

**ALGORITMOS PARA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE
DISTRIBUCIÓN USANDO MÁQUINAS DE SOPORTE VECTORIAL**

JUAN CARLOS BEDOYA CEBALLOS

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA
MAESTRÍA EN INGENIERÍA ELÉCTRICA
PEREIRA
2010**

**ALGORITMOS PARA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE
DISTRIBUCIÓN USANDO MÁQUINAS DE SOPORTE VECTORIAL**

JUAN CARLOS BEDOYA CEBALLOS

**Proyecto de grado para optar al título de
MAGISTER EN INGENIERÍA ELÉCTRICA**

**DIRECTOR
PhD. JUAN JOSÉ MORA FLORES**

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA
MAESTRÍA EN INGENIERÍA ELÉCTRICA
PEREIRA
2010**

Nota de aceptación:

Presidente del Jurado

Director

Jurado

Jurado

TABLA DE CONTENIDO

INTRODUCCION	7
1.1 CALIDAD DE LA ENERGÍA EN SISTEMAS DE DISTRIBUCIÓN	7
1.2 METODOLOGÍAS EMPLEADAS PARA LA LOCALIZACIÓN DE FALLAS.....	7
1.3 PROPUESTA METODOLÓGICA PARA LA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN	9
1.4 OBJETIVOS DE LA INVESTIGACIÓN	11
1.4.1 Objetivo general	11
1.4.2 Objetivos específicos	11
1.5 RESULTADOS OBTENIDOS.....	12
1.6 ESTRUCTURA DEL DOCUMENTO.....	12
2. MÁQUINAS DE SOPORTE VECTORIAL.....	14
2.1 INTRODUCCIÓN.....	14
2.2 MÁQUINAS DE SOPORTE VECTORIAL PARA CLASIFICACIÓN (SVMc)	15
2.2.1 Problema básico de las SVM para problemas de biclasificación	15
2.2.2 Caso linealmente no separable para SVM para problemas de biclasificación.....	18
2.2.3 Máquinas de soporte vectorial no lineales para biclasificación	18
2.2.4 SVM para clasificación multiclase.....	20
2.2.4.1 Multiclasificación a partir de SVM biclasificadoras.....	20
2.2.4.2 Métodos de reconstrucción.....	22
2.2.4.3 Esquemas de Votación.....	23
2.2.5 Ejemplo de aplicación de una máquina de soporte vectorial para clasificación no lineal.	23
2.3 MÁQUINAS DE SOPORTE VECTORIAL PARA REGRESIÓN.....	29
2.3.1 Problema básico	29
2.3.2 Máquinas de soporte vectorial no lineales para regresión	31
2.4 SINTONIZACIÓN DE PARÁMETROS PARA LA SVM	32
2.5 NORMALIZACIÓN DE LOS DATOS	33
3. METODOLOGIA PROPUESTA PARA LA LOCALIZACION DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN	34
3.1 INTRODUCCION.....	34

3.2	DESCRIPCION DE LA METODOLOGIA	34
3.3	MODELAMIENTO DEL SISTEMA DE DISTRIBUCIÓN	36
3.3.1	Modelamiento de líneas	37
3.3.2	Modelo de la carga.....	37
3.3.3	Compensadores de potencia reactiva.....	38
3.3.4	Reguladores de tensión.....	38
3.3.5	Protecciones	39
3.3.5.1	Fusibles	39
3.3.5.2	Reconectores	42
3.3.5.3	Seccionalizadores	44
3.3.5.4	Esquema de Salvamento de Fusibles.....	46
3.3.6	Transformadores.....	47
3.4	SIMULACIÓN AUTOMATICA DE FALLAS.....	48
3.4.1	Generar lista de fallas	49
3.4.2	Generar los diferentes Sistemas en el ATP	50
3.4.3	Generar el archivo de ejecución automática de los archivos ATP de fallas	52
3.4.4	Convertir los resultados de los archivos .pl4 a archivos .mat.....	53
3.5	INFORMACIÓN REQUERIDA PARA LA CARACTERIZACIÓN DE LOS EVENTOS DE FALLA.....	54
3.5.1	Caracterización de la profundidad de los huecos de tensión.....	54
3.5.2	Variación de la magnitud de la corriente.....	55
3.5.3	Análisis de la señal transitoria	56
3.5.4	Análisis de coeficientes α y β	60
3.5.5	Análisis de los patrones de la corriente de carga (firma del sistema)	60
3.5.6	Zonificación del sistema.....	64
3.6	LOCALIZACIÓN DE FALLAS COMO PROBLEMA DE CLASIFICACION	65
3.7	PROBLEMA DE LOCALIZACION DE FALLAS COMO PROBLEMA DE REGRESION ..	67
4.	APLICACIÓN DE LAS SVM AL PROBLEMA DE LA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN.....	72
4.1	INTRODUCCIÓN.....	72
4.2	SISTEMA DE DISTRIBUCION PARA PRUEBAS.....	72
4.3	SIMULACIÓN DE FALLAS EN EL SISTEMA DE DISTRIBUCIÓN DE PRUEBAS	80
4.4	PRUEBAS PARA LA LOCALIZACION DE LA ZONA DE FALLA	84
4.5	PRUEBAS PARA LA ESTIMACION DE LA DISTANCIA DE FALLA	97
5.	CONCLUSIONES Y TRABAJOS FUTUROS.....	119
5.1	CONCLUSIONES GENERALES	119

5.2 CONCLUSIONES ASOCIADAS A LA CARACTERIZACION DE LAS SEÑALES DE TENSION Y DE CORRIENTE.....	120
5.3 CONCLUSIONES ASOCIADAS A LA MBC PARA LA LOCALIZACION DE FALLAS.	121
5.4 TRABAJOS FUTUROS	122
BIBLIOGRAFIA.....	124
ANEXO 1	128
ANEXO 2	134

INTRODUCCION

1.1 CALIDAD DE LA ENERGÍA EN SISTEMAS DE DISTRIBUCIÓN

La calidad de la energía ha sido objeto de amplias investigaciones con significativos avances en los últimos años por ser éste un tema de bastante interés tanto para los operadores de red como para los usuarios finales. Entre los diferentes aspectos que se consideran en la calidad de la energía se encuentran, la forma de onda, la continuidad del servicio, aspectos económicos, entre otras [Bollen 2000].

En lo relacionado con la continuidad del servicio, existen a nivel colombiano dos indicadores que reflejan el número de interrupciones y la duración de las mismas en el servicio eléctrico que son los índices FES (frecuencia equivalente de interrupción del servicio) y DES (duración equivalente de interrupción del servicio) respectivamente. El incumplimiento de los valores máximos establecidos por el ente regulador (para el caso colombiano, la CREG: Comisión de Regulación de Energía y Gas) por parte de los agentes prestadores del servicio, lleva consigo sanciones de tipo económico [CREG 1998] [CREG 2003] [CREG 2005].

Entre las múltiples causas que alteran la continuidad del servicio se encuentran fallas provocadas por diversos factores tales como condiciones ambientales, aves u otros, los cuales afectan el comportamiento de la red eléctrica. La duración de la interrupción se ve afectada de manera directa por el tiempo de localización de la falla y es por ello que este trabajo de tesis está basado en el supuesto de que la *oportuna localización de las fallas ocurridas en la red permitirá a las empresas prestadoras del servicio mejorar sus indicadores de continuidad del servicio*. La rápida localización de fallas permite disminuir la duración de la interrupción (DES) y adicionalmente la continua localización de las mismas permite detectar las debilidades de la red y fortalecer el sistema para que las interrupciones sean cada vez menores [Short 2003]

1.2 METODOLOGÍAS EMPLEADAS PARA LA LOCALIZACIÓN DE FALLAS

Las metodologías utilizadas en la actualidad para localizar fallas en sistemas de distribución están basadas en la inspección visual, observación de los medidores ubicados en la subestación y de algunos métodos algorítmicos que presentan el inconveniente de la múltiple estimación; éste problema consiste básicamente en que la distancia de falla medida desde el alimentador principal puede coincidir con varios puntos dentro del sistema. La gran mayoría de los sistemas de distribución poseen una estructura que consiste de una subestación, un alimentador principal y múltiples ramificaciones tal como se presenta en la figura 1.1

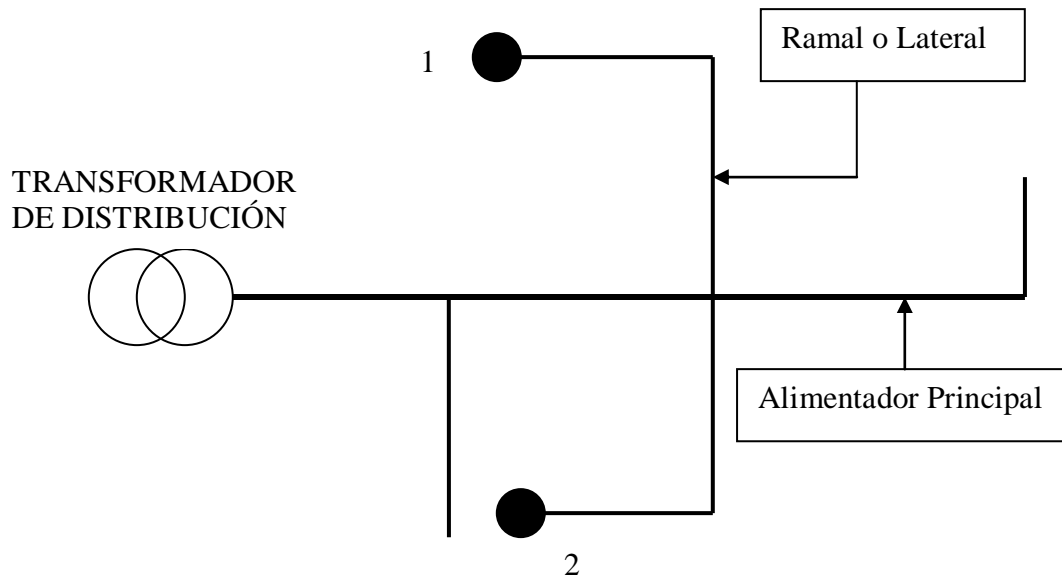


Figura 1.1. Estructura de los sistemas de distribución.

Para la solución del problema de localización se han utilizado diferentes metodologías que se pueden clasificar en dos grandes grupos: las que usan el modelo de la red y las que aprovechan otro tipo de información del sistema.

Los métodos basados en el modelo de la red (MBM) más comúnmente usados para localización de fallas, utilizan el valor RMS del fundamental de tensión y de corriente para estimar la impedancia medida desde la subestación del sistema de distribución hasta el lugar donde se encuentra localizada la falla. Esta impedancia es convertida a una distancia equivalente a partir de los parámetros eléctricos de la red; sin embargo, debido a la alta ramificación del sistema, existe una múltiple estimación de la sección que está bajo falla y además presentan el inconveniente de la alta dependencia del modelo del sistema. [Mora *et al* (2005a)]

Por ejemplo, frente a la ocurrencia de una falla en el punto 1 de la figura 1.1, la distancia medida desde el transformador de distribución al lugar de falla, podría coincidir en dos puntos (punto 1 y 2), esta particularidad en la estimación del lugar de la falla se conoce como el problema de la múltiple estimación. En la práctica, comprobar en cuál de los puntos ocurrió el evento, puede tomar un tiempo considerable, lo que trae como consecuencia pérdidas económicas para los usuarios. El problema en el proceso de localización de fallas en los sistemas de distribución radica esencialmente en la dificultad de implementar los algoritmos y esquemas existentes para la identificación y localización de la misma, en razón de sus implicaciones técnicas, económicas y de operación.

Por otra parte, los métodos que aprovechan otro tipo de información se conocen como basados en el conocimiento (MBC) y usan la señal en el dominio del tiempo de tensión y corriente, el tipo de carga conectada en cada nodo, ubicación de las protecciones e información de las fallas ocurridas anteriormente en la red, entre otras, con el fin de localizar una región o zona de falla. Estos métodos son un excelente complemento de los MBM y permiten reducir el problema de la múltiple

estimación del sitio de falla. No obstante, para una eficiente implementación de los MBC se necesita una gran cantidad de información de fallas ocurridas en los diferentes nodos del sistema; normalmente las empresas distribuidoras de energía u operadores de red no poseen dicha información y por tanto se hace necesario recurrir a técnicas de simulación que permitan obtener una base de datos con registros de fallas [Mora et al (2005b)].

Nuevamente, si se toma como ejemplo una falla en el punto 1 de la figura 1.1, y adicionalmente se ha dividido el sistema en dos zonas (la primera zona contiene los ramales que se encuentran en la parte norte de la red y la segunda los que se encuentran en la parte sur), mediante la implementación de una metodología basada en el conocimiento, podría clasificarse en cuál de las dos zonas ha ocurrido el evento; uniendo la información de ambas metodologías (las basadas en el modelo de la red y las que usan otro tipo de información) podría estimarse de manera eficiente, el punto de localización de la falla.

1.3 PROPUESTA METODOLÓGICA PARA LA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN

Este trabajo propone una metodología para resolver el problema de localización de fallas en sistemas de distribución; para ello, se construye una base de datos con las características más importantes del sistema en estado de prefalla y postfalla, obtenidas mediante el modelamiento y simulación eléctrica de la red. La información almacenada en la base de datos es adaptada y aprovechada por una metodología basada en el conocimiento para luego estimar, de manera oportuna, la zona y la distancia de localización de la falla.

La figura 1.2 presenta las etapas de las cuales consta la metodología propuesta para la localización de fallas en sistemas de distribución:

- *Modelamiento de los elementos del sistema de distribución.*

Modelar el sistema de distribución de energía es de vital importancia para la solución del problema de localización de fallas, ya que mediante la implementación del modelo, en un software de simulación de sistemas eléctricos, es posible reproducir el comportamiento de las variables del sistema real. Para la correcta implementación de una metodología de localización de fallas es necesario disponer de un volumen de información bastante amplio de tal forma que la simulación y el adecuado modelamiento de la red juegan un papel importante en el desarrollo de la metodología.

- *Clasificación de las áreas o zonas del sistema de distribución*

La clasificación o división en áreas o zonas del sistema de distribución es la base para la implementación de la MBC, pues estas áreas se convertirán posteriormente en las "clases" a las cuales la técnica de clasificación podrá asociar el evento de falla.

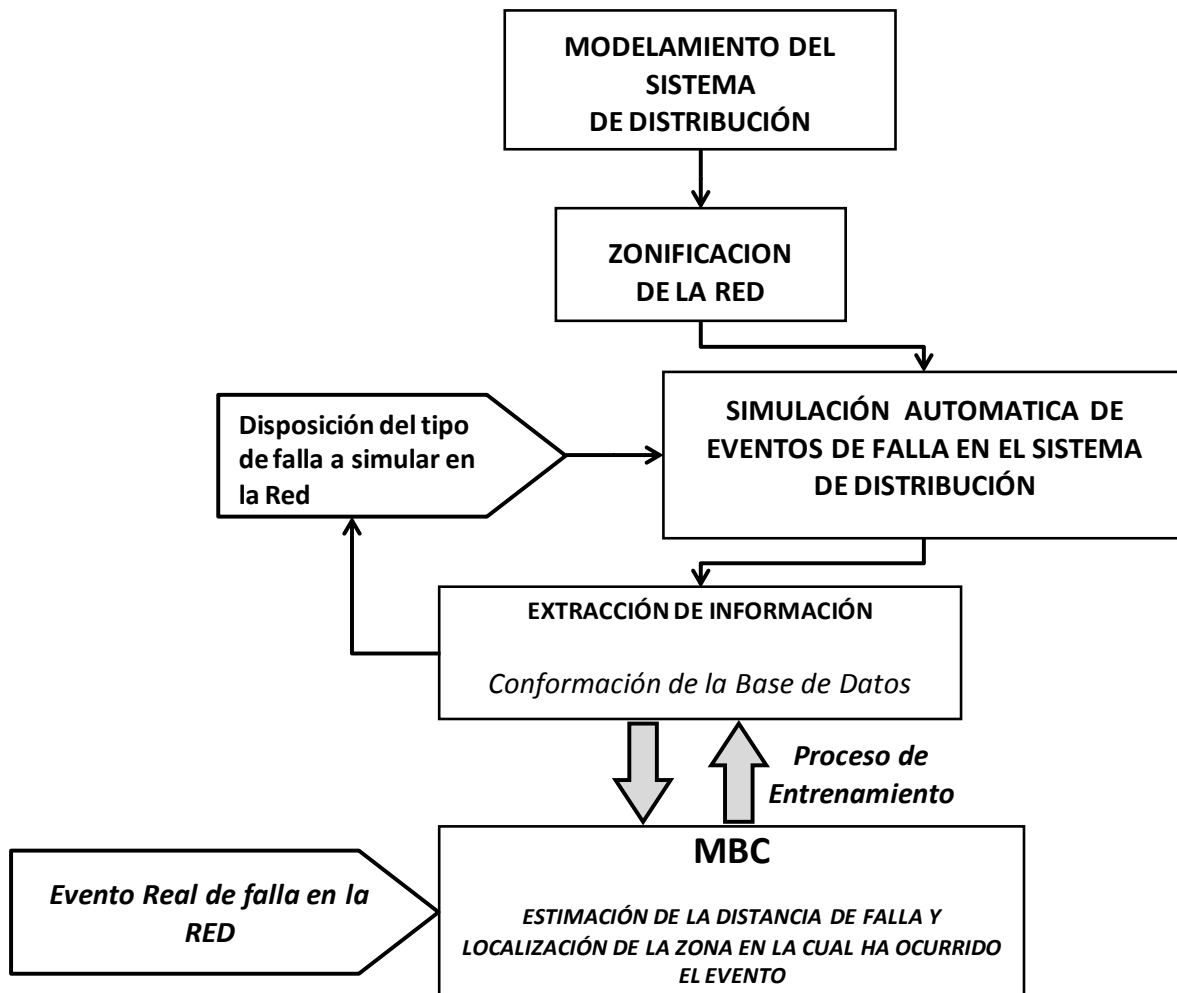


Figura 1.2. Esquema de las etapas metodológicas necesarias para construir una herramienta que permita modelar los conocimientos extraídos de las señales de tensión y corriente para resolver el problema de localización de fallas.

- *Simulación automática de las fallas*

Anteriormente se mencionó que para la correcta implementación de una MBC es necesario disponer de un conjunto de información que describa al sistema de distribución bajo estado de falla para los diferentes puntos donde es factible que ocurra un evento; generalmente, esta información se obtiene de las señales de tensión y corriente (medidas en la subestación del sistema de distribución) y variables adicionales; sin embargo, la mayoría de empresas distribuidoras de energía no poseen esta información y se hace necesario recurrir a las técnicas de simulación que permitan obtener una base de datos con registros de las características del sistema fallado; de allí la importancia del adecuado modelamiento eléctrico del sistema y la automatización del proceso de simulación que permita obtener la información requerida en tiempos reducidos.

- *Extracción de las características del sistema bajo estado de falla.*

Las señales obtenidas en la etapa anterior deberán ser sometidas a un tratamiento que permita extraer las características que describan el comportamiento del sistema de distribución bajo estado

de falla. A este tipo de información se le conoce como “descriptores” del sistema de distribución bajo estado de falla. Esta información debe ser analizada, validada y normalizada de tal manera que no impacte o altere el proceso asociado a la metodología basada en el conocimiento (MBC).

- *Implementación de la MBC para la estimación de la zona y la distancia de ocurrencia del evento.*

Una vez obtenida y organizada toda la información disponible, se procede con la implementación de la metodología basada en el conocimiento (MBC) que se encargará de estimar la zona de ocurrencia del evento de falla. En principio, se utiliza un subconjunto de la información para la etapa de aprendizaje de la MBC, es decir mediante el subconjunto de datos mencionado (datos de entrenamiento) la MBC reconocerá el comportamiento del sistema de distribución para una zona de falla en particular. Así como las MBC permiten estimar variables discretas (en este caso, la zona del evento de falla) también permiten estimar variables continuas, así no sólo se puede encontrar la zona de ocurrencia de la falla sino también la distancia, medida a partir de la subestación, a la cual a ocurrido el evento.

Posteriormente, se evalúa la capacidad de la MBC para estimar la zona y la distancia de falla para el subconjunto de datos que aun no se ha utilizado (datos de validación).

La MBC propuesta en este trabajo está basada en Máquinas de Soporte Vectorial (*SVM del Inglés Support Vector Machine*) las cuales poseen la versatilidad de clasificar (*SVMc del Inglés Support Vector Machine for Clasification*) y realizar regresión de datos (*SVMr del Inglés Support Vector Machine for Regression*); se propone una metodología híbrida entre ambas herramientas para realizar la localización de la zona y la estimación de la distancia de ocurrencia de la falla, respectivamente.

1.4 OBJETIVOS DE LA INVESTIGACIÓN

1.4.1 Objetivo general

Desarrollar una metodología basada en el conocimiento la cual, frente un evento de falla en la red, permita localizar la zona de falla y la distancia, medida a partir de la subestación del sistema, en la cual ha ocurrido dicho evento.

1.4.2 Objetivos específicos

- Estudiar la teoría matemática que soporta el funcionamiento de las máquinas de soporte vectorial tanto para clasificación de variables discretas como regresión de variables continuas.
- Desarrollar el modelo eléctrico de las protecciones comúnmente aplicadas en sistemas de distribución al interior del software de simulación escogido para el desarrollo de las pruebas.
- Desarrollar un algoritmo de simulación automática de fallas en un sistema de distribución específico que al mismo tiempo permita extraer eficientemente las señales de tensión y corriente asociadas con el evento de falla.
- Caracterizar la información asociada a los tipos de fallas de un sistema de distribución de energía eléctrica de característica radial, que permita obtener la localización de la falla.

- Implementar la metodología de localización de fallas en un sistema de distribución ampliamente difundido entre los investigadores de esta temática.
- Validar la eficiencia de la metodología desarrollada mediante el uso de esquemas de análisis tales como las matrices de confusión y herramientas de análisis estadístico.

1.5 RESULTADOS OBTENIDOS

El trabajo realizado permitió crear una metodología basada en el conocimiento la cual, ante un evento de falla en la red, permite localizar la zona de falla dentro del sistema y posteriormente, con base en esta información es posible determinar la distancia de ocurrencia del evento. La base sobre la cual se desarrolló esta metodología consiste en la correcta definición de las zonas del sistema de distribución, de manera tal que la metodología encargada de estimar la zona de falla posea altos índices de acierto; posteriormente, a través de una herramienta híbrida que combina las máquinas de soporte vectorial para regresión con las máquinas de soporte vectorial para clasificación, es posible desarrollar un estimador de la distancia a la cual ha ocurrido la falla.

Los aportes más importantes asociados con este trabajo son:

- Estudio de la teoría matemática que soporta el funcionamiento de las máquinas de soporte vectorial para y regresión.
- Desarrollo del modelo eléctrico de los fusibles, reconectores y seccionalizadores en el software de simulación EMTP/ATP mediante el uso de los MODELS.
- Desarrollo de un algoritmo de simulación automática de fallas y extracción de señales en el sistema de distribución IEEE de 123 Nodos.
- Desarrollo de una metodología que caracteriza la información asociada a los tipos de fallas de una red de distribución, mediante el análisis de las señales de tensión y corriente medidas en la subestación principal del sistema.
- Implementación de la metodología de localización de fallas en el sistema de distribución de 123 nodos de la IEEE.
- Pruebas y análisis de los resultados de la metodología de localización de fallas desarrollada mediante el uso de herramientas como las matrices de confusión y análisis estadístico.

1.6 ESTRUCTURA DEL DOCUMENTO

Este documento está organizado en cinco capítulos. En el capítulo inicial se discuten las generalidades de la calidad de la energía y la manera en la que este trabajo contribuye con el desarrollo de esta temática; también se presenta de manera formal la metodología propuesta para la localización de fallas en sistemas de distribución y se exponen brevemente los aspectos más importantes.

En el capítulo dos de este trabajo, se presenta la fundamentación teórica de la metodología basada en el conocimiento que se propone como herramienta de solución al problema de locación de fallas; se desarrollan ejemplos básicos orientados a la solución de dicho problema mediante las máquinas de soporte vectorial para clasificación.

En el capítulo tres, se presenta de manera detallada la metodología propuesta para la localización de fallas en sistemas de distribución, iniciando por la etapa de modelamiento de los elementos de la red, posteriormente la simulación automática de fallas, la extracción de información, la caracterización del sistema, el ordenamiento de información y finalizando en la configuración de la metodología basada en el conocimiento para resolver el problema de localización de la zona y distancia de falla.

En el capítulo cuatro, se presentan los resultados obtenidos al aplicar la metodología descrita para el sistema de distribución IEEE de 123 nodos; se discuten de manera detallada los resultados obtenidos con las aplicaciones de mejor funcionalidad; dicha discusión se encuentra enmarcada en las características reales del sistema y en los parámetros eléctricos de la red simulada.

En el capítulo cinco se presenta las conclusiones del trabajo y los trabajos futuros que devienen del desarrollo de esta metodología.

Finalmente, se relacionan las referencias bibliográficas y dos anexos, el primero presenta el detalle del sistema de distribución escogido para pruebas y el segundo presenta el código fuente de una de las protecciones desarrolladas mediante los MODELS del software de simulación ATP el cual es uno de los aportes de este trabajo.

2. MÁQUINAS DE SOPORTE VECTORIAL

2.1 INTRODUCCIÓN

Las múltiples ramificaciones presentes en los sistemas de distribución, el reducido número de medidores, las condiciones del terreno y del medio que rodean las redes, entre otras, son la causa de la complejidad de la localización oportuna de las faltas que ocurren en el sistema en un instante de tiempo determinado. Es por ello que para la solución de este problema se han utilizado diferentes metodologías, las cuales se pueden clasificar en dos grandes grupos: las metodologías basadas en el modelo de la red (MBM), y las metodologías basadas en el conocimiento (MBC); las primeras emplean los valores RMS de las señales de tensión y corriente medidas en la cabecera del sistema de distribución para estimar la impedancia vista desde el punto de medición hasta el punto de localización de la falta y por medio de los parámetros eléctricos de la red se calcula la distancia de localización de la falla; las segundas aprovechan información propia del sistema tal como la señal de tensión y corriente en el dominio del tiempo, el tipo de carga conectada en cada nodo, ubicación de las protecciones, información de las fallas ocurridas anteriormente en la red, registros de reportes de fallas informados por los mismos usuarios, entre otras, para conformar un volumen amplio de información y reproducir, por medio de este conocimiento, el comportamiento de la red bajo estado de falla.

Por otra parte, la principal desventaja de las MBM radica en la múltiple estimación del punto de falla debido al alto número de ramificaciones de los sistemas de distribución, haciendo necesario el uso de otras técnicas tales como la incorporación de ecuaciones adicionales que consideran las fases no falladas, entre otras [Morales et al (2008)]. En cuanto a las MBC, la principal desventaja está asociada con el volumen elevado de información necesario para el correcto funcionamiento de las mismas, el cual generalmente no lo poseen las empresas prestadoras del servicio, haciendo necesario recurrir a herramientas de simulación; es por ello que en la literatura moderna se encuentran diversos trabajos en los que se proponen metodologías híbridas que combinan las principales ventajas de las MBM y las MBC [Mora et al (2005)].

Algunas de las MBC empleadas comúnmente son: El método *LAMDA* (Learning Algorithm for Multivariate Data Analysis) que es un algoritmo de aprendizaje que permite el análisis de datos multivariados mediante una metodología de clasificación estadística [Waissman et al (2000)]; las *Redes Neuronales Artificiales* (RNA) [Martin et al (2001)] [Li-Xing (1997)]; los *Sistemas de Inferencia Difusa* (FIS) que permiten codificar el conocimiento de un problema específico a través de reglas o etiquetas lingüísticas del tipo IF-THEN [Martin et al (2001)] [Li-Xing (1997)]; los *Sistemas Adaptativos de Inferencia Neuro-Difusa* (ANFIS) que envuelven las ventajas de las RNA y los FIS [Martin et al (2001)] [Li-Xing (1997)]; las *Máquinas de Soporte Vectorial*, (SVMs, Support Vector Machines) que son una evolución de las RNA, que clasifican con un hiperplano óptimo de separación (OSH) los datos de entrada y los cataloga en alguna de las clases establecidas. [Russell et al (2000)] [Burgess (2001)]; las *Redes Bayesianas* que son un modelo probabilístico multivariado que relaciona un conjunto de variables aleatorias mediante un grafo dirigido [Cormane et al (2005)], entre otros.

En este capítulo se presenta la Metodología Basada en el Conocimiento (MBC) llamada Máquinas de Soporte Vectorial (SVM del inglés Support Vector Machine) la cual ha sido de amplia implementación en las últimas décadas para la solución de problemas de clasificación y de estimación de variables discretas y/o continuas. Para el problema de localización de fallas, las SVM utilizan patrones o características significativas del sistema de distribución y son de gran aplicabilidad dado el tamaño de dicha información y la complejidad para el análisis de la misma.

Las SVM fueron desarrolladas por Vladimir Vapnik usando herramientas estadísticas de la optimización y de las máquinas de aprendizaje; su estudio es útil debido a características en el proceso de aprendizaje (entrenamiento) como la ausencia de mínimos locales, alta capacidad de generalización y el uso de funciones de transformación (funciones de kernel). Para el caso de clasificación de variables discretas, la estructura de las SVM solo depende de un parámetro de penalización C en la función objetivo del problema de optimización en el aprendizaje y de los parámetros que posea la función de kernel; para el caso de estimación de variables continuas se posee adicionalmente un parámetro de control que define el margen de aceptación de la estimación.

Las Máquinas de Soporte Vectorial representan una nueva técnica de aprendizaje introducido en el marco de la Minimización del Riesgo Estructural (SRM del inglés Structural Risk Minimization) y la teoría de las fronteras VC [Burgos (2001)]

2.2 MÁQUINAS DE SOPORTE VECTORIAL PARA CLASIFICACIÓN (SVMc)

2.2.1 Problema básico de las SVM para problemas de biclasificación

Dado un conjunto de datos distribuidos en dos clases, una SVM biclasificadora y lineal, es aquella que determina el hiperplano que separa ambas clases de tal manera que la mayor cantidad de datos de la misma clase se encuentran a un mismo lado del hiperplano, mientras se maximiza la distancia de separación entre dichas clases.

Para un grupo de datos de entrenamiento de tamaño N compuesto por pares atributo-etiqueta $(\mathbf{x}_i, \mathbf{y}_i)$ con $1 \leq i \leq N$, donde $\mathbf{x}_i \in \mathbf{R}^p$ y $\mathbf{y}_i \in \{-1, 1\}$, se desea obtener la ecuación del hiperplano que separe dicho grupo de entrenamiento, de manera que aquellos puntos con igual etiqueta se encuentren a un mismo lado del hiperplano. Esto significa encontrar un vector \vec{w} (vector normal que define al hiperplano) y un escalar b tal que

$$y_i (\langle \vec{w}, \vec{x}_i \rangle + b) \geq 0 \quad i = 1, 2, \dots, N \quad (2.1)$$

Donde $\langle \cdot, \cdot \rangle$ denota el producto interior Euclídeo para el espacio \mathbf{R}^p .

En la figura 2.1 se muestran diferentes soluciones al problema de biclasificación; mediante un hiperplano (para este caso, el hiperplano consiste de una línea recta) se pueden separar los datos de cada clase. Las SVM biclasificadoras permiten encontrar un único hiperplano conocido como el Hiperplano Óptimo de Separación (OSH del inglés Optimal Separating Hiperplane) el cual maximiza el margen de separación entre ambas clases. La figura 2.2 muestra el OSH para el problema de biclasificación, su ecuación cartesiana se muestra en la ecuación 2.2.

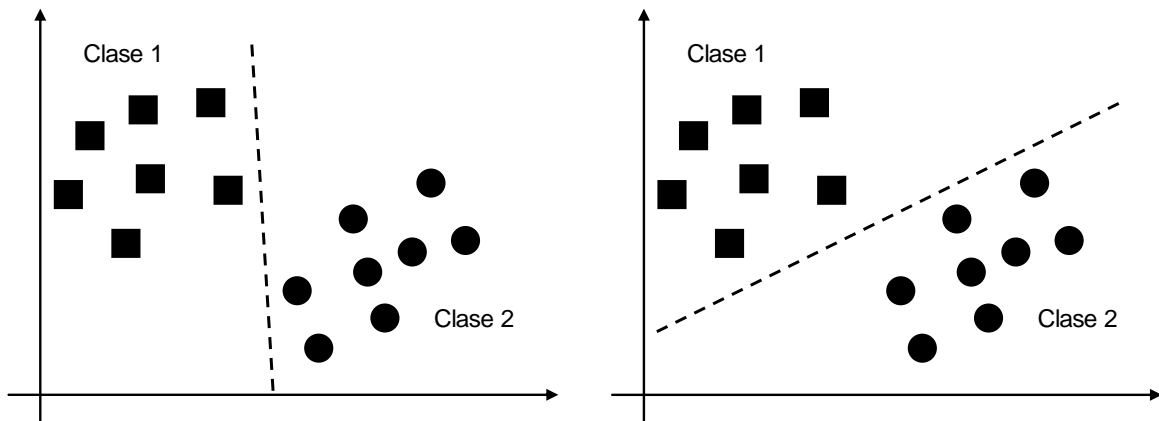


Figura 2.1. Planos de Separación para el problema de Biclasiación.

Hiperplano Optimo de Separación:

$$(\langle \bar{w}, \bar{x}_i \rangle + b) = 0 \quad (2.2)$$

Paralelo al OSH se encuentran dos hiperplanos definidos ó soportados por datos pertenecientes a cada clase respectivamente; estos datos son conocidos como vectores de soporte y su nombre obedece a que el vector normal al OSH está dado por la combinación lineal de los vectores de soporte tal como se muestra en la ecuación 2.3, los escalares λ_k que se muestran allí son las variables duales del problema de optimización de la ecuación 2.5 afectados por el signo de la clase a la que corresponde cada vector de soporte.

Combinación Lineal de los Vectores de Soporte

$$w = \sum_{k \in \substack{\text{Vectores} \\ \text{Soporte}}} \lambda_k x_k \quad (2.3)$$

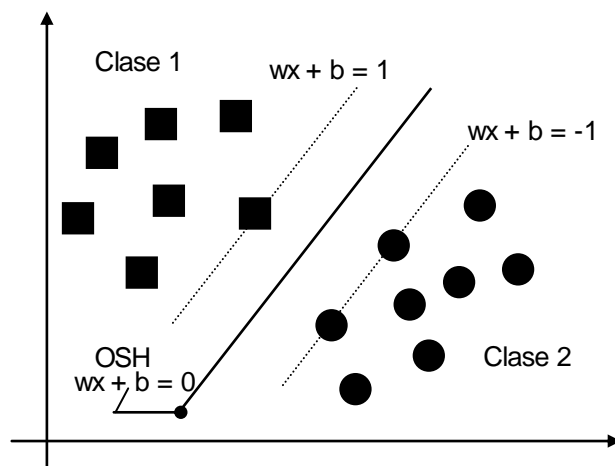


Figura 2.2. Plano de Separación Optimo para el problema de Biclasiación.

Por medio de algunas manipulaciones algebraicas las ecuaciones cartesianas de los dos hiperplanos paralelos al OSH pueden escribirse como se muestra en la ecuación 2.4. La distancia entre estos dos planos es conocida como el margen de separación entre las clases y el problema de maximización de este margen puede escribirse como el problema de minimización del inverso del cuadrado del margen, sujeto a las restricciones descritas en la ecuación 2.1, tal como se muestra en la ecuación 2.5.

Planos paralelos al OSH.

$$(\langle \vec{w}, \vec{x}_i \rangle + b) = 1 \quad \text{y} \quad (\langle \vec{w}, \vec{x}_i \rangle + b) = -1 \quad (2.4)$$

Problema de Optimización Asociado a las SVM lineal biclasificadora

$$\begin{aligned} \underset{w,b}{\text{Min}} \quad & \frac{\|\vec{w}\|^2}{2} \\ & y_i (\langle \vec{w}, \vec{x}_i \rangle + b) \geq 0 \quad i = 1, 2, \dots, N \end{aligned} \quad (2.5)$$

El problema de optimización de la ecuación 2.5 es un problema de programación cuadrático; para resolverlo, generalmente se utiliza su formulación dual (Condiciones de Karush-Kuhn-Tucker. *KKT*) tal como se muestra en la ecuación 2.6. Las variables α_i son las variables duales o multiplicadores de Lagrange del problema de optimización y son diferentes de cero sólo aquellos multiplicadores asociados a las restricciones asociadas a los vectores de soporte (Holguras Complementarias).

$$\text{Max} \quad \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle$$

$$\begin{aligned} \text{Sujeto a} \quad & \sum_{i=1}^N y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C \quad i = 1, 2, \dots, N \end{aligned} \quad (2.6)$$

De las condiciones KKT se tiene que (\mathbf{x}_1 y \mathbf{x}_{-1} son un vector de soporte cualquiera de cada una de las clases **1** y **-1** respectivamente)

$$\begin{aligned} \vec{w} &= \sum_{j=1}^n \alpha_j y_j \vec{x}_j \\ b &= -\frac{1}{2} \langle \vec{w}, (\vec{x}_1 + \vec{x}_{-1}) \rangle \end{aligned} \quad (2.7)$$

Donde \mathbf{x}_1 es un vector de soporte de la clase **1** y \mathbf{x}_{-1} son los vectores de soporte de la clase **-1**.

Una vez hallada la solución del problema de maximización de la ecuación 2.6, la manera de validar la pertenencia de un dato \mathbf{x}_m es

$$f(\bar{x}_m) = \text{signum}\left(\sum_{i=1}^n \alpha_i y_i \langle \bar{x}_i, \bar{x}_m \rangle + b\right) \quad (2.8)$$

2.2.2 Caso linealmente no separable para SVM para problemas de biclasificación

Para datos no son linealmente separables (ver Figura 2.), la solución al problema formulado en la ecuación 2.6 sería vacía ya que la región factible sería nula; con la finalidad de ampliar esta región factible, se pueden introducir variables de relajación $\xi_1, \xi_2, \dots, \xi_n$ con $\xi_i \geq 0, \forall i$ así

$$y_i (\langle \bar{w}, \bar{x}_i \rangle + b) \geq 1 + \xi_i \quad i = 1, 2, \dots, N \quad (2.9)$$

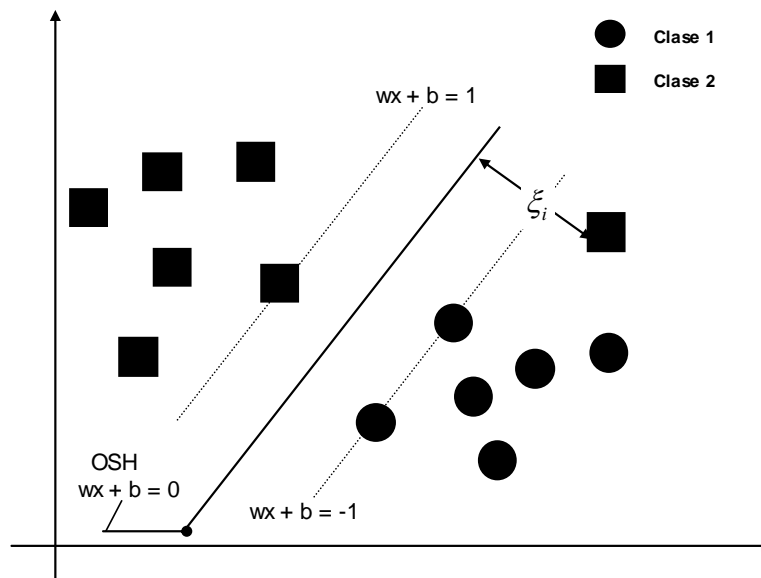


Figura 2.3. Datos No-linealmente separables.

El OSH puede ser encontrado mediante una variación a la función objetivo de la ecuación 2.5 tal como se presenta en la ecuación 2.10.

$$\text{Min} \quad \frac{\|\bar{w}\|^2}{2} + C \sum_{i=1}^N \xi_i \quad (2.10)$$

Sujeto a

$$y_i (\langle \bar{w}, \bar{x}_i \rangle + b) \geq 1 + \xi_i \quad i = 1, 2, \dots, N$$

El parámetro **C** debe ser de gran magnitud (elegida por el usuario) y su propósito es el de penalizar la función objetivo con cada dato que queda ubicado en una categoría a la cual no corresponda.

2.2.3 Máquinas de soporte vectorial no lineales para biclasificación

En la mayoría de aplicaciones de los problemas de clasificación, la separación de clases por medio de una formulación Lineal no es apropiada debido a la estructura de los datos. La idea de una SVM no lineal es la de transformar el espacio de entrada en un espacio de representación de dimensión mayor a través de una función no lineal elegida a priori. Si denominamos esta función como $\Phi(\cdot)$ para transformar cada dato del problema, la formulación de ecuación 2.6 da como resultado la ecuación 2.11.

$$\text{Max} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle \phi(x_i), \phi(x_j) \rangle$$

Sujeto a
$$\sum_{i=1}^N y_i \alpha_i = 0 \quad (2.11)$$

$$0 \leq \alpha_i \quad i = 1, 2, \dots, N$$

Y en forma matricial, puede escribirse como

$$\text{Max} [1 \dots 1] \cdot \vec{\alpha} - \frac{1}{2} \cdot \vec{\alpha}^T \cdot [H] \cdot \vec{\alpha}$$

Sujeto a

$$[y_1 \quad \dots \quad y_n] \cdot \vec{\alpha} = 0 \quad (2.12)$$

$$0 \leq \alpha_i \quad i = 1, 2, \dots, N$$

Donde H es una matriz $N \times N$ y su elemento i, j equivale a

$$H(i, j) = y_i \cdot y_j \cdot \langle \phi(x_i), \phi(x_j) \rangle$$

El producto interior $\langle \Phi(x_i), \Phi(x_j) \rangle$ mostrado en la ecuación 2.11 viene dado por las funciones de *Kernel*, de este modo siempre se evita el cálculo explícito de la función de transformación $\Phi(\cdot)$. Es por esta razón que para las SVM no lineales, el problema de optimización debe ser resuelto por medio de la formulación Dual ya que la implementación del problema primal requiere de la formulación explícita de la función de transformación. Entre las funciones de *kernel* más utilizadas tenemos:

Función de base radial

$$K(x_i, x_j) = \langle X_i, X_j \rangle = \langle \phi(x_i), \phi(x_j) \rangle = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (2.13)$$

Función Polinomial

$$K(x_i, x_j) = \langle X_i, X_j \rangle = \langle \phi(x_i), \phi(x_j) \rangle = (\langle x_i, x_j \rangle + 1)^p \quad (2.14)$$

$$(2.15)$$

Función Sigmoidal

$$K(x_i, x_j) = \langle X_i, X_j \rangle = \langle \phi(x_i), \phi(x_j) \rangle = \text{tgh}(\beta \langle x_i, x_j \rangle - \delta)$$

Para que una función “Kernel” represente el producto interior en el nuevo espacio es necesario que esta cumpla con las *Condiciones de Mercer* [Burges (2001)], las cuales son condiciones generalizadas de las condiciones de un producto interior como el definido en el álgebra lineal básica.

Una vez resuelto el problema de optimización de la ecuación 2.12, es posible encontrar la función de biclasificación, al extender las ecuaciones 2.7 en la ecuación 2.8 así:

$$\begin{aligned} \vec{w} &= \sum_{j=1}^n \alpha_j y_j \vec{x}_j \\ b &= -\frac{1}{2} \sum_{j=1}^n \alpha_j y_j \langle \vec{x}_j, (\vec{x}_1 + \vec{x}_{-1}) \rangle \\ f(\vec{x}) &= \text{signum} \left(\sum_{i=1}^n \alpha_i y_i \langle \vec{x}_i, \vec{x} \rangle + b \right) \end{aligned} \quad (2.16)$$

2.2.4 SVM para clasificación multiclase.

Hasta el momento sólo se han estudiado las SVM para clasificación binaria; sin embargo, en la práctica se encuentra que en la mayoría de las ocasiones se requiere hacer clasificación en conjuntos que poseen más de dos clases. Para realizar esta labor se podría pensar en crear una SVM que logré hacer multclasificación en un solo paso, sin embargo, esta opción no es la mejor dado que se debe solucionar un problema de optimización sujeto a muchas restricciones lo cual genera un alto costo computacional.

2.2.4.1 Multclasificación a partir de SVM biclasificadoras.

Se ha encontrado que la mejor opción en cuanto a la multclasificación con máquinas de soporte vectorial es crear una función clasificadora global a partir de funciones biclasificadoras binarias generalizadas, es decir, extender la herramienta de biclasificación a una herramienta de multclasificación que sea más ajustada a las necesidades reales. En la actualidad existen varias arquitecturas básicas acerca de la multclasificación con máquinas de soporte vectorial biclasificadoras, de las cuales las que más sobresalen son

- Uno contra uno

Esta arquitectura de clasificación se realiza implementando k clasificadores binarios o biclasificadores donde $k = l(l-1)/2$ y l corresponde al número de clases que se desean separar ($l > 2$). De esta manera para el entrenamiento de la i -ésima SVM biclasificadora se toman solo dos de las l clases existentes, asignando etiqueta positiva a los datos \mathbf{n}_i que pertenecen a la clase i ($\mathbf{t}_i = +1$) y etiqueta negativa a los datos \mathbf{n}_j que pertenecen a la clase j ($\mathbf{t}_j = -1$), los datos de las demás clases $(1, 2, \dots, i-1, i+1, j-1, j+1, \dots, l)$ denominados \mathbf{n}_r , no son utilizadas en este

entrenamiento y se les asigna un cero, creándose así la matriz de descomposición cuyos elementos se definen por la ecuación 2.17.

$$D = \begin{cases} +1 & \text{si } n_F \in n_i \\ -1 & \text{si } n_F \in n_j \\ 0 & \text{si } n_F \in n_r \end{cases} \quad (2.17)$$

Por ejemplo para una clasificación de cuatro clases la matriz de descomposición que se genera es la siguiente

$$D_{O-v-o} = \begin{bmatrix} +1 & -1 & 0 & 0 \\ +1 & 0 & -1 & 0 \\ +1 & 0 & 0 & -1 \\ 0 & +1 & -1 & 0 \\ 0 & +1 & 0 & -1 \\ 0 & 0 & +1 & -1 \end{bmatrix}$$

- Uno contra el resto

En esta arquitectura de clasificación se implementan tantos clasificadores binarios como clases o categorías existan en el problema, en otras palabras se requieren k clasificadores binarios ($k = l$), los cuales se encargan de diferenciar cada clase de las demás.

La clasificación se realiza tomando un dato de la clase j ($j \in \{1, 2, \dots, l\}$) y asignándole una etiqueta positiva ($t_j = +1$), y para los demás datos pertenecientes a las $l-1$ clases restantes se les asigna una etiqueta negativa ($t_r = -1$). Así, los elementos que conforman esta matriz se definen por la ecuación 2.18.

$$D = \begin{cases} +1 & \text{si } n_F \in n_j \\ -1 & \text{si } n_F \in n_r \end{cases} \quad (2.18)$$

De esta manera, para el ejemplo anterior (4 clases), la matriz de descomposición que se genera en esta arquitectura es

$$D_{O-v-R} = \begin{bmatrix} +1 & -1 & -1 & -1 \\ -1 & +1 & -1 & -1 \\ -1 & -1 & +1 & -1 \\ -1 & -1 & -1 & +1 \end{bmatrix}$$

- Arquitectura tipo ECOC

Este método utiliza la codificación estándar para obtener robustez contra fallos en las arquitecturas expuestas anteriormente. Esta arquitectura utiliza una codificación estándar de cada una de las posibles particiones de todo el conjunto de clases. En esta arquitectura se asigna etiqueta positiva ($t_i = +1$) a un subconjunto de clases N_i , y etiqueta negativa ($t_r = -1$) al subconjunto compuesto por

el resto de clases N_r . Así, los términos de la matriz de descomposición se obtienen a partir de la ecuación 2.19.

$$D = \begin{cases} +1 & \text{si } N_F \in N_i \\ -1 & \text{si } N_F \in N_r \end{cases} \quad (2.19)$$

De esta manera la matriz de descomposición generada debe tener sus columnas tan diferentes una de otra como sea posible, esto con el fin de crear redundancia; el numero de clasificadores binarios que se generan es $k = 2^{l-1} - 1$, donde l corresponde al número de clases ($l > 2$). Para el ejemplo que se ha venido trabajando con cuatro clases ($k = 7$) se tiene que la matriz de descomposición que se genera es la siguiente:

$$D_{\text{ococ}} = \begin{bmatrix} +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & -1 & +1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & +1 & -1 & +1 \\ +1 & +1 & +1 & -1 \end{bmatrix}$$

2.2.4.2 Métodos de reconstrucción.

Con la utilización de las arquitecturas de descomposición estándares se ha dividido el problema de multclasificación en un problema de k clasificadores binarios que entregan una respuesta numérica para cada entrada; estos resultados se deben interpretar de manera adecuada para obtener la respuesta general al problema de clasificación.

Las respuestas numéricas producidas por los clasificadores binarios son +1 para elementos que se clasifican con etiqueta positiva, -1 para elementos clasificados con etiqueta negativa y 0 para elementos que no son tenidos en cuenta en la clasificación.

Después de interpretar las predicciones hechas por la arquitectura de Descomposición, es necesario realizar un recuento de las clases votadas para cada dato, lo que se conoce como esquema de reconstrucción. En este paso se interpretan los signos de las predicciones, según el método de descomposición empleado así:

- Asignación de Votos para arquitectura O-v-O

Si denominamos como S^{i-j} al signo de la predicción del i - j -ésimo clasificador binario y denominamos como C_i y C_j a las Clases involucradas tenemos que la asignación de Votos será

$$\Theta = \begin{cases} C_i & \text{si } S^{i-j} = +1 \\ C_j & \text{si } S^{i-j} = -1 \end{cases} \quad (2.20)$$

La asignación de Voto será para la clase i cuando el signo de validación de un dato en el biclasificador i - j sea +1; para el caso en el que la validación sea -1 la asignación de voto será para

la clase j . Las demás clases no recibirán votos, dado que no están involucradas en el proceso del biclasificador i - j .

- Asignación de Votos para arquitectura O-v-O

Si denominamos como S^i al signo de la predicción del i -ésimo clasificador binario y denominamos como C_i a la Clase involucrada tenemos que la asignación de Votos será

$$\Theta = \begin{cases} C_i & \text{si } S^i = +1 \\ 0 & \text{si } S^i = -1 \end{cases} \quad (2.21)$$

La asignación de Voto será sólo para la clase que define el biclasificador cuando el signo de de la validación de un dato en éste sea +1; para el caso en el que la validación sea -1 la asignación de voto no se hará a ninguna de las clases.

- Asignación de Votos para arquitectura ECOC

La asignación de Votos en esta arquitectura es similar a la arquitectura O-v-O, pues asigna voto a las clases involucradas de acuerdo al signo de la validación de un dato en el biclasificador.

$$\Theta = \begin{cases} C_i & \text{si } S^{i-j} = +1 \\ C_j & \text{si } S^{i-j} = -1 \end{cases} \quad (2.22)$$

La asignación de Voto será sólo para la(s) clase(s) que define el biclasificador cuando el signo de la validación de un dato en éste sea +1; para el caso en el que la validación sea -1 la asignación de voto no se hará a ninguna de las clases.

2.2.4.3 Esquemas de Votación

Después de la implementación del método de reconstrucción (asignación de Votos) según la arquitectura implementada, se procede a tomar decisiones a partir de alguna de las siguientes posibilidades

- *Votación por Unanimidad*: se determina como respuesta aquella clase que haya obtenido todos los votos posibles en las predicciones.
- *Votación por mayoría absoluta*: Se determina como respuesta aquella única clase que haya obtenido más de la mitad de los votos posibles.
- *Votación por mayoría simple*: Se determina como respuesta aquella única clase que haya obtenido más votos que las demás.

2.2.5 Ejemplo de aplicación de una máquina de soporte vectorial para clasificación no lineal.

El siguiente ejemplo presenta una máquina de soporte vectorial biclasificadora aplicada para la solución del problema de localización de fallas en el sistema de distribución monofásico presentado en la figura 2.4.

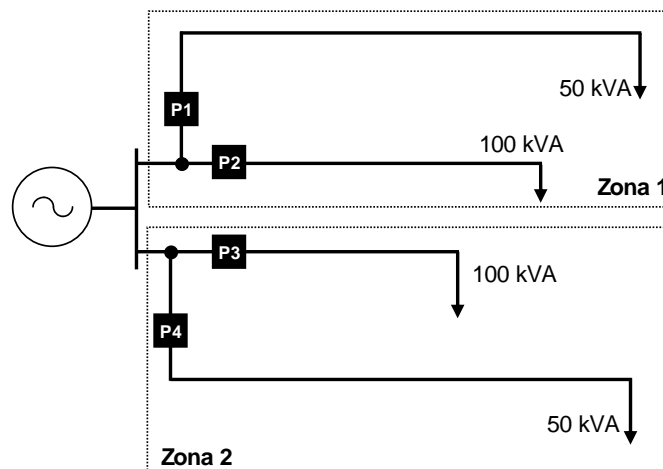


Figura 2.4. Sistema de distribución elemental para ejemplificar la estimación de la zona de falla.

El voltaje en la cabecera del sistema es de 10 kV; al interior del sistema existen cuatro protecciones P1, P2, P3 y P4, cada una con un tiempo de acción de 400, 200, 400 y 200 milisegundos para las fallas que ocurren aguas abajo de cada una de ellas respectivamente. La corriente de estado estable medida en la cabecera del sistema es aproximadamente 30 Amperios (despreciando la caída de tensión a lo largo de los ramales), y ante fallas ocurridas aguas abajo de las protecciones P1, P2, P3 y P4 la corriente de estado estable experimentará una reducción de 5, 10, 10 y 5 Amperios respectivamente (asociados a la liberación de la falla por parte de las protecciones). Adicionalmente el sistema de distribución ha sido dividido en dos zonas, la primera contiene los ramales que se encuentran aguas abajo de las protecciones P1 y P2, y la segunda contiene los ramales que se encuentran aguas abajo de las protecciones P3 y P4.

Con la información anterior podemos definir cuatro tipos de falla monofásicas que pueden ocurrir en el sistema; cada una de ellas está asociada a los ramales que se encuentran aguas abajo de cada una de las protecciones. Las características de interés que describen cada tipo de falla se presentan a continuación.

Protección asociada	Magnitud de reducción de corriente. [A]	Tiempo de acción de la protección [ms]	Zona
P1	5.0	400	1
P2	10.0	200	1
P3	10.0	400	2
P4	5.0	200	2

Estas características tomadas del sistema de distribución bajo estado de falla se conocen como descriptores del sistema y se presentan con profundidad en el capítulo 3; las características de reducción de corriente y tiempo de acción de las protecciones pueden obtenerse fácilmente a partir del patrón de corriente medida en la cabecera del sistema, tal como se presenta en la sección 3.5.

La figura 2.5 presenta la información anterior en un gráfico bidimensional donde las abscisas representan la magnitud de reducción de corriente, las ordenadas representan el tiempo de acción y los símbolos “■” y “●” representan las zonas 1 y 2 del sistema respectivamente.

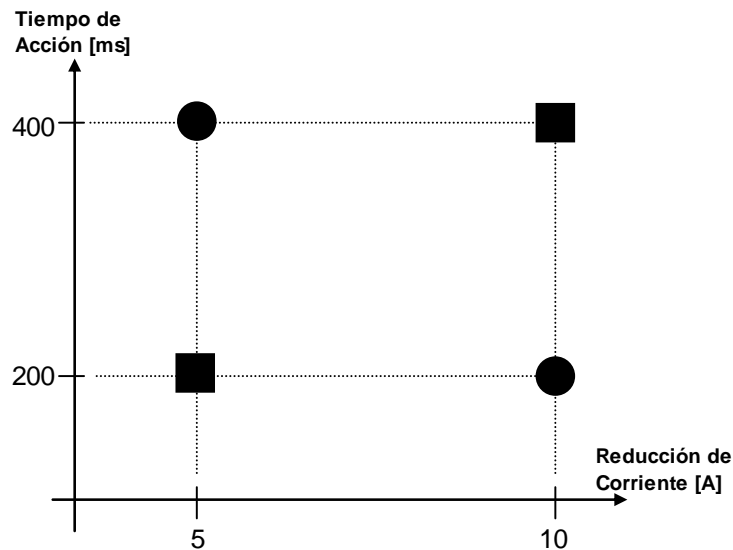


Figura 2.5 Representación gráfica de las características de falla.

Se observa que no existe un hiperplano (específicamente una línea recta) que permita separar los datos de cada zona del sistema de distribución (clases). Mediante un escalamiento¹ adecuado, el problema anterior puede ser representado como el problema de clasificación XOR, es decir, donde los cuatro datos presentados ocupan las posiciones (0,1), (1,0), (1,1) y el origen, y la primera clase está asociada con el primer par de datos y la segunda clase está asociada con el segundo par de datos, tal como se presenta en la figura 2.6.

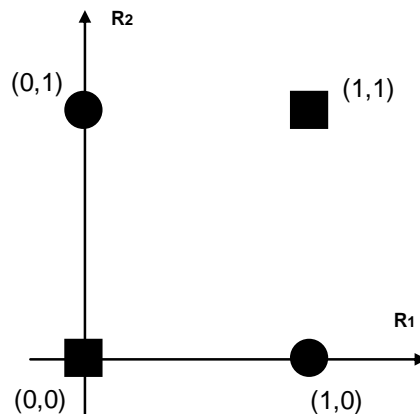


Figura 2.6. Representación gráfica de las características de falla. Problema de clasificación XOR

Para dar solución al problema de clasificación de los datos de la figura 2.6 se debe tener presente que no existe un hiperplano que permita separar dichas clases (zonas del sistema de distribución). Para superar esta característica, emplearemos una de las funciones kernel presentadas en la sección 0, específicamente emplearemos el kernel polinomial de segundo orden ($p=2$) para

¹ Si definimos \mathbf{R} como la reducción de corriente y t al tiempo de acción, escalamiento que transforma los datos de la Figura 2. en los datos de la Figura 2.6 es:

$$\begin{bmatrix} X1 \\ X2 \end{bmatrix} = \begin{bmatrix} 1/5 & 0 \\ 0 & 1/200 \end{bmatrix} \begin{bmatrix} R \\ t \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

representar el producto interior de los datos transformados desde el espacio $\mathbf{R}_1-\mathbf{R}_2$ a un espacio de mayor dimensión donde los datos si podrán ser separados. Los datos del problema son:

Datos a clasificar: $x_1 = (0,0); x_2 = (1,1); x_3 = (0,1); x_4 = (1,0);$

Clases a las cuales está asociado cada dato: $y_1 = y_2 = 1; y_3 = y_4 = -1;$

Elementos de la matriz \mathbf{H} (ecuación 2.14): $H(i, j) = y_i y_j (\bar{x}_i \cdot \bar{x}_j + 1)^2$

Al representar de forma matricial el problema de clasificación tenemos:

$$\underset{\alpha_1, \alpha_2, \alpha_3, \alpha_4}{Max} \sum_{i=1}^4 \alpha_i - \frac{1}{2} \cdot [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4] \cdot \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 9 & -4 & -4 \\ -1 & -4 & 4 & 1 \\ -1 & -4 & 1 & 4 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}$$

Sujeto a

$$\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0$$

$$0 \leq \alpha_i \quad i = 1, \dots, 4$$

Para la solución de este problema, se construye la función de Lagrange llevando la restricción a la función objetivo afectada por el multiplicador de Lagrange λ respectivo.

$$L(\alpha_1, \dots, \alpha_4, \lambda) = (\alpha_1 + \alpha_2)(1 - \lambda) + (\alpha_3 + \alpha_4)(1 + \lambda) - \frac{\alpha_1^2 + 9\alpha_2^2 + 4\alpha_3^2 + 4\alpha_4^2}{2} - \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_1\alpha_4 + 4\alpha_2\alpha_3 + 4\alpha_2\alpha_4 - \alpha_3\alpha_4$$

Al igualar a cero el gradiente de la función L se tiene el siguiente sistema de ecuaciones lineales:

$$\begin{bmatrix} 1 & 1 & -1 & -1 & 1 \\ 1 & 9 & -4 & -4 & 1 \\ -1 & -4 & 4 & 4 & -1 \\ -1 & -4 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

Y la solución del sistema de ecuaciones es:

$$\alpha_1 = 10/3; \alpha_2 = 2; \alpha_3 = 8/3; \alpha_4 = 8/3; \lambda = 1$$

Una vez halladas las variables α_i es posible encontrar la función de clasificación de todo el espacio $\mathbf{R}_1-\mathbf{R}_2$ mediante las ecuaciones 2.16 así.

$$b = -\frac{1}{2} \sum_{j=1}^n \alpha_j y_j (\bar{x}_j \cdot ((0,0) + (1,0)) + 1)^2 = 1$$

$$f(\bar{r}) = \text{signum} \left(\sum_{i=1}^n \alpha_i y_i (\bar{x}_i \cdot (r_1, r_2) + 1)^2 + b \right)$$

$$= \text{signum} \left(-\frac{2}{3} (r_1^2 + r_2^2) - \frac{4}{3} (r_1 + r_2) + 4r_1 r_2 + 1 \right)$$

El argumento de la función de signo mostrada en la ecuación anterior corresponde con un hiperboloide, el cual es la función de separación de ambas clases. La figura 2.7 presenta esta función y se observa que los datos (0,0) y (1,1) correspondientes a la primera clase, han sido transformados en los datos (0, 0,1) y (1, 1,1) y que los datos (0,1) y (1,0) han sido transformados en los datos (0, 1,-1) y (1, 0,-1). La figura 2.8 muestra el resultado de aplicar la función de signo a los datos transformados por la función de separación.

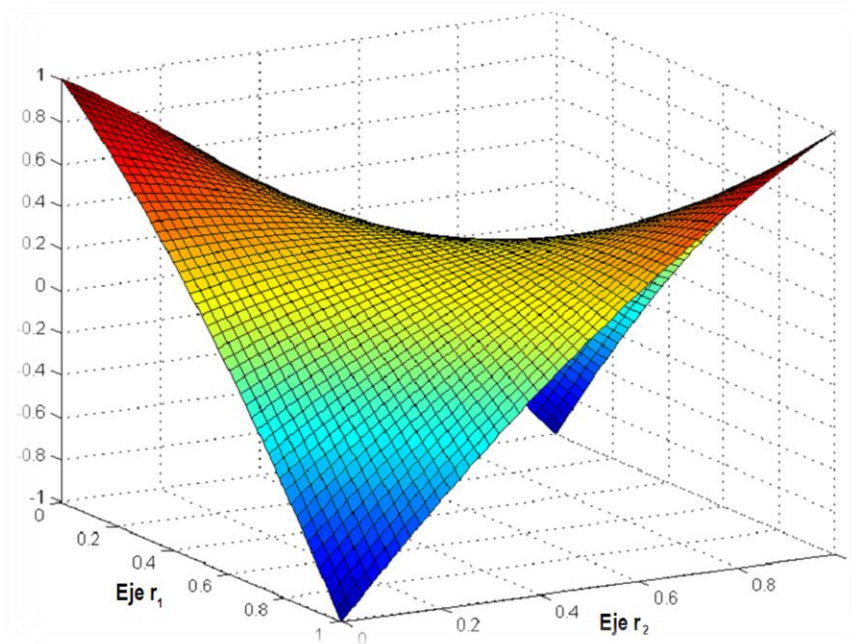


Figura 2.7. Hiperboloide de separación entre clases, para el problema de clasificación XOR

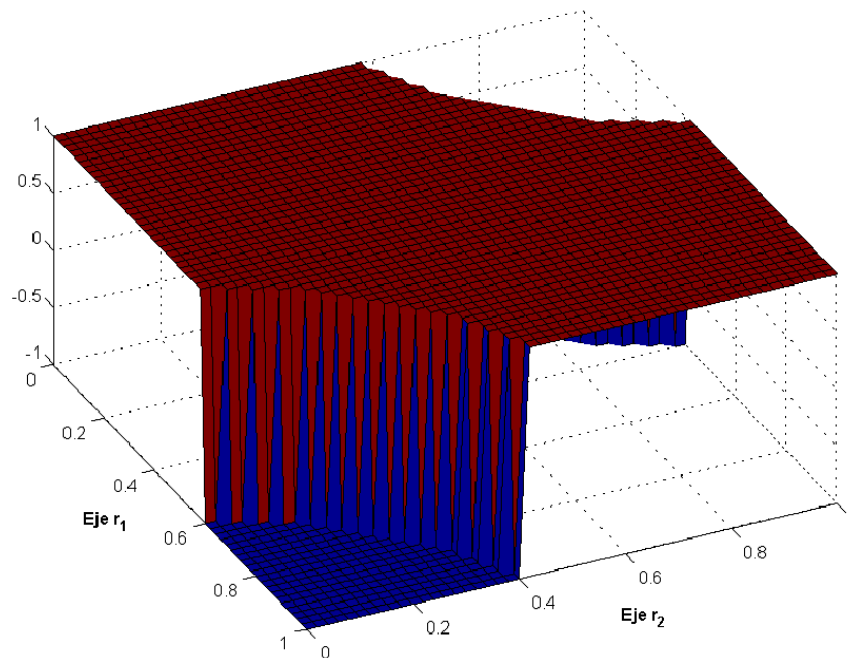


Figura 2.8. Función de Clasificación. Se presenta cómo es clasificado el espacio definido por $[0, 1] \times [0, 1]$.

Mediante un cambio de variables adecuado se puede eliminar el término que involucra el producto de las variables r_1, r_2 ; concluyendo finalmente que las curvas de nivel de la función de separación corresponden con hipérbolas con centro en el punto $(1/2, 1/2)$ y sus ejes están rotados 45° respecto de los ejes r_1 y r_2 tal como se observa en la figura 2.9.

$$r_1 = \frac{\sqrt{2}}{2} \bar{r}_1 - \frac{\sqrt{2}}{2} \bar{r}_2 \quad r_2 = \frac{\sqrt{2}}{2} \bar{r}_1 + \frac{\sqrt{2}}{2} \bar{r}_2$$

$$\frac{\left(\bar{r}_1 - \frac{\sqrt{2}}{2}\right)^2}{\frac{3}{4}} - \frac{\bar{r}_2^2}{\frac{3}{8}} = C - \frac{1}{3}$$

El parámetro C del lado derecho de la ecuación anterior representa los diferentes valores con los que se definen las curvas de nivel de la función de separación de ambas clases. Se observa que las hipérbolas cambian su eje focal para un valor de C superior a $1/3$.

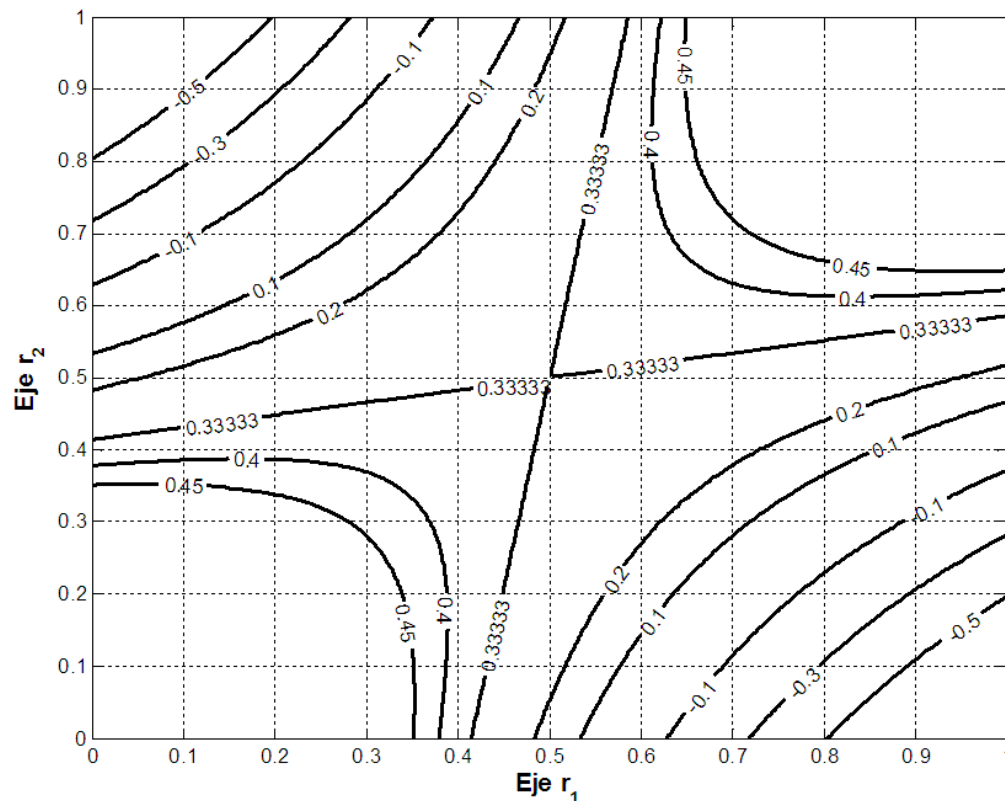


Figura 2.9. Curvas de Nivel de la Función de Clasificación.

De esta forma hemos encontrado una función de localización de fallas, la cual asocia un par ordenado (“magnitud de reducción de corriente”, “tiempo de acción de la protección”) y lo asocia a una de las zonas definidas en el sistema de distribución de la figura 2.4. En la práctica los sistemas de distribución poseen un numero de ramales amplio, por tal motivo es necesario emplear una máquina de soporte vectorial multiclase que permita clasificar fallas en un sistema de distribución con más de dos zonas.

2.3 MÁQUINAS DE SOPORTE VECTORIAL PARA REGRESIÓN

Las SVM no sólo han sido aplicadas al problema de clasificación de variables sino que también ha sido aplicada exitosamente al problema de regresión o de estimación de variables, por medio de la introducción de una función de pérdidas tal como se presenta a continuación:

2.3.1 Problema básico

La formulación de las SVM para regresión (SVMr) puede abordarse de la siguiente manera. Considere un conjunto de \mathbf{N} datos de la forma $(\mathbf{x}_i, \mathbf{y}_i)$ con $1 \leq i \leq \mathbf{N}$, donde $\mathbf{x}_i \in \mathbf{R}^p$ y $\mathbf{y}_i \in \mathbf{R}$; se desea obtener una función lineal como la que se muestra en la ecuación 2.23 que produzca el mínimo error de estimación para los \mathbf{N} datos.

$$f(\vec{x}) = \langle \vec{w}, \vec{x} \rangle + b \quad (2.23)$$

Donde $\langle \cdot \rangle$ representa el producto interior Euclídeo para \mathbf{R}^P . La obtención de esta función puede formularse como un problema donde se minimice el cuadrado de la norma Euclídea del vector \vec{w} sujeto a que los datos se encuentren al interior de un margen definido por dicha función y un parámetro de precisión ε , escogido por el usuario, así

$$\underset{w}{Min} \frac{\|\vec{w}\|^2}{2}$$

Sujeto a:

$$\begin{aligned} y_i - \langle \vec{w}, \vec{x}_i \rangle - b &\leq \varepsilon \\ \langle \vec{w}, \vec{x}_i \rangle + b - y_i &\leq \varepsilon \end{aligned} \quad (2.24)$$

Las restricciones mostradas en la ecuación 2.24 hacen que la función lineal se aproxime a los datos del problema de tal manera que ellos estén contenidos en un margen de ε unidades por encima y por debajo de la función misma (figura 2.10).

En la mayoría de las aplicaciones, es difícil que el problema formulado en la ecuación 2.24 posea solución factible dado que pueden existir *outliers* o datos que no se pueden ajustar a un margen ε pequeño alrededor de la función; por tal razón es importante incluir en el problema de minimización las variables de relajación ξ y ξ^* (ecuación 2.25) que permitan cierto error en la estimación y que además conviertan en un espacio factible al espacio de soluciones del problema.

La introducción de las variables de holgura en las restricciones del problema de la ecuación 2.24, viene acompañada de una penalización a la función objetivo mediante una función de penalización; de acuerdo con la función de penalización escogida será de mayor o menor complejidad la solución del problema de regresión; entre las funciones de penalización reportadas en la literatura se encuentran la *función cuadrática*, *función de Laplace*, *función de Hubber* y la *función de Margen Suave*, siendo esta última la de mayor implementación debido a la reducción matemática simple a la que se puede llevar el problema de optimización dual. Es importante anotar que por cada dato en el problema de regresión existen dos restricciones que a su vez implican dos variables en el problema dual; la solución al problema de optimización mostrado en la ecuación 2.25 siempre se realiza a través del problema dual ya que permite también la implementación de funciones de regresión no lineales.

$$\underset{w, \xi, \xi^*}{Min} \frac{\|\vec{w}\|^2}{2} + C \sum_{i=1}^N (\xi_i + \xi_i^*)$$

Sujeto a:

$$\begin{aligned} y_i - \langle \vec{w}, \vec{x}_i \rangle - b &\leq \varepsilon + \xi_i \\ \langle \vec{w}, \vec{x}_i \rangle + b - y_i &\leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0 \end{aligned} \quad (2.25)$$

La figura 2.10 presenta una interpretación del problema para el caso bidimensional; se puede observar que para los datos al interior del margen definido por ε , las restricciones de la ecuación 2.24 no se encontraran activas y para los datos que se encuentran en el límite del margen y por fuera de él las restricciones de esta misma ecuación estarán activas (holgura cero) y por lo tanto, la variable dual asociada será diferente de cero (condiciones KKT).

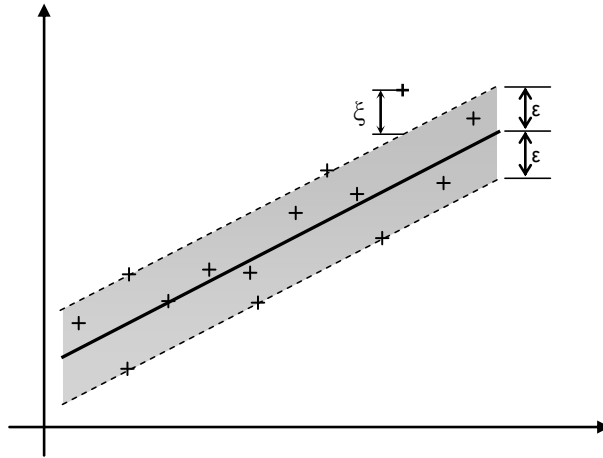


Figura 2.10. Margen definido para el problema de regresión de las SVMr

La formulación dual del problema mostrado en la ecuación 2.25 se muestra en la ecuación 2.26 y la función de regresión, también conocida como la *función de regresión expandida en Vectores de Soporte* se muestra en la ecuación 2.27.

$$\text{Max}_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle \bar{x}_i, \bar{x}_j \rangle - \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) + \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*)$$

Sujeto a:

$$\sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0$$

$$\alpha_i, \alpha_i^* \in [0, C]$$

(2.26)

$$f(\bar{x}) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \langle \bar{x}_i, \bar{x} \rangle + b$$

(2.27)

2.3.2 Máquinas de soporte vectorial no lineales para regresión

En la mayoría de aplicaciones de los problemas de regresión, la estimación de un conjunto de datos por medio de una formulación Lineal no es apropiada debido a la estructura de los mismos. La idea de una SVM no lineal para regresión es la de transformar el espacio de entrada en un espacio de representación de dimensión mayor a través de una función no lineal elegida a priori. Si denominamos esta función como $\Phi(\cdot)$ para transformar cada dato del problema, la formulación de ecuación 2.28 es

$$\text{Max}_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle \Phi(\bar{x}_i), \Phi(\bar{x}_j) \rangle - \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) + \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*)$$

(2.28)

Sujeto a:

$$\sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0$$

$$\alpha_i, \alpha_i^* \in [0, C]$$

El producto interior $\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$ viene dado por las funciones de *Kernel* y de este modo siempre se evita el cálculo explícito de la función de transformación $\Phi(\cdot)$. Es por esta razón que para las SVM no lineales, el problema de optimización debe ser resuelto por medio de la formulación Dual ya que la formulación primal requiere de una función de transformación en forma explícita la cual es de difícil obtención. Entre las funciones de *kernel* más utilizadas tenemos las funciones mostradas en las ecuaciones 2.13 a 2.15.

La matriz Hessiana de la función objetivo del problema presentado en 2.28 es una matriz cuya dimensión es dos veces el número de datos de entrenamiento (existe un α y un α^* por cada dato x_i en el problema). Adicionalmente esta matriz presenta singularidades y una de las maneras de eludirlas, para el proceso de solución del problema de optimización, es mediante la introducción de un pequeño ruido Gaussiano a las componentes de dicha matriz [Gunn 1998].

2.4 SINTONIZACIÓN DE PARÁMETROS PARA LA SVM

La sintonización de parámetros está asociada a los parámetros de penalización C de la función objetivo de las ecuaciones 2.11 y 2.28, adicionalmente los parámetros de los diferentes kernels presentados en las ecuaciones 2.13 a 2.15. En el proceso de entrenamiento, estos parámetros deberán ser fijados por el usuario de manera tal que el desempeño de la SVM estará supeditado al valor de dichos parámetros. Por esta razón se realiza una etapa de sintonización de los parámetros de penalización y de kernel; una de las estrategias de mayor implementación es la *Búsqueda en Malla* y la *Validación Cruzada* [Lin 2004].

La validación cruzada consiste en que para un mismo conjunto de datos de entrenamiento (por ejemplo un 70% de la totalidad de datos disponibles), se debe entrenar múltiples veces la SVM, es decir, se divide el conjunto de datos de entrenamiento en m subconjuntos y realizan m entrenamientos para la SVM; para cada entrenamiento se utilizan $m-1$ subconjuntos de los m disponibles y el subconjunto restante se utiliza como subconjunto de datos para la validación. Finalmente se promedian los m errores de validación arrojados por cada una de las SVM.

La *Búsqueda en Malla* consiste en repetir el procedimiento de la *Validación Cruzada* para cada punto de una rejilla; la *Búsqueda en Malla* consiste en realizar un barrido a lo largo de una rejilla que involucre en una dimensión el parámetro de penalización C y en las otras dimensiones los parámetros del *kernel* (ver figura 2.11); para el caso de un *kernel* de Base Radial, el parámetro involucrado es σ . La rejilla escogida debe garantizar que el dominio de los parámetros sea apropiado de acuerdo con el objetivo y función del mismo.

Finalmente, se escoge el punto de la Rejilla para el cual el Error de Validación Promedio en el proceso de la *Validación Cruzada* es mínimo y alrededor de éste se elabora una nueva rejilla de menores dimensiones pero con la misma cantidad de divisiones de la rejilla anterior, para nuevamente repetir el proceso de la *Validación Cruzada* con los nuevos puntos de la nueva rejilla.

El proceso se repite hasta alcanzar un nivel de precisión deseado, o hasta determinar que el error de validación no mejora una cantidad preestablecida.

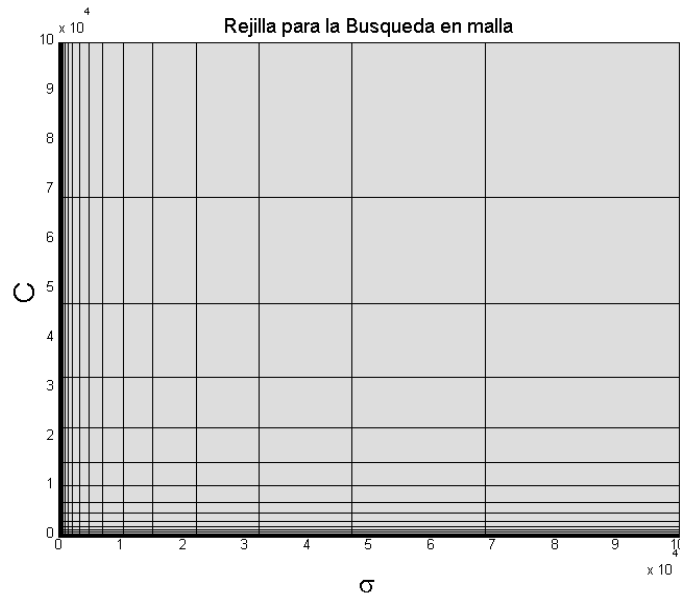


Figura 2.11. Rejilla para la sintonización de parámetros a través de la búsqueda en malla.

Una propuesta para la sintonización de parámetros para *kernels* que poseen varios de ellos (como el kernel sigmooidal), es la implementación de técnicas heurísticas y/o evolutivas; éstas requieren de una función de adaptación que evalúe el desempeño para cada iteración del método (en cada iteración se escogen nuevos parámetros para la SVM) y en este caso la función de adaptación debe ser el error de validación. Las variables deben codificarse de tal manera que el dominio de los parámetros sea apropiado de acuerdo con el objetivo y función del mismo (una alternativa es la codificación binaria; mediante la implementación de un número determinado de bits se puede barrer de forma discreta el intervalo de dominio de un parámetro).

2.5 NORMALIZACIÓN DE LOS DATOS

Como se acostumbra para la gran mayoría de las MBC utilizadas para trabajar con información diversa, es necesario realizar una normalización de los datos para optimizar el desempeño de las mismas. Una clase de normalización de amplia difusión, por su fácil implementación y fácil reconstrucción, es la Normalización Lineal y de escala $[-1,1]$; esta normalización consiste en dividir al conjunto de todos los datos entre el mayor de los valores absolutos de cada uno de ellos; teniendo presente la dependencia entre las variables; no es conveniente normalizar cada variable de manera independiente sin antes validar la relación existente con otras variables, pues podría omitirse información importante y descriptiva del fenómeno en estudio; un caso típico en el problema de localización de fallas sucede cuando se posee información sobre una característica asociada a cada fase (por ejemplo la corriente de falla), normalizar los datos por fase sería incorrecto ya que se omitiría la información que diferencia el comportamiento de las variables de las fases entre sí (para el ejemplo citado, se perdería la información que permitiría detectar la fase bajo falla).

3. METODOLOGIA PROPUESTA PARA LA LOCALIZACION DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN

3.1 INTRODUCCION

A partir de la metodología basada en el conocimiento (MBC) presentada en el capítulo anterior, en este capítulo se presenta el problema de localización de fallas desde el punto de vista de las técnicas inteligentes. Se plantean aquí las condiciones necesarias requeridas para resolver este problema a través de las máquinas de soporte vectorial para clasificación (SVMc) y regresión (SVMr).

En este capítulo se describe cada una de las etapas que componen la metodología propuesta, iniciando por el modelamiento del sistema de distribución, continuando con la zonificación de la red, la simulación automática de fallas, la caracterización del sistema bajo falla y la estructura soportada por una MBC que permitirá localizar la zona y estimar la distancia de ocurrencia de los eventos de falla.

Las secciones siguientes describen de manera sistemática las etapas metodológicas propuestas para la solución del problema de localización de fallas en sistemas de distribución. La sección 3.3 muestra la etapa de modelamiento del sistema, detallando los diferentes elementos que poseen las redes de distribución y el papel que desempeñan las protecciones para la localización de la zona de falla. La sección 3.4 describe los aspectos relativos a la simulación eléctrica de la red, las herramientas empleadas para realizar dichas simulaciones de manera automática y la obtención de la información del comportamiento del sistema para los diferentes puntos donde pueda ocurrir una falla. La sección 3.5 describe la información necesaria que debe ser extraída de los resultados de la simulaciones automáticas para caracterizar al sistema de distribución bajo estado de falla; asimismo se describe cómo organizar la información en la forma de datos Entrada-Salida con el fin de preparar la información para el proceso de entrenamiento de la MBC. Finalmente, las secciones 3.6 en adelante describen la MBC propuesta como herramienta de solución al problema de localización de la zona y estimación de la distancia de falla

A lo largo de este capítulo se hace referencia al sistema de distribución prototipo que fue escogido para realizar las pruebas e implementación de la metodología aquí expuesta, sin embargo será sólo hasta el capítulo 4 donde se exponga en detalle dicha red.

3.2 DESCRIPCION DE LA METODOLOGIA

El desarrollo de una metodología basada en el conocimiento (MBC) para la localización de fallas requiere de un planteamiento metodológico por etapas que contemple las herramientas, tanto para obtención y extracción de la información como el entrenamiento y validación de la MBC escogida. La figura 3.1 presenta el esquema de etapas mencionado.

Para la correcta implementación de la MBC es necesario contar con un volumen amplio de información del sistema de distribución en estado de falla; generalmente las empresas prestadoras del servicio no poseen dicha información y obtenerla en forma real es impráctico, es por esto que las primeras dos etapas de esta metodología consisten en el modelamiento y la simulación de los sistemas eléctricos como herramienta fundamental para resolver el problema de localización de fallas. En el numeral 3.3 se explica de manera detallada los elementos que conforman un sistema de distribución y los aspectos más importantes a tener en cuenta para su adecuado modelamiento y simulación. De allí se extrae información de las condiciones impuestas para la falla (tipo de falla, zona de falla, etc.) y los resultados de interés, como las señales de tensión y corriente, entre otra información que se describirá a profundidad en el numeral 3.5.

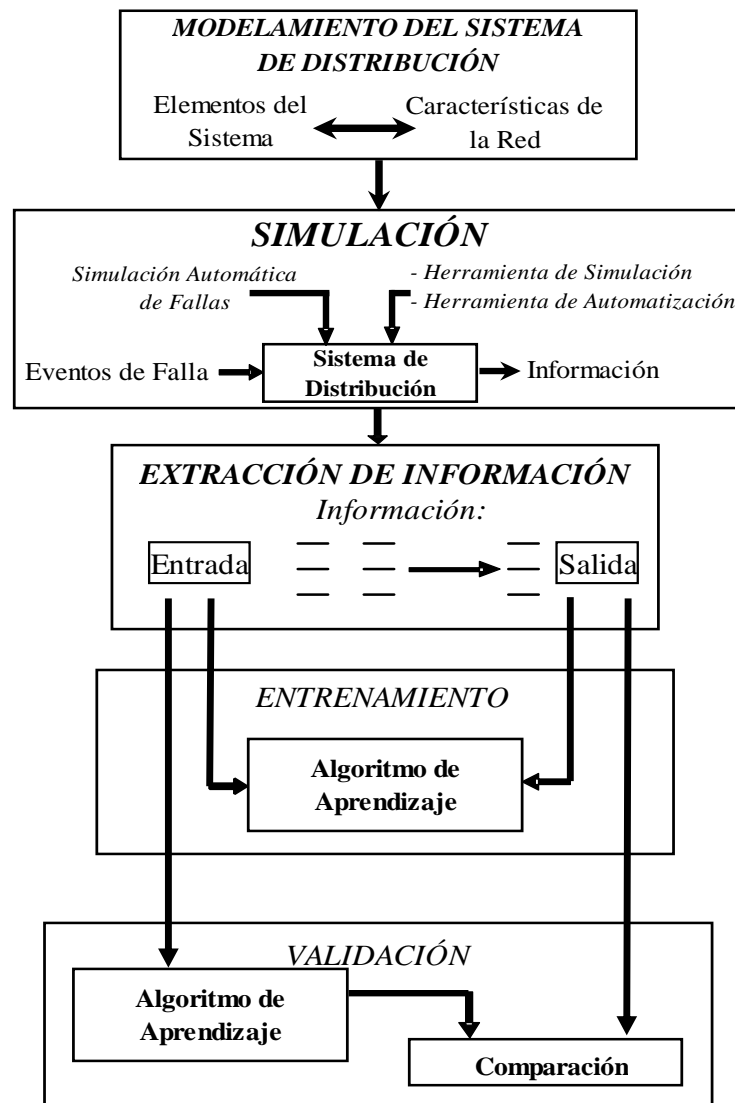


Figura 3.1. Esquema de las etapas metodológicas necesarias para construir una herramienta que permita modelar los conocimientos extraídos de las señales de tensión y corriente para resolver el problema de localización de fallas.

La tercera etapa consiste en formar una base de datos en la que esté contenida la información del sistema bajo diferentes condiciones de falla. Para obtener un volumen de información importante es necesario recurrir a la *simulación automática* de este tipo de eventos de forma tal que se pueda obtener información sobre el comportamiento de la red para cada tipo de falla. El proceso de simulación automatizado se describe en el numeral 3.4. La simulación de distintos eventos de fallas en diversos puntos de la red permitirá construir una base de datos con información que será organizada en el formato Entrada-Salida antes de iniciar la etapa de entrenamiento de la metodología basada en el conocimiento (MBC).

Los datos “Entrada” corresponden con los “Descriptoros del sistema” y los datos de “Salida” corresponden con la zona de Falla (para el caso de localización de la zona de falla) y/o con la distancia de falla (para el caso de estimación de la distancia de falla); este tipo de información se describen en el numeral 3.5.

La cuarta etapa hace referencia al algoritmo capaz de adquirir el conocimiento a partir de ejemplos; para ello es necesario realizar un proceso conocido como aprendizaje supervisado, el cual consiste en presentarle al algoritmo el conjunto de datos de entrada y salida organizados en la base de datos para que éste reconozca las características de la red ante cualquier tipo de falla. La metodología basada en el conocimiento propuesta en este trabajo se presentó en el capítulo 2, su aplicación al problema de la localización de la zona de fallas y al problema de la estimación de la distancia de falla se presentan en los numerales 3.6 y 3.7 respectivamente.

Finalmente la quinta etapa presenta el proceso de validación en el cual se busca comprobar la efectividad de la MBC. El error se estima por la comparación entre la respuesta del algoritmo ante una entrada específica y el valor de salida esperado.

Para la implementación práctica de este tipo de metodologías se requiere de un conocimiento amplio del sistema de distribución y una lectura en línea de las variables o información requerida del mismo, de tal forma que al ocurrir un evento de falla al interior de la red, la MBC, una vez entrenada, permita la localización apropiada del evento. De esta manera se puede proceder de una manera efectiva, con las labores de reparación y restauración del sistema.

3.3 MODELAMIENTO DEL SISTEMA DE DISTRIBUCIÓN

Los sistemas de distribución son el conjunto de elementos interconectados al sistema eléctrico cuyo objetivo es llevar el servicio al usuario final. En Colombia, las redes de distribución hacen parte de las redes de transmisión regional (STR) y los sistemas de distribución local (SDL), donde las primeras se encuentran conectadas al sistema de transmisión nacional (STN) y este último integra en una sola red los grandes centros de generación.

Los sistemas de distribución están conformados por un gran número de elementos eléctricos y mecánicos, de los cuales se destacan

- Líneas de distribución aérea y subterránea.
- Cargas que consumen potencia activa y reactiva.
- Compensadores de potencia reactiva.
- Reguladores de tensión.
- Protecciones.
- Transformadores

Para los sistemas de distribución radial se conoce como cabecera o subestación principal al punto en el cual el sistema de transmisión o subtransmisión cambia el nivel de tensión, para convertirse en un sistema de distribución.

De acuerdo con las características del problema de localización de fallas en sistemas de distribución, se propone como herramienta de simulación el software ATP (del inglés *Analysis Transients Program*), que es ampliamente utilizado a nivel mundial por los investigadores en el área de sistemas eléctricos de potencia y especialmente para el análisis de fallas y fenómenos transitorios. Una ventaja adicional es que el ATP es de libre uso para propósitos académicos y permite que el usuario pueda hacer aportes para su mejoramiento, lo que ha permitido su rápida evolución [CAUE (2002)]

Gracias a la interfaz gráfica del ATP conocida como ATPDraw [CAUE (2002)], la inclusión de los elementos del sistema de distribución para realizar la simulación del sistema es bastante rápida. Se pueden seleccionar los diferentes elementos de una lista muy completa, la cual incluye generadores, líneas de transmisión (existen diferentes modelos de la línea), elementos pasivos, conmutadores, entre otros. La implementación de las protecciones (las cuales no se encuentran dentro del listado mencionado anteriormente) se hace a través de los MODELS del ATP los cuales consisten en una herramienta que integra un lenguaje de código como un nuevo elemento eléctrico dentro del sistema que se está simulando en el ATPDraw. Los MODELS permiten crear nuevos elementos mediante una estructura de programación en lenguaje de código de propósito general, soportado por un conjunto de herramientas de simulación para representación y estudio de diferentes variables [CAUE (2002)].

A continuación se describen las generalidades de los elementos más comunes que poseen los sistemas de distribución, y para este trabajo en particular, los elementos por los cuales está conformado el sistema de distribución escogido para las pruebas de localización de fallas.

3.3.1 Modelamiento de líneas

Para el modelo de las líneas de distribución es conveniente utilizar un modelo eléctrico de línea corta o un modelo de línea media; utilizar un modelo de línea larga o de parámetros distribuidos puede convertir la simulación del sistema de distribución en un proceso de bastante exigencia computacional. Para el propósito de este trabajo no se ha incluido el efecto de las puestas a tierra para el sistema de distribución (en cables de guarda, estructuras, subestaciones, entre otras).

Dado que las líneas de los sistemas de distribución generalmente no llevan transposición y su disposición geométrica no es simétrica, los acoplos capacitivos e inductivos entre fases son de diferente magnitud. El modelamiento empleado en este trabajo hace uso de las matrices de impedancia y admitancia de línea media [Gonen 1988].

3.3.2 Modelo de la carga

En [Mora 2006] y [Gallego 2006] se presentan diferentes metodologías para el modelamiento de la carga; entre los modelos más utilizados para los sistemas de distribución se encuentran los modelos exponenciales y los polinomiales de la carga en función del voltaje; este último se muestra en la ecuación 3.1.

$$P = a_0 + a_1|V| + a_2|V|^2 \dots ; Q = b_0 + b_1|V| + b_2|V|^2 \dots \quad (3.1)$$

Para el modelo polinomial se referencia tres casos dependiendo de los valores de los coeficientes de los términos en la ecuación 3.1; por ejemplo, si sólo a_0 y b_0 son diferentes de cero se tiene un modelo de potencia constante; si sólo a_1 y b_1 son diferentes de cero se tiene un modelo de corriente constante; y si sólo a_2 y b_2 son diferentes de cero se tiene un modelo de impedancia constante. Existen otros modelos de carga basados en funciones racionales de la tensión, pero no son requeridos para el propósito de este trabajo.

En este trabajo se ha implementado el modelo de impedancia constante, pues la respuesta de las señales de tensión y corriente refleja un comportamiento bastante cercano a la respuesta del sistema real. En el problema de la localización de fallas, la implementación de un modelo de carga más elaborado se ve reflejado en un proceso de alto esfuerzo computacional, pues se hace necesario resolver un conjunto de ecuaciones no lineales para hallar las señales de tensión y corriente en el dominio del tiempo; es necesario la representación de las señales de tensión y corriente en función del tiempo ya que otro tipo de representación (por ejemplo fasorial) no permite obtener la suficiente información que caracteriza al sistema bajo estado de falla.

3.3.3 Compensadores de potencia reactiva

En los sistemas de distribución, los elementos compensadores de potencia reactiva consisten de capacitores conectados en diferentes subestaciones o nodos del sistema cuyo objeto es corregir factor de potencia mediante la inyección de potencia reactiva al sistema y mejorar los perfiles de tensión y cargabilidad de las líneas. El sistema de distribución de pruebas escogido para este trabajo incluye varios de estos elementos; su inclusión dentro de la simulación y su correcto modelamiento permite desarrollar una mejor metodología para la localización de fallas.

La ubicación de los capacitores dentro del sistema de distribución obedece a estudios de planeamiento donde se tienen presentes los costos de inversión, los costos operativos y adicionalmente estudios armónicos de la red para garantizar el buen funcionamiento de los mismos [Gonen 1986].

Dado que el alcance de este trabajo está orientado al problema de localización de fallas y no a la ubicación de capacitores, dichos elementos fueron ubicados en los nodos referenciados por el sistema de pruebas escogido.

3.3.4 Reguladores de tensión.

Los reguladores de tensión son elementos que generalmente consisten de autotransformadores de alta capacidad de corriente, cuya función es la de mejorar el perfil de tensión de la red. Al igual que los compensadores de potencia reactiva, su ubicación obedece a estudios de planeamiento de la distribución. El modelamiento de este tipo de elementos aún se encuentra en estudio, sin embargo se han logrado resultados que reflejan fielmente el comportamiento de ellos en el sistema, mediante modelos eléctricos similares al modelo del autotransformador [Kagan 1998].

Durante el desarrollo de este trabajo, en la etapa de obtención de los descriptores de la red bajo estado de falla, se encontró una gran similitud entre los resultados obtenidos cuando se incluyen los reguladores de tensión y cuando éstos no son incluidos. Por otra parte, asociado con la inclusión de este tipo de elementos se observa un mayor esfuerzo computacional para la herramienta de simulación escogida; así que para los propósitos de este trabajo se ha omitido la implementación de los reguladores del sistema de distribución.

3.3.5 Protecciones

Los elementos de protección están presentes en todos los niveles de tensión del sistema eléctrico de potencia, unos más especializados que otros dependiendo de la aplicación. Entre las protecciones de mayor implementación en sistemas de distribución, cuyo fin es la liberación de fallas por medio de la apertura de circuitos, son Reconectores, Seccionalizadores y Fusibles [Phadke 1995].

Para el modelamiento de estos elementos en el software de simulación EMTP/ATP se utilizan los MODELS ya que dentro del listado general de elementos de la librería del software no se encuentran de forma explícita las protecciones descritas anteriormente y con las características de operación requeridas. Es por esto que a través de elementos de programación disponible en los MODELS del EMTP/ATP se logra programar la operación de estas protecciones con las características de operaciones requeridas.

La simulación de las protecciones dentro del sistema de distribución será de gran utilidad para el problema de localización de fallas, pues la señal de corriente que resulta de la acción de las protecciones describe patrones característicos asociados con cada una de las zonas donde se encuentra localizada cada protección (ver 0); así se hace posible asociar cada una de las zonas principales de cada elemento de protección con la zonificación de la red. A continuación se describen los aspectos más relevantes asociados con las protecciones implementadas en este trabajo.

3.3.5.1 Fusibles

Los fusibles son elementos que no necesitan dispositivos adicionales para su correcto funcionamiento como transformadores de corriente (CT's); son elementos que ofrecen protección, de acuerdo a una banda de característica inversa tiempo contra corriente. Ante una falla en el sistema de distribución, estos elementos actúan solamente una vez por expulsión o fundición; el tiempo de disparo está dado por la característica del fusible, los parámetros con los que ha sido programado y la magnitud de la corriente de falla. [Lewis (1998)].

Para efectos de la simulación del comportamiento de este elemento en el sistema de distribución, la característica de operación inversa de tiempo contra corriente implementada es la misma que se utiliza para los relés de protección (ecuación 3.2) [Relaying 1999]. Esta aproximación permite modelar de manera apropiada el comportamiento de los fusibles.

$$t_D = D \cdot \left(B + \frac{A}{\left(I/I_{PK} \right)^P - 1} \right) \quad (3.2)$$

Donde:

- t_D : Tiempo de Disparo para la protección.
- D : Valor del Dial de la protección.
- I : Corriente medida en el lado de la protección
- I_{PK} : Corriente de activación (*Corriente de Pick Up*)

A, B y P son constantes que para este trabajo han sido fijadas en 0.0147, 0.183 y 0.02 respectivamente y corresponden con la curva de operación inversa tiempo-corriente de los fusibles empleados en sistemas de distribución.

La figura 3.2 presenta el esquema del funcionamiento y la lógica de programación utilizada para simular la operación del fusible. En la figura 3.3 se presenta una descripción del código de programación elaborado con MODELS del ATPDraw, para el modelamiento del fusible (ver también anexo 2).

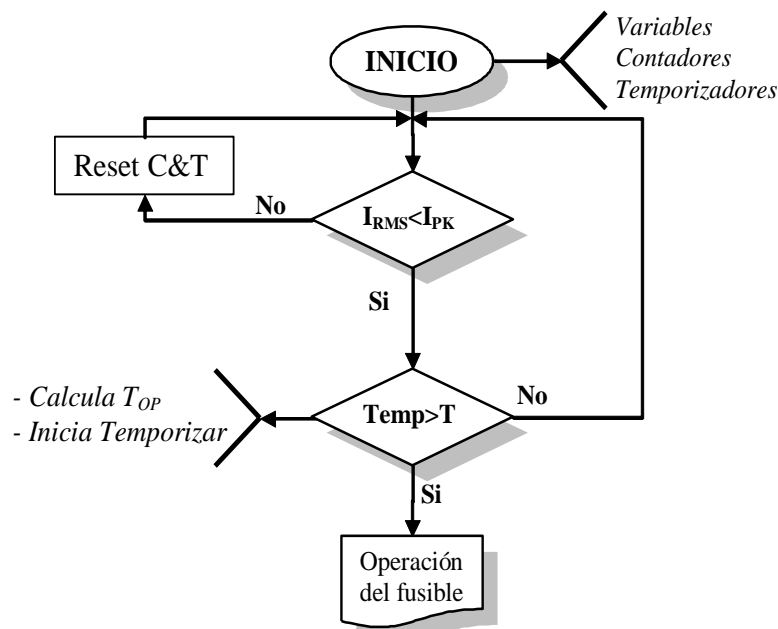


Figura 3.2. Esquema de funcionamiento del modelo implementado en ATP para la simulación del fusible.


```

MODEL FUSE

INPUT      (Señal de Entrada del MODEL)

DATA      (Datos de la curva del fusible y que pueden ser cambiados por el usuario)

VAR       (Variables para la medición del valor RMS de la Señal)
             (Variables de instantes de tiempo, temporización, y detección de cruce por cero)
             (Variables de Salida)

OUTPUT    (Variable para la apertura de la protección)

EXEC      (Inicio de la Ejecución)

IF T=0 THEN (Se inicializan las variables en el instante de la simulación T=0)
ENDIF

IF (T>0) THEN (Lectura de la Corriente RMS)

IF IRMS>=IPK THEN      (Detección de sobrecorriente, cálculo de tiempo de
                          operación del fusible según DIAL y parámetros de
                          la curva A,B,P...
                          Inicializar Temporización)

ENDIF

IF TEMP> Top THEN (Ha terminado la temporización, el fusible debe actuar en el
                     sigüiente cruce por cero de la señal)

ENDIF

ENDIF

ENDEXEC    (Termina la ejecución)

```

Figura 3.3. Descripción del MODEL implementado en ATP para la simulación del fusible.

La figura 3.4 muestra la característica de la señal de corriente luego de ocurrida una falla y su respectiva liberación debido a la acción del fusible.

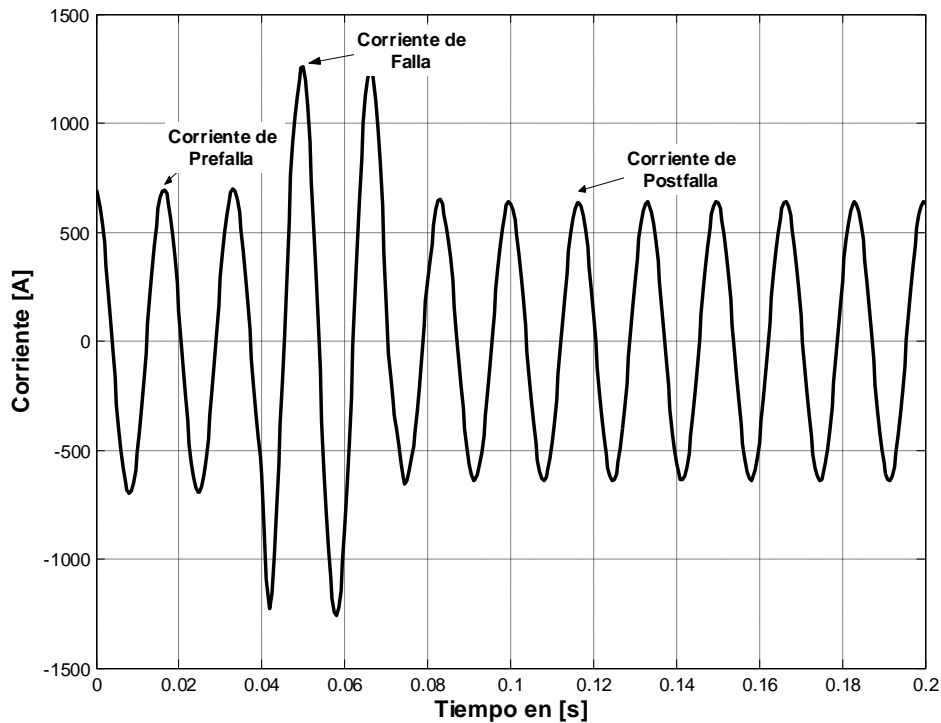


Figura 3.4. Patrón de la corriente de falla liberada por un fusible

Cuando ocurre la falla en el sistema, la corriente que circula en el lado de la protección experimenta un incremento; si el valor de esta corriente es superior a la corriente de activación del fusible, se inicia el proceso de operación de la protección, el cual abrirá el circuito luego de transcurrido el tiempo dado por los parámetros del fusible y la corriente de falla evaluados en la ecuación 3.2.

3.3.5.2 Reconectores

Los reconectores cuentan con funciones de protección (50-51), interrupción (52), integradas con control de recierre del circuito. Estos elementos establecen una secuencia de recierre típica de las estructuras de automatización de la distribución. Los reconectores tienen comúnmente dos curvas, una rápida y una lenta. Ante una falla en el sistema, se mide la corriente de falla y esta protección actúa un número determinado de veces con la curva rápida y un tiempo de recierre t_R . Si la falla persiste, el reconnector actúa un número determinados de veces con la curva lenta, recerrando el circuito con un tiempo t_L . Estos elementos son normalmente utilizados en los sistemas de distribución para reducir el efecto de las fallas transitorias en el índice equivalente de interrupciones DES. [Lewis (1998)]

Debido al desbalance de cargas en las fases de los sistemas de distribución, los reconectores experimentan corrientes distintas en cada una de ellas; en la práctica debido al desbalance, el proceso de coordinación de protecciones suele emplear una corriente de activación diferente para cada una de las tres fases.

El código implementado dentro del MODEL para la simulación de la operación real del reconectador, es muy similar al código del MODEL implementado para fusible (figura 3.5). El reconectador mide permanentemente el valor RMS de la corriente en cada una de las fases. En caso de existir sobrecorriente en el sistema, el reconectador inicia su proceso de operación, que consiste de operaciones rápidas y lentas. Cuando hay fallas monofásicas, el reconectador tiene la posibilidad de hacer disparo monofásico. Para fallas bifásicas o trifásicas, el disparo se realiza en las tres fases. Si durante la operación de apertura y recierre del reconectador, la corriente del sistema vuelve a la normalidad, este inhibirá el conteo y ejecución de disparos.

La figura 3.5 muestra el funcionamiento y la lógica de programación del MODEL utilizado para simular la operación del reconectador.

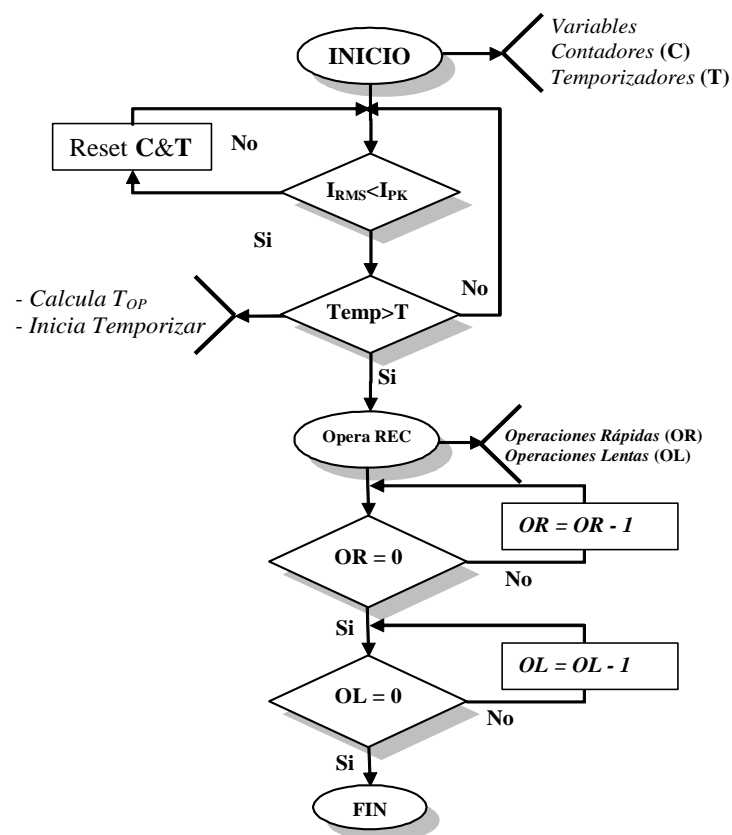


Figura 3.5. Esquema de funcionamiento del modelo implementado en ATP para la simulación del reconectador.

La figura 3.6 muestra la característica de la señal de corriente luego de ocurrida una falla y su respectiva liberación debido a la acción del reconectador.

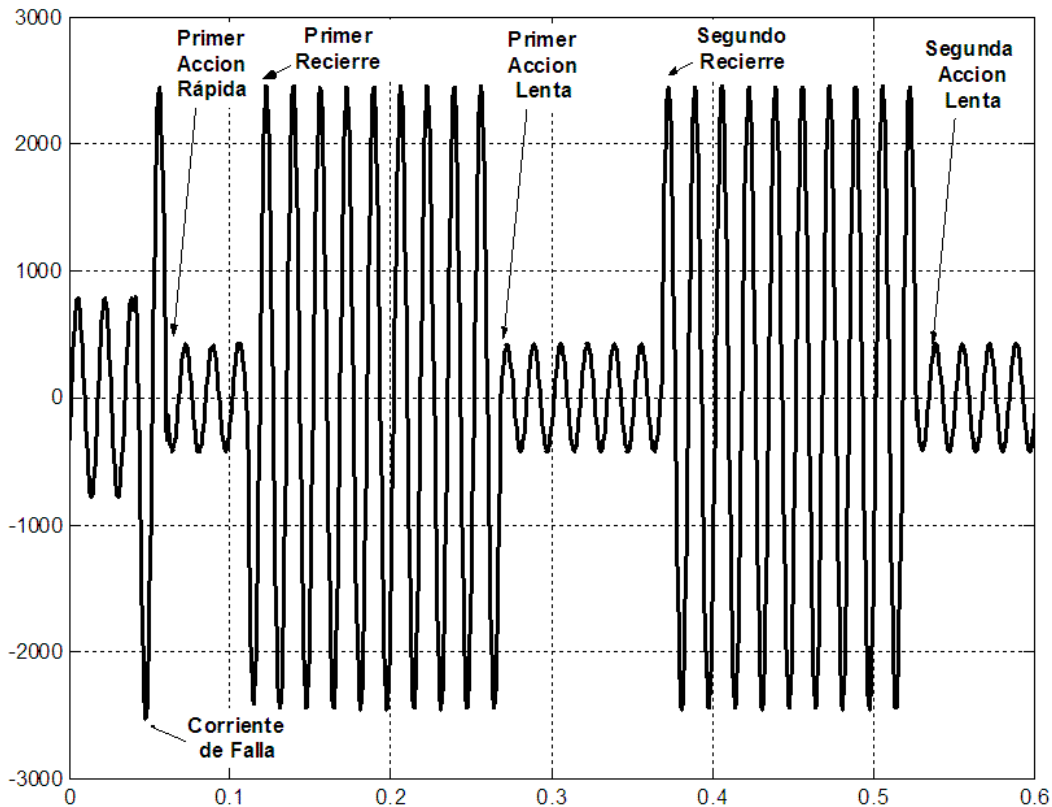


Figura 3.6. Patrón de la corriente de falla liberada por un reconectador

3.3.5.3 Seccionalizadores

Los seccionalizadores son dispositivos para aislar o efectuar corte del flujo de corriente en el sistema de distribución; vienen acompañados por elementos de medición que le permiten contar el número de veces que la corriente a través de él ha sido nula durante intervalos de tiempo extendidos a causa de los elementos de protección ubicados aguas arriba de su ubicación en el sistema de distribución; luego de alcanzado un nivel máximo en el conteo se activa la función de interrupción que abre el circuito. A diferencia de los fusibles y los reconectadores, quienes despejan la falla en presencia de corriente, los seccionalizadores requieren de su ausencia para su correcto disparo; es decir, estos elementos no actúan bajo carga; por esta razón es necesario que los seccionalizadores estén dispuestos aguas abajo de los reconectadores del sistema de distribución para que en la liberación de la falla por parte del reconectador, el seccionalizador pueda despejar el circuito fallado aguas abajo de él [Lewis 1998]

La figura 3.7 muestra el funcionamiento y la lógica de programación del MODEL utilizado para simular la operación del seccionalizador.

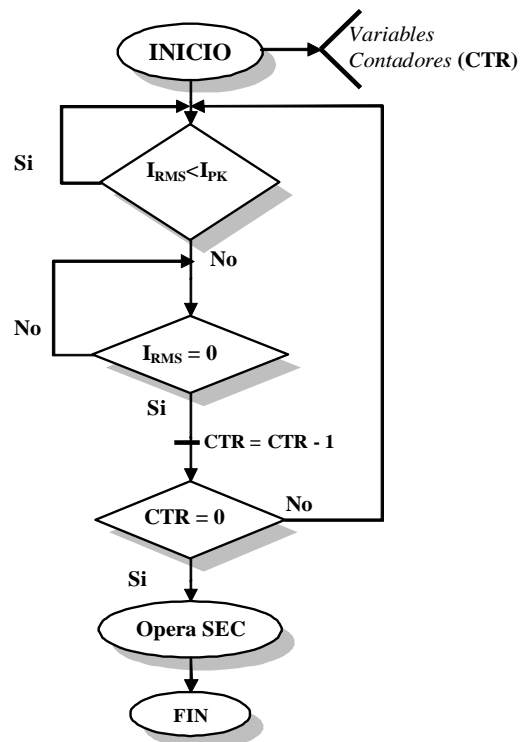


Figura 3.7. Esquema de funcionamiento del modelo implementado en ATP para la simulación del seccionizador

La figura 3.8 muestra la característica de la señal de corriente luego de ocurrida una falla y su respectiva liberación debido a la acción del seccionizador. Para la correcta acción de estos elementos es necesario que aguas arriba se encuentre ubicado un reconectador. Específicamente, la figura 3.8 muestra la señal de corriente para una falla ocurrida aguas abajo de un seccionizador que a su vez está precedido por un reconectador; en este caso el reconectador está programado con una acción rápida y dos acciones lentas; el seccionizador está programado con un nivel de conteo máximo de dos.

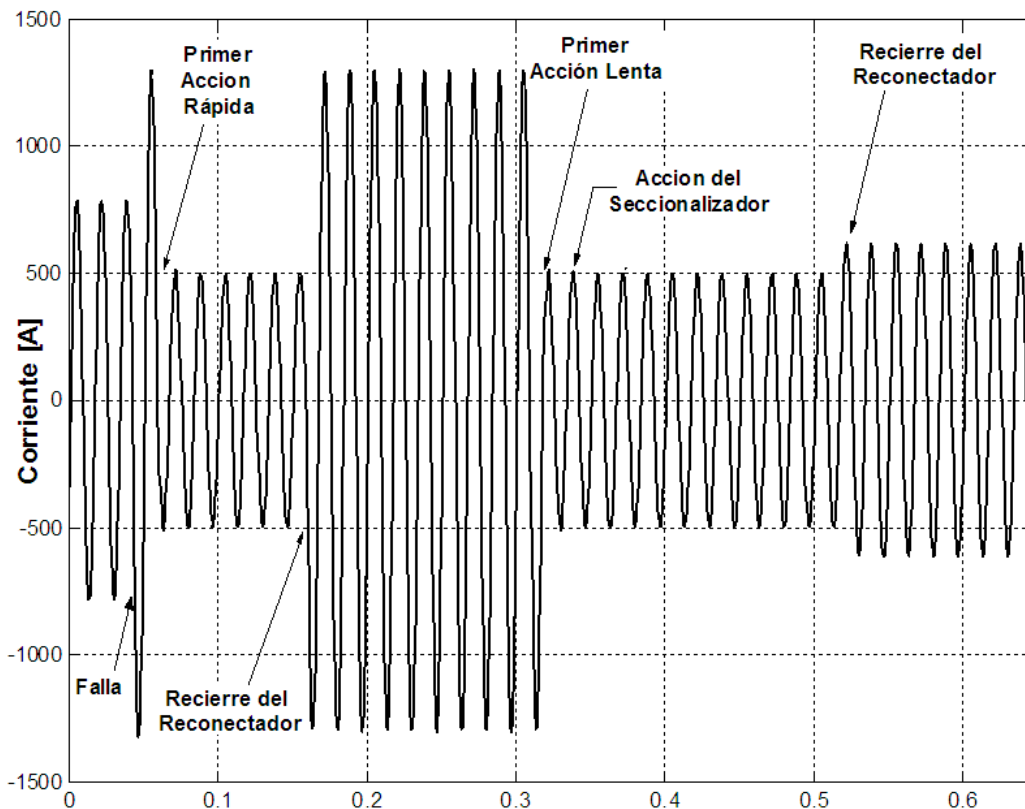


Figura 3.8. Patrón de la corriente de falla liberada por un seccionizador

3.3.5.4 Esquema de Salvamento de Fusibles

El esquema de salvamento de fusibles corresponde con la disposición de los elementos de protección (reconectores y fusibles) de forma que se optimice el desempeño de ellas dentro del sistema de distribución [Seung *et al* 2004].

El esquema de salvamento de fusibles consiste en la conexión de fusibles aguas debajo de los reconectores; el tiempo de acción rápido del reconector debe ser menor que el tiempo de apertura con el fin de que la sección fallada sea despejada inicialmente por el reconector, posteriormente se reconecta la sección para examinar si la falla ya ha sido despejada. La estructura de salvamento de fusibles evita la acción de los fusibles del sistema ante fallas transitorias; es decir, las fallas de carácter transitorio son despejadas por los reconectores y no por los fusibles ya que estos últimos quedan inutilizables luego de su operación. Una configuración básica para el esquema de salvamento de fusibles se muestra en la figura 3.9.

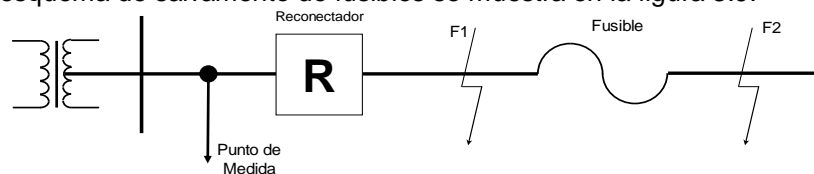


Figura 3.9. Esquema básico de salvamento de fusibles básico.

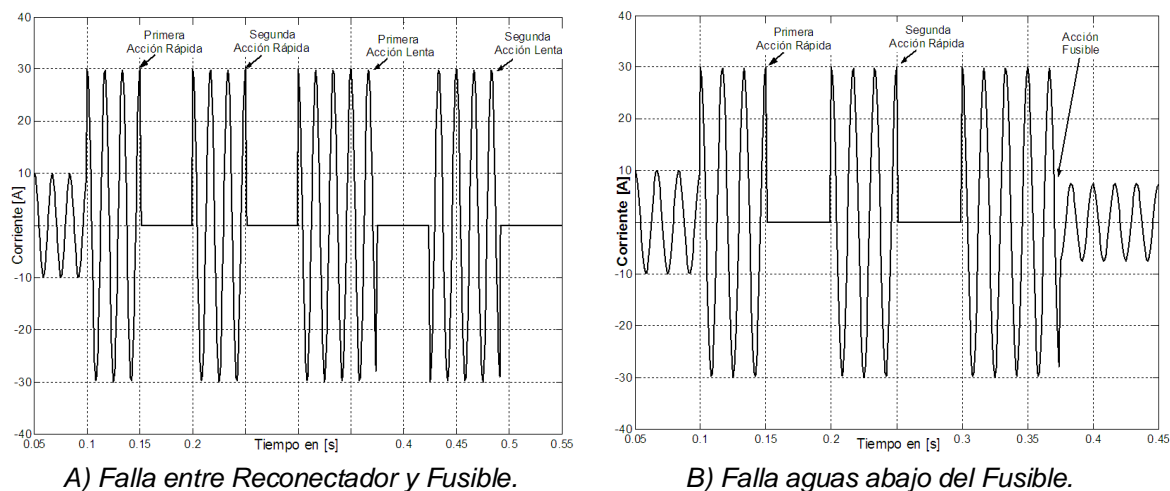


Figura 3.10. Patrón de Corriente para una falla ocurrida en un punto específico del sistema

Para detallar el funcionamiento del esquema de salvamento de fusibles, el reconector de la figura 3.10 ha sido programado con dos acciones rápidas y dos acciones lentas. Para una falla permanente del tipo F1 (como se muestra en la figura 3.10a) la corriente en el punto de medida es como se muestra en la figura 3.10a (para cada apertura del reconector la corriente se hace súbitamente nula), mientras que para una falla permanente del tipo F2, la corriente en el punto de medida sería como se muestra en la figura 3.10b (para la apertura final del fusible se observa la corriente asociada con la carga que no ha sido desconectada).

Con este tipo de coordinación se logra que ante una falla de carácter transitorio, aguas abajo del fusible, el primer disparo lo realiza el reconector y se evita el disparo del fusible (el cual es un disparo permanente mientras que el de los relectores no necesariamente lo es), pero si la falla es de carácter permanente, necesariamente será el fusible quien finalmente despeje la falla.

Para el problema de localización de fallas, el esquema básico de salvamento de fusibles es de gran utilidad ya que el número de conmutaciones presentes en la señal de corriente de una fase fallada está asociado con el tipo, número y características de las protecciones que se encuentran a lo largo del ramal y los laterales asociados con el evento; por ejemplo, para una falla que ocurre en un punto del sistema de distribución donde la única protección, entre la subestación principal y el punto de falla, es un fusible, se presentará un solo incremento en la corriente el cual será liberado posteriormente por dicho fusible; pero si la falla ocurre en un punto del sistema de distribución donde las protecciones, entre la subestación principal y el punto de falla, son un reconector y un fusible, se presentarán $n+1$ conmutaciones donde n corresponde al número de acciones rápidas con el cual ha sido programado el reconector y el incremento adicional ($+1$) corresponde a la acción final del fusible dentro del esquema de salvamento de fusibles.

3.3.6 Transformadores

Los transformadores hacen parte de los elementos vitales dentro de un sistema de potencia. Su objetivo es el modificar el nivel de tensión de un punto de la red al otro, transfiriendo los niveles de potencia adecuados para el correcto funcionamiento de la misma. En los sistemas de distribución, estos transformadores, se encargan de reducir el nivel de tensión a niveles aptos y seguros para el usuario final.

El modelamiento de los transformadores en los sistemas de distribución se suele hacer mediante impedancias que representan los efectos resistivos e inductivos de los devanados y núcleo

respectivamente; esta impedancia puede referirse a cualquiera de los lados del transformador facilitando el trabajo en uno u otro nivel de tensión. Los transformadores de distribución vienen dotados de un sistema de TAP para modificar en pequeños porcentajes (alrededor de la tensión nominal) el nivel de tensión en la salida. [Chapman 1987] [Rodas 2004]

3.4 SIMULACIÓN AUTOMÁTICA DE FALLAS

En esta sección se describe la metodología propuesta para la simulación automática de todos los tipos de fallas para diferentes valores típicos de la resistencia de falla para cada uno de los nodos del sistema de pruebas.

Para la obtención de una base de datos completa con registros de las señales de tensión y corriente en el dominio del tiempo, se simulan diferentes tipos de fallas en el sistema de distribución para pruebas. Esto es necesario ya que el conocimiento empírico necesita de una gran cantidad de ejemplos para el aprendizaje. La generación automática de fallas en ATP consiste en generar y simular automáticamente una gran cantidad de copias del sistema de distribución, sometido cada vez a una falla diferente.

En total son once el tipo de fallas que se simulan; a continuación se muestra cada una de ellas

Fallas monofásicas a tierra

- Falla fase A tierra - *Falla01*.
- Falla fase B tierra - *Falla02*.
- Falla fase C tierra - *Falla03*.

Fallas Bifásicas a tierra

- Falla entre fases A y B y tierra - *Falla07*.
- Falla entre fases B y C y tierra - *Falla08*.
- Falla entre fases A y C y tierra - *Falla09*.

Fallas Bifásicas

- Falla entre fases A y B - *Falla04*.
- Falla entre fases B y C - *Falla05*.
- Falla entre fases A y C - *Falla06*.

Fallas Trifásicas

- Falla entre fases A, B y C - *Falla10*.
- Falla entre fases A, B y C y tierra - *Falla11*.

Para tener una base de datos completa, estas fallas se deben simular en cada uno de los nodos del sistema; se debe tener presente que en cada nodo perteneciente a un ramal monofásico sólo se puede simular un tipo de falla, en los nodos pertenecientes a ramales bifásicos se pueden simular cuatro tipos de falla y en los nodos pertenecientes a ramales trifásicos se pueden simular hasta once tipos de fallas; por último, cada uno de estos eventos, se debe simular con diferentes valores de las resistencias de fallas [Mora et al (2005a)].

La figura 3.11 muestra las etapas o bloques operativos para la simulación automática de fallas. Como se explicó anteriormente, la herramienta de simulación escogida es el software ATP (Analysis Transients Program). Adicionalmente para la simulación automática de todos los eventos se empleó el software **Matlab®** el cual es de amplio uso entre los investigadores de las diferentes disciplinas, con la ventaja de que permite interactuar con un amplio conjunto de aplicativos distintos a los ya implementados por *The Mathworks Inc* (representantes de **Matlab®**)

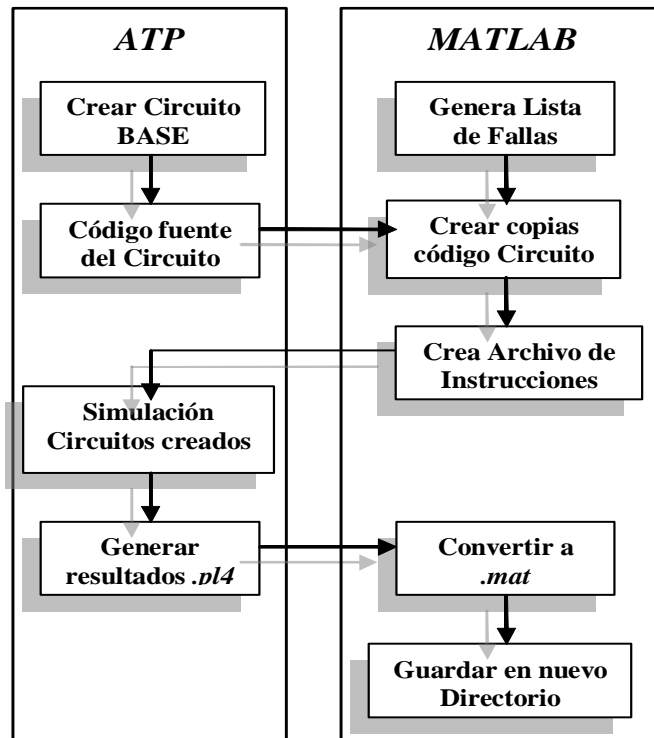


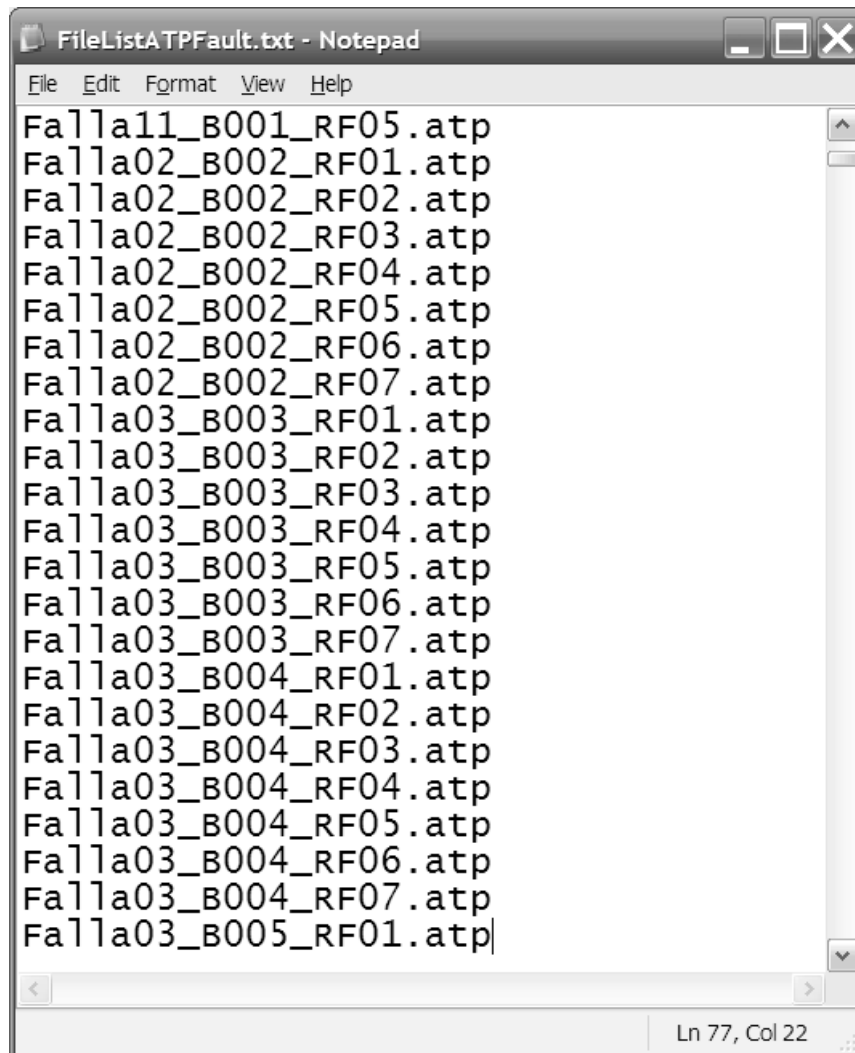
Figura 3.11. Esquema de interacción implementada entre ATP y MATLAB® para la generación automática de fallas.

3.4.1 Generar lista de fallas

Para la generación del listado de fallas se creó un aplicativo con la ayuda del software **Matlab®**. Esta aplicación se encarga de generar un solo archivo de texto (con extensión **.txt**), con un listado de fallas. La forma de generar este listado, es mediante la generación y concatenación de texto y números convertidos a caracteres.

El nombre de cada archivo ATP contenido en el archivo de texto es “Falla**XX**Bus**YYY**R**FZZ**.atp”, donde **XX** corresponde al tipo de falla. (De la **01** a la **11**), **YY** es el nodo donde se simula la falla (Del **01** a **123** nodos del sistema de prueba), y **ZZ** es el valor de la resistencia de falla (Numerados de **01** a **06**).

La figura 3.12 muestra el archivo de texto, generado mediante **Matlab®**, el cual contiene el listado de fallas que se deben simular. Se aprecian los diferentes nodos en los que se simularán las fallas (en la figura 3.12 se muestran los nodos 001, 002, 003, 004 y 005, de los 123 nodos que posee el sistema de pruebas escogido), los tipos de fallas que se simularán (en la figura 3.12 se muestran las fallas tipo 11, 02 y 03 de los 11 tipos de falla descritos anteriormente) y las resistencias de falla (en la figura 3.12 se muestran las resistencias de falla 01, 02, 03, 04, 05, 06 y 07 las cuales corresponden con valores específicos de resistencia que pueden ir desde 1 Ω hasta 40 Ω).



```

FileListATPFault.txt - Notepad
File Edit Format View Help
Falla11_B001_RF05.atp
Falla02_B002_RF01.atp
Falla02_B002_RF02.atp
Falla02_B002_RF03.atp
Falla02_B002_RF04.atp
Falla02_B002_RF05.atp
Falla02_B002_RF06.atp
Falla02_B002_RF07.atp
Falla03_B003_RF01.atp
Falla03_B003_RF02.atp
Falla03_B003_RF03.atp
Falla03_B003_RF04.atp
Falla03_B003_RF05.atp
Falla03_B003_RF06.atp
Falla03_B003_RF07.atp
Falla03_B004_RF01.atp
Falla03_B004_RF02.atp
Falla03_B004_RF03.atp
Falla03_B004_RF04.atp
Falla03_B004_RF05.atp
Falla03_B004_RF06.atp
Falla03_B004_RF07.atp
Falla03_B005_RF01.atp
Ln 77, Col 22

```

Figura 3.12. Archivo de texto, generado mediante MATLAB®, con el listado de fallas a simular.

3.4.2 Generar los diferentes Sistemas en el ATP

En esta etapa se genera un gran número de copias de un mismo sistema de distribución base donde cada copia se encuentra asociada con una falla diferente. Este proceso se realiza a partir del archivo base, el cual contiene el código fuente del sistema de potencia dibujado en el ATPDraw complementado con las protecciones desarrolladas con los MODELS. Basta con modificar algunas líneas del código del archivo para cambiar el valor de la resistencia de falla y la ubicación del nodo de falla.

El archivo que contiene el sistema de pruebas se genera en un ambiente gráfico del ATP conocido como el ATPDraw, el cual es una interfaz que facilita el trabajo entre el usuario y la máquina. Al compilar y ejecutar el archivo creado en el ATPDraw, se genera automáticamente el archivo ATP, el cual (al abrirlo como un archivo de texto) contiene el código fuente del sistema eléctrico creado gráficamente. Para generar una falla en el sistema de distribución basta con conectar la resistencia de falla (que generalmente varía entre 1 Ω y 40 Ω) a tierra; también se debe conectar un interruptor

en serie con la resistencia de falla para controlar el instante de entrada del evento de falla (ver figura 3.13).

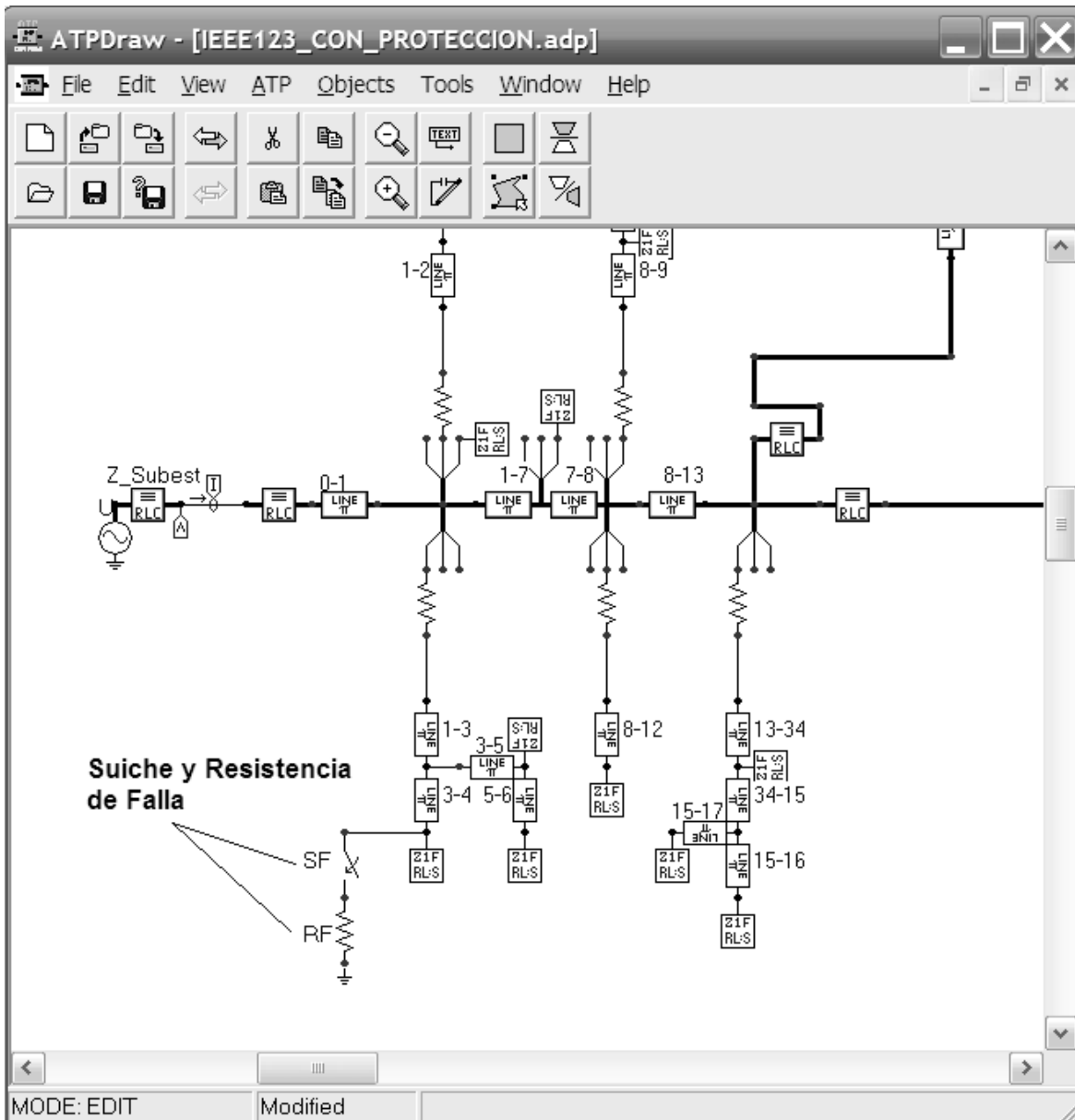
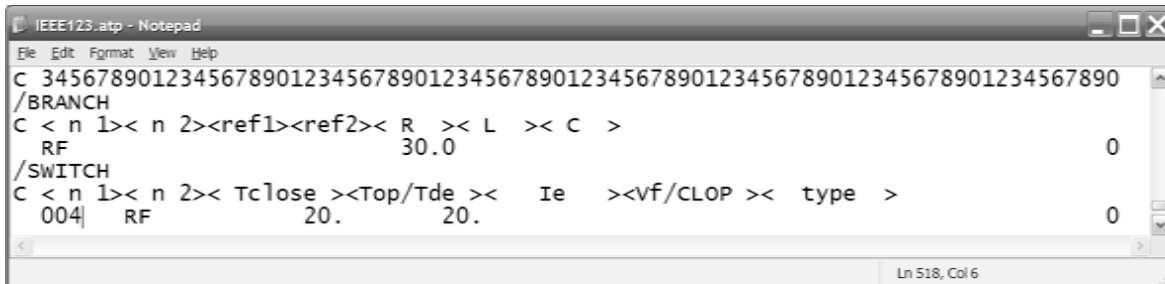


Figura 3.13. Sistema de distribución de pruebas creado mediante el ATPDraw. Se destacan los elementos que permiten generar el evento de falla: Resistencia de falla e interruptor de control de entrada de la misma.

En la figura 3.14 se enfocan las líneas de texto del archivo base que contienen la información de la resistencia de falla y el interruptor que controla la entrada del evento en el sistema de distribución. Las líneas que comienzan por la letra “C” son comentarios las cuales no tienen ningún efecto en el circuito y cuyo propósito es facilitar la visualización del código; por ejemplo, la primera línea muestra las columnas (de la 1 a la 80) del archivo de texto (el software ATP siempre maneja un ancho de 80 columnas); la tercera línea indica el nodo inicial y el nodo final entre los cuales están conectados los elementos (“BRANCH”) y sus respectivos valores: Resistivo (R), inductivo (L) o

capacitivo (C); para este caso se muestra una resistencia de 30.0Ω que va conectada entre el nodo "RF" y el nodo de tierra (el espacio en blanco es procesado como nodo de tierra); igualmente la quinta línea indica el nodo inicial y el nodo final entre los cuales están conectados los interruptores ("SWITCH") del sistema, en este caso hay un interruptor con un tiempo para cierre de 20 segundos, que se encuentra conectado entre el nodo "004" (que es el nodo en el cual se desea simular la falla) y el nodo "RF".



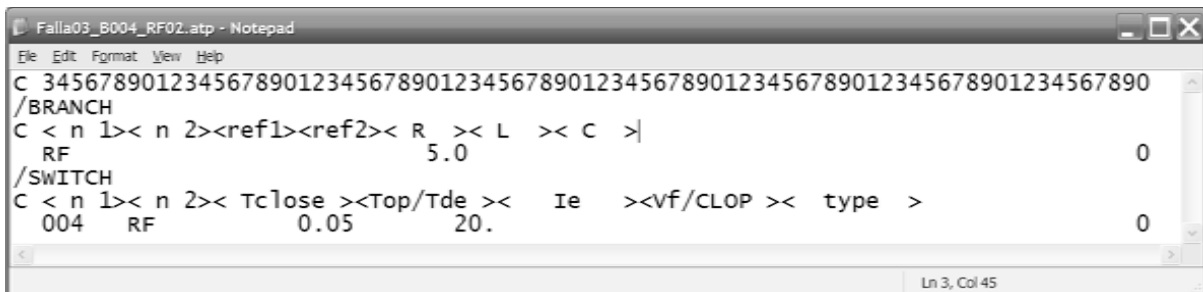
```

IEEE123.atp - Notepad
File Edit Format View Help
C 34567890123456789012345678901234567890123456789012345678901234567890
/BRANCH
C < n 1>< n 2><ref1><ref2>< R >< L >< C >
  RF          30.0          0
/SWITCH
C < n 1>< n 2>< Tclose ><Top/Tde >< Ie ><vf/CLOP >< type >
  004 RF          20.          20.          0
Ln 518, Col 6

```

Figura 3.14. Código fuente del archivo ATP original. Se destacan las líneas de código que contienen la resistencia de falla y el interruptor que controla la entrada de la misma.

Mediante el software **Matlab®**, y el listado de fallas generado en la etapa anterior, se crea una copia del archivo original (mostrado figura 3.14) y se modifican las líneas de texto que contienen la información de la resistencia de falla y el interruptor que controla la entrada del evento de falla. La figura 3.15 muestra el nuevo archivo generado con los valores deseados para la resistencia de falla (para este caso RF02 en el título del archivo corresponde con una resistencia de 5Ω) y el tiempo de entrada del evento de falla (para este caso, el interruptor crea una falla que inicia a los 0.05 s y termina a los 20 s; el tiempo total de la simulación es de dos segundos así que 20 s para la apertura del interruptor significa que la falla es de carácter permanente.)



```

Falla03_B004_RF02.atp - Notepad
File Edit Format View Help
C 34567890123456789012345678901234567890123456789012345678901234567890
/BRANCH
C < n 1>< n 2><ref1><ref2>< R >< L >< C >
  RF          5.0          0
/SWITCH
C < n 1>< n 2>< Tclose ><Top/Tde >< Ie ><vf/CLOP >< type >
  004 RF          0.05         20.          0
Ln 3, Col 45

```

Figura 3.15. Archivo ATP que contiene codificado el sistema de distribución escogido para pruebas. Se destacan las líneas de código que contienen las resistencias de Falla y los suiches para la entrada de las fallas.

3.4.3 Generar el archivo de ejecución automática de los archivos ATP de fallas

Para la ejecución automática de los circuitos generados anteriormente, se utiliza un archivo de instrucciones de procesamiento por lotes (**batch**) de extensión **.bat** que ejecuta todas las simulaciones. Las instrucciones que contiene este archivo son:

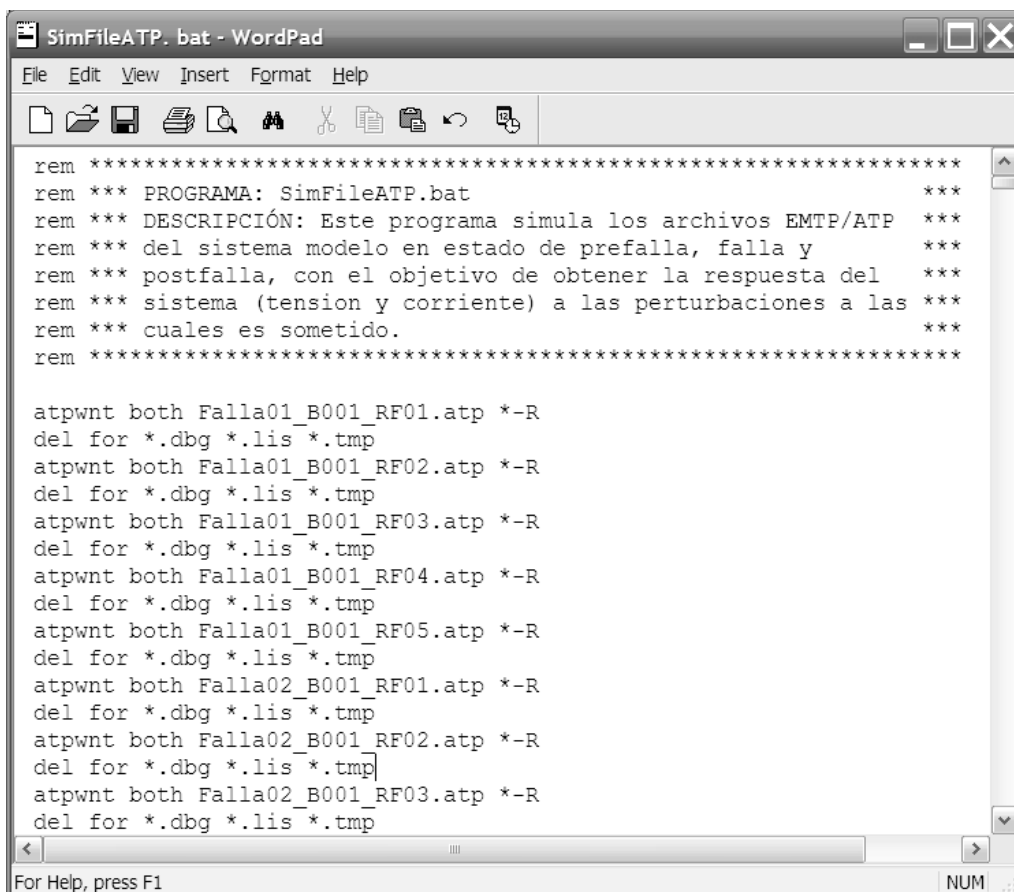
```

atpwnt archivo.atp *-R
del for *.tmp *.dbg *.lis

```

La primera instrucción se encarga de ejecutar los archivos **.atp** generados en la etapa anterior. En la simulación de un archivo **.atp** se generan otros archivos de extensión **.lis**, **.dbg**, **.tmp** y **.pl4** de los cuales sólo interesa este último, de modo que la segunda instrucción se encargará de borrar los demás archivos. Los archivos **.pl4** corresponden a los resultados de la simulación de los archivos **.atp** allí se encuentran almacenadas los valores de las señales de tensión y de corriente en el dominio del tiempo.

Para este caso y acorde con los datos que se disponen en los sistemas de distribución reales, la base de datos de registros de fallas queda conformada por las señales de tensión y corriente en el dominio del tiempo y sus valores RMS, medidos en el nodo de la subestación. La figura 3.16 muestra el código que debe ejecutar el archivo de procesamiento por lotes (archivo batch) generado en esta etapa.



```

rem *****
rem *** PROGRAMA: SimFileATP.bat ***
rem *** DESCRIPCIÓN: Este programa simula los archivos EMTP/ATP ***
rem *** del sistema modelo en estado de prefalla, falla y ***
rem *** postfalla, con el objetivo de obtener la respuesta del ***
rem *** sistema (tension y corriente) a las perturbaciones a las ***
rem *** cuales es sometido. ***
rem *****

atpwnt both Falla01_B001_RF01.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla01_B001_RF02.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla01_B001_RF03.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla01_B001_RF04.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla01_B001_RF05.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla02_B001_RF01.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla02_B001_RF02.atp *-R
del for *.dbg *.lis *.tmp
atpwnt both Falla02_B001_RF03.atp *-R
del for *.dbg *.lis *.tmp

```

Figura 3.16. Archivo Batch que posee codificadas las instrucciones para la simulación automática de todos los circuitos que se generaron en la etapa anterior.

3.4.4 Convertir los resultados de los archivos **.pl4** a archivos **.mat**.

Los archivos de extensión **.mat** son archivos de datos en código ASCII, que pueden ser manipulados por el **Matlab®**. En este archivo se guardan las señales de tensión y corriente en el dominio del tiempo y sus valores RMS. Teniendo en cuenta la potencia de manipulación de datos de Matlab®, se creó un archivo que se encarga de convertir los archivos **.pl4** en archivos **.mat**.

Una vez se tienen los resultados como archivos de extensión *.mat* se puede realizar con el software *Matlab*® los respectivos análisis y procesar las señales de tensión y corriente obtenidas del ATP para extraer la información que se describe en 3.5.

3.5 INFORMACIÓN REQUERIDA PARA LA CARACTERIZACIÓN DE LOS EVENTOS DE FALLA.

Los métodos de localización de fallas basados en el conocimiento (MBC) normalmente formulan la solución de un problema mediante un *algoritmo* de clasificación que usa un conjunto de datos de Entrada-Salida cuyo número es suficientemente significativo. El objetivo es conseguir que el MBC *aprenda* automáticamente las propiedades deseadas a partir de estos ejemplos, mediante una modificación apropiada de sus parámetros. El proceso en el cual se adecuan dichos parámetros a la resolución de cada problema, se denomina *aprendizaje ó entrenamiento*.

La mayor parte de la información que describe al sistema de distribución bajo estado de falla, pertenece básicamente a los registros en el dominio del tiempo de las señales de tensión y corriente medidas en el nodo de la subestación principal. También existe información adicional que es útil para resolver el problema de localización de fallas, tales como la localización de los elementos de protección del sistema, la tipificación de la carga (tipo de carga que existe en cada nodo, industrial comercial y residencial, y para este último la estratificación social), el registro de llamadas telefónicas hechas por los clientes informando sobre fallas en la red, y las condiciones climáticas en función del tiempo. [Mora et al 2005b]

Mediante las señales de tensión y de corriente medidas en la subestación del sistema durante el estado de prefalla, falla y postfalla se puede caracterizar el sistema de distribución; a este tipo de información se le conoce como descriptor del sistema. Posteriormente se obtiene un valor de cada descriptor para cada estado diferente de falla en el sistema de distribución [Mora et al 2005c]. A continuación se presentan los descriptores empleados en este trabajo y se muestran ejemplos para su extracción.

3.5.1 Caracterización de la profundidad de los huecos de tensión

La profundidad del hueco de tensión se define como la variación del valor eficaz de la señal de tensión entre los estados estables de prefalla y falla. La señal de tensión registrada en la subestación tiene una magnitud variable durante los intervalos de prefalla, falla y postfalla; la profundidad del hueco de tensión que se presenta entre los dos primeros intervalos permite extraer características que se pueden asociar con las fallas tal como se presenta en [Mora et al 2003]. La magnitud del hueco de tensión en estado de falla depende de la impedancia vista desde la subestación hasta el sitio de localización del evento y de la propia impedancia de falla; de esta forma se pueden asociar la magnitud del hueco de tensión a la localización del evento.

La figura 3.17 muestra la señal de tensión en cada fase para una falla ocurrida en el nodo 01, del sistema de distribución escogido para pruebas, con una resistencia de falla a tierra de 2Ω . La variación de la tensión en las fases A, B y C debido a la falla son 192, 196 y 194 V respectivamente.

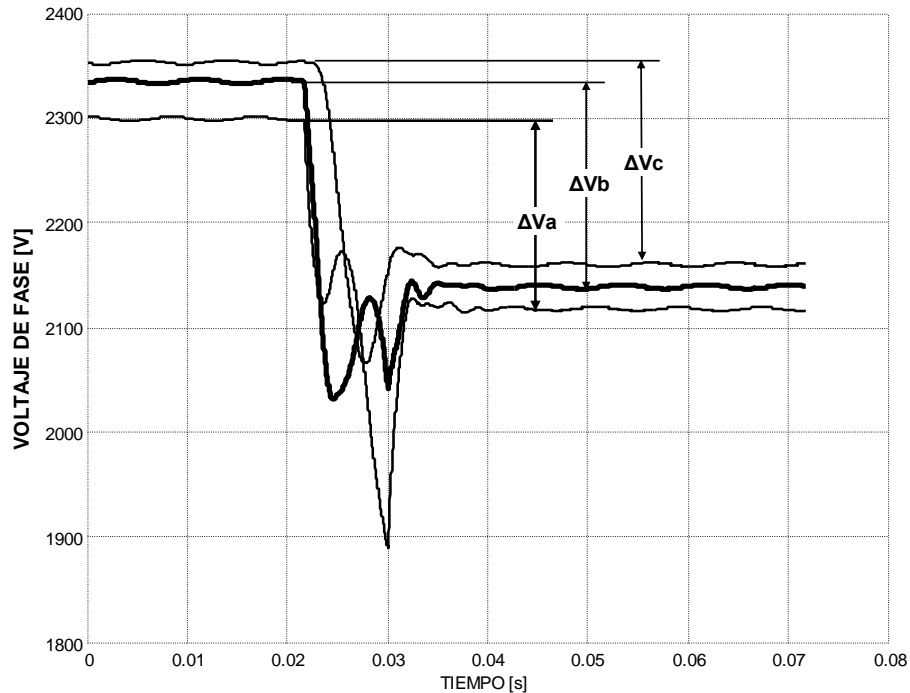


Figura 3.17. Voltaje RMS de cada fase para falla trifásica en el nodo 1 del sistema de Pruebas.

3.5.2 Variación de la magnitud de la corriente

Al igual que el descriptor presentado en 0, se puede cuantificar la magnitud del cambio de corriente entre los estado de prefalla y de postfalla. La variación de la magnitud de la corriente se define como el cambio del valor RMS de la señal de corriente en estado de prefalla y falla, para cada fase.

La variación de corriente es una variable que a pesar de estar asociada a los parámetros de los elementos de la red de distribución depende fuertemente del valor de la impedancia de falla; para valores muy altos de la impedancia de falla, la variación de la corriente es pequeña y para valores pequeños de la impedancia de falla la variación de corriente es alta; una consecuencia adicional al efecto de la impedancia de falla es que existirán fallas cercanas a la cabecera de la subestación con una impedancia de falla grande que producirá una variación de corriente similar al valor que generan fallas en nodos muy alejados de la cabecera del sistema y con impedancia de falla pequeña.

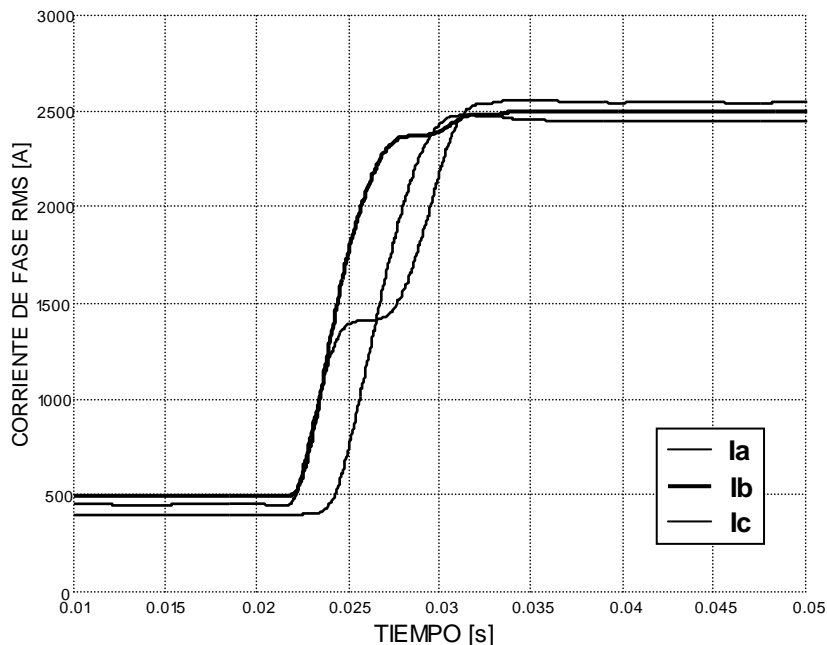


Figura 3.18. Corriente RMS de cada fase para falla trifásica en el nodo 1.

3.5.3 Análisis de la señal transitoria

Las componentes de frecuencia y duración del transitorio, se pueden relacionar con la localización de la falla tal como se muestra en [Martínez (2002)]. De la teoría de circuitos eléctricos se conoce que la frecuencia de oscilación asociada a una señal transitoria en un sistema eléctrico está asociada a los parámetros capacitivos e inductivos del mismo; para cada evento y para cada sitio de la localización de la falla estos parámetros tendrán valores distintos dado que los elementos del sistema de distribución involucrados desde la subestación al sitio de ocurrencia de la falla son diferentes para cada evento posible; así, la frecuencia de oscilación de la señal transitoria al igual que la duración de la misma puede asociarse de forma directa con la localización de la falla.

La transformada de Fourier ha sido ampliamente utilizada como herramienta para el análisis de frecuencia de señales eléctricas; sin embargo, para el problema de localización de fallas la implementación de esta herramienta por sí sola no es adecuada, ya que no tiene en cuenta el intervalo de tiempo dentro del cual ocurren los eventos de falla.

La transformada Wavelet es una herramienta muy apropiada para el análisis de transitorios ya que tiene presente tanto la frecuencia de oscilación como el intervalo de tiempo asociado a la señal transitoria originada por la falla. Su funcionamiento está basado en la descomposición de la señal en bandas de frecuencia (filtración sucesiva de la señal) y operaciones de modulación, traslación y dilatación en el tiempo.

La metodología propuesta consiste del uso de las dos herramientas para la extracción de la frecuencia y la duración del transitorio en la señal. El análisis Wavelet permite realizar un filtrado múltiple de la señal en un conjunto de anchos de banda determinado y luego se obtiene la duración

del estado transitorio; posteriormente, mediante un análisis de Fourier de las señales filtradas es posible obtener la frecuencia de oscilación de la señal transitoria.

La metodología para la extracción de los descriptores de frecuencia y duración de la señal transitoria se muestra en la figura 3.19 la cual comprende las siguientes actividades

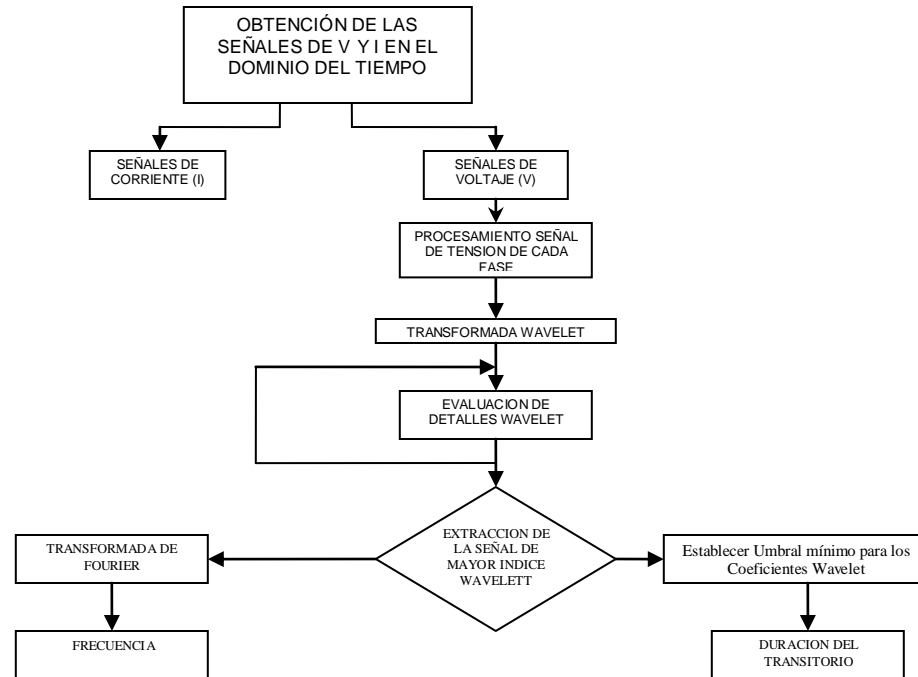


Figura 3.19. Diagrama del proceso de extracción de los descriptores de Frecuencia y duración del transitorio.

- Extracción de señales: Se extraen las señales de tensión y corriente en el dominio del tiempo para el estado de prefalla y falla.
- Estudio del estado transitorio: Se realiza para las señales de tensión ya que las protecciones del sistema están programadas para realizar su apertura al paso de la corriente por cero. Las señales de corriente no permiten observar el fenómeno transitorio con suficiente claridad, mientras que este fenómeno es más notorio en las señales de voltaje.
- Análisis de las señales de tensión de cada fase mediante la Transformada Wavelet: Según el nivel de descomposición deseado se establece el número de detalles a obtener.
- Determinación del detalle que posee el mayor coeficiente Wavelet: Se toma la señal filtrada por el detalle que posee el mayor coeficiente Wavelet y posteriormente se realiza el análisis de Fourier para obtener la frecuencia de la señal transitoria.
- Obtención de la duración del estado transitorio: A partir del tiempo en el cual los coeficientes Wavelet del detalle seleccionado (en la etapa previa) alcanzan un nivel de extinción inferior al valor umbral establecido se determina la duración del estado transitorio.

La figura 3.20 muestra la señal de tensión en la fase A para una falla monofásica con resistencia de falla a tierra de 5Ω ocurrida en el nodo 35 del sistema de pruebas prototipo. Se destaca el estado transitorio en la señal debido al evento de falla.

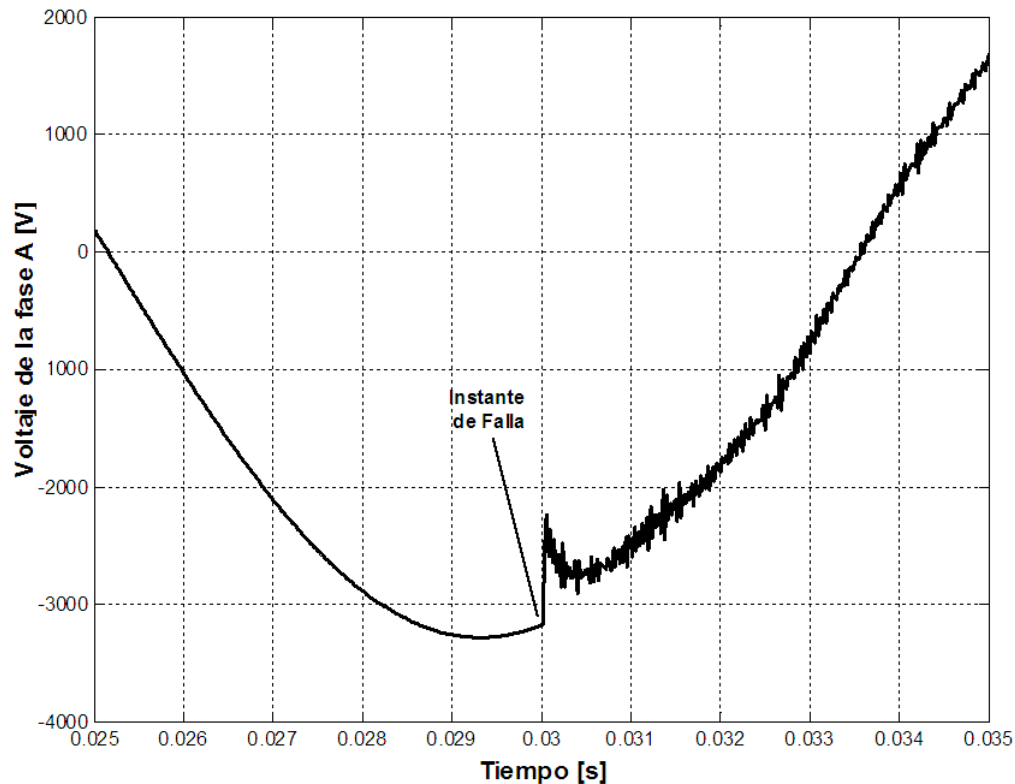


Figura 3.20. Señal de tensión de la fase A para falla Monofásica de la fase A a tierra

La figura 3.21 muestra los detalles Wavelet extraídos de la señal de tensión. La función empleada como Wavelet madre ha sido la Daubechies de orden 4, comúnmente utilizada para el análisis de eventos transitorios de señales eléctricas [Martínez 2002].

La frecuencia de muestreo de la señal es de 5 kHz, como consecuencia el detalle número uno contiene la señal de tensión de la fase A filtrada en el ancho de banda de los 2500 y 1250 Hz, el detalle dos contiene la señal filtrada entre los 1250 y 625 Hz, el detalle tres contiene la señal filtrada entre los 625 y 313 Hz, el detalle 4 contiene la señal filtrada entre los 313 y 156 Hz y finalmente el detalle 5 contiene la señal filtrada entre los 156 y 78 Hz [Martínez 2002].

El primer detalle Wavelet es el que posee el coeficiente de mayor magnitud, es entonces en este ancho de banda ([2500 1250] Hz) en donde se encuentra la principal componente de la frecuencia de oscilación de la señal transitoria. Al realizar el análisis de Fourier a la señal del detalle #1 se obtiene la principal frecuencia de oscilación del transitorio en la señal de tensión con un valor de 2050 Hz; la figura 3.22 muestra su espectro de frecuencias.

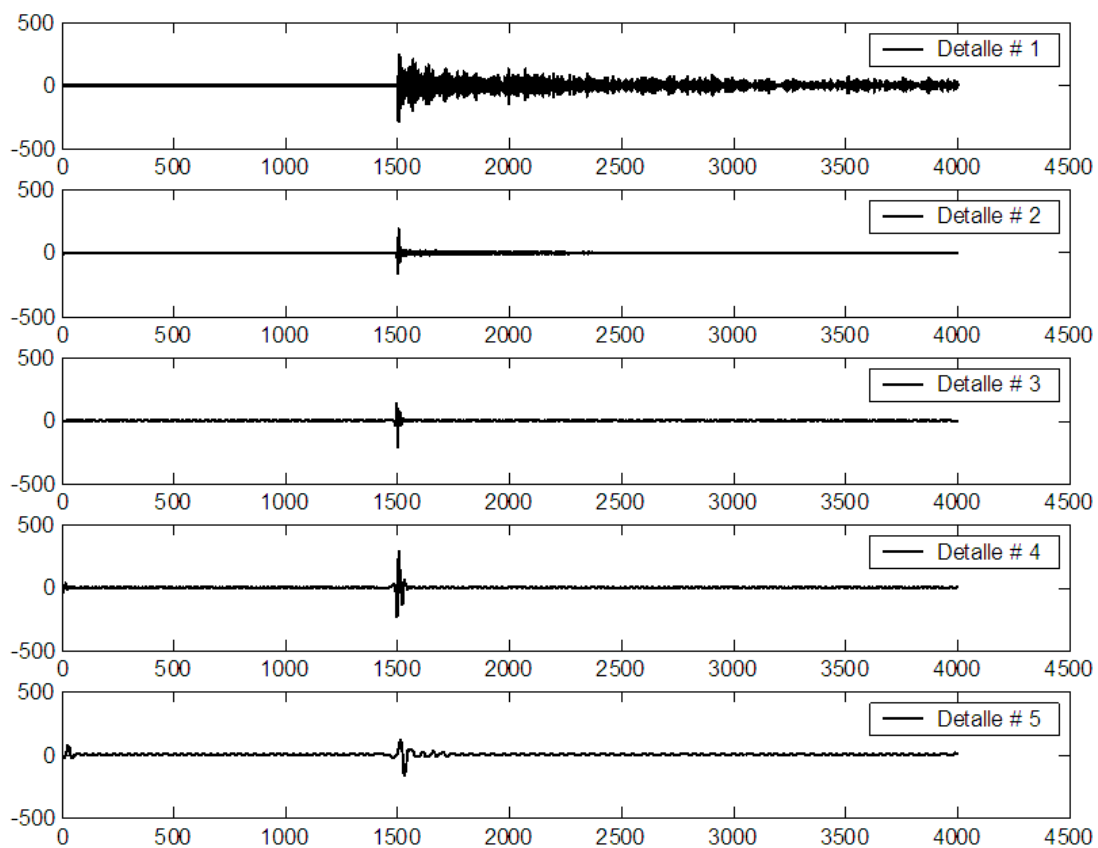


Figura 3.21. Detalles Wavelet para la señal de tensión.

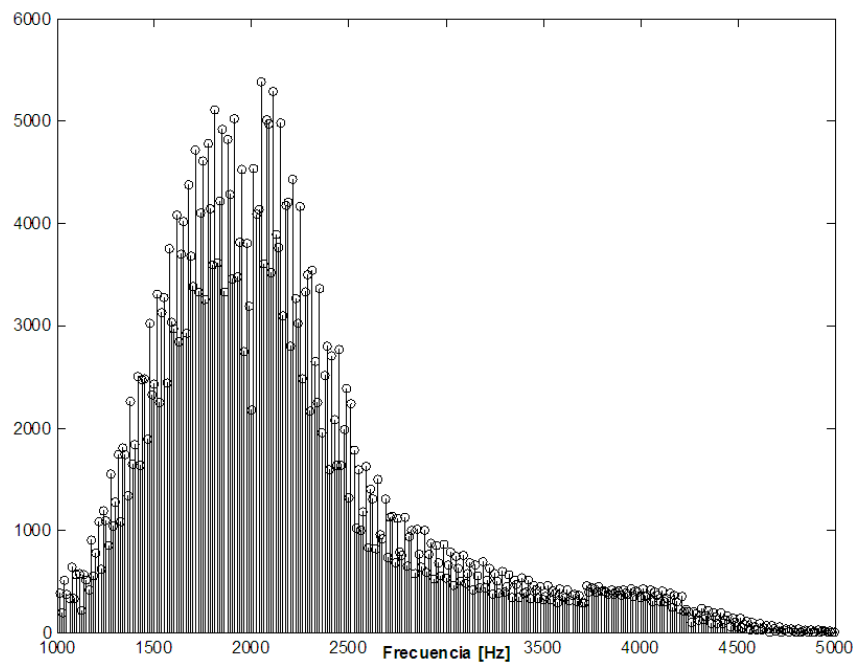


Figura 3.22. Análisis de Fourier para el detalle de mayor energía.

3.5.4 Análisis de coeficientes α y β

Esta metodología se fundamenta en la transformación de las muestras de las tres señales de corriente de falla, utilizando la transformada de Clark-Concordia [Martins *et al* (2002)]. Existe una relación entre la distancia al punto de falla y los valores propios de una matriz de correlación \mathbf{B} . La metodología para la obtención de esta matriz y sus valores propios consiste en transformar las corrientes de post falla (en forma de vector fila) en sus componentes $\alpha\beta$, aplicando una versión simplificada de la transformación Clarke-Concordia en la cual no se incluye la componente de secuencia cero (ecuación 3.3).

A partir de las señales de corriente transformadas se obtiene la matriz \mathbf{A} donde su número de filas corresponde con el número de muestras significativas, t_0 es el tiempo inicial de la muestra de datos y Δt corresponde con el tiempo de muestreo (ecuación 3.4). La matriz de correlación \mathbf{B} se calcula como se muestra en la ecuación 3.5. De los dos valores propios de esta matriz, únicamente el mayor de ellos se utiliza como descriptor del sistema bajo estado de falla y se utiliza para la definición de una función de distancia de falla (ecuación 3.6). En [Martins *et al* (2003)] se presenta una relación lineal entre el valor propio obtenido y la distancia de falla.

De esta forma, el descriptor obtenido mediante la metodología presentada, será útil para ser evaluado mediante una metodología basada en el conocimiento (MBC) para resolver el problema de localización de fallas.

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & \sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix} \begin{bmatrix} I_A \\ I_B \\ I_C \end{bmatrix} \quad (3.3)$$

$$[\mathbf{A}] = \begin{bmatrix} i_\alpha(t_0) & i_\beta(t_0) \\ i_\alpha(t_0 + \Delta t) & i_\beta(t_0 + \Delta t) \\ \vdots & \vdots \\ i_\alpha(t_0 + (n-1) \cdot \Delta t) & i_\beta(t_0 + (n-1) \cdot \Delta t) \end{bmatrix}^T \quad (3.4)$$

$$[\mathbf{B}] = [\mathbf{A}] \cdot [\mathbf{A}]^T \quad (3.5)$$

$$\lambda = \text{Max}(\text{eigvalor}[\mathbf{B}]) \quad (3.6)$$

3.5.5 Análisis de los patrones de la corriente de carga (firma del sistema)

Esta metodología permite caracterizar los registros de conmutación de los elementos de protección del sistema de distribución. La actuación de reconectores, seccionadores, interruptores en conjunto con relés de sobrecorriente instantáneos y de relación inversa tiempo corriente y fusibles, asociados a su tiempo de calibración permiten identificar cada zona del sistema y a su vez caracterizar el sistema de distribución bajo falla. [Lewis (1998)][Seung *et al* (2004)].

Este tipo de caracterización también es conocida como la “firma del sistema”, pues para cada elemento del conjunto de fallas que pueden ocurrir en el sistema de distribución, existe un patrón particular en el comportamiento de la señal de corriente el cual se encuentra ligado a la acción de las protecciones existentes en el sistema. Los descriptores asociados a los patrones de la corriente de carga son:

- Número de conmutaciones presentes en la señal.
- Paridad del número de conmutaciones.
- Tiempo de recierre de los reconectadores.
- Variación de las corrientes en estado de liberación de falla.
- Corriente de postfalla luego de la Acción de las protecciones.

A continuación se presenta de forma detallada cada uno de ellos; para un mejor entendimiento es necesario ver el sistema de pruebas descrito en la sección 4.1 y las protecciones implementadas descritas en la sección 3.3.5.

3.5.1.1 Número de conmutaciones presentes en la señal.

El número de conmutaciones presentes en las señales de corriente medidas en la subestación es el número de cambios abruptos de dicha señal; estos varían dependiendo del sitio de localización de la falla y la programación de las protecciones que definen dicha zona. Se muestra a continuación un ejemplo del comportamiento de la señal de corriente para una de las fases y se destaca allí el número de conmutaciones presentes para el caso en el que la protección primaria es un fusible y las protecciones aguas arriba corresponden a un reconectador (figura 3.23).

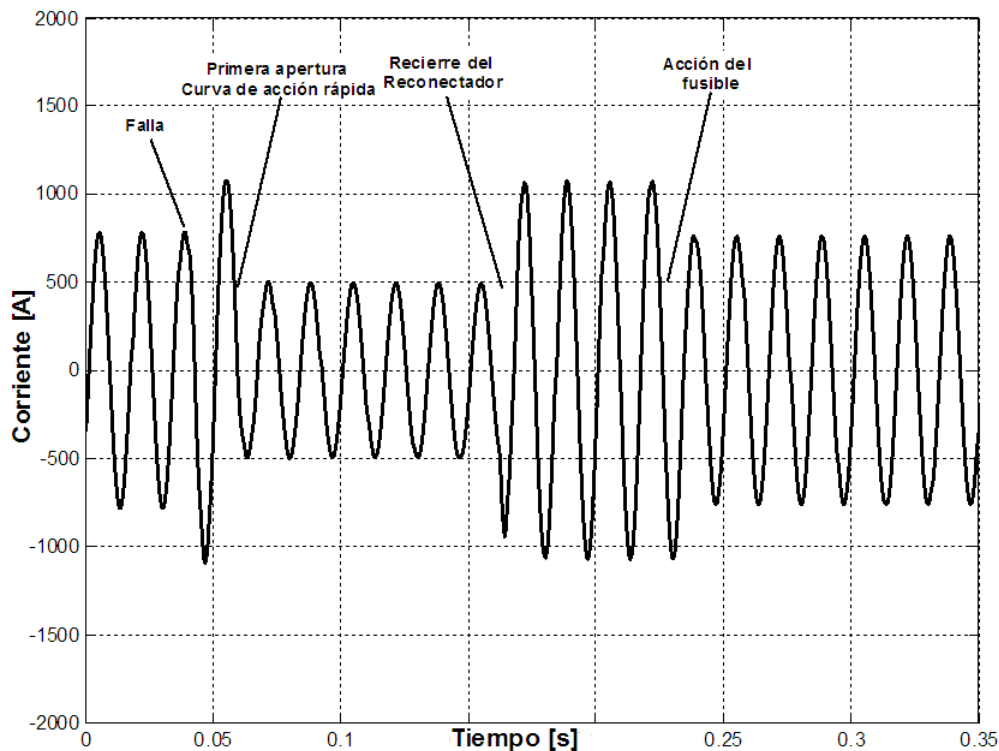


Figura 3.23. Comportamiento de la señal para falla aguas abajo de un fusible cuya protección de respaldo es un Reconectador.

3.5.1.2 Paridad del número de conmutaciones presentes en la señal

Este descriptor muestra si el número de conmutaciones es o no un número par; su utilidad radica en que discrimina de forma inmediata si la falla está localizada en una de las zonas definidas por los seccionadores del sistema.

Como aclaración, se puede observar en la figura 3.24 que el número de conmutaciones de la señal de corriente es cuatro (falla, apertura por parte del reconectador, recierre del reconectador y apertura por parte del fusible), mientras que en la figura 3.8 se observa que el número de conmutaciones presentes en la señal son cinco (falla, apertura del reconectador con la curva de acción rápida, recierre del reconectador, apertura del reconectador con la curva de acción lenta acompañada por la acción del seccionador y recierre del reconectador). Siempre que exista un seccionador, el número de conmutaciones presentes en la señal (debidas a la acción de los elementos de protección asociados a una falla en la zona del seccionador) será impar, puesto que el seccionador siempre evita una apertura final del reconectador.

3.5.1.3 Corriente de postfalla luego de la acción de las protecciones

Consiste en detectar el valor RMS de la corriente de cada una de las fases que queda después de la acción completa del conjunto de protecciones del sistema. Dado que el conjunto de protecciones desconecta cierta carga del sistema, la corriente de postfalla corresponde a la carga que permanece conectada luego de la ocurrencia de la falla (figura 3.24).

3.5.1.4 Variación de la corriente en estado de liberación de la falla

Este descriptor se encargará de medir la diferencia entre la corriente de postfalla y la corriente que queda durante la primera apertura de uno de los reconectores, para cada una de las fases. Asimismo permite discriminar que elementos de protección han actuado en el proceso de liberación de la falla. La figura 3.24 ilustra el significado de este descriptor.

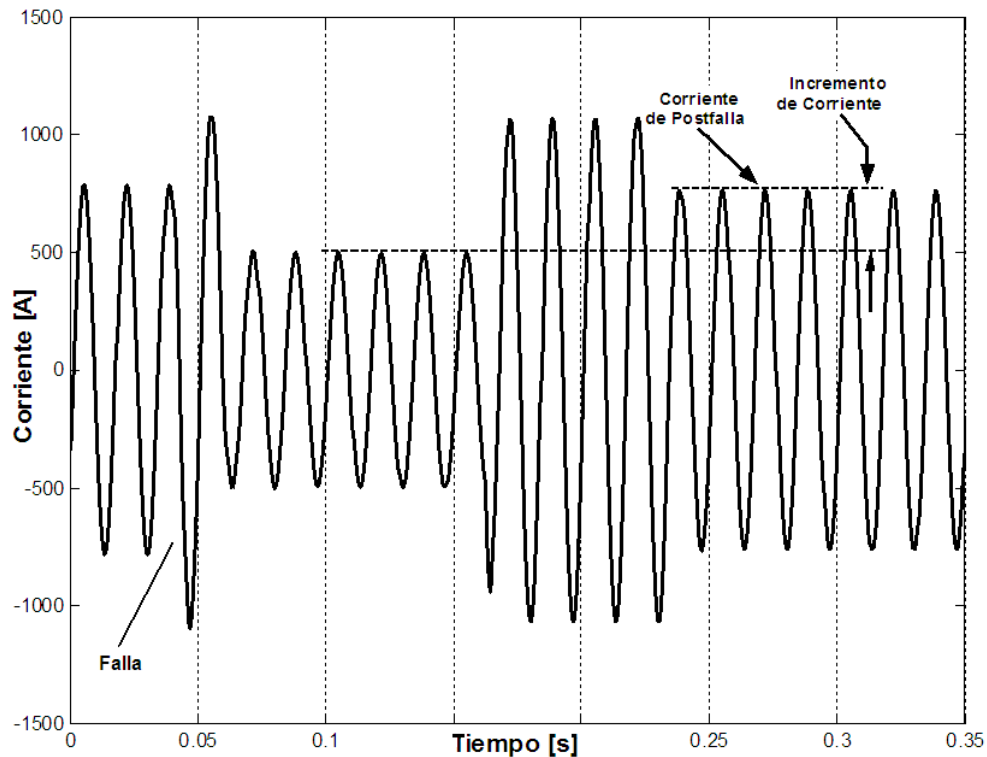


Figura 3.24. Diferencia de las corrientes durante los estados de liberación de la falla y corriente de postfalla

3.5.1.5 Tiempo de recierre de los reconectores

Este descriptor consiste en determinar el tiempo para el recierre de los Reconectores. Se debe calcular el tiempo de recierre después de la primera acción rápida, ya que no todas las fallas que pueden ocurrir en el sistema se extienden de manera que el reconector inicie su temporización dada por la curva de acción lenta. La figura 3.25 ilustra el significado de este descriptor.

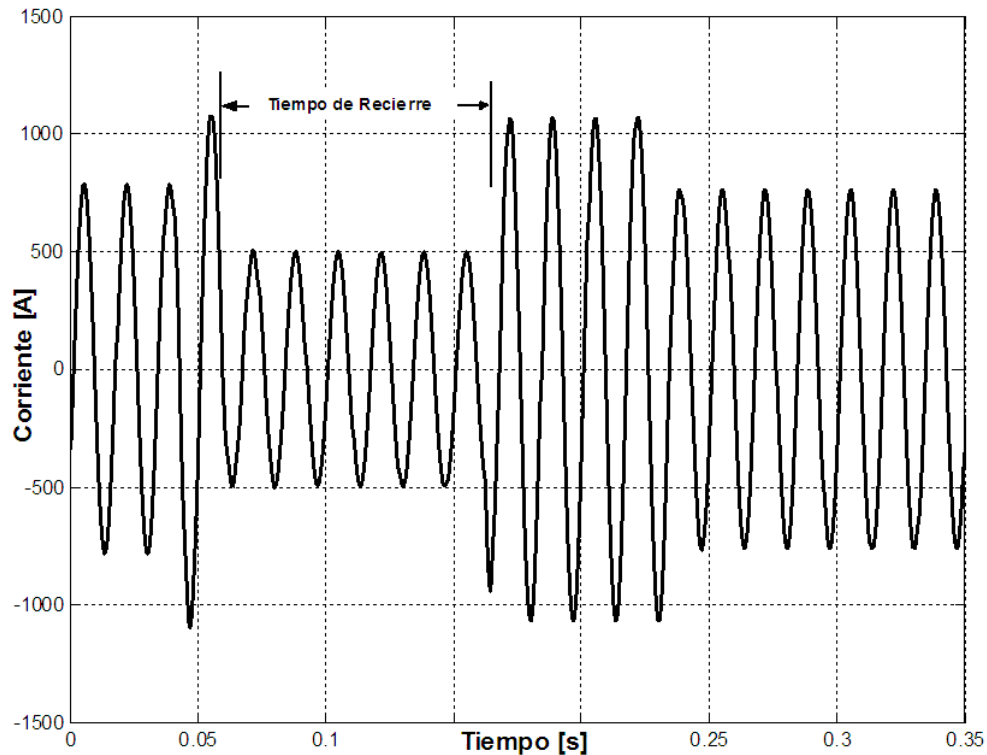


Figura 3.25. Tiempo para el recierre de los Reconectores

3.5.6 Zonificación del sistema

La división del sistema de distribución por zonas hace parte de la información necesaria para la solución del problema de Localización de fallas. Este tipo de información, hará parte de los datos de entrenamiento de la MBC escogida.

Como se mencionó anteriormente, en este trabajo se propone una metodología basada en el conocimiento (MBC) que se encargará de estimar la zona de ocurrencia de la falla y para ello es necesario realizar una etapa de entrenamiento para la MBC en la cual se requieren como datos de entrenamiento ***n-uplas*** de la forma $(\mathbf{x}_i, \mathbf{y}_i)$ donde \mathbf{x}_i corresponde con un subconjunto de la información presentada en los numerales 0 a 0 y \mathbf{y}_i corresponden con la zona del sistema asociada a la falla.

En este trabajo se proponen como zonas de la red a cada una de las zonas primarias que protegen las protecciones en el sistema de distribución; de ahí la importancia de poseer un sistema de protecciones que sea selectivo. Como ejemplo, si se tiene un ramal de un sistema de distribución que posee un reconector y un fusible conectados en un esquema de salvamento de fusibles (fusible aguas abajo del reconector) se deberán escoger dos zonas, la primera constará de los ramales que se encuentran aguas abajo del fusible y la segunda de los ramales que se encuentran entre el reconector y el fusible. Esto se puede apreciar en mayor detalle en el capítulo 4 donde se especifican cada una de las zonas escogidas en el sistema de distribución.

3.6 LOCALIZACION DE FALLAS COMO PROBLEMA DE CLASIFICACION

Dentro de la temática de localización de fallas existen diferentes variables de carácter discreto y de carácter continuo; las primeras permiten resolver el problema de localización como un problema de clasificación y las segundas permiten resolverlo como un problema de regresión ó estimación.

Dos variables discretas importantes para el enfoque de clasificación del problema de localización de fallas son las “zonas” del sistema de distribución y el tipo de falla o evento ocurrido. El objetivo de la división del sistema por zonas es crear clases o categorías para la MBC, es decir cada zona del sistema es una clase ó categoría de la MBC a las cuales pertenecerá cada patrón de entrenamiento y de validación para posterior clasificación.

En este trabajo se han tomado como MBC las Máquinas de Soporte Vectorial (SVM) y se han enfocado como la herramienta fundamental de solución al problema de localización de fallas en sistemas de distribución. Su efectividad está fundamentada en la teoría de optimización cuadrática que define su aprendizaje. Asimismo la inclusión de funciones de transformación en el problema de las SVM permite resolver problemas linealmente no separables para el espacio de las variables de entrada. Esta técnica ha sido implementada con excelentes resultados en problemas similares y de gran complejidad, donde otras metodologías han fracasado o han alcanzado inferiores resultados mediante un esfuerzo computacional mayor [Burges 2001].

A continuación se explica la metodología propuesta para dar solución al problema de la *localización de la zona* de la falla. Las SVM clasificadoras (SVMc) son la herramienta escogida para reproducir el conocimiento del sistema bajo estado de falla.

3.6.1 Zonificación del Sistema

Como se explicó en la sección 0, para poder resolver el problema de localización de fallas como un problema de clasificación, es necesario crear categorías o clases a las cuales pertenezca cada patrón o dato del problema. Para este problema, los patrones o datos corresponden con la información requerida descrita en la sección 3.5 y las categorías corresponden con las zonas en las cuales debe ser dividido el sistema de distribución; así, cada patrón o dato pertenecerá a una clase de acuerdo a la ubicación de la falla dentro del sistema de distribución.

El criterio para la zonificación del sistema de distribución debe tener presente que la identificación la zona fallada le permita a los operarios reducir el tiempo de la localización del evento para posterior reparación y mantenimiento. La zonificación del sistema está basada en la topología de la red, la longitud de los alimentadores, la ubicación de las protecciones, el tipo de carga conectada y las políticas de la empresa de distribución.

Cada zona poseerá un número determinado de nodos; en cada uno de ellos se realizó la simulación de fallas de las cuales se extrae un conjunto de datos (descriptores o información requerida); de esta forma, dentro de la base de datos de las fallas del sistema, existen subconjuntos de datos que describen el comportamiento del sistema de distribución cuando el evento de falla ha sucedido en una zona en particular.

3.6.2 Estructura de la SVMc para el problema de localización de fallas

Para el proceso de entrenamiento y validación de la SVMc para la localización de fallas, se debe implementar una estructura de siete SVMc especializadas en la localización de la zona de falla

para un tipo particular de falla (tres SVMc para tres tipos de fallas monofásicas, tres SVMc para tres tipos de fallas bifásicas y una SVMc para un tipo de falla trifásica). Cada una de las SVMc es precedida por una SVMc general que reconoce el tipo de falla y lo asigna a la correspondiente SVMc. El esquema de la figura 3.26 muestra la estructura de SVMc [Mora 2006].

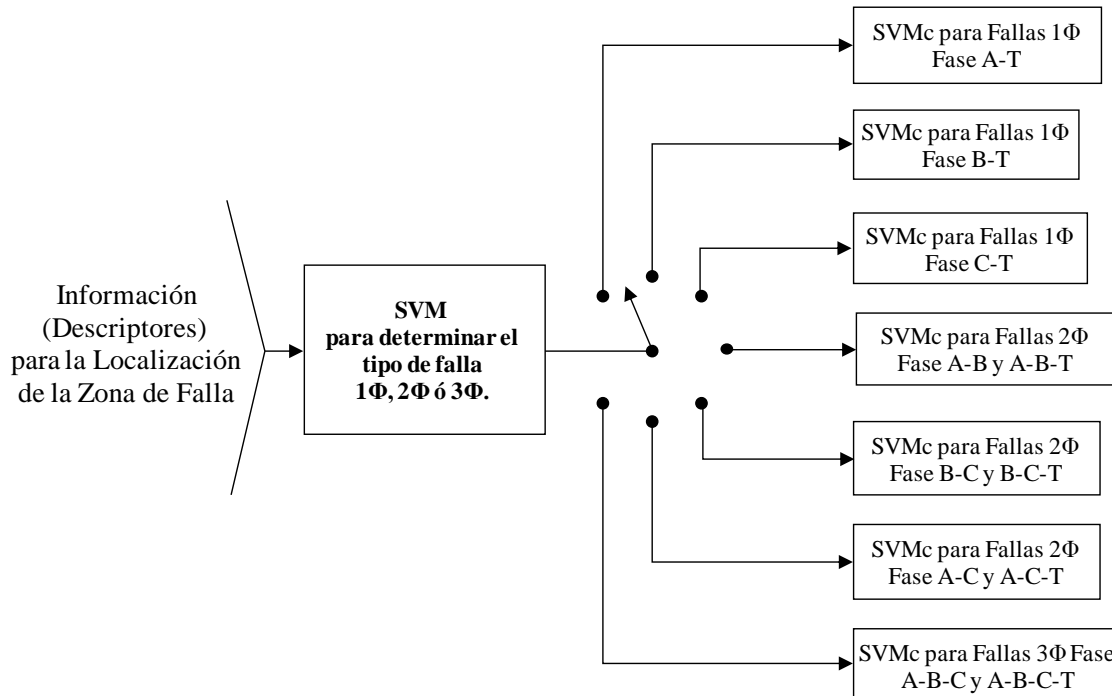


Figura 3.26. Estructura de SVMc para la localización de Fallas

Los parámetros de entrenamiento para cada SVMc corresponden a datos de la forma (\vec{x}_i, y_i) , que para el caso de localización de fallas en sistemas de distribución corresponde a un subconjunto de la información requerida (descriptor) y la zona del sistema de la cual fue obtenida dicha información.

Para la gran mayoría de sistemas de distribución reales, el resultado de la zonificación del sistema da como resultado un número de zonas grande (más de dos zonas por lo menos), por tal razón, cada una de las siete SVMc es de característica Multiclase. Se propone una SVMc multiclase basada en SVM biclasificadoras, donde el proceso de entrenamiento y votación de cada una es del tipo *Uno contra Uno* (*OvO: One vs. One*). Asimismo, cada una de las siete SVMc poseerá $\mathbf{N} \cdot (\mathbf{N}-1)/2$ SVM biclasificadoras, donde \mathbf{N} es el número total de zonas (clases) involucradas en cada SVMc. [Burges 2001] [Gunn 1998]

Se propone una estructura estándar para todas las SVM biclasificadoras, donde la función Kernel es la Función de Base Radial **RBF**, por ser la de mejor comportamiento para este tipo de aplicaciones y su fácil implementación computacional; los parámetros sigma (σ) y de Penalización (**C**) tendrán el mismo valor para todas las SVM biclasificadoras pertenecientes a una de las siete SVMc; en esta forma se evita la manipulación un número de variables bastante alto al intentar independizar los parámetros de cada SVMc biclasificadora, se agiliza el proceso de entrenamiento al aliviar el esfuerzo computacional y los resultados obtenidos son bastante satisfactorios justificando así la homogenización de los parámetros para cada SVM biclasificadora [Henao 2004].

3.6.3 Esquema de entrenamiento de la estructura de SVMc

Para el entrenamiento de cada SVMc se propone una sintonización de parámetros (σ , C) a través de una búsqueda en malla con nivel de profundidad uno; es decir, se hace una exploración en una rejilla amplia para los parámetros de cada SVMc y finalmente se debe seleccionar el valor para el cual se obtiene el menor error de validación [Lin 2004].

Para un entrenamiento apropiado de las siete SVMc se debe utilizar la metodología de validación cruzada mediante la subdivisión del conjunto de datos de entrenamiento. Para cada SVMc se debe dividir el conjunto de datos de entrenamiento en N grupos de forma que se realicen N entrenamientos (Se toman $N-1$ grupos y se valida con el grupo restante) y se promedia el error de validación de ellos. La búsqueda en malla en compañía de la validación cruzada permite seleccionar, como parámetros de la SVMc, los parámetros (σ y C) que presenten el menor error promedio de validación.

Finalmente se proponen como datos de entrada a la estructura de SVMc propuesta la información obtenida del Análisis de los patrones de la corriente de carga (firma del sistema) que son:

- Número de conmutaciones presentes en la señal.
- Paridad del número de conmutaciones.
- Tiempo de recierre de los reconectores.
- Variación de las Corrientes en estado de liberación de falla.
- Corriente de postfalla luego de la acción de las protecciones.

No todos los descriptores que se presentaron en la sección 3.5 fueron utilizados para la localización de la zona de falla, los descriptores que no se han utilizado aquí son de utilidad para la etapa de estimación de la *distancia* de falla (sección 3.7) e introducirlos y el haberlos incluido en la etapa de clasificación (localización de la zona de falla) ocasionaría que el proceso de entrenamiento de las SVMc fuese de mayor exigencia computacional ya que varios de estos descriptores poseen tres valores (uno para cada fase) por cada patrón.

Es necesario tener presente que existirán descriptores que tienen un único valor para un evento de falla en particular (para todos los tipos de falla, los primeros tres de los cinco descriptores escogidos poseen un único valor y los dos descriptores finales poseen un único valor sólo para fallas monofásicas), y existen otros descriptores que poseen uno, dos ó tres valores distintos para otros eventos de falla (los dos últimos descriptores de los cinco descriptores escogidos poseen un dato por cada fase del sistema que se encuentre involucrada en la falla). Si en el proceso de entrenamiento de las SVMc se utiliza un descriptor perteneciente a este último grupo de descriptores, la SVMc tendrá una o dos entradas más, para fallas bifásicas y trifásicas respectivamente; debido a esta situación se debe tener cuidado con los tiempos de cómputo y la exigencia computacional que puede alcanzar el proceso de entrenamiento.

Una vez seleccionados los descriptores a utilizar para cada prueba se procede con la selección de los datos de entrenamiento y los datos para la validación. El criterio para la selección de los datos de entrenamiento debe ser tal que garantice la presencia de datos de todas las zonas del sistema para cada SVMc.

3.7 PROBLEMA DE LOCALIZACION DE FALLAS COMO PROBLEMA DE REGRESION

Como se describió al inicio de la sección 3.6, la temática de localización de fallas envuelve variables de carácter discreto y de carácter continuo, donde estas últimas permiten resolver el

problema de localización como un problema de regresión ó estimación. Una variable continua muy importante para el enfoque del problema de localización de fallas como problema de regresión o estimación es la “*Distancia*” de localización de la falla respecto de la cabecera. El objeto de utilizar la distancia de falla para el problema de localización, consiste en implementar una MBC que estime, con base en una información de entrada, la distancia a la cual está localizada la falla.

Como complemento a la metodología propuesta en la sección 3.6 para la *Localización de la Zona* de falla (problema de clasificación), se propone una metodología para la *Estimación de la Distancia* de falla (problema de regresión) la cual determina la distancia eléctrica (a lo largo de los ramales del sistema de distribución) desde el nodo de la subestación o cabecera del sistema de distribución hasta el sitio donde ha ocurrido la falla. La metodología propuesta está basada en SVM para Regresión (SVMr).

Como se explico en el capítulo 2, el modelo de optimización de las SVM no sólo se ha aplicado a problemas de clasificación sino que se ha extendido a problemas de regresión o estimación de variables logrando resultados satisfactorios que otras metodologías han alcanzado con esfuerzos computacionales más altos.

3.7.1 Distancia de la localización de la falla en el sistema.

La distancia de la localización de la falla es definida como la longitud a lo largo de los ramales del circuito que unen la subestación principal o cabecera del sistema con el nodo del sistema en el cual ha ocurrido el evento de falla. Dado que el sistema de distribución es radial, el conjunto de ramales que unen el nodo de falla con la cabecera del sistema es único, de modo que la distancia de la localización de la falla tiene un único valor para cada evento.

Recorrer una distancia determinada desde la cabecera del sistema no siempre conduce al punto de localización de la falla. Si el sistema de distribución radial posee múltiples ramales, existirán diversas alternativas para el recorrido dependiendo del número de ramales. Esta dificultad en la localización del sitio de falla a partir de la distancia se conoce como el problema de la múltiple estimación [Das 1998].

Dado que las fallas realizadas para el sistema de distribución fueron simuladas en cada uno de los nodos del sistema, la distancia de localización de la falla corresponde con la distancia de localización de cada uno de los nodos en el sistema. Estas distancias vienen dictadas por la longitud de las líneas de potencia que conectan la cabecera del sistema con el nodo.

La estructura para la estimación de la distancia de falla consiste de varias SVMr que tienen como entradas los datos extraídos del sistema de distribución; sin embargo una estructura que sólo consista de SVMr es impráctica y poco eficiente en sistemas reales debido a la cantidad de información y al esfuerzo computacional requerido para el entrenamiento de las mismas. Es por esta razón que se propone una estructura que subdivida el trabajo por cada tipo de falla y que además subdivida el problema por cada zona del sistema de distribución (se deben trabajar con las mismas zonas implementadas en la metodología propuesta en la sección 3.6). Esta estructura híbrida es explicada en 3.8 y su implementación está justificada a partir de la especialización de tareas para cada SVMr; asimismo cada SVMr está asociada a cada zona del sistema de distribución y se encargará de estimar la distancia de falla únicamente para su zona respectiva.

3.7.2 Esquema de entrenamiento de la estructura de la estructura basada en SVMr para la estimación de la distancia de falla.

Los parámetros de entrenamiento para cada SVMr corresponden a datos de la forma (\mathbf{x}_i, y_i) , que para el caso de la estimación de la distancia, \mathbf{x}_i corresponde a un subconjunto de la información requerida (descriptores) y y_i a la distancia de la ubicación de la falla de la cual fue obtenida dicha información.

Se propone utilizar como función kernel la función de base radial **RBF**, por ser la de mejor comportamiento para este tipo de aplicaciones y su fácil implementación computacional [Russell et al (2000)] [Burges 2001]

Para el entrenamiento de cada SVMr se propone una sintonización de parámetros (σ, \mathbf{C}) a través de una búsqueda en malla con nivel de profundidad tres; es decir, se hace una exploración en una rejilla amplia para los parámetros $(\sigma$ y $\mathbf{C})$, se selecciona el punto para el cual se obtiene el menor error de validación y allí se establece una nueva rejilla de menor amplitud para realizar una nueva búsqueda; este procedimiento se realiza tres veces para completar la búsqueda en malla y seleccionar los mejores parámetros (dentro del espacio explorado) para cada SVMr. Para un entrenamiento apropiado de las SVMr se debe utilizar la metodología de validación cruzada mediante la subdivisión del conjunto de datos de entrenamiento. Para cada SVMr se debe dividir el conjunto de datos de entrenamiento en \mathbf{N} grupos de forma que se realicen \mathbf{N} entrenamientos (Se toman $\mathbf{N}-1$ grupos y se valida con el grupo restante) y se promedia el error de validación de ellos. La búsqueda en Malla en compañía de la validación cruzada permite seleccionar, como parámetros de la SVMr, los parámetros $(\sigma$ y $\mathbf{C})$ que presenten el menor error promedio de validación.

Finalmente se proponen como datos de entrada a la estructura de SVMr, la información obtenida de:

- Huecos de tensión con Voltajes de fase.
- Huecos de tensión con Voltajes de línea.
- Frecuencia de oscilación de la señal transitoria del estado de falla.
- Duración de la señal transitoria del estado de falla.

Es necesario tener presente para el proceso de entrenamiento de las SVMr los siguientes aspectos:

- Los descriptores que se proponen para esta metodología de estimación de la distancia de falla se encuentran en el espacio de dimensión tres, pues todos ellos poseen un valor por cada fase del sistema.
- El proceso de entrenamiento involucra una matriz Jacobbiana en la función objetivo (proceso de optimización asociado) cuyo número de filas y columnas es del doble del número de datos utilizados para dicho entrenamiento (ver capítulo 2)

De este modo, para conjuntos de entrenamiento bastante grandes y datos de entrada en espacios de gran dimensión, el proceso de entrenamiento puede requerir de elevados tiempos de cómputo haciéndose en ocasiones inapropiado. Se recomienda que para el proceso de entrenamiento de las SVMr se implementen cuatro estructuras de SVMr donde cada una opere con uno sólo de los cuatro descriptores propuestos y se valide el desempeño de la metodología propuesta para cada caso. En cuanto a la selección de los datos de entrenamiento se debe garantizar la presencia de datos de todos los nodos de cada zona asociado a cada SVMr.

3.8 ESTRUCTURA HÍBRIDA PROPUESTA PARA LA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN.

Este trabajo propone una nueva metodología para la localización de fallas en sistemas de Distribución radial. La localización será descrita por dos variables que son la *Zona* y la *Distancia* de localización de la falla. Esta nueva metodología basada en Máquinas de Soporte Vectorial, reúne las ventajas de las SVMc y de las SVMr como MBC básicas para la localización.

Como se describió en la sección 3.6, la herramienta para la Localización de la zona de falla es una estructura basada en SVMc. En la sección 3.7 se describieron las bases de la estructura basada en SVMr para la estimación de la distancia de falla. La metodología propuesta reúne las mejores características de las estructuras propuestas creando una metodología híbrida cuyo funcionamiento se muestra en la figura 3.27. Con este nuevo esquema, la *Estimación de la Distancia* de falla queda complementada con la *Localización de la Zona* de falla, y cada SVMr del para el problema de estimación de la distancia se especializa en la estimación para cada zona del sistema de distribución; un esquema más detallado se presenta en la figura 3.28.

Al igual que para el proceso de Localización de la zona de falla, también se ha propuesto una estructura de SVM para facilitar el proceso de entrenamiento. En esta estructura deben interactuar tanto SVM para Clasificación como SVM para regresión, para el correcto desempeño de la metodología propuesta para la estimación de la distancia de falla.

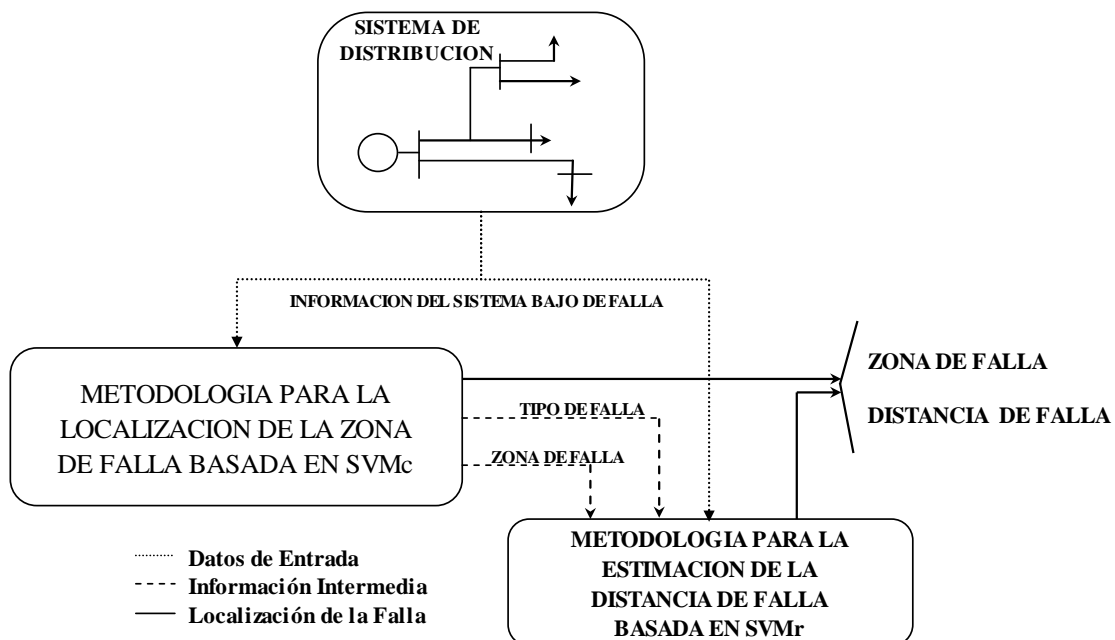


Figura 3.27. Estructura Híbrida propuesta para la Localización de Fallas en Sistemas de Distribución.

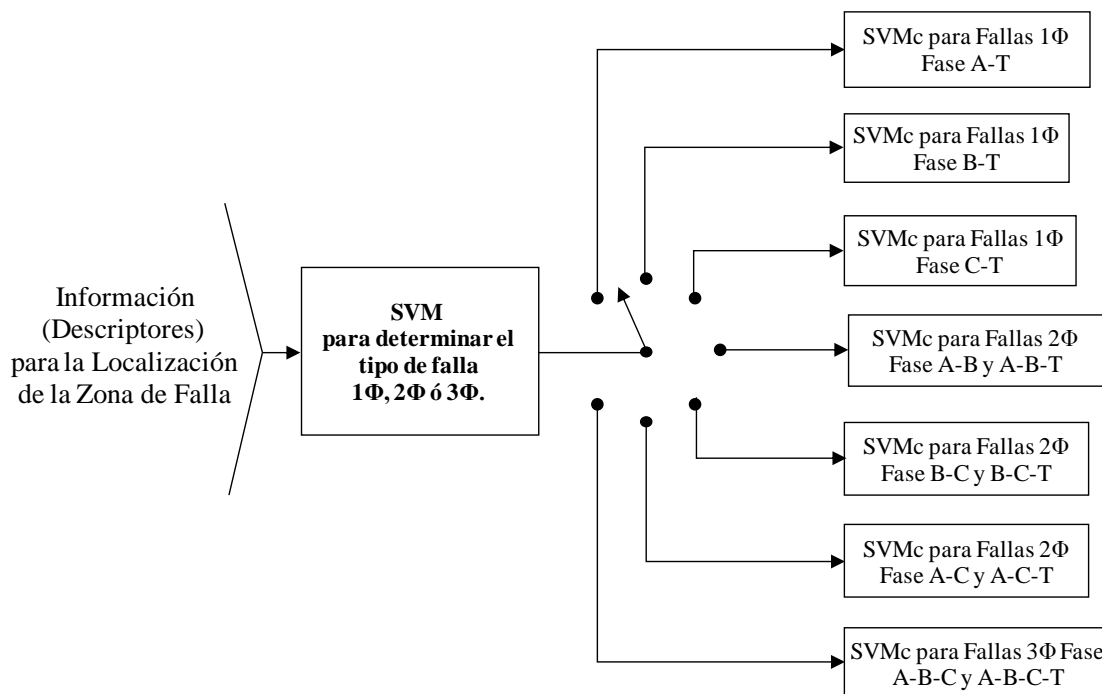


Figura 3.28. Estructura de las SVMr para la estimación de la distancia de Falla.

Como se muestra en la figura 3.28, el trabajo para la estimación de la distancia de la falla es llevado a cabo por un conjunto de SVMr, donde cada una se especializa en estimar distancias para un tipo de falla y una zona en particular.

Esta propuesta se fundamenta en que se logran obtener resultados satisfactorios a un menor costo computacional, cuando se subdivide el problema y se asigna la estimación a pequeñas SVMr especializadas por zonas, que cuando se implementan grandes SVMr que generalizan el aprendizaje para un tipo de falla en particular para todo el sistema de distribución completo. Esta estructura híbrida propuesta combina las bondades de las SVM para clasificación y Regresión para la estimación de la distancia de falla.

4. APLICACIÓN DE LAS SVM AL PROBLEMA DE LA LOCALIZACIÓN DE FALLAS EN SISTEMAS DE DISTRIBUCIÓN

4.1 INTRODUCCIÓN

Como se presentó en el capítulo 2, diversas metodologías basadas en el modelo de la red (MBM) pueden ser empleadas para resolver el problema de localización de fallas, dichas metodologías traen consigo el inconveniente de la múltiple estimación; también se presentaron diversas metodologías basadas en el conocimiento (MBC) y en particular se presentaron las Máquinas de Soporte Vectorial para clasificación (SVMc) y regresión (SVMr) como herramienta de solución al problema de localización de fallas en sistemas de distribución.

Posteriormente, en el capítulo 3 se mostraron los aspectos necesarios, tales como la zonificación del sistema, la simulación automática de fallas, la extracción de los descriptores entre otras, y se planteó la estructura basada en SVMc y SVMr como herramienta de solución del problema de localización de fallas.

En este capítulo se muestra la aplicación de la herramienta propuesta para la solución del problema de localización de fallas en sistemas de distribución. Se tomo como modelo la red de distribución de 123 nodos del Subcomité de Análisis de Sistemas de Distribución de la IEEE Power Engineering Society; inicialmente se describen los elementos que lo conforman y el modelamiento eléctrico que se le dio a cada uno de ellos; entre ellos se muestran las protecciones del sistema y el esquema de salvamento de fusibles implementado en el proceso de calibración de las mismas (las protecciones juegan un papel importante en la solución del problema a través de las MBC); se describen también el proceso de sintonización de los parámetros de la MBC y el análisis de los resultados obtenidos al implementar la metodología de localización de fallas. Para el modelamiento del sistema de distribución se empleó el software ATP (Analysis Transients Program); el sistema de distribución de 123 Nodos se presenta en la figura 4.1 [IEEE 123 Node Test].

4.2 SISTEMA DE DISTRIBUCION PARA PRUEBAS

El sistema de distribución escogido para pruebas consta de una subestación principal o cabecera del sistema, interruptores para reconfiguración de la red, compensadores de potencia reactiva, reguladores de tensión, protecciones y finalmente las cargas respectivas; a continuación se describen cada uno de estos elementos.

4.2.1 Subestación

La cabecera del sistema de distribución consiste de una subestación de 5 MVA 115/4.16 kV. Todos los nodos del sistema se encuentran a una tensión de 4160 V a excepción del nodo 610 que opera a 480 V.

4.2.2 Interruptores

El sistema de distribución de pruebas posee 11 interruptores de los cuales 6 se encuentran permanentemente cerrados y los cinco restantes permanecen abiertos; la función de estos

interruptores es permitir la reconfiguración o modificación de la topología del sistema de distribución. En este trabajo se simuló el sistema de distribución con la topología original, es decir la establecida por los interruptores que normalmente permanecen cerrados (topología fija que se presenta en la figura 4.1). La aplicabilidad de la metodología propuesta para la localización de fallas también es aplicable para diversas topologías del sistema de distribución.

4.2.3 Cargas

El conjunto de cargas del sistema de distribución incluye cargas monofásicas, bifásicas y trifásicas cuyo comportamiento puede modelarse como cargas de potencia constante, corriente constante ó impedancia constante; en este trabajo se consideraron todas las cargas con un modelo de impedancia constante, ya que otro tipo de modelamiento implicaría la solución de un sistema no lineal de ecuaciones (para encontrar el voltaje y la corriente en el dominio del tiempo) haciendo que el proceso de simulación de fallas fuese de difícil implementación computacional; los resultados obtenidos con el modelo de impedancia constante representa de manera muy aproximada el comportamiento real de la red; adicionalmente, para los propósitos establecidos en este trabajo, el modelo de impedancia constante es suficiente para los objetivos planteados.

4.2.4 Inyección de Reactivos

El sistema de distribución posee elementos de inyección de potencia reactiva (capacitores) ubicados en los nodos 83, 88, 90 y 92. Las inyecciones de cada uno se presentan en el anexo 1.

4.2.5 Reguladores de Tensión

El sistema de distribución escogido para pruebas posee cuatro reguladores de tensión, ubicados así: En la cabecera del sistema, entre los nodos 9-14, 25-26 y los nodos 160-67. El modelo del regulador de tensión obedece a una variante del modelo del autotransformador y su implementación en el software de simulación (*ATP*) eleva los tiempos de simulación y la complejidad computacional. Para los propósitos de este trabajo, las señales de tensión y de corriente medidas en la cabecera de la subestación, producto de la simulación de fallas, cuando se han incluido este tipo de elementos son bastante aproximadas a las señales obtenidas cuando estos elementos no son tenidos en cuenta.

De acuerdo con lo anterior, la metodología de localización de fallas propuesta en este trabajo no se ve alterada de manera significativa con la ausencia de estos elementos debido a que la solución del problema de localización de fallas está soportada por una metodología basada en el conocimiento (MBC), que si bien está fundamentada en el modelo eléctrico de la red y la adecuada forma en la que éste represente el sistema real, se ajusta para diversas variaciones del modelo de red escogido, gracias a su robustez obtenido del aprendizaje a partir de ejemplos. Adicionalmente la inclusión de los reguladores de tensión dentro de la simulación del sistema de pruebas aumenta la exigencia computacional del proceso; como se mencionó anteriormente, para la correcta implementación de esta metodología se requiere hacer una simulación automática de fallas y en la medida que el modelo de la red se torne complejo el tiempo de cómputo puede ser impráctico; más adelante se explicará con mayor detalle que el número de simulaciones a realizar para el sistema

de pruebas escogido es de 4339, así que el tiempo de cómputo para cada simulación es de vital importancia en este problema.

4.2.6 Protecciones

El conjunto de las protecciones implementadas dentro del sistema de distribución consta de tres reconectores, cuatro seccionalizadores y 19 fusibles ubicados tanto en nodos monofásicos como trifásicos. La ubicación de las protecciones se hizo bajo criterios de selectividad, confiabilidad y económicos y han sido ubicados en un esquema de salvamento de fusibles.

El reconector principal R0 se encuentra ubicado en el nodo cabecera del sistema de distribución y se ha configurado para que ejecute hasta dos aperturas, es decir siempre que una falla en el sistema genere una sobrecorriente superior a la corriente de activación de R0, este reconector, luego de superado el tiempo de disparo (dado por la curva inversa de tiempo-corriente), realizará la apertura de los interruptores desenergizando el circuito por completo (una sola fase en caso de falla monofásica o las tres fases para otros eventos de falla); posteriormente la protección efectúa un recierre de los interruptores (el tiempo para el recierre es de 300 ms) y el circuito se energiza nuevamente; si la falla persiste se realiza la segunda operación de apertura (luego de transcurrido el tiempo de disparo) y el circuito quedará desenergizado permanentemente.

Los reconectores R1 y R2 ubicados entre los nodos 13-142 y 13-143 respectivamente, constituyen la protección inicial de los dos ramales principales del circuito de distribución; poseen una operación similar al reconector R0, excepto que R1 y R2 poseen tres acciones de apertura, la primera de ellas es de acción rápida (tiempo de disparo y de recierre reducidos ó ligeramente menores a los tiempos de acción de los fusibles ubicados aguas abajo de ellos) y las dos restantes son lentas (tiempos de disparo y recierre elevados o ligeramente mayores a los tiempos de acción de fusibles ubicados aguas abajo de ellos). El tiempo de acción de los fusibles ubicados aguas abajo de los reconectores se encuentra dentro del intervalo definido por los tiempos de acción de la curva rápida y la curva lenta de estos últimos; esto es con el fin de que el esquema de salvamento de fusibles tenga un funcionamiento apropiado (ver sección 3.4.5.4). También se aclara que el tiempo de acción para el reconector R0 siempre es mayor que el tiempo de disparo de R1 y R2 con el fin de garantizar la selectividad de las protecciones en el sistema de distribución. La ubicación de todas las protecciones, la corriente RMS que cada una experimenta y la corriente de activación con las cuales fueron calibradas se muestran en la tabla 4.1

	Nodo Inicial	Nodo Final	Corriente Nominal			Corriente de Activación		
			Fase A	Fase B	Fase C	Fase A	Fase B	Fase C
F1	1	119	-	9,1	-	-	13,7	-
F2	1	120	-	-	44,8	-	-	67,2
F3	8	121	43,5	-	-	65,3	-	-
F4	8	122	-	9,1	-	-	13,6	-
F5	47	123	32,7	34,5	33,6	49,1	51,8	50,4
F8	47	124	24,8	34,6	34,2	37,2	51,9	51,3
F11	44	125	17,0	-	-	25,5	-	-
F12	42	126	-	18,0	-	-	26,9	-
F13	35	127	17,1	18,0	-	25,6	27,0	-
F15	25	128	17,1	-	17,5	25,6	-	26,2

	Nodo Inicial	Nodo Final	Corriente Nominal			Corriente de Activación		
			Fase A	Fase B	Fase C	Fase A	Fase B	Fase C
F17	25	129	33,9	-	17,5	50,9	-	26,2
F20	18	130	34,2	-	0,0	51,4	-	-
F21	97	131	17,0	17,0	17,0	50,9	50,9	50,9
F24	67	132	50,6	-	-	75,9	-	-
F25	76	133	30,4	56,5	19,5	45,7	84,8	29,3
F28	76	134	68,3	72,3	69,9	102,4	108,5	104,9
F31	108	135	58,6	-	-	87,9	-	-
F32	101	136	-	-	43,3	-	-	86,6
F33	13	137	-	-	44,2	-	-	66,3
SC1	18	138	133,2	122,0	93,2	159,9	146,4	111,9
SC2	60	139	33,1	50,1	82,2	39,8	60,1	98,6
SC3	67	140	106,4	121,9	124,9	127,7	146,3	149,9
SC4	67	141	75,5	53,7	60,7	90,6	64,4	72,8
R1	13	142	218,6	139,9	145,5	262,4	167,9	174,6
R2	13	143	270,8	241,8	265,2	325,0	302,3	331,5
S/E	SBST	0	544,5	385,4	483,0	626,2	443,2	555,4

Tabla 4.1. Ubicación de las protecciones en el sistema y corrientes de activación

El sistema de distribución es desbalanceado y por cada una de las tres fases circulan corrientes de magnitudes diferentes. La calibración y la coordinación del conjunto de protecciones se realizó teniendo presente este fenómeno, de modo que las corrientes de activación (I_{PK}) de una protección trifásica difieren entre sí para cada fase. Por otra parte, es importante tener presente que las fallas de alta impedancia que ocurren en los nodos más alejados de la protección generan corrientes de activación muy pequeñas y por tal razón los tiempos de disparo de la protecciones pueden convertirse en tiempos bastante elevados. Se debe ajustar la corriente de activación de tal forma que no se generen tiempos de apertura extensos para fallas de alta impedancia asociadas a los nodos lejanos a la protección.

La calibración de las corrientes de activación de las protecciones se realizó por medio de un proceso de simulación de fallas de alta impedancia en el nodo más lejano de la zona de cada protección y con la medición de las corrientes que circulan por cada uno de los elementos involucrados con la falta simulada. De la misma manera, se simularon fallas de baja impedancia en el nodo más cercano a cada protección. Con base en estos dos escenarios se calibraron las corrientes de activación y los tiempos de disparo para cada una de las protecciones. Por ejemplo, en la tabla 4.2 se muestra que el fusible F25 se encuentra asociado a la zona 23 del sistema de distribución; el nodo trifásico más alejado de esta protección es el nodo 95 (ver figura 4.1) y en él se simuló una falla trifásica a tierra a través de una resistencia de 30Ω (falla de alta impedancia); las corrientes medidas en el punto donde se encuentra el fusible F25 son de 101,8 A, 130,1 A y 89,8 A para cada fase respectivamente; aguas arriba del fusible F25 se encuentra ubicado el reconectador R2 quien también está involucrado con el evento de falla; las corrientes medidas en el punto donde se encuentra este reconectador son 338,0 A, 314,7 A y 336,6 A; con base en estos valores se logró calibrar al fusible F25 y adicionalmente, teniendo presente la información de todos los fusibles ubicados aguas abajo del reconectador R2, se logró calibrar la curva de acción lenta de este último elemento. Por otra parte, se simuló una falla de baja impedancia en el nodo más cercano al fusible F25 (nodo 86), empleando un valor de $0,1 \Omega$ como resistencia de falla; las corrientes medidas en el punto donde se encuentra el fusible F25 son 2462,9 A, 2619,8 A y 2616,3 A y las corrientes medidas en el punto donde se encuentra el reconectador R2 son 2520,8 A, 2672,6 A y 2677,1 A; con base en esta información y la asociada a todos los fusibles conectados aguas abajo de R2 se logró calibrar la curva de acción rápida de este reconectador.

Por otra parte, los seccionadores de la red de distribución fueron programados para despejar la falta bajo la ausencia de carga luego de dos aperturas de los reconectores que se encuentran aguas arriba de él. Los seccionadores del sistema de distribución fueron ubicados en los nodos 65, 62, 72 y 97.

PROTECCION	ZONA	NODOS													
R0	1	1	7	8	13										
F1	2	2	119												
F2	3	3	120	4	5	6									
F3	4	9	10	121	11	14									
F4	5	12	122												
F33	6	34	15	16	17	137									
R1	7	18	21	22	23	24	25	142							
SC1	8	35	40	41	42	44	47	138							
F5	9	48	123												
F8	10	49	50	51	124										
F11	11	45	46	125											
F12	12	43	126												
F13	13	36	37	38	39	127									
F15	14	26	27	31	32	33	128								
F17	15	28	29	30	117	129									
F20	16	19	20	130											
R2	17	52	53	54	55	56	57	58	59	60	61	67	143		
SC2	18	62	63	64	65	66	139								
SC3	19	72	73	74	75	76	140								
SC4	20	97	101	105	106	107	108	118	141						
F21	21	98	99	100	116	131									
F24	22	68	69	70	71	132									
F25	23	86	87	88	89	90	91	92	93	94	95	96	133		
F28	24	77	78	79	80	81	82	83	84	85	134				
F31	25	109	110	111	112	113	114	135							
F32	26	102	103	104	136										

Tabla 4.2. Zonas del sistema y Nodos pertenecientes a cada una.

La inclusión de las protecciones en el sistema de distribución ha aumentado el número de nodos de 123 a un total de 142. La tabla 4.3 presenta los parámetros con los cuales fueron calibradas cada una de las protecciones del sistema.

	A	B	P	Dial
F1	0,0470	0,183	0,02	0,020
F2	0,0470	0,183	0,02	0,020
F3	0,0470	0,183	0,02	0,020
F4	0,0470	0,183	0,02	0,020
F5	0,0470	0,183	0,02	0,035
F8	0,0470	0,183	0,02	0,050
F11	0,0470	0,183	0,02	0,060
F12	0,0470	0,183	0,02	0,060
F13	0,0470	0,183	0,02	0,060
F15	0,0470	0,183	0,02	0,060
F17	0,0470	0,183	0,02	0,070
F20	0,0470	0,183	0,02	0,040
F21	0,0470	0,183	0,02	0,100
F24	0,0470	0,183	0,02	0,080
F25	0,0470	0,183	0,02	0,140
F28	0,0470	0,183	0,02	0,050
F31	0,0470	0,183	0,02	0,070
F32	0,0470	0,183	0,02	0,050
F33	0,0470	0,183	0,02	0,020

SC1	Número de Recierres: 2
SC2	Número de Recierres: 2
SC3	Número de Recierres: 2
SC4	Número de Recierres: 2

	A	B	P	DIAL	Número de Acciones de Apertura	Tiempo de Recierre
R1	Parámetros de Acción Rápida	0.047	0.183	0.02	0.004	Rápida: 1
	Parámetros de Acción Lenta	0.047	0.183	0.02	0.050	Lenta: 2
R2	Parámetros de Acción Rápida	0.047	0.183	0.02	0.008	Rápida: 1
	Parámetros de Acción Lenta	0.047	0.183	0.02	0.090	Lenta: 2
R0		0.0147	0.183	0.02	0.050	2
						300 ms

$$t_D = D \cdot \left(B + \frac{A}{(I/I_{PK})^P - 1} \right)$$

- t_D : Tiempo de Disparo para la protección.
D: Valor del Dial de la protección.
I: Corriente medida en el lado de la protección
 I_{PK} : Corriente de activación (*Corriente de Pick Up*)

Tabla 4.3. Parámetros de las protecciones del sistema de distribución.

Las MBC propuesta para la solución del problema de localización de la zona de falla está basada en Máquinas de Soporte Vectorial para Clasificación (SVMc ver sección 3.6) las cuales requieren que los datos a clasificar pertenezcan a una categoría o clase en particular; para este tipo de trabajo, el conjunto de categorías corresponde con cada una de las 26 zonas del sistema de distribución y cada una de estas zonas se encuentra definida por cada protección del sistema tal como se presenta en la tabla 4.2.

4.2.7 Distribución de nodos

Cada uno de los nodos está definido por las líneas del sistema de distribución y las cargas conectadas en el sistema. Los parámetros eléctricos, la configuración de cada línea y la carga conectada en cada nodo se muestran en el anexo 1 y en [IEEE 123 Node Test].

El sistema de distribución posee 71 nodos trifásicos; tres nodos bifásicos pertenecientes a las fases A-C, dos nodos bifásicos pertenecientes a las fases A-B; y 66 nodos monofásicos de los cuales 27, 15 y 24 nodos pertenecen a las fases A, B y C respectivamente. En total existen 142 nodos en el circuito. Para efectos de la simulación y el modelamiento, no se incluyó el transformador conectado al nodo 60 (cuyo lado secundario define el nodo 115); de acuerdo con [IEEE 23 Node Test] este transformador no posee carga alguna conectada en su lado secundario

así que para propósitos del problema de localización de fallas, no es necesario incluirlo dentro de la simulación de fallas (en el software ATP) y no se incurrir en errores del modelamiento de la red. La distribución de los nodos se resume así:

- Nodos Monofásicos Fase A: 9, 37, 94, 10, 45, 109, 11, 46, 110, 14, 68, 111, 19, 69, 112, 20, 70, 113, 33, 71, 114, 121, 125, 130, 132, 135, 88.
- Nodos Monofásicos Fase B: 2, 59, 12, 22, 90, 38, 96, 39, 106, 43, 107, 58, 119, 122, 126.
- Nodos Monofásicos Fase C: 3, 24, 75, 4, 31, 84, 5, 32, 85, 6, 34, 92, 15, 41, 102, 16, 73, 103, 17, 74, 104, 120, 136, 137.
- Nodos Bifásicos Fases A-C: 26, 27, 128.
- Nodos Bifásicos Fases A-B: 36, 127.
- Nodos Trifásicos: 1, 30, 52, 64, 81, 98, 7, 35, 53, 65, 82, 99, 8, 40, 54, 66, 83, 100, 13, 42, 55, 67, 86, 101, 18, 44, 56, 72, 87, 105, 21, 47, 57, 76, 89, 108, 23, 48, 60, 77, 91, 116, 25, 49, 61, 78, 93, 117, 28, 50, 62, 79, 95, 118, 29, 51, 63, 80, 97, 134, 123, 124, 129, 131, 133, 143, 138, 139, 140, 141, 142.

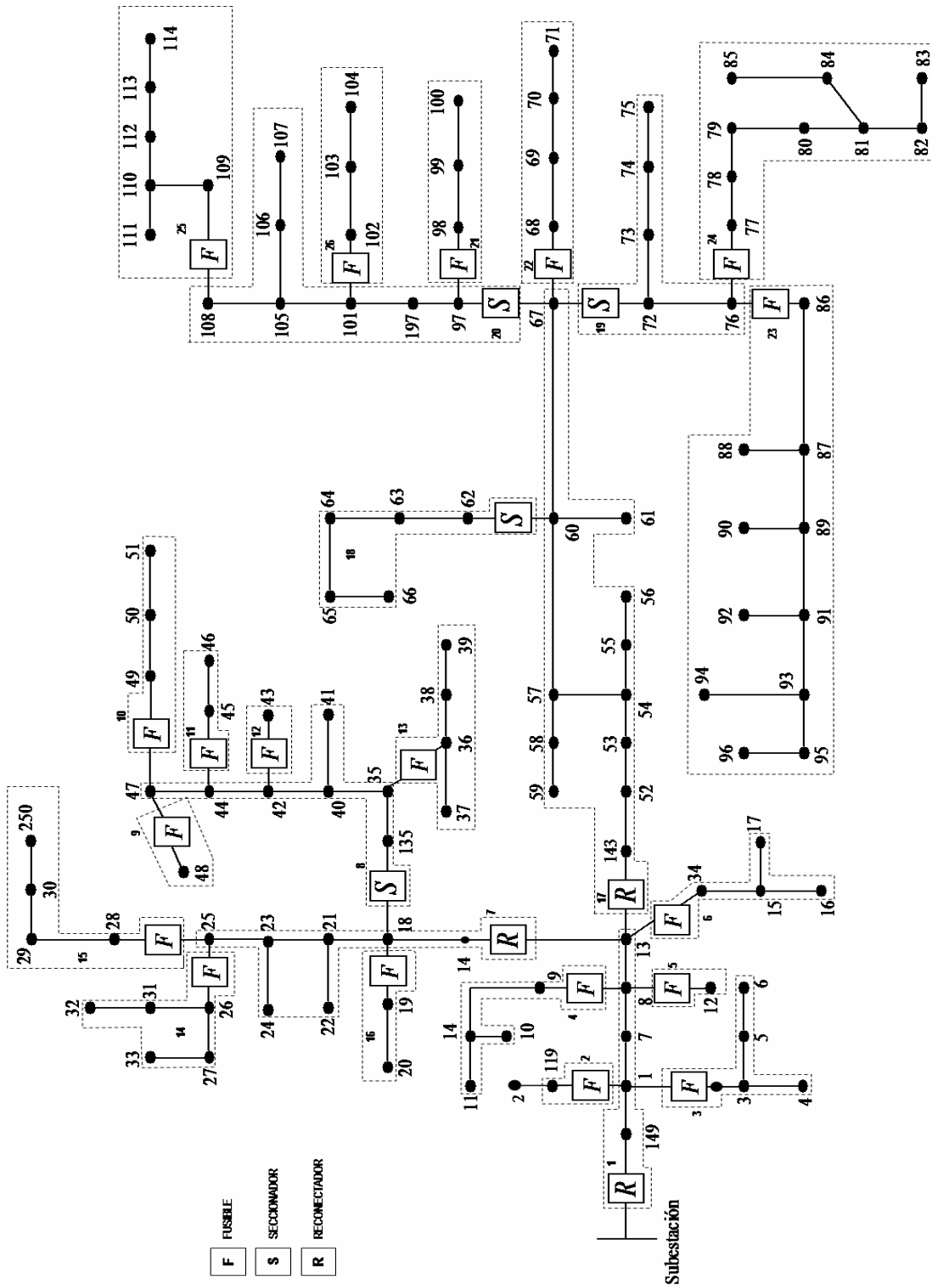


Figura 4.1. Sistema de Distribución para pruebas. Tomado de IEEE 123 Nodes Test System Feeder

4.3 SIMULACIÓN DE FALLAS EN EL SISTEMA DE DISTRIBUCIÓN DE PRUEBAS

El objetivo de la simulación de eventos eléctricos en la red es reproducir el comportamiento del sistema bajo los posibles estados de falla. Mediante la simulación automática de fallas en cada nodo de la red y la caracterización de las señales de tensión-corriente asociadas al evento (medidas en el nodo de subestación) se pueden reconocer los diferentes estados del sistema por medio de una metodología basada en el conocimiento.

Los diferentes tipos de eventos a simular en la red de distribución dependen del tipo de nodo en el cual se de la falla; por ejemplo:

- Para los nodos monofásicos sólo se puede generar una falla fase tierra.
- Para los nodos bifásicos se pueden generar nueve tipos de falla: fase-tierra, falla línea-línea, y falla línea-línea-tierra, todas para cada una de las tres fases.
- Para los nodos trifásicos se pueden generar once tipos de fallas: falla línea-tierra en cada fase, falla línea-línea y línea-línea-tierra por cada par de fases, y falla trifásica y trifásica a tierra.

En resumen, si denominamos las fases del sistema trifásico por A, B y C respectivamente, el conjunto de fallas que pueden ser simuladas son:

Falla Fase A-T.	Falla entre fases A y C
Falla Fase B-T.	Falla entre fases A, B y tierra.
Falla Fase C-T.	Falla entre fases B, C y tierra.
Falla entre fases A y B.	Falla entre fases A, C y tierra.
Falla entre fases B y C.	Falla entre fases A, B y C.
Falla entre fases A, B, C y tierra.	

A cada uno de los eventos están asociadas varias impedancias de falla, estas impedancias son de carácter resistivo [Mora et al (2005a)]. El valor de estas resistencias de falla se ha fijado en los siguientes valores 1 Ω , 5 Ω , 10 Ω , 20 Ω , 25 Ω y 30 Ω para todos los tipos de nodos con excepción de los monofásicos a los cuales se les ha adicionado una resistencia de falla más cuyo valor es de 15 Ω . El valor de resistencia adicional para los nodos monofásicos obedece a que en ellos sólo se puede simular un único tipo de falla mientras que para los nodos trifásicos existen once tipos de fallas posibles; con el valor resistivo adicional se busca aumentar el número de muestras para los nodos monofásicos y posterior uso en el entrenamiento y validación de la metodología basada en el conocimiento.

Para los nodos más alejados de cada protección que define cada una de las zonas del sistema, no es factible simular fallas de alta impedancia dado que los valores de impedancia tanto de las líneas como de la impedancia de falla y nivel de tensión, hace que la corriente medida en el punto donde se encuentra la protección no sea suficiente como para activar la liberación de la misma, aún si se lograra activar, el tiempo de acción sería excesivamente elevado de modo que el tiempo de simulación elevaría excesivamente la exigencia computacional; en los sistemas reales, existen protecciones adicionales especializadas para despejar fallos de alta impedancia. Por este motivo en los nodos 13 y 67 sólo se incluyeron resistencias de falla con un valor hasta de 25 Ω y para el nodo 66 sólo se incluyeron resistencias de falla con un valor hasta de 20 Ω . Con las restricciones descritas, el número total de eventos simulados ha sido de 4399, dentro de los cuales se

Nodos	Tipo de Fallas Simuladas											Fallas Simuladas en cada Nodo
	1	2	3	4	5	6	7	8	9	10	11	
132	7											7
133	5	5	5	5	5	5	5	5	5	5	5	55
134	5	5	5	5	5	5	5	5	5	5	5	55
135	7											7
136			7									7
137			7									7
138	5	5	5	5	5	5	5	5	5	5	5	55
139	4	4	4	4	4	4	4	4	4	4	4	44
140	5	5	5	5	5	5	5	5	5	5	5	55
141	5	5	5	5	5	5	5	5	5	5	5	55
142	5	5	5	5	5	5	5	5	5	5	5	55
143	5	5	5	5	5	5	5	5	5	5	5	55
Total	563	462	532	359	349	364	359	349	364	349	349	4399

Tabla 4.4. Resumen de las Pruebas Realizadas en cada nodo del sistema

4.4 PRUEBAS PARA LA LOCALIZACION DE LA ZONA DE FALLA

La primera etapa de pruebas consiste en utilizar las Máquinas de Soporte Vectorial para Clasificación (SVMc) como herramienta de solución al problema de localización de zona de falla en el sistema de distribución escogido para pruebas.

4.4.1 Características de las pruebas

El sistema de distribución fue dividido en las 26 zonas descritas y definidas por cada protección del circuito de distribución. Las categorías o clases asignadas para las SVMc corresponden con cada una de ellas.

Se implementó la metodología descrita en la sección 3.5 de modo que el máximo número de SVM biclasificadoras que puede poseer alguna de las SVMc multiclase es de 325 para 26 clases ($N(N-1)/2$, bajo el esquema O-v-O).

Para la búsqueda en malla se tomaron como dominios los intervalos $[10^3, 10^5]$ y $[0.5, 5]$ para el parámetro de Penalización C y el parámetro del Kernel RBF σ respectivamente; la rejilla de búsqueda se formó mediante 5 particiones linealmente distribuidas para cada intervalo para un total de 25 puntos de búsqueda.

Para la validación cruzada, se dividió al conjunto de datos de entrenamiento en tres grupos y con ellos se realizó el proceso de sub-entrenamiento (dos de los tres grupos cada vez) y el proceso de sub-validación (el tercer grupo restante cada vez).

Para las pruebas de localización de la zona de falla se implementaron los cinco descriptores asociados al Análisis de los patrones de la Corriente de Falla, que son

- A. Número de conmutaciones presentes en la señal.
- B. Paridad del número de conmutaciones:

- C. Tiempo de recierre de los reconectores
- D. Variación de las corrientes en estado de liberación de falla
- E. Corriente de postfalla luego de la acción de las protecciones

Se hicieron tantas pruebas de localización de fallas como subconjuntos de dos o más descriptores (entre los cinco escogidos) se pueden obtener; en total se realizaron 26 pruebas. Los resultados y el análisis de los mismos se presentan a continuación.

4.4.2 Resultados de pruebas de la localización de la zona de falla

A continuación se resumen los resultados de validación de las 26 pruebas realizadas; la Tabla 4.6 muestra los resultados del proceso de validación o reconocimiento de la localización de la zona de falla para cada uno de las 26 pruebas realizadas.

Se puede observar que de las pruebas realizadas para la localización de fallas, las SVMc que presentan un mejor reconocimiento de la zona son aquellas que involucran los descriptores B, C y E es (*paridad del número de conmutaciones, tiempo para el recierre de reconectores, y variación de la corriente de postfalla luego de la acción de las protecciones*). Este resultado es interesante ya que confirma que el hacer uso de toda la información para la localización de la zona de falla no es lo más apropiado; de hecho, se puede observar que al involucrar toda la información para la localización de la falla presenta un acierto del 91.5% ocupando el onceavo lugar de las 26 pruebas realizadas. Esto permite afirmar que para el proceso de entrenamiento en este tipo de problemas donde se involucran las metodologías basadas en el conocimiento (MBC), en este caso las SVMc, siempre se puede agrupar o “clusterizar” la información ya que en la mayoría de las veces existe información redundante que puede aumentar el esfuerzo computacional en la solución del problema o disminuir la calidad de la solución del problema de optimización asociado con las SVMc.

	Descriptores					% Acierto de Validación
	A	B	C	D	E	
1		■	■		■	98,3
2	■	■	■		■	97,4
3			■		■	96,7
4	■		■		■	96,3
5	■				■	95,3
6	■	■			■	95,1
7		■			■	94,7
8				■	■	92,2
9	■	■		■	■	91,9
10		■	■	■	■	91,7
11	■	■	■	■	■	91,5
12		■		■	■	90,7
13	■			■	■	90,5
14	■	■	■	■		90,5
15		■	■	■		90,4
16	■		■	■	■	90,0
17			■	■	■	89,5

	Descriptorios					% Acierto de Validación
	A	B	C	D	E	
18			■	■		88,8
19	■		■	■		88,4
20	■	■		■		85,3
21	■			■		84,4
22		■		■		82,5
23	■		■			58,9
24	■	■	■			58,7
25		■	■			42,5
26	■	■				29,4

A: Numero de Conmutaciones presentes en la señal

B: Paridad del Numero de Conmutaciones en la señal

C: Tiempo de recierre de los Reconectores

D: Variación de las corrientes en estado de liberación de falla

E: Corriente de Postfalla luego de la acción de las protecciones

Tabla 4.5. Resumen de Resultados de las Pruebas Realizadas

Otro patrón destacado es que para las primeras trece mejores pruebas de las 26 realizadas involucran el descriptor **E**, *Corriente de postfalla luego de la acción de las protecciones*, lo que concluye que la corriente que queda al final del periodo de acción de las protecciones es una variable que discrimina la zona de localización de la falla, pero que por sí sola no es concluyente ya que existen diversas zonas en el sistema de pruebas que poseen cargas de magnitudes similares y que geográficamente son lejanas entre sí. A continuación se detallan los resultados para las cuatro pruebas que arrojaron los mejores resultados en la localización de la zona de falla; se hace uso de las matrices de confusión con el fin de presentar el detalle de los aciertos y de los datos que fueron mal clasificados.

Prueba 1:

Descriptores:

Paridad del Numero de Conmutaciones

Tiempo para el recierre de Reconectores

Variación de la Corriente de Postfalla luego de la acción de las protecciones.

			Datos de Entrenamiento		Datos de Validación		Información sobre la SVMc			
			Numero de Datos	Validación [% Acierto]	Numero de Datos	Validación [% Acierto]	# SVM biclasificación	Número Promedio de Vectores de Soporte	Parámetros	
σ	C									
SVMc para Fallas Monofásicas	Fase A-T	SVMc 1	412	100	563	99.3	190	4	5.0	1.0e5
	Fase B-T	SVMc 2	352	96.3	462	93.5	136	7	0.5	1.0e3
	Fase C-T	SVMc 3	392	95.7	532	92.1	136	9	0.5	5.1e4
SVMc para Fallas Bifásicas	Fases A-B ó A-B-T	SVMc 4	584	100	718	100	91	4	5.0	5.1e4
	Fases B-C ó B-C-T	SVMc 5	568	100	698	100	78	5	2.8	1.0e5
	Fases A-C ó A-C-T	SVMc 6	592	100	728	100	91	4	5.0	5.1e4
SVMc para Fallas Trifásicas	Fases A-B-C ó A-B-C-T	SVMc 7	568	100	698	100	78	4	5.0	1e3

Tabla 4.6. Resultados de la Validación en el reconocimiento de la zona de falla del conjunto de SVMc

		Zonas estimadas por la SVMc																											
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
Zona originales	1	220																											
	2		14																										
	3			35																									
	4				35																								
	5					14																							
	6						35																						
	7							289																					
	8								337																				
	9									108	2																		
	10										220																		
	11											21																	
	12												14																
	13													61															
	14														68	13													
	15														10	265													
	16																21												
	17																	562											
	18																		264										
	19																			1	179	6							
	20																			10	1	333							
	21																						275						
	22																							35					
	23																				15				405				
	24																				10		8			436			
	25																										49		
	26																												28

Acierto de Validación: 98.3

Tabla 4.7. Matriz de Confusión de la validación del entrenamiento de las SVMc para la localización de fallas

Se puede apreciar que 76 datos de fallas, de 4399 datos en total utilizados para la validación del desempeño del localizador de la zona de falla, son mal clasificados; por ejemplo, de los 186 datos disponibles para la zona 19, 179 de ellos fueron clasificados correctamente, uno es clasificado en la zona 18 y seis son clasificados en la zona 20. Esto se debe a que las zonas 18, 19 y 20 tienen tres características comunes que influyen de manera directa en los descriptores y son:

- La protección que define cada zona corresponde con un seccionalizador, de modo que la *Paridad del número de conmutaciones* es el mismo para cada zona.
- Las tres zonas, debido a su cercanía en el sistema de distribución, poseen al reconector R2 como la protección aguas arriba de cada seccionalizador de modo que *el tiempo de recierre de los Reconectores* es el mismo para las tres zonas.
- La carga que es despejada con la acción de los seccionalizadores para las zonas 19 y 20 es de magnitud similar de modo que la *Variación de la Corriente de Postfalla luego de la acción de las protecciones* es similar para estas zonas.

Las características descritas anteriormente son las que provocan la clasificación errónea de la SVMc, pues en el proceso de validación estas SVMc multiclase presentan “empate” en la votación asociada al esquema O-v-O, tal como se explicó en el capítulo 2.

La Tabla 4.8 muestra los datos que fueron mal clasificados. Se muestra tanto el nodo y la resistencia de falla con la cual fue simulada la falla, así como la zona a la que pertenece dicho nodo y la zona en la cual fue mal clasificado. Se puede apreciar que el 75% de los datos mal clasificados corresponden a las fallas simuladas con valores altos para la resistencia de falla.

En particular para las resistencias de falla de 25 y 30 Ω (en nodos monofásicos) y de 30 Ω (en nodos trifásicos), los descriptores no hacen una distinción suficiente entre las zonas como para que la SVMc multiclase clasifique la zona de falla correctamente (a pesar de fijar el hiperplano de separación entre clases se presentan problemas de empate en la votación). Esto se debe a que no existe una diferencia fuertemente marcada entre los descriptores involucrados en esta prueba pues dependen en gran parte del nivel de la corriente de falla; para estos casos dichos niveles de corriente son muy reducidos debido a la alta impedancia vista desde la subestación hasta el sitio de la localización de la falla y la misma resistencia de falla.

Las tablas 4.5 a 4.10 muestran los resultados de validación de la segunda a la cuarta prueba de acuerdo al éxito de la validación. Los resultados son igualmente satisfactorios, y las razones por las cuales algunos de los datos han sido mal clasificados se deben a razones similares a las descritas para la primera prueba.

NODO	Resistencia de Falla *	Zona Adecuada	Zona Errada	NODO	Resistencia de Falla *	Zona Adecuada	Zona Errada
72	6	19	20	28	2	15	14
76	6	19	20	28	3	15	14
108	6	20	19	29	2	15	14
140	6	19	20	29	3	15	14
48	1	9	10	30	2	15	14
72	6	19	20	30	3	15	14
76	6	19	20	31	5	14	15
77	6	24	20	31	6	14	15
78	6	24	20	31	7	14	15
79	6	24	20	32	1	14	15
80	6	24	20	32	5	14	15

NODO	Resistencia de Falla *	Zona Adecuada	Zona Errada	NODO	Resistencia de Falla *	Zona Adecuada	Zona Errada
81	6	24	20	32	6	14	15
82	6	24	20	32	7	14	15
83	6	24	20	75	5	19	18
86	6	23	18	77	6	24	18
90	7	23	18	78	6	24	18
91	6	23	18	79	6	24	18
93	6	23	18	80	6	24	18
95	6	23	18	81	6	24	18
96	7	23	18	82	6	24	18
97	6	20	18	83	6	24	18
101	6	20	18	84	7	24	18
105	6	20	18	85	7	24	18
106	6	20	18	86	6	23	18
106	7	20	18	87	6	23	18
107	6	20	18	89	6	23	18
107	7	20	18	91	6	23	18
108	6	20	18	92	7	23	18
118	6	20	18	93	6	23	18
123	1	9	10	95	6	23	18
133	6	23	18	117	2	15	14
134	6	24	20	117	3	15	14
140	6	19	20	128	4	14	15
141	6	20	18	128	6	14	15
26	4	14	15	129	2	15	14
26	6	14	15	129	3	15	14
27	4	14	15	133	6	23	18
27	6	14	15	134	6	24	18

(*) Resistencias de falla para nodos Trifásicos

1: 1.0 Ω .

2: 5.0 Ω .

3: 10.0 Ω .

4. 20.0 Ω

5. 25.0 Ω

6. 30.0 Ω .

(*) Resistencias de falla para nodos Monofásicos

1: 1.0 Ω

2: 5.0 Ω .

3: 10.0 Ω .

4. 15.0 Ω

5. 20.0 Ω

6. 25.0 Ω .

7. 30.0 Ω

Tabla 4.8. Resumen de los datos mal clasificados para la prueba con mejores resultados.

Prueba 2:

Descriptores:

Número de conmutaciones.

Paridad del número de conmutaciones.

Tiempo para el recierre de reconectores

Variación de la corriente de postfalla luego de la acción de las protecciones.

			Datos de Entrenamiento		Datos de Validación		Información sobre la SVMc			
			Numero de Datos	Validación [% Acierto]	Numero de Datos	Validación [% Acierto]	# SVM biclasificación	Número Promedio de Vectores de Soporte	Parámetros	
									σ	C
SVMc para Fallas Monofásicas	Fase A-T	SVMc 1	412	100	563	99.3	190	5	2.8	5.1e4
	Fase B-T	SVMc 2	352	96.3	462	94.6	136	4	5.0	1.0e3
	Fase C-T	SVMc 3	392	95.7	532	92.7	136	5	5.0	1.0e3
SVMc para Fallas Bifásicas	Fases A-B ó A-B-T	SVMc 4	584	100	718	100	91	4	5.0	5.1e4
	Fases B-C ó B-C-T	SVMc 5	568	100	698	100	78	5	5.0	1.0e5
	Fases A-C ó A-C-T	SVMc 6	592	100	728	100	91	4	5.0	1.0e5
SVMc para Fallas Trifásicas	Fases A-B-C ó A-B-C-T	SVMc 7	568	100	698	93.1	78	13	0.5	1.0e3

Tabla 4.9. Resultados de la Validación en el reconocimiento de la zona de falla del conjunto de SVMc

		Zonas estimadas por la SVMc																										
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
Zonas Originales	1	220																										
	2		14																									
	3			35																								
	4				35																							
	5					14																						
	6						35																					
	7							289																				
	8								325													12						
	9									100	10																	
	10										220																	
	11											21																
	12												14															
	13													61														
	14														80	1												
	15														20	255												
	16																21											
	17																	562										
	18								1											259		4						
	19																				177	9						
	20																				11	1	332					
	21																						275					
	22																							35				
	23																8						14			398		
	24																9						16				429	
	25																										49	
	26																											28

Acierto de Validación: 97.4

Tabla 4.10. Matriz de Confusión de la validación del entrenamiento de las SVMc para la localización de fallas

Prueba 3:

Descriptores:

Tiempo para el recierre de reconectores

Variación de la corriente de postfalla luego de la acción de las protecciones.

			Datos de Entrenamiento		Datos de Validación		Información sobre la SVMc			
			Numero de Datos	Validación [% Acierto]	Numero de Datos	Validación [% Acierto]	# SVM biclasificación	Número Promedio de Vectores de Soporte	Parámetros	
									σ	C
SVMc para Fallas Monofásicas	Fase A-T	SVMc 1	412	99.5	563	97.0	190	5	5.0	1.0e3
	Fase B-T	SVMc 2	352	100	462	93.1	136	7	0.5	5.1e4
	Fase C-T	SVMc 3	392	95.4	532	91.9	136	5	5.0	1.0e5
SVMc para Fallas Bifásicas	Fases A-B ó A-B-T	SVMc 4	584	100	718	92.5	91	11	0.5	5.1e4
	Fases B-C ó B-C-T	SVMc 5	568	100	698	100	78	5	5.0	1.0e3
	Fases A-C ó A-C-T	SVMc 6	592	100	728	100	91	5	2.8	1.0e5
SVMc para Fallas Trifásicas	Fases A-B-C ó A-B-C-T	SVMc 7	568	100	698	100	78	5	2.8	1.0e3

Tabla 4.11. Resultados de la Validación en el reconocimiento de la zona de falla del conjunto de SVMc

		Zonas estimadas por la SVMc																										
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
Zonas Originales	1	220																										
	2		14																									
	3			35																								
	4				35																							
	5					14																						
	6						35																					
	7							289																				
	8								331											6								
	9									108	2																	
	10										220																	
	11											21																
	12												14															
	13													61														
	14														70	11												
	15														15	260												
	16																21											
	17																	562										
	18																			263				1				
	19																			6	174	6						
	20																			22	1	321						
	21																						275					
	22																			5				30				
	23																8			14		9			389			
	24																9			16		8				421		
	25																			7							42	
	26																											28
Acierto de Validación:		96.7																										

Tabla 4.12. Matriz de confusión de la validación del entrenamiento de las SVMc para la localización de fallas

Prueba 4:

Descriptores:

Número de conmutaciones presentes en la señal

Tiempo para el recierre de reconectores

Variación de la corriente de postfalla luego de la acción de las protecciones.

			Datos de Entrenamiento		Datos de Validación		Información sobre la SVMc			
			Numero de Datos	Validación [% Acierto]	Numero de Datos	Validación [% Acierto]	# SVM biclasificación	Número Promedio de Vectores de Soporte	Parámetros	
									σ	C
SVMc para Fallas Monofásicas	Fase A-T	SVMc 1	412	99.8	563	99.1	190	4	5.0	1.0e5
	Fase B-T	SVMc 2	352	100	462	93.5	136	7	0.5	1.0e5
	Fase C-T	SVMc 3	392	95.7	532	92.7	136	5	5.0	1.0e3
SVMc para Fallas Bifásicas	Fases A-B ó A-B-T	SVMc 4	584	100	718	92.5	91	11	0.5	5.1e4
	Fases B-C ó B-C-T	SVMc 5	568	100	698	94.8	78	12	5.0	1.0e5
	Fases A-C ó A-C-T	SVMc 6	592	100	728	100	91	4	2.8	1.0e5
SVMc para Fallas Trifásicas	Fases A-B-C ó A-B-C-T	SVMc 7	568	100	698	100	78	4	2.8	1.0e3

Tabla 4.13. Resultados de la Validación en el reconocimiento de la zona de falla del conjunto de SVMc

		Zonas estimadas por la SVMc																											
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
Zonas Originales	1	220																											
	2		14																										
	3			35																									
	4				35																								
	5					14																							
	6						35																						
	7							289																					
	8								331											6									
	9									108	2																		
	10										220																		
	11											21																	
	12												14																
	13													61															
	14														80	1													
	15														20	255													
	16																21												
	17																	562											
	18								1											263									
	19																			6	168	12							
	20																			22	2	320							
	21																						275						
	22																							35					
	23																8			21		14				377			
	24																9			16		24					405		
	25																											49	
	26																												28
Acierto de validación:		96.3																											

Tabla 4.14. Matriz de confusión para de la validación del entrenamiento de las SVMc para la localización de fallas.

4.5 PRUEBAS PARA LA ESTIMACION DE LA DISTANCIA DE FALLA

La segunda etapa de pruebas consiste en utilizar las Máquinas de Soporte Vectorial para regresión (SVMr) como herramienta para la localización de fallas estimando la distancia desde el nodo de la subestación hasta el sitio de ocurrencia del evento.

Tal como se mostró en el capítulo 3, la metodología planteada para la estimación de la distancia de la localización de la falla hace uso del estimador de la zona de falla, es decir que se hace uso del clasificador expuesto en el numeral 4.4, es decir, se emplean tanto el clasificador del tipo de falla como el clasificador de la zona de falla como datos de entrada para la SVMr que estimará la distancia de falla.

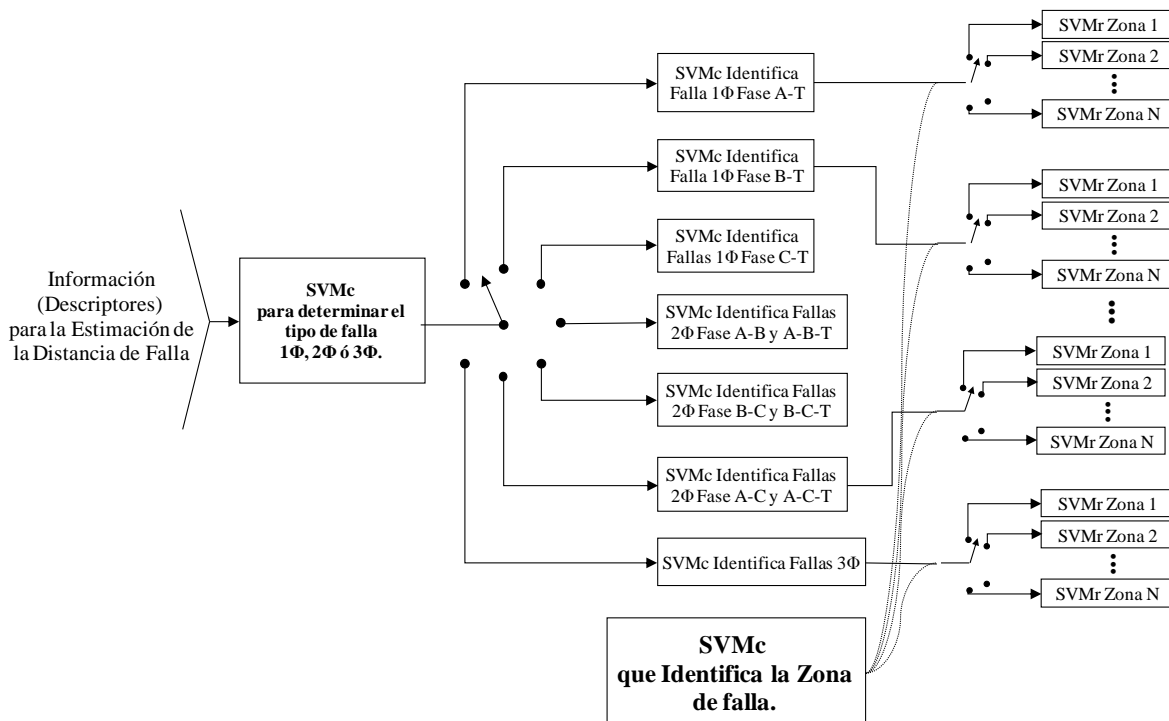


Figura 4.2. Estructura de las SVMr para la estimación de la distancia de Falla.

4.5.1 Características de las pruebas

La distancia con respecto a la subestación a la cual está localizado cada nodo se muestra en [IEEE 123 Node Test] y en la Tabla 4.15; allí también se muestran los resultados de la estimación de la SVMr para regresión.

Se implementó la metodología descrita en sección 3.7 (*metodología híbrida para la localización de fallas en sistemas de distribución*). Para la búsqueda en malla se tomaron como dominios los intervalos $[10^6, 10^8]$ y $[5, 15]$ para el parámetro de Penalización C y el parámetro del Kernel RBF σ respectivamente; la rejilla de búsqueda se formó mediante **tres** particiones linealmente distribuidas para cada intervalo para un total de **nueve** puntos de búsqueda.

Para la validación cruzada, se dividió al conjunto de datos de entrenamiento en **tres** grupos y con ellos se realizó el proceso de sub-entrenamiento (dos de los tres grupos cada vez) y el proceso de sub-validación (el tercer grupo restante cada vez).

Una vez encontrado el punto de la rejilla que producía el menor error, se trazó una rejilla de menor dimensión alrededor de dicho punto y se repetía el proceso de la validación cruzada para cada uno de los nueve puntos de la nueva rejilla. Este proceso se realizó tres veces, es decir, se crearon tres rejillas durante el proceso de entrenamiento y en cada punto de ellas se realizó el proceso de validación cruzada con el fin de sintonizar lo mejor posible los parámetros de la SVMr.

Para las pruebas de estimación de la distancia de falla se escogieron los siguientes descriptores asociados al análisis de la señal transitoria y el análisis de los huecos de Tensión que son:

- A. Caída de tensión en las señales de voltaje de línea en prefalla y postfalla
- B. Caída de tensión en las señales de voltaje de línea tierra en prefalla y postfalla
- C. Duración de la señal transitoria.
- D. Frecuencia de la señal transitoria.

Los descriptores escogidos poseen dos características: la primera es que no están asociados al análisis de los patrones de la corriente de carga presentados en la sección 3.5.5 (firma del sistema) y la segunda es que representan variables continuas, a diferencia de los descriptores usados en la sección 4.4 para estimar la zona de falla que emplean variables discretas. Esto permite poseer un dominio mucho más amplio en los parámetros de entrada que si se empleasen variables discretas.

También se aclara que aunque para el entrenamiento de las SVMr no se usa de manera explícita la información de los descriptores asociada con el análisis de los *patrones de la corriente de carga* (firma del sistema), la estructura híbrida propuesta para la estimación de la distancia de falla emplea la información de la SVMc que estima la zona de falla para luego direccionar la información a una SVMr en especial; así que en la estimación de la distancia de falla se emplea implícitamente la información de los *patrones de corriente de carga* y el éxito de dicha estimación se encuentra asociado con el éxito de la localización de la zona de falla.

Se realizó una prueba con cada descriptor, es decir que se realizaron un total de cuatro pruebas. En el desarrollo de este trabajo se realizaron pruebas que combinaron parejas de descriptores (con los cuatro descriptores presentados se pueden construir hasta 15 pruebas); sin embargo, los tiempos de cómputo para este proceso de entrenamiento fueron absurdamente extendidos y no se logró obtener resultados; esto se debe a que cada descriptor posee un dato asociado a cada una de las fases del sistema, es decir que cada descriptor pertenece al espacio tridimensional (cada muestra de cada descriptor pertenece a \mathbf{R}^3); al combinar dos o más descriptores se crea un espacio de entrada de dimensión seis o mayor lo cual dificulta el proceso de entrenamiento y si a esto sumamos el hecho de que el problema de optimización de las SVMr comprende un matriz de $2N \times 2N$ (donde N corresponde con el número de datos de entrenamiento) se puede concluir que los tiempos de cómputo se extienden hasta el punto en el que es mucho más práctico emplear un único descriptor por prueba. Los resultados obtenidos en este trabajo muestran que mediante el uso de un solo descriptor se obtienen una correcta estimación de la distancia de falla.

4.5.2 Resultados de pruebas de la estimación de la zona de falla

- Prueba 1

Los resultados de las pruebas de la estimación de la distancia de falla por medio del descriptor asociado al **Hueco en las Señales de Voltaje de Línea** se presentan a continuación.

La primera columna muestra todos los nodos del sistema; en cada uno de ellos se simularon todos los tipos de fallas que se presentaron en el capítulo 3, de modo que para un mismo nodo se

poseen varios datos para la estimación de la distancia de falla. La segunda columna muestra la distancia a la que está ubicado cada nodo con respecto a la cabecera del sistema o subestación principal del sistema de distribución. La tercera y cuarta columna muestran el promedio y la desviación estándar respectivamente, de las distancias estimadas por la metodología propuesta, para cada dato de falla disponible en cada uno de los nodos.

Nodo	Distancia Nodo – Subestación [ft]	Distancia Estimada [ft]	Desviación Estándar de Estimación [ft]	Nodo	Distancia Nodo – Subestación [ft]	Distancia Estimada [ft]	Desviación Estándar de Estimación [ft]
1	400	664	180.3	37	3350	3027	336.4
2	575	525	78.4	38	3300	3237	100.8
3	650	601	100.4	39	3625	3563	81.3
4	850	911	124.1	40	2650	2708	189.2
5	975	996	27.7	41	2975	3228	113.8
6	1225	1093	68.3	42	2900	2793	202.7
7	700	735	109.2	43	3400	3460	118.9
8	900	797	121.4	44	3100	2835	225.7
9	1125	1537	189.2	45	3300	3275	23.6
10	1800	1434	325.3	46	3600	3468	187.7
11	1800	1435	325.7	47	3350	2956	221.4
12	1125	1097	32.9	48	3500	3474	42.0
13	1200	970	182.5	49	3600	3674	102.1
14	1550	1521	62.0	50	3850	3774	124.2
15	1450	1494	113.7	51	4100	3810	182.0
16	1825	1723	88.0	52	1600	2146	371.3
17	1800	1676	90.1	53	1800	2258	307.5
18	2025	2195	235.0	54	1925	2183	261.8
19	2275	2275	20.1	55	2200	2296	225.4
20	2600	2606	9.1	56	2475	2254	282.0
21	2325	2213	263.6	57	2275	2257	342.0
22	2850	2728	147.9	58	2525	2530	35.6
23	2575	2449	221.0	59	2775	2641	165.1
24	3125	2756	256.4	60	3025	2763	261.5
25	2850	2564	257.6	61	3575	2773	578.5
26	3200	3227	85.7	62	3275	3442	219.8
27	3475	3355	145.2	63	3450	3562	234.5
28	3050	3327	250.1	64	3800	3767	153.2
29	3350	3333	175.5	65	4225	3937	355.8
30	3700	3528	238.7	66	4550	4013	364.0
31	3425	3129	147.7	67	3375	2906	330.2
32	3725	3667	65.4	68	3575	3724	101.4
33	3975	3956	150.4	69	3850	3719	104.6
34	1350	1591	167.4	70	4175	4102	51.6
35	2400	2640	221.8	71	4450	4410	79.4

Nodo	Distancia Nodo – Subestación [ft]	Distancia Estimada [ft]	Desviación Estándar de Estimación [ft]	Nodo	Distancia Nodo – Subestación [ft]	Distancia Estimada [ft]	Desviación Estándar de Estimación [ft]
36	3050	3025	113.0	72	3650	3678	106.4
73	3925	4033	81.6	109	4925	5317	288.3
74	4275	4295	44.1	110	5225	5277	209.3
75	4675	4591	110.6	111	5800	5811	11.2
76	3850	3755	115.3	112	5350	5449	180.2
77	4250	4501	285.8	113	5875	5889	28.0
78	4350	4547	211.1	114	6200	5776	303.6
79	4575	4625	249.3	116	5550	4892	336.0
80	4825	4826	266.5	117	3900	3577	236.9
81	5300	4950	354.4	118	5475	4601	535.0
82	5550	5020	385.0	119	400	391	14.7
83	5800	5279	556.4	120	400	731	208.0
84	5975	5751	165.8	121	900	915	15.3
85	6450	5453	387.1	122	900	934	36.8
86	4550	5158	411.3	123	3350	3380	46.1
87	5000	5163	277.4	124	3350	3619	163.8
88	5175	5172	9.2	125	3100	3107	35.2
89	5275	5284	178.2	126	2900	3250	352.1
90	5500	5521	78.8	127	2400	2691	380.2
91	5500	5404	161.3	128	2850	3091	231.7
92	5800	5928	95.9	129	2850	3158	237.0
93	5725	5496	195.9	130	2025	2049	30.4
94	6000	5691	216.7	131	3800	4016	350.7
95	6025	5542	338.0	132	3375	3356	27.1
96	6225	5566	381.1	133	3850	4666	626.0
97	3625	3821	256.5	134	3850	4296	371.9
98	3900	4191	396.8	135	4475	5315	497.5
99	4450	4543	232.4	136	3875	4410	564.1
100	4750	4696	219.2	137	1200	1502	161.9
101	3875	4070	289.1	138	2025	2461	272.4
102	4100	4222	269.2	139	3025	3490	347.8
103	4425	4589	181.2	140	3375	3460	126.4
104	5125	5147	42.6	141	3375	3685	311.8
105	4150	4139	221.6	142	1200	1858	442.1
106	4375	4570	238.4	143	1200	1979	510.9
107	4950	5085	150.1				
108	4475	4290	273.1				

Tabla 4.15. Resultados de la estimación de la distancia de falla para el descriptor de caída de tensión en la señal de voltaje de línea.

A continuación, las figura 4.3 a la figura 4.28 presentan la distribución de la distancia estimada de falla con la estructura de SVMr para el descriptor del **Hueco en las Señales de Voltaje de Línea** en cada uno de los nodos de cada zona del sistema de distribución.

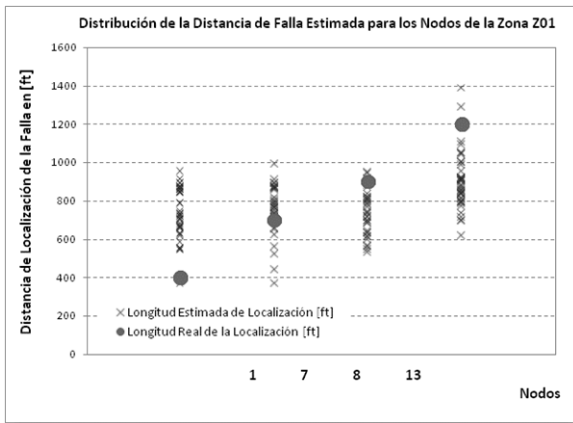


Figura 4.3 Distribución de la falla estimada en la Zona 1

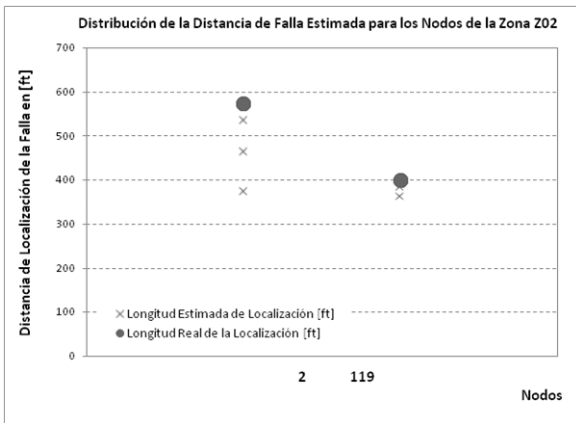


Figura 4.4 Distribución de la falla estimada en la Zona 2

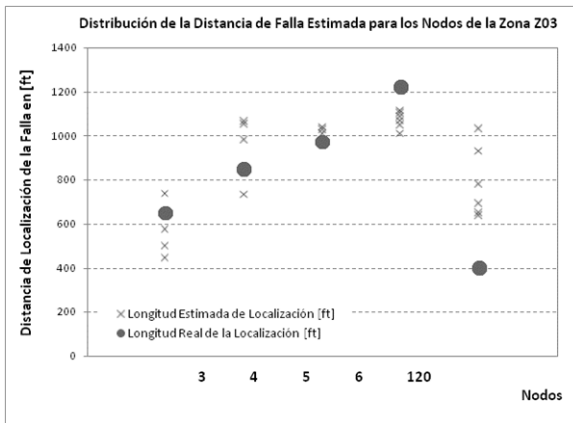


Figura 4.5 Distribución de la falla estimada en la Zona 3

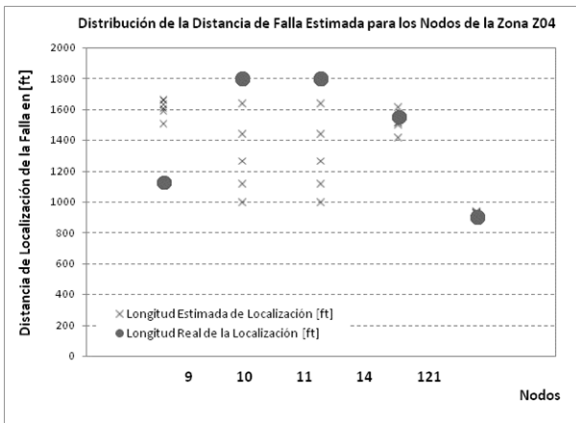


Figura 4.6 Distribución de la falla estimada en la Zona 4

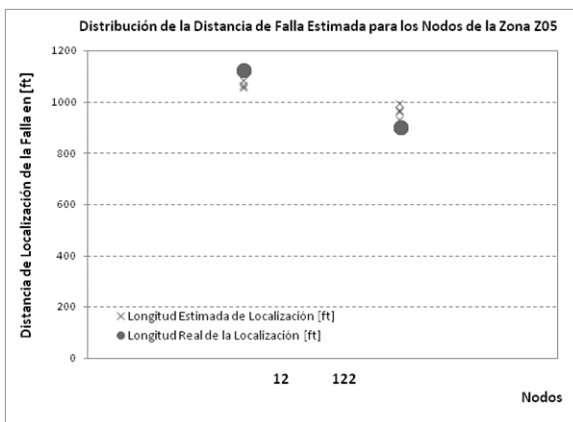


Figura 4.7 Distribución de la falla estimada en la Zona 5

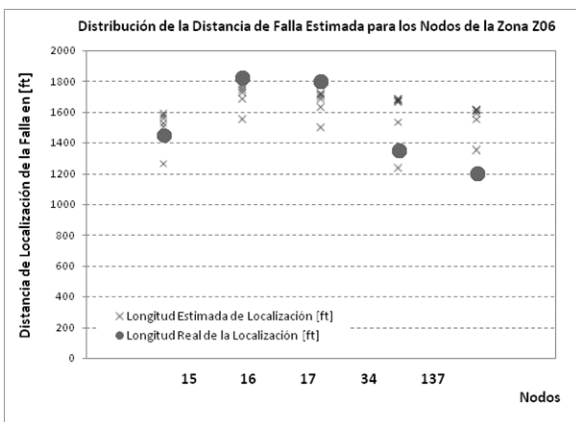


Figura 4.8 Distribución de la falla estimada en la Zona 6

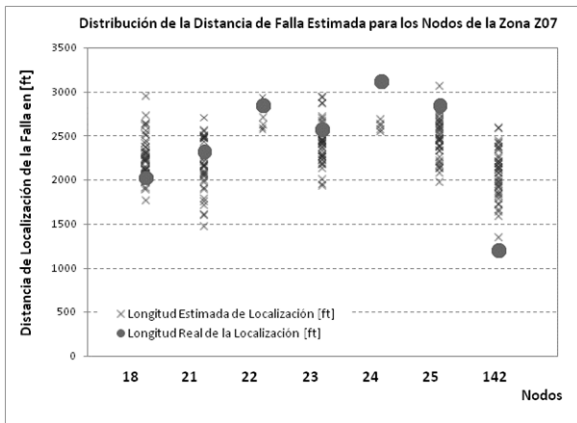


Figura 4.9 Distribución de la falla estimada en la Zona 7

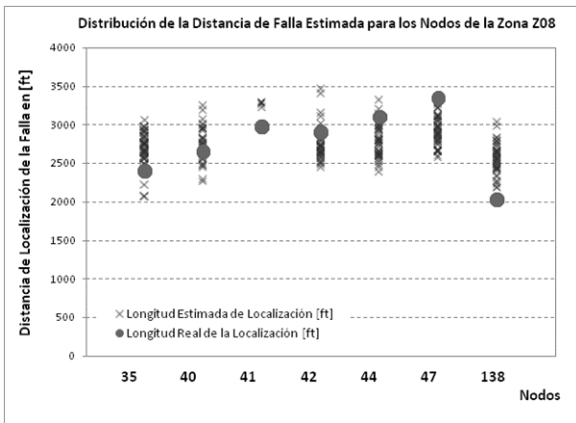


Figura 4.10 Distribución de la falla estimada en la Zona 8

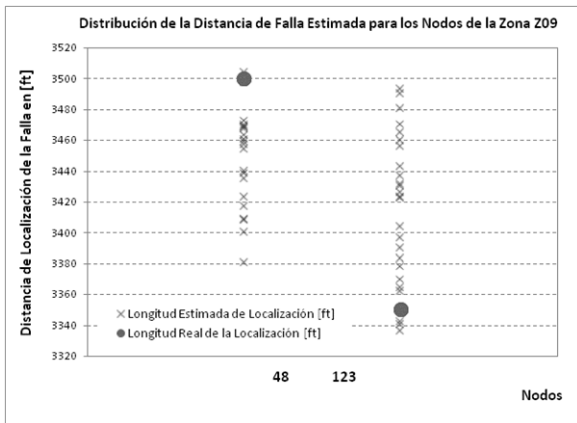


Figura 4.11 Distribución de la falla estimada en la Zona 9

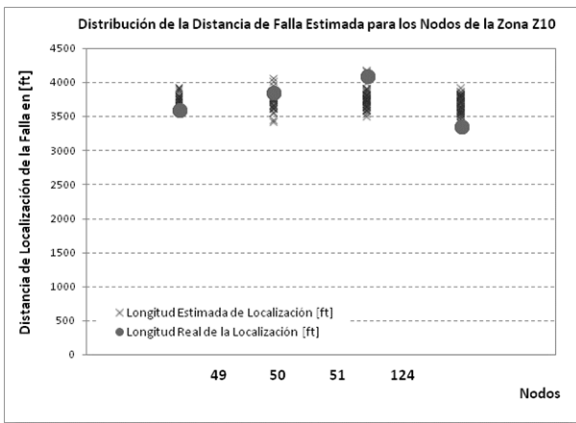


Figura 4.12 Distribución de la falla estimada en la Zona 10

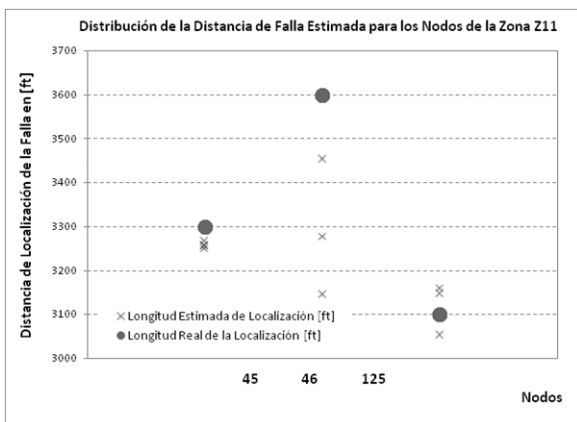


Figura 4.13 Distribución de la falla estimada en la Zona 11

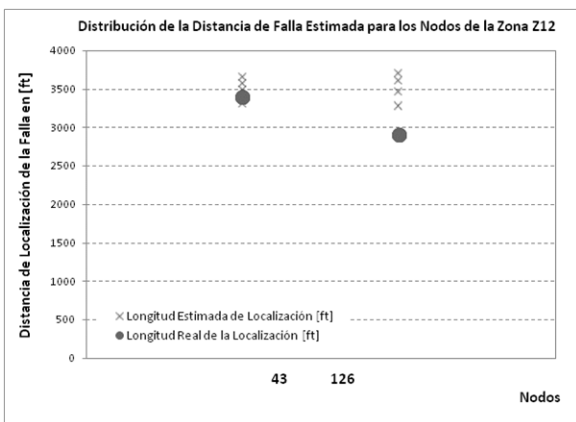


Figura 4.14 Distribución de la falla estimada en la Zona 12

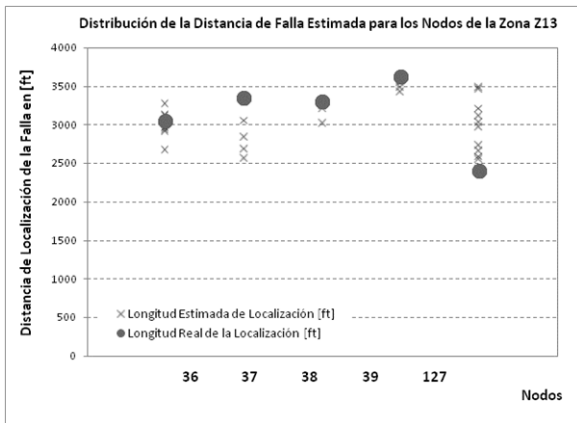


Figura 4.15 Distribución de la falla estimada en la Zona 13

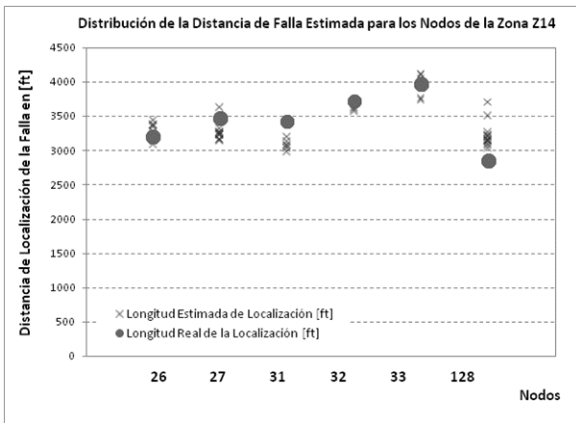


Figura 4.16 Distribución de la falla estimada en la Zona 14

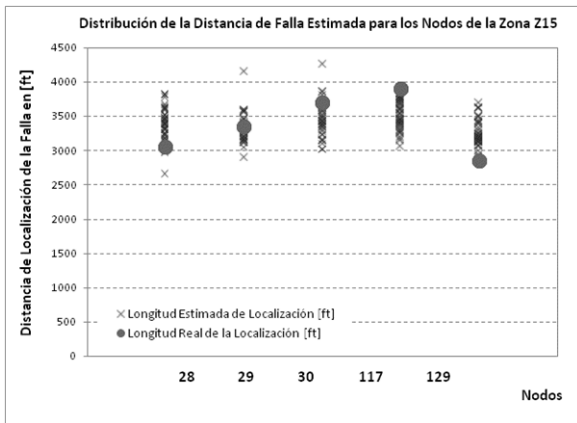


Figura 4.17 Distribución de la falla estimada en la Zona 15

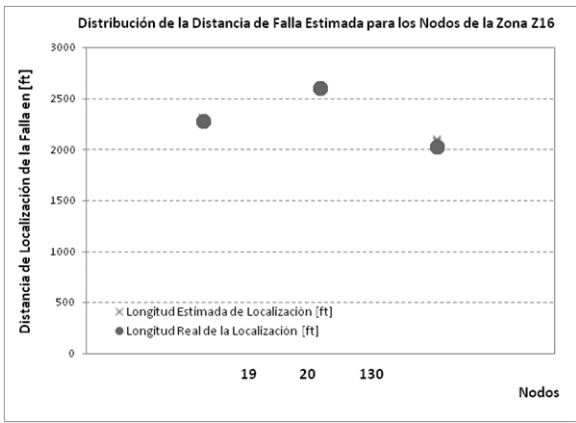


Figura 4.18 Distribución de la falla estimada en la Zona 16

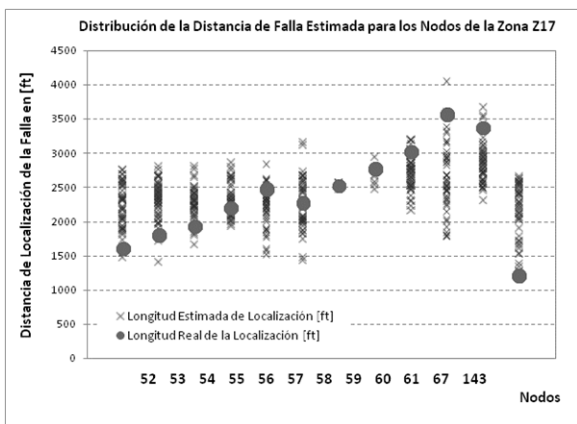


Figura 4.19 Distribución de la falla estimada en la Zona 17

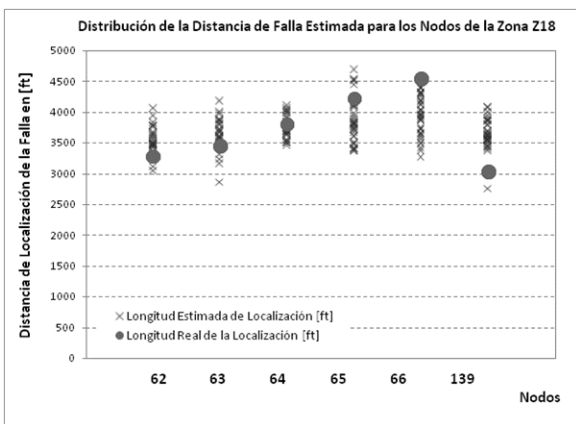


Figura 4.20 Distribución de la falla estimada en la Zona 18

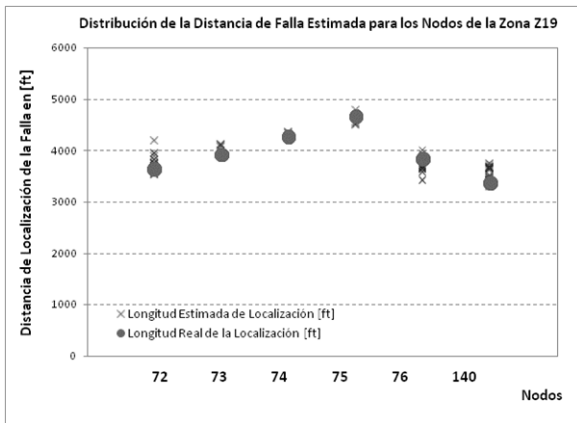


Figura 4.21 Distribución de la falla estimada en la Zona 19

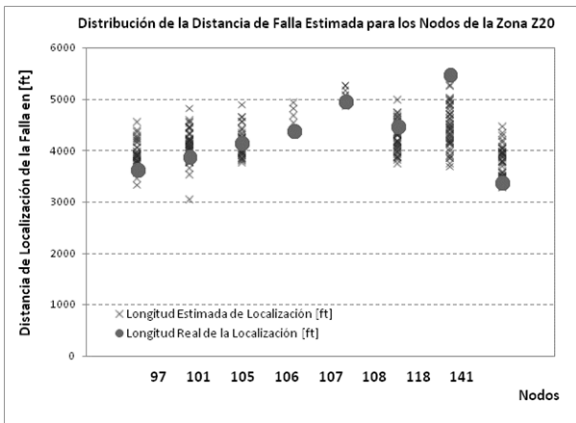


Figura 4.22 Distribución de la falla estimada en la Zona 20

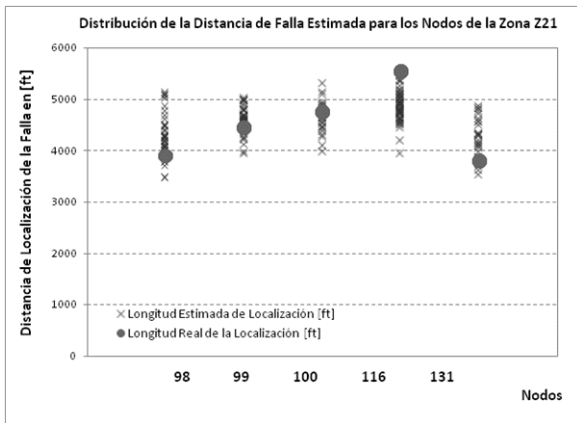


Figura 4.23 Distribución de la falla estimada en la Zona 21

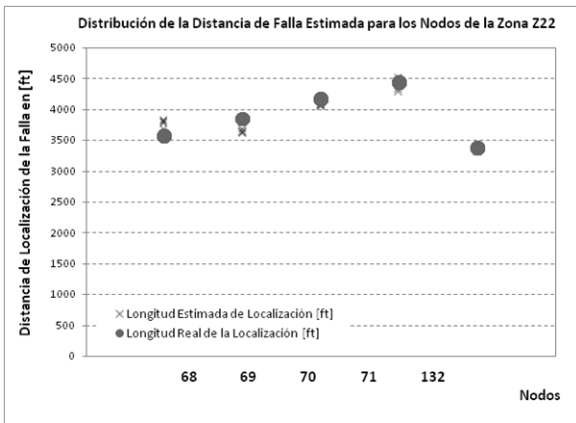


Figura 4.24 Distribución de la falla estimada en la Zona 22

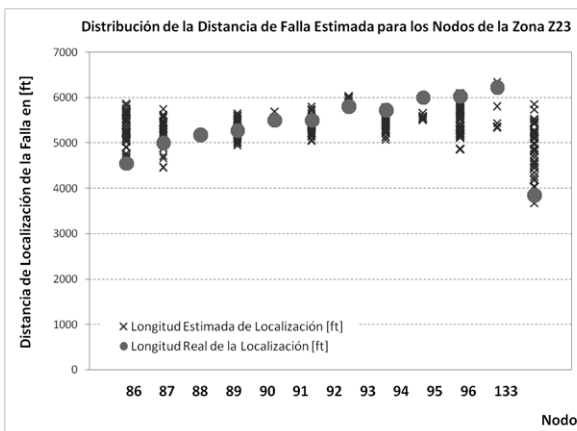


Figura 4.25 Distribución de la falla estimada en la Zona 23

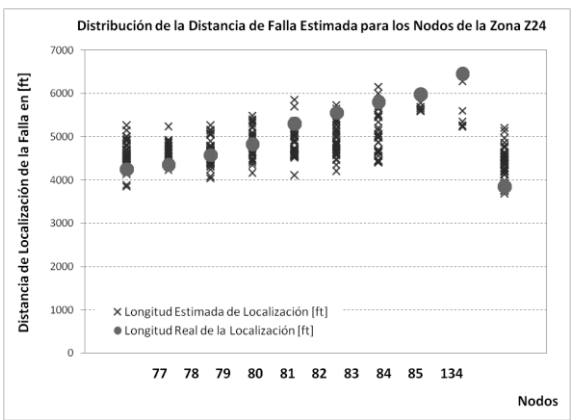


Figura 4.26 Distribución de la falla estimada en la Zona 24

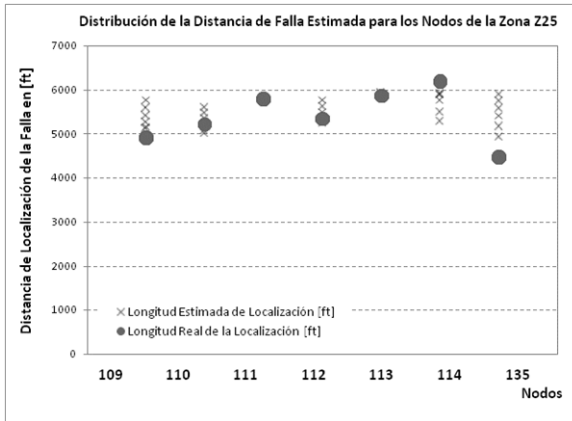


Figura 4.27 Distribución de la falla estimada en la Zona 25

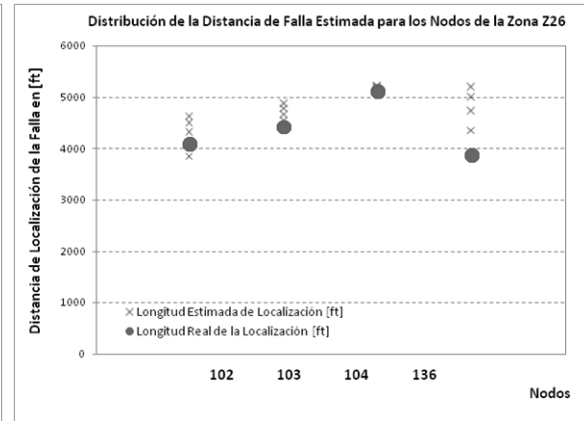


Figura 4.28 Distribución de la falla estimada en la Zona 26

De los resultados obtenidos se destacan los siguientes aspectos.

Para los nodos más cercanos a la subestación la estimación de la distancia presenta una desviación estándar alta; esto se debe a que la magnitud del hueco de tensión en estos nodos (descriptor empleado para las primeras pruebas) está fuertemente ligada al valor de la resistencia de falla, la cual predomina sobre la impedancia del sistema de distribución.

Como ejemplo, se tomó el nodo inicial de la red y se desarrolló el equivalente trifásico del sistema de distribución; se modeló la carga como impedancia constante y mediante un proceso iterativo que inicia en los nodos finales, calculando la impedancia equivalente de la carga y de la línea que conecta dicha carga hasta alcanzar el nodo inicial, se encontró la matriz de impedancia equivalente:

$$\begin{bmatrix} Z_{aa} & Z_{ac} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} = \begin{bmatrix} 3.7376 + 1.9483j & 0.0422 + 0.1638j & 0.0734 + 0.4050j \\ 0.0422 + 0.1638j & 5.9555 + 2.0860j & 0.0376 + 0.1519j \\ 0.0734 + 0.4050j & 0.0376 + 0.1519j & 4.6552 + 2.1288j \end{bmatrix} [\Omega]$$

Las magnitudes de las componentes de la matriz equivalente son:

$$\begin{bmatrix} |Z_{aa}| & |Z_{ab}| & |Z_{ac}| \\ |Z_{ba}| & |Z_{bb}| & |Z_{bc}| \\ |Z_{ca}| & |Z_{cb}| & |Z_{cc}| \end{bmatrix} = \begin{bmatrix} 4.2149 & 0.1691 & 0.4116 \\ 0.1691 & 6.3103 & 0.1564 \\ 0.4116 & 0.1564 & 5.1188 \end{bmatrix} [\Omega]$$

Se puede observar que del rango de valores escogido (desde 1Ω hasta 30Ω) para la resistencia de falla durante las simulaciones existen algunos valores inferiores y comparables con la impedancia equivalente del circuito y existen otros valores superiores. El comportamiento del descriptor empleado para esta primera prueba (caída de tensión en las señales de voltaje de línea en prefalla y postfalla) puede asemejarse al comportamiento de un divisor de voltaje tal como se presentan en la figura 4.29.

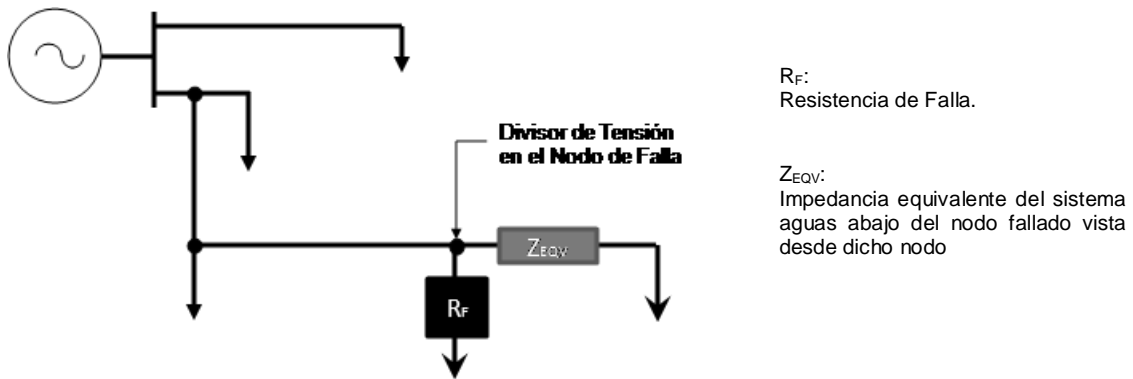


Figura 4.29. Representación del sistema de distribución para falla en un nodo del sistema. Se destaca el equivalente del sistema en el nodo de falla y el divisor de tensión equivalente

De esta forma, para los casos en los que la resistencia de falla es inferior o comparable con la impedancia equivalente de la Red (vista en el nodo de falla), el descriptor “Caída de Tensión” permite realizar una correcta estimación de la distancia de falla; esto sucede porque existen valores fuertemente diferenciados entre sí, asociados con cada valor de la resistencia de falla. Por otra parte, para los casos en los que la resistencia de falla es superior a la impedancia equivalente del sistema, el descriptor en cuestión no presenta variaciones significativas con cada valor de la resistencia de falla dificultando el proceso de estimación de la distancia de falla.

A continuación la Tabla 4.16 presenta para cada nodo del sistema de distribución, el promedio de las impedancias propias $((|Z_{aa}|+ |Z_{bb}| + |Z_{cc}|)/3)$ y de las impedancias mutuas $((|Z_{ab}|+ |Z_{bc}| + |Z_{ac}|)/3)$; también se expone el error medio cuadrático de la distancia de falla estimada por la estructura de SVMr agrupado por cada valor de la resistencia de falla. Se observa en la mayoría de los nodos, que para valores de la resistencia de falla inferiores a la impedancia equivalente, el error cuadrático medio es inferior al error obtenido con valores más altos de la resistencia de falla.

Nodo	$\frac{Z_{aa} + Z_{bb} + Z_{cc}}{3}$	$\frac{Z_{ab} + Z_{bc} + Z_{ac}}{3}$	R_1 1Ω	R_2 5Ω	R_3 10Ω	R_4 15Ω	R_5 20Ω	R_6 25Ω	R_7 30Ω
1	5.0549	0.1182	17%	68%	84%	82%	80%		
2	85.9930	-	0%	0%	0%	0%	7%	19%	35%
3	17.2201	-	0%	0%	13%	0%	12%	23%	31%
4	42.9963	-	0%	0%	14%	0%	15%	24%	25%
5	28.6770	-	0%	0%	0%	0%	3%	5%	6%
6	42.9964	-	0%	18%	9%	10%	11%	13%	14%
7	5.2739	0.1047	8%	8%	11%	11%	16%		
8	5.2654	0.0883	5%	11%	14%	14%	14%		
9	17.2217	-	0%	34%	42%	48%	47%	45%	41%
10	85.9933	-	0%	0%	9%	20%	30%	38%	45%
11	42.9964	-	0%	0%	9%	20%	30%	38%	45%
12	85.9932	-	0%	7%	0%	0%	1%	4%	6%
13	5.3908	0.0688	8%	22%	20%	23%	27%		
14	28.6805	-	9%	0%	4%	0%	3%	3%	2%
15	28.6881	-	4%	13%	0%	5%	7%	9%	10%
16	42.9965	-	0%	15%	8%	5%	4%	4%	3%

Nodo	$\frac{Z_{aa} + Z_{bb} + Z_{cc}}{3}$	$\frac{Z_{ab} + Z_{bc} + Z_{ac}}{3}$	R ₁ 1Ω	R ₂ 5Ω	R ₃ 10Ω	R ₄ 15Ω	R ₅ 20Ω	R ₆ 25Ω	R ₇ 30Ω
17	85.9937	-	1%	17%	9%	6%	5%	5%	4%
18	14.1599	0.0470	8%	12%	9%	8%		10%	
19	21.5075	-	0%	0%	1%	0%	2%	0%	1%
20	42.9965	-	0%	0%	0%	0%	0%	0%	1%
21	71.8124	0.0193	4%	9%	9%	9%		10%	
22	42.9967	-	0%	3%	0%	5%	8%	9%	10%
23	224,925.5716	4.9731	4%	8%	7%	8%		10%	
24	42.9967	-	0%	0%	14%	16%	16%	18%	19%
25	274,646.8119	5.8901	3%	5%	11%	15%		17%	
26	86.1025	0.0005	0%	2%	3%	2%		1%	
27	2,475,672.7641	8.7842	1%	2%	3%	6%		8%	
28	343,328.0218	10.2165	5%	10%	9%	11%		14%	
29	465,943.4411	14.8450	2%	4%	3%	3%		4%	
30	1,726,908.5894	176,761.4906	3%	6%	7%	6%		7%	
31	43.0055	-	0%	13%	7%	9%	10%	11%	11%
32	85.9935	-	0%	3%	0%	0%	1%	3%	4%
33	42.9966	-	0%	5%	6%	0%	2%	3%	3%
34	17.2112	-	9%	13%	24%	24%	24%	24%	25%
35	20.0163	0.0558	7%	13%	12%	12%		12%	
36	86.0670	0.0004	0%	0%	5%	0%		5%	
37	42.9964	-	0%	0%	0%	9%	15%	20%	23%
38	43.0063	-	0%	8%	0%	0%	1%	2%	2%
39	85.9936	-	0%	0%	0%	0%	2%	4%	6%
40	22.9667	0.0531	4%	11%	5%	1%		3%	
41	85.9936	-	0%	8%	10%	10%	11%	10%	11%
42	23.7559	0.0346	2%	3%	7%	8%		8%	
43	42.9966	-	0%	0%	3%	0%	2%	5%	7%
44	25.9445	0.0241	4%	9%	9%	11%		12%	
45	43.0054	-	0%	0%	2%	0%	1%	1%	1%
46	85.9935	-	0%	0%	0%	0%	4%	9%	13%
47	27.7453	0.0056	7%	9%	12%	15%		16%	
48	67.0580	0.0001	0%	1%	0%	1%		2%	
49	73.8405	0.0019	1%	2%	3%	2%		3%	
50	784,845.9916	29.9169	3%	2%	2%	3%		3%	
51	5,364,935.1269	843,096.7619	6%	7%	7%	8%		8%	
52	8.9231	0.1211	16%	29%	32%	46%	51%		
53	9.0519	0.1082	12%	23%	25%	35%	37%		
54	9.2078	0.1017	6%	12%	9%	21%	26%		
55	790,717.5824	29.6868	4%	9%	9%	7%	9%		
56	4,701,440.1444	296,474.3429	6%	12%	16%	10%	8%		
57	9.3854	0.0759	7%	12%	12%	12%	13%		
58	43.0040	-	1%	0%	2%	0%	1%	1%	

Nodo	$\frac{Z_{aa} + Z_{bb} + Z_{cc}}{3}$	$\frac{Z_{ab} + Z_{bc} + Z_{ac}}{3}$	R ₁ 1Ω	R ₂ 5Ω	R ₃ 10Ω	R ₄ 15Ω	R ₅ 20Ω	R ₆ 25Ω	R ₇ 30Ω
59	85.9933	-	6%	9%	11%	7%	4%	4%	
60	9.5766	0.0178	5%	11%	13%	10%	9%		
61	3,871,632.3356	1,051,602.2002	3%	19%	29%	32%	32%		
62	46.5587	0.0500	2%	8%	6%	6%			
63	48.9430	0.0404	2%	7%	6%	5%			
64	71.6427	0.0253	2%	2%	3%	5%			
65	100.9081	0.0000	5%	7%	9%	13%			
66	322,251.5159	-	7%	11%	14%	16%			
67	12.3173	0.0212	6%	8%	17%	21%	21%		
68	14.3665	-	0%	1%	3%	5%	6%	7%	7%
69	17.2169	-	0%	0%	2%	5%	6%	6%	5%
70	28.6783	-	0%	2%	0%	1%	2%	3%	3%
71	42.9964	-	0%	0%	1%	0%	1%	3%	4%
72	19.0933	0.0579	1%	2%	2%	1%		2%	
73	14.3548	-	0%	0%	2%	4%	4%	4%	5%
74	21.5097	-	0%	0%	1%	0%	1%	2%	2%
75	42.9965	-	0%	2%	1%	3%	3%	4%	4%
76	21.1946	0.0722	2%	4%	3%	2%		3%	
77	31.9158	0.1454	7%	9%	6%	6%		6%	
78	31.6243	0.1346	2%	6%	4%	5%		6%	
79	5,756,268.8820	589,152.9529	3%	5%	5%	3%		3%	
80	31.3985	0.0917	4%	5%	3%	3%		5%	
81	30.6843	0.0470	3%	7%	7%	9%		10%	
82	30.0527	0.0231	3%	9%	11%	13%		14%	
83	29.2930	0.0000	2%	10%	11%	12%		12%	
84	28.6887	-	0%	0%	3%	5%	5%	6%	7%
85	42.9966	-	3%	13%	17%	18%	19%	19%	19%
86	75.9327	0.0563	4%	18%	16%	15%		15%	
87	78.2966	0.0232	3%	6%	5%	5%		6%	
88	38.4567	-	0%	0%	0%	0%	0%	0%	0%
89	105.3746	0.0197	2%	3%	3%	2%		3%	
90	38.4566	-	0%	3%	1%	0%	0%	1%	1%
91	124.5346	0.0078	3%	2%	3%	2%		3%	
92	38.4565	-	0%	0%	2%	3%	3%	4%	4%
93	713,465.5549	21.5563	1%	4%	4%	5%		6%	
94	42.9964	-	0%	0%	6%	7%	7%	8%	8%
95	4,470,752.0108	702,568.0164	3%	7%	9%	10%		10%	
96	85.9931	-	2%	7%	13%	14%	14%	14%	14%
97	36.3365	0.0191	3%	8%	6%	6%		9%	
98	129.0812	0.0232	5%	6%	9%	10%		14%	
99	237,249.1893	14.8478	2%	5%	3%	3%		6%	
100	680,499.4801	42,928.3276	1%	5%	4%	2%		2%	

Nodo	$\frac{Z_{aa} + Z_{bb} + Z_{cc}}{3}$	$\frac{Z_{ab} + Z_{bc} + Z_{ac}}{3}$	R ₁ 1Ω	R ₂ 5Ω	R ₃ 10Ω	R ₄ 15Ω	R ₅ 20Ω	R ₆ 25Ω	R ₇ 30Ω
101	51.2230	0.0082	7%	5%	7%	7%		8%	
102	17.2350	-	0%	0%	7%	0%	5%	9%	13%
103	21.5183	-	0%	10%	0%	0%	3%	5%	8%
104	42.9969	-	0%	0%	1%	0%	1%	1%	2%
105	201,243.1181	9.0113	3%	4%	3%	3%		5%	
106	21.5147	-	0%	1%	1%	4%	7%	10%	13%
107	42.9967	-	6%	2%	2%	0%	2%	4%	6%
108	576,877.6924	90,656.9358	3%	5%	6%	7%		7%	
109	12.3198	-	4%	0%	4%	7%	10%	14%	17%
110	17.2335	-	0%	4%	3%	0%	2%	5%	7%
111	85.9947	-	0%	0%	0%	0%	0%	0%	0%
112	21.5342	-	0%	0%	2%	0%	2%	5%	8%
113	28.6687	-	0%	0%	1%	0%	0%	0%	1%
114	85.9936	-	0%	15%	11%	7%	5%	5%	4%

Tabla 4.16. Comparación del error en la estimación de la distancia de falla en función de la resistencia de falla y la impedancia equivalente del sistema vista en el nodo de falla

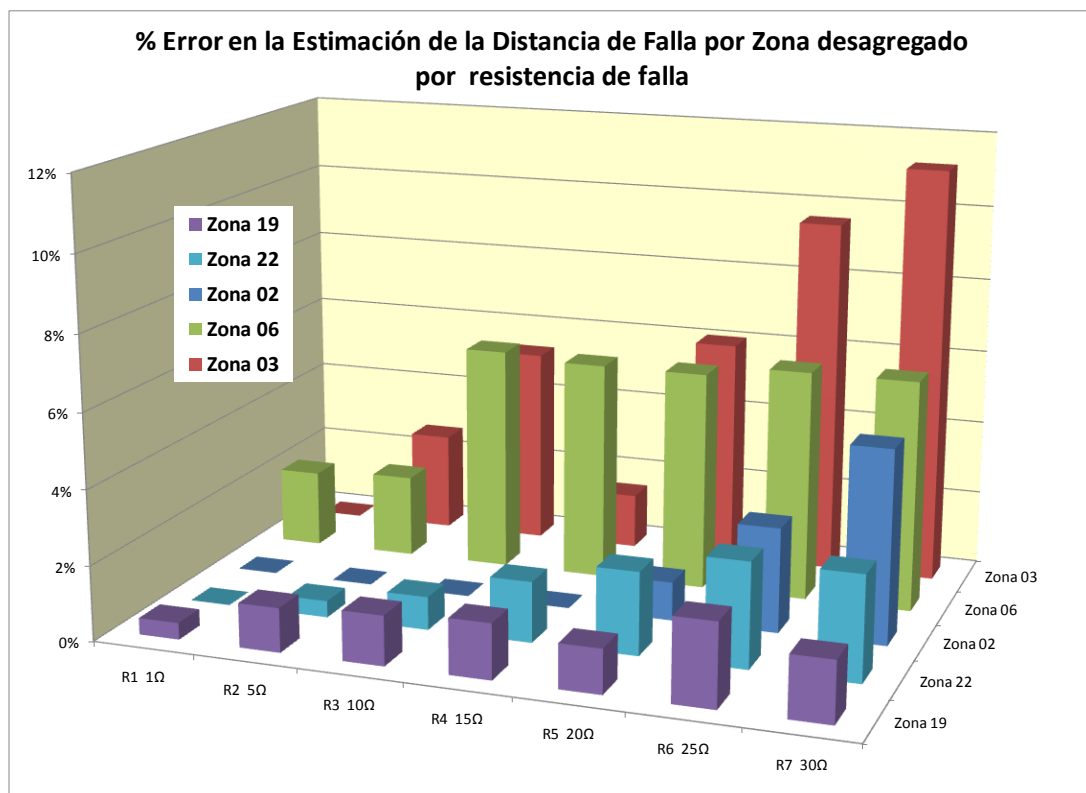


Figura 4.2 Distribución del error en la estimación de la distancia de falla para las Zonas 19, 22, 2,6 y 3 en función de la resistencia de falla.

Para la mayoría de los nodos monofásicos, la estimación de la distancia de falla es satisfactoria. Esto se debe a que en los nodos monofásicos, el número de datos de entrenamiento y validación es reducido; la varianza de la magnitud del hueco de tensión es bastante reducida comparada con la varianza del mismo descriptor en nodos trifásicos; adicionalmente en la estructura propuesta (de acuerdo al tipo de falla y la zona de falla) la estimación de la distancia está establecida por zonas del sistema, de modo que una misma SVMr, sólo manipula información de una misma zona, facilitándose así el proceso de aprendizaje de cada una de ellas; sin embargo, como se ha creado un gran número de SVMr especializadas por cada zona y por cada tipo de falla, se presenta el inconveniente del número máximo de SVMr ya que durante el procesamiento de la información (entrenamiento y validación) se debe almacenar continuamente en memoria la información correspondiente a los parámetros de cada SVMr lo cual podría traducirse en dificultades para sistemas de distribución mucho más grandes donde el número de zonas es superior.

La decisión de emplear SVMr especialización por zonas y por tipo de falla para cada una de las SVMr fue llevada a cabo dado que los resultados obtenidos con un número menor de SVMr que involucraran datos de distintas zonas fueron bastante alejados de la realidad debido a la diversidad de la información; adicionalmente el proceso de entrenamiento demanda mayor tiempo y esfuerzo computacional ya que, como se muestra en el capítulo 2, el problema de optimización que se resuelve en el proceso de entrenamiento de una SVMr involucra una matriz (matriz Hessiana de la función objetivo) de $2N \times 2N$ donde N representa el número de datos de entrenamiento, de modo que entre mayor sea el número de datos mayor será el esfuerzo computacional.

A continuación se muestran los parámetros finales σ y C , con los que cada SVMr asociada al tipo de falla y a la zona del sistema, estimó la distancia de localización de las fallas.

ZONA	Tipo de Falla	SVM #	Parámetros		ZONA	Tipo de Falla	SVM #	Parámetros	
			σ	C				σ	C
1	Fase A-T	SVMr 1	5.00	5.1E+07	15	Fases A-B-C ó A-B-C-T	SVMr 7	10.00	5.1E+07
1	Fase B-T	SVMr 2	3.33	1.0E+06	16	Fase A-T	SVMr 1	3.33	1.0E+06
1	Fase C-T	SVMr 3	3.33	1.0E+06	17	Fase A-T	SVMr 1	3.33	1.0E+06
1	Fases A-B ó A-B-T	SVMr 4	3.33	1.0E+08	17	Fase B-T	SVMr 2	3.33	1.0E+06
1	Fases B-C ó B-C-T	SVMr 5	3.33	5.1E+07	17	Fase C-T	SVMr 3	3.33	1.0E+06
1	Fases A-C ó A-C-T	SVMr 6	5.00	5.1E+07	17	Fases A-B ó A-B-T	SVMr 4	11.67	1.0E+08
1	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	1.0E+06	17	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08
2	Fase B-T	SVMr 2	3.33	1.0E+08	17	Fases A-C ó A-C-T	SVMr 6	3.33	5.1E+07
3	Fase C-T	SVMr 3	3.33	5.1E+07	17	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	1.0E+08

ZONA	Tipo de Falla	SVM #	Parámetros		ZONA	Tipo de Falla	SVM #	Parámetros	
			σ	C				σ	C
4	Fase A-T	SVMr 1	5.00	1.0E+06	18	Fase A-T	SVMr 1	3.33	5.1E+07
5	Fase B-T	SVMr 2	8.33	1.0E+08	18	Fase B-T	SVMr 2	3.33	1.0E+06
6	Fase C-T	SVMr 3	16.67	1.0E+06	18	Fase C-T	SVMr 3	3.33	5.1E+07
7	Fase A-T	SVMr 1	3.33	1.0E+06	18	Fases A-B ó A-B-T	SVMr 4	10.00	1.0E+08
7	Fase B-T	SVMr 2	5.00	1.0E+06	18	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08
7	Fase C-T	SVMr 3	3.33	1.0E+06	18	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+08
7	Fases A-B ó A-B-T	SVMr 4	11.67	1.0E+06	18	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	1.0E+06
7	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08	19	Fase A-T	SVMr 1	3.33	5.1E+07
7	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+08	19	Fase B-T	SVMr 2	5.00	1.0E+06
7	Fases A-B-C ó A-B-C-T	SVMr 7	6.67	1.0E+08	19	Fase C-T	SVMr 3	3.33	1.0E+08
8	Fase A-T	SVMr 1	3.33	5.1E+07	19	Fases A-B ó A-B-T	SVMr 4	3.33	5.1E+07
8	Fase B-T	SVMr 2	3.33	5.1E+07	19	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08
8	Fase C-T	SVMr 3	3.33	1.0E+08	19	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+08
8	Fases A-B ó A-B-T	SVMr 4	8.33	1.0E+08	19	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	1.0E+08
8	Fases B-C ó B-C-T	SVMr 5	5.00	5.1E+07	20	Fase A-T	SVMr 1	3.33	5.1E+07
8	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+06	20	Fase B-T	SVMr 2	3.33	1.0E+06
8	Fases A-B-C ó A-B-C-T	SVMr 7	8.33	5.1E+07	20	Fase C-T	SVMr 3	3.33	1.0E+06
9	Fase A-T	SVMr 1	3.33	5.1E+07	20	Fases A-B ó A-B-T	SVMr 4	3.33	5.1E+07
9	Fase B-T	SVMr 2	3.33	1.0E+08	20	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08
9	Fase C-T	SVMr 3	3.33	1.0E+08	20	Fases A-C ó A-C-T	SVMr 6	6.67	1.0E+08
9	Fases A-B ó A-B-T	SVMr 4	3.33	1.0E+06	20	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	1.0E+08
ZONA	Tipo de	SVM #	Parámetros		ZONA	Tipo de	SVM #	Parámetros	

	Falla		σ	C		Falla		σ	C
9	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+06	21	Fase A-T	SVMr 1	3.33	5.1E+07
9	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+06	21	Fase B-T	SVMr 2	3.33	5.1E+07
9	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	1.0E+06	21	Fase C-T	SVMr 3	3.33	1.0E+08
10	Fase A-T	SVMr 1	10.00	5.1E+07	21	Fases A-B ó A-B-T	SVMr 4	3.33	1.0E+06
10	Fase B-T	SVMr 2	3.33	5.1E+07	21	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08
10	Fase C-T	SVMr 3	3.33	1.0E+08	21	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+08
10	Fases A-B ó A-B-T	SVMr 4	10.00	1.0E+06	21	Fases A-B-C ó A-B-C-T	SVMr 7	3.33	5.1E+07
10	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+08	22	Fase A-T	SVMr 1	5.00	1.0E+06
10	Fases A-C ó A-C-T	SVMr 6	5.00	1.0E+06	23	Fase A-T	SVMr 1	3.33	1.0E+08
10	Fases A-B-C ó A-B-C-T	SVMr 7	13.33	1.0E+08	23	Fase B-T	SVMr 2	3.33	1.0E+06
11	Fase A-T	SVMr 1	3.33	1.0E+06	23	Fase C-T	SVMr 3	3.33	1.0E+08
12	Fase B-T	SVMr 2	3.33	1.0E+06	23	Fases A-B ó A-B-T	SVMr 4	3.33	1.0E+06
13	Fase A-T	SVMr 1	3.33	5.1E+07	23	Fases B-C ó B-C-T	SVMr 5	3.33	5.1E+07
13	Fase B-T	SVMr 2	3.33	1.0E+06	23	Fases A-C ó A-C-T	SVMr 6	6.67	1.0E+06
13	Fases A-B ó A-B-T	SVMr 4	3.33	1.0E+08	23	Fases A-B-C ó A-B-C-T	SVMr 7	5.00	1.0E+06
14	Fase A-T	SVMr 1	5.00	5.1E+07	24	Fase A-T	SVMr 1	3.33	5.1E+07
14	Fase C-T	SVMr 3	3.33	5.1E+07	24	Fase B-T	SVMr 2	3.33	5.1E+07
14	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+08	24	Fase C-T	SVMr 3	3.33	5.1E+07
15	Fase A-T	SVMr 1	11.67	1.0E+08	24	Fases A-B ó A-B-T	SVMr 4	3.33	1.0E+08
15	Fase B-T	SVMr 2	3.33	1.0E+06	24	Fases B-C ó B-C-T	SVMr 5	3.33	1.0E+06
15	Fase C-T	SVMr 3	3.33	1.0E+06	24	Fases A-C ó A-C-T	SVMr 6	3.33	5.1E+07
15	Fases A-B ó A-B-T	SVMr 4	11.67	1.0E+06	24	Fases A-B-C ó A-B-C-T	SVMr 7	5.00	1.0E+08

ZONA	Tipo de Falla	SVM #	Parámetros		ZONA	Tipo de Falla	SVM #	Parámetros	
			σ	C				σ	C
15	Fases B-C ó B-C-T	SVMr 5	3.33	5.1E+07	24	Fase A-T	SVMr 1	3.33	5.1E+07
15	Fases A-C ó A-C-T	SVMr 6	3.33	1.0E+08	26	Fase C-T	SVMr 3	3.33	1.0E+06

Tabla 4.17. Parámetros σ y C de cada SVMr para la estimación de la Distancia de Falla

- Prueba 2

Los resultados de las pruebas de la estimación de la distancia de falla por medio del descriptor asociado al **hueco en las señales de voltaje de línea tierra** se presentan a continuación.

Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación	Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación
1	400	605	191.6	37	3350	3357	47.0
2	575	576	1.3	38	3300	3222	55.3
3	650	629	29.6	39	3625	3632	8.8
4	850	789	101.4	40	2650	2694	207.7
5	975	892	89.0	41	2975	3067	70.2
6	1225	817	190.0	42	2900	2807	163.1
7	700	739	132.7	43	3400	3286	121.5
8	900	837	135.4	44	3100	2849	242.4
9	1125	1371	239.1	45	3300	3305	8.1
10	1800	1690	222.3	46	3600	3595	15.4
11	1800	1689	222.2	47	3350	3003	250.3
12	1125	1146	40.2	48	3500	3473	44.9
13	1200	1013	164.3	49	3600	3662	148.5
14	1550	1791	171.8	50	3850	3798	121.5
15	1450	1389	82.8	51	4100	3878	187.9
16	1825	1847	56.6	52	1600	2015	431.7
17	1800	1746	88.1	53	1800	2031	247.9
18	2025	2129	284.2	54	1925	2106	374.1
19	2275	2305	24.0	55	2200	2231	284.8
20	2600	2488	78.8	56	2475	2468	358.6
21	2325	2371	219.7	57	2275	2270	273.2
22	2850	2481	295.5	58	2525	2234	190.7
23	2575	2487	229.8	59	2775	2841	74.0
24	3125	2492	281.8	60	3025	2636	299.0
25	2850	2626	188.5	61	3575	3001	447.4
26	3200	3193	141.4	62	3275	3570	272.4
27	3475	3413	135.7	63	3450	3573	194.9
28	3050	3322	271.1	64	3800	3748	184.5

Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación	Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación
29	3350	3400	191.3	65	4225	3907	252.3
30	3700	3565	249.2	66	4550	3959	361.6
31	3425	3441	35.0	67	3375	2603	482.6
32	3725	3564	117.1	68	3575	3733	117.1
33	3975	3910	72.0	69	3850	3905	128.1
34	1350	1439	67.1	70	4175	4202	47.7
35	2400	2631	234.2	71	4450	4377	51.8
36	3050	3017	114.8	72	3650	3691	89.4
73	3925	3955	53.5	109	4925	5108	169.2
74	4275	4317	48.7	110	5225	5516	99.5
75	4675	4552	85.1	111	5800	5307	89.8
76	3850	3768	98.9	112	5350	5418	115.6
77	4250	4542	246.3	113	5875	5419	35.7
78	4350	4597	233.1	114	6200	5526	223.3
79	4575	4695	254.0	116	5550	4986	422.8
80	4825	4783	144.0	117	3900	3524	243.3
81	5300	4850	259.1	118	5475	4789	529.9
82	5550	4941	357.3	119	400	407	10.7
83	5800	5649	394.3	120	400	662	128.6
84	5975	5740	166.7	121	900	1342	295.7
85	6450	5317	407.3	122	900	972	71.0
86	4550	5047	351.6	123	3350	3362	25.7
87	5000	5103	206.5	124	3350	3504	177.7
88	5175	5166	23.6	125	3100	3155	62.2
89	5275	5223	153.3	126	2900	2939	52.0
90	5500	5586	259.2	127	2400	2597	302.0
91	5500	5427	223.6	128	2850	3011	163.0
92	5800	5821	99.9	129	2850	3102	248.9
93	5725	5576	226.0	130	2025	2140	107.6
94	6000	5788	103.3	131	3800	4084	362.5
95	6025	5626	271.4	132	3375	3508	143.5
96	6225	5422	374.6	133	3850	4358	555.0
97	3625	3904	295.3	134	3850	4299	378.2
98	3900	4237	320.7	135	4475	5301	373.8
99	4450	4547	222.6	136	3875	4066	181.6
100	4750	4694	264.4	137	1200	1351	154.4
101	3875	4235	355.9	138	2025	2322	296.7
102	4100	4226	216.9	139	3025	3378	366.7
103	4425	4520	108.2	140	3375	3442	96.6
104	5125	5139	34.4	141	3375	3596	312.3
105	4150	4184	250.9	142	1200	1834	523.3
106	4375	4540	291.9	143	1200	1989	505.0
107	4950	4989	56.7				
108	4475	4226	307.6				

Tabla 4.18. Resultados de la estimación de la distancia de falla para el descriptor de caída de tensión en la señal de voltaje de fase.

- Prueba 3

Los resultados de las pruebas de la estimación de la distancia de falla por medio del descriptor asociado a la Duración de la señal transitoria debido a la falla se presentan a continuación.

Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación	Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación
1	400	559	157.9	37	3350	3050	1.1
2	575	548	43.6	38	3300	3211	0.0
3	650	739	142.6	39	3625	3211	0.0
4	850	800	161.2	40	2650	2759	188.6
5	975	1002	22.9	41	2975	2893	268.3
6	1225	1062	52.8	42	2900	2883	145.8
7	700	764	103.6	43	3400	2900	0.0
8	900	869	134.9	44	3100	2886	172.3
9	1125	1182	73.7	45	3300	3300	0.0
10	1800	1799	2.9	46	3600	3300	0.0
11	1800	1799	2.9	47	3350	2928	106.4
12	1125	1101	10.7	48	3500	3422	65.2
13	1200	968	195.5	49	3600	3714	110.8
14	1550	1380	128.6	50	3850	3737	116.4
15	1450	1499	103.2	51	4100	3729	106.8
16	1825	1709	6.3	52	1600	2194	337.0
17	1800	1740	85.9	53	1800	2253	335.5
18	2025	2275	246.4	54	1925	2212	275.5
19	2275	2360	43.3	55	2200	2238	170.0
20	2600	2493	140.0	56	2475	2340	310.5
21	2325	2421	145.3	57	2275	2326	148.6
22	2850	2325	0.0	58	2525	2383	224.5
23	2575	2447	138.6	59	2775	2475	0.0
24	3125	2775	198.7	60	3025	2401	268.2
25	2850	2497	155.2	61	3575	2567	493.7
26	3200	3277	107.0	62	3275	3528	299.2
27	3475	3266	134.7	63	3450	3544	230.9
28	3050	3324	84.9	64	3800	3717	239.3
29	3350	3344	36.3	65	4225	3815	302.0
30	3700	3379	108.2	66	4550	3947	369.0
31	3425	3425	0.0	67	3375	2584	417.4
32	3725	3428	8.0	68	3575	3574	6.0
33	3975	3298	103.9	69	3850	3846	11.1
34	1350	1646	202.7	70	4175	4270	307.3
35	2400	2718	213.0	71	4450	4450	0.0
36	3050	2930	261.1	72	3650	3623	292.8
73	3925	4118	207.8	109	4925	5350	0.0

Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación	Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación
74	4275	4279	1.8	110	5225	5350	0.0
75	4675	4279	1.8	111	5800	5350	0.0
76	3850	3735	138.4	112	5350	5350	0.0
77	4250	4492	311.7	113	5875	5350	0.0
78	4350	4549	315.4	114	6200	5350	0.0
79	4575	4817	417.1	116	5550	4893	428.8
80	4825	4921	376.2	117	3900	3376	105.1
81	5300	5121	346.8	118	5475	4699	517.2
82	5550	5226	396.8	119	400	442	64.7
83	5800	5435	495.9	120	400	676	194.9
84	5975	5611	160.7	121	900	974	133.4
85	6450	5640	237.4	122	900	1045	99.3
86	4550	4918	447.3	123	3350	3403	56.3
87	5000	5258	392.3	124	3350	3670	132.1
88	5175	5500	0.0	125	3100	3271	75.6
89	5275	5392	249.0	126	2900	2900	0.0
90	5500	5500	0.0	127	2400	2727	323.6
91	5500	5495	198.5	128	2850	3237	144.8
92	5800	5504	10.3	129	2850	3280	178.8
93	5725	5546	198.1	130	2025	2187	191.1
94	6000	5500	0.0	131	3800	4095	383.6
95	6025	5593	235.3	132	3375	3427	95.7
96	6225	5500	0.0	133	3850	4288	622.3
97	3625	3865	374.0	134	3850	4242	420.6
98	3900	4208	324.0	135	4475	5350	0.0
99	4450	4526	184.0	136	3875	4279	249.9
100	4750	4650	257.3	137	1200	1154	248.2
101	3875	4055	325.6	138	2025	2509	382.3
102	4100	4358	177.1	139	3025	3402	275.3
103	4425	4426	3.1	140	3375	3502	158.8
104	5125	4426	3.1	141	3375	3784	355.5
105	4150	4223	300.5	142	1200	1576	515.9
106	4375	4150	0.0	143	1200	1957	486.8
107	4950	4150	0.0				
108	4475	4330	208.6				

Tabla 4.19. Resultados de la estimación de la distancia de falla para el descriptor de duración de la señal transitoria.

- *Prueba 4*

Los resultados de las pruebas de la estimación de la distancia de falla por medio del descriptor asociado a la frecuencia de oscilación de la señal transitoria debido a la falla se presentan a continuación.

Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación	Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación
1	400	440	111.0	37	3350	3050	0.1
2	575	432	39.9	38	3300	3050	0.0
3	650	683	56.4	39	3625	3050	0.0
4	850	954	96.4	40	2650	2745	190.3
5	975	975	0.0	41	2975	2888	322.8
6	1225	1011	237.6	42	2900	2844	166.2
7	700	711	96.9	43	3400	2900	0.0
8	900	877	125.2	44	3100	2859	173.5
9	1125	1055	55.3	45	3300	3300	0.0
10	1800	1648	403.0	46	3600	3300	0.0
11	1800	1648	403.0	47	3350	2922	150.6
12	1125	1125	0.0	48	3500	3336	440.1
13	1200	1100	123.9	49	3600	3487	1126.8
14	1550	1610	236.6	50	3850	3540	1130.8
15	1450	1565	3.8	51	4100	3679	916.9
16	1825	1605	86.3	52	1600	2005	250.2
17	1800	1605	85.9	53	1800	2081	263.4
18	2025	2210	264.3	54	1925	2117	233.2
19	2275	2275	0.0	55	2200	2216	142.5
20	2600	2462	171.6	56	2475	2284	184.3
21	2325	2382	388.8	57	2275	2221	207.5
22	2850	2325	0.0	58	2525	2275	0.0
23	2575	2519	167.6	59	2775	2275	0.0
24	3125	3125	0.0	60	3025	2479	460.2
25	2850	2631	198.6	61	3575	2848	655.9
26	3200	3224	120.5	62	3275	3507	240.5
27	3475	3247	128.9	63	3450	3576	241.9
28	3050	3322	93.5	64	3800	3765	247.0
29	3350	3342	55.2	65	4225	3941	275.4
30	3700	3374	129.7	66	4550	3884	354.4
31	3425	3425	0.0	67	3375	2617	540.2
32	3725	3454	76.1	68	3575	3542	41.0
33	3975	3200	0.0	69	3850	3850	0.2
34	1350	1476	137.7	70	4175	4281	289.7
35	2400	2660	214.6	71	4450	4450	0.4
36	3050	2861	214.3	72	3650	3653	528.7
73	3925	4095	234.6	109	4925	5350	0.1
74	4275	4275	0.0	110	5225	5350	0.1
75	4675	4275	0.0	111	5800	5350	0.1
76	3850	3834	663.1	112	5350	5350	0.1
77	4250	4459	296.5	113	5875	5350	0.1
78	4350	4495	314.2	114	6200	5350	0.0
79	4575	4794	428.0	116	5550	4920	425.1
80	4825	4873	414.8	117	3900	3373	115.7

Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación	Nodo	Distancia Nodo - Subestación	Distancia Estimada	Desviación Estándar de Estimación
81	5300	5164	394.4	118	5475	4858	568.1
82	5550	5337	404.2	119	400	411	28.3
83	5800	5448	461.0	120	400	386	28.1
84	5975	5549	3.0	121	900	944	74.7
85	6450	5549	3.0	122	900	1061	109.7
86	4550	4912	521.2	123	3350	3307	435.0
87	5000	5141	465.2	124	3350	3469	989.2
88	5175	5500	0.7	125	3100	3271	75.6
89	5275	5466	213.9	126	2900	2900	0.0
90	5500	5500	0.0	127	2400	2735	263.3
91	5500	5572	157.6	128	2850	3174	161.0
92	5800	5500	0.0	129	2850	3267	167.8
93	5725	5622	165.2	130	2025	1972	66.1
94	6000	5500	0.7	131	3800	3988	277.4
95	6025	5677	194.4	132	3375	3403	74.4
96	6225	5500	0.0	133	3850	4259	677.0
97	3625	3842	286.1	134	3850	4203	380.3
98	3900	4180	346.8	135	4475	5350	0.1
99	4450	4541	183.2	136	3875	4268	268.3
100	4750	4688	269.1	137	1200	1234	149.4
101	3875	4008	274.1	138	2025	2406	332.3
102	4100	4351	194.8	139	3025	3440	250.6
103	4425	4425	0.0	140	3375	3617	478.6
104	5125	4425	0.0	141	3375	3713	298.3
105	4150	4172	233.8	142	1200	1379	471.5
106	4375	4150	0.0	143	1200	1651	394.8
107	4950	4150	0.0				
108	4475	4308	315.4				

Tabla 4.20. Resultados de la estimación de la distancia de falla para el descriptor de frecuencia de la señal transitoria.

5. CONCLUSIONES Y TRABAJOS FUTUROS

5.1 CONCLUSIONES GENERALES

La calidad de la energía es y seguirá siendo un tema de amplio interés para los operadores de red, no sólo a nivel de distribución sino también a nivel de transmisión de energía. La importancia actual de estos aspectos se encuentra ligada, y más aún en caso colombiano, a la alta participación del capital privado en los sistemas eléctricos y las exigencias regulatorias que demandan una mejor prestación del servicio y menores costos de operación. Para el caso colombiano, la remuneración de la mayoría de los activos de niveles de tensión III y IV se da a través de un mecanismo de remuneración único directamente proporcional con el tiempo de disponibilidad de cada activo de la red, de tal forma que la remuneración de los activos puede verse afectada dependiendo de la duración de los eventos de indisponibilidad de los activos. De allí la importancia y el interés que tienen los operadores de red en mantener en servicios sus activos y diagnosticar de manera efectiva el estado de las redes.

Los índices de continuidad del servicio se encuentran directamente asociados con el tiempo de localización de fallas y aunque a nivel de transmisión el problema de localización de fallas se encuentra casi resuelto, gracias a la implementación de protecciones en los extremos de cada línea, a nivel de distribución el problema se ha tornado mucho más complejo debido al problema de la múltiple estimación asociado con la alta ramificación de las redes, el reducido número de elementos de medición y la escasa información que poseen los mismos operadores sobre sus redes en estado de falla.

El problema de la localización de fallas a nivel de sistemas de distribución no es nuevo, es posible encontrar aplicaciones implementadas inicialmente para las redes de transmisión pero con poca efectividad. Con la aparición y el aumento de las penalizaciones que ha establecido la regulación para los sistemas de distribución ha aumentado el número de investigaciones en la temática de la localización de fallas en este tipo de redes.

La duración de una interrupción a nivel de distribución se ve afectada de manera directa por el tiempo de localización de esta falla y es por ello que en este trabajo de tesis se ha desarrollado una metodología eficiente y funcional que permite localizar la zona y estimar la distancia a la cual ha ocurrido un evento de falla en redes de distribución. Bajo el supuesto de que la *oportuna localización de las fallas ocurridas en la red permitirá a las empresas prestadoras del servicio mejorar sus indicadores de continuidad en el servicio* la metodología propuesta en este trabajo permite disminuir considerablemente los tiempos de localización de fallas por otra parte adicionalmente la continua localización de las mismas permite detectar las debilidades de la red y fortalecer el sistema de manera tal que las interrupciones puedan ser cada vez menores.

5.2 CONCLUSIONES ASOCIADAS A LA CARACTERIZACION DE LAS SEÑALES DE TENSION Y DE CORRIENTE.

Este trabajo está fundamentado en metodologías basadas en el conocimiento y como tal se encuentra soportada por la información que permite reproducir el conocimiento empírico de la información extraída en la subestación principal o cabecera de la red.

Como se definió en este trabajo, un descriptor del sistema bajo estado de falla es una variable que posee información asociada no sólo con las características del evento sino también con la localización de la misma. El hecho de extraer la información de la falla a partir de las señales medidas en la cabecera de la red, obedece al hecho de que para la mayoría de las redes de distribución de nuestro país, la medición de tensión y corriente se hace en el punto de inicio de la red. Se mostró que los descriptores básicos de mayor impacto para la localización de la zona de falla se encuentran asociados con el análisis de los patrones de la corriente de carga o firma del sistema; asimismo se mostró que los descriptores de mayor aplicabilidad para la estimación de la distancia de falla son los asociados a las características continuas de prefalla, falla y postfalla de las señales de tensión y corriente tanto de línea como de fase.

Es un hecho de que la localización de la zona de falla obedece a la solución de un problema cuya variable de salida es de carácter discreto (*zona de falla*); asimismo el tipo de información que se propone en este trabajo para la solución de esta etapa del problema es de tipo discreto (*análisis de los patrones de la corriente de carga*).

Los descriptores que se propusieron fueron el *número de conmutaciones presentes en la señal*, la *paridad del número de conmutaciones*, el *tiempo de recierre de los reconectores*, la *variación de las corrientes en estado de liberación de falla* y la *corriente de postfalla luego de la acción de las protecciones*. Se mostró a partir de las pruebas realizadas para la localización de fallas, empleando la combinación de dos o más de estos descriptores que el mejor acierto en la estimación de la zona de falla se da cuando se involucran los descriptores de la *paridad del número de conmutaciones*, el *tiempo para el recierre de reconectores* y la *variación de la corriente de postfalla luego de la acción de las protecciones*. Este resultado es interesante ya que confirma que el hacer uso de toda la información para la localización de la zona de falla no es lo más apropiado.

También se mostró que un descriptor importante que ayuda al mejoramiento de la localización de la zona de falla es el de la *Corriente de postfalla luego de la acción de las protecciones*, es decir que la corriente que queda al final del periodo de acción de las protecciones es una variable que discrimina la zona de localización de la falla; no obstante este descriptor sólo, no proporciona la información necesaria para la correcta localización, pues existen diversas zonas en los sistemas de distribución que pueden poseer cargas conectadas cuyas magnitudes sean muy similares, dando lugar a la estimación incorrecta de la zona de falla (problemas de empate en el esquema de votación de las SVMc). Para dar solución a la estimación incorrecta mencionada anteriormente, es necesario realizar una adecuada calibración de las protecciones, de tal manera que si las zonas son de idéntica carga, otro tipo de descriptor permita discriminar la zona en la cual ha sucedido el evento, tales como el tiempo de recierre del reconector o el número de suiches presentes en la señal tal como se desarrolló en este trabajo.

Por otra parte, así como la información para la localización de la zona de falla proviene de la interacción de las protecciones de la red, la zonificación del sistema de distribución se hizo con base en las zonas que respalda cada una de las protecciones; el hecho de emplear otro tipo de zonas no se ajustaría para los descriptores propuestos en este trabajo; sin embargo una alternativa adicional podría basarse en la definición de sub-zonas que constituyan un subconjunto de las

zonas aquí propuestas, no obstante, el proceso de solución es de mayor exigencia al requerir información adicional a los descriptores del *análisis de los patrones de la corriente de carga* pues éstos por sí solos no permiten hacer una distinción entre sub-zonas que protege una misma protección.

De la misma manera, es un hecho de que la estimación de la distancia de falla obedece a la solución de un problema cuya variable de salida es de carácter continuo (*distancia de falla*); asimismo el tipo de información que se propone en este trabajo para la solución de esta etapa del problema es de tipo continuo (*Caída de tensión en las señales de voltaje de línea en prefalla y postfalla, Caída de tensión en las señales de voltaje de línea tierra en prefalla y postfalla, duración de la señal transitoria y frecuencia de la señal transitoria*).

Para las pruebas realizadas se mostró, en cuanto a los descriptores asociados con la *Caída de tensión*, que la estimación de la distancia de falla se encuentra estrechamente relacionado a la impedancia de la red y al valor de la impedancia de falla; este fenómeno obedece al hecho de que la magnitud del hueco de tensión depende de gran manera del valor de la resistencia de falla, la cual, según su valor, puede predominar sobre la impedancia del sistema de distribución. Se mostró adicionalmente que para valores de la resistencia de falla inferiores a la impedancia equivalente de la red (medida en el punto de falla), los descriptores asociados con la caída de tensión presentan un buen comportamiento para la estimación de la distancia de falla, pero no así para los casos en los que la resistencia de falla posee una magnitud mayor a la impedancia del sistema de distribución. Esta tarea requirió del cálculo de la impedancia equivalente de la red, medida desde cada uno de los nodos del sistema de distribución.

Por otra parte y aunque no se presenta formalmente en el capítulo cuatro, el descriptor asociado con la frecuencia de falla está directamente relacionado con la distancia de falla, al encontrarse determinado por los parámetros inductivos y capacitivos de la red; no obstante debido al modelamiento del sistema de distribución (las cargas son de impedancia constante) el comportamiento de este descriptor podría ser distinto al emplear un modelo diferentes para las cargas. En cuanto a la duración del evento transitorio en la red, se encuentra que éste está fuertemente supeditado al valor de la resistencia de falla dado que el factor de amortiguamiento de las señales transitorias obedece a parámetros no solo inductivos y capacitivos sino también resistivos [Greenwood 1990].

Los resultados obtenidos con estos dos descriptores, hacen suponer que los sistemas de medición a nivel de la subestación principal del sistema, deben poseer una frecuencia de muestreo de cien mil muestras por segundo, sin embargo, esto podría ser económicamente infactible para las redes de distribución del país.

Los descriptores asociados al análisis de los patrones de la corriente de carga juegan un papel muy importante en la tarea de localización de la zona de falla; estos descriptores están asociados con las protecciones del sistema de distribución y para su simulación en la herramienta ATP es necesario hacer uso de los MODELS. Se debe tener cuidado con este tipo de implementaciones dado que el tiempo y el esfuerzo computacional en la simulación de fallas puede incrementarse excesivamente.

5.3 CONCLUSIONES ASOCIADAS A LA MBC PARA LA LOCALIZACION DE FALLAS.

El problema de localización de fallas en sistemas de distribución puede ser abordado por Metodologías Basadas en el Conocimiento (MBC). Para la estimación de la zona de falla se requiere una MBC capaz de resolver problemas de clasificación; este tipo de métodos permiten

establecer una relación entre un conjunto de datos de entrada (descriptores) y un conjunto de salida (zona de ocurrencia de la falla).

Los métodos de clasificación se caracterizan por tener dos etapas básicas: Entrenamiento y validación. En la primera etapa se sintonizan todos los parámetros asociados a la MBC y para tal fin existen diversas herramientas; en este trabajo se empleó la validación cruzada y la búsqueda en malla. En la etapa de validación se verifica la efectividad del método de clasificación; nuevos valores de los descriptores se le presentan a la MBC para que ésta localice la zona de falla.

Asimismo se pueden emplear MBC para la estimación de la distancia de falla, a este tipo de metodologías se les conoce con el nombre de estimadores o máquinas de regresión, pues permiten establecer una relación (generalmente no lineal) entre un conjunto de datos de entrada (descriptores) y una variable continua (distancia de falla).

En este trabajo se propone una MBC fundamentada en Máquinas de Soporte Vectorial (SVM), tanto para la localización de la zona como la estimación de la distancia de falla respectivamente. Las máquinas de soporte vectorial para clasificación, presentan un alto desempeño en la estimación de la zona de falla debido a la fuerte correlación que existe entre los descriptores empleados (*Análisis de los patrones de la corriente de carga*) para la localización y las zonas del sistema de distribución (*las zonas del sistema corresponden con las zonas primarias de cada una de las protecciones del sistema*). Las máquinas de soporte vectorial para regresión presentan una muy buena estimación de la distancia de falla para aquellos eventos donde el valor de la resistencia de falla es pequeño pero para los casos donde el valor de la resistencia de falla es alto se presenta dispersión en la distancia estimada, esto se debe a la fuerte dependencia que existe entre los descriptores empleados (*caída de tensión en las señales de voltaje de línea y línea a tierra*) y el valor de la resistencia de falla (para valores elevados de la resistencia de falla se presenta una estimación de distancia de falla errónea).

La solución del problema de localización de la zona de falla mediante SVMc se resuelve a través de un problema de optimización cuadrático, donde el número de variables equivale al número de datos de entrenamiento, tal como se presentó en el capítulo 2. Dado que el problema de localización de la zona de falla en el sistema de distribución escogido para pruebas (y en general para la mayoría de los sistemas de distribución) presenta un elevado número de datos, fue necesario dividir el problema de clasificación entre siete SVMc donde cada una de ellas está especializada en la localización de un tipo de falla específica.

Por otra parte, para la solución del problema de estimación de la distancia de falla por medio de SVMr se involucra un problema de optimización cuadrático donde el número de variables equivale a dos veces el número de datos de entrenamiento tal como se presentó en el capítulo 2; adicionalmente, la matriz Hessiana de la formulación del problema de regresión presenta singularidades que dificultan el proceso de entrenamiento. Dado que para la estimación de la distancia de falla en el sistema de distribución escogido para pruebas (y en general para la mayoría de los sistemas de distribución) se posee un elevado número de datos de entrenamiento fue necesario dividir el problema mediante dos criterios: el tipo de falla y la zona de falla.

5.4 TRABAJOS FUTUROS

A partir de la investigación realizada en este trabajo se presentan nuevas alternativas para ampliar los aportes de ésta; algunos de los más importantes se presentan a continuación.

En el desarrollo de este trabajo se empleó como sistema de pruebas un sistema de distribución de gran escala propuesto por IEEE, sin embargo es necesario aplicar la metodología propuesta para sistemas de distribución reales; para ello se deben establecer acuerdos con las empresas de distribución para que suministren información de las señales de tensión y corriente en estado de falla, medidas en la subestación principal del sistema.

En este trabajo se escogieron los MODELS del ATP para la simulación de las protecciones del sistema de distribución; sin embargo existen otras alternativas para dicha tarea; con esto queda abierta la posibilidad de hacer más eficiente el proceso de la simulación automática de fallas en sistemas de distribución a gran escala.

Bajo los parámetros con los cuales fue realizada la simulación automática de fallas en este trabajo, se abre una nueva posibilidad para la implementación eficiente de metodologías basadas en el modelo eléctrico de la red (MBM), y más aún para el caso de sistemas de distribución real. La implementación eficiente consiste en la simulación del sistema de distribución real en el ATPDraw que a su vez genera la tarjeta que contiene todos los datos de los elementos del sistema, posteriormente, mediante un algoritmo que interprete todos los parámetros de la tarjeta, es posible crear una base de datos donde se encuentren cada uno de los elementos de la red, sus parámetros y el ordenamiento nodal, así es posible llevar esta información en un formato estándar para la implementación de cualquier metodología basada en el modelo, las cuales van recorriendo de manera iterativa cada uno de los tramos del sistema de distribución, partiendo desde la subestación principal y finalizando con cada uno de los ramales laterales.

BIBLIOGRAFIA

- [Aguilar 1982] Aguilar, J. Lopez, N. "The process of classification and learning the meaning of linguistic descriptors of concepts". Approximate Reasoning in Decision Analysis, 1982. p. 165-175. North Holland.
- [Angulo 2001] Angulo, C. "Aprendizaje con Máquinas Núcleo en entornos de Multiclasificación". Tesis Doctoral
- [Billinton 1984] Billinton, R. Allan, R. "Reliability Evaluation of Power Systems". New York, Plenum Press , 1984.
- [Bollen 2000] Bollen , M. "understanding Quality Problems: Voltage Sags and Interruptions". IEEE Press. 2000.
- [Burges (2001)] Burges, C. (2001). A tutorial on Support Vector Machines for Patter Recognition. <http://www.kernel-machines.org/>
- [CAUE (2002)] ATP Rulebook. Comité Argentino de Usuarios del EMTP/ATP, Argentina (CAUE 2002). Marzo de 2002.
- [Chapman 1987] CHAPMAN, S. "Máquinas Eléctricas". Mc Graw Hill. S.A. Bogotá, Colombia, 1987
- [Cormane et al (2005)] Cormane, A., Vargas, H., Ordóñez, G. (2005). Modelo Estadístico para la Localización de Fallas en Sistemas de Distribución. III Simposio Internacional sobre Calidad de la Energía Eléctrica SICEL 2005 Bogotá D.C., Colombia.
- [CREG 1998] Comisión de Regulación de Energía y Gas. Resolución 070 del año 1998. Por la cual se establece el reglamento de distribución de energía eléctrica como parte del reglamento de operación del sistema interconectado nacional.
- [CREG 2003] Comisión de Regulación de Energía y Gas. Resolución 113 del año 2003. Por la cual se dictan normas en materia de calidad del servicio de energía eléctrica prestado en el sistema de interconexión nacional del que trata el reglamento de distribución de energía eléctrica.
- [CREG 2005] Comisión de Regulación de Energía y Gas. Resolución 024 del año 2005. Por la cual se modifican las normas de calidad de la potencia eléctrica aplicables a los servicios de distribución de energía.

- [Cristianini 2000] Cristianini, N. Shawe-Taylor, J. "An introduction to Support Vector Machines", Cambridge University Press, 2000
- [Das 1998] Das, R. "Determining the Locations of Faults in Distribution Systems", Doctoral Thesis, University of Saskatchewan Saskatoon, Canada, Spring 1998.
- [Gallego 2006] Gallego, R. Granada M, Escobar, A. "Reducción de Pérdidas por Etapas en Sistemas de Distribución". ISBN 958-8272-27-0 Universidad Tecnológica de Pereira. 2006.
- [Gonen 1986] Turan Gonen, "Electric Power Distribution System Engineering". Mc Graw-Hill Publishers.1986
- [Gonen 1988] Turan Gonen, "Electric Power Transmission System Engineering: Analysis and Design", John Wiley & Sons, 1988
- [Greenwood 1990] Greenwood, Allan, "Electrical Transients in Power Systems", Electric Power Engineering Department, Troy, New York. Second edition, March 1990.
- [Gunn 1998] Gunn, S. "Support Vector Machine for Classification and Regression" Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science. 1998.
- [Henao 2004] Henao, R. "Selección de Hiperparámetros en Máquinas de Soporte Vectorial". Tesis de Maestría en Control y Automatización Industrial. Universidad Nacional de Manizales. 2004
- [IEEE 123 Node Test] IEEE POWER ENGINEERING SOCIETY. "IEEE 123 Node Test Feeder". Power System Analysis Subcommittee, Computing and Economics Committee.
- [Kagan 1998] Kagan N., De Oliveira C., Méffe A., Alocação de Bancos de Capacitores e reguladores de tensão em redes de distribuição. University of São Paulo. Brazil. Department of Electrical Engineering. 1998.
- [Lewis (1998)] Lewis, J. (1998). Protective Relaying Principles and Applications. Marcel Dekker Inc.
- [Lin 2004] Lin, C. Chang, C. Hsu, C. "A practical guide to support vector clasification". National Taiwan University. 2004
- [Li-Xing (1997)] Li-Xing W. (1997). A Course in Fuzzy Systems and Control. Prentice Hall.
- [Martin et al (2001)] Martin, B., Sanz, A. (2001). Redes Neuronales y Sistemas Borrosos. RA-MA.
- [Martínez 2002] Martínez Malo Julio. De Castro Fernández Rosa María. Análisis de la teoría de Ondículas Orientada a las aplicaciones en Ingeniería Eléctrica: Fundamentos. E.T.S.I. Industriales departamento de Ingeniería Eléctrica.2002
- [Martins et al (2002)] Martins J. Sousa L. Fernao V. Alegria V. "A New Accurate Fault Location Method Using ab Space Vector Algorithm". XIV Power System Control Conference. 2002.

- [Martins et al (2003)] Martins J. Sousa L. Fernao V. Alegria V. "A Network Distribution Power System Fault Location based on neural Eigenvalue Algorithm". IEEE Bologna Power Tech. 2003
- [Mora 2006] Mora, J. "Localización de faltas en sistemas de distribución de Energía Eléctrica usando métodos basados en el modelo y métodos basados en el conocimiento", Tesis Doctoral, Universidad de Girona, España, 2006.
- [Mora et al (2005a)] Mora, J., Muñoz S., Carrillo C. (2005a). Técnicas algorítmicas de localización de fallas como alternativa para reducir el efecto de las salidas en sistemas de potencia - Una revisión. III Simposio Internacional sobre Calidad de la Energía Eléctrica SICEL 2005 Bogotá D.C., Colombia.
- [Mora et al (2005b)] Mora, J., Serrano, H., Ordóñez, G. (2005c). Mejoramiento de los Índices de Calidad del Servicio mediante la Caracterización de Señales para Localización de Fallas en Distribución. III International Symposium of Power Quality 2005. IEEE Branch - Colombia, Bogotá.
- [Mora et al (2005c)] Mora, J., Meléndez, J., Carrillo, G. (2005b). Una Arquitectura Genérica para el Desarrollo de Sistemas Híbridos. Propuesta de localizador de fallas para mejorar los índices de calidad. III Simposio Internacional sobre Calidad de la Energía Eléctrica SICEL 2005 Bogotá D.C., Colombia.
- [Mora et al 2003] Mora, J. Colomer, J. "Voltage Sag Characterization and Classification for Diagnosis in Electric Power Quality Domain", Research Report, Department of Electronics, Computer Science and Automatic Control – Group eXIT, University of Girona, España – 2003.
- [Morales et al (2008)] G. Morales, J. Florez, H. Vargas. "Elimination of multiple estimation for fault location in radial power systems by using fundamental single end measurements".
- [Phadke 1995a] Phadke, A., Horowitz, S. (1995). Power System Relaying. Research studies press. Baldock, England 1995.
- [Phadke 1995b] Phadke, A., Power System Protection. IEEE Press Power Engineering Series. Mc Graw-Hill. 1995.
- [Relaying 999] Power System Relaying Comitee. "IEEE Standard Inverse-Time Characteristic equations for overcurrent relays". IEEE Transactions on Power Delivery, Vol 14, No. 3, July 1999
- [Rodas 2004] Rodas, Darío, Gallego, Ramón, Padilla-Feltrin A., Ochoa L. F. "Three-phase autotransformers modeling". IEEE Power Delivery, enviado para evaluación (2004).
- [Russell et al (2000)] Russell, S., Norvig, P. (2002). Artificial Intelligence: A Modern Approach (Second Edition). Prentice Hall.
- [Seung et al 2004] Seung Jae Lee, Myeon Song Choi, "An Intelligent and Efficient Fault Location and Diagnosis Scheme for Radial Distribution Systems". IEEE Transactions on Power Delivery, vol. 19 No 2, April 2004.
- [Short 2003] T. A. Short. "Electric Power Distribution Handbook". CRC Press. 2003.

- [SMO 2000] "Sequential Minimal Optimization for SVM", <http://www.ics.uci.edu/~xge/svm/>
- [Smola 1998] Smola, J. Scholkopf, B. "A Tutorial on Support Vector Regression" part of the ESPRIT Working Group in Neural and Computational Learning II. NeuroCOLT. 1998.
- [Waissman et al (2000)] Waissman, J., Aguilar, M., Dahhou, B. (2000). Construction d'un Modele Comportemental pour la Supervision de Procedes: Application a une Station de Traitement des Eaux. Doctoral Thèse, Institut National Polytechnique de Toulouse. France.

ANEXO 1 INFORMACION DEL SISTEMA DE DISTRIBUCIÓN ESCOGIDO PARA PRUEBAS

Datos de los segmentos de línea

Nodo A	Nodo B	Long (ft)	Config.
1	2	175	10
1	3	250	11
1	7	300	1
3	4	200	11
3	5	325	11
5	6	250	11
7	8	200	1
8	12	225	10
8	9	225	9
8	13	300	1
9	14	425	9
13	34	150	11
13	18	825	2
14	11	250	9
14	10	250	9
15	16	375	11
15	17	350	11
18	19	250	9
18	21	300	2
19	20	325	9
21	22	525	10
21	23	250	2
23	24	550	11
23	25	275	2
25	26	350	7
25	28	200	2
26	27	275	7
26	31	225	11
27	33	500	9
28	29	300	2
29	30	350	2
30	250	200	2
31	32	300	11
34	15	100	11
35	36	650	8
35	40	250	1
36	37	300	9
36	38	250	10
38	39	325	10
40	41	325	11
40	42	250	1
42	43	500	10
42	44	200	1
44	45	200	9

Datos de los segmentos de línea
(Continuación)

42	44	200	1
44	45	200	9
44	47	250	1
45	46	300	9
47	48	150	4
47	49	250	4
49	50	250	4
50	51	250	4
52	53	200	1
53	54	125	1
54	55	275	1
54	57	350	3
55	56	275	1
57	58	250	10
57	60	750	3
58	59	250	10
60	61	550	5
60	62	250	12
62	63	175	12
63	64	350	12
64	65	425	12
65	66	325	12
67	68	200	9
67	72	275	3
67	97	250	3
68	69	275	9
69	70	325	9
70	71	275	9
72	73	275	11
72	76	200	3
73	74	350	11
74	75	400	11
76	77	400	6
76	86	700	3
77	78	100	6
78	79	225	6
78	80	475	6
80	81	475	6
81	82	250	6
81	84	675	11
82	83	250	6
84	85	475	11
86	87	450	6
87	88	175	9
87	89	275	6

Datos de los segmentos de línea
(Continuación)

89	90	225	10
89	91	225	6
91	92	300	11
91	93	225	6
93	94	275	9
93	95	300	6
95	96	200	10
97	98	275	3
98	99	550	3
99	100	300	3
100	450	800	3
101	102	225	11
101	105	275	3
102	103	325	11
103	104	700	11
105	106	225	10
105	108	325	3
106	107	575	10
108	109	450	9
108	300	1000	3
109	110	300	9
110	111	575	9
110	112	125	9
112	113	525	9
113	114	325	9
135	35	375	4
149	1	400	1
152	52	400	1
160	67	350	6
197	101	250	3

Configuración de las líneas Aéreas

Conf.	Fases	Conductor de Fase	Conductor de Neutro	Espaciamento
		ACSR	ACSR	ID
1	A B C N	336,400 26/7	4/0 6/1	500
2	C A B N	336,400 26/7	4/0 6/1	500
3	B C A N	336,400 26/7	4/0 6/1	500
4	C B A N	336,400 26/7	4/0 6/1	500
5	B A C N	336,400 26/7	4/0 6/1	500
6	A C B N	336,400 26/7	4/0 6/1	500
7	A C N	336,400 26/7	4/0 6/1	505
8	A B N	336,400 26/7	4/0 6/1	505
9	A N	1/0	1/0	510
10	B N	1/0	1/0	510
11	C N	1/0	1/0	510

Configuración de las líneas Subterráneas

Conf.	Fases	Cable	Espaciamento
12	A B C	1/0 AA, CN	515

Datos de los Transformadores

	kVA	kV del Lado de Alta	kV del Lado de Baja	R - %	X - %
Subestación	5.000	115 - D	4.16 Gr-W	1	8
XFM - 1	150	4.16 - D	.480 - D	1,27	2,72

INYECTORES DE POTENCIA REACTIVA
(CAPACITORES)

Nodo	Fs A kVAr	Fs B kVAr	Fs C kVAr
83	200	200	200
88	50		
90		50	
92			50
Total	250	250	250

Información de las Cargas

Nodo	Carga	Fs 1	Fs 1	Fs 2	Fs 2	Fs 3	Fs 3
	Modelo	kW	kVAr	kW	kVAr	kW	kVAr
1	Y-PQ	40	20	0	0	0	0
2	Y-PQ	0	0	20	10	0	0
4	Y-PQ	0	0	0	0	40	20
5	Y-I	0	0	0	0	20	10
6	Y-Z	0	0	0	0	40	20
7	Y-PQ	20	10	0	0	0	0
9	Y-PQ	40	20	0	0	0	0
10	Y-I	20	10	0	0	0	0
11	Y-Z	40	20	0	0	0	0
12	Y-PQ	0	0	20	10	0	0
16	Y-PQ	0	0	0	0	40	20
17	Y-PQ	0	0	0	0	20	10
19	Y-PQ	40	20	0	0	0	0
20	Y-I	40	20	0	0	0	0
22	Y-Z	0	0	40	20	0	0
24	Y-PQ	0	0	0	0	40	20
28	Y-I	40	20	0	0	0	0
29	Y-Z	40	20	0	0	0	0
30	Y-PQ	0	0	0	0	40	20
31	Y-PQ	0	0	0	0	20	10
32	Y-PQ	0	0	0	0	20	10
33	Y-I	40	20	0	0	0	0
34	Y-Z	0	0	0	0	40	20
35	D-PQ	40	20	0	0	0	0
37	Y-Z	40	20	0	0	0	0
38	Y-I	0	0	20	10	0	0
39	Y-PQ	0	0	20	10	0	0
41	Y-PQ	0	0	0	0	20	10
42	Y-PQ	20	10	0	0	0	0
43	Y-Z	0	0	40	20	0	0
45	Y-I	20	10	0	0	0	0
46	Y-PQ	20	10	0	0	0	0
47	Y-I	35	25	35	25	35	25
48	Y-Z	70	50	70	50	70	50
49	Y-PQ	35	25	70	50	35	25
50	Y-PQ	0	0	0	0	40	20
51	Y-PQ	20	10	0	0	0	0
52	Y-PQ	40	20	0	0	0	0
53	Y-PQ	40	20	0	0	0	0
55	Y-Z	20	10	0	0	0	0
56	Y-PQ	0	0	20	10	0	0
58	Y-I	0	0	20	10	0	0
59	Y-PQ	0	0	20	10	0	0

Información de las Cargas (Cont.)

58	Y-I	0	0	20	10	0	0
59	Y-PQ	0	0	20	10	0	0
60	Y-PQ	20	10	0	0	0	0
62	Y-Z	0	0	0	0	40	20
63	Y-PQ	40	20	0	0	0	0
64	Y-I	0	0	75	35	0	0
65	D-Z	35	25	35	25	70	50
66	Y-PQ	0	0	0	0	75	35
68	Y-PQ	20	10	0	0	0	0
69	Y-PQ	40	20	0	0	0	0
70	Y-PQ	20	10	0	0	0	0
71	Y-PQ	40	20	0	0	0	0
73	Y-PQ	0	0	0	0	40	20
74	Y-Z	0	0	0	0	40	20
75	Y-PQ	0	0	0	0	40	20
76	D-I	105	80	70	50	70	50
77	Y-PQ	0	0	40	20	0	0
79	Y-Z	40	20	0	0	0	0
80	Y-PQ	0	0	40	20	0	0
82	Y-PQ	40	20	0	0	0	0
83	Y-PQ	0	0	0	0	20	10
84	Y-PQ	0	0	0	0	20	10
85	Y-PQ	0	0	0	0	40	20
86	Y-PQ	0	0	20	10	0	0
87	Y-PQ	0	0	40	20	0	0
88	Y-PQ	40	20	0	0	0	0
90	Y-I	0	0	40	20	0	0
92	Y-PQ	0	0	0	0	40	20
94	Y-PQ	40	20	0	0	0	0
95	Y-PQ	0	0	20	10	0	0
96	Y-PQ	0	0	20	10	0	0
98	Y-PQ	40	20	0	0	0	0
99	Y-PQ	0	0	40	20	0	0
100	Y-Z	0	0	0	0	40	20
102	Y-PQ	0	0	0	0	20	10
103	Y-PQ	0	0	0	0	40	20
104	Y-PQ	0	0	0	0	40	20
106	Y-PQ	0	0	40	20	0	0
107	Y-PQ	0	0	40	20	0	0
109	Y-PQ	40	20	0	0	0	0
111	Y-PQ	20	10	0	0	0	0
112	Y-I	20	10	0	0	0	0
113	Y-Z	40	20	0	0	0	0
114	Y-PQ	20	10	0	0	0	0
Total		1420	775	915	515	1155	635

Impedancias del sistema de distribución para pruebas de la IEEE de 123 nodos

Configuración 1

Z (R +jX) [Ω /milla]					
0.4576	1.0780	0.1560	0.5017	0.1535	0.3849
		0.4666	1.0482	0.1580	0.4236
				0.4615	1.0651
B [μ S/milla]					
5.6765	-1.8319	-0.6982			
	5.9809	-1.1645			
		5.3971			

Configuración 2

Z (R +jX) [Ω /milla]					
0.4666	1.0482	0.1580	0.4236	0.1560	0.5017
		0.4615	1.0651	0.1535	0.3845
				0.4576	1.0780
B [μ S/milla]					
5.9809	-1.1645	-1.8319			
	5.3971	-0.6982			
		5.6765			

Configuración 3

Z (R +jX) [Ω /milla]					
0.4615	1.0651	0.1535	0.3849	0.1580	0.4236
		0.4576	1.0780	0.1560	0.5017
				0.4666	1.0482
B [μ S/milla]					
5.3971	-0.6982	-1.1645			
	5.6765	-1.8319			
		5.9809			

Configuración 4

Z (R +jX) [Ω /milla]					
0.4615	1.0651	0.1535	0.3849	0.1580	0.4236
		0.4666	1.0482	0.1560	0.5017
				0.4576	1.0780
B [μ S/milla]					
5.3971	-1.1645	-0.6982			
	5.9809	-1.8319			
		5.6765			

Configuración 5

Z (R +jX) [Ω /milla]					
0.4666	1.0482	0.1560	0.5017	0.1580	0.4236
		0.4576	1.0780	0.1535	0.3849
				0.4615	1.0651
B [μ S/milla]					
5.9809	-1.8319	-1.1645			
	5.6765	-0.6982			
		5.3971			

Configuración 6

Z (R +jX) [Ω /milla]					
0.4576	1.0780	0.1535	0.3849	0.1560	0.5017
		0.4615	1.0651	0.1580	0.4236
				0.4666	1.0482
B [μ S/milla]					
5.6765	-0.6982	-1.8319			
	5.3971	-1.1645			
		5.9809			

Configuración 7

Z (R +jX) [Ω /milla]					
0.4576	1.0780	0.0000	0.0000	0.1535	0.3849
		0.0000	0.0000	0.0000	0.0000
				0.4615	1.0651
B [μ S/milla]					
5.1154	0.0000	-1.0549			
	0.0000	0.0000			
		5.1704			

Configuración 8

Z (R +jX) [Ω /milla]					
0.4576	1.0780	0.1535	0.3849	0.0000	0.0000
		0.4615	1.0651	0.0000	0.0000
				0.0000	0.0000
B [μ S/milla]					
5.1154	-1.0549	0.0000			
	5.1704	0.0000			
		0.0000			

Configuración 9

Z (R +jX) [Ω /milla]					
1.3292	1.3475	0.0000	0.0000	0.0000	0.0000
		0.0000	0.0000	0.0000	0.0000
				0.0000	0.0000
B [μ S/milla]					
4.5193	0.0000	0.0000			
	0.0000	0.0000			
		0.0000			

Configuración 10

Z (R +jX) [Ω /milla]					
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		1.3292	1.3475	0.0000	0.0000
				0.0000	0.0000
B [μ S/milla]					
0.0000	0.0000	0.0000			
	4.5193	0.0000			
		0.0000			

Configuración 11

Z (R +jX) [Ω /milla]					
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		0.0000	0.0000	0.0000	0.0000
				1.3292	1.3475
B [μ S/milla]					
0.0000	0.0000	0.0000			
	0.0000	0.0000			
		4.5193			

Configuración 12

Z (R +jX) [Ω /milla]					
1.5209	0.7521	0.5198	0.2775	0.4924	0.2157
		1.5329	0.7162	0.5198	0.2775
				1.5209	0.7521
B [μ S/milla]					
67.2242	0.0000	0.0000			
	67.2242	0.0000			
		67.2242			

ANEXO 2

CÓDIGO FUENTE DEL MODEL DEL FUSIBLE MONOFÁSICO IMPLEMENTADO EN ATPDraw

MODEL FUSE

INPUT XXIN

DATA IPK
A
B
P
DIAL

VAR af[1..500], -- Vector que recoge las muestras de corrientes.
YY, -- Salida Booleana para el suiche que despejará la Falla.
IRMS, -- Valor RMS de la corriente.
PP, KK, JJ, -- Ancho de Ventana RMS. Realiza desplazamiento. Contador
para Inicio Irms
AUX, FLGD, -- Variables para medir Irms
TEMR, -- Temporizador Tiempo de APERTURA.
TSET, -- T1: Tiempo para ABRIR protección
MM, -- Múltiplo de la corriente de Arranque.
IT1 -- Valor de la Corriente en el momento de la orden de apertura

OUTPUT YY

EXEC

```
comment *****
***** INICIALIZACION DE VARIABLES *****
***** endcomment
```

```
IF T=0 THEN
  af[1..500]:=0 -- Vector que recoge las muestras de corrientes.
  YY:=BOOL(1) -- Salida Booleana para el suiche que despejará la Falla.
  IRMS:=0 -- Valor RMS de la corriente.
  PP:=1/120 -- Ancho de Ventana RMS
  KK:=1 -- Para hacer desplazamiento de Componentes del vector "a".
  JJ:=2 -- Contador para inicio de la IRMS
  AUX:=0 -- Variable auxiliar para medir la IRMS
  FLGD:=0 -- Bandera Auxiliar para medir la IRMS

  TEMR:= 0 -- Temporizador Tiempo de APERTURA, curva RAPIDA
  TSET:=stoptime -- Tiempo de apertura. Se calcula con la Curva del Fuse.
  IT1:=1 -- Valor de la Corriente en el instante de la orden de disparo
  -- Inicia valiendo un valor cualquiera (en este caso: 1)
  MM:=0 -- Múltiplo de la corriente de Arranque
ENDIF
```

```
IF YY=1 THEN -- Condición General
```

```

comment *****
*****      LECTURA DE LA CORRIENTE RMS      *****
***** endcomment
IF (T>0) AND (T<=PP) THEN -- Primeras muestras mientras se llena el vector
  af[JJ]:=XXIN**2*timestep
  AUX:=AUX+af[JJ]
  JJ:=JJ+1
ENDIF

IF T>PP THEN
  IF FLGD=0 THEN
    IRMS:=sqrt(AUX/PP)
    FLGD:=1
  ENDIF
  IF FLGD=1 THEN
    IRMS:= ( ( IRMS**2*PP ) + timestep*XXIN**2 - af[1] ) / PP
    IF IRMS<=0 THEN
      IRMS:=0
    ENDIF
    IRMS:=SQRT(IRMS)
    af[JJ-1]:=XXIN**2*timestep
    FOR KK:=1 to JJ+2 DO -- Se actualiza el vector así:
      -- En la primera posición quedará la segunda muestra
      -- En la segunda posición quedará la tercera muestra
      -- En la n,esima posición quedará la n+1,esi muestra
      af[KK]:=af[KK+1]
    ENDFOR
  ENDIF
ENDIF

comment *****
*****      ACTUACION DEL FUSE      *****
***** endcomment
MM:=IRMS/IPK
IF MM<1 THEN
  TEMR:=0
  TSET:=stoptime
  IT1:=1
ENDIF
IF (MM>=1) AND (TEMR<=TSET) THEN
  TSET:=DIAL*(B+ A/(MM**P-1))
  TEMR:=TEMR+timestep
  IF TEMR>TSET THEN
    IT1:=XXIN
  ENDIF
ENDIF
IF (IRMS>IPK) AND (TEMR>TSET) AND (IT1*XXIN<0) THEN
  YY:=BOOL(-1)
ENDIF

ENDIF -- End de la Condición General
ENDEXEC
ENDMODEL

```