y otro con respecto a la cantidad de muestras por fuera del triángulo de referencia dados por los vértices mencionados anteriormente, una explicación detallada de este descriptor es dada en [23].

### 3.2.20 Desviación de los estados ( $\sigma E$ )

Durante el análisis de los registros suministrados se tiene la existencia de algunos eventos de tensión multiestado, (ver Figura 19). En esta grafica inicialmente se aprecia una falla monofásica, luego una falla bifásica y finalmente una falla

trifásica. Con este descriptor es posible estimar la magnitud de los cambios de tensión presentes en eventos multiestado o evolutivos, por tanto permitirá discriminar los huecos multiestado de los monoestado [23].

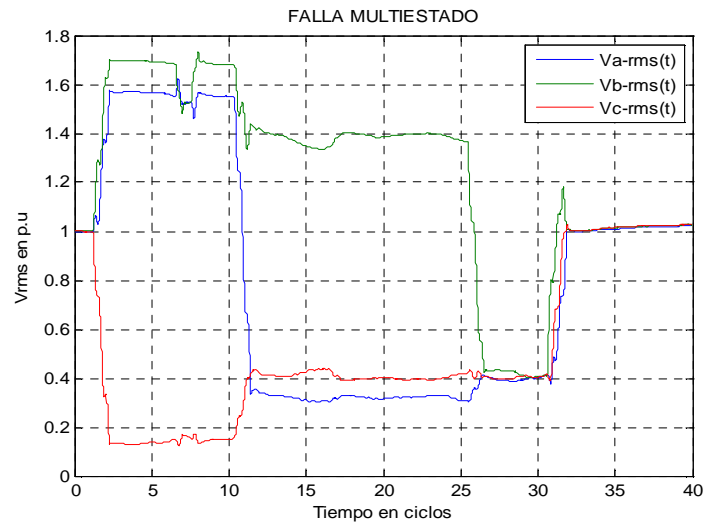


Figura 19. Hueco de tensión Multiestado

La *Desviación de los Estados*  $\sigma E$  es definida como el delta máximo de tensión en p.u que se presenta entre estados estables consecutivos en cualquiera de las tres tensiones eficaces. Para estimar los posibles estados estables del hueco de tensión se aplica la derivada a cada una de las señales de tensión eficaz.

Este cálculo se desarrolla en el período durante el cual las tensiones son menores al 90% de la tensión nominal [23].

El valor del descriptor es el mayor de todos los deltas de tensión estimados, todos los cálculos se realizan para cada tensión de fase eficaz. Esto se define de la siguiente manera:

$$\Delta_{aN} = |V_{aN} - V_{aN+1}| \quad (3.49)$$

$$\Delta_{bN} = |V_{bN} - V_{bN+1}| \quad (3.50)$$

$$\Delta_{cN} = |V_{cN} - V_{cN+1}| \quad (3.51)$$

$V_{a,b,cN}$ : Valor medio de la tensión eficaz en el estado estable N.

$$\sigma E = \max[\Delta_{aN} \ \Delta_{bN} \ \Delta_{cN}] \quad (3.52)$$

### 3.2.21 Relación angular de las longitudes ( $\delta L$ )

Es posible evaluar el carácter desbalanceado de un hueco mediante la estimación del número de fases que intervienen en dicho desbalance. La aplicabilidad de este descriptor radica en la identificación de huecos de tensión originados por fallas de red monofásicas y bifásicas. En la Figura 20 se toma un punto de referencia para cada tensión de fase y desde este punto se trazan rectas a cada uno de los puntos consecutivos a éste, determinando el valor de la proyección vertical de cada una de las rectas. El punto de referencia corresponde a la primera muestra de tensión eficaz de la respectiva fase, y así para el resto de las fases. Con estas proyecciones se realiza una sumatoria, una para cada fase, y a cada resultado se le denomina “longitud de la fase ( $l_A, l_B, l_C$ )” las cuales se podrían denominar como “longitudes eléctricas” [23] y su cálculo se muestra en las Ecuaciones 3.53, 3.54, 3.55 para cada fase.

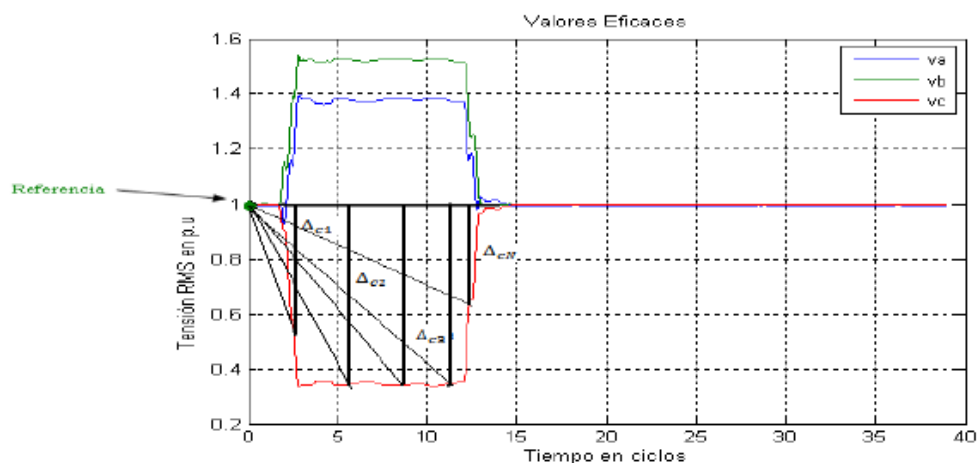


Figura 20. Metodología para el cálculo  $\delta L$

$$\ell_A = \sum_{i=1}^n \Delta_{ai} = \sum_{i=1}^n V_a(ref) - V_a(i) \quad (3.53)$$

$$\ell_B = \sum_{i=1}^n \Delta_{bi} = \sum_{i=1}^n V_b(ref) - V_b(i) \quad (3.54)$$

$$\ell_C = \sum_{i=1}^n \Delta_{ci} = \sum_{i=1}^n V_c(ref) - V_c(i) \quad (3.55)$$

Donde  $n$  es el número total de muestras de la tensión eficaz, así se obtienen tres longitudes totales  $\ell_A$ ,  $\ell_B$ ,  $\ell_C$ , de las cuales se selecciona la mayor como base para obtener un sistema en p.u. Después de esto se forman un par de triángulos, como se muestra en la Figura 21.

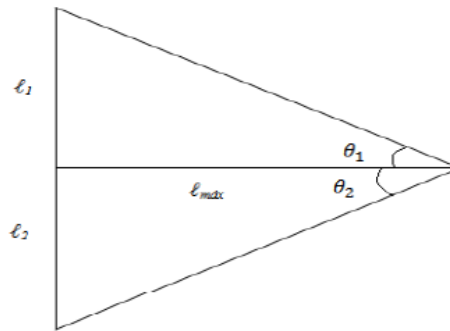


Figura 21. Triángulo de longitudes eléctricas

En la Figura 21 la mayor de las tres longitudes  $\ell_{max}$  corresponde a 1p.u. y sería este valor la base común de los triángulos en los que sus catetos serán  $\ell_1$  y  $\ell_2$ . Posteriormente se calculan los ángulos  $\theta_1$  y  $\theta_2$ . Como la mayor longitud se conoce solo queda por hallar si hay otra fase con una longitud similar o si las otras dos fases tienen longitudes más pequeñas. Por tanto el máximo valor que se puede presentar es  $45^\circ$ . De los ángulos hallados en p.u, el mayor de ellos correspondiente al descriptor  $\delta L$  [23].

$$\theta_1 = \tan^{-1}\left(\frac{\ell_1}{\ell_{max}}\right) \quad (3.56)$$

$$\theta_2 = \tan^{-1}\left(\frac{\ell_2}{\ell_{max}}\right) \quad (3.57)$$

$$(\theta_{1 pu}, \theta_{2 pu}) = \left(\frac{\theta_1}{45^\circ}, \frac{\theta_2}{45^\circ}\right) \quad (3.58)$$

$$\delta_L = \max(\theta_{1 pu}, \theta_{2 pu}) \quad (3.59)$$

Este descriptor sirve para identificar si la falla de red es monofásica o bifásica. Si es monofásica entonces  $\theta_1$  y  $\theta_2$  están por debajo de  $45^\circ$  y el descriptor tendrá un valor pequeño. Y si se trata de una falla bifásica, por lo menos uno de los dos ángulos se acercará a  $45^\circ$  haciendo que el descriptor  $\delta_L$  tienda a 1 p.u.

### **3.2.22 Desviación de las longitudes eléctricas ( $\sigma\delta_L$ )**

Con el objeto de evaluar el carácter balanceado o desbalanceado de un hueco de tensión se plantea un descriptor que estima el grado de desbalance de la perturbación para caracterizar el tipo de falla [23].

$$\sigma\delta_L = \sigma(\ell_A, \ell_B, \ell_C) \quad (3.60)$$

En la Figura 22 se muestra la secuencia de valores eficaces para un hueco de tensión originado por una falla de red trifásica. Se puede corroborar un comportamiento similar de las tres tensiones de fase durante la evolución del hueco, por lo que las tres longitudes eléctricas van a tener valores similares; lo anterior se traduce en un valor muy cercano a cero para la desviación estándar de las longitudes  $\sigma\delta_L$ . En el caso de los huecos de tensión desbalanceados, las longitudes eléctricas difieren significativamente, obteniéndose así valores representativos para este descriptor [23].

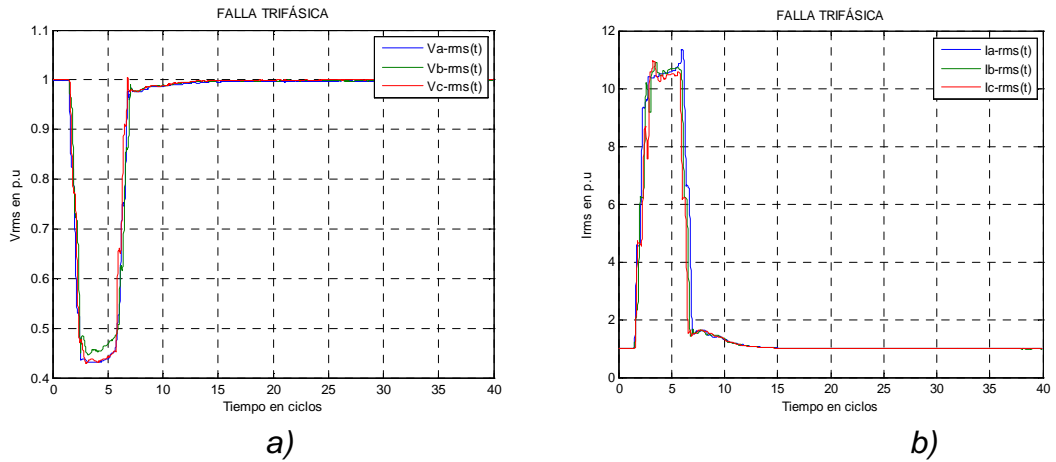


Figura 22. Huevo de tensión originado por una falla de red trifásica  
a) Tensión RMS b) Corriente RMS

### 3.2.23 Delta de longitudes ( $\Delta L_V$ y $\Delta L_I$ )

El descriptor  $\Delta L$  estima la diferencia porcentual entre las longitudes  $\ell_A$ ,  $\ell_B$ ,  $\ell_C$  anteriormente descritas.  $\Delta L$  es un complemento del *Delta Angular de Longitudes*  $\delta L$  en la identificación de huecos monofásicos o bifásicos. Se necesita caracterizar correctamente algunos huecos de tensión monofásicos en los cuales la longitud de la fase que falla no es tan significativa en comparación con las otras longitudes, pudiéndose considerar esa perturbación como un hueco bifásico.

Para su estimación se determina la diferencia entre la longitud máxima  $\ell_{max}$  y la subsiguiente mayor; de igual forma se hace con las dos longitudes de menor magnitud ( $\ell_1$  y  $\ell_2$ ). A partir del cálculo de estas dos diferencias, se calcula una relación porcentual entre ellas, la cual corresponde al valor del *Delta de Longitudes* ( $\Delta L$ ) [23].

Teniendo el orden decreciente de las tres longitudes [ $\ell_{max}$   $\ell_1$   $\ell_2$ ]:

$$\Delta L_1 = \ell_{max} - \ell_1 \quad (3.61)$$

$$\Delta L_2 = \ell_1 - \ell_2 \quad (3.62)$$

En donde la toma de decisión para el valor del descriptor se presenta de la siguiente manera:

$$\Delta L_1 > \Delta L_2 \Rightarrow \Delta L = 1 - \frac{\Delta L_2}{\Delta L_1} \quad (3.63)$$

$$\Delta L_1 < \Delta L_2 \Rightarrow \Delta L = \frac{\Delta L_1}{\Delta L_2} \quad (3.64)$$

Este descriptor ha sido implementado para entradas de tensión y corriente eficaz, en donde el análisis de las condiciones de decisión se explica a continuación.

Para el análisis con la tensión eficaz de un hueco producido por una falla bifásica, la condición correspondería a la Ecuación 3.64 ya que la diferencia entre las dos fases con una caída más severa  $\ell_1$  y  $\ell_2$  no sería tan significativa, por lo que  $\Delta LV$  tomaría un valor cercano a cero. Si se considera una falla monofásica, la condición será la Ecuación 3.63, en donde solo una fase tomaría un valor significativo de caída, por lo que  $\Delta L$  tomará un valor cercano a la unidad. Para el caso del análisis de la señal del valor eficaz de la corriente, un hueco producido por una falla bifásica, arroja como descriptor  $\Delta LI$  un valor cercano a la unidad. Para una perturbación producida por una falla monofásica se obtendrá un valor cercano a la cero, debido a la naturaleza de la gráfica de corriente para cada una de las fases.

### **3.2.24 Pérdida de tensión ( $L_V$ , $L_{V3}$ )**

Estos descriptores consideran la pérdida de tensión causada por la perturbación como un indicador para identificar fallas de red. Se calculan a partir del *minRCV* y las tres tensiones trifásicas respectivamente [29]. El descriptor  $L_V$  se basa en la profundidad de la tensión compleja remanente o *minRCV* a través del tiempo, mientras que el  $L_{V3}$  se basa en la profundidad promedio de las tres tensiones de fase. Están definidos por las siguientes expresiones:

$$LV = \sum_{Sag} (1 - RCV(rms)) \quad (3.65) \quad LV_3 = \sum_{Sag} (1 - V_{a,b,c}(rms)) \quad (3.66)$$

Donde  $minRCV$ ,  $V_{a,b,c}(rms)$  deben estar en por unidad. Nótese que  $Lv$  y  $Lv3$  calculan el área comprendida entre el valor nominal de la tensión (1 p.u) y la caída de tensión a lo largo de la duración de la perturbación ( $T_{sag}$ ).

Tanto  $Lv$  como  $Lv3$  están asociados a fallas de red debido a la caída de la tensión ante la ocurrencia de estas perturbaciones. Estos descriptores permiten caracterizar fallas monofásicas, bifásicas y trifásicas. La pérdida de tensión en una falla monofásica se debe solo a una fase, motivo por el cual  $Lv$  y  $Lv3$  convergen a valores mínimos comparados con pérdidas de tensión debido a fallas bifásicas o trifásicas.

## 4. ANÁLISIS DE LOS DESCRIPTORES MEDIANTE ESTADÍSTICA MULTIVARIANTE

### 4.1 ANÁLISIS MULTIVARIANTE DE LA VARIANZA – MANOVA

MANOVA realiza el análisis de la varianza para más de una variable dependiente. En este caso las variables dependientes corresponden a los descriptores definidos en el capítulo anterior. MANOVA determina la relevancia de cada descriptor con respecto a la localización relativa y causa de los eventos de tensión. En este caso, los eventos de tensión corresponden a las variables independientes en el análisis MANOVA. Como resultado, es posible conocer el grado de influencia o efecto de la causa/localización de los eventos de tensión sobre cada uno de los descriptores planteados.

MANOVA detecta si hay diferencias entre grupos a través de una combinación lineal de las variables dependientes, concluyendo si existen diferencias significativas entre las medias de dichas variables dependientes en los niveles de las variables independientes.

Al igual que en el análisis ANOVA (Análisis Univariante de la Varianza), MANOVA primero ejecuta un test global de significación y luego ejecuta un análisis ANOVA en cada una de las variables dependientes.

Una de las ventajas del análisis MANOVA es que tiene en cuenta la correlación entre todas las variables dependientes.

El modelo interno manejado por MANOVA, define que para K variables dependientes ( $k=1,2,\dots, K$ ), y G grupos ( $g=1,2,\dots, G$ ) realiza el siguiente análisis:

$$\begin{bmatrix} Y_{1gi} \\ Y_{2gi} \\ \cdot \\ \cdot \\ Y_{kgi} \end{bmatrix} = \begin{bmatrix} \mu_{1g} \\ \mu_{2g} \\ \cdot \\ \cdot \\ \mu_{kg} \end{bmatrix} + \begin{bmatrix} \epsilon_{1gi} \\ \epsilon_{2gi} \\ \cdot \\ \cdot \\ \epsilon_{kgi} \end{bmatrix} \quad g = 1,2, \dots, G, \quad i = 1, \dots, n_g \quad (4.1)$$

$$y_g = \mu_g + \epsilon_g \quad (4.2)$$

En donde la variable  $y_g$  es la media teórica del grupo, más una variable aleatoria. MANOVA trabaja con hipótesis o supuestos estadísticos. En el análisis multivariante, se trata una igualdad vectorial y no un número escalar a diferencia del análisis univariante [30].

En el planteamiento de las hipótesis sobre el proceso de obtención de la muestra, se asume que se ha extraído una muestra aleatoria multivariante independiente de cada una de las G poblaciones. En las hipótesis sobre la población, la matriz de covarianzas es la misma para todas las poblaciones (homocedasticidad). Estas poblaciones tienen una distribución normal multivariante distribuida así  $y_g \sim N(\mu_g)$ .

Usando el contraste de estas hipótesis es posible determinar si la variable independiente (causa o localización) tiene influencia significativa sobre las variables dependientes (descriptores).

La hipótesis nula ( $H_0$ ) argumenta que todas las medias grupales son las mismas y la hipótesis no nula ( $H_1$ ) rechaza la anterior proposición.

$$H_0: \mu_1 = \mu_2 = \mu_3 \dots \mu_g = \dots = \mu_G \quad (4.3)$$

$$H_1: \text{No todas las } \mu_G \text{ son iguales} \quad (4.4)$$

En este análisis estadístico se realiza una prueba para verificar la presencia de grupos o clases en los datos, esto quiere decir que existen nubes de datos en el espacio de descriptores (variables dependientes) correspondientes a grupos de eventos de tensión que pudieron ser producidos por diversas causas: fallas de red, energización de un transformador, bancos de capacitores, arranque de motores, etcétera, o por el contrario, su localización relativa aguas arriba o aguas abajo. Si se comprueba la existencia de estas clases, será posible identificar los descriptores más relevantes al aplicar el Análisis estadístico de Varianza Multivariante (MANOVA).

Para verificar si hay diferencias significativas entre los centros de los grupos/clases, se puede realizar la prueba llamada *Contraste Multivariado*, que trabaja sobre la distribución estadística *Lambda de Wilks* [30].

En un *Contraste Multivariado*, la hipótesis nula  $H_0$ , puede ser rechazada o aceptada, dependiendo del margen de confianza Lambda de Wilks, que se toma habitualmente a 0,05. Si el nivel de significancia es mayor que el margen de confianza (0,05),  $H_0$  es rechazada, esto implica la no existencia de grupos en los datos.

La prueba se realizó con un total de 522 eventos de tensión, los cuales fueron previamente clasificados según su causa y localización relativa. A partir de ellos se obtuvieron 46 descriptores (variables dependientes), de los cuales 38 descriptores están relacionados con la causa y 8 descriptores según la localización relativa de los eventos de tensión.

El valor de ajuste tomado como margen de confianza de la prueba Lambda de Wilks fue del 5% (0,05). Por tanto, al realizar un contraste de hipótesis y el nivel de significancia es mayor que 0,05, la hipótesis nula es rechazada.

Previo a cualquier análisis MANOVA, se recomienda realizar un contraste multivariado con la intención de verificar la existencia de grupos en el espacio de los descriptores. Si no existen grupos/clases ( $H_0$  rechazada) no se recomienda realizar MANOVA, los resultados no serán fiables. Si existen grupo ( $H_0$  aceptada) se recomienda la realización de MANOVA.

En el análisis estadístico efectuado en MANOVA la calidad de cada descriptor es definida por el estadístico  $R^2$ -corregida (%). Una  $R^2$ -corregida igual a 0% indica que el descriptor tiene una relevancia nula con respecto a la clase o variable independiente que se esté analizando. Una  $R^2$ -corregida igual al 100% indica que el descriptor posee un alto grado de relevancia con respecto a la variable independiente que se esté analizando (causa, origen, etc). Un descriptor con alta relevancia indica que con él es posible clasificar, distinguir o discriminar entre los posibles valores tomados por la variable independiente.

#### **4.2 IDENTIFICACIÓN DE LOS DESCRIPTORES RELEVANTES DE ACUERDO A LA CAUSA DE LOS EVENTOS DE TENSIÓN.**

Dentro de la identificación de los descriptores relevantes con respecto a la causa, se excluyeron todos aquellos que no están relacionados directamente con ella, para evitar la introducción de ruido en el análisis estadístico de los eventos. Por ejemplo, el descriptor  $V_{arc_{max}}$  no tendrá sentido que se tenga en cuenta en el análisis de eventos balanceados, ya que al energizar un banco de capacitores o al efectuar el arranque de un motor, no existe la posibilidad de aparición de un arco eléctrico debido a esta acción. Por otro lado, el descriptor  $Z_{falla}$  igualmente no se debe tener en cuenta en el análisis de eventos balanceados, debido a que en el momento de la energización de un banco de capacitores no se puede calcular su impedancia en el momento de la energización.

Para la identificación de los descriptores relevantes se establecieron 4 escenarios así:

1. Eventos balanceados y desbalanceados: Los eventos causados por energización de capacitores, arranque de motores y fallas de red trifásicas fueron etiquetadas como "Evento balanceado", mientras que las fallas de red monofásicas, bifásicas y energización de transformadores como "Evento desbalanceado". Este escenario es útil para la identificación de aquellos descriptores que permitirán distinguir eventos balanceados de desbalanceados.
2. Eventos desbalanceados: Sólo se introdujeron al análisis MANOVA aquellos descriptores obtenidos a partir de los eventos etiquetados como desbalanceados descritos en el escenario anterior. Por lo tanto, los descriptores relevantes identificados en este escenario, permiten discriminar entre eventos causados por falla monofásica, bifásica y energización de un transformador.

3. Eventos balanceados: Similarmente, los descriptores relevantes identificados en este escenario permitirán distinguir entre eventos causados por el arranque de un motor, la energización de un banco de capacitores y una falla trifásica.
  
4. Eventos monoestado y multiestado: Los eventos fueron etiquetados según la cantidad de estados cuasi-estables. Los descriptores identificados permitirán identificar los eventos multiestado.

A continuación se presenta el estadístico  $R^2$ -corregida<sup>1</sup> para cada uno de los descriptores respecto al escenario previamente establecido.

<sup>1</sup>Coficiente calculado por MANOVA que indica el grado de importancia de un descriptor respecto al escenario establecido, causa o localización relativa.

Tabla 1. Resultados MANOVA según el grado de desbalance del evento de tensión

<b>Descriptor</b>	<b>Definición</b>	<b>R<sup>2</sup>-corregida</b>
<b>CFasesV</b>	Comparación de fases de tensión	<b>92,40%</b>
<b>CFasesI</b>	Comparación de fases de corriente	<b>78,80%</b>
<b><math>\sigma_{PV}</math></b>	Desviación estándar pendientes de tensión	<b>91,40%</b>
<b><math>\sigma_{PI}</math></b>	Desviación estándar pendientes de corriente	4,00%
<b>RCV</b>	Voltaje complejo remanente	<b>90,40%</b>
<b>PNF</b>	Factor positivo - negativo	<b>90,00%</b>
<b><math>\sigma_{\max}V</math></b>	Máxima desviación estándar de la señal de tensión	<b>96,60%</b>
<b><math>\sigma_{\max}I</math></b>	Máxima desviación estándar de la señal de corriente	5,00%
<b><math>\sigma_{\delta L}</math></b>	Desviación de las longitudes eléctricas	<b>95,70%</b>
<b>Vnratio</b>	Máxima tensión de neutro	<b>83,00%</b>
<b>Inratio</b>	Máxima corriente de neutro	<b>71,60%</b>
GVo	Ganancia de tensión de secuencia cero	20,00%
Lv	Pérdida de tensión basada en RCV	11,00%
Lv3	Pérdida de tensión basada en las tres tensiones	2,00%
II2I	Segunda componente armónica de la corriente	22,60%
MaxVo	Amplitud máxima de tensión de secuencia cero	30,00%

Nota. Evento desbalanceado: Fallas monofásicas, bifásicas y energización de transformador.  
 Evento balanceado: Arranque de motores, bancos de capacitores y fallas trifásicas.

En la Tabla I es posible apreciar la relevancia o calidad ( $R^2$ -corregida) de cada descriptor con respecto a su grado de desbalance. Por ejemplo, el descriptor  $\sigma_{max}V$  (96,6%) obtuvo la mayor relevancia seguido por  $\sigma\delta L$  (95,7%) y  $CfasesV$  (92,4%).

También es posible apreciar que los descriptores basados en la corriente no tienen una relevancia significativa con respecto al grado de desbalance del evento de tensión, excepto para  $Inratio$  (71,6%) y  $CfasesI$  (78,8%). Por lo tanto es posible afirmar que la señal tensión contiene mayor información sobre el grado de desbalance de los eventos que la señal de corriente.

Los descriptores referentes a la pérdida de tensión  $Lv$  y  $Lv3$  obtuvieron niveles de relevancia insignificantes, aparentemente esto se debe a que un evento de tensión profundo y de poca duración, tiene la misma pérdida de tensión que un evento poco profundo y de larga duración.

Tabla 2. Resultados MANOVA para eventos desbalanceados

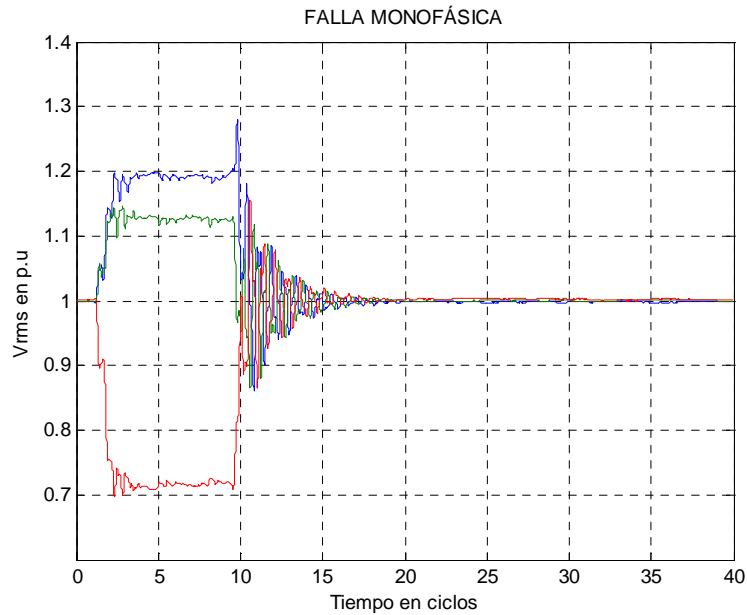
Descriptor	Definición	R <sup>2</sup> -corregida
<i>CFasesV</i>	Comparación de fases de tensión	60,00%
<b><i>CFasesI</i></b>	Comparación de fases de corriente	<b>84,60%</b>
<i>σ<sub>PV</sub></i>	Desviación estándar pendientes de tensión	63,00%
<i>σ<sub>PI</sub></i>	Desviación estándar pendientes de corriente	14,00%
<b>RCV</b>	Voltaje complejo remanente	<b>98,00%</b>
<b>PNF</b>	Factor positivo – negativo	<b>98,00%</b>
<b>σ<sub>maxV</sub></b>	Máxima desviación estándar de la señal de tensión	<b>71,00%</b>
<i>σ<sub>maxI</sub></i>	Máxima desviación estándar de la señal de corriente	14,00%
<b>δL</b>	Relación angular de las longitudes	<b>89,00%</b>
<b>ΔLV</b>	Delta de longitudes de V	<b>85,00%</b>
<b>ΔLI</b>	Delta de longitudes de I	<b>82,60%</b>
<b>CF</b>	Coefficiente de forma	<b>85,40%</b>
<b>Vnratio</b>	Máxima tensión de neutro	<b>72,00%</b>
<b>Inratio</b>	Máxima corriente de neutro	<b>68,70%</b>
<b>HighOscFrec</b>	Oscilaciones en alta frecuencia	<b>97,00%</b>
<i>GVo</i>	Ganancia de tensión de secuencia cero	58,00%
<i>Lv</i>	Pérdida de tensión basada en RCV	8,00%
<i>Lv3</i>	Pérdida de tensión basada en las tres tensiones	15,00%
<b>II2I</b>	Segunda componente armónica de la corriente	<b>68,60%</b>
<b>MaxVo</b>	Amplitud máxima de tensión de secuencia cero	<b>88,50%</b>

Nota. Se incluyeron sólo los eventos etiquetados como: Fallas monofásicas, bifásicas y energización de transformadores.

En la Tabla II se observa que el descriptor  $CFasesI$  (84,6%) y  $\delta L$  (89%) tienen una relevancia adecuada con respecto a los eventos desbalanceados. Este hecho indica que los eventos desbalanceados pueden ser discriminados tanto con la corriente ( $CfasesI$ ) como con la tensión ( $\delta L$ ).

Por otro lado, los descriptores  $RCV$  y  $PNF$  obtuvieron los índices de calidad más alto (98%), esto es debido a que los dos son capaces de discriminar las fallas monofásicas de las bifásicas y las causadas por energización de transformadores, ya que  $RCV$  y  $PNF$  toman valores cercanos a la unidad. Tales componentes aparecen durante el despeje de la falla y son causadas debido al transitorio entre la impedancia inductiva equivalente de la red y la equivalente capacitiva de la red subterránea. Una proporción considerable de los circuitos de ENDESA es subterránea, ver Figura 23.

Entonces,  $HighOscFrec$  obtiene una alta relevancia debido a que este descriptor es capaz de discriminar gran parte de las fallas de red, ya que la mayoría contienen componentes de alta frecuencia durante el despeje de la falla. Sin embargo, esto no quiere decir que en general todas las fallas de red tengan componentes de alta frecuencia. Esto es un caso particular de ENDESA y de circuitos subterráneos (ver Figura 23).



*Figura 23. Hueco de tensión monofásico con oscilación de alta frecuencia durante el despeje de la falla*

Para el descriptor  $GVo$  la duración de la perturbación afecta la calidad del descriptor, por ejemplo si se tiene un hueco de corta duración debido a una falla monofásica,  $GVo$  toma un valor bajo, y esto contradice la hipótesis que las fallas monofásicas tienen una considerable componente de secuencia cero; tal comportamiento se ve reflejado en los resultados estadísticos con un bajo porcentaje de  $R^2$  – corregida (58%). Por el contrario, el descriptor  $MaxVo$  que no tiene en la cuenta la duración del evento, obtiene una relevancia del 88,5%.

Tabla 3. Resultados MANOVA para eventos balanceados

Descriptor	Definición	R <sup>2</sup> -corregida
<b>Sratio</b>	Relación entre la potencia aparente de posfalla y prefalla	<b>98,00%</b>
<b>Vratio</b>	Relación de tensión de posfalla y prefalla	<b>76,30%</b>
<b>Iratio</b>	Relación de corriente de posfalla y prefalla	<b>98,00%</b>
<i>FPratio<sub>FFT</sub></i>	Relación entre el factor de potencia de posfalla y prefalla calculada con ángulos por FFT	42,20%
<i>FPratio<sub>P&amp;S</sub></i>	Relación entre el factor de potencia de posfalla y prefalla calculada por la FFT de la relación entre P y S	4,00%
<b>FPratio<sub>cruce</sub></b>	Relación entre el factor de potencia de posfalla y prefalla calculada por el cruce por cero de la tensión y la corriente	<b>85,00%</b>
<i>PVmax</i>	Pendiente máxima de tensión	43,80%
<i>PImax</i>	Pendiente máxima de corriente	43,00%
$\theta_{VI}$	Diferencia de los ángulos de la impedancia de posfalla y prefalla calculada por la FFT	5,00%
$\theta_{VI\_cruce}$	Diferencia de los ángulos de la impedancia de posfalla y prefalla calculada por el cruce por cero de la tensión y la corriente	21,00%
HighOscFrec	Oscilaciones en alta frecuencia	62,00%
<b><math>\Delta Q</math></b>	Diferencia de potencias reactivas de prefalla y posfalla	<b>70,00%</b>

Nota. Se incluyeron sólo los eventos etiquetados como: Fallas trifásicas, energización de banco de capacitores y arranque de motores.

En Tabla III se muestra que los descriptores basados en la relación de potencias, tensiones y corrientes son los que tienen mayor relevancia con respecto a los eventos balanceados. Esto era de esperarse, debido a que los eventos balanceados son simétricos y no es sencillo encontrar información valiosa durante

el evento que permita discriminar entre los distintos tipos de eventos de naturaleza balanceada.

Por otro lado, se evidencia que la relación de factores de potencia estimada utilizando los cruces por cero ( $FPratio_{cruce}$ ) de las tensiones y corrientes instantáneas tiene mayor efecto sobre los eventos balanceados que  $FPratio_{P\&S}$  y  $FPratio_{FFT}$ .

Tabla 4. Resultados MANOVA para eventos monoestado y multiestado

Descriptor	Definición	Estadístico R2-correctada
$\sigma E$	Desviación de los estados	<b>71,70%</b>
<b>Segmentos<math>_{\delta=0.1, m=128}</math></b>	Número de segmentos de transición con umbral de 0.1%	<b>77,60%</b>
<b>Segmentos<math>_{\delta=8, m=64}</math></b>	Número de segmentos de transición con umbral de 8%	<b>90,70%</b>

Nota. Evento monoestado: Fallas monofásicas, bifásicas y trifásicas. Evento multiestado: Fallas de red multiestado.

En Tabla IV se observa que para distinguir eventos monoestado de multiestados, es mejor utilizar el descriptor basado en segmentación  $Segmentos_{\delta=8, m=64}$  (90,7%) que el descriptor  $\sigma E$  (71,7%). Lo importante con el descriptor  $Segmentos_{\delta=8, m=64}$  es la selección de los parámetros de segmentación. Como se mencionó anteriormente, en este trabajo se utilizó una tasa de submuestreo de  $m=64$  y un umbral para la segmentación  $\delta=8\%$ .

De igual manera se nota que el descriptor  $Segmentos_{\delta=0.1, m=128}$  tiene menor relevancia que  $Segmentos_{\delta=8, m=64}$ , ya que los parámetros de segmentación del primero son adecuados para discriminar eventos de un solo estado transitorios, como energización de transformadores, energización de capacitores, cambios de carga, entre otros.

### 4.3 CLASIFICACIÓN DE LOS DESCRIPTORES SEGÚN EL ORIGEN DE LOS EVENTOS DE TENSIÓN

En la Tabla V se listan los descriptores asociados con la localización relativa de los eventos y su respectiva  $R^2$ -corregida.

Tabla 5. Resultado MANOVA según la localización relativa de los eventos de tensión

Descriptor	Definición	$R^2$ -corregida
<b><i>Rex</i></b>	Parte real de la impedancia	<b>72,30%</b>
<b><i>Rey</i></b>	Parte imaginaria de la impedancia	<b>76,30%</b>
<i>Re</i>	Resistencia equivalente	46,20%
<b><i>Zratio</i></b>	Relación de la impedancia durante la falla y la prefalla	<b>87,00%</b>
<i>AngleZag</i>	Angulo de la impedancia durante la falla	60,70%
<b><i>I1ratio</i></b>	Relación entre la magnitud de la corriente de falla de secuencia positiva y la corriente de prefalla de secuencia positiva	<b>98,00%</b>
<b><i>PCSC</i></b>	Diferencia del ángulo de fase de la corriente de falla de secuencia positiva y la corriente de prefalla de secuencia positiva	<b>70,90%</b>
<b><i>RCC</i></b>	Componente real de la corriente	<b>79%</b>

Nota. Los eventos incluidos en el análisis fueron etiquetados previamente como “Aguas arriba” y “Aguas abajo”.

Los resultados obtenidos son similares a los presentados en [7] y resaltan los descriptores  $I1_{ratio}$  y  $Zratio$ . Este último es un descriptor complementario con  $AngleZag$ , los cuales en conjunto incrementan la relevancia en la estimación del origen del evento de tensión.

## 5. EXTRACCIÓN DE REGLAS DE DECISIÓN

Una vez se tiene la lista de los descriptores con mayor relevancia respecto a la causa y al origen de los eventos de tensión, se aplicaron los algoritmos de extracción de conocimiento con el objetivo de extraer reglas de decisión que permitan diagnosticar de manera automática eventos de tensión.

Los algoritmos de extracción de reglas de las herramientas WEKA y Orange Canvas definidas en el Capítulo 2 fueron seleccionados por su robustez [8] y por su recomendación en algunos textos de la literatura consultada. En donde se introducen a estos algoritmos los datos de los descriptores seleccionados del análisis MANOVA de acuerdo al escenario que identifica la causa y la localización previamente establecido.

Los siguientes son algunos criterios tenidos en la cuenta para su selección:

- Crean reglas que relacionan varios descriptores, encontrando el mínimo número de errores.
- Entregan conjunto de reglas “SI...ENTONCES”, que son útiles en aplicaciones de sistemas expertos de minería de datos, mostrando reglas de fácil comprensión, que aquellos algoritmos que arrojan arboles de decisión.
- Cubren un mayor número de descriptores.
- Usan antecedentes “Y” juntos para relacionar descriptores.
- Trabajan con procesos iterativos hasta disminuir la tasa de error.
- Poseen alta precisión cuando predicen la clase con mayor cobertura.
- Algunos de ellos están relacionados con las mismas técnicas heurísticas, siendo confiables para comparar las reglas extraídas por ellos.

### 5.1 ALGORITMOS DE LA HERRAMIENTA COMPUTACIONAL WEKA

Los diferentes algoritmos seleccionados en WEKA que extraen reglas de decisión, son: *Conjunctive Rule*, *JRip (RIPPER)*, *OneR*, *PART* y *Ridor* mencionados en el Capítulo 2. En la figura 24 se muestra el entorno WEKA en la aplicación *Explorer*

para el análisis de extracción de reglas de eventos: balanceados y desbalanceados y los respectivos descriptores identificados como relevantes en el análisis MANOVA del capítulo anterior.

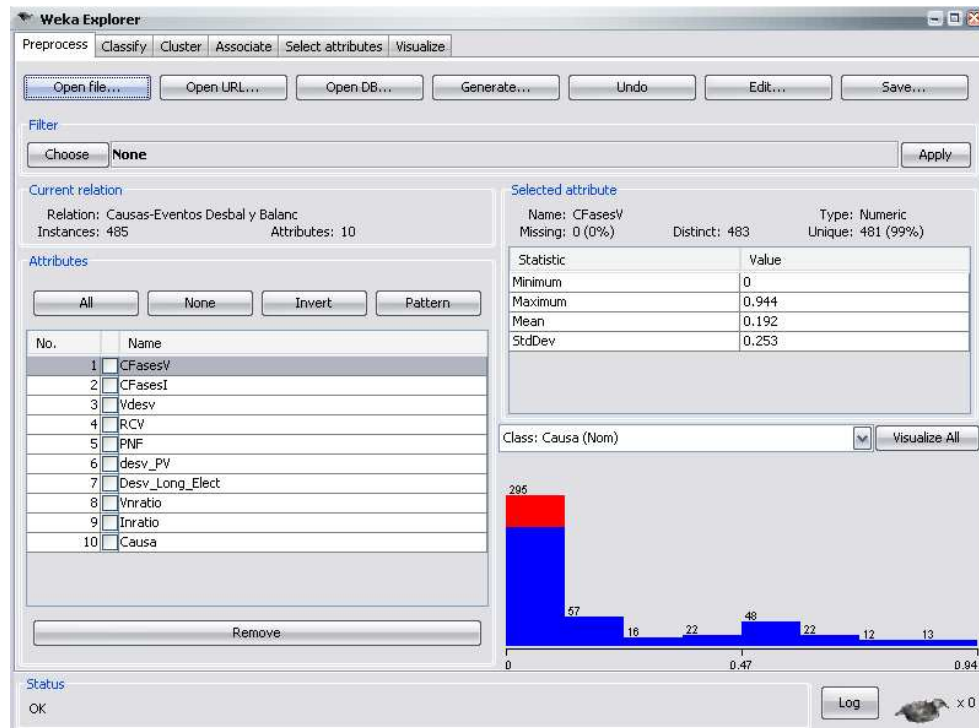


Figura 24. Entorno WEKA en la aplicación Explorer

Dentro de la aplicación *Classify de WEKA*, ver Figura 24, en donde aparece el entorno para ejecutar los diferentes algoritmos que extraen las reglas de decisión, se debe ajustar además de los valores predeterminados, como lo son el número de variables independientes (instancias totales) y número total de descriptores que describen el evento analizado (atributos).

Dentro de las opciones de prueba de la herramienta está la opción “*Cross-validation*”, validación cruzada, que permite analizar el conjunto de datos, dividiendo una muestra de los mismos en subconjuntos en la etapa inicial de la prueba, y posteriormente valida otros grupos de datos junto con los anteriores para crear nuevas reglas y así obtener una combinación final de las mismas.

Como validación cruzada se le debe asignar como número límite un valor de diez (para formar diez subconjuntos para efectuar las combinaciones de las reglas) y así obtener un buen resultado en las reglas de decisión arrojadas por el algoritmo. En la Figura 25 se ilustra las reglas extraídas por el algoritmo *JRip* para el escenario de eventos desbalanceados y balanceados.

```

Test mode:      10-fold cross-validation

=== Classifier model (full training set) ===

JRIP rules:
=====

(Desv_Long_Elect <= 0.089704) => Causa=Falla Balanceada      (61.0/0.0)
(Desv_Long_Elect <= 0.116368) => Causa=Falla Balanceada      (2.0/0.0)
=> Causa=Falla Desbalanceada    (422.0/0.0)

Number of Rules : 3

Time taken to build model: 0.06 seconds

```

*Figura 25. Tres reglas de decisión arrojadas por el algoritmo JRip para la clase eventos desbalanceados y balanceados.*

Para comparar los resultados obtenidos por las reglas y seleccionar aquellas con mejor cobertura y precisión, se necesita conocer la matriz de confusión de la prueba estadística. Esta matriz es una representación de los resultados de clasificación, mostrando las diferencias entre la clase verdadera y la estimada para un conjunto de observaciones que están previamente etiquetadas.

La diagonal principal de dicha matriz corresponde a los elementos que ha acertado el clasificador y los demás son errores, es decir elementos que no han sido clasificados correctamente. Las columnas indican las categorías clasificadas por el clasificador y las filas son las categorías reales de los datos.

En la Figura 26 se muestra la matriz de confusión para la clase evento desbalanceado y balanceado citado en el ejemplo anterior.

```

=== Confusion Matrix ===

  a  b  <-- classified as
422  0 |  a = Falla Desbalanceada
  1 62 |  b = Falla Balanceada

```

*Figura 26. Matriz de confusión para las reglas del algoritmo JRip, evento desbalanceado y balanceado.*

En la Figura 26, la suma de los términos de la primera fila de la matriz de confusión son el total de eventos desbalanceados e igualmente la suma de los términos de la segunda fila corresponde a los eventos balanceados. Así, solo los elementos de la diagonal principal son los eventos clasificados correctamente por el algoritmo, habiendo solo un evento (1) de los sesenta y tres (63) balanceados que fue clasificado como evento desbalanceado, reflejándose en un porcentaje de clasificación del 98,41%.

## **5.2 ALGORITMOS DE LA HERRAMIENTA COMPUTACIONAL Orange Canvas**

En la Figura 27 se ilustran las reglas extraídas por el algoritmo CN2 para el escenario de eventos desbalanceados y balanceados, en donde se aprecia la longitud de cada regla, su calidad, su respectiva clase y cobertura.

La *longitud* de la regla es el número de descriptores que emplea la regla dentro de la conjunción. La *calidad* de la regla se define como la utilidad que tienen los descriptores empleados que forman dicha regla y que provocan que esta misma cumpla con el objetivo planteado. La *cobertura* de una regla es el número de observaciones o descriptores para las cuales ella predice correctamente.

Length	Quality	Coverage	Class	Distribution	Rule
1	0.998	422.0	Falla Desbalanceada	<0.0,422.0>	IF Desv_Long_Elect>0.116 THEN Causa=Falla Desbalanceada
1	0.985	63.0	Falla Balanceada	<63.0,0.0>	IF Desv_Long_Elect<=0.116 THEN Causa=Falla Balanceada

Figura 27. Dos reglas de decisión arrojadas por el algoritmo CN2 para el escenario de eventos desbalanceados y balanceados.

### 5.3 RESULTADOS DE LOS ALGORITMOS DE EXTRACCIÓN DE REGLAS DE DECISIÓN

En las siguientes tablas se muestran los resultados de los algoritmos de extracción de reglas de decisión para cada uno de los escenarios descritos en el capítulo anterior. En dichas tablas, no aparecen todos los descriptores mencionados durante el análisis MANOVA, debido a que los descriptores extraídos por los algoritmos de extracción de reglas toman aquellos que maximizan la cobertura y calidad de las reglas.

Que algún descriptor no haya sido incluido en las reglas no implica que sea inadecuado, por el contrario, implica que existe otro que es capaz de cubrir con mejor calidad la misma cantidad o más eventos.

En estas tablas se muestran datos importantes que sirven para seleccionar reglas de decisión robustas para el diagnóstico de eventos de tensión de acuerdo a su causa y localización.

Luego se seleccionan las reglas de decisión a partir de la comparación de los resultados obtenidos con los diferentes algoritmos en cada uno de los escenarios. Seguidamente se analiza la confiabilidad de cada una de estas reglas, observando los puntos de corte de cada descriptor en cada una de las reglas seleccionadas por los algoritmos.

En la Figura 28 se muestra la metodología propuesta para el diagnóstico de los eventos de las Tablas VI, VII y VIII, tomando las mejores reglas extraídas.

De igual manera se ilustra esta metodología para los eventos multiestado y monoestado (Tabla IX), ver Figura 29, e igualmente para la localización relativa (Tabla X) en la Figura 30.

Tabla 6. Resultados de los algoritmos de extracción de reglas para la clase: eventos desbalanceados y balanceados

ALGORITMO	TOTAL DE HUECOS (Instancias)	TOTAL HUECOS POR CLASE DE EVENTO	COBERTURA (Huecos clasificados por clase)	CALIDAD DE LA REGLA	CLASE DE EVENTO	LONGITUD DE LA REGLA	REGLA DE DECISIÓN
Conjuntive Rule	485	422	422	100%	Evento desbalanceado	1	SI $\sigma\delta L > 0.131022$ ENTONCES Evento Desbalanceado SINO Evento Balanceado
		63	63	100%	Evento balanceado		
JRip	485	422	422	100%	Evento desbalanceado	1	SI $\sigma\delta L \leq 0.116368$ ENTONCES Evento Balanceado SINO Evento Desbalanceado
		63	62	98%	Evento balanceado		
OneR	485	422	422	100%	Evento desbalanceado	1	SI $\sigma\delta L \geq 0.13102$ ENTONCES Evento Desbalanceado SINO Evento Balanceado
		63	63	100%	Evento balanceado		
PART	485	422	422	100%	Evento desbalanceado	1	SI $\sigma\delta L > 0.116368$ ENTONCES Evento Desbalanceado SINO Evento Balanceado
		63	62	98%	Evento balanceado		
Ridor	485	422	422	100%	Evento desbalanceado	1	SI $\sigma\delta L \leq 0.11769$ ENTONCES Evento Balanceado SINO Evento Desbalanceado
		63	63	100%	Evento balanceado		
CN2	485	422	422	100%	Evento desbalanceado	1	SI $\sigma\delta L > 0.116368$ ENTONCES Evento Desbalanceado SINO Evento Balanceado
		63	63	100%	Evento balanceado		

En la Tabla VI se ilustran todas las reglas de decisión extraídas por los algoritmos de decisión para el escenario eventos desbalanceados y balanceados. Todas estas reglas coinciden en la misma longitud, es decir contienen la misma cantidad de descriptores en las reglas extraídas. Adicionalmente todos los algoritmos extrajeron el mismo descriptor  $\sigma\delta L$ . Por tanto, se puede afirmar que el descriptor  $\sigma\delta L$  es idóneo para discriminar eventos balanceados y desbalanceados.

El punto de corte de los algoritmos Conjunctive Rule y OneR (0,13102) es el mismo para una cobertura total de todos los datos. En cuanto a los algoritmos JRip (0,11636), PART (0,11636), Ridor (0,1176) y CN2 (0,1163), presentan un punto de corte menor y muy cercano a los mencionados inicialmente, pero no erróneo, debido a que la regla cumple tiene una cobertura y calidad adecuada, es decir que aunque dicho punto de corte no coincide, son útiles para realizar el diagnóstico de eventos de tensión.

Tabla 7. Resultados de los algoritmos de extracción de reglas para la clase: eventos desbalanceados

ALGORITMO	TOTAL DE HUECOS (Instancias)	TOTAL HUECOS POR CLASE DE EVENTO	COBERTURA (Huecos tomados por evento)	CALIDAD DE LA REGLA	CLASE DE EVENTO	LONGITUD DE LA REGLA	REGLA DE DECISIÓN
Conjunctive Rule	422	203	203	100%	Falla monofásica	1	SI $RCV > 0.864599$ ENTONCES Falla Monofásica SINO Falla Bifásica
		208	208	100%	Falla bifásica		
		11	0	0%	Energización de un transformador		
JRip	422	203	136	67%	Falla monofásica	1	SI $RCV > 0.92154$ ENTONCES Falla Monofásica SINO Falla Bifásica
		208	132	64%	Falla bifásica		
		11	9	82%	Energización de un transformador	1	SI $RCV \leq 0.02631$ ENTONCES Energización de un Transformador SINO Falla Bifásica
OneR	422	203	203	100%	Falla monofásica	1	SI $RCV > 0.864599$ ENTONCES Falla Monofásica
		208	208	100%	Falla bifásica	1	SI $RCV < 0.864599$ ENTONCES Falla Bifásica
		11	11	100%	Energización de un transformador	1	SI $RCV < 0.24176$ ENTONCES Energización de un Transformador
PART	422	203	203	100%	Falla monofásica	1	SI $PNF > 0.859$ ENTONCES Falla Monofásica
		208	208	100%	Falla bifásica	2	SI $PNF > 0.22224$ Y $PNF \leq 0.7384$ ENTONCES Falla Bifásica
		11	10	91%	Energización de un transformador	1	SI $CF > 0.613312$ ENTONCES Energización de un Transformador

Ridor	422	203	203	100%	Falla monofásica	1	SI <b>RCV</b> $\leq 0.8645$ ENTONCES Falla Bifásica SINO Falla Monofásica
		208	208	100%	Falla bifásica	1	SI <b>RCV</b> $\leq 0.864592$ ENTONCES Falla Bifásica
		11	10	91%	Energización de un transformador	1	SI <b>CF</b> $> 0.605565$ ENTONCES Energización de un Transformador
CN2	422	203	129	64%	Falla monofásica	2	SI $\delta L \leq 0.2524041$ Y <b>maxVo</b> $> 0.0276782$ ENTONCES Falla Monofásica
		208	163	78%	Falla bifásica	2	SI $\delta L > 0.847079$ Y $\Delta LV \leq 0.3419195$ ENTONCES Falla Bifásica
		11	10	91%	Energización de un transformador	2	SI <b>CF</b> $> 0.61331242$ Y <b>HighOscFree</b> $> 0.0012053$ ENTONCES Energización de un Transformador

En la Tabla VII se muestran las reglas de decisión extraídas por los algoritmos de conocimiento para el escenario desbalanceados. Los algoritmos Conjunctive Rule, OneR y Ridor tienen reglas con una misma cobertura del 100%. En este caso el descriptor *RCV* proporciona información adecuada para clasificar o discriminar entre eventos monofásicos y bifásicos.

Para el caso de las reglas obtenidas por el algoritmo JRip, el punto de corte estimado para el descriptor *RCV* no coincide con los demás, afectando la cobertura y por tanto la calidad de la regla. Entonces se descartan las reglas obtenidas por este algoritmo debido a su baja calidad.

El algoritmo PART extrae reglas combinando los descriptores *PNF* y *CF*, en donde la calidad de la regla es alta y confiable. Este algoritmo tiene en cuenta el descriptor *PNF* ideal para la clasificación de fallas monofásicas y bifásicas cumpliendo con el punto de corte especificado teóricamente. Igualmente para el descriptor *CF* se tiene un intervalo de corte capaz de identificar eventos por energización de transformadores.

Los intervalos de corte de las reglas extraídas por el algoritmo CN2 para el caso de la identificación de fallas monofásicas tienen una calidad del 64%, lo cual corresponde a un bajo porcentaje. Esto es debido a que el punto de corte del descriptor *MaxVo* en esta regla no es el adecuado para discriminar las fallas monofásicas de las bifásicas, porque que este intervalo no es lo suficientemente bueno para distinguir de forma adecuada los valores estimados por este descriptor.

Tabla 8. Resultados de los algoritmos de extracción de reglas para la clase: eventos balanceados

ALGORITMO	TOTAL DE HUECOS (Instancias)	TOTAL HUECOS POR CLASE DE EVENTO	COBERTURA (Huecos tomados por evento)	CALIDAD DE LA REGLA	CLASE DE EVENTO	LONGITUD DE LA REGLA	REGLA DE DECISIÓN
Conjuntive Rule	63	25	9	36%	Falla trifásica	1	SI $FPratio\_cruce > 1.014881$ ENTONCES Energización de un Banco de Capacitores
		28	18	64%	Energización de un banco de capacitores		
		10	8	80%	Arranque de motores		
JRip	63	25	20	80%	Falla trifásica	1	SI $FPratio\_cruce = 1.000112$ ENTONCES Falla trifásica
		28	25	89%	Energización de un banco de capacitores	1	SI $FPratio\_cruce >= 1.01959$ ENTONCES Energización de un Banco de Capacitores
		10	9	90%	Arranque de motores	1	SI $Iratio >= 1.575331$ ENTONCES Arranque de Motores
PART	63	25	19	76%	Falla trifásica	1	SI $FPratio\_cruce = 1.000211$ ENTONCES Falla Trifásica
		28	26	93%	Energización de un banco de capacitores	2	SI $Vratio > 0.99954$ Y $Iratio <= 0.99646$ ENTONCES Energización de un Banco de Capacitores
		10	9	90%	Arranque de motores	1	SI $Vratio <= 0.99954$ ENTONCES Arranque de Motores

Ridor	63	25	18	72%	Falla trifásica	1	SI $Iratio \leq 1.36718$ ENTONCES Falla Trifásica SINO Arranque de Motores
		28	25	89%	Energización de un banco de capacitores	2	SI $FPratio_{cruce} > 1.024622$ Y $Iratio \leq 0.99667$ ENTONCES Energización de un Banco de Capacitores
		10	9	90%	Arranque de motores	1	SI $Iratio < 0.997744$ ENTONCES Energización de un banco de capacitores SINO Arranque de motores
CN2	63	25	21	75%	Falla trifásica	1	SI $Sratio = 0.968$ ENTONCES Falla Trifásica
		28	23	82%	Energización de un banco de capacitores	1	SI $FPratio_{cruce} > 1.02065$ Y $\Delta Q > 0.023$ ENTONCES Energización de un Banco de Capacitores
		10	9	90%	Arranque de motores	1	SI $Iratio > 1.159035$ ENTONCES Arranque de Motores

En la Tabla VIII se ilustran las reglas de decisión extraídas por los algoritmos de conocimiento para los eventos balanceados.

Para el algoritmo Conjunctive Rule, la regla asociada para las tres clases de eventos mostrados, presenta diferentes grados de cobertura respecto a cada clase de evento.  $FPratio_{cruce}$  no puede discriminar para fallas trifásicas y con un resultado no confiable para energización de banco de capacitores, pero si tiene buena calidad para identificar arranque de motores. Al compararlo con los algoritmos Ridor y CN2, se puede concluir que el punto de corte para  $FPratio_{cruce}$  es muy similar para identificar energización de bancos de capacitores, siendo las reglas extraídas por Ridor y CN2 apropiadas para este fin.

En los algoritmos JRip y PART en la identificación de fallas trifásicas, se aprecia que el descriptor  $FPratio_{cruce}$  detecta para este tipo de evento valores cercanos a la unidad, debido a que no existen cambios de carga en el sistema.

El algoritmo PART y Ridor extraen reglas con buena cobertura y calidad que relacionan el descriptor  $Vratio$  e  $Iratio$ . Se observa que estos descriptores son capaces de caracterizar cambios de carga debido a la energización de bancos de capacitores y arranque de motores, porque el punto de corte que emplea para la respectiva determinación es el adecuado. Es decir, en el caso de la energización

de capacitores el incremento en la tensión se refleja en un  $Vratio$  mayor que la unidad, y para el caso de  $Iratio$  menor que la unidad.

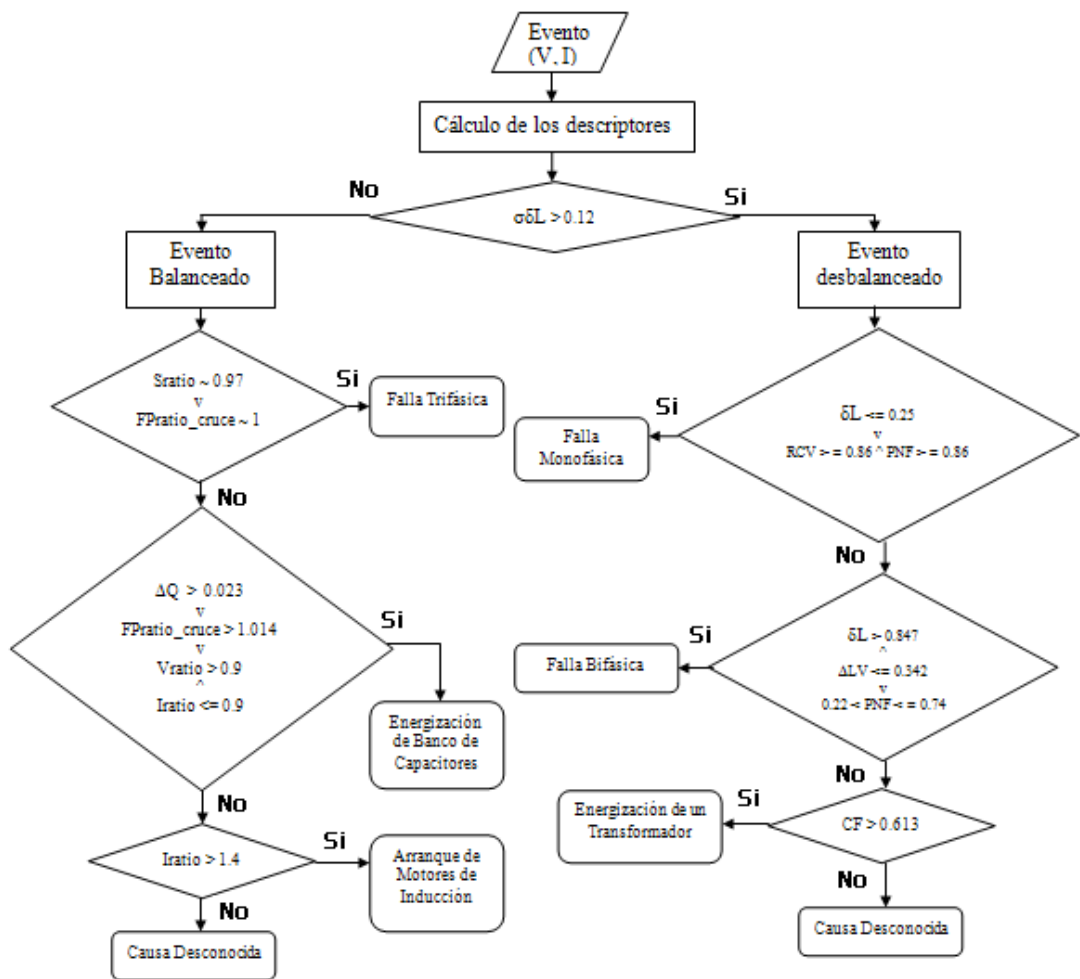


Figura 28. Pasos propuestos para el diagnóstico de los eventos que involucran el grado de desbalance

Tabla 9. Resultados de los algoritmos de extracción de reglas para la clase: eventos monoestado y multiestado

ALGORITMO	TOTAL DE HUECOS (Instancias)	TOTAL HUECOS POR CLASE DE EVENTO	COBERTURA (Huecos tomados por evento)	CALIDAD DE LA REGLA	CLASE DE EVENTO	LONGITUD DE LA REGLA	REGLA DE DECISIÓN
Conjuntive Rule	473	436	429	98%	Falla monoestado	1	SI $\sigma E \leq 0.239749$ ENTONCES Falla monoestado SINO Falla multiestado
		37	31	84%	Falla multiestado		
JRip	473	436	429	98%	Falla monoestado	1	SI $\sigma E \geq 0.240654$ ENTONCES Falla multiestado SINO Falla monoestado
		37	30	81%	Falla multiestado		
OneR	473	436	427	98%	Falla monoestado	1	SI $\sigma E < 0.57209$ ENTONCES Falla monoestado SINO Falla multiestado
		37	30	81%	Falla multiestado		
PART	473	436	429	98%	Falla monoestado	1	SI $\sigma E \leq 0.238845$ ENTONCES Falla monoestado
		37	27	73%	Falla multiestado	1	SI $\sigma E > 0.557234$ ENTONCES Falla multiestado
Ridor	473	436	432	99%	Falla monoestado	2	SI $\sigma E > 0.239119$ Y $\sigma E > 0.5721$ ENTONCES Falla multiestado SINO Falla monoestado
		37	28	77%	Falla multiestado		
CN2	473	436	279	65%	Falla monoestado	2	SI $\sigma E \leq 0.238845$ Y $Segmentos(\delta=8, m=64) < 5.0$ ENTONCES Falla monoestado
		37	26	71%	Falla multiestado	2	SI $Segmentos(\delta=8, m=64) > 4.0$ Y $\sigma E > 0.557233$ ENTONCES Falla multiestado

En la Tabla IX se muestran las reglas de decisión extraídas por los algoritmos de conocimiento para los eventos monoestado y multiestado.

Los algoritmos Conjuntive Rule, JRip, PART y Ridor extraen reglas con alta calidad y por tanto cubren la mayoría de observaciones relacionadas con el descriptor  $\sigma E$ . Su punto de corte es similar en todas las reglas.

El algoritmo CN2 incluye en sus reglas los descriptores  $\sigma E$  y  $Segmentos_{\delta=8, m=64}$ . Respecto a la identificación de fallas monoestado, se aprecia que la cobertura

para esta regla no es tan alta, sin embargo se decide tomar su veracidad porque los puntos de corte indican claramente su validez en el diagnóstico.

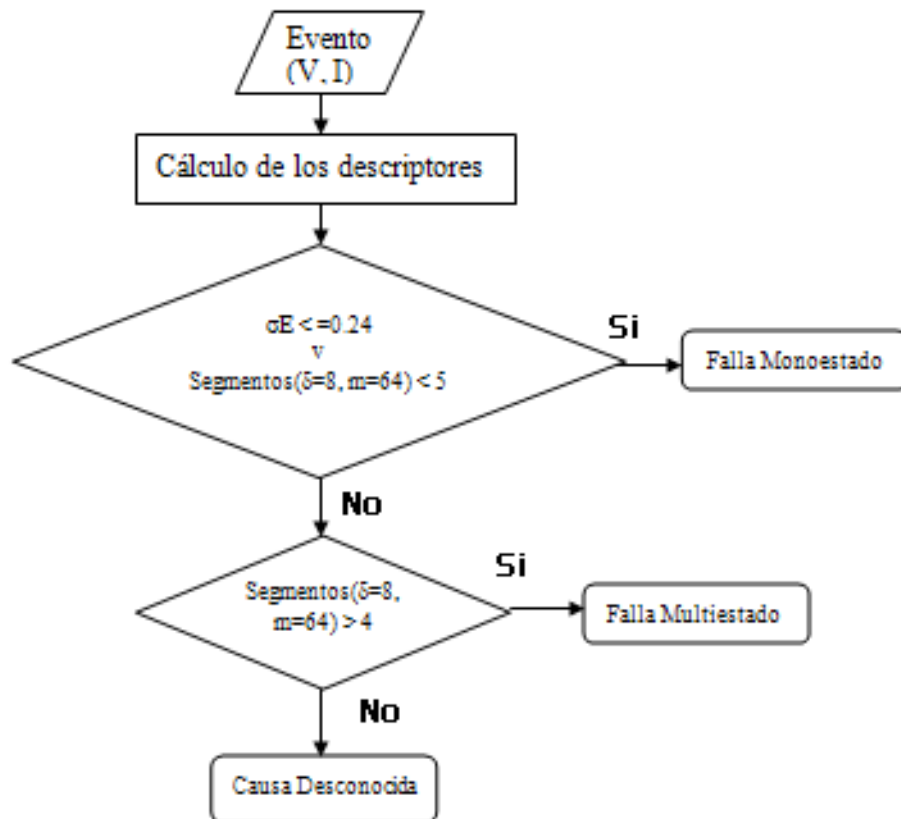


Figura 29. Pasos propuestos para el diagnóstico de la causa para eventos multiestado y monoestado

Tabla 10. Resultados de los algoritmos de extracción de reglas para la clase: localización aguas arriba y aguas abajo

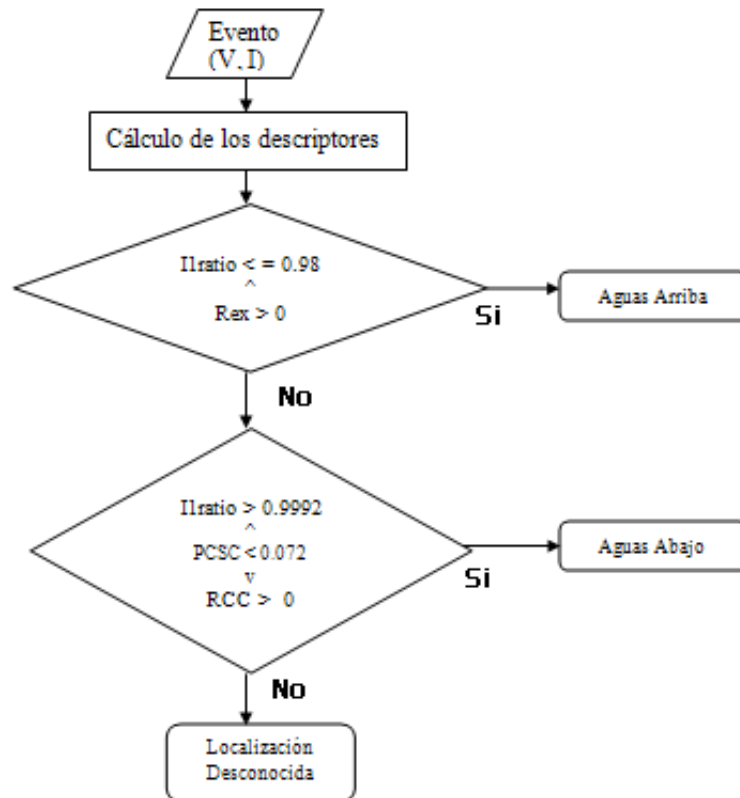
ALGORITMO	TOTAL DE HUECOS (Instancias)	TOTAL HUECOS POR CLASE DE EVENTO	COBERTURA (Huecos tomados por evento)	CALIDAD DE LA REGLA	CLASE DE EVENTO	LONGITUD DE LA REGLA	REGLA DE DECISIÓN
Conjuntive Rule	484	257	247	96%	Aguas arriba	1	Si <b>lratio</b> <= <b>0.987821</b> ENTONCES Aguas arriba SINO Aguas abajo
		227	219	96%	Aguas abajo		
JRip	484	257	243	95%	Aguas arriba	2	Si <b>RCC</b> >= <b>14.67612</b> Y <b>lratio</b> <= <b>0.803981</b> ENTONCES Aguas abajo SINO Aguas arriba
		227	219	96%	Aguas abajo	2	Si <b>lratio</b> >= <b>0.996286</b> Y <b>RCC</b> >= <b>0</b> ENTONCES Aguas abajo
OneR	484	257	245	95%	Aguas arriba	1	Si <b>lratio</b> < <b>0.988850</b> ENTONCES Aguas arriba SINO Aguas abajo
		227	218	96%	Aguas abajo		
PART	484	257	244	95%	Aguas arriba	3	Si <b>lratio</b> <= <b>0.981414</b> Y <b>RCC</b> <= <b>34.150345</b> Y <b>Res</b> > <b>0</b> ENTONCES Aguas arriba
		227	219	96%	Aguas abajo	2	Si <b>PCSC</b> <= <b>0.069869</b> Y <b>RCC</b> > <b>6.663819</b> ENTONCES Aguas abajo
Ridor	484	257	247	96%	Aguas arriba	2	Si <b>PCSC</b> <= <b>0.072408</b> Y <b>RCC</b> > <b>37.6818</b> ENTONCES Aguas abajo SINO Aguas Arriba
		217	217	96%	Aguas abajo	3	Si <b>lratio</b> > <b>0.999251</b> Y <b>PCSC</b> <= <b>-0.035653</b> Y <b>lratio</b> <= <b>2.242781</b> ENTONCES Aguas abajo SINO Aguas arriba

CN2	484	257	226	88%	Aguas arriba	3	SI $I_{ratio} \leq 0.981414$ Y $RCC \leq 14.226112$ Y $I_{ratio} > 0.3240797$ ENTONCES Aguas arriba
		227	190	84%	Aguas abajo	3	SI $PCSC \leq -0.03686$ Y $I_{ratio} \leq 2.221087$ Y $I_{ratio} > 0.542091$ ENTONCES Aguas abajo

En la Tabla X indica las reglas de decisión extraídas por los algoritmos de conocimiento respecto a la localización relativa de los eventos de tensión.

En el algoritmo JRip en el escenario de localización relativa para eventos aguas abajo, los descriptores  $I_{ratio}$  y  $RCC$  contienen la suficiente información para identificar eventos aguas abajo. La calidad de esta regla es similar a las arrojadas por otros algoritmos, pero esta tiene un grado de confianza superior, debido a que el punto de corte para el descriptor  $RCC$  cubre los valores estimados en las demás reglas.

El algoritmo PART es el único que involucra al descriptor  $Rex$ , con un alto porcentaje de cobertura para la regla extraída respecto a la localización relativa para los eventos aguas arriba.



*Figura 30. Pasos propuestos para el diagnóstico de la localización relativa*

Las anteriores metodologías mostradas en las figuras 28, 29 y 30, dan las pautas necesarias para el diagnóstico de las causas y la localización relativa de los eventos de tensión.

Una vez propuesta la metodología para el diagnóstico de la causa y origen de un evento de tensión, se realizan algunas pruebas con los resultados obtenidos en los programas de extracción de reglas para validar dichas reglas en la estimación de la causa y localización relativa de perturbaciones eléctricas.

#### **5.4 VALIDACIÓN DE ALGUNAS REGLAS QUE DIAGNOSTICAN LA CAUSA DE LOS EVENTOS DE TENSIÓN**

Algunas reglas son evaluadas para verificar su validez al momento de estimar los eventos de tensión tabulados en la Tabla XI. Se extraen 63 perturbaciones de la base de ENDESA para la identificación de cambios de carga que corresponden a energización de bancos de capacitores, transformadores y arranque de motores de inducción. Igualmente se emplean 422 perturbaciones para evaluar eventos monofásicos, bifásicos y trifásicos.

Tabla 11. Resultados de la clasificación obtenida por algunas reglas que diagnostican eventos de tensión

TIPO DE EVENTO	ENERGIZACIÓN BANCOS DE CAPACITORES	ARRANQUE DE MOTORES DE INDUCCIÓN	ENERGIZACIÓN DE TRANSFORMADORES	FALLA BIFÁSICA	FALLA MONOFÁSICA	FALLA TRIFÁSICA
Verdaderos positivos (VP)	21	8	18	163	129	20
Falsos positivos (FP)	7	2	10	45	74	5
Verdaderos negativos (VN) & Falsos negativos (FN)	35	53	35	214	219	397
Total de eventos	63	63	63	422	422	422

Donde:

- VP, Verdaderos positivos. Casos correctamente clasificados y que pertenecen a la clase de referencia.
- VN, Verdaderos negativos. Casos correctamente clasificados y que pertenecen a la clase no referencia.
- FP, Falsos positivos. Casos clasificados como pertenecientes a la clase referencia pero su clase real es la clase no referencia.
- FN, Falsos negativos. Casos clasificados como pertenecientes a la clase no referencia pero su clase real es la clase referencia [7].

Se evalúa el comportamiento de algunas reglas extraídas por el algoritmo Conjunctive Rule para la identificación de cambios de carga. Se emplean 28 eventos correspondientes a energización de capacitores, de los cuales 21 son clasificados correctamente y 7 eventos no causados por energización de un capacitor fueron clasificados a este hecho. La cobertura y calidad de la regla es adecuada, es decir, es útil para realizar un diagnóstico de este evento de tensión.

En cuanto a los eventos debidos al arranque de motores de inducción, energización de transformadores, fallas de red monofásica, bifásica y trifásica, se presentan eventos que son mal clasificados, pero no representan una cantidad significativa con respecto al total de eventos. Este hecho indica que las reglas obtienen buenos resultados en cuanto a la identificación del grado de desbalance de los eventos de tensión.

La energización de un transformador es detectada mediante el *CF*, este descriptor determina la tendencia triangular de una señal cuando se energiza un transformador. La cobertura y calidad de la regla para este caso no permiten hacer una buena estimación de esta clase de evento debido a que no se cuenta con los puntos de cortes ideales para la respectiva determinación.

Una falla monofásica y bifásica fácilmente puede identificarse con la combinación de los descriptores  $\delta L$ ,  $MaxV_0$ ,  $minPNF$  y  $minRCV$ . Según los valores de cobertura establecidos, se observa que las reglas asociadas a estos eventos no muestran una buena clasificación ya que la regla empleada para este hecho no cuenta con la suficiente calidad para dicha estimación.

Algunas reglas son evaluadas para verificar su validez al momento de identificar eventos multiestado de monoestado. Los resultados obtenidos son tabulados en la Tabla XII. Se extraen 473 perturbaciones de la base de ENDESA para la identificación de dichos eventos.

Tabla 12. Resultados de la clasificación obtenida por algunas reglas que identifican eventos monoestado y multiestado

TIPO DE EVENTO	FALLA MONOESTADO	FALLA MULTIESTADO
Verdaderos positivos (VP)	427	26
Falsos positivos (FP)	9	11
Verdaderos negativos (VN) & Falsos negativos (FN)	37	436
Total de eventos	473	473

Una manera de determinar si una falla es multiestado o monoestado, es detectando el número de segmentos transitorios, las reglas extraídas para este hecho presentan dos alternativas para detectar los cambios de estado transitorio mediante la segmentación. En la Tabla XII se registran 427 eventos monoestado y 26 eventos multiestado correctamente clasificados. Los resultados de clasificación se consideran adecuados, debido a que la estimación del número de segmentos transitorios fue realizada a partir de dos reglas, una basada en  $\sigma E$  y otra en *Segmentos* <sub>$\delta=8, m=64$</sub> .

Algunas reglas extraídas para el diagnóstico de la localización relativa son evaluadas mediante los resultados tabulados en la Tabla XIII. Se emplean 484 eventos de tensión, provenientes de la base de datos de ENDESA para analizar que tan adecuada es la cobertura y calidad de las reglas al momento de realizar la respectiva clasificación aguas arriba o aguas abajo.

Tabla 13. Resultados de la clasificación obtenida por algunas reglas que diagnostican la localización relativa

TIPO DE EVENTO	AGUAS ARRIBA	AGUAS ABAJO
Verdaderos positivos (VP)	243	190
Falsos positivos (FP)	14	37
Verdaderos negativos (VN) & Falsos negativos (FN)	227	246
Total de eventos	484	473

Se puede ver en los resultados de la Tabla XIII, 243 eventos clasificados correctamente como aguas arriba, 14 mal clasificados al ser registrados como aguas arriba cuando no lo son. Por otro lado, 190 pertenecen correctamente a huecos de tensión cuya localización relativa es aguas abajo y 37 se tabulan como mal clasificados, por aparecer registrados a esta clase sin en verdad pertenecer a ella. La calidad de la regla aunque no es la adecuada sirve para una buena estimación de la localización relativa de los eventos de tensión.

## 6. CONCLUSIONES

Se presentan a continuación algunos aspectos puntuales que enmarcan el desarrollo de este trabajo de grado, evidenciando cada una de las etapas cubiertas con sus respectivos resultados. Igualmente se proponen recomendaciones para posteriores estudios relacionados con el tema de los eventos de tensión.

- En el análisis bibliográfico y estado del arte de los algoritmos de extracción de reglas se revisaron aquellos que crean reglas relacionando varios descriptores con un mínimo de error, que cubren el mayor número de descriptores, tienen alta precisión cuando predicen la clase, tienen mayor cobertura y relacionan las mismas técnicas heurísticas.
- La selección de los descriptores se hizo de acuerdo a su relación con el análisis de cada clase propuesta para la localización relativa y la causa de los eventos de tensión, evitando la introducción de ruido al estimar los valores de  $R^2$ -corregida y significancia en el análisis multivariable realizado por MANOVA.
- La calidad de cada descriptor es estimada de acuerdo al análisis estadístico, mostrando la información suministrada por cada descriptor propuesto en las diferentes clases que evalúan la localización relativa y causa de los eventos de tensión.
- Se contó con un conjunto de perturbaciones reales registradas en subestaciones de la empresa de distribución de energía eléctrica ENDESA. De forma adicional se obtuvieron por simulación algunos registros de huecos de tensión a partir de un sistema eléctrico tomado del artículo IEEE [32].

- El descriptor  $GVo$  que estima el grado de desbalance de las diferentes perturbaciones, no muestra una buena clasificación en el análisis estadístico, debido a que las señales presentan diferente duración en la evolución del hueco de tensión, distorsionándose así la tendencia marcada para este descriptor evidenciada por la respectiva conceptualización.
- En los registros suministrados por ENDESA algunas fallas de red tienen altas frecuencias de oscilación en la señal de tensión, que son estimadas por el descriptor  $HighOscFrec$ . Esta apreciación es un caso particular del sistema de potencia ya que existen tramos subterráneos, cuya impedancia capacitiva es importante, produciendo un transitorio durante el despeje de la falla.
- Los descriptores  $Varc_{max}$  y  $FIPA$  son excluidos del análisis de este trabajo de grado ya que sólo discriminan fallas en el sistema debido a contactos de animales, ramas o caída de un rayo, y no existe un previo seguimiento para estos tipos de causas en la base de datos suministrada por ENDESA.
- Las reglas arrojadas por el algoritmo JRip y que son validadas por WEKA presentan una estimación errónea al decidir el tipo de causa en la clase desbalanceada, ya que arroja puntos de corte establecidos por la regla que no corresponden a la respectiva conceptualización del evento ocurrido.
- Las reglas obtenidas con el algoritmo PART tienen una correcta estimación en la decisión del tipo de desbalance del evento, ya que estas reglas establecen rangos de confianza que corresponden con la causa descrita por la literatura.
- Las reglas obtenidas por los diferentes algoritmos emplean pocos descriptores, las cuales muestran resultados adecuados para el análisis

previo MANOVA. Suministrando información relevante a la hora de estimar eventos de tensión.

- Se presentó una metodología genérica de eventos de tensión a manera de ejemplo para ilustrar de forma práctica la utilidad de los descriptores y reglas previamente clasificadas, y así facilitar la interpretación de los resultados obtenidos en este trabajo de grado.
- La validación de algunas reglas obtenidas para cada escenario, fue realizada mediante los resultados de los programas de extracción de reglas WEKA y Orange Canvas.
- Como futuros desarrollos se propone validar las metodologías mencionadas en el Capítulo 5 a través de la herramienta computacional MATLAB para automatizar el proceso que estima la causa y la localización relativa de la falla.

Los resultados obtenidos en el escenario de localización relativa, son el punto de partida para estudios de localización precisa, en la que se pretende estimar la distancia en la que ocurre la falla respecto al equipo de medición.

Todos los resultados en este proyecto complementan los estudios de monitorización de huecos de tensión en sistemas de distribución de energía eléctrica llevados a cabo en la Universitat de Girona en España.

- Este trabajo de grado aporta a los autores estrategias que permiten estudiar la calidad de la energía eléctrica mediante información extraída de datos reales, relacionando distintos conocimientos adquiridos durante el proceso de formación académica que facilitan entender y dar explicación a hechos presentes en la industria.

## BIBLIOGRAFÍA

- [1] J. Stones, A. Collinson. "Power quality" Power Engineering Journal, Volume: 15 Issue: 2, pp. 58 –64, 2001
- [2] K. J. Cornick y H. Q. Li. "Power Quality and Voltage Dips". IEEE, Development Department Singapore Power Systems Lt, vol. 1, pp. 149-156, Oct/Nov 2000.
- [3] Seon-Ju Ahn , Dong-Jun Won, D-Yop Chung y Seung-U Moon. "Determination of the Relative Location of Voltage Sag Source According to Event Cause".IEEE, Power Engineering Society G.M, vol.1, pp. 620-625, June 2004.
- [4] K. Kim, J. Park, J. Lee, S. Ahn y S. Moon. "A Method to Determine the Relative Location of Voltage Sag Source for PQ Diagnosis". IEEE. Seoul National University, Korea. Págs. 2192-2197.
- [5] M.H.J. Bollen y L. Zhan. "Characteristic of Voltage Dips (Sags) in Power System". IEEE Trans. Power Del. vol. 15, no. 2, pp. 827-832, Apr. 2000.
- [6] ZHANG, Lidong y BOLLEN, Math H. J. "A Method for Characterisation of Three-Phase Unbalanced Dips from Recorded Voltage Waveshapes".IEEE.Telecommunications Energy Conference, Chalmers of Electric of Technology, 1999.
- [7] V. Barrera, X. Berjaga, G. Ordonez y J.Melendez. "Estrategias para la localización relativa de la fuente de huecos de tensión en sistemas de distribución de energía eléctrica". Revista de Ingeniería de la Univiversidad de Antioquia (UdeA), 2007.
- [8] Hernandez J, Quintana M<sup>a</sup>.J, Ramírez C. Introducción a la minería de datos. Pearson Prentice Hall. (2004), cap 9, pp 237-254.
- [9] J. Cendrowska. PRISM: An algorithm for inducing modular rules. International Journal of Man-Machine Studies 27:4 (1987), pp. 349-370.
- [10] R.C. Holte. Very simple classification rules perform well on most commonly used datasets. Machine Learning 11 (1993), pp. 63-91.
- [11] R. S. Michalksi. On the quasi-minimal solution of the general covering problem. Proceedings of the First International Simposium of Information Processing (Bled. Yugoslavia), A3 (1969), pp. 125-128.
- [12] Peter Clark and Tim Niblett. The CN2 induction algorithm. Machine Learning Journal 3:4 (1989), pp. 261-283.

- [13]. W.W. Cohen. Fast Effective Rule Induction. Machine Learning: Proceedings of the Twelfth International Conference. Lake Tahoe California (United States of America, 1995), pp. 1-10.
- [14] William W. Cohen. Fast Effective Rule Induction. In: Twelfth International Conference on Machine Learning, 115-123, 1995.
- [15] R. Agrawal, R. Srikant: Fast Algorithms for Mining Association Rules in Large Databases. In: 20th International Conference on Very Large Data Bases, 478-499, 1994.
- [16] Bing Liu, Wynne Hsu, Yiming Ma: Integrating Classification and Association Rule Mining. In: Fourth International Conference on Knowledge Discovery and Data Mining, 80-86, 1998.
- [17] Brian R. Gaines, Paul Compton (1995). Induction of Ripple-Down Rules Applied to Modeling Large Databases. *J. Intell. Inf. Syst.* 5(3):211-228.
- [18] A.K Pradhan, A. Routray, S. Madhan, "Fault direction estimation in radial distribution system using phase change in sequence current", *IEEE TPWDelivery*, vol. 22, No. 4, pp. 2065-2071, october 2007.
- [19] C. Li, T. Tayjasanant, W. Xu, and X. Liu, "Method for voltage-sag source detection by investigating slope of the system trajectory," *IEEE Proceedings-Generation, Trans. and Distr.*, vol. 150, pp. 367-72, 2003.
- [20] V. Barrera, X. Berjaga, J. Melendez, S. Herraiz, J. Sánchez and M. Castro, "Two New Methods for Voltage Sag Source Location", *ICHQP 2008 - 13th International Conference on Harmonics & Quality of Power*, 28th September – 1st October, Australia.
- [21] P. K. Dash, A. K. Pradhan, G. Panda, and A. C. Liew, "Adaptive relay setting for Flexible AC Transmission Systems (FACTS)," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 38–43, Jan. 2000.
- [22] N. Hamzah, A. Mohamed and A. Hussain. "Locating Voltage Sag Source At The Point Of Common Coupling In Industrial Distribution Systems," *IEEE PEDS*, pp. 2–3, 2005.
- [23] J. Blanco, J. Jagua. "Metodología para el diagnóstico de la causa de huecos de tensión: análisis de fallas", Proyecto de grado Universidad Industrial de Santander, pp. 36-37, 2009.
- [24] M. Bollen, "Algorithms for Characterizing Measured three-Phase Unbalanced Voltage Sags", *IEEE Trans. Power Del.*, vol. 18, no. 3, July 2003.

- [25] M. Bollen, D. Sabin, "International Coordination for Voltage Sag Indices", PES TD 2005/2006, May 21-24, 2006, pp.:229 – 234.
- [26] V. Barrera, J. Meléndez, S. Herraiz, "Evaluation of Fault Relative Location Algorithms using Voltage Sag Data Collected at 25-kV Substations", Special Issue on Power Quality, European Transactions on Electrical Power (ETEP), DOI 10.1002/etep.393, October, 2009.
- [27] V. Barrera, I. Yu-Hua Gu, M Bollen, "Feature Characterization of Power Quality Events According to Their Underlying Causes", Septiembre, 2009, pp.1-8.
- [28] V. Barrera, S. Kulkarni, S. Santoso, J. Melendez, "Feature Analysis and Classification Methodology for Overhead Distribution Fault Events", July, 2009, pp. 1-8.
- [29] V. Barrera, J. Meléndez, S. Jaramillo, J. Sánchez, "Unusual Voltage Sag Event Detection in Power Systems", August, 2008, pp. 1-5.
- [30] Barker, H.R. and Barker B.M. "Multivariate Analysis of Variance (MANOVA) a Practical Guide to its use in Scientific Decision Making ". Alabama, 1984
- [31] Ordoñez Gabriel; "Perturbaciones en la onda de tensión: huecos y sobretensiones". Universidad Industrial de Santander. Colombia 2002.  
Disponible en Internet en: [eia.udg.es/~secse/huecos\\_de\\_tension\\_creg\\_alure\\_gop.pdf](http://eia.udg.es/~secse/huecos_de_tension_creg_alure_gop.pdf)
- [32] Radial Test Feeders - IEEE Distribution System Analysis Subcommittee.  
<http://ewh.ieee.org/soc/pes/dsacom/testfeeders.html>. Consultado Noviembre de 2009

## ANEXOS

### A. ALGORITMOS DE APRENDIZAJE

#### A.1 Algoritmo de aprendizaje de reglas por cobertura

Estos algoritmos generan un conjunto de reglas partiendo de un conjunto de ejemplos, utilizando la técnica de cobertura. Dicho algoritmo tiene la característica que añade condiciones (pero no un conjunto de condiciones alternativas) a las reglas. Ya que este algoritmo no deshace o rectifica las elecciones hechas, el criterio para seleccionar dichas condiciones, ya sea el número de ejemplos negativos excluidos, es relevante y se deben evaluar bien y a priori para ver que particiones son las mejores [8].

A manera de ejemplo se muestra el algoritmo citado como sigue:

```
ALGORITMO Cobertura(Epos, Eneg: conjunto de ejemplos)
Reglas := ∅
MIENTRAS Epos ≠ ∅ Y NO ParadaReglas HACER // Aprender una nueva regla
  NuevaRegla := ∅;
  Eneg_Act := Eneg;
  MIENTRAS Eneg_Act ≠ ∅ Y NO ParadaCondiciones HACER // Aprender una nueva condición.
    Cond := Seleccionar una condición según criterio (elimina muchos negativos)
    NuevaRegla := NuevaRegla ∪ { Cond } // Añadimos la nueva condición a la regla
    Eneg_Act := Ejemplos negativos consistentes con NuevaRegla;
  FIN MIENTRAS; // La regla está especializada suficientemente. No quedan ejemplos negativos.
  Reglas := Reglas ∪ { NuevaRegla };
  Epos := Epos – Ejemplos cubiertos por NuevaRegla
FIN MIENTRAS; // El conjunto de reglas ya cubre todos los ejemplos positivos.
RETORNA Reglas;
FIN ALGORITMO
```

*Figura A.1. Algoritmo de aprendizaje de reglas por “Cobertura”*

#### A.2 Algoritmos de aprendizaje de extracción de reglas de decisión

##### A.2.1 Algoritmo de aprendizaje de reglas PRIMS

En la siguiente figura se muestra el algoritmo de PRISM.

```

Para cada clase  $C$ 
  Sea  $E$  = ejemplos de entrenamiento
  Mientras  $E$  tenga ejemplos de clase  $C$ 
    Crea una regla  $R$  con LHS vacío y clase  $C$ 
    Until  $R$  es perfecta. do
      Para cada atributo  $A$  no incluido en  $R$  y cada valor  $v$ ,
        Considera añadir la condición  $A = v$  al LHS de  $R$ 
        Selecciona el par  $A = v$  que maximice  $p/t$ 
        (en caso de empates, selecciona la que tenga  $p$  mayor)
        Añade  $A = v$  a  $R$ 
      Elimina de  $E$  los ejemplos cubiertos por  $R$ 

```

Figura A.2. Algoritmo de PRISM

A continuación se explica brevemente el funcionamiento de este algoritmo:

- Sea  $\underline{t}$  el número de ejemplos cubiertos por la regla y  $\underline{p}$  el número de ejemplos positivos cubiertos por la regla.
- Lo que hace PRISM es añadir condiciones a reglas que maximicen la relación  $\underline{p}/\underline{t}$  (relación entre ejemplos positivos cubiertos y ejemplos cubiertos en total) sea mayor.
- Este algoritmo, como va eliminando los ejemplos que va cubriendo cada regla, las reglas que se construyen tienen que interpretarse en orden (las nuevas reglas se diseñan solo para cubrir los casos que faltan).
- Reglas que dependen del orden para su interpretación se conocen como *decision lists* o listas de decisión.
- Con varias clasificaciones se puede seleccionar la regla que cubra más ejemplos, y cuando no se tiene una clasificación, escoger la clase mayoritaria [9].

### A.2.2 Algoritmo de aprendizaje de reglas 1R

En la siguiente figura se muestra el algoritmo de 1R.

**Para cada atributo**

**Para cada** valor de cada atributo, crea una regla:

cuenta cuántas veces ocurre la clase

encuentra la clase más frecuente

asigna esa clase a la regla.

Calcula el error de todas las reglas

Selecciona las reglas con el error más bajo

*Figura A.3. Algoritmo de 1R*

Explicación del funcionamiento del algoritmo:

- Los valores faltantes en el algoritmo 1R se tratan como un nuevo valor y para los atributos continuos se hace una división simple.
- Primero se ordenan los atributos con respecto a la clase.
- Se sugieren puntos de partición en cada lugar donde cambia la clase.
- Si existen dos clases diferentes con el mismo valor, se mueve el punto de partición a un punto intermedio con el siguiente valor hacia arriba o abajo dependiendo de donde está la clase mayoritaria.
- Lo que se hace es que se exige que cada partición tenga un número mínimo de ejemplos de la clase mayoritaria.
- Cuando hay clases adyacentes con la misma clase mayoritaria, estas se juntan [10].

### A.2.3 Algoritmo de aprendizaje de reglas AQ

```
Este algoritmo genera reglas para clasificar ejemplos de clase C
Sea: POS el conjunto de ejemplos de clase C (positivos) y NEG el
conjunto de ejemplos que no pertenecen a C (negativos)

Procedimiento AQ(POS,NEG)
  Sea. cover = { }
  REPITE hasta que POS sea vacío
    Selecciona. una SEMILLA (un ejemplo positivo no cubierto por
    cover)
    Sea. STAR = STAR(SEMILLA,NEG) (un conjunto de complexes que
    cubren a SEMILLA y que no cubren a ejemplos negativos)
    Sea. MEJOR el mejor complex en STAR
    Añade MEJOR como una disjunción extra. a cover
    POS = POS - Ejemplos cubiertos por MEJOR
  Regresa. cover

Procedimiento STAR(SEMILLA,NEG)
  Sea. STAR = conjunto con el complex vacío
  WHILE exista un complex en STAR que cubra algún ejemplo en NEG,
  Selecciona. el ejemplo negativo (Eneg) más cercano1, cubierto por un
  complex en STAR
  Especializar complexes en STAR para excluir a Eneg:
    Sea. EXTENSION el conjunto de selectores que cubren a semilla.
    pero no a Eneg
    Sea. STAR el conjunto {  $x \wedge y \mid x \in STAR, y \in EXTENSION$  }
    Quitar todos los complexes en STAR subsumidos por otros
    complexes.
  Repetir hasta que el tamaño de STAR sea  $\leq maxstar$ 
  Quitar el peor complex de STAR
  Regresa. STAR
```

Figura A.4. Algoritmo de AQ

Explicación del funcionamiento del algoritmo AQ:

La *condición* de la regla de producción '*if condición then predice clase*', donde *clase* es la clase más común en los ejemplos (de entrenamiento) que satisface la *condición*.

La generación de reglas se hace en etapas; cada etapa genera una sola regla y luego elimina los ejemplos de entrenamiento cubiertos por la regla. Este proceso se repite hasta que se encuentran suficientes reglas para cubrir todos los ejemplos de la clase escogida. Este proceso se repite para cada clase.

AQ garantiza encontrar un conjunto completamente consistente de reglas con los datos de entrenamiento, si es que estas reglas existen. Esta es una propiedad deseable en dominios sin ruido [11].

#### A.2.4 Algoritmo de aprendizaje de reglas APRIORI

El proceso de iteración del primer paso se llama *level-wise* y va considerando los superconjuntos nivel por nivel. Lo que se tiene es una propiedad anti-monótona: si un conjunto no pasa una prueba, ninguno de sus superconjuntos la pasan.

Si un conjunto de *ítems* no pasa la prueba de soporte, ninguno de sus superconjuntos la pasan. Esto se aprovecha en la construcción de candidatos para no considerarlos todos [15] y [16].

En las siguientes figuras se muestra el algoritmo Apriori con las reglas utilizadas.

```
Apriori()
 $L_1 = \text{find-frequent-1-itemsets}(D)$ 
for ( $k = 2$ ;  $L_{k-1} \neq \text{NULL}$ ;  $k++$ )
    % generate & prune candidate k-itemsets
     $C_k = \text{AprioriGen}(L_{k-1})$ 
    forall transactions  $t \in D$ 
         $C_t = \text{subset}(C_k, t)$ 
        forall candidates  $c \in C_t$ 
             $c.\text{count}++$ 
     $L_k = \{c \in C_k \mid c.\text{count} \geq \text{minsup}\}$ 
Return  $\cup_k L_k$ 
```

Figura A.5. Algoritmo de APRIORI

```

AprioriGen( $L$ ) – Assume transactions in lexicographic order
insert into  $C_k$  all  $p.item_1, p.item_2, \dots, p.item_{k-1}, q.item_{k-1}$  from  $p, q \in L$ 
where  $p.item_1 = q.item_1, p.item_2 = q.item_2, \dots, p.item_{k-1} < q.item_{k-1}$ 
%– Prune itemsets s.t. some  $(k-1)$ -subset of  $c$  is  $\notin L$ 
%– A  $(k-1)$  itemset that is not frequent cannot be a subset of
%– a frequent  $k$ -itemset, then it is removed
forall itemsets  $c \in C_k$ 
  forall  $(k-1)$ -subsets  $s$  of  $c$  do
    if ( $s \notin L_{k-1}$ ) then
      delete  $c$  from  $C_k$ 

```

*Figura A.6. Función Apriori Genera*

```

AssocRules()
forall large itemsets  $l_k, k \geq 2$ 
  GenRules( $l_k, l_k$ )

```

*Figura A.7. Función Reglas de Asociación*

```

GenRules( $l_k, a_m$ ) –Generate all valid rules  $a \rightarrow (l_k - a)$ , for all  $a \subset a_m$ 
 $A = \{(m - 1) - \text{itemsets } a_{m-1} | a_{m-1} \subset a_m\}$ 
forall  $a_{m-1} \in A$ 
   $conf = \text{support}(l_k) / \text{support}(a_{m-1})$ 
  if ( $conf \geq min\_conf$ ) then
    output rule  $a_{m-1} \rightarrow (l_k - a_{m-1})$  with confidence  $conf$ ,  $support = \text{support}(l_k)$ 
    if ( $m - 1 > 1$ ) then
      GenRules( $l_k, a_{m-1}$ ) – Generate rules with subsets of  $a_{m-1}$  as antecedents

```

*Figura A.8. Función Genera Reglas*

## B. HUECOS DE TENSIÓN

El hueco de tensión es la reducción brusca de la señal de tensión en una o varias fases de la red eléctrica, esta disminución ocurre entre el 10% y el 99% de la amplitud, en un intervalo de tiempo de 10ms a 1min.

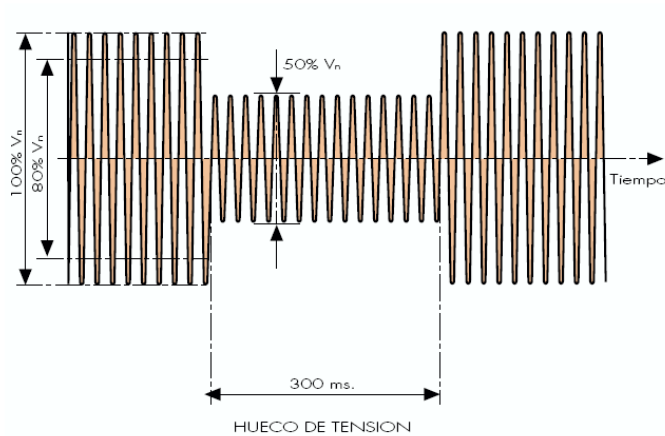


Figura B.1. Hueco de tensión a un 50% de la amplitud de la señal

Se suele asimilar a una onda cuadrada en tanto por ciento de la nominal y que dura entre 100 y 600 milisegundos, volviendo a los valores anteriores de tensión al cabo de entre 1 y 3 segundos.

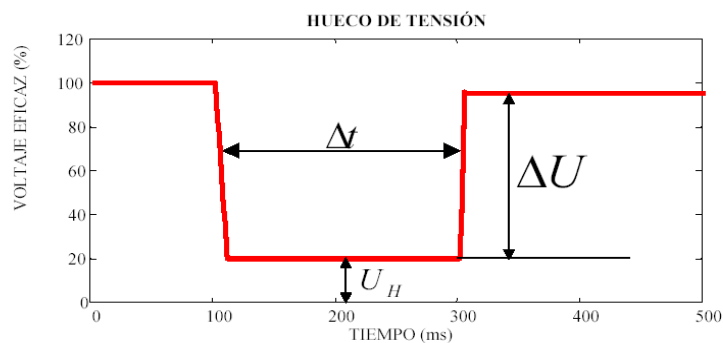


Figura B.2. Características de un hueco de tensión

La Figura B-2 ilustra las características que tiene un hueco de tensión, donde la magnitud de la tensión del hueco es  $U_H$ , la caída y la tensión del hueco es  $\Delta U$  y su duración es  $\Delta t$ . Usualmente un hueco de tensión es caracterizado por su magnitud y duración.

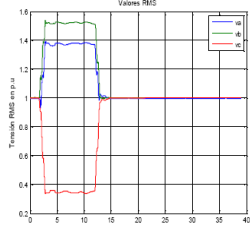
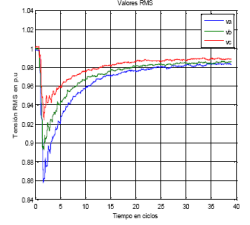
El hueco de tensión es a menudo causado por fallas, aumento de la demanda y eventos transitorios tales como arranque de motores grandes o energización de transformadores. Adicionalmente, según la Comisión Electrotécnica Internacional (IEC-CEI), un “dip” o hueco, es una reducción repentina de la tensión, seguida por una recuperación de la misma, después de un período de tiempo corto que varía entre medio ciclo y algunos segundos. Para el estándar 1346 de IEEE, un “sag” o hueco, es la disminución del valor RMS de la tensión o la corriente a la frecuencia del sistema, de duración de entre 0.5 ciclos a 1 min. [31].

### **B.1 Clasificación de los huecos de tensión**

De acuerdo a su duración y frecuencia, La IEEE std 1159,95 y la NTC 5000,02 clasifica los huecos de tensión.

La Tabla B.1 muestra las características de los huecos de tensión según su duración y la causa general que lo produjo.

Tabla B.1. Clasificación general de los huecos de tensión

Clasificación	Duración y causa
Momentáneos	Duración: Entre 30 ciclos y 3s.
Instantáneos	Duración: Entre 0.5 ciclos y 30 ciclos
Temporales	Duración: Entre los 3s y 1min.
<p data-bbox="448 815 699 848">Hueco rectangular</p> 	<p data-bbox="879 748 1441 1016">Debidos a fallas en el sistema eléctrico las cuales producen un aumento de la corriente y por consiguiente ocurre una fuerte caída de tensión en los elementos que componen el sistema. Este comportamiento va fuertemente ligado a la operación de las protecciones eléctricas del sistema.</p>
<p data-bbox="429 1133 718 1167">Hueco no rectangular</p> 	<p data-bbox="879 1133 1441 1402">Causados por fallas en un sistema eléctrico con cargas rotativas o por el arranque de motores de inducción. Este evento tiene dos componentes, un hueco debido a una falla y un post-hueco debido a la reacción de los motores. De igual forma los huecos causados por energización de transformador.</p>

## **C. SIMULACIÓN DE HUECOS DE TENSIÓN PARA VALIDAR LA BASE DE DATOS SUMINISTRADA**

### **C.1 Simulación en ATP del sistema de potencia IEEE de 34 nodos**

Con el objeto de validar los datos de las perturbación suministradas, se decide simular algunos huecos de tensión causados por fallas de red, energización de banca de capacitores y arranque de motores. Esto pretende aumentar en nivel de confiabilidad de las señales adquiridas para el diagnóstico de la causa de los huecos de tensión.

El sistema utilizado para efectuar las simulaciones fue seleccionado de artículos IEEE que contaran con los parámetros necesarios para el modelado en el software. A continuación se hace una breve descripción de la simulación de este sistema.

Se toma el artículo [32] como referencia que cuenta con un sistema de distribución de 34 nodos, localizado en Arizona EEUU con una tensión nominal de 24.9 kV. Este circuito se caracteriza por ser de gran longitud, ligeramente cargado y por tener cargas desequilibradas.

Se realizaron simulaciones de tal sistema en el ATPDraw para obtener y validar las perturbaciones originadas por fallas de red, arranque de motores de inducción y energización de bancos de capacitores.

El diagrama esquemático realizado en ATPDraw del sistema se muestra en la Figura C.1.

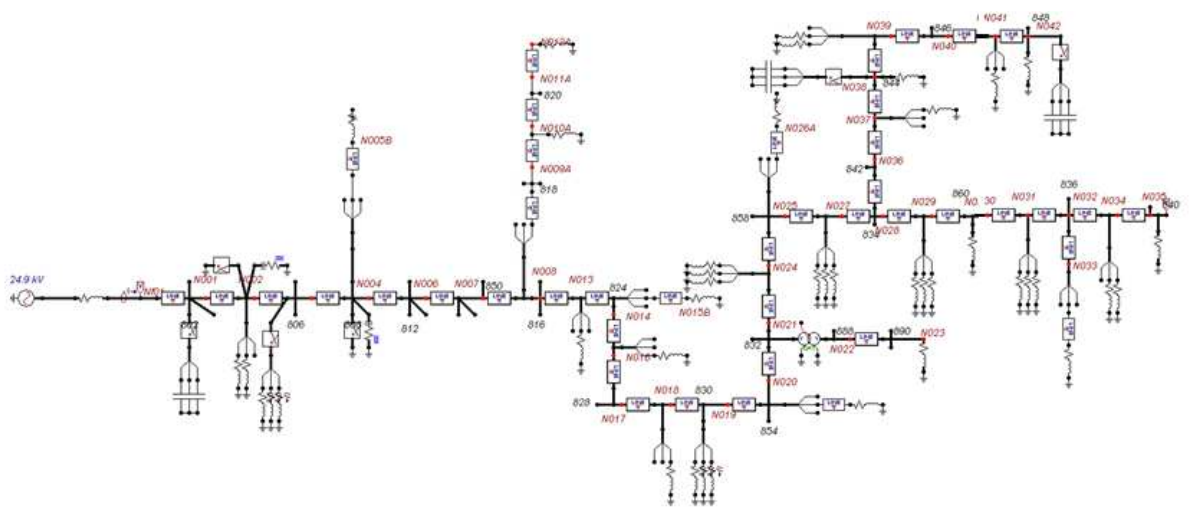


Figura C.1. Sistema modelado en ATP

Los parámetros del sistema eléctrico de distribución utilizados en la simulación en ATP se muestran a continuación en la Tabla C.1.

Tabla C.1. Configuraciones de las líneas de transmisión eléctrica

Config.	Phasing	Phase	Neutral	Spacing ID
		ACSR	ACSR	
300	B A C N	1/0	1/0	500
301	B A C N	#2 6/1	#2 6/1	500
302	A N	#4 6/1	#4 6/1	510
303	B N	#4 6/1	#4 6/1	510
304	B N	#2 6/1	#2 6/1	510

En la Tabla C.1 se indica si el circuito es trifásico tetrafilar, bifásico, monofásico. De igual forma se especifican los calibres de los conductores y se asocia el espaciamiento entre conductores en la quinta columna (*Spacing ID*).

En la Tabla C.2 se muestra la configuración por tramo del sistema de distribución.

Tabla C.2. Configuración de la línea de transmisión eléctrica

Datos de las Líneas			
Node A	Node B	Length(ft.)	Config.
800	802	2580	300
802	806	1730	300
806	808	32230	300
808	810	5804	303
808	812	37500	300
812	814	29730	300
814	850	10	301
816	818	1710	302
816	824	10210	301
818	820	48150	302
820	822	13740	302
824	826	3030	303
824	828	840	301
828	830	20440	301
830	854	520	301
832	858	4900	301
834	860	2020	301
834	842	280	301
836	840	860	301
836	862	280	301
842	844	1350	301
844	846	3640	301
846	848	530	301
850	816	310	301
852	832	10	301
854	856	23330	303
854	852	36830	301
856	864	1620	303
858	834	5830	301
860	836	2680	301
862	838	4880	304
866	890	10560	300

Para ajustar el modelo respecto al tipo de configuración, se tiene una matriz de impedancia asociada. Con esta información es posible llevar a cabo la simulación del modelo en ATP [32] indicándolas a continuación:

Configuration 300:

----- Z & B Matrices Before Changes -----

```

      Z (R +jX) in ohms per mile
1.3368  1.3343  0.2101  0.5779  0.2130  0.5015
          1.3238  1.3569  0.2066  0.4591
                          1.3294  1.3471
    
```

```

B in micro Siemens per mile
    5.3350  -1.5313  -0.9943
           5.0979  -0.6212
                   4.8880

```

Configuration 301:

```

      Z (R +jX) in ohms per mile
1.9300  1.4115  0.2327  0.6442  0.2359  0.5691
           1.9157  1.4281  0.2288  0.5238
                   1.9219  1.4209
B in micro Siemens per mile
    5.1207  -1.4364  -0.9402
           4.9055  -0.5951
                   4.7154

```

Configuration 302:

```

      Z (R +jX) in ohms per mile
2.7995  1.4855  0.0000  0.0000  0.0000  0.0000
           0.0000  0.0000  0.0000  0.0000
                   0.0000  0.0000
B in micro Siemens per mile
    4.2251  0.0000  0.0000
           0.0000  0.0000
                   0.0000

```

Configuration 303:

```

      Z (R +jX) in ohms per mile
0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
           2.7995  1.4855  0.0000  0.0000
                   0.0000  0.0000
B in micro Siemens per mile
    0.0000  0.0000  0.0000
           4.2251  0.0000
                   0.0000

```

Configuration 304:

```

      Z (R +jX) in ohms per mile
0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
           1.9217  1.4212  0.0000  0.0000
                   0.0000  0.0000
B in micro Siemens per mile
    0.0000  0.0000  0.0000
           4.3637  0.0000
                   0.0000

```

En las siguientes tablas se especifican los parámetros de los transformadores, los valores de cargas concentradas, cargas distribuidas y la especificación de los capacitores conectados al sistema. Cada una de las cargas se modeló como una impedancia respecto al nivel de tensión y a la potencia demandada.

Tabla C.3. Especificaciones de los transformadores

Transformer Data					
	kVA	kV-high	kV-low	R - %	X - %
Substation:	2500	69 - D	24.9 - Gr. W	1	8
XFM -1	500	24.9 - Gr.W	4.16 - Gr. W	1.9	4.08

Tabla C.4. Cargas concentradas

Spot Loads							
Node	Load	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-4
	Model	kW	kVAr	kW	kVAr	kW	kVAr
860	Y-PQ	20	16	20	16	20	16
840	Y-I	9	7	9	7	9	7
844	Y-Z	135	105	135	105	135	105
848	D-PQ	20	16	20	16	20	16
890	D-I	150	75	150	75	150	75
830	D-Z	10	5	10	5	25	10
Total		344	224	344	224	359	229

Tabla C.5. Cargas distribuidas

Distributed Loads								
Node	Node	Load	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-3
A	B	Model	kW	kVAr	kW	kVAr	kW	kVAr
802	806	Y-PQ	0	0	30	15	25	14
808	810	Y-I	0	0	16	8	0	0
818	820	Y-Z	34	17	0	0	0	0
820	822	Y-PQ	135	70	0	0	0	0
816	824	D-I	0	0	5	2	0	0
824	826	Y-I	0	0	40	20	0	0
824	828	Y-PQ	0	0	0	0	4	2
828	830	Y-PQ	7	3	0	0	0	0
854	856	Y-PQ	0	0	4	2	0	0
832	858	D-Z	7	3	2	1	6	3
858	864	Y-PQ	2	1	0	0	0	0
858	834	D-PQ	4	2	15	8	13	7
834	860	D-Z	16	8	20	10	110	55
860	836	D-PQ	30	15	10	6	42	22
836	840	D-I	18	9	22	11	0	0
862	838	Y-PQ	0	0	28	14	0	0
842	844	Y-PQ	9	5	0	0	0	0
844	846	Y-PQ	0	0	25	12	20	11
846	848	Y-PQ	0	0	23	11	0	0
Total			262	133	240	120	220	114

Tabla C.6. Especificaciones de los capacitores

Node	Ph-A	Ph-B	Ph-C
	kVAr	kVAr	kVAr
844	100	100	100
848	150	150	150
Total	250	250	250

Las señales simuladas hacen referencia a la secuencia de ondas de tensión eficaces. En la figura C.2 se ilustran algunas de las formas de onda obtenidas de la simulación.

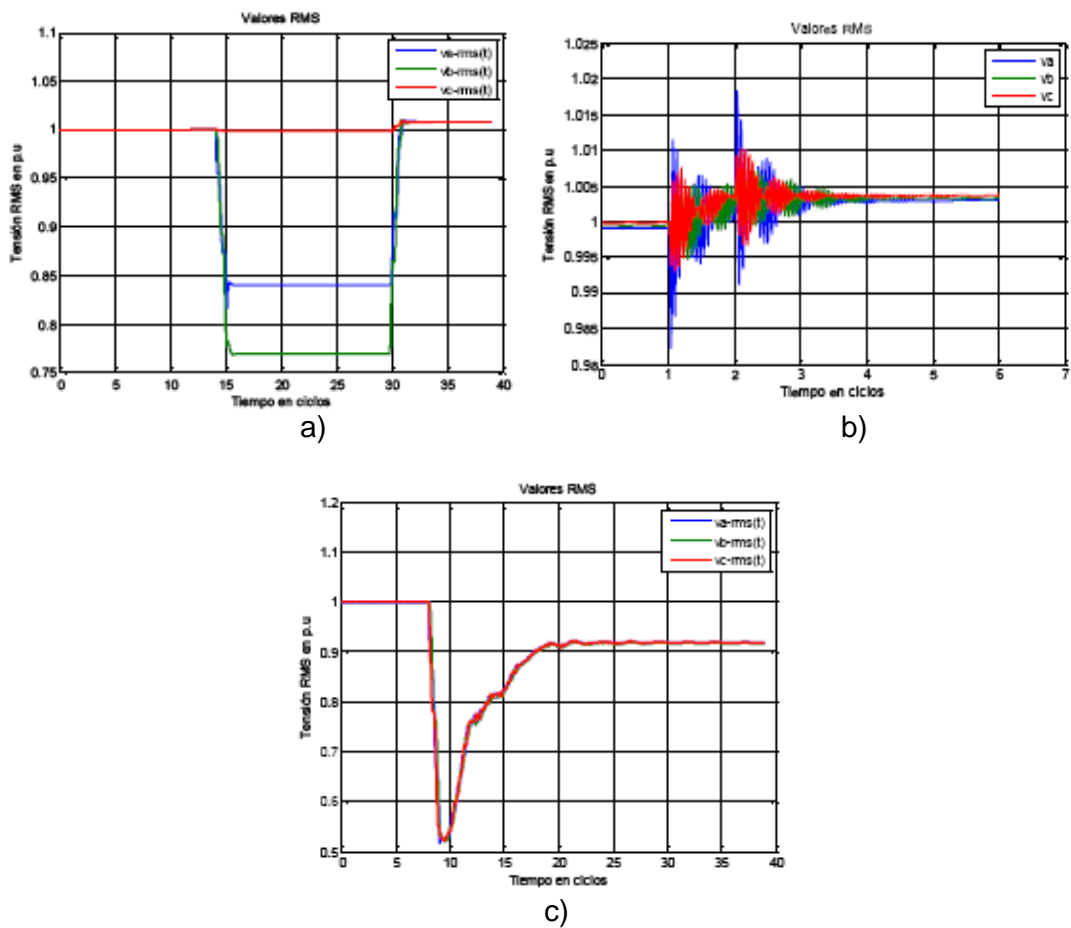


Figura C.2. a) Falla bifásica. b) Energización de banco de capacitores. c) Arranque de motor de inducción