

Estrategias para la localización relativa de la fuente de huecos de tensión en sistemas de distribución de energía eléctrica

Strategies for Voltage Sag Source Relative Location in Power Distribution Networks

Victor Augusto Barrera Nuñez^{1}, Xavier Berjaga Moliner¹, Gabriel Ordóñez Plata², Joaquim Meléndez Frigola¹*

¹Grupo eXiT, Instituto de Informática Industrial y Automática – IiiA, Universitat de Av. Lluís Santalo, E-17071, Girona, España.

²Grupo GISEL, Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones Universidad Industrial de Santander, Cra. 27 Calle 9, Bucaramanga, Colombia.

(Recibido el 30 de octubre de 2008. Aceptado el 26 de mayo de 2009)

Resumen

En este artículo se presentan y comparan dos estrategias para estimar el origen (aguas arriba o aguas abajo) de huecos de tensión registrados en subestaciones de distribución. La primera estrategia se basa en la aplicación de una regla booleana que conjuga impedancias obtenidas antes y durante la perturbación. La segunda estrategia utiliza la forma de onda completa de las tensiones y corrientes trifásicas registradas durante el evento. Las dos estrategias fueron verificadas con huecos de tensión registrados en subestaciones de distribución.

----- *Palabras clave:* Localización de fallas, monitorización de la calidad de la potencia, hueco de tensión.

Abstract

In this paper two strategies to estimate the origin (upstream or downstream) of voltage sags gathered in distribution substation are presented and compared. The first strategy is based on boolean rule that contains impedances obtained before and during the perturbation. The second strategy uses the three-phase voltage and current complete waveforms gathered during the event. Both strategies were verified with voltage sag events gathered in distribution substations.

----- *Keywords:* Fault location, power quality monitoring, voltage sag (dip).

* Autor de correspondencia: teléfono: + 57 + 7+ 634 40 00 ext: 2703, correo electrónico: victor.barrera@udg.edu.co (V. Barrera).

Introducción

Debido al impacto en las cargas industriales y los costos producidos por posibles daños y gastos de mantenimiento, el sector industrial está prestando atención al estudio de los huecos de tensión. Estas perturbaciones se propagan a través del sistema eléctrico afectando las cargas conectadas a la red. Por lo tanto, la localización del origen de huecos de tensión es un aspecto de interés, debido a la necesidad de valorar la responsabilidad en la generación de huecos de tensión. Una vez registrado un hueco de tensión, inmediatamente se deben realizar los dos siguientes pasos: determinar si la fuente del hueco de tensión está aguas arriba o aguas abajo con respecto al punto de registro de la perturbación; luego, un algoritmo de localización debe ser utilizado para estimar la localización exacta de la fuente que causó la perturbación registrada. Este artículo se centra en el primer paso. Estudios recientes demuestran la importancia de desarrollar algoritmos robustos y sencillos para la estimación de la localización relativa de la fuente que originó un determinado hueco de tensión 1.

La primera estrategia para estimar el origen de los huecos de tensión está basada en una regla obtenida a partir de un análisis estadístico multivariable de distintos parámetros (impedancias, ángulos de fase) utilizados en algoritmos de localización propuestos en la literatura. La segunda estrategia utiliza la información proporcionada por las ondas de tensión y corriente del hueco de tensión. Esta estrategia se fundamenta en un modelo basado en el *Análisis de componentes principales* (PCA) capaz de discriminar a partir de las ondas de tensión y corriente el origen de un hueco de tensión.

El problema de estimar la localización de la fuente de una perturbación a partir de las ondas registradas no es un tema reciente. Chouhy [1] comparó cinco algoritmos de localización utilizando ondas obtenidas de simulación [2-5]. En esta comparación el *Algoritmo relé distancia* (DR) obtuvo los mejores resultados de estimación [5], mientras que el *Algoritmo del signo de la resis-*

tencia (RS) obtuvo los resultados más deficientes [3]. El *Algoritmo del cambio de fase de las corrientes de secuencia* (PCSC) ha sido verificado únicamente por su autor en [6]. Los descriptores utilizados por los algoritmos DR, RS y PCSC fueron utilizados para realizar un análisis estadístico multivariable y proponer la regla de localización utilizada en la primera estrategia.

Por otro lado, Khosravi [7, 8] presenta una metodología basada en MPCA (*Análisis de componentes principales multicanal*) para crear un modelo estadístico capaz de discriminar el origen de huecos de tensión a partir de las ondas de tensión y corriente. La clasificación es realizada basándose en la proyección de nuevos huecos en el espacio de proyección definido por el modelo y analizando la distancia al centro del modelo (estadístico T^2) y el error de la proyección (estadístico Q). Posteriormente, Meléndez [9] redefinió el método agregando en el proceso de decisión un paso basado en *Razonamiento basado en casos*. La agregación de este último paso permitió mejorar los resultados de clasificación.

Este artículo está organizado en seis secciones adicionales. En la tercera sección se describe el conjunto de huecos de tensión utilizado para las pruebas. Luego, una breve descripción de la matriz de confusión y de las curvas ROC (*Receiver Operation Characteristics*), las cuales son utilizadas en la comparación de las estrategias. En la quinta y sexta sección se presentan ambos métodos. Después, en la séptima sección se comparan los resultados obtenidos por ambas estrategias. Finalmente, en la última sección se dan algunas conclusiones.

Huecos de tensión utilizados en el estudio

ENDESA, Empresa de distribución de energía de Barcelona-España, suministró un conjunto de huecos de tensión registrados en tres subestaciones (25kV) clasificados previamente según su origen. Cada perturbación contiene 128 muestras por periodo (50Hz) y 40 ciclos (Figura 1).

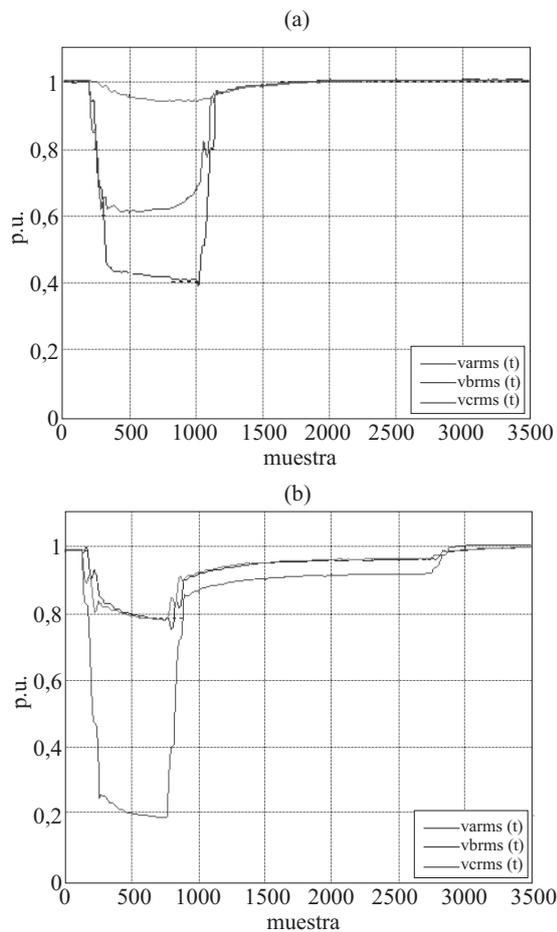


Figura 1 Ejemplo de hueco de tensión: (a) Origen aguas abajo, (b) Origen aguas arriba. Secuencia de tensión eficaz calculada con una ventana deslizante de 128 muestras

Para los registros suministrados se desconoce el punto exacto de la fuente que produjo el hueco de tensión. Por tal motivo, no se consideró conveniente presentar el diagrama unifilar de las tres redes donde fueron registradas las perturbaciones. Sin embargo, es importante resaltar que las tres subestaciones son radiales y no poseen cogeneración aguas abajo con respecto al punto de registro de los huecos de tensión analizados.

Preprocesado de los huecos de tensión

Para el cálculo de la secuencia de tensión y corriente eficaz a frecuencia nominal (50 Hz) se utilizó la Transformada Corta de Fourier - SFT con

una ventana deslizante de un ciclo y 128 muestras de longitud.

Selección de los huecos de tensión

El conjunto inicial de huecos de tensión contenía 221 perturbaciones. De este conjunto fue necesario excluir 48 registros de huecos de tensión porque carecían de estado de pre-falla o post-falla [10] y los algoritmos de localización requieren de uno o ambos estados. En la Figura 2 se presenta un hueco de tensión excluido del análisis por que carece de estado de pre-falla. Finalmente, en el estudio fueron utilizados 173 huecos de tensión, correspondientes a tres subestaciones, de los cuales 73 tienen su origen aguas abajo (\downarrow) y 100 aguas arriba (\uparrow) con respecto al punto de registro, tabla 1.

Matriz de confusión y curvas ROC

Para comparar los resultados obtenidos con las dos estrategias, se utiliza la matriz de confusión y las curvas ROC - *Receiver Operation Characteristic*. Una matriz de confusión es una representación de los resultados de clasificación como se muestra en la tabla 2. Esta muestra las diferencias entre la clase verdadera y la estimada para un conjunto de observaciones previamente etiquetadas [11].

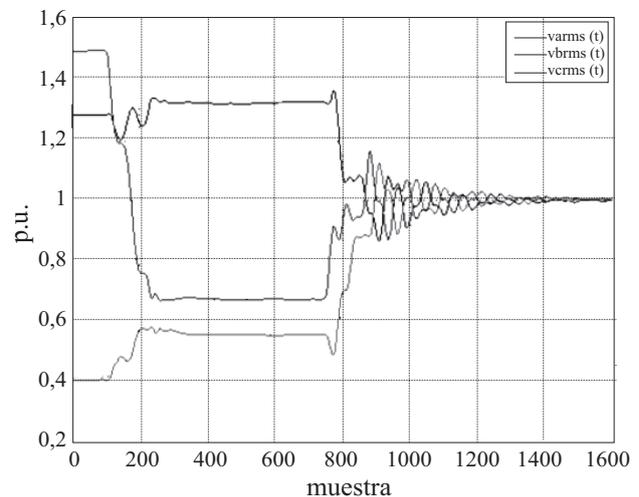


Figura 2 Ejemplo de hueco de tensión excluido del estudio (25 kV). Este evento no posee estado de pre-falla. Los algoritmos de localización de la fuente del hueco de tensión, requieren el estado de pre-falla

Tabla 1 Huecos de tensión utilizados en el estudio

Subestación	Cantidad inicial		Excluidos			Total	
	↓	↑	↓	↑	↓	↑	↓+↑
A	43	49	0	15	43	34	77
B	26	35	8	8	18	27	45
C	12	56	0	17	12	39	51
	81	140	8	40	73	100	173

↓: origen aguas abajo, ↑: origen aguas arriba

Tabla 2 Matriz de confusión

		Clase real	
		Clase referencia	Clase no referencia
Clase estimada	Clase referencia	TP	FP
	Clase no referencia	FN	TN

Donde,

TP, Verdaderos positivos. Casos correctamente clasificados y que pertenecen a la clase de referencia.

TN, Verdaderos negativos. Casos correctamente clasificados y que pertenecen a la clase no referencia.

FP, Falsos positivos. Casos clasificados como pertenecientes a la clase referencia pero su clase real es la clase no referencia.

FN, Falsos negativos. Casos clasificados como pertenecientes a la clase no referencia pero su clase real es la clase referencia.

La evaluación de estos índices permite obtener varios parámetros del clasificador. En este artículo se presta especial atención a la *sensibilidad* (ecuación 1) y *especificidad* (ecuación 2) para obtener la curva ROC.

$$Sensibilidad = \frac{TP}{TP + FN} \quad (1)$$

$$Especificidad = \frac{TN}{TN + FP} \quad (2)$$

De donde *TP*, *FN* y *FP* corresponden a las definiciones dadas anteriormente.

Una curva ROC es una gráfica bidimensional donde el eje vertical representa la *sensibilidad* y el horizontal el complemento de la *especificidad* (*1-especificidad*) del clasificador. El área bajo las curvas ROC puede ser utilizada para comparar el comportamiento de varios modelos o algoritmos de clasificación (*Area Under de ROC Curve - AUC*). Las curvas ROC y AUC son utilizadas para seleccionar los mejores parámetros para el modelo PCA en la segunda estrategia planteada en este artículo.

Estrategia basada en una regla booleana

La regla propuesta utiliza tres descriptores para determinar el origen del hueco de tensión, y es la siguiente:

Regla propuesta:

“SI $Rey < 0 \cap Z_{sag} < Z_{ss}$ ENTONCES Aguas abajo SINO Aguas arriba”

Donde *Z_{sag}* es la impedancia durante el evento y *Z_{ss}* es la impedancia de estado estable o de pre-falla. *Rey* es estimada a partir de una transformación de las componentes de secuencia positiva de tensión y corriente del sistema trifásico. *Z_{sag}* y *Z_{ss}* son impedancias propuestas en el algoritmo DR, mientras que *Rey* en el algoritmo RS.

Esta regla fue obtenida a partir de un análisis estadístico a través del *Análisis multivariante de la varianza* (MANOVA) y el *Análisis discriminante* (DA).

Algoritmos de estimación del origen de huecos de tensión

En esta sección se describen brevemente los tres algoritmos de donde se seleccionaron los descrip-

tores utilizados en el análisis estadístico multivariante.

1) Phase Change in Sequence Current - PCSC: Este algoritmo estima el origen del hueco de tensión utilizando el cambio del ángulo de fase $\Delta\phi$ entre la componente de secuencia positiva de la corriente en condiciones de falla y pre-falla. La regla utilizada por este algoritmo es:

Regla PCSC:

“SI $\Delta\phi > 0$ ENTONCES Aguas arriba SINO Aguas abajo”

Donde $\Delta\phi$ varía entre $-\pi$ y π .

El ciclo anterior a la falla es utilizado para calcular el fasor de pre-falla y el posterior para estimar el fasor de corriente de falla [6].

2) Resistance Sign - RS: El origen del hueco es obtenido a partir del signo de la parte real de la impedancia estimada. El algoritmo RS utiliza las componentes de tensión y corriente de secuencia positiva para calcular los valores de R_{ex} y R_{ey} a través de las ecuaciones 3 y 4. Las impedancias son estimadas tomando n ciclos de las ondas de tensión y corriente incluyendo ciclos de pre-falla y falla. El número de ciclos es determinado por la reversión del flujo de potencia. R_{ex} es debida a la impedancia estimada a partir de la parte real de las componentes de secuencia positiva de la tensión, mientras que R_{ey} a la parte imaginaria, ecuaciones 3 y 4.

$$\begin{bmatrix} \text{Re } x \\ X_{ex} \\ E_{RE} \end{bmatrix} = \begin{bmatrix} I_{RE}(1) & I_{IM}(1) & 1 \\ \vdots & \vdots & \vdots \\ I_{RE}(n) & I_{IM}(n) & 1 \end{bmatrix}^{-1} \begin{bmatrix} V_{RE}(1) \\ \vdots \\ V_{RE}(n) \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} \text{Re } y \\ X_{ey} \\ E_{IM} \end{bmatrix} = \begin{bmatrix} I_{IM}(1) & I_{RE}(1) & 1 \\ \vdots & \vdots & \vdots \\ I_{IM}(n) & I_{RE}(n) & 1 \end{bmatrix}^{-1} \begin{bmatrix} V_{IM}(1) \\ \vdots \\ V_{IM}(n) \end{bmatrix} \quad (4)$$

Donde,

$I_{RE}(n), I_{IM}(n), V_{RE}(n), V_{IM}(n),$

Parte real e imaginaria de la corriente y tensión en el ciclo n .

R_{ex} y R_{ey} ,

Resistencias equivalentes en eje horizontal y vertical respectivamente.

$X_{ex}, X_{ey}, E_{RE}, E_{IM}$

Reactancia y tensión de fuente

Una explicación detallada sobre éste algoritmo se presenta en la referencia [3]. La regla del algoritmo RS es la siguiente:

RS Rule:

“SI $R_{ex} > 0 \cap R_{ey} > 0$ ENTONCES Aguas arriba SINO SI $R_{ex} < 0 \cap R_{ey} < 0$ ENTONCES Aguas abajo SINO Origen indeterminado”

Por lo tanto, si ambas resistencias tienen diferente signo el algoritmo no determina el origen del hueco de tensión. Signos positivos significa que la fuente del hueco está aguas arriba y signo negativo lo contrario [3].

3) Distance Relay - DR: Se basa en el cambio de la magnitud y ángulo de la impedancia antes y después de la perturbación. Para un hueco con origen aguas abajo la impedancia durante la perturbación decrecerá con respecto a la impedancia en estado estable, y su ángulo también cambiará. La regla del algoritmo DR es:

Regla DR:

“SI $|Z_{sag}| < |Z_{ss}| \cap \Delta Z_{sag} > 0$ ENTONCES Aguas abajo SINO Aguas arriba”

Z_{sag} y Z_{ss} son estimados teniendo en cuenta los diferentes tipos de fallas, es decir, son estimadas considerando las tensiones y corrientes que están presentes durante la falla [5].

Resultados de los algoritmos

Los algoritmos fueron implementados en Matlab®. Los resultados obtenidos para cada uno de ellos se muestran en la tabla 3. La clase de referencia corresponde a los huecos de tensión cuyo origen es aguas arriba. Los porcentajes de clasificación obtenidos con los tres algoritmos son bajos: 51,4%, 27,7% y 65,9% respectivamente.

Tabla 3 Matriz de confusión de los algoritmos localización

Clase real		Clase real		Clase real		Total			
PCSC	RS	DR	TP	FN	TP		FN		
(51,4%)	(27,7%)	(65,9%)	Aguas arriba		100				
56	44	48	0	87	13	100			
FP		TN		FP		TN		Aguas abajo	
40	33	0	0	46	27	73			

Los peores resultados se dan con el algoritmo RS. Esto es debido a que *Rex* y *Rey* tienen diferentes signos en muchos casos. *Rex* no toma valores negativos, por tal razón FP y FN toman valores iguales a cero (RS arroja *origen indeterminado*). En la figura 3 se muestra este comportamiento. El algoritmo PCSC tiene un porcentaje de clasificación, 51,4%. Con este porcentaje de clasificación no es posible garantizar una adecuada clasificación. El algoritmo DR tiene el mejor índice de clasificación (65,9%), sin embargo porcentaje de clasificación es bajo para aplicaciones prácticas. Este bajo porcentaje de clasificación es debido a que la condición $\Delta Zsag > 0$ fue considerada como falsa en muchos casos de huecos de tensión cuyo origen es aguas abajo, ya que para huecos aguas abajo $\Delta Zsag > 0$ es una condición verdadera.

Herramientas estadísticas: justificación de la regla de localización

A continuación se describirán las herramientas estadísticas multivariantes utilizadas para proponer la regla de estimación del origen de huecos de tensión.

1) MANOVA: El propósito de MANOVA es explorar como las variables independientes influyen en el comportamiento de las variables dependientes.

Para este caso, el origen del hueco de tensión es la variable independiente, mientras que los descriptores $\Delta\phi$, *Rex*, *Rey*, *Zsag/Zss*, $\Delta Zsag$ son las variables dependientes. MANOVA permite responder esta pregunta: ¿Cuál es la importancia de cada descriptor con respecto al origen de hueco de tensión? Por tanto, es posible conocer el grado de incidencia (*Calidad*) del origen de hueco de tensión sobre cada uno de los cinco descriptores utilizados para el análisis. Finalmente, fueron seleccionados los descriptores con mayor grado influencia.

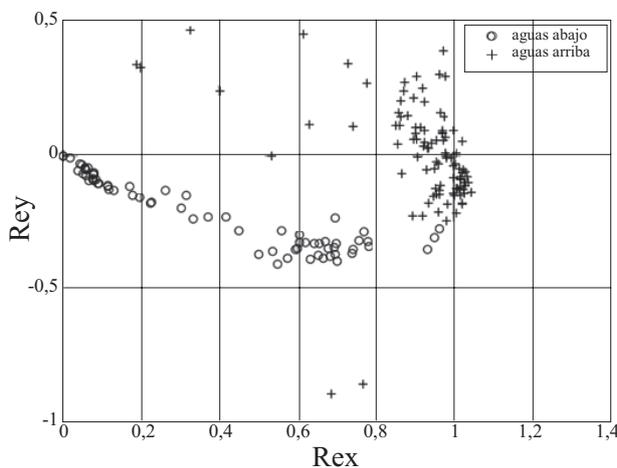


Figura 3 *Rex* Vs *Rey*. Descriptores extraídos del algoritmo RS

La tabla 4 muestra la calidad del efecto del origen del hueco de tensión (variable independiente) sobre cada uno de los cinco descriptores. Valores próximos al 100% indican que la mayor parte de la variabilidad en el descriptor está asociada con el origen del hueco de tensión.

La prueba de los efectos inter-sujetos indica que el origen del hueco de tensión no incide significativamente en $\Delta\phi$ y $\Delta Zsag$ (tabla 5), debido a que en el contraste de hipótesis el valor de significancia de éstos dos descriptores supera el p-valor 0,05; el cual corresponde a un 95% de confianza. En tabla 5, *F* corresponde al valor tomado por la variable aleatoria con *distribución F de Fisher-Snedecor*.

Tabla 4 Calidad del efecto del origen del hueco de tensión sobre los descriptores

Descriptor	Definición	Algoritmo	Calidad
$\Delta\phi$	Diferencia en el ángulo de fase entre la componente de secuencia positiva de las corrientes en condiciones de falla y pre-falla.	PCSC	1,2%
<i>Rex</i>	Parte real de la impedancia estimada a partir de la parte real de componentes de secuencia.	RS	91,2%
<i>Rey</i>	Parte real de la impedancia estimada a partir de la parte imaginaria de las componentes de secuencia.	RS	36,1%
<i>Zsag/Zss</i>	Cociente entre la impedancia en condiciones de falla y la condiciones de pre-falla.	DR	44,1%
$\Delta Zsag$	Ángulo de fase de la impedancia en condiciones de falla.	DR	0,4%

Tabla 5 Prueba de los efectos inter-sujetos

Descriptor	F	Significancia
$\Delta\phi$	2,02	0,146
<i>Rex</i>	891,95	0,00
<i>Rey</i>	49,91	0,00
<i>Zsag/Zss</i>	69,12	0,00
$\Delta Zsag$	0,647	0,53

2) DA [12]: Esta es una técnica para la clasificación de un conjunto de observaciones en clases previamente definidas. El propósito DA es determinar la clase de una observación teniendo en cuenta un conjunto de variables de entrada (descriptores). El modelo es construido en función de un conjunto de observaciones para las cuales las clases son conocidas. Luego, se consideran este conjunto de observaciones para que DA construya un conjunto de funciones lineales de las variables de entrada (*funciones discriminantes*).

DA posee un método llamado *Adicción por pasos* implementado en SPSS® (herramienta informática para el análisis estadístico, www.spss.com), el cual agrega paso por paso cada descriptor con el objetivo de determinar los descriptores más discriminantes.

El método de *Adicción por pasos* fue aplicado al conjunto de huecos de tensión. *Rex* y *Rey* fueron

obtenidas por DA como las variables más discriminadoras.

DA y MANOVA tienen resultados similares, excepto que MANOVA adiciona el descriptor *Zsag/Zss*. Consecuentemente, los tres descriptores fueron seleccionados y se analizó el grado de influencia del origen del hueco sobre cada uno de ellos.

Selección de descriptores

El comportamiento de *Rex* Vs *Rey* es mostrado en la figura 3. *Rex* sólo toma valores positivos mientras que *Rey* toma valores positivos y negativos. El rango de valores de éstas resistencias es importante porque el algoritmo RS está basado en el signo de ambas resistencias. En consecuencia, el algoritmo RS no arrojó ningún resultado correcto para los huecos de tensión con origen aguas abajo, por el contrario el algoritmo arrojó *origen indeterminado* (signos diferentes entre *Rex* y *Rey*). En la referencia [13] se obtuvieron resultados similares para este algoritmo, razón por la cual *Rex* no se considera en el análisis.

Un clasificador lineal (con pendiente negativa) sería una solución adecuada para utilizar los descriptores *Rex* y *Rey* (Figura 3), pero el objetivo de este análisis es proponer un regla, ya que con ella no es necesario realizar procesos de entrenamiento y validación.

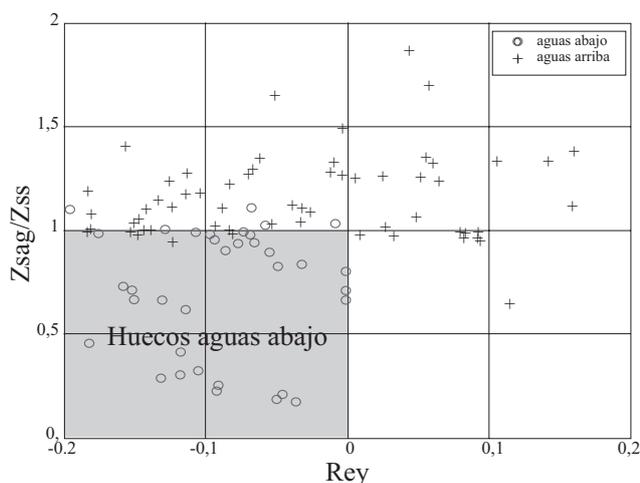


Figura 4 Rey Vs Zsag/Zss. Descriptores extraídos de los algoritmos RS y DR. Los huecos de tensión con origen aguas abajo se encuentran dentro del área $Rey < 0$ y $Zsag/Zss < 1$

La gráfica de *Rey* vs. *Zsag/Zss* permite establecer el origen de hueco de tensión. Observando la figura 4 es posible obtener la regla propuesta. Entonces, el origen del hueco de tensión será aguas abajo si $Rey < 0$ y $Zsag < Zss$. En otro caso, el origen será aguas arriba. En la misma figura, los huecos de tensión que no cumplen con esta regla, es decir, huecos aguas abajo fuera del área o aguas arriba dentro del área, es debido a la mala estimación del tipo de falla, la cual es utilizada para estimar el *Zsag/Zss*.

Resultados de la estrategia basada en la regla

La regla propuesta fue aplicada al conjunto de huecos de tensión utilizado para el estudio. Los resultados son mostrados en la tabla 6.

Tabla 6 Matriz de confusión de la regla propuesta

	Clase real		Total
	TP	FP	Aguas arriba
(Predicción)	94	6	100
	FN	TN	Aguas abajo
(Predicción)	8	65	73

Se observa que los resultados obtenidos con la regla propuesta son mejores que los resultados obtenidos previamente con los algoritmos de localización. El porcentaje de clasificación es 91,9%, el cual es 26 unidades mayor que el porcentaje de clasificación del algoritmo DR (65,9%).

Estrategia basada en la forma de onda

Esta estrategia consiste en construir una modelo basado en PCA a partir de huecos de tensión con origen conocido. En este análisis el conjunto de huecos de tensión con origen aguas arriba fue seleccionado para construir el modelo, debido a que la cantidad de éstos en el conjunto de estudio es mayor que los huecos de tensión con origen aguas abajo. Por consiguiente, el objetivo es extraer la información relevante de los huecos aguas arriba útil para discriminarlos de los huecos con origen aguas abajo. La metodología tiene dos pasos generales [9]:

1) Preparación de la base de casos y construcción del modelo en el espacio de las componentes principales: Se calcula la secuencia RMS de las señales de tensión y corriente, y luego, éstas secuencias son normalizadas, después se crea el modelo PCA [12] y los huecos de tensión aguas arriba y aguas abajo son proyectados en el espacio de las componentes principales.

2) Utilización del modelo: Este paso se basa en los criterios de similitud entre los nuevos huecos de tensión y aquellos almacenados previamente (huecos aguas arriba) en la base de casos. Un hueco de tensión nuevo es proyectado en el modelo-aguas-arriba. Luego, en el espacio PCA los *k1* vecinos más cercanos son identificados basándose en los valores tomados por el estadístico Q. Después, a partir de los *k1* huecos identificados, se toman los *k2* vecinos más cercanos de acuerdo a los valores tomados por el estadístico T². Finalmente, para estimar la clase del hueco de tensión, se calcula el cociente entre las sumatoria de las distancias a los *k1* y los *k2* huecos de tensión más cercanos con respecto al nuevo hueco a clasificar. Si este cociente supera un umbral de decisión *Th*, el nuevo hueco de tensión pertenecerá a la clase de referencia (aguas arriba) [14, 15].

T^2 y Q son dos estadísticos utilizados en el Análisis de Componentes Principales para estimar la adecuación de una observación al modelo PCA (parecido con respecto al modelo). T^2 permite medir el grado de ajuste de una observación al modelo, mientras que el estadístico Q , la carencia de ajuste al modelo [16].

Resultados de la estrategia basada en la forma onda

La validación de la estrategia se realizó utilizando validación cruzada de 4 particiones y calculando la *sensibilidad* y *especificidad* para cada experimento.

El modelo-aguas-arriba fue construido. El modelo PCA explica el 95% de la variabilidad de los datos en su espacio original, resultando un modelo con las 10 primeras componentes principales.

La metodología fue verificada utilizando diferentes pares de valores para $k1$ y $k2$ y diferentes umbrales de decisión (Th) para calcular las curvas ROC de la Figura 5. Se observa que todos los experimentos tienen buen desempeño debido a que el área bajo la curva es próximo a la unidad tabla 7.

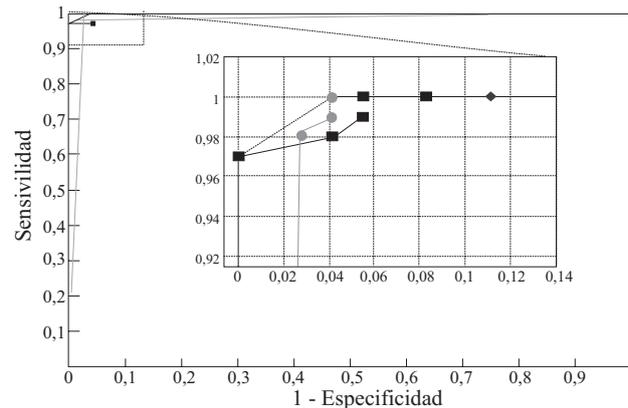


Figura 5 Curvas ROC

Tabla 7 Matrices de confusión, sensibilidad, especificidad y valores AUC de los clasificadores

Parámetros	TP	FP	FN	TN	Sensibilidad	Especificidad	AUC
$k1=15, k2=1$	100	1	0	72	1	0,986	0,993
$k1=15, k2=3$	100	3	0	70	1	0,958	0,999
$k1=15, k2=5$	100	5	0	68	1	0,931	0,998
$k1=10, k2=3$	100	2	0	71	1	0,973	0,999
$k1=10, k2=5$	100	4	0	69	1	0,945	0,990

De donde,

$k1$ y $k2$,

Cantidad de vecinos más cercanos identificados a partir de los estadísticos Q y T^2 , respectivamente.

TP, FP, FN, TN ,

Elementos de la matriz de confusión.

Sensibilidad, Especificidad,

Parámetros del clasificador obtenidos a partir de los elementos de la matriz de confusión.

AUC,

Área bajo las curvas ROC.

De acuerdo a la tabla 7, el mejor clasificador tiene $k1=10$ y $k2=3$ porque presenta el área bajo la

curva ROC más alto $AUC=0,999$. Para seleccionar el valor Th para este clasificador, se realizó una prueba con varios valores Th entre 0 y 1. La mejor *sensibilidad* y *especificidad* fue obtenida con $Th=0,2$.

Comparación de las dos estrategias

La estrategia basada en la forma de onda obtuvo el mejor resultado (98,8%), debido a que clasificó correctamente 171 huecos de tensión de un total de 173. En la tabla 8 se presenta la matriz de confusión y en ella se observa que esta estrategia presenta dos errores, mientras que la basada en la regla booleanas presenta 14 errores en la estimación del origen de los huecos de tensión.

Tabla 8 Matriz de confusión de las dos estrategias

<i>Estrategia</i>	<i>TP</i>	<i>FP</i>	<i>FN</i>	<i>TN</i>	<i>Sensibilidad</i>	<i>Especificidad</i>	<i>Acierto</i>
Regla booleana	94	6	8	65	0,922	0,910	91,9%±4,1%
Forma de onda	100	2	0	71	1,000	0,973	98,8%±1,6%

Adicionalmente los errores de clasificación son 8,1% y 1,2% para cada una de las estrategias, siendo menor la de la estrategia basada en la forma de onda. En ésta estrategia, con un intervalo de confianza del 95% (distribución normal) la incertidumbre es del 1,6% [17], lo cual significa que para otras pruebas similares es muy probable

(0,95) que se obtengan porcentajes de clasificación en el intervalo 97,2% a 100%.

A pesar de la diferencia entre los porcentajes de clasificación, cada estrategia tiene ventajas y desventajas que se deben tener presentes. En la tabla 9 se resumen algunas de ellas.

Tabla 9 Ventajas y desventajas de las dos estrategias

<i>Primera estrategia (Regla)</i>	<i>Segunda estrategia (Onda)</i>
<ul style="list-style-type: none"> - El porcentaje de clasificación es bueno. - La costo computacional es bajo. - No requiere procesos de entrenamiento y validación. - Es posible implementarla online en equipos de monitorización de energía. 	<ul style="list-style-type: none"> - El porcentaje de clasificación es excelente. - El costo computacional es alto. - Requiere de procesos de entrenamiento y validación, por tanto de datos sin procesar. - Es posible implementarla online en equipos de monitorización de energía.

Conclusiones

Se verificaron dos estrategias para la localización del origen de huecos de tensión en sistemas de distribución de energía eléctrica. Una estrategia utiliza un modelo estadístico basado en el Análisis de Componentes Principales de las formas de onda, mientras que la otra, sólo utiliza una regla booleana de tres impedancias obtenidas de las ondas de tensión y de corriente. Ambos métodos presentan resultados aceptables en la estimación del origen de los huecos de tensión. Algunas de sus ventajas y desventajas fueron identificadas.

Actualmente, la estrategia basada en la regla booleana está siendo analizada con una mayor cantidad de huecos de tensión. Con el objetivo de obtener una regla con mejor capacidad de generalización.

Agradecimientos

Los autores ofrecen sus agradecimientos a las siguientes entidades por la financiación de esta investigación: Ministerio de Educación y Ciencias de España (beca BES-2007-14942, proyecto DPI2006-09370), FIE y ENDESA DISTRIBUCIÓN.

Nomenclatura

PCSC	Algoritmo del cambio de fase de las corrientes de secuencia
RS	Algoritmo del signo de la resistencia
DR	Algoritmo del relé distancia
MANOVA	Análisis multivariante de la varianza
AD	Análisis discriminante
PCA	Análisis de componentes principales
MPCA	Análisis de componentes principales multicanal

Referencias

1. R. Chouhy Leborgne. *Voltage Sags: Single event characterisation, system performance and source location*. PhD thesis, Chalmers University of Technology. 2007. pp. 73-92.
2. C. Li, T. Tayjasanant, W. Xu, X. Liu. "Method for voltage-sag source detection by investigating slope of the system trajectory". *IEEE Proceedings-Generation, Trans. and Distr.* Vol. 150. 2003. pp. 367-72.
3. T. Tayjasanant, C. Li, W. Xu. «A resistance sign-based method for voltage sag source detection». *IEEE TPWD*. Vol. 20. 2005. pp. 2544-51.
4. N. Hamzah, A. Mohamed, A. Hussain. "A new approach to locate the voltage sag source using real current component". *Electric Power Systems Research*. Vol. 72. 2004. pp. 113-23.
5. A.K. Pradhan, A. Routray. "Applying distance relay for voltage sag source detection". *IEEE Transactions on Power Delivery*. Vol. 20. 2005. pp. 529-31.
6. A.K Pradhan, A. Routray, S. Madhan. "Fault direction estimation in radial distribution system using phase change in sequence current". *IEEE TPWDelivery*. Vol. 22. 2007. pp. 2065-2071.
7. A. Khosravi, J. Meléndez, J. Colomer. "Classification of Voltage Sags Based on MPCA Models". *Lecture Notes in Computer Science (Online)*. Vol. 4477. 2007. pp. 362-369.
8. A. Khosravi, J. Melendez, J. Colomer. "A Hybrid Method for Sag Source Location in Power Network". *9th International Conference. Electrical Power Quality and Utilisation*. Barcelona. 2007. pp. 1-6.
9. J. Meléndez, X. Berjaga, S. Herraiz, J. Sánchez, M. Castro. "Classification of Voltage Sags based on k-NN in the Principal Component Space". *International Conference on Renewable Energies and Power Quality – ICREPQ'08*. 2008. pp. 1-6.
10. V. Barrera, J. Meléndez, S. Herraiz. "Unusual voltage sag event detection in power systems" *Transmission and Distribution Conference and Exposition: Latin America*. Bogotá. 2008. pp. 1-3.
11. A. P. Bradley. "The use of the area under the ROC curve in the evaluation of machine learning algorithms". *Pattern Recognition*. Vol. 30. 1996. pp. 1145-1159.
12. E.L. Russell, L.H. Chiang, R.D. Braatz. *Data-Driven Methods for Fault Detection and Diagnosis in Chemical Processes*. London. Springer-Verlog, Berlin Heidelberg New York. 2000. pp: 33-65.
13. R. Chouhy Leborgne, D. Karlsson, J. Daalder. "Voltage Sag Source Location Methods Performance Under Symmetrical and Asymmetrical Fault Conditions". *T&D Conference and Exposition: Latin America. IEEE/PES*. 2006. pp. 1-6.
14. R. L. Mantaras, E. Plaza. "Case-based reasoning: An overview". *AI Communications*. Vol. 10. 1997. pp 21-29.
15. A. Aamodt, E. Plaze. "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches". *AI Communications*. Vol. 7. 1994. pp. 39-59.
16. Alberto J. Ferrer Riquelme, "Control estadístico multivariante de procesos por lotes (BMSPC) - Una potente herramienta estratégica para los procesos por lotes", *Revista Automática e Instrumentación*. Vol. 377. 2006. pp. 72-76.
17. J. Hernández O. M. Ramírez., C. Ferri. *Introducción a la minería de datos*. Ed. Prentice Hall. Madrid. 2004. pp. 466-467.