



UNIVERSIDAD NACIONAL DE COLOMBIA

Detección y clasificación automática de registros sísmicos en el Observatorio Vulcanológico y Sismológico de Pasto utilizando redes neuronales artificiales

Oscar Ernesto Cadena Ibarra

Universidad Nacional de Colombia
Facultad de Geociencias, postgrado Geofísica
Bogotá, Colombia
2011

Detección y clasificación automática de registros sísmicos en el Observatorio Vulcanológico y Sismológico de Pasto utilizando redes neuronales artificiales

Oscar Ernesto Cadena Ibarra

Tesis presentada como requisito parcial para optar al título de:
Magister en Ciencias - Geofísica

Directora:
Ph.D. María Cristina Dimaté Castellanos

Línea de Investigación:
Sismología

Universidad Nacional de Colombia
Facultad de Geociencias, postgrado Geofísica
Bogotá, Colombia
2011

A Lyna y Sofía.

No hay que confundir nunca el conocimiento con la sabiduría. El primero nos sirve para ganarnos la vida; la sabiduría nos ayuda a vivir.

Sorcha Carey

Agradecimientos

Quiero agradecer en primer lugar al Ing. John Jeu Meneses Muñoz del Observatorio Vulcanológico y Sismológico de Pasto-INGEOMINAS por sus aportes en el diseño y codificación del programa LAKiY. A la Ph. D. María Cristina Dimaté Castellanos de la Universidad Nacional de Colombia por su tiempo y dedicación en este proyecto. Al INGEOMINAS por facilitarme todos los recursos necesarios para el desarrollo de este trabajo.

Resumen

En este trabajo se aplica una metodología que combina Análisis de predicción lineal (LPC) y Redes neuronales artificiales (RNA) con el objetivo de detectar y clasificar tres tipos de sismos volcánicos: volcano-tectónicos (VT) tipo A, largo periodo (LP) y tremor (TRE). Se construye una herramienta computacional orientada a la detección y clasificación automática de estas señales sísmicas, generando información en tiempo real. Segmentos cortos de la traza continua se representan utilizando la técnica LPC que también cumple la función de reducir la dimensión del vector que contiene los datos de la forma de onda. Para las fases de detección y clasificación se utilizan RNA tipo perceptrón multicapa con una capa oculta. Entre las ventajas de la metodología está la posibilidad de detectar sismos con arribos emergentes, de corta duración o de bajo nivel energético. La aplicación del método muestra cerca de 100% de acierto para la distinción entre ruido y señal sísmica (fase de detección), 92 % de acierto para la clasificación entre sismos LP y VT y cerca del 100% en la detección de sismos tipo TRE.

Palabras clave: sismos volcánicos, detección automática, clasificación automática, predicción lineal, redes neuronales artificiales.

Abstract

In this work is applied a methodology that to combine Linear Prediction Coding (LPC) and Artificial Neural Networks (ANN) in order to detect and classify three types of volcanic seismic signals: Type A Volcano-Tectonic events (VT), Long Period (LP) and volcanic Tremor (TRE). It was built a software oriented to detection and automatic classification of these kinds signals, generating information in Real-Time. Short segments of the continuous streams are represented using the LPC technique, also carry out the reduction of the vector containing waveform data. For the detection and classification processing-phases are used ANN type Multi-Layer Perceptron (MLP). Among the advantages of this methodology are the possibility of detect seismic events with emergent arrives, short duration or low energetic level. The application of this method shows near to 100 % of good choices for to distinguish between signal and noise level (detection phase); 92% of good choices for classification between LP and VT events; and near to 100 % for TRE-type seismicity detection.

Keywords: volcanic seismicity, automatic detection, automatic classification, linear prediction coding, artificial neural networks.

Contenido

	Pág.
Resumen	IX
Lista de figuras	XIII
Lista de tablas	XVII
Introducción.....	1
1. Sismicidad en el volcán Galeras.....	3
1.1 La red sismológica del volcán Galeras	3
1.2 Clasificación de sismos en el VG.....	5
1.2.1 Sismos volcánicos de fuente transitoria.....	6
☒ Sismos volcano-tectónicos (VT)	6
1.2.2 Señales sísmicas volcánicas continuas.....	12
☒ Tremor volcánico (TRE)	12
☒ Tremor armónico.....	13
☒ Ruido.....	15
2. Coeficientes de predicción lineal (LPC)	16
2.1. Modelo de análisis por LPC	16
2.2. Estimación de los coeficientes de predicción lineal (Jackson, 1989)	18
2.2.3. Método de autocorrelación.....	18
3. Redes neuronales artificiales	20
3.1. La Neuron Artificial.....	20
3.1.1. Entradas y salidas.....	21
3.1.2. Pesos sinápticos	21
3.1.3. Regla de propagación.....	21
3.1.4. Función de activación	22
3.1.5. Función de salida	22
3.2. Arquitectura de las Redes Neuronales	22
3.3. Definición de red neuronal artificial	23
3.4. Operación de las redes neuronales artificiales	24
3.4.1. Fase de aprendizaje	24
3.4.2. Aprendizaje supervisado	24
3.4.3. Aprendizaje no supervisado.....	24
3.4.4. Retropropagación del error.....	25
3.4.5. Fase de recuerdo.....	26
3.4.6. Memoria asociativa.....	26
3.5. Capacidad de generalización	29

3.6.	Validación cruzada.....	30
3.7.	El Perceptrón multicapa.....	31
3.8.	Número de ejemplos de entrenamiento	32
4.	Metodología	34
4.1.	Detección y clasificación.....	34
4.1.1.	Tratamiento previo de las señales, pre-procesamiento.....	35
4.2.	Representación de los registros	37
4.2.1.	LPC para la representación de los registros	37
4.2.2.	Elección del tamaño de las unidades	38
4.3.	Desarrollo e implementación de la herramienta computacional.....	38
4.3.1.	El programa LAKiY	39
4.3.2.	Implementación de una base de datos.....	43
4.4.	Determinación de topologías adecuadas para la detección y clasificación.....	43
4.4.1.	Selección de los datos.....	43
4.4.2.	Generación de las matrices de entrenamiento	45
4.5.	Validación de resultados	46
5.	Resultados	47
5.1.	Representación de las señales	47
5.2.	Elección de la topología de las RNA.....	50
5.2.1.	Unidades de Sismos vs unidades de ruido.....	50
5.2.2.	Clasificación VT vs LP.....	52
5.2.3.	Clasificación de Tremor.....	57
5.3.	Validación de resultados	57
5.3.1.	Unidades de sismos vs. unidades de ruido y tremor	57
5.3.2.	Clasificación LP vs VT.....	61
6.	Discusión.....	64
7.	Conclusiones y recomendaciones.....	66
7.1.	Conclusiones.....	66
7.2.	Recomendaciones.....	66
A.	Anexo: Guía para el uso de LAKiY versión 1.0	68
	Instalación.....	68
	El menú principal	68
	Generar matrices de entrenamiento	69
	Entrenar RNA.....	71
	Ejecución.....	72
	Modo Fuera de línea.....	72
	En línea – tiempo real y tiempo pasado.....	74
B.	Anexo: Estructura de la base de datos.....	78
	Archivos adjuntos.....	81
	Bibliografía.....	82

Lista de figuras

	Pág.
Figura 1-1: Red sismológica del VG.....	4
Figura 1-2: Sismograma, espectrograma y espectro (de arriba hacia abajo) de un evento tipo VT-A registrado por la componente EW de la estación Urcunina, ubicada a 2.2 Km al sur del foco del sismo. El eje de las ordenadas, está en cuentas (sin respuesta instrumental) para el sismograma; en el espectrograma está en Hz; el eje de las ordenadas en el espectro es de potencia. Esta convención servirá para todas las figuras de este capítulo.....	6
Figura 1-3: Sismogramas (izquierda) y espectros (derecha) de un evento tipo VT-A registrado por las componentes verticales de algunas estaciones de la red del VG ordenados de arriba hacia abajo según la distancia fuente-estación foco.	7
Figura 1-4: Sismogramas de un evento tipo VT-B en las estaciones, Anganoy y Cufiño.....	8
Figura 1-5: Detalles del registro de la componente NS de la estación Cufiño, del sismo presentado en la Figura 1-4. Se muestra el sismograma, espectrograma y espectro de la señal.	8
Figura 1-6: Sismograma, espectrograma y espectro de un evento tipo LP en la componente vertical de la estación Urcunina.	9
Figura 1-7: Sismogramas (izquierda) y espectros (derecha) de un evento tipo LP registrado en las estaciones Anganoy, Cráter, Cufiño, Urcunina, Condor y Nariño, ordenadas de más cercana a más lejana del cráter principal. Obsérvese la diferencia en los tiempos de arribo a medida que las estaciones se alejan del cráter y la forma en que varía el arribo de la onda P.	10
Figura 1-8: Sismograma de un evento tipo Tornillo en la estación Anganoy. Obsérvese el lento decaimiento de sus ondas de coda y el pico dominante en los 1.38 Hz.	11
Figura 1-9: Sismogramas de un evento tipo HYB en las estaciones Cráter-2, Urcunina NS y Calabozo EW. Obsérvese que el inicio del evento está dominado por altas frecuencias y su coda por armónicos de baja frecuencia.....	11
Figura 1-10: Detalles en la componente NS de la estación Urcunina para el sismo mostrado en la Figura 1-9. Sismograma (superior), espectrograma (centro) y espectro (inferior).....	12
Figura 1-11: Segmento de tremor volcánico registrado en la estación Anganoy.	13
Figura 1-12: Tremor armónico registrado por la estación Anganoy. Sismograma (superior), espectrograma (centro) y espectro (inferior). Obsérvese los armónicos que sobresalen en su espectro.....	13
Figura 1-13: Sismograma y espectro de un tremor armónico registrado por la estación Anganoy.....	14

Figura 1-14: Sismograma y ampliación de un segmento de un episodio de tremor armónico producido por en el VG. Se observan modulaciones en la señal ocasionadas por armónicos con frecuencias cercanas entre sí.	14
Figura 1-15: Sismograma y ampliación de un segmento de un episodio de tremor armónico producido por en el VG. Se observan modulaciones en la señal ocasionadas por armónicos con frecuencias cercanas entre sí.	15
Figura 2-1: Flujo de la señal desde su excitación hasta su salida.....	16
Figura 2-2: Elementos involucrados en la extracción de los LPC.....	16
Figura 3-1: La neurona artificial.....	21
Figura 3-2: Red neuronal artificial.....	23
Figura 3-3: Esquema de descenso por el gradiente para una función de error. La esfera azul indica el punto donde se evalúa el error y la flecha la dirección hacia donde se debe mover la esfera para alcanzar un mínimo de la función (Tomado de JOONE, 2007).....	25
Figura 3-4: Errores global y de generalización en un problema ideal, la flecha verde indica el momento en que el aprendizaje es óptimo.....	30
Figura 3-5: La función escalón o Heaviside.....	31
Figura 3-6: Funciones de activación utilizadas en las neuronas del MLP. (a) a la izquierda la correspondiente a la ecuación 3.30 y a la derecha a la ecuación 3.31.....	32
Figura 4-1: Elementos que intervienen en la generación del sismograma.....	34
Figura 4-2: Localización aproximada de la sismicidad de fluidos, C1, C2 y C3 representan el recorrido de un rayo entre la fuente y las respectivas estaciones.	35
Figura 4-3: Localización de dos sismos de fractura, obsérvese que los recorridos a cada estación son diferentes.	35
Figura 4-4: Sismograma y espectro de un segmento muestreado a 100 Hz. (izquierda) y del mismo segmento re-muestreado a 70 Hz. (derecha).	36
Figura 4-5: Ventana de Hamming para 300 muestras.....	37
Figura 4-6: Segmento sísmico sin ventanear (Izquierda), Segmento sísmico ventaneado (Derecha).	37
Figura 4-7: Estructura del programa LAKiY.....	39
Figura 4-8: Esquema de salvado de diferentes configuraciones de red a medida que avanza el proceso de entrenamiento, en este ejemplo se guarda un clon cada 5 iteraciones.	41
Figura 4-9: Esquema de extracción y conversión de señales desde la base de datos del OVSP.	45
Figura 5-1: Varianza vs. orden de LPC para una muestra de 40 unidad de sismos, cada curva representa una unidad.	47
Figura 5-2: Varianza vs. orden de LPC para una muestra de 40 unidades de ruido, cada curva representa un unidad.....	47
Figura 5-3: Comportamiento de la varianza promedio de la muestra de unidades de sismos y de ruido respecto al orden de la predicción.....	49
Figura 5-4: Comportamiento de las diferencias entre varianzas promedio sucesivas de la muestra de unidades de sismos y de ruido respecto al orden de la predicción. Obsérvese que después del orden 5 la curva es asintótica para el caso de las unidades de sismos.	49

Figura 5-5: Comportamiento de los errores mínimos global y de generalización de cada entrenamiento.....	52
Figura 5-6: Curvas de error global y de generalización para el entrenamiento SRS34.....	53
Figura 5-7: Diagrama de frecuencias que muestra la cantidad de sismos VT por duración, se ha tomado un valor por cada múltiplo de 3 segundos.....	54
Figura 5-8: Diagrama de frecuencias que muestra la cantidad de sismos LP por duración, se ha tomado un valor por cada múltiplo de 3 segundos.....	54
Figura 5-9: Errores mínimos de generalización y error global en la etapa de entrenamiento.....	56
Figura 5-10: Variaciones del error global y de generalización con el número de iteraciones para el entrenamiento con RNA de topología 35 - 3 - 1.....	56
Figura 5-11: Diagrama de frecuencias que muestra la cantidad de sismos tipo TRE por duración.....	57
Figura 5-12: Resultados de la ejecución de la RNA09 con 50 ejemplos, 25 unidades de ruido y 25 unidades de sismos. Los puntos sobre las líneas representan las diferencias entre el objetivo y el valor calculado por la RNA. En la gráfica estas diferencias se miden alrededor de 0 para las unidades de ruido y alrededor de 1 para las unidades de sismos. En la tabla se pueden observar las diferencias exactas en la última columna para algunas unidades. Para este caso las curvas son líneas casi rectas indicando un buen desempeño de la red.	59
Figura 5-13: Valores absolutos de las diferencias obj-cal para las 50 unidades ejecutadas por las redes RNA09 (eje izquierdo) y RNA10 (eje derecho).....	60
Figura 5-14: Trazas utilizadas para ejecución en línea de las redes pre-seleccionadas, la marca azul corresponde al segmento de traza que contiene 7 unidades para el sismo 1 registrado el 18 de febrero de 2011 (arriba) y 16 para el sismo 2 registrado el 21 de febrero de 2011 (abajo).....	60
Figura 5-15: Resultados de la ejecución de la RNA15 con 50 ejemplos, 25 LP y 25 VT. Los puntos sobre las líneas representan las diferencias entre el objetivo y el valor calculado por la RNA. En la gráfica estas diferencias se miden alrededor de 0 para el tipo VT y alrededor de 1 para los sismos LP. Para este caso los puntos rojos número 1 y 6 fueron los desaciertos de la red. Existen valores cuyas diferencias no sobrepasan ± 0.5 , por tanto no se consideran errores.	63
Figura A-1: Panel inicial y menú principal del programa LAKiY.....	68
Figura A-2: Contenido de la pestaña Entrenamiento del menú principal.....	69
Figura A-3: Contenido de la pestaña Ejecución del menú principal.....	69
Figura A-4: Ventana de configuración de parámetros para generar matrices de entrenamiento.....	70
Figura A-5: Ventana de configuración de los parámetros de entrenamiento.....	71
Figura A-6: Información de salida de entrenamiento.....	72
Figura A-7: Ventana de ejecución en el modo fuera de línea.....	73
Figura A-8: Información de salida del modo ejecución fuera de línea.....	74
Figura A-9: Ventana de configuración del modo en línea tiempo real.....	75
Figura A-10: Ventana del modo en línea tiempo pasado.....	76
Figura A-11: Información de salida del modo ejecución en línea.....	77

Figura B-1: Diagrama de la tabla Proyecto.....	78
Figura B-2: Diagrama de la tabla Matriz entrenamiento.....	78
Figura B-3: Diagrama de la tabla Objetivos matriz.....	79
Figura B-4: Diagrama de la tabla Entrenamiento red.....	79
Figura B-5: Diagrama de la tabla Configuración estación.....	79
Figura B-6: Diagrama de la tabla Estaciones clasificación.....	80
Figura B-7: Diagrama de la tabla Set estaciones.....	80

Lista de tablas

	Pág.
Tabla 1-1: Estaciones de la red del VG y sus principales características.....	4
Tabla 4-1: Número de ejemplos extraídos.	44
Tabla 4-2: Extensión que identifica los diferentes tipos de señales.	44
Tabla 5-1: Resultados del análisis para establecer un orden de LPC.	48
Tabla 5-2: Configuración de las matrices de entrenamiento de reconocimiento entre unidades de ruido y sismos.....	50
Tabla 5-3: Valores de los errores global y de generalización para las RNA entrenadas. Se han marcado de color cian los entrenamientos con los errores mínimos y de rojo los entrenamientos con los errores máximos.	51
Tabla 5-4: Configuración de las matrices de entrenamiento para clasificación entre LP y VT.	53
Tabla 5-5: Valores mínimos de los errores global y de generalización para cada entrenamiento, las filas en color naranja indican las mejores redes con base en el criterio de mínimo error de generalización.....	55
Tabla 5-6: Resultados de la ejecución fuera de línea de las RNA para unidades de señal y de ruido.....	58
Tabla 5-7: Resumen de resultados de ejecución en línea para las cuatro RNA preseleccionadas, la columna verde corresponde a la red con mejor desempeño.	61
Tabla 5-8: Resultados de la ejecución fuera de línea de las RNA entrenadas para la clasificación entre los tipos LP y VT.	62

Introducción

El volcán Galeras (VG) es uno de los más activos de Colombia. Se encuentra localizado en el departamento de Nariño en el suroccidente colombiano en $1^{\circ}13.26' N$ y $77^{\circ}21.54' W$, con una altitud máxima de 4276m sobre el nivel del mar. Su cono activo está ubicado a 9km al occidente de Pasto, la capital del departamento, con 400.000 habitantes aproximadamente. El VG es un estrato-volcán conformado por depósitos de flujos de lava andesítica, flujos piroclásticos y depósitos de caída. Después del ciclo eruptivo entre 1942 y 1944, el VG mostró nuevamente actividad a finales de 1988; el INGEOMINAS en 1989 empezó la observación permanente de su actividad con el establecimiento del Observatorio Vulcanológico y Sismológico de Pasto (OVSP).

En las rutinas de operación del OVSP la preparación de la información sismológica como insumo para el análisis e interpretación de los niveles de actividad volcánica inicia con la generación de histogramas a partir de la clasificación de la sismicidad. Esto se lleva a cabo manualmente, analizando uno a uno los registros sísmicos, lo que implica, de una parte, una componente subjetiva importante, y de otra, una gran inversión de tiempo del personal de turno. Estas dificultades son especialmente agudas en épocas de crisis volcánica cuando se registran centenares de sismos diariamente.

La generación de técnicas de detección y clasificación automáticas es necesaria para el monitoreo volcánico, especialmente en épocas de crisis, en los que es imperativo tomar decisiones rápidas y efectivas para informar de manera oportuna a las personas que habitan en la zona de influencia del volcán. Un procedimiento de clasificación automática robusto permitiría al personal de turno liberar parte del tiempo del trabajo de clasificación, sin comprometer la precisión de los resultados. De esta manera, ese tiempo podría ser utilizado en un procesamiento orientado a la interpretación del fenómeno y la toma de decisiones.

Una dificultad en la detección de la sismicidad volcánica es que con frecuencia esta se manifiesta con eventos de bajo contenido energético, con arribos emergentes o con una variedad de formas en sus envolventes, esto hace que la detección de sismos sea complicada con algoritmos clásicos como el STA/LTA (*short term average/long term average*). Por otro lado, hay casos en los que varios temáticos tienen opiniones diferentes sobre la clasificación de un mismo sismo.

El objetivo de este estudio es diseñar una metodología basada en la combinación de Análisis de predicción lineal (LPC) y Redes neuronales artificiales (RNA) e implementar una herramienta computacional que permita detectar y clasificar los sismos del VG en tiempo real. La representación de las señales por medio de LPC se justifica porque reduce la dimensión de los vectores de entrada al sistema clasificador. RNA adecuadamente entrenadas son utilizadas para la clasificación.

En la fase de detección se diferencia entre cortos segmentos de señales asociadas con sismos y ruido y en una fase posterior se realiza la clasificación en tres tipos de eventos VT, asociados con fracturamiento de material cortical; LP, relacionados con movimiento de fluidos de fuente transitoria y tremor (TRE), que corresponde a movimiento de fluidos de fuente persistente en el tiempo. También es posible detectar eventos de bajo contenido energético.

Este trabajo no tiene en cuenta sismos no volcánicos, ni segmentos de ruido que haya sido originado por factores como tormentas eléctricas, inducción de voltajes no deseados, etc.

Los investigadores del observatorio del volcán Vesubio (Del Pezzo *et al.*, 2003) fueron los primeros que combinaron los LPC con las RNA para dar solución al problema de clasificación en sismología volcánica. En su trabajo se clasifican *a posteriori* (no en tiempo real) sismos VT y explosiones en canteras. Los trabajos de clasificación de sismos para el VG de Rugeles & Cadena (2006) y Cadena & Arcos (2008) contienen parte de la metodología utilizada en la presente investigación; en el primero se clasifica entre sismicidad tipo LP y VT, mientras que en el segundo se discrimina entre segmentos cortos de ruido y segmentos cortos de sismos de los tipos VT y LP; en ambos trabajos se clasifica *a posteriori* y se utiliza software comercial.

Este documento está organizado de la siguiente manera: en el Capítulo 1 se tratan los aspectos generales de la sismicidad volcánica, clasificación de la sismicidad en el VG y una descripción de las señales sísmicas en los dominios del tiempo y la frecuencia, en el Capítulo 2 se presentan los fundamentos del modelo de análisis por predicción lineal, la estimación de los coeficientes de predicción y su aplicación en señales sísmicas, el Capítulo 3 concierne a las redes neuronales artificiales, el tipo de problemas a los cuales es aplicable esta herramienta, los tipos de arquitectura de RNA, las fases de aprendizaje y de recuerdo, la generalización del aprendizaje y la RNA perceptrón multicapa, que es el tipo de red utilizada en este estudio, en el Capítulo 4 se detalla la metodología aplicada, en el Capítulo 5 se presenta los resultados y su validación, en el Capítulo 6 se realiza una discusión basada en la metodología y los resultados obtenidos y, finalmente, en el Capítulo 7 se consignan las conclusiones y recomendaciones del trabajo.

1. Sismicidad en el volcán Galeras

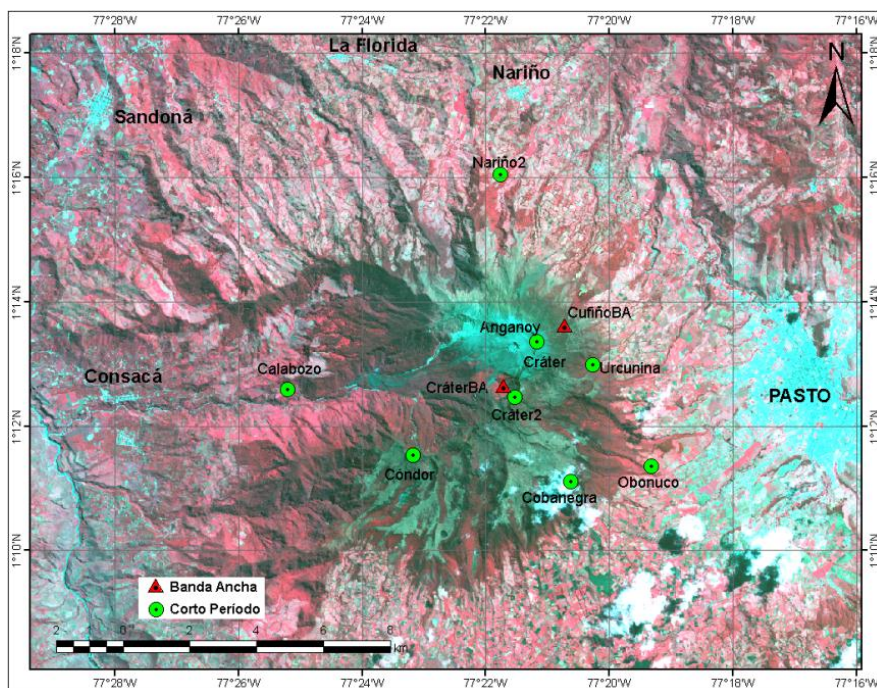
La sismicidad es un reflejo de procesos asociados con liberación de energía en el interior de la Tierra, en regiones volcánicas estos procesos obedecen principalmente a la dinámica del magma. Su movimiento, cambios de fase, cambios de temperatura, desprendimiento de volátiles, son algunos de los fenómenos que producen cambios de presión que son transmitidos por la roca y registrados por los sismómetros. En ocasiones estos cambios de presión son transitorios y en otros se sostienen produciendo sismos de mayor duración. Suele suceder que dichas presiones son suficientes para rebasar el nivel de fragilidad de la roca, fracturándola, originando sismos bastante similares a los sismos tectónicos.

Los sismómetros captan el movimiento del terreno cuya señal dibuja sobre los sismogramas diferentes formas denominadas firmas o registros sísmicos. Cada firma es diferente de otra, sin embargo, existen características comunes que se pueden encontrar en sismos originados por condiciones similares de fuente, de manera que un análisis de los registros y una clasificación que atienda a la fuente que los produjo son claves para lograr una adecuada interpretación del fenómeno volcánico.

Entre 1989 y 2004, la clasificación de eventos en el OVSP (Gómez *et al.*, 2005) se basó en las clasificaciones de Lahr *et al.*, (1989), Latter, 1979, Minakami, (1960, 1974), Malone, (1983), Newhall, (1984) y Chouet (1981, 1988, 1992 1996), aunque se incluyeron algunas particularidades propias del VG como el tipo *Tornillo* y el tipo *Mariposa*. Como herramientas de clasificación para la elaboración del catálogo se tuvo en cuenta registros en sismogramas de papel, así como también registros digitales; en cuanto al análisis espectral aplicado se utilizaron espectros y espectrogramas basados en la transformada discreta de Fourier. Adicionalmente se compararon eventos sísmicos con resultados de estudios que proponen modelos de fuente para cada tipo de sismicidad (Chouet, 1981, 1988, 1992 1996).

1.1 La red sismológica del volcán Galeras

La red sismológica instalada en el VG ha tenido algunas variaciones desde su despliegue inicial en el año 1989, actualmente está constituida por sensores tanto de Banda ancha (BA) como de Corto periodo (CP), se cuenta con cinco sensores triaxiales y cuatro de componente vertical. Los sensores de CP utilizados son tipo L4C y L4C-3D, mientras que los de BA son tipo Guralp de 60 y 120 segundos. Todos estos sensores son transductores de velocidad. La disposición de los instrumentos está fuertemente condicionada por la topografía de la zona (Figura 1-1 y Tabla 1-1). Las señales de los sensores de BA se digitalizan en campo y se transmiten teleméricamente a través de radios de doble vía y espectro ensanchado. La información que registran los sensores de CP es enviada a través de la banda UHF y se digitaliza en la sede del OVSP, con excepción de las estaciones Calabozo y Cráter2 cuya información se digitaliza en campo.

Figura 1-1: Red sismológica del VG.

Tabla 1-1: Estaciones de la red del VG y sus principales características.

Estación	Tipo de datos a transmitir	Tipo de sensor	Componentes	Distancia al cráter (km)	Altitud (msnm)
Anganoy	Analógico	Corto periodo	Vertical	1	4227
Cráter-2	Analógico	Corto periodo	Vertical	1.4	4063
Cráter-2	Digital	Banda ancha	Triaxial	1.5	4052
Cufiño	Digital	Banda ancha	Triaxial	1.8	3828
Urcunina	Analógico	Corto periodo	Triaxial	2.8	3494
Cobanegra	Analógico	Corto periodo	Vertical	4.8	3625
Obonuco	Analógico	Corto periodo	Triaxial	5.9	3021
Cóndor	Analógico	Corto periodo	Vertical	4.4	3977
Nariño-2	Analógico	Corto periodo	Vertical	4.8	2846
Calabozo	Digital	Corto periodo	Triaxial	6.7	2313

La digitalización se realiza mediante tarjetas National Instruments, Data Translation y digitalizadores Guralp. La adquisición se hace con los programas Scream© (Guralp Systems Ltd.), Earthworm (EW) del United States Geological Survey (USGS) y TJ (Raigosa, 2008). El sistema EW funciona como sistema de adquisición principal, el sistema TJ funciona como un sistema de respaldo. La información de las señales digitalizadas en campo, por los digitalizadores Guralp, se envían en paralelo a los sistemas principal y de respaldo. Las trazas se guardan en formato *SUDS*, con archivos de dos minutos de duración; cada archivo contiene la información de todas las estaciones.

En el caso de las estaciones de BA el fabricante provee las curvas de respuesta y dado que la información está digitalizada no se ve afectada por la telemetría. Por otro lado en el caso de los sensores de CP es necesario hacer pruebas desde el sitio de la estación debido a que la telemetría afecta la información que se envía, la razón de esto es que la señal no viaja digitalizada y los componentes del sistema de telemetría tales como el *VCO (Oscilador controlado por voltaje)* o la pareja de radios (transmisor y receptor) afectan la señal. Periódicamente se realizan pruebas de respuesta y se generan curvas para su respectiva corrección.

1.2 Clasificación de sismos en el VG

En los volcanes del mundo se presentan sismos que tienen *firmas* semejantes y que se han asociado con el tipo de fuente que los origina. Por ejemplo, la fractura de material cortical en el edificio volcánico, sin que importe la causa que la produjo, genera eventos sísmicos que tienen arribos impulsivos y sus fases P y S son fácilmente reconocidas, estos sismos son llamados *volcano-tectónicos* (Latter, 1979; Lahr *et al.*, 1994 y Chouet, 1996), se distinguen por las siglas VT. Puede suceder que los sismos VT sean producto de actividad volcánica o de acumulación de esfuerzos tectónicos que se transmiten hacia la región volcánica; el hecho de que su foco esté localizado en inmediaciones de la estructura volcánica es suficiente para clasificarlo como VT.

Se observan también eventos sísmicos volcánicos relacionados con la dinámica del sistema magma-gas e hidrotermal (Chouet, 1981, 1988, 1992 y 1996). Este tipo de sismos se divide en dos grupos diferenciados por el tiempo de excitación de la fuente generatriz. Se llaman sismos de *largo periodo* o LP aquellos en que su fuente es transitoria y *tremor* o *TRE* cuando la fuente es persistente en el tiempo. Chouet considera que una concatenación de LP puede ser tratada como un solo episodio de *tremor*. Otro tipo de sismos involucra tanto fractura de material sólido como movimiento de fluidos, estos son llamados *Híbridos* o HYB.

En algunos volcanes se observan también tipos propios, que responden a particularidades en su estructura o dinámica, por ejemplo los llamados *tornillos* o TO (Gómez *et al.*, 1997, 1999) registrados en el VC y en contados volcanes del mundo.

La clasificación de sismos en el VG ha sufrido modificaciones. Cuando se inició el monitoreo sismológico del volcán y teniendo en cuenta que no se contaba con una línea base de información sobre la cual referir la sismicidad, se decidió adoptar la clasificación que hasta el momento se utilizaba en el volcán Nevado del Ruiz, basada en la de Minakami, (1960, 1974) y Latter, (1979), y modificada por Malone, (1983) y Newhall, (1984), que era similar a la que se utilizaba en el volcán Santa Helena. Luego, con los estudios de Lahr *et al.*, (1994) y Chouet, (1981, 1988, 1992, 1996), se adoptó la clasificación de: largo periodo, volcano-tectónico, *tremor armónico* y se incluyeron además el *tremor espasmódico*, el *tornillo* y el *mariposa* (Gómez *et al.*, 2005).

Las definiciones que se describen a continuación se han tomado del *Catálogo de señales sísmicas volcánicas de Colombia* (Gómez *et al.*, 2005), que describe los tipos de sismos volcánicos dentro de una división superior que contiene: señales sísmicas volcánicas transitorias y señales sísmicas volcánicas continuas.

1.2.1 Sismos volcánicos de fuente transitoria

Hace referencia a señales cuya fuente actúa en un tiempo relativamente corto, generalmente unos pocos segundos, luego de esto se generan las ondas de coda del evento. Existen señales transitorias asociadas con fractura de material cortical y también con movimiento de fluidos, dentro de estas señales se encuentran los tipos: VT, LP, HYB y TO.

▪ Sismos volcano-tectónicos (VT)

Dentro de este tipo de sismos podemos distinguir dos clases, VT-A y VT-B:

Sismos volcano-tectónicos tipo A o VT-A

En general ocurren a profundidades entre 2 y 10 km desde la cima del volcán. En su registro en el dominio del tiempo es posible distinguir con facilidad las fases P y S. El decaimiento de la envolvente de sus ondas de coda muestra una forma exponencial disminuyendo su amplitud rápidamente con el tiempo. El espectro de este tipo de eventos contiene altas frecuencias, mayores que 5 Hz; los picos que aportan con mayor energía a la señal están generalmente en una banda entre los 6 y 8 Hz y sus ondas de coda pueden contener frecuencias significativamente energéticas, incluso hasta los 15 o 20 Hz.

En el campo de los desplazamientos el contenido espectral no supera los 6 Hz. En la (Figura. 1-2) se muestra el registro de un evento de este tipo en una de las estaciones de la red del VG y en la (Figura. 1-3) otro ejemplar en varias estaciones de la red. Esta clase de sismos corresponde a los del tipo A de Minakami (Minakami *et al.*, 1950; Minakami, 1974), a los sismos volcano-tectónicos de Latter, (1979), y a los de *tipo h* de Malone, (1983). La fase inicial de los sismos tipo VT es frecuentemente impulsiva, ocasionalmente se observan arribos emergentes en algunas estaciones como consecuencia de pérdidas de energía en la onda P o por la interacción de esta con material más fluido dentro del sistema volcánico (Latter, 1984), también influye el patrón de radiación de la fuente.

La distribución de eventos VT en el tiempo es, por lo general, a manera de enjambre, con magnitudes aleatorias, en contraste con los eventos no volcánicos que muestran una secuencia de réplicas que siguen al evento principal y van disminuyendo su magnitud en el tiempo.

Figura 1-2: Sismograma, espectrograma y espectro (de arriba hacia abajo) de un evento tipo VT-A registrado por la componente EW de la estación Urcunina, ubicada a 2.2 Km al sur del foco del sismo. El eje de las ordenadas, está en cuentas (sin respuesta instrumental) para el sismograma; en el espectrograma está en Hz; el eje de las ordenadas en el espectro es de potencia. Esta convención servirá para todas las figuras de este capítulo.

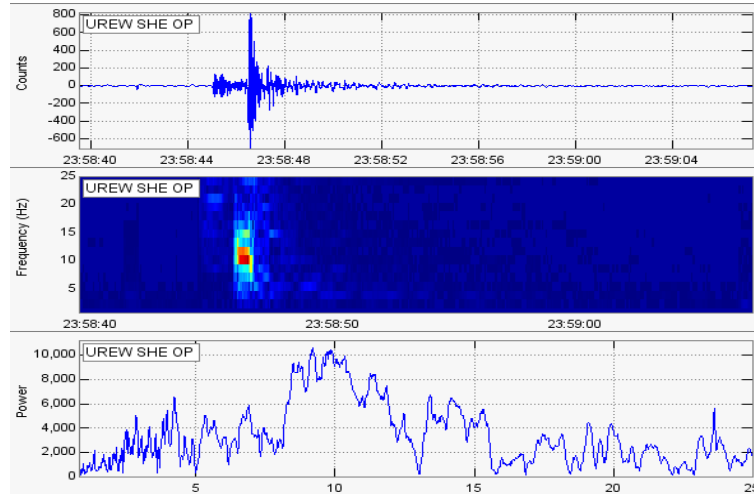
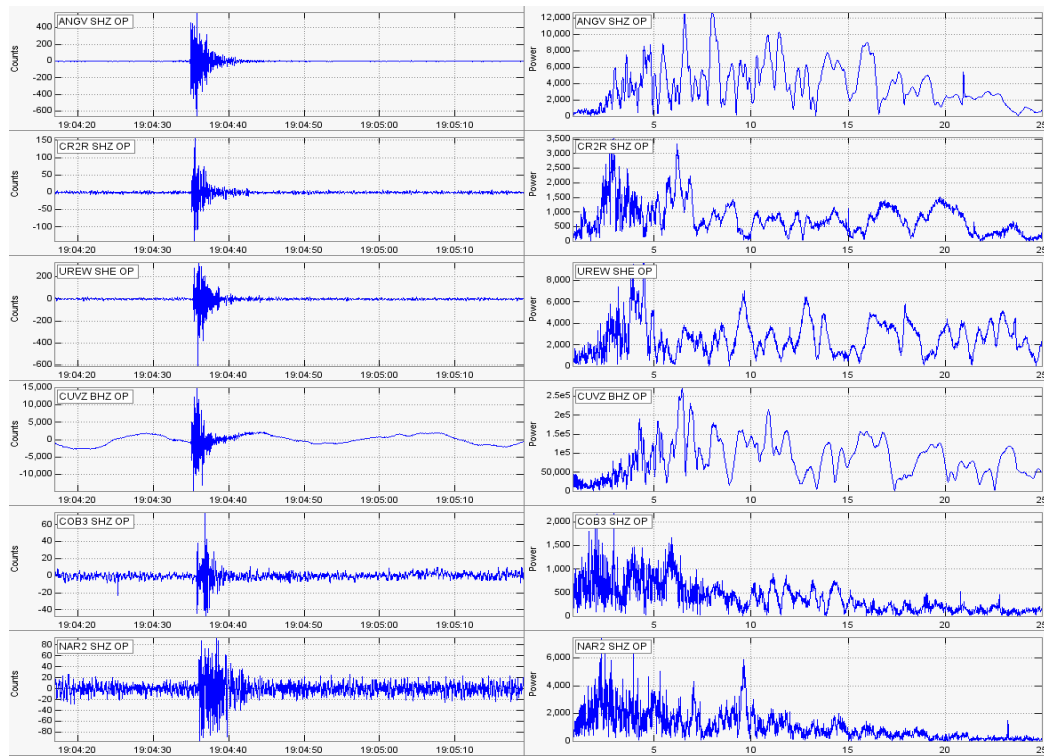


Figura 1-3: Sismogramas (izquierda) y espectros (derecha) de un evento tipo VT-A registrado por las componentes verticales de algunas estaciones de la red del VG ordenados de arriba hacia abajo según la distancia fuente-estación foco.



Sismos volcano-tectónicos tipo B o VT-B

Este tipo de eventos se presentan generalmente a profundidades menores que 2 Km. Entre las características más destacadas de su forma de onda están sus duraciones relativamente largas, el arribo de la fase P es emergente y es bastante difícil reconocer la fase S (Figuras. 1-4 y 1-5). En su espectro se observa que la energía del sismo está concentrada en una banda de bajas frecuencias, generalmente menores que 5 Hz, por otro lado, en sus ondas de coda es

posible encontrar indicios, de dispersión de ondas superficiales. Una posible explicación es que por tratarse de focos poco profundos encuentran una gran cantidad de dispersores en su trayecto que afectan las frecuencias más altas durante la propagación de la onda.

Figura 1-4: Sismogramas de un evento tipo VT-B en las estaciones, Anganoy y Cufiño.

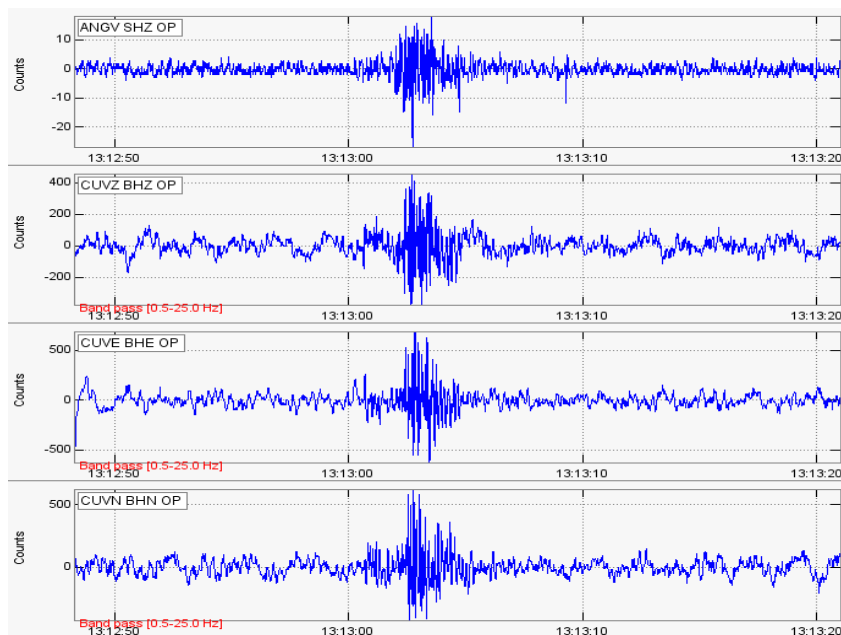
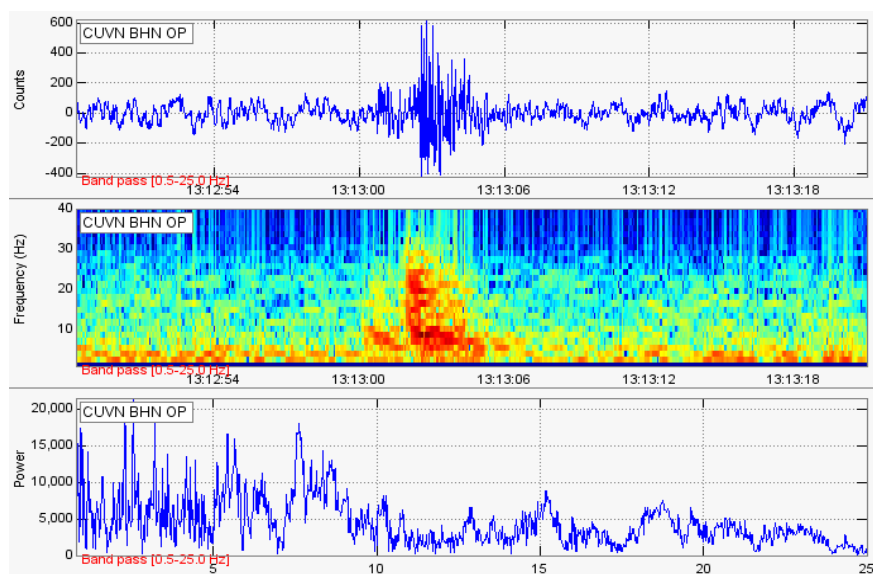


Figura 1-5: Detalles del registro de la componente NS de la estación Cufiño, del sismo presentado en la Figura 1-4. Se muestra el sismograma, espectrograma y espectro de la señal.



Se piensa que estos sismos, al igual que los volcano-tectónicos tipo A, son causados por fracturamiento de material cortical cuya fuente puede consistir principalmente en un mecanismo de doble-par de fuerzas, sin embargo la diferencia con los registros de los sismos VT-A podría reflejar efectos de camino ya que al ubicarse en capas superficiales estas contienen heterogeneidades de piroclastos y lavas, incrementando la dispersión (*scattering*). Los VT-B son los sismos clasificados como *tipo B* por Minakami *et al.*, (1950) o *tipo m* en Santa Helena por Malone, (1983).

▪ Sismos tipo Largo periodo (LP)

Los sismos LP se caracterizan porque en su firma generalmente es posible distinguir la forma de onda de sus armónicos dominantes, que generalmente se encuentran por debajo de los 5 Hz, superpuestos a estos armónicos se encuentran componentes menos energéticas que alcanzan los 13 Hz (Figura 1-6 y 1-7). Cuando el foco del evento es cercano a la estación de registro se observa un arribo impulsivo, similar al de un VT-A, que en las estaciones más lejanas se observa como emergente.

En los sismos LP el decaimiento de ondas de coda es relativamente lento por lo cual la duración de estos eventos es prolongada, es difícil distinguir la onda S. Las formas de onda son a veces similares a las del tipo VT-B. Su firma en ocasiones exhibe pulsos que pueden originarse por la superposición de armónicos de frecuencias cercanas o ser originados por vibraciones pulsátiles de la fuente misma.

Los sismos tipo LP en ocasiones muestran una señal emergente de bajo nivel energético, a modo de precursor, seguido a este se presenta el evento que es mucho más energético respecto a su precursor y con arribo impulsivo. Es común observar componentes de alta frecuencia al inicio de la señal.

Figura 1-6: Sismograma, espectrograma y espectro de un evento tipo LP en la componente vertical de la estación Urcunina.

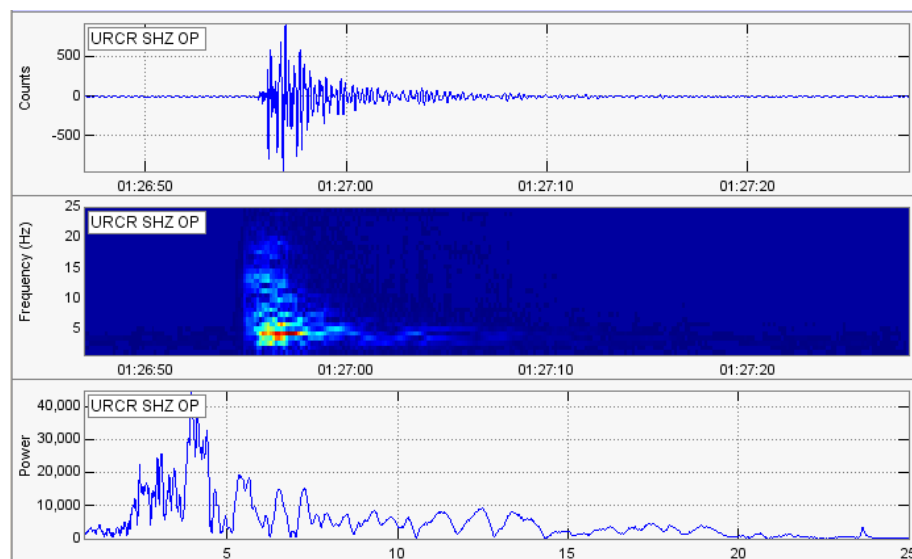
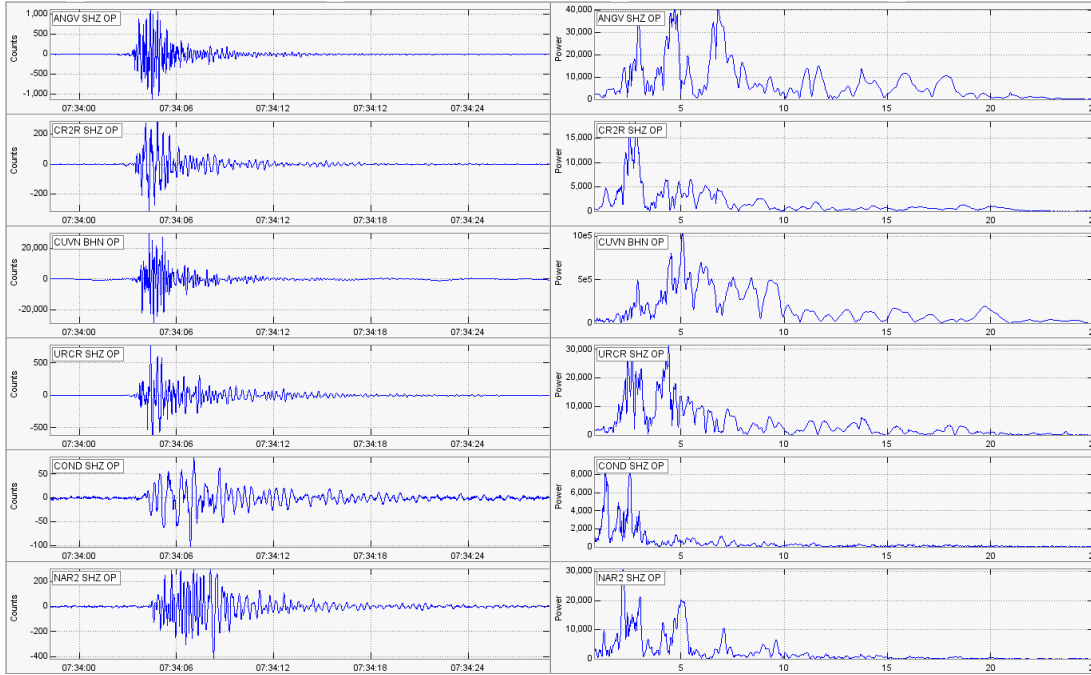


Figura 1-7: Sismogramas (izquierda) y espectros (derecha) de un evento tipo LP registrado en las estaciones Anganoy, Cráter, Cufiño, Urcunina, Condor y Nariño, ordenadas de más cercana a más lejana del cráter principal. Obsérvese la diferencia en los tiempos de arribo a medida que las estaciones se alejan del cráter y la forma en que varía el arribo de la onda P.

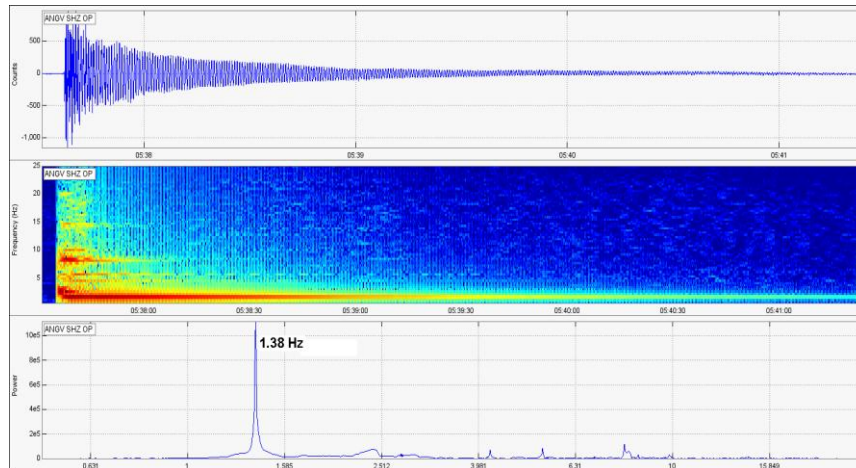


En general, los focos de estos sismos se ubican a profundidades menores que 2 km. Según Chouet, (1985, 1988) estos sismos pueden ser causados por resonancia de conductos o fracturas llenas de fluidos e iniciados por transitorios de presión. Existe una variedad de modelos para explicar la génesis de los sismos LP, estos van desde la resonancia en conductos o grietas (Chouet, 1988), hasta fenómenos de resonancia dentro del propio magma (Seidl *et al.*, 1981). Estos modelos explican gran parte de las características espectrales de este tipo de eventos.

▪ Sismos tipo Tornillo (TO)

Este nombre fue acuñado en Colombia por el personal del OVSP para distinguir un tipo especial de sismos de largo periodo cuya fuente está asociada con procesos de resonancia (Sakuraba *et al.*, 2002). Su firma muestra un inicio impulsivo tras el cual se observa que uno o varios armónicos alargan su coda por tiempos relativamente largos. La manera en que decae la amplitud del sismograma a lo largo del tiempo es lenta, lo que implica que la tasa de pérdida de energía en la cavidad resonante es lenta también. Su espectro muestra uno o pocos picos en las frecuencias de los armónicos que aportan la energía del sismo (Figura. 1-8).

Figura 1-8: Sismograma de un evento tipo Tornillo en la estación Anganoy. Obsérvese el lento decaimiento de sus ondas de coda y el pico dominante en los 1.38 Hz.



En el VG generalmente se presentan en el contexto de un sistema cerrado. Un transitorio de presión dentro de una grieta, por ejemplo, podría ser el inicio de la resonancia que dependiendo de la composición y propiedades físicas de los fluidos involucrados determinarían los armónicos dominantes en su espectro y la duración del evento. Este tipo de sismos ha precedido la mayoría de las erupciones registradas entre 1992 y 1993 (Torres *et al.*, 1996; Narváez *et al.*, 1997; Gómez & Torres, 1997) y en las erupciones recientes, entre 2004 y 2010. Es notable el decrecimiento de las frecuencias dominantes y el aumento de la duración de los tornillos en cada una de las etapas pre-eruptivas. Estas observaciones se han interpretado como el enriquecimiento de la fracción gaseosa en el fluido perturbado y el incremento del contraste de impedancias respectivamente.

▪ Sismos tipo Híbrido (HYB)

Muestran características originadas por la combinación de procesos que involucran movimiento de fluidos y fracturamiento de material cortical por lo cual se ha sugerido que se trata de la combinación de eventos tipo VT-A o VT-B con sismos tipo LP (Figura 1-9 y 1-10). Generalmente su arribo es impulsivo, dominado altas frecuencias y en muchas ocasiones es posible distinguir la fase S. Sus ondas de coda muestran patrones armónicos no dispersivos, característicos de los eventos tipo LP. Como los sismos VT, los sismos HYB muestran diferentes polaridades en sus arribos en las diferentes estaciones, mientras que los LP tienen igual polaridad en las diferentes estaciones de registro. Miller *et al.*, (1998) concluyeron que los eventos híbridos son bastante superficiales y asociados al crecimiento de domos.

Figura 1-9: Sismogramas de un evento tipo HYB en las estaciones Cráter-2, Urcunina NS y Calabozo EW. Obsérvese que el inicio del evento está dominado por altas frecuencias y su coda por armónicos de baja frecuencia.

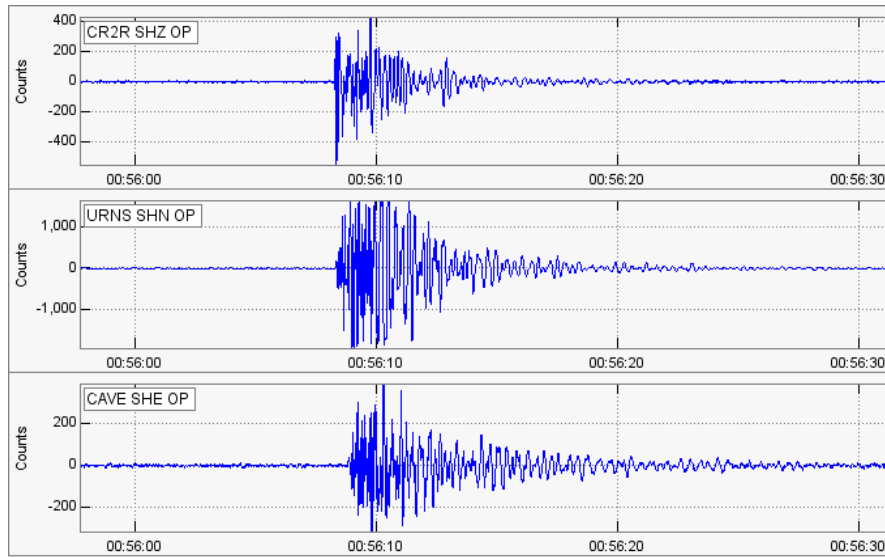
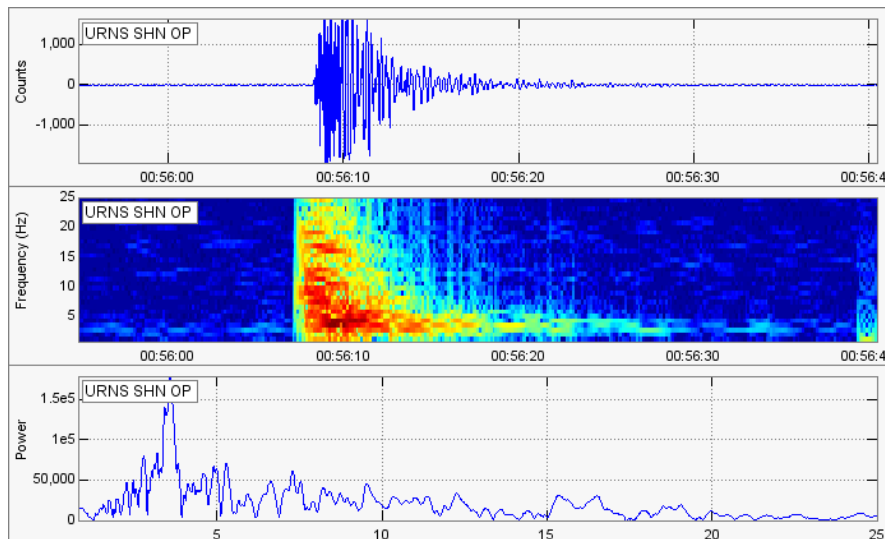


Figura 1-10: Detalles en la componente NS de la estación Urcunina para el sismo mostrado en la Figura 1-9. Sismograma (superior), espectrograma (centro) y espectro (inferior).



1.2.2 Señales sísmicas volcánicas continuas

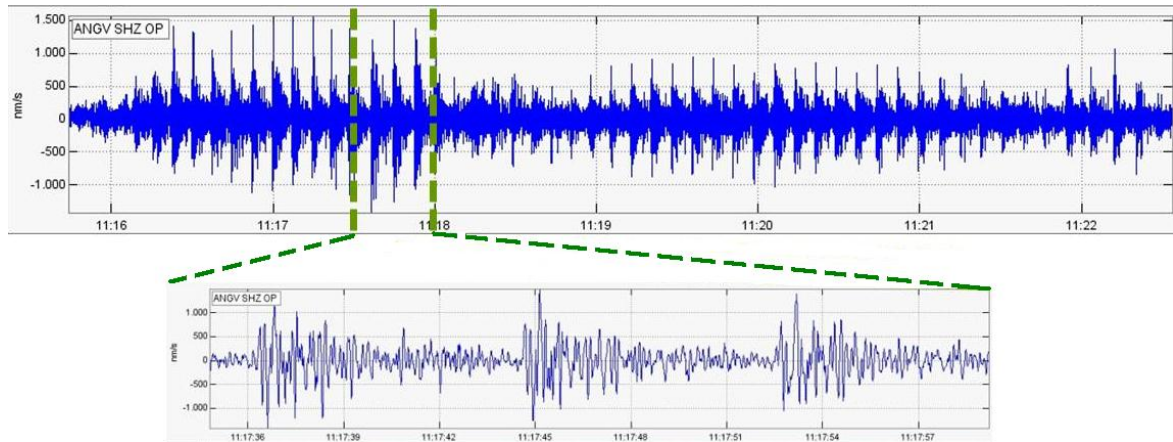
El sismograma de las señales volcánicas continuas puede estar asociado con una fuente persistente en el tiempo relacionada con dinámica de fluidos o con procesos de deslizamiento de roca o *lahares*. En este trabajo es de interés el tremor generado por actividad volcánica.

- **Tremor volcánico (TRE)**

Este tipo de sismicidad presenta una firma continua en el tiempo que refleja una excitación persistente de la fuente. El sismograma del tremor volcánico puede tener variaciones en

amplitud producidas por cambios en la energía del fenómeno que lo produce, aunque en otros casos la amplitud es estable. La duración del tremor puede variar desde unos pocos minutos hasta meses. En ocasiones es posible distinguir una secuencia de sismos, de tipo LP por ejemplo, que se concatenan formando una señal continua que también es considerada como tremor (Figura 1-11). Por sus características espectrales se puede diferenciar el tremor armónico y por su secuencia temporal el tremor espasmódico.

Figura 1-11: Segmento de tremor volcánico registrado en la estación Anganoy.



▪ Tremor armónico

Su principal característica es la presencia dominante de uno o varios armónicos, mostrando en su espectro picos muy angostos en sus respectivas frecuencias (Figura 1-12). Aunque en general los picos dominantes están por debajo de los 10 Hz, suelen ser comunes armónicos con frecuencias mayores, inclusive hasta 20 Hz (Figura 1-12). La amplitud en las crestas de sus ondas puede ser relativamente estable (Figura 1-13) o mostrar modulaciones suaves debido a la cercanía de sus armónicos en el espectro (Figura 1-14).

Las observaciones más comunes de tremor armónico se han hecho en volcanes en los que la lava es de alta viscosidad, varios autores proponen como fuente un medio resonante con un factor de calidad Q relativamente alto sobre el que actúa un mecanismo de retroalimentación preciso (Hellweg *et al.*, 1994; Schlindwein *et al.*, 1995). El tremor armónico es considerado como uno de los eventos premonitorios de erupciones en el VG ya que ha estado presente días o semanas antes de erupciones. Una hipótesis sobre su origen es que en el contexto de actividad pre-eruptiva el tremor armónico se relaciona con recarga de material magmático hacia niveles más superficiales.

Figura 1-12: Tremor armónico registrado por la estación Anganoy. Sismograma (superior), espectrograma (centro) y espectro (inferior). Obsérvense los armónicos que sobresalen en su espectro.

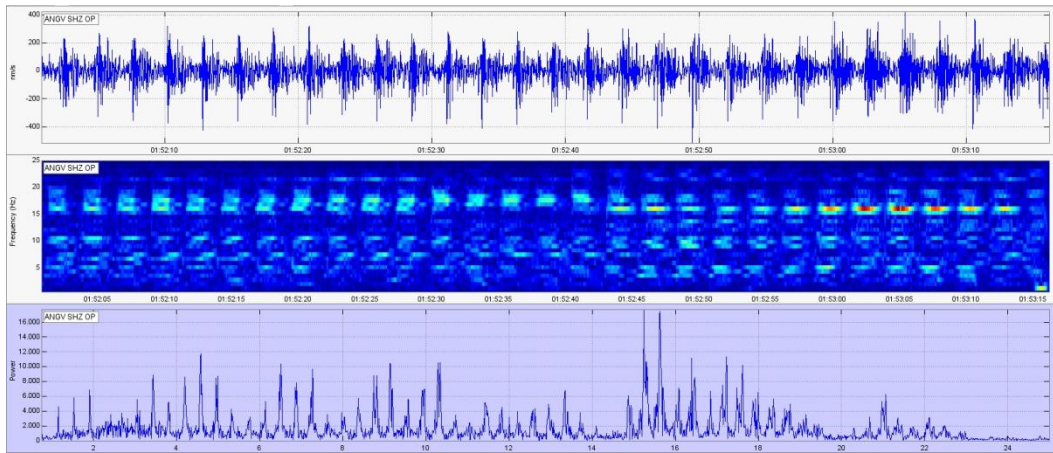


Figura 1-13: Sismograma y espectro de un tremor armónico registrado por la estación Anganoy.

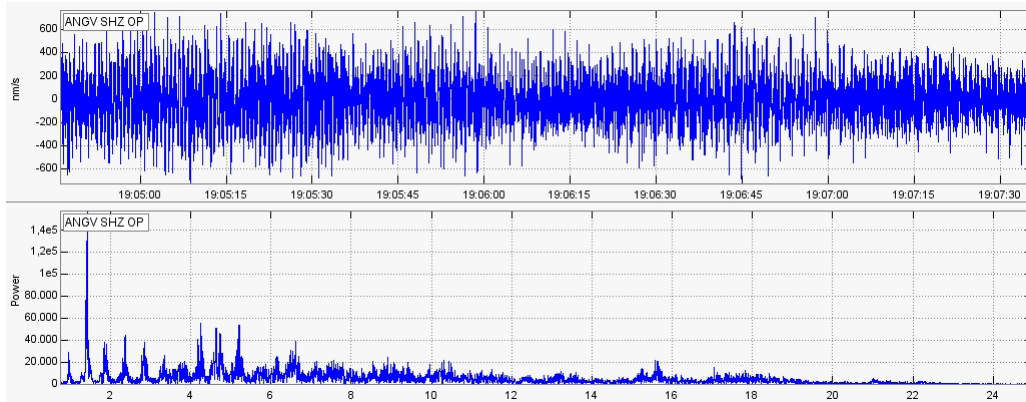
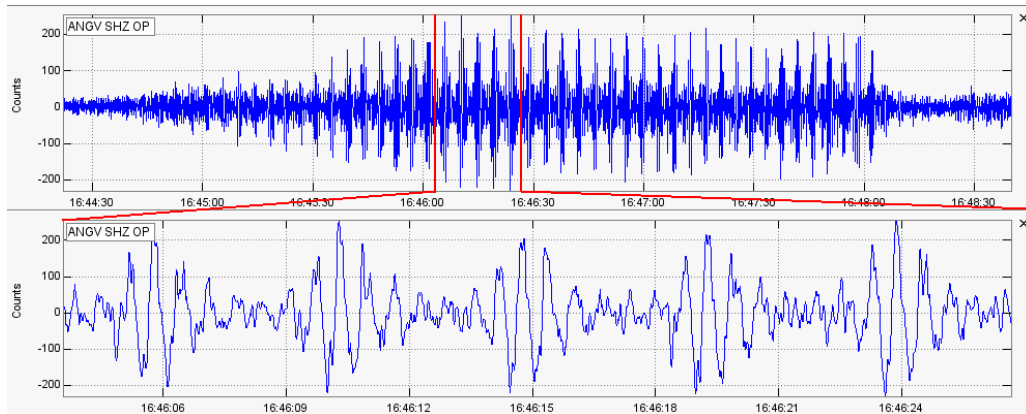


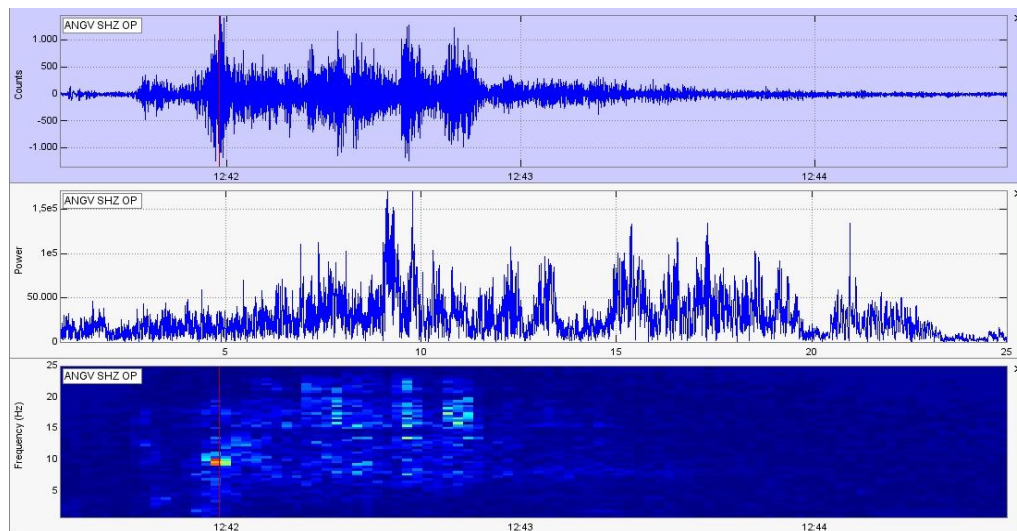
Figura 1-14: Sismograma y ampliación de un segmento de un episodio de tremor armónico producido por en el VG. Se observan modulaciones en la señal ocasionadas por armónicos con frecuencias cercanas entre sí.



▪ Tremor espasmódico

En este tipo de tremor la amplitud varía significativamente con el tiempo (Figura 1-15), el calificativo de espasmódico hace referencia a la aleatoriedad de la distribución de sus frecuencias, las cuales no presentan armónicos claramente dominantes. En su contenido espectral se observan bandas amplias que pueden ocupar cualquier sector del espectro hasta los 25 Hz (Figura 1-15). En el VG este tipo de tremor se ha correlacionado con emisión de gases o ceniza, suele ser común que los episodios de incremento de presión en el sistema volcánico culminen con el advenimiento de tremor espasmódico y la consecuente liberación de presiones. En concordancia con lo observado el contenido espectral del tremor espasmódico sugiere la interacción entre fluido y volátiles dentro del sistema.

Figura 1-15: Sismograma y ampliación de un segmento de un episodio de tremor armónico producido por en el VG. Se observan modulaciones en la señal ocasionadas por armónicos con frecuencias cercanas entre sí.



▪ No clasificables

Sismos que por su bajo contenido energético es difícil establecer su origen y por consiguiente de incluirlos en alguno de los tipos mencionados. Sin embargo son un recurso importante en el análisis de la actividad volcánica ya que nos dan una idea de la recurrencia con la que actúan determinadas fuentes. El problema en la caracterización de este tipo de sismos radica en la dificultad para distinguirlos del ruido base.

▪ Ruido

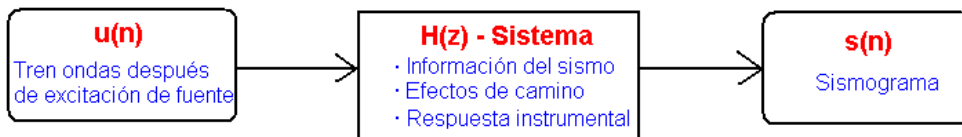
Para este estudio se define como ruido cualquier señal que no sea un sismo de los tipos mencionados anteriormente. De acuerdo con esto, inclusive los eventos tectónicos que se localicen fuera del edificio volcánico se consideran ruido. Entre los factores de mayor incidencia sobre el ruido base están el clima y la acción del hombre cerca de las estaciones sismológicas.

2. Coeficientes de predicción lineal (LPC)

2.1. Modelo de análisis por LPC

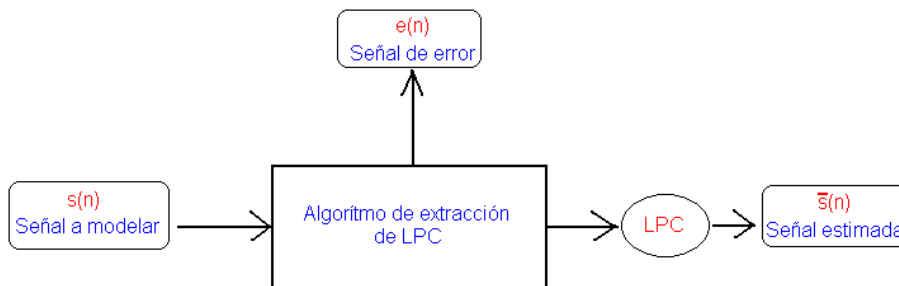
Supongamos una señal de excitación $u(n)$; esta señal está afectada por un sistema que modelaremos a través de su función de transferencia $H(z)$; la señal de salida del sistema es $s(n)$. El sistema, representado por $H(z)$, incluye características de fuente y todos los factores que afectan la señal hasta obtener el registro sismográfico (Figura. 2-1).

Figura 2-1: Flujo de la señal desde su excitación hasta su salida.



Los LPC se calculan utilizando un algoritmo que actúa sobre la señal $s(n)$ y arroja los vectores que contienen los coeficientes de predicción y una señal de error $e(n)$. Es posible recuperar la señal estimada $\tilde{s}(n)$ a partir de la señal de error aplicando sobre $e(n)$ un filtro inverso que contenga los LPC como coeficientes del filtro (Figura. 2-2).

Figura 2-2: Elementos involucrados en la extracción de los LPC.



La muestra $\tilde{s}(n)$ se obtiene mediante la combinación lineal de las salidas pasadas y las entradas presente y pasadas. Esta representación corresponde a la siguiente ecuación de diferencias:

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n-k) + G \sum_{l=0}^q b_l u(n-l), \quad b_0 = 1 \quad (2.1)$$

donde G , $\{a_k\}$ y $\{b_l\}$ son los parámetros del sistema. El valor de p es el número de muestras de salidas pasadas involucradas en el cálculo e indica el orden de la predicción lineal. La función de transferencia del sistema se obtiene aplicando la transformada Z a la ecuación (2.1) y calculando la relación entre la salida y la entrada en este dominio:

$$H(z) = \frac{S(z)}{U(z)} = G \frac{1 + \sum_{l=1}^q b_l z^{-l}}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (2.2)$$

La transformada \mathbf{Z} de una señal de tiempo discreto $x(n)$ está definida como la serie de potencias (Proakis, 1996):

$$X(z) = \sum_{n=-\infty}^{n=\infty} x(n)z^{-n} \quad (2.3)$$

donde \mathbf{z} es una variable compleja.

La ecuación 2.2 representa un *modelo polo-cero*. Sin embargo, la información más importante se encuentra en las bandas espectrales en que se concentra la mayor cantidad de energía y estas están representadas por los *polos* de $H(z)$. En teoría de filtros (Proakis, 1996) existen las siguientes posibilidades:

Si $b_l = 0$, para $1 \leq l \leq q$, $H(z)$ se reduce a un *modelo todo-polo* que se conoce como *modelo autoregresivo*.

Si $a_k = 0$, para $1 \leq k \leq p$, $H(z)$ se reduce a un *modelo todo-cero* o de *promedio móvil*.

Para resolver un *modelo polo-cero* es necesario dar solución a un conjunto de ecuaciones no lineales, mientras que la solución de un *modelo autoregresivo* requiere resolver un conjunto de ecuaciones lineales, facilitando su cálculo.

La función de transferencia para un *modelo todo-polo* es (Proakis, 1996):

$$H(z) = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (2.4)$$

y su transformada \mathbf{Z} inversa es:

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n) \quad (2.5)$$

Si imponemos que el factor de ganancia G sea igual a 1, tenemos:

$$H(z) = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} = \frac{1}{A(z)} \quad (2.6)$$

donde el polinomio $1 - \sum_{k=1}^p a_k z^{-k}$ está representado por $A(z)$. Los coeficientes $\{a_k\}$ son llamados *coeficientes de predicción lineal* ó LPC.

El error $e(n)$ asociado a la estimación de los coeficientes está dado por la diferencia entre la señal adquirida y la señal estimada mediante el filtro

$$e(n) = \tilde{s}(n) - \sum_{k=1}^p a_k s(n-k). \quad (2.7)$$

que en el dominio z es equivalente a:

$$E(z) = S(z)A(z). \quad (2.8)$$

2.2. Estimación de los coeficientes de predicción lineal (Jackson, 1989)

Para estimar los LPC en este estudio se utiliza el método de *autocorrelación*. En este método se buscan los coeficientes $\{a_k\}$ tales que la energía residual $e(n)$ sea mínima, esto se hace mediante la técnica de *mínimos cuadrados*.

2.2.3. Método de autocorrelación

Se pretende minimizar la energía de la señal residual que está definida como:

$$E = \sum_{n=-\infty}^{\infty} e^2(n) = \sum_{n=-\infty}^{\infty} \left(s_w(n) - \sum_{k=1}^p a_k s_w(n-k) \right)^2 \quad (2.9)$$

Donde s_w es la señal que se ha sometido a un proceso de *ventaneo*. Para encontrar los valores de $\{a_k\}$ que minimizan la función E es necesario encontrar las derivadas parciales $\frac{\partial E}{\partial a_k}$ para $k = 1, \dots, p$ e igualarlas a cero. Lo anterior da lugar a p ecuaciones con p variables $\{a_k\}$ por determinar, así:

$$\sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s_w(n-i)s_w(n-k) = \sum_{n=-\infty}^{\infty} s_w(n-i)s_w(n), \quad 1 \leq i \leq p, \quad (2.3)$$

Estas ecuaciones lineales pueden expresarse en términos de la función de *autocorrelación*. La *función de autocorrelación* del segmento ventaneado $s_w(n)$ está definida como:

$$R(i) = \sum_{n=i}^{N_w-1} s_w(n)s_w(n-i), 0 \leq i \leq p. (2.4)$$

Se sabe que la función de *autocorrelación* es una función par $R(i) = R(-i)$ Aplicando esta condición en la ecuación 2.14 tenemos:

$$\sum_{k=1}^p R(|i-k|)a_k = R(i), 1 \leq i \leq p. (2.5)$$

Que en forma matricial se representa como:

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix} (2.6)$$

o que en notación matricial sería igual a:

$$\mathbf{R}\mathbf{a}=\mathbf{r} \quad (2.7)$$

La matriz \mathbf{R} es una matriz de *Toeplitz* en la que los elementos de la diagonal principal son iguales. Esta característica permite que las ecuaciones puedan resolverse a través del algoritmo de *Levinson-Durbin*.

El algoritmo de *Levinson-Durbin* resuelve sistemas de ecuaciones de la forma $\mathbf{Ax}=\mathbf{b}$, donde \mathbf{A} es una matriz de *Toeplitz*, simétrica y positiva y \mathbf{b} es un vector cualquiera. *Durbin* encontró un algoritmo más eficiente para resolver el sistema en el caso en que \mathbf{b} tenga algunos elementos de \mathbf{A} y las ecuaciones de *autocorrelación* también satisfagan esta condición.

Sea $a_k(m)$ el k -ésimo coeficiente para un segmento en particular en la m -ésimo iteración. El algoritmo de *Levinson-Durbin* resuelve el siguiente conjunto de ecuaciones ordenadas recursivamente por $m = 1, 2, \dots, p$:

$$k(m) = R(m) - \sum_{k=1}^{m-1} a_k(m-1)R(m-k), (2.8)$$

$$a_m(m) = k(m), (2.9)$$

$$a_k(m) = a_k(m-1) - k(m)a_{m-k}(m-1), 1 \leq k \leq m. (2.17)$$

$$E(m) = (1 - k(m)^2)E(m-1). (2.18)$$

donde inicialmente $E(0) = R(0)$ y $a(0) = 0$. En cada iteración el m -ésimo coeficiente $a_k(m)$ para $k = 1, 2, \dots, m$ describe el m -ésimo predictor lineal óptimo y el error $E(m)$ se reduce en un factor de $(1 - k(m)^2)$.

3. Redes neuronales artificiales

Las *redes neuronales artificiales* (en adelante RNA) nacen como el producto de un intento por emular ciertas características del cerebro humano desde el punto de vista sensorial. Las *neuronas*, procesadores elementales y relativamente lentos se agrupan formando *capas* y estas a su vez forman *redes* que funcionan como un todo que nosotros conocemos como cerebro. El estilo de procesamiento que el cerebro utiliza es en *paralelo*, esto es, varias neuronas pueden activarse al mismo tiempo sin que el sistema colapse por ello. El conocimiento está ampliamente *distribuido* en la red lo que evita grandes pérdidas de información en caso de daños parciales. El cerebro humano tiene la capacidad de *adaptarse* a las condiciones cambiantes del medio exterior, aprendiendo nuevas cosas y aplicándolas sobre ese nuevo entorno. En el cerebro existe una gran *tolerancia a los fallos* ya que si una neurona deja de funcionar no es motivo para que el sistema completo interrumpa su funcionamiento. El tipo de control de los procesos es *auto-organizado*, lo que implica que existen grupos de neuronas especializadas en tipos específicos de procesos.

En la descripción anterior, de las características del cerebro que se pretende emular mediante RNA se encuentran las siguientes: *procesamiento en paralelo, almacenamiento distribuido, adaptabilidad, tolerancia a fallos y control de proceso auto-organizado*.

Las RNA se han convertido en una alternativa para solucionar problemas con las siguientes características (Del Brío, 2002):

- Falta de un conjunto de reglas sistemáticas que describan completamente el problema.
- Disponibilidad de muchos ejemplos o casos históricos.
- Datos imprecisos o ruidosos.
- Elevada *dimensionalidad*.
- Condiciones de trabajo cambiantes.

3.1. La Neurona Artificial

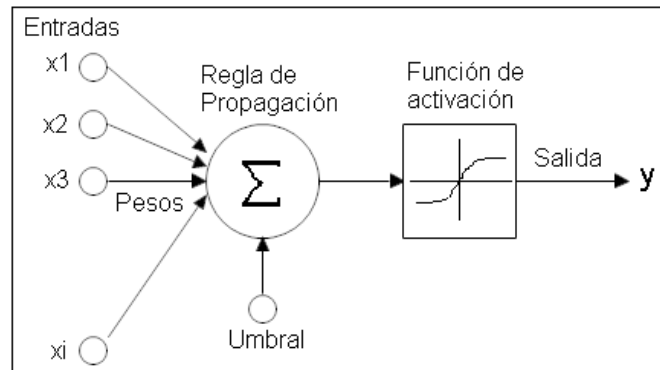
Una neurona artificial es *una unidad elemental de proceso que realiza cálculos* a partir de un vector de entrada procedente del exterior o de otras neuronas. La neurona artificial es la base estructural de las redes neuronales artificiales.

En una neurona i pueden distinguirse los siguientes componentes (Figura. 3-1):

- Conjunto de **entradas** x_j .
- **Pesos sinápticos** de la neurona i , w_{ij} que representan la fuerza de la interacción entre cada neurona *presináptica* j (anterior a los pesos sinápticos) y la neurona *postsináptica* i (posterior a los pesos sinápticos).
- **Regla de propagación** que proporciona el valor del potencial *postsináptico* de la neurona i en función de sus pesos sinápticos y entradas.

- **Función de activación** que proporciona el estado de activación actual de la neurona i en función de su estado anterior y de su potencial *postsináptico*.
- **Función de salida** que proporciona la salida actual de la neurona i en función de su estado de activación.

Figura 3-1: La neurona artificial.



Frecuentemente se añade al conjunto de pesos sinápticos de la neurona un parámetro adicional que se denomina umbral θ_i , que reduce el valor de entrada a la *función de activación*.

3.1.1. Entradas y salidas

La forma de las entradas y salidas en las redes neuronales puede variar dependiendo del tipo de aplicación. Cuando se quiere resolver un problema de clasificación es conveniente utilizar como entradas un vector que contiene información acerca del patrón que se desea discriminar y como salidas vectores de baja *dimensionalidad* cuyos elementos suelen estar en los rangos $\{0,+1\}$ o $\{-1,+1\}$. Los vectores de entrada pueden también normalizarse teniendo en cuenta los conjuntos completos de los diferentes tipos de patrones a clasificar.

3.1.2. Pesos sinápticos

Los pesos sinápticos condicionan el poder de la interacción entre la neurona *presináptica* j y la neurona *postsináptica* i . Si la entrada es positiva el peso sináptico que interactúa con dicha entrada producirá efectos que realzarán la excitación (peso positivo) o inhibición (peso negativo) de la neurona dependiendo de si el peso es positivo o negativo respectivamente.

3.1.3. Regla de propagación

La regla de propagación determina el *potencial postsináptico* $h_i(t)$ a partir del vector de entrada y los pesos sinápticos, generalmente se utiliza el producto escalar entre el vector de entrada y el de pesos sinápticos:

$$h_i(t) = \sum_j w_{ij} x_j \quad (3.1)$$

3.1.4. Función de activación

De esta función depende el estado de activación de la neurona, su argumento es el potencial *postsináptico*. Las funciones utilizadas suelen ser continuas y monótonas crecientes; en problemas de clasificación los algoritmos utilizados para el aprendizaje de la red exigen también que la función de activación sea derivable. El estado de activación actual de la red sería entonces:

$$a_i(t) = f_i(h_i(t)) \quad (3.2)$$

3.1.5. Función de salida

Esta función es la que presenta el estado global de la neurona, depende del estado de activación; lo más común es utilizar este estado como la propia salida de la neurona, por tanto:

$$y_i(t) = F_i(a_i(t)) = a_i(t) \quad (3.3)$$

3.2. Arquitectura de las Redes Neuronales

Se denomina *arquitectura* de una red neuronal a la estructura que conecta las neuronas entre sí para formar una red neuronal. Las arquitecturas de red son de diversos tipos, unas son más apropiadas que otras para la solución de determinados problemas. La arquitectura de red que se escoja depende entre, otras cosas, de las entradas que se usen para representar los patrones y de la manera como se desee obtener las salidas. En nuestro caso, en que se necesita clasificar entre conjuntos con diferentes características, se suele utilizar la arquitectura *perceptrón multicapa* (Del Pezzo, 2003; Rugeles & Cadena 2006; Cadena & Arcos 2008).

En la arquitectura *perceptrón multicapa* las neuronas se agrupan en unidades denominadas capas, una capa generalmente contiene neuronas de un mismo tipo. El número de capas y neuronas por capa se denomina topología de la red. Las capas son de tres tipos: *de entrada*, *de salida* y *ocultas*. Las capas de entrada y salida tienen contacto directo con el exterior, es decir, se comportan como *sensores*, en el caso de las capas de entrada, o *afectores* en el caso de las capas de salida. Las capas ocultas no tienen contacto con el exterior, pero se comportan como los elementos computacionales más importantes al momento de buscar las relaciones entre las entradas y las salidas deseadas.

Una vez conectadas las capas para formar una red neuronal (Figura. 3-2) es posible que la información fluya únicamente hacia adelante, es decir, desde las neuronas de entrada hacia las ocultas y finalmente a las de salida; en este caso estamos hablando de *redes unidireccionales*, también puede suceder que la información fluya de tal manera que se devuelva entre las capas o inclusive desde la capa de salida hacia la de entrada, en este caso estaremos tratando *redes neuronales recurrentes*.

3.3. Definición de red neuronal artificial

Una definición formal de red neuronal fue formulada por Muller, (1990) que acude al concepto de *grafo*, así:

Un *grafo* G es un par ordenado $G = (V, E)$, donde:

V es un conjunto de vértices o nodos, y

E es un conjunto de aristas o arcos, que relacionan estos nodos.

Normalmente V suele ser finito.

Un *grafo dirigido* es un grafo $G = (V, E)$ donde:

$V \neq \emptyset$

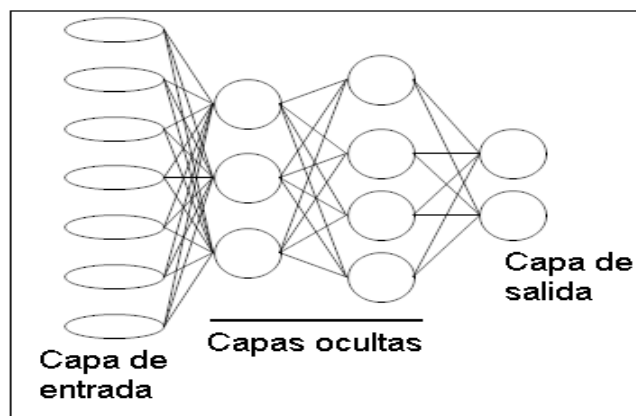
$E \subseteq \{(a, b) \in V \times V : a \neq b\}$ es un conjunto de pares ordenados de elementos de V .

Dada una arista (a, b) , a es su *nodo inicial* y b su *nodo final*. Por definición, los grafos dirigidos no contienen *bucles*.

Por tanto, una red neuronal es un grafo dirigido, con las siguientes propiedades:

1. A cada nodo i se asocia una variable de estado x_i .
2. A cada conexión (i, j) de los nodos i y j se asocia un peso $w_{ij} \in \mathfrak{R}$.
3. A cada nodo i se asocia un umbral θ_i .
4. Para cada nodo i se define una función $f_i(x_i, w_{ij}, \theta_i)$, que depende de los pesos de sus conexiones, del umbral y de los estados de los nodos j a él conectados. Esta función proporciona el nuevo estado del nodo.

Figura 3-2: Red neuronal artificial.



3.4. Operación de las redes neuronales artificiales

Existen dos modos en los que operan las RNA, *aprendizaje o entrenamiento y recuerdo o ejecución*. En la fase de entrenamiento la RNA aprenderá una tarea deseada hasta que el entrenador (quien opera la red) considere que la tarea puede ser ejecutada sin la necesidad de ser supervisada, es entonces cuando se pone en marcha la fase de ejecución en que la RNA entrenada aplica los conocimientos adquiridos para realizar la labor para la que fue creada.

3.4.1. Fase de aprendizaje

A nivel operativo la fase de aprendizaje o entrenamiento consiste en ajustar los pesos sinápticos w_{ij} de la red a través de la aplicación de un algoritmo que minimiza una función de error hasta satisfacer un criterio establecido anticipadamente. Este algoritmo llamado también *regla de aprendizaje* es el encargado de tomar las señales del exterior o ejemplos y en función de estos ajustar en un valor Δw_{ij} el estado actual de los pesos sinápticos:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (3.4)$$

donde $w_{ij}(t)$ representa el valor actual del peso sináptico que une las neurona presináptica j con la postsináptica i y $\Delta w_{ij}(t)$ representa el ajuste.

Este proceso se realiza de manera iterativa hasta que se considere que la red ha alcanzado cierto nivel de aprendizaje. En este punto del proceso la RNA tiene una percepción del entorno que utilizará para realizar la tarea que le fue asignada. De los tipos de aprendizaje que existen los más utilizados son el *supervisado* y el *no supervisado*.

3.4.2. Aprendizaje supervisado

En este tipo de aprendizaje el entrenador le presenta a la red patrones de un tipo junto con una etiqueta del conjunto al que pertenece el patrón. Por ejemplo, se presenta un conjunto de vectores como entrada y un *objetivo* común como salida; este último suele ser un vector de menor dimensión con el cual se desea asociar las entradas. El proceso de presentación de los patrones es iterativo y la regla de aprendizaje ajusta los pesos sinápticos de la red en cada iteración hasta encontrar una relación entre las entradas y las salidas.

3.4.3. Aprendizaje no supervisado

Este tipo de aprendizaje se basa en el cálculo de la densidad de probabilidad $P(x)$ dada por la distribución de los patrones de entrada x . La regla de aprendizaje busca características similares entre los ejemplos que se le presentan y los agrupa en conjuntos dependiendo de sus características.

La escogencia de una arquitectura de red y una regla de aprendizaje adecuados evita problemas de convergencia en el algoritmo al momento de minimizar la función de error.

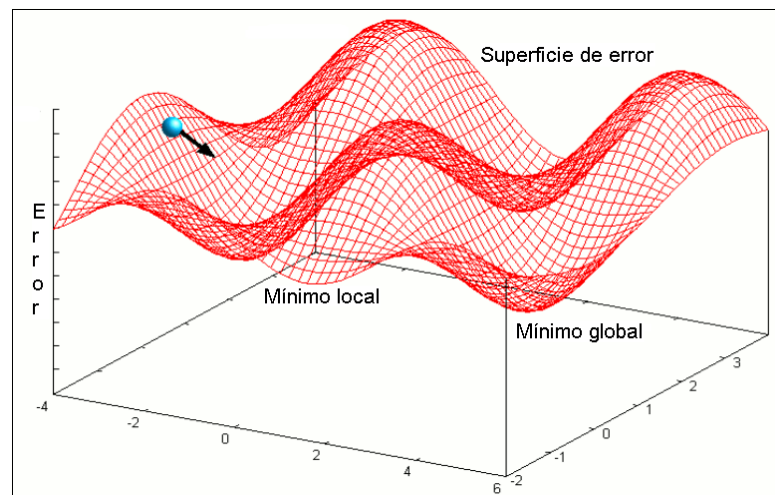
3.4.4. Retropropagación del error

Atendiendo al tipo de problema a resolver existe una variedad de algoritmos de aprendizaje, entre los más utilizados están: el aprendizaje *Jebbiano*, la regla *LMS* (*Least Mean Squares*), la regla de aprendizaje de *Boltzman*, el de corrección de errores que incluye la *Adalina* y el de *retropropagación del error* (Rumelhart, 1986; Hecht-Nielsen, 1991).

El algoritmo por retropropagación del error o *BP* por sus siglas en inglés (*Backpropagation*) está basado en la regla *Widrow-Hoff* o *LMS* que conduce a actualizaciones de los pesos sinápticos proporcionales al error de la red neuronal.

Sea E una función de error $E = E[W], E = R^n \rightarrow R$, función de los pesos sinápticos que mide la eficacia de la red. Se requiere encontrar un proceso que optimice la función de error buscando un mínimo. Para ello se utiliza el método de *descenso por el gradiente* que se ilustra en la Figura. 3-3 Si la función utilizada dependiera solamente de dos variables (en problemas reales su dimensión es mucho mayor) su gráfica sería semejante a una sábana plegada con crestas y valles; estar sobre uno de sus valles corresponde a estar ubicado en un mínimo local óptimo. El objetivo, que no siempre es posible de lograr, es encontrar un mínimo global; sin embargo un buen mínimo local muy frecuentemente es suficiente para abordar el problema. Este proceso se lleva a cabo iterativamente con ayuda de los patrones de aprendizaje y los objetivos de cada conjunto de entrenamiento.

Figura 3-3: Esquema de descenso por el gradiente para una función de error. La esfera azul indica el punto donde se evalúa el error y la flecha la dirección hacia donde se debe mover la esfera para alcanzar un mínimo de la función (Tomado de JOONE, 2007).



En el método de *descenso por el gradiente* los pesos sinápticos para $t=0$, $W(0)$ (inicializados con valores aleatorios pequeños) se ajustan una vez se ha calculado el gradiente de la función de error $E(W)$. El gradiente proporciona la dirección de la máxima variación de la función, por tanto el ajuste se hace en sentido contrario a esta dirección y en un valor adecuado, no tan grande de tal manera que no se salten características importantes de esa superficie tales como mínimos locales cercanos, ni tan pequeño que el tiempo de computo sea excesivamente largo. Cada una de estas actualizaciones tendría la forma:

$$W(t+1) = W(t) - \varepsilon \cdot \nabla E(w) \quad (3.5)$$

donde ε indica el tamaño del paso tomado en cada iteración. Para la iteración t la variación en la función de error $E(W)$ es:

$$\delta(E(w_{ij})) = \sum_{ij} \frac{\delta E(w_{ij})}{\delta w_{ij}} \delta w_{ij} \quad (3.6)$$

donde los w_{ij} son los elementos de la matriz de pesos sinápticos. Al multiplicar por el gradiente obtenemos:

$$\delta(f(w_{ij})) = \sum_{ij} \frac{\delta E(w_{ij})}{\delta w_{ij}} \left(-\varepsilon \frac{\delta E(w_{ij})}{\delta w_{ij}} \right) = -\varepsilon \sum_{ij} \left(\frac{\delta E(w_{ij})}{\delta w_{ij}} \right)^2 \leq 0 \quad (3.7)$$

Así se demuestra que la actualización es siempre menor que cero, asegurando que a través de este método se encontrarán mínimos ya sean locales o globales.

En este método la forma como se alcanza un mínimo es asintótica, ya que con el transcurrir de las iteraciones las variaciones son cada vez más pequeñas dado que son proporcionales al error asociado al paso por la red.

La desventaja de la regla LMS reside en que la función de error debe estar en función únicamente de los pesos sinápticos de la capa de salida. En el caso de una arquitectura neuronal con capas ocultas es necesario generalizar la regla, esto se logra con el algoritmo *BP*. Debido a que en el proceso de ajuste de los pesos sinápticos es necesario derivar las funciones según la regla de la cadena se exige que las funciones de transferencia de las neuronas sean derivables.

3.4.5. Fase de recuerdo

Una vez ha concluido la fase de aprendizaje, la RNA está lista para responder a estímulos con los que nunca ha trabajado, pero que tienen características comunes a los conjuntos con los que se entrenó. Inicia entonces la fase de recuerdo, que depende de la manera en que están *memorizadas* las características de los patrones con que se entrenó a la red.

3.4.6. Memoria asociativa

La simulación de una red neuronal biológica conlleva al desarrollo de una memoria asociativa. La principal propiedad de este tipo de memoria es su habilidad para relacionar (*asociar*) un patrón de salida, como consecuencia de cierta actividad neuronal, con un patrón de entrada.

Durante la fase de aprendizaje se presenta a la red un estímulo de entrada o patrón de entrada y la memoria lo transforma en uno memorizado, mientras que en la fase de recuerdo

se presenta a la red una versión ruidosa del patrón presentado anteriormente y la memoria se encarga de asociarlo con el memorizado.

La memoria asociativa puede ser *lineal* o *no lineal*, esto depende del tipo de neuronas que se utilicen en la arquitectura de red. Sean \mathbf{a} y \mathbf{b} los vectores de entrada y salida respectivamente. La memoria asociativa es de tipo lineal cuando la relación entre estos vectores es de la forma:

$$\mathbf{b} = \mathbf{M}\mathbf{a} \quad (3.8)$$

donde \mathbf{M} es la matriz de memoria. Si la relación entre \mathbf{b} y \mathbf{a} tiene la forma:

$$\mathbf{b} = \varphi(\mathbf{M}, \mathbf{a})\mathbf{a} \quad (3.9)$$

donde φ es una función que depende de la matriz de memoria, \mathbf{M} , y del vector de entradas \mathbf{a} , entonces se dice que la memoria es de tipo no lineal.

Supongamos que como respuesta a un vector de estímulo de entrada \mathbf{a}_k se obtiene simultáneamente un vector \mathbf{b}_k , donde el subíndice k denota el número de patrón o vector que se asociará. Para nuestro problema contaremos con q patrones:

$$\mathbf{a}_k = [a_{k1}, a_{k2}, \dots, a_{kp}]^T \quad (3.10)$$

$$\mathbf{b}_k = [b_{k1}, b_{k2}, \dots, b_{kp}]^T \quad (3.11)$$

Por simplicidad convenimos que la dimensión de los vectores de entrada y salida es igual, digamos p . Si asumimos un tipo de memoria lineal podemos escribir:

$$\mathbf{b}_k = \mathbf{W}(k)\mathbf{a}_k, \quad k = 1, 2, \dots, q \quad (3.12)$$

donde $\mathbf{W}(k)$ es la matriz de pesos sinápticos. La salida b_{ki} de la neurona i debida a la acción combinada de los elementos del conjunto de patrones \mathbf{a}_k aplicados como estímulo en la capa de entrada de la red es:

$$b_{ki} = \sum_{j=1}^p w_{ij}(k)a_{kj}, \quad i = 1, 2, \dots, q \quad (3.13)$$

$$j = 1, 2, \dots, p$$

Los w_{ij} son los pesos sinápticos de la neurona i correspondientes al k -ésimo par de patrones asociados. La ecuación 3.13 en notación matricial es:

$$b_{ki} = [w_{i1}(k) \quad w_{i2}(k) \quad \dots \quad w_{ip}(k)] \begin{bmatrix} a_{k1} \\ a_{k2} \\ \cdot \\ \cdot \\ a_{kp} \end{bmatrix} \quad i = 1, 2, \dots, q \quad (3.14)$$

que se puede escribir como:

$$\begin{bmatrix} b_{k1} \\ b_{k2} \\ \cdot \\ \cdot \\ b_{kp} \end{bmatrix} = \begin{bmatrix} w_{11}(k) & w_{12}(k) & \dots & w_{1p}(k) \\ w_{21}(k) & w_{22}(k) & \dots & w_{2p}(k) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ w_{p1}(k) & w_{p1}(k) & \dots & w_{pp}(k) \end{bmatrix} \begin{bmatrix} a_{k1} \\ a_{k2} \\ \cdot \\ \cdot \\ a_{kp} \end{bmatrix} \quad (3.15)$$

expandiendo las matrices de pesos sinápticos:

$$\mathbf{W}(k) = \begin{bmatrix} w_{11}(k) & w_{12}(k) & \dots & w_{1p}(k) \\ w_{21}(k) & w_{22}(k) & \dots & w_{2p}(k) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ w_{p1}(k) & w_{p1}(k) & \dots & w_{pp}(k) \end{bmatrix} \quad (3.16)$$

La asociación de los q pares individuales produce q matrices del tipo $\mathbf{W}(k)$. De tal manera que se puede definir una matriz que contenga la suma de las matrices de pesos sinápticos que asocian cada par. Esta matriz de memoria tendrá la forma:

$$\mathbf{M} = \sum_{k=1}^p \mathbf{W}(k) \quad (3.17)$$

que define la conectividad total entre las capas de entrada y de salida de la *memoria asociativa*.

La ecuación 3.17 puede escribirse en forma recursiva, para hacer más evidente la forma como se construye la matriz de memoria asociativa \mathbf{M} , acudiendo a las actualizaciones de la matriz a medida que los pares de entrada y salida entran en juego. Así tenemos:

$$\mathbf{M}_k = \mathbf{M}_{k-1} + \mathbf{W}(k), \quad k = 1, 2, \dots, q \quad (3.18)$$

donde \mathbf{M}_k es la matriz ajustada por la matriz $\mathbf{W}(k)$ que contiene la información la nueva (k -ésima) asociación.

Consideremos ahora un estimado de la matriz \mathbf{M} que denotaremos por $\hat{\mathbf{M}}$ (Anderson, 1972, 1983 y Cooper, 1973), como sigue:

$$\hat{\mathbf{M}} = \sum_{k=1}^q \mathbf{b}_k \mathbf{a}_k^T \quad (3.19)$$

donde el producto $\mathbf{b}_k \mathbf{a}_k^T$ es un estimado de la matriz $\mathbf{W}(k)$. En forma matricial $\hat{\mathbf{M}}$ puede escribirse como:

$$\hat{\mathbf{M}} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_q] \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{a}_q^T \end{bmatrix} \quad (3.20)$$

finalmente podemos expresar la ecuación 3.20 en forma recursiva:

$$\hat{\mathbf{M}}_k = \hat{\mathbf{M}}_{k-1} + \mathbf{b}_k \mathbf{a}_k^T, \quad k = 1, 2, \dots, q \quad (3.21)$$

que es una ecuación análoga a 3.18, pero con la diferencia de que en este caso tenemos un *estimado de la matriz M*.

Entonces, el problema en la fase de recuerdo es solucionar:

$$\mathbf{b} = \hat{\mathbf{M}} \mathbf{a}_j \quad (3.22)$$

para dar respuesta a un patrón nuevo, que no fue utilizado por el conjunto de ejemplos de entrenamiento a partir del que se estimó dicha matriz.

3.5. Capacidad de generalización

Si el problema a solucionar fuese el de verificar que el patrón desconocido es exactamente uno de los presentados en la fase de entrenamiento, sería mejor utilizar una base de datos que contuviera todos los vectores con los que se habría aplicado el entrenamiento y un algoritmo que comparara elemento a elemento los vectores memorizados con el que se desea dar respuesta. Una de las ventajas de las RNA es que no necesita memorizar uno a uno los

patrones de entrenamiento sino el de reconocer características claves en ellos y extender ese conocimiento a estímulos desconocidos, esto es lo que se conoce como *generalización*. Para detener el entrenamiento es necesario medir el nivel de generalización de la red en función de su nivel de aprendizaje.

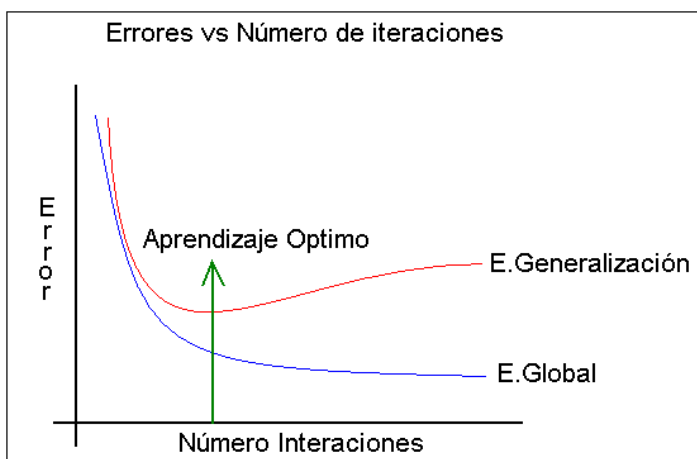
3.6. Validación cruzada

Durante el proceso de entrenamiento aparecen dos tipos de error: 1) el error de ajuste de cada patrón a su respectivo objetivo, llamado *error global*, que se calcula como el error cuadrático medio del conjunto de patrones y objetivos (para el caso del aprendizaje supervisado), 2) el error de respuesta a patrones desconocidos, llamado *error de generalización*.

Se considera que una RNA ha tenido una fase de aprendizaje exitosa cuando extiende la aplicación de lo aprendido a patrones desconocidos, esto se puede evaluar realizando pruebas de la respuesta de la red a ejemplos de entrada no conocidos. Es preferible seleccionar una RNA que muestre un error de generalización aceptable a una que arroje errores globales pequeños, ya que estos últimos pueden reducirse tanto como se quiera con solo incrementar el número de iteraciones. Un hecho observado en experimentos con simuladores de RNA es que cuando se alcanza un error global bastante pequeño, digamos menor que 1%, la generalización de la red se degrada (Del Brío, 2002).

En la Figura 3-4 se observa como el error global disminuye asintóticamente a medida que el número de iteraciones crece, mientras que el error de generalización alcanza un mínimo. Detener el proceso de entrenamiento o guardar la configuración de pesos sinápticos y umbrales en ese instante en que el aprendizaje es óptimo, nos asegura que la configuración de la matriz de memoria es la más adecuada para utilizarla en la fase de recuerdo, esta técnica se conoce como validación cruzada.

Figura 3-4: Errores global y de generalización en un problema ideal, la flecha verde indica el momento en que el aprendizaje es óptimo.



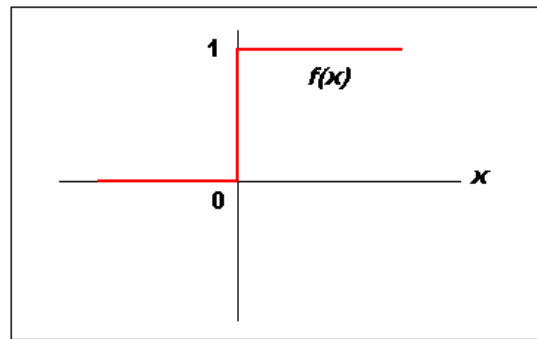
3.7. El Perceptrón multicapa

El *perceptrón multicapa* o MLP (*Multilayer perceptron*) tiene sus raíces en una arquitectura de red más simple denominada *perceptrón simple*, que consta de dos capas, una de entrada, que no realiza cálculo alguno y otra de salida que es la que realiza las operaciones encargadas de dar respuesta a un estímulo determinado. La función de activación de un *perceptrón simple* es la función escalón o Heaviside (Figura. 3-5).

$$f(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases} \quad (3.23)$$

Esta arquitectura básica puede utilizarse para problemas simples de clasificación debido a que su salida es esencialmente booleana. La gran limitación del *perceptrón simple* es que puede discriminar dos conjuntos *linealmente separables*, esto es, que se puedan distinguir con ayuda de un *hiperplano* para el caso de más de dos dimensiones o una línea recta para el caso de dos dimensiones.

Figura 3-5: La función escalón o Heaviside.

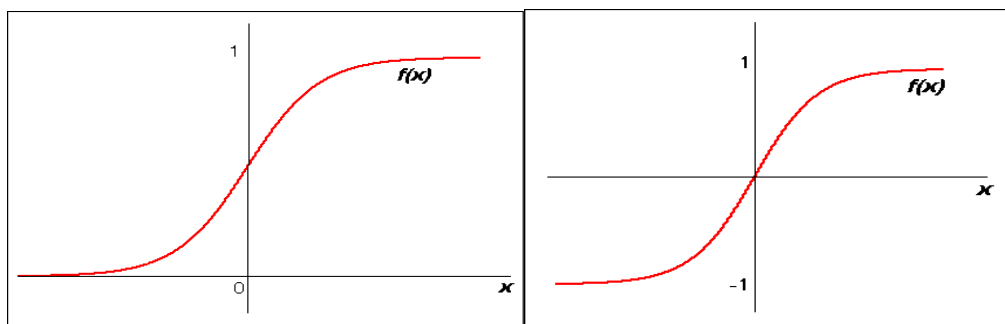


El MLP (Rumelhart, David E., 1986), supera las limitaciones del *perceptrón simple* y permite abordar problemas complejos de clasificación para conjuntos no necesariamente linealmente separables. Esta RNA se entrena con ayuda del algoritmo de retropropagación del error y tiene la característica de que la función de activación que se utiliza en sus neuronas ocultas y de salida es de tipo sigmoideo (Figura. 3-6):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.24)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \tanh(x) \quad (3.25)$$

Figura 3-6: Funciones de activación utilizadas en las neuronas del MLP. (a) a la izquierda la correspondiente a la ecuación 3.30 y a la derecha a la ecuación 3.31.



El entrenamiento de un MLP, siguiendo el procedimiento que aplica el algoritmo de retropropagación del error, se resume en los siguientes pasos (Del Brío, 2002):

2. Establecer aleatoriamente los pesos y umbrales iniciales.
3. Para cada patrón μ del conjunto de aprendizaje:
 - Llevar a cabo una fase de ejecución para obtener la respuesta de la red ante el patrón μ -ésimo.
 - Calcular las señales de error asociadas Δ'_k y Δ''_j según (3.11 - 3.12).
 - Calcular el incremento parcial de los pesos y umbrales debidos a cada patrón μ (elemento de los sumatorios 3.11 y 3.12).
4. Calcular el incremento total (es decir, para todos los patrones) actual de los pesos δw_{kj} y δw_{ji} según (3.11 - 3.12). Hacer lo mismo para los umbrales.
5. Actualizar pesos y umbrales.
6. Calcular el error actual, $t = t + 1$, y volver a 2) si el error todavía no es satisfactorio.

Existen variaciones al algoritmo en las que se reemplaza el método de descenso por el gradiente por otras técnicas llamadas de *segundo orden* como son los *gradientes conjugados*, *gradientes conjugados escalados* y *Levenberg-Marquardt*, que agilizan la convergencia hacia un mínimo. El objetivo principal del uso de estas variantes es hacer más eficiente el recurso computacional, pero en términos de la efectividad alcanzada, en general, ninguno es superior (Del Brío, 2002).

3.8. Número de ejemplos de entrenamiento

Haykin, (1999) explica que la capacidad de generalización de la red está determinada por tres aspectos: la arquitectura de la red, el número de ejemplos de entrenamiento y la complejidad del problema.

Mientras más complejo sea el problema que se desea abordar, la RNA contará con más parámetros por ajustar y por tanto serán necesarios más ejemplos en la fase de entrenamiento de la red. Baum, (1989) demostró que para una red con n entradas y h

neuronas ocultas con w pesos en total, se requiere un número de patrones de aprendizaje del orden de $p = \frac{w}{\varepsilon}$ para proporcionar un error de aprendizaje del orden de ε , esto fue corroborado por Haykin, (1999). Así, por ejemplo, para un MLP con 15 neuronas en la capa de entrada, cinco en la capa oculta y dos en la de salida (15-5-2) el número de ejemplos necesarios para entrenar la red será de $p = 1840$.

Para calcular el número de patrones de entrenamiento Haykin, (1999) sugiere reservar un 20% para aplicar la técnica de *validación cruzada*.

De otra parte, si el número de neuronas de entrada es elevado, el número de parámetros de la red lo será aún mucho más, por tanto es necesario contar con un número de patrones de aprendizaje que en la mayoría de los casos es imposible de tener. A modo de ejemplo: si se cuenta con una red de 200 neuronas de entrada, cinco en su capa oculta y 2 en la salida (200-5-2) el número de ejemplos de entrenamiento para lograr un adecuado aprendizaje (digamos con un error del 5%) será cercano a los 40.000, que es considerablemente grande. Para evitar esto es necesario reducir la dimensión del vector de entrada utilizando técnicas que representen la señal en una dimensión más pequeña que la original; esto conlleva a realizar un *pre-procesamiento* de los patrones de entrenamiento.

4. Metodología

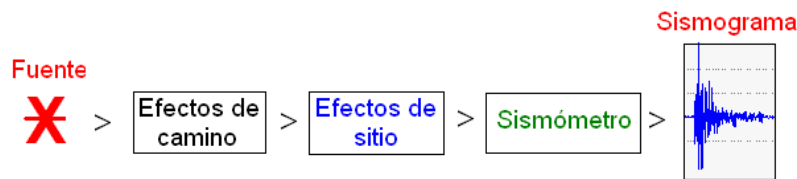
La estrategia para resolver el problema contempla cinco pasos: en el primer paso se plantea la manera de abordar la detección y la clasificación de los sismos, los tipos de sismos a clasificar y las señales que están fuera del alcance de este trabajo; en el segundo paso, relacionado con la representación y reducción de *dimensionalidad*, se estudia el orden de la predicción lineal para una adecuada representación de las señales; en el tercer paso se diseña y desarrolla un programa que descienda a un nivel operativo la aplicación de la detección y la clasificación y se implementa una base de datos relacional para organizar la información generada por el programa; el cuarto paso aborda el estudio de las topologías de las RNA para las fases de detección y clasificación mediante el entrenamiento de la RNA con diferentes topologías y ejecuta fuera de línea para evaluar los resultados y, finalmente, en el quinto paso se procede a la ejecución en línea del proceso de clasificación. Como ejemplo de prueba de la metodología y de la herramienta computacional se realiza un ejercicio de clasificación de eventos en una de las estaciones de la red sismológica del VG, sin embargo, su aplicación se puede extender a todas las estaciones de la red.

4.1. Detección y clasificación

En la fase de detección se pretende reconocer señales asociadas con sismicidad volcánica registradas en la traza continua de un sismograma.

El sismograma contiene información de la fuente, de las características del camino (*scattering*, atenuación intrínseca, atenuación geométrica, etc.), efectos de sitio y respuesta del instrumento (Figura 4-1). El problema que se plantea es extraer del sismograma aquellas características de las fuentes que permitan su clasificación.

Figura 4-1: Elementos que intervienen en la generación del sismograma.



El efecto del sismómetro se maneja removiendo la respuesta instrumental del sismograma según las curvas proporcionadas por el fabricante y el efecto de sitio realizando la clasificación estación por estación. Los efectos de camino, para sismos relativamente cercanos al observador, tienen una menor incidencia sobre el sismograma que los efectos de la fuente. Esto podría explicarse por las cortas distancias recorridas por las ondas; en el caso de los sismos tipo LP y TRE, en su mayoría se localizan en la zona epicentral cercana al cráter a profundidades menores que 2 km (Figura. 4-2) y en el caso de los sismos VT, aunque más dispersos, localizados en todo caso dentro del edificio volcánico (Figura 4-3) en un área de cerca de 10 km de radio.

La evidencia de lo expuesto anteriormente está en que la *firma* de eventos del mismo tipo contiene características generales similares, sin importar que las ondas hayan seguido diferentes caminos hasta llegar a la estación.

Figura 4-2: Localización aproximada de la sismicidad de fluidos, C1, C2 y C3 representan el recorrido de un rayo entre la fuente y las respectivas estaciones.

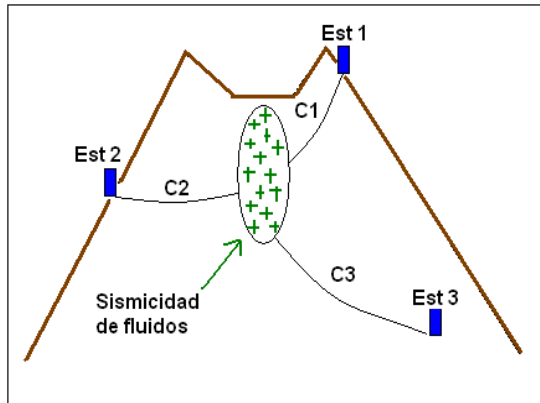
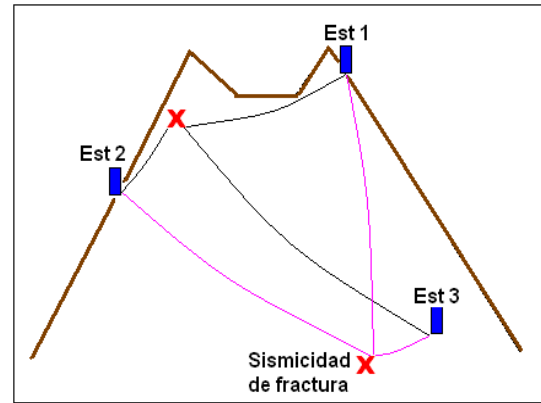


Figura 4-3: Localización de dos sismos de fractura, obsérvese que los recorridos a cada estación son diferentes.



Dado que las RNA funcionan como clasificadoras y no activando algún tipo de *trigger*, para la fase de *detección* se propone *clasificar* en dos conjuntos de señales, las que están asociadas a *ruido* y las que están asociadas a los sismos volcánicos. Los elementos de cada conjunto corresponden a segmentos de traza, denominados *unidades*, que contienen un número determinado de muestras. La duración de una unidad se mide en segundos y debe ser lo más corta posible para que la aplicación se acerque al *tiempo real*. El conjunto de ruido contiene unidades sin actividad sísmica, mientras que el conjunto de sismicidad volcánica contiene unidades de sismos de los tipos LP, VT-A, TRE y no clasificables (en adelante se entenderá cuando se encuentre las siglas VT se hace alusión al tipo A).

Una ventaja de este esquema de detección es que puede ser utilizado para el reconocimiento de sismo tipo TRE si se aprovechan las características de duración de estos eventos, que en general es mayor que la de los tipos LP y VT. Se plantea reconocer tremor volcánico a partir de un umbral mínimo de unidades consecutivas clasificadas como segmentos de sismo.

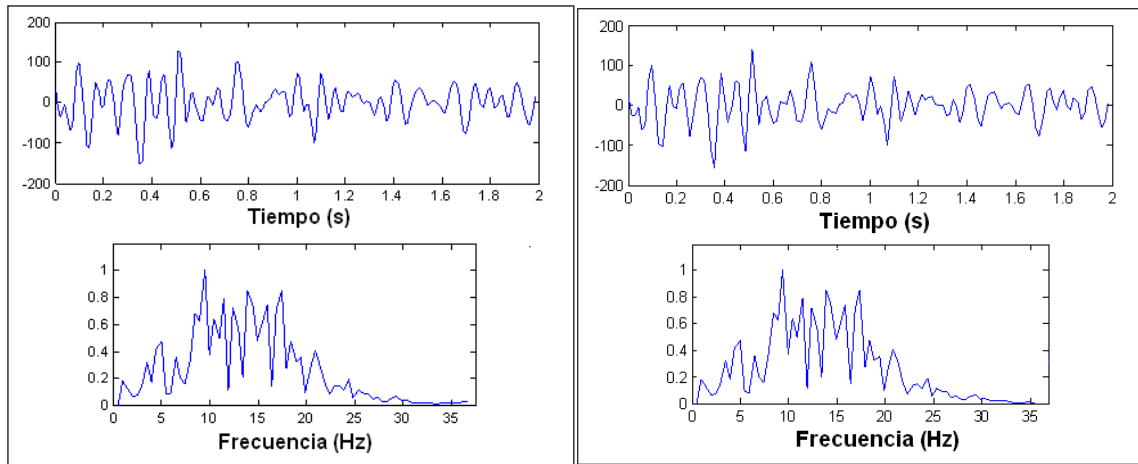
Para la *clasificación* entre LP y VT se juntan varias unidades que contengan segmentos consecutivos de sismo sin sobrepasar el umbral mínimo de unidades que clasificaría la señal como un episodio de tremor, posteriormente se somete al proceso de clasificación entre estos dos tipos.

4.1.1. Tratamiento previo de las señales, pre-procesamiento

El análisis espectral de los sismogramas de banda ancha de eventos en el VG, en su gran mayoría, no contiene información por encima de los 35 Hz, lo anterior reducir la cantidad de muestras a 70 por segundo (re-muestreo) para agilizar los cálculos y eliminar ruido de altas frecuencias.

En la Figura 4-4 se observa que el efecto de re-muestro es mínimo, tanto en el dominio del tiempo, como en el contenido espectral de la señal en el rango de frecuencias menores a la mitad de la frecuencia de muestreo, sin embargo, remuestrear la señal disminuye el tiempo de cómputo, especialmente durante la fase de ejecución de las RNA.

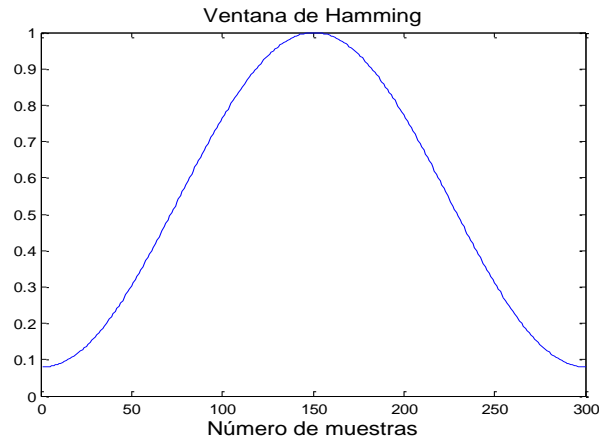
Figura 4-4: Sismograma y espectro de un segmento muestreado a 100 Hz. (izquierda) y del mismo segmento re-muestreado a 70 Hz. (derecha).



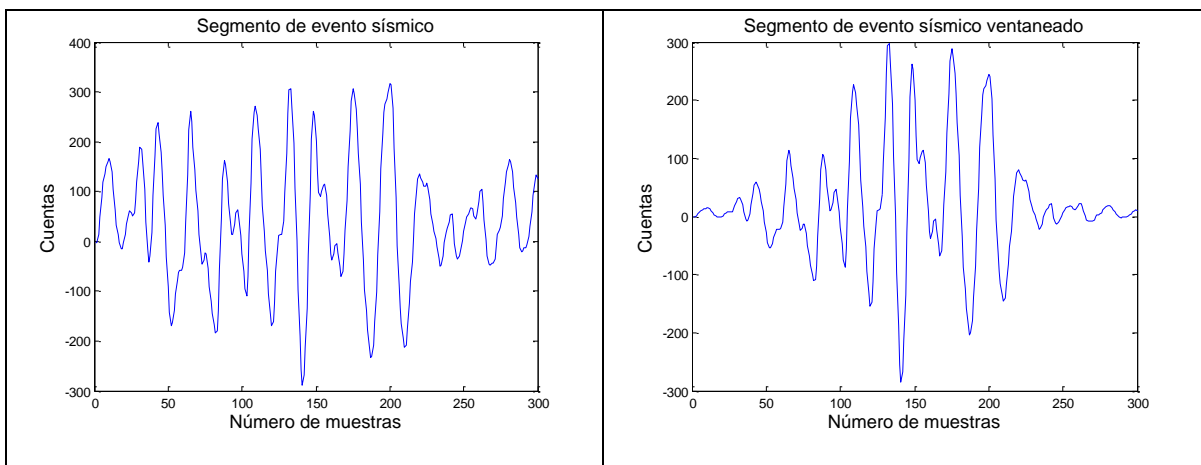
Después del re-muestreo se cortan las señales en unidades del mismo número de muestras, se remueve el offset y se normalizan las unidades. Se ha incluido la posibilidad de normalizar cada unidad, un conjunto consecutivo de unidades y el vector de LPC, esta normalización puede hacerse en los rangos $[0,1]$ y $[-1,1]$.

Finalmente se aplica sobre las unidades un proceso de ventaneo. Se sabe que la señal correspondiente a un sismo es no estacionaria, sin embargo, como el análisis de la señal se realiza en ventanas cortas de tiempo se puede suponer que los tramos son estacionarios dado que el contenido espectral de la señal varía lentamente en un lapso corto de tiempo. Se emplea una ventana suavizada en sus extremos como la de *Hamming* para evitar el efecto de borde. (Figura 4-5):

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N_w - 1}\right), & 1 \leq n \leq N_w - 1, \\ 0, & \text{en cualquier otro caso.} \end{cases} \quad (2.10)$$

Figura 4-5: Ventana de Hamming para 300 muestras.

En la Figura 4-6 se muestra un tramo de evento sísmico sin ventanear y el mismo segmento ventaneado con la función de Hamming.

Figura 4-6: Segmento sísmico sin ventanear (Izquierda), Segmento sísmico ventaneado (Derecha).

4.2. Representación de los registros

4.2.1. LPC para la representación de los registros

La representación de la señal por medio de los LPC se justifica porque reduce la dimensión de los vectores de entrada al sistema clasificador.

La representación de la señal en términos de sus LPC requiere determinar el orden adecuado de la predicción lineal de tal manera que el número de coeficientes no sea tan grande como para incluir detalles de la señal que entorpecerían la clasificación ni tan pequeño que resulte

en la pérdida de información relevante. La escogencia del orden se basa en el análisis de error de la representación: a medida que el orden de la predicción lineal aumenta, la varianza disminuye hasta un punto en que se torna asintótica, se trata de identificar el orden a partir del cual la varianza cambia muy poco, digamos un 2% a 5%.

4.2.2. Elección del tamaño de las unidades

La frecuencia de muestro pone un límite superior a la banda espectral de las señales. El límite inferior está determinado por el ancho de banda de los sensores que se utilice para registrarlas, que pueden ser de corto periodo o de banda ancha, estos últimos detectan armónicos con periodos de hasta 120 segundos en dos de las estaciones de la red de Galeras. Las señales usadas en este trabajo contienen información por debajo de 0.5 Hz, sin embargo frecuencias menores de este valor no poseen información de nuestro interés.

Para observar una frecuencia mínima de 0.5 Hz se debe tener una ventana de 2 segundos de registro, sin embargo, los espectros de señales capturadas con sensores banda ancha muestran frecuencias menores que ese valor. Una ventana de 3 segundos de duración nos permitiría ver armónicos con frecuencias de hasta 0.33 Hz, que es un límite razonable para el espectro de los registros volcánicos utilizados en este trabajo.

Entre los sismos que se puede detectar y clasificar están los de tipo Tornillo, y como se mencionó en el capítulo 1, estos se caracterizan por presentar contenido espectral monocromático o cuasi-monocromático con la mayoría de su contenido energético concentrado en el pico de frecuencia dominante. En el VG Galeras se han registrado tornillos con armónicos de hasta 0.7 Hz, por tanto, el tamaño de la ventana que se propone incluiría esta frecuencia.

Tamaños más grandes de ventana mejorarían la resolución en el dominio de la frecuencia, pero incrementarían el esfuerzo computacional del sistema, volviéndolo lento y la respuesta se alejaría más del “tiempo real”.

4.3. Desarrollo e implementación de la herramienta computacional

La herramienta computacional debe cumplir con las siguientes características:

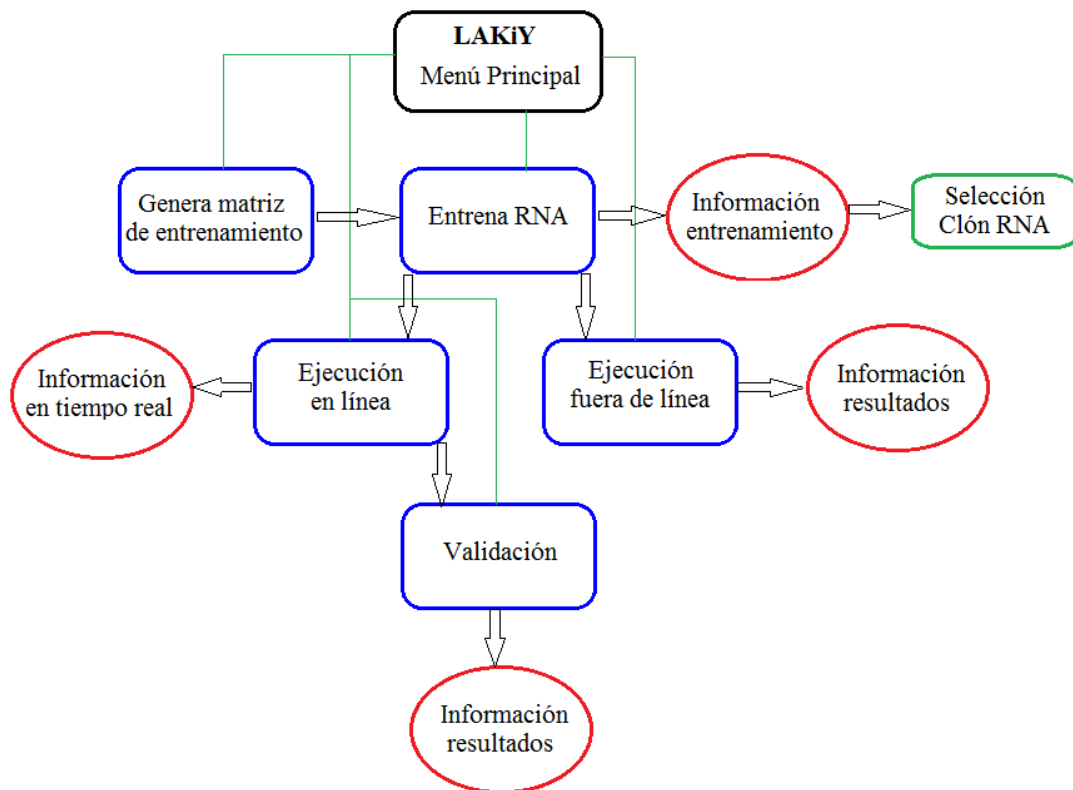
- Se debe contemplar cada una de las fases dentro de su desarrollo y crear módulos lo suficientemente flexibles de tal manera que facilite la manipulación de las variables en cada prueba que se ejecute. Elementos como los parámetros libres de la red, pre-procesamiento de las señales, representación de los registros y modo de entrenamiento, entre otros, son algunos de los factores a explorar para explotar eficientemente las RNA.
- Por otro lado, dada la cantidad de pruebas, variables a ajustar y sus respectivos resultados, es estratégico implementar una base de datos que almacene de manera sistemática las entradas y salidas del programa. Se plantea utilizar software no comercial tanto en el desarrollo del programa como en la base de datos.

4.3.1. El programa LAKiY

El programa desarrollado para este trabajo se ha llamado LAKiY, que en el idioma quechua significa clasificar. La guía de usuario se presenta en el ANEXO A. En el diseño y desarrollo del programa se utiliza la programación orientada a objetos a través del lenguaje libre JAVA (JDK 6.0 y JRE 6), con ayuda del entorno de desarrollo Netbeans IDE 6.9.1. Para las librerías no comerciales de RNA en JAVA se seleccionó el paquete **Joone 2.0.0RC1** creado por Jeff Heaton, cuya versión final es del año 2006. Joone consiste en un conjunto de *clases* basadas en *componentes* que deben ser enlazados adecuadamente para obtener los resultados deseados. Puede ser extendido construyendo nuevas arquitecturas de redes neuronales y algoritmos de aprendizaje.

En el programa LAKiY se crean clases para la aplicación de las diferentes etapas, como son: construcción de las matrices de entrenamiento, entrenamiento de las RNA, ejecución en línea, fuera de línea y validación. Se crea también una interfaz gráfica para facilitar su operación. El diagrama de la Figura 4-7 muestra la estructura básica del programa y a continuación se detallan cada uno de sus componentes.

Figura 4-7: Estructura del programa LAKiY.



- **Módulo de construcción de las matrices de entrenamiento**

Una vez seleccionados los segmentos de ruido, segmentos de sismos y sismos de los tipos VT, LP y TRE, se aplica el pre-procesamiento y extracción de los coeficientes de predicción lineal y se ubica cada vector de LPC. Las variables de entrada que se incluyen en este módulo son:

- Nombre de la matriz
- Ubicación de las señales
- Estación
- Tasa de muestreo
- Tasa de re-muestreo
- Duración de la unidad
- Número de unidades
- Orden de LPC
- Normaliza traza
- Normaliza LPC
- Duración traza
- Porcentaje de traslapo
- Relación entre objetivos y extensiones
- Ubicación matriz generada

Este módulo origina como salida una matriz, dentro de un archivo plano, que contiene los vectores de LPC intercalados entre los conjuntos que se van a clasificar con sus respectivos objetivos y ubica la matriz en el directorio que se le designe.

▪ **Módulo de entrenamiento de RNA**

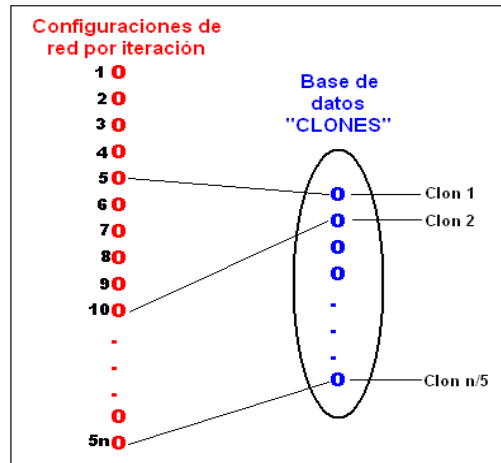
En este módulo se utilizan las matrices generadas en el paso anterior como entrada para las RNA. La configuración y los parámetros libres de las RNA pueden variarse. Las variables de entrada a este módulo son:

- Nombre del entrenamiento
- Selección de una matriz de entrenamiento
- Número de neuronas en la capa de entrada
- Número de neuronas en la capa oculta
- Número de neuronas en la capa de salida
- Modo de entrenamiento
- Tasa de aprendizaje
- Momentum
- Número de patrones de entrenamiento
- Número de iteraciones
- Número de patrones de validación
- Número de iteraciones por validación
- Ubicación de archivos de salida

Como salida este módulo arroja información gráfica y textual del proceso de entrenamiento. La elección de una configuración de RNA se hace a partir del análisis de los resultados por el método de ensayo-error. Uno de los primeros controles que se hace sobre cada entrenamiento, es el de seleccionar la mejor configuración de pesos sinápticos, para esto se plantea la generación de copias de red con diferentes configuraciones de pesos. A medida que un entrenamiento transcurre, sus errores, tanto global como de aprendizaje, varían debido al ajuste en los pesos sinápticos en cada iteración. Para determinar la mejor configuración dentro de un gran número de iteraciones se guardan *clones* de la red cada cierto número de

iteraciones dentro del que la configuración de pesos no cambia significativamente (Figura 4-7) y se escoge haciendo un análisis gráfico de errores.

Figura 4-8: Esquema de salvado de diferentes configuraciones de red a medida que avanza el proceso de entrenamiento, en este ejemplo se guarda un clon cada 5 iteraciones.



Otro de los factores determinante para conseguir buenos resultados en la aplicación de las RNA es la cantidad de neuronas en cada capa. Aquí se trata de encontrar una red con la menor cantidad de neuronas que realice el trabajo de discriminación. Las pruebas inician con una capa oculta en la que el número de neuronas varía desde una hasta un número en el que los resultados no mejoren considerablemente, teniendo en cuenta no incurrir en el sobreentrenamiento. En cuanto a los objetivos, para el caso de la discriminación entre unidades de ruido y sismo se utiliza un vector con un solo elemento, 0 para ruido y 1 para unidad de sismo. Para el caso de reconocimiento entre los tipos LP y VT se propone, 0 para LP y 1 para VT.

Otros parámetros ajustables en los entrenamientos son la *tasa de aprendizaje* y el *momentum*. Para la tasa de aprendizaje es común el uso de valores entre 0.05 y 0.75 (Bishop, 1995), aunque su cota máxima es 1, sin embargo un valor cercano a 1 podría producir que el algoritmo de aprendizaje se salte mínimos locales o un mínimo global, por el contrario un valor cercano a su cota inferior podría producir desplazamientos muy pequeños sobre la superficie de error a tal punto de no encontrar ningún mínimo. Por otro lado, valores recomendados para el *momentum* varían entre 0.3 y 0.5 siendo sus límites 0 y 1; este parámetro ayuda a la velocidad de convergencia y a evitar mínimos locales (Bishop, 1995). El núcleo de control de aprendizaje de las RNA es el proceso de validación. En este proceso se utilizan registros con los que la red no fue entrenada para cuantificar los resultados del entrenamiento.

▪ Módulo de ejecución fuera de línea

Este módulo está pensado para clasificar los diferentes conjuntos en un ambiente controlado, sirve especialmente para probar la efectividad de las redes entrenadas. Este módulo es clave en la estrategia que se plantea en el sexto paso de la metodología para validar los resultados. Las entradas a este módulo son:

- Nombre de del clon de RNA seleccionado
- Ubicación de la matriz que contiene los elementos a clasificar
- Número de elementos a clasificar
- Ubicación del archivo de salida

Como salidas el módulo despliega información gráfica y textual de la clasificación.

▪ **Módulo de ejecución en línea**

Las RNA seleccionadas tanto para detección como para clasificación de sismos se ejecutan en tiempo real o tiempo pasado extrayendo las trazas desde un servidor de ondas que puede ser el *Waveserver* que le pertenece al sistema EW o el *Winston* que genera una base de datos de formas de onda a través del gestor *MySQL*. Este último puede contener información desde que se puso en funcionamiento, para el caso del OVSP está activo desde abril de 2010.

En el módulo de ejecución en línea se configura un set que corresponde a una pareja de RNA, una especializada en detección y otra en clasificación de los tipos LP y VT. Las entradas a este módulo son:

- Nombre del Set
- IP del servidor de ondas (*Waveserver* o *Winston*)
- Puerto por donde se transmite la información
- Estación
- Nombre de la RNA para detección
- Nombre de la RNA para clasificación LP, VT
- Ubicación de un archivo de resultados
- La salida de este módulo corresponde a información gráfica y textual

▪ **Módulo de validación**

Este módulo aprovecha la propiedad que tiene el *Waveserver* de almacenar datos de días pasados, que dependiendo de la configuración de su buffer puede contener información de hasta 20 días aproximadamente (para el sistema de adquisición del OVSP), si el tiempo que se quiere consultar excede 20 días es posible hacerlo mediante el servidor de ondas *Winston* que como ya se mencionó contiene información desde que fue implementado. En este módulo se puede validar la detección y clasificación utilizando sets de RNA diferentes al que se utilizó en línea tiempo real. Este módulo emula el tiempo real pero con una fecha y hora de inicio pasadas. Sus entradas son:

- Fecha
- Hora
- Set de clasificación
- Sus salidas son de manera gráfica y textual iguales a las utilizadas en línea tiempo real.

4.3.2. Implementación de una base de datos

La cantidad de pruebas requiere el almacenamiento de un gran volumen de datos de entrada y de salida (respuestas y sus errores y tiempos asociados) y la operación con una gran cantidad de combinación de variables y un gran volumen de resultados, tanto intermedios como finales. Para ello se ha implementado una base de datos relacional utilizando el gestor *PostgreSQL 8.1.4* que está conectada al sistema de detección y clasificación en todas sus etapas donde se escribe la información. En el ANEXO B se detalla la estructura de la base de datos.

4.4. Determinación de topologías adecuadas para la detección y clasificación

La experiencia de otros autores (e.g., del Pezzo, Rugeles & Cadena, Cadena & Arcos) sugiere que la arquitectura del tipo *perceptrón multicapa con una capa oculta* es adecuada para clasificar señales sísmicas volcánicas, ya que los porcentajes de acierto alcanzados en estos trabajos supera el 90%.

Ya que el número de neuronas en la capas oculta de una red define la complejidad de la *hiper-superficie* que constituye el límite entre los conjuntos a separar (Del Brío, 2002), es necesario determinar el número de neuronas en la capa oculta que permita un mejor desempeño de la RNA, haciendo un balance entre su capacidad de acertar en la detección y clasificación y su capacidad de generalización.

El número óptimo de neuronas se determina por ensayo y error ya que no existe una regla que permita agilizar este proceso (Del Brío, 2002). Esto se realiza fuera de línea, clasificando segmentos de ruido y segmentos de sismos para el caso de la detección y eventos completos de los tipos LP y VT para el caso de la clasificación. El proceso se repite dos o tres veces para evitar los efectos de inicialización de los pesos sinápticos.

4.4.1. Selección de los datos

Se utilizan los siguientes tipos de señales:

- Segmentos de ruido
- Segmentos de sismos volcánicos de los tipos LP, VTA y TRE.
- Sismos tipo LP
- Sismos tipo VTA
- Sismos tipo TRE

Las señales seleccionadas se encuentran en el periodo comprendido entre enero de 2007 y agosto de 2010.

Los archivos que contienen segmentos de sismos se generan extrayendo pedazos de eventos de los tipos LP, VT, TRE y no clasificables. Se deben generar segmentos que contengan información del inicio de los sismos y otros con información de la coda, tratando de extraer señal sísmica hasta donde se distinga del ruido base.

Los archivos que contienen sismos de los tipos LP o VT deben contener un solo evento y se extraen de la base de eventos ya clasificados, de esta manera se asegura que en el proceso de entrenamiento se transmita la experiencia de la clasificación compilada por los temáticos durante los últimos años. En lo posible se escogen eventos de los que no se tenga dudas del tipo al que pertenecen.

Los trabajos de Del Pezzo, et al., (2003); Rugeles & Cadena, (2006); Cadena & Arcos (2008), obtuvieron buenos resultados con un número igual o menor a 500 ejemplos por tipo para discriminar dos conjuntos diferentes, sin embargo, lo mejor es tener la mayor cantidad de ejemplos posible, de esta manera la cantidad de registros extraídos para este estudio se muestra en la Tabla 4-1.

Tabla 4-1: Número de ejemplos extraídos.

Tipo de Señal	Número de ejemplos
Segmentos de ruido	3120
Segmento de sismos	1550
Sismo tipo LP	450
Sismo tipo VTA	494
Sismo tipo TRE	523

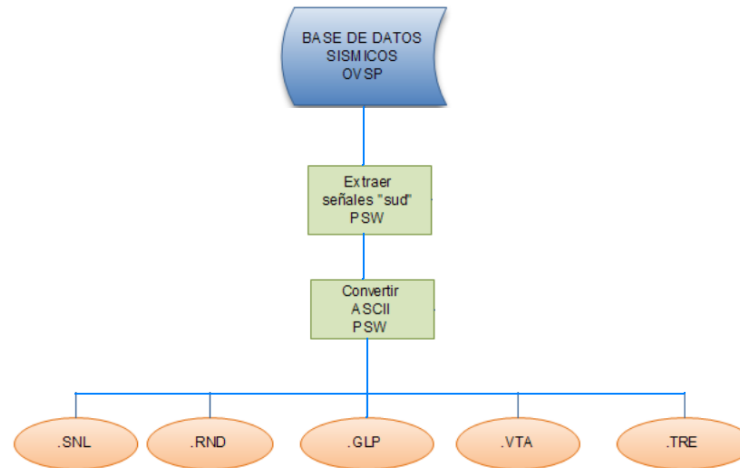
▪ Aspectos operacionales de los datos

En la base de datos del OVSP las trazas se encuentran almacenadas en formato *suds* demultiplexado, una de las ventajas de este formato es que en un mismo archivo está contenida la información de 2 minutos de registro de las estaciones de la red. Con ayuda del programa PSW (Raigoza, 2009) se transforma cada uno de estos archivos al formato ASCII. (Figura 4-8).

Para identificar el tipo de información que contiene cada archivo se agregó una extensión a la nomenclatura del nombre. El nombre de cada archivo contiene información de la fecha y hora del registro, las extensiones utilizadas se muestran en Tabla 4-2.

Tabla 4-2: Extensión que identifica los diferentes tipos de señales.

Tipo de señal	Extensión
Segmento de ruido	.RND
Segmento de sismos	.SNL
Sismo tipo LP	.GLP
Sismo tipo VTA	.GVA
Sismo tipo TRE	.GTR

Figura 4-9: Esquema de extracción y conversión de señales desde la base de datos del OVSP.

Así un archivo que contenga un segmento de ruido registrado el día 20 de febrero a las 05:23 a.m. llevaría por nombre "02200523.RND".

Esta nomenclatura no solo permite la organización de los mismos, sino que facilita las fases de entrenamiento y ejecución ya que a través de las extensiones es posible relacionar el tipo de señal con los objetivos de las RNA.

4.4.2. Generación de las matrices de entrenamiento

Los vectores que contienen la información espectral de los sismos se incluyen en matrices con un formato específico para la aplicación de la fase de entrenamiento. Se deben generar lotes para realizar pruebas del tipo ensayo-error tanto para la fase de detección como para la fase de clasificación.

Adicional a las matrices de entrenamiento es necesario generar matrices de prueba, estas se aplican en la etapa de validación de las redes seleccionadas. Los ejemplos para la etapa de validación también pueden formar parte de las matrices de entrenamiento, esto es, en una sola matriz se pueden contener los ejemplos de entrenamiento y validación.

Las matrices de entrenamiento y validación deben contener los objetivos para cada uno de los vectores, el tamaño de los vectores de objetivos define el número de neuronas en la capa de salida. Ya que la función de transferencia de las neuronas de salida es de tipo sigmoideo los objetivos estarán limitados al intervalo [0,1] y generalmente se escogen como objetivos de cada conjunto los límites del intervalo.

Bishop, (1995) recomienda ordenar los vectores en la matriz de entrenamiento de manera aleatoria o intercalada para evitar vicios en el proceso de entrenamiento de RNA.

En esta fase se entrenan diferentes topologías de red. Dentro de valores razonables se varía el número de neuronas en la capa oculta, el número de iteraciones, la *tasa de aprendizaje*, el *momentum*, el número de patrones de validación, etc. El resultado del proceso de entrenamiento es un conjunto de RNA para detección y otro para clasificación que posteriormente serán sometidos a un proceso de validación.

4.5. Validación de resultados

La estrategia para validar los resultados tiene dos posibilidades, o clasificar fuera de línea con ejemplos que no se incluyeron en las matrices de entrenamiento de los cuales se tiene certeza del tipo al que pertenecen o clasificar en línea y comparar los resultados con la clasificación que los temáticos concluyan. La desventaja de la segunda opción es que no se tiene control sobre el tipo de señales que se le presentan a la red, aspecto que si se puede controlar en la primera trabajando con ejemplos del pasado de los cuales se tiene certeza del tipo al que pertenecen.

5. Resultados

5.1. Representación de las señales

Para establecer el mejor orden de LPC para representar las señales se escogió una muestra de 40 unidades de ruido y 40 unidades de sismos, se calculó la varianza para los segmentos entre los órdenes 2 y 25 (el valor del primer orden es siempre 1) (Figura 5-1 y 5-2), se extrajo un promedio de las varianzas para cada orden, se calculó la correspondiente desviación estándar y se cuantificó la diferencia de las varianzas entre un orden y el siguiente en términos porcentuales. Los resultados se resumen en la Tabla 5-1.

Figura 5-1: Varianza vs. orden de LPC para una muestra de 40 unidades de sismos, cada curva representa una unidad.

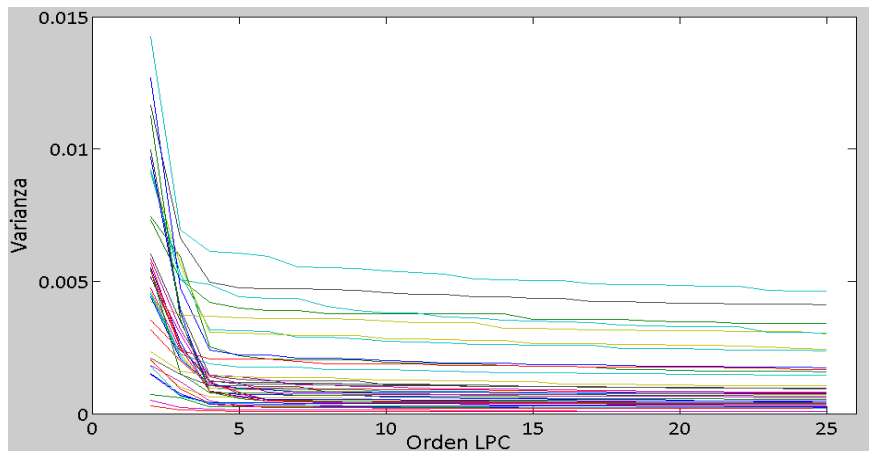


Figura 5-2: Varianza vs. orden de LPC para una muestra de 40 unidades de ruido, cada curva representa un unidad.

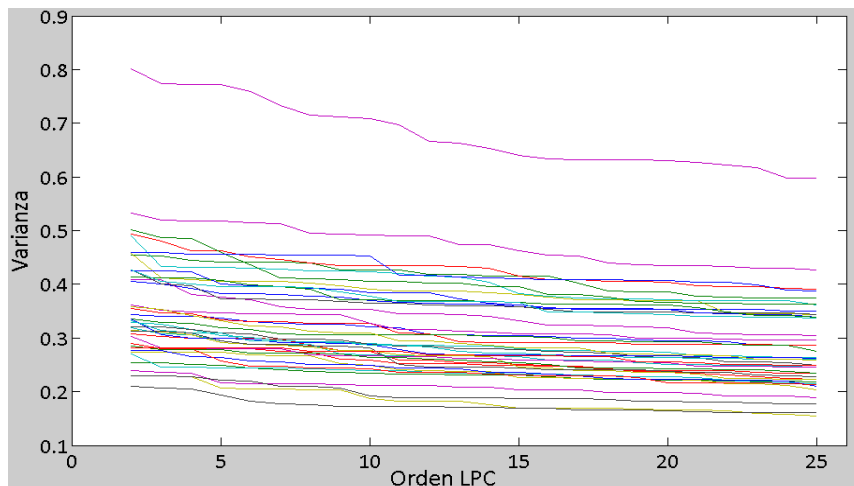


Tabla 5-1: Resultados del análisis para establecer un orden de LPC.

Orden LPC	Unidades de sismos				Unidades de ruido			
	Promedio varianza	Desviación estándar	Diferencia	Diferencia (%)	Promedio varianza	Desviación estándar	Diferencia	Diferencia (%)
2	0.0054	0.0035			0.3611	0.1100		
3	0.0028	0.0018	0.0026	100.0	0.3512	0.1048	0.0099	100.0
4	0.0016	0.0014	0.0012	46.1	0.3473	0.1041	0.0039	39.2
5	0.0015	0.0014	0.0001	4.9	0.3399	0.1051	0.0074	74.1
6	0.0014	0.0014	0.0001	2.3	0.3355	0.1038	0.0044	44.3
7	0.0014	0.0014	0.0000	1.7	0.3319	0.1012	0.0036	36.4
8	0.0013	0.0014	0.0000	1.1	0.3283	0.0989	0.0036	36.1
9	0.0013	0.0013	0.0000	0.7	0.3246	0.0986	0.0037	37.2
10	0.0013	0.0013	0.0000	1.2	0.3211	0.0991	0.0035	34.9
11	0.0013	0.0013	0.0000	0.5	0.3153	0.0982	0.0058	58.7
12	0.0013	0.0013	0.0000	0.5	0.3126	0.0951	0.0027	26.7
13	0.0012	0.0013	0.0000	0.8	0.3101	0.0941	0.0025	24.7
14	0.0012	0.0013	0.0000	0.7	0.3078	0.0930	0.0023	23.2
15	0.0012	0.0012	0.0000	0.7	0.3040	0.0910	0.0038	38.6
16	0.0012	0.0012	0.0000	0.4	0.3007	0.0890	0.0033	33.5
17	0.0012	0.0012	0.0000	0.6	0.2987	0.0888	0.0019	19.2
18	0.0012	0.0012	0.0000	0.6	0.2965	0.0882	0.0022	22.2
19	0.0012	0.0012	0.0000	0.3	0.2946	0.0884	0.0020	20.0
20	0.0012	0.0012	0.0000	0.4	0.2934	0.0885	0.0011	11.4
21	0.0011	0.0012	0.0000	0.2	0.2913	0.0877	0.0021	20.8
22	0.0011	0.0012	0.0000	0.3	0.2887	0.0874	0.0027	27.1
23	0.0011	0.0012	0.0000	0.6	0.2870	0.0870	0.0017	16.8
24	0.0011	0.0012	0.0000	0.3	0.2852	0.0851	0.0018	18.0
25	0.0011	0.0011	0.0000	0.3	0.2831	0.0853	0.0021	21.5

El comportamiento de las varianzas es diferente para las unidades de ruido que para las unidades de sismos (Figura 5-3). Para el primer caso la curva de varianza promedio decrece suavemente a medida que el orden de la predicción se incrementa, mientras que para las unidades de sismos la curva cae rápidamente hasta el orden 5 (Figura 5-3). Por otro lado, la curva de las diferencias entre varianzas consecutivas muestra un comportamiento oscilante para las unidades de ruido, aunque su tendencia es de decrecimiento. Para las unidades de sismos esta curva se muestra más estable y tiene un fuerte decaimiento para los órdenes más bajos (Figura 5-4).

Figura 5-3: Comportamiento de la varianza promedio de la muestra de unidades de sismos y de ruido respecto al orden de la predicción.

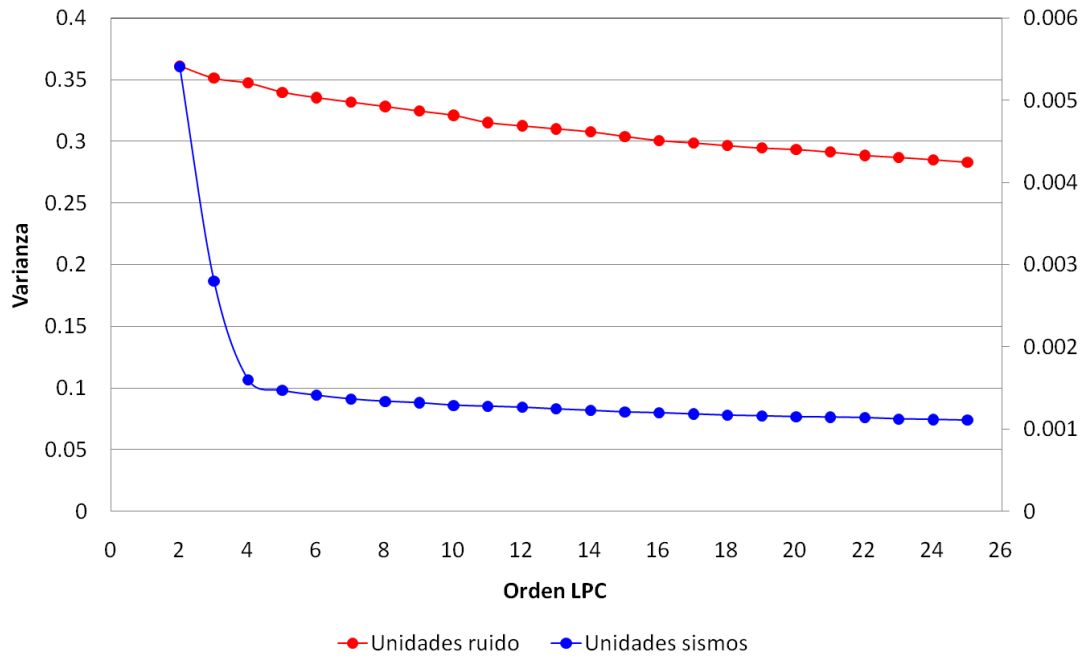
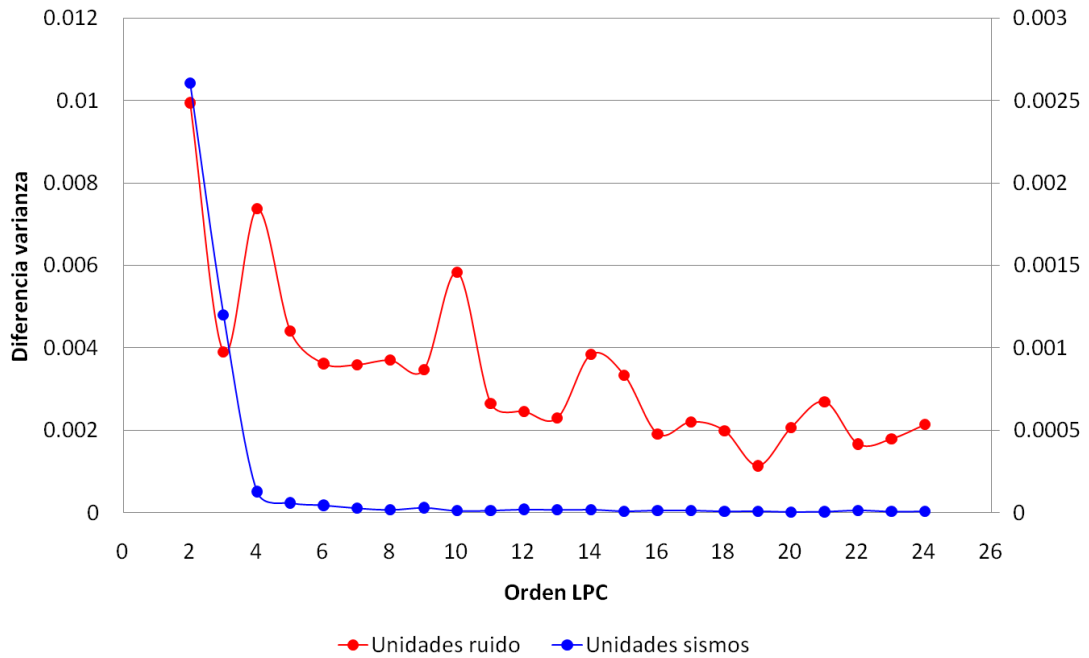


Figura 5-4: Comportamiento de las diferencias entre varianzas promedio sucesivas de la muestra de unidades de sismos y de ruido respecto al orden de la predicción. Obsérvese que después del orden 5 la curva es asintótica para el caso de las unidades de sismos.



5.2. Elección de la topología de las RNA

5.2.1. Unidades de Sismos vs unidades de ruido

Para entrenar las RNA se utilizaron 6 matrices de entrenamiento, en todas se incluyeron las mismas unidades de sismos, pero unidades diferentes de ruido, esto con el fin de construir matrices que incluyeran ruido actual para transmitirle a las RNA información lo más actualizada posible y así mejorar su desempeño en tiempo real.

Cada matriz fue construida con vectores de LPC de diferentes órdenes, 5, 6 y 7. La Tabla 5-2 muestra la estructura de las matrices de entrenamiento.

Tabla 5-2: Configuración de las matrices de entrenamiento de reconocimiento entre unidades de ruido y sismos.

Nombre matriz	Unidades ruido (2009-2010)	Unidades ruido (2011)	Unidades sismos	Total unidades	Orden LPC
SRL01.mtz	1397	0	1397	2794	5
SRL02.mtz	1397	0	1397	2794	6
SRL03.mtz	1397	0	1397	2794	7
SRS07.mtz	712	800	1512	3024	5
SRS08.mtz	712	800	1512	3024	6
SRS09.mtz	712	800	1512	3024	7

Los objetivos de estas matrices fueron: 0 para unidades de ruido y 1 para unidades de sismos.

En esta fase no se normalizaron los vectores ya los valores extremos del intervalo en el que se normaliza están presentes en todos los vectores de ambos conjuntos, esto no contribuye a diferenciarlos. La normalización podría utilizarse para vectores de órdenes mayores sin pérdida de información relevante.

Un 20 % de los elementos de cada matriz se utiliza como conjunto de validación y 50 elementos por matriz no fueron utilizados en el entrenamiento con el objetivo de usar estas unidades en la fase ejecución de las RNA.

Se entrenaron 25 RNA con una capa oculta. El número de neuronas de entrada varía entre 5 y 7 dependiendo del orden de la predicción, y se mantiene el número de neuronas ocultas por debajo del número de neuronas de entrada. Los valores de la tasa de aprendizaje y momentum se fijaron ambos en 0.5, el número de iteraciones en 1000 y el número de iteraciones por cada validación en 10.

La Tabla 5-3 y la Figura 5-5 muestran el comportamiento de los mínimos errores global y de generalización por entrenamiento.

Se observa que el error de generalización crece para las redes que fueron entrenadas con matrices que contienen ruido de 2011, lo cual es razonable ya que contienen mayor diversidad de ejemplos, por tanto, las RNA tienen mayor dificultad para ajustar sus pesos

sinápticos (Figura 5-5). También se observa que el error global tiende a decrecer a medida que el número de neuronas de entrada se incrementa.

Los mejores resultados en términos del error de generalización se obtuvieron para las redes con las siguientes topologías: 5 - 2 - 1, 5 - 3 - 1, 5 - 4 - 1 y 7 - 4 - 1, (neuronas de entrada - neuronas ocultas - neuronas de salida), correspondientes a los entrenamientos SRL01, SRL02, SRL03 y SRS10. De estas topologías las dos últimas alcanzaron los valores mínimos en el error de generalización y tienen en común una capa oculta con cuatro neuronas.

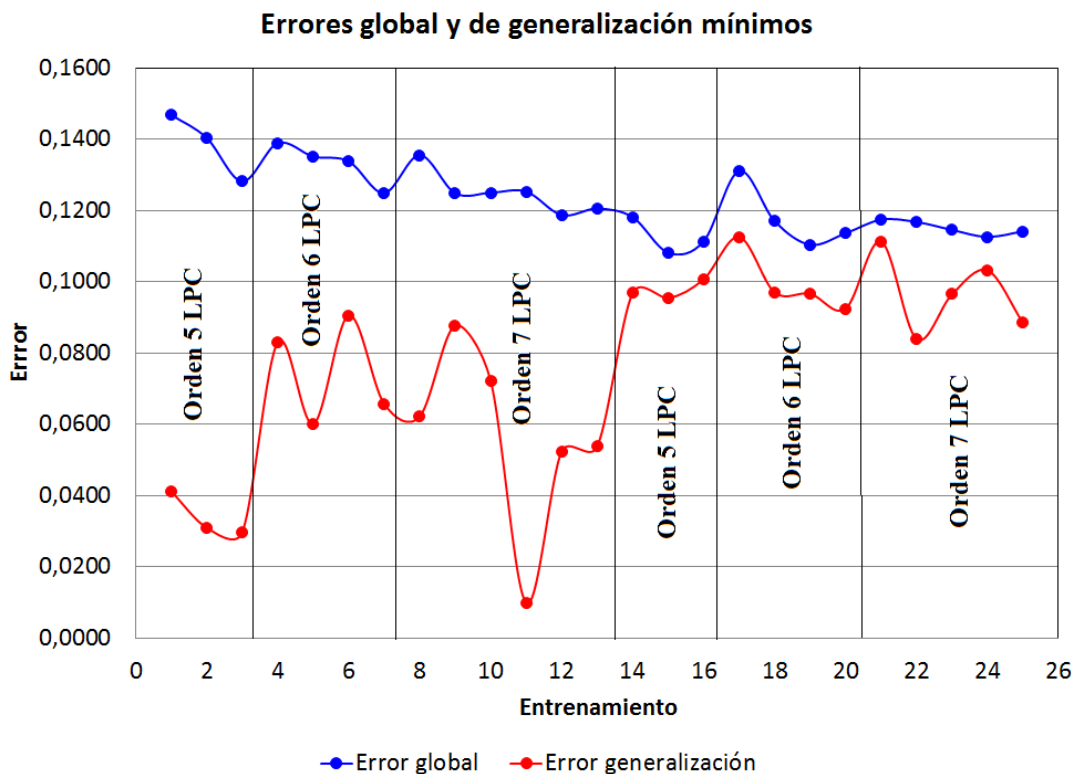
Tabla 5-3: Valores de los errores global y de generalización para las RNA entrenadas. Se han marcado de color cian los entrenamientos con los errores mínimos y de rojo los entrenamientos con los errores máximos.

Consecutivo	Nombre entrenamiento	Matriz de entrenamiento utilizada	Error global (mínimos)	Error generalización (mínimos)
1	SRL01	SRL01.mtz	0,1469	0,0412
2	SRL02	SRL01.mtz	0,1404	0,0310
3	SRL03	SRL01.mtz	0,1282	0,0296
4	SRL04	SRL02.mtz	0,1389	0,0830
5	SRL05	SRL02.mtz	0,1352	0,0601
6	SRL06	SRL02.mtz	0,1338	0,0906
7	SRL07	SRL02.mtz	0,1249	0,0657
8	SRL08	SRL03.mtz	0,1355	0,0622
9	SRL09	SRL03.mtz	0,1249	0,0878
10	SRL09b	SRL03.mtz	0,1250	0,0722
11	SRL10	SRL03.mtz	0,1253	0,0099
12	SRL11	SRL03.mtz	0,1187	0,0524
13	SRL12	SRL03.mtz	0,1205	0,0537
14	SRS26	SRL07.mtz	0,1181	0,0969
15	SRS27	SRL07.mtz	0,1082	0,0954
16	SRS28	SRL07.mtz	0,1114	0,1007
17	SRS29	SRL08.mtz	0,1311	0,1126
18	SRS30	SRL08.mtz	0,1170	0,0970
19	SRS31	SRL08.mtz	0,1104	0,0966
20	SRS32	SRL08.mtz	0,1136	0,0924
21	SRS33	SRL09.mtz	0,1174	0,1114
22	SRS34	SRL09.mtz	0,1168	0,0839
23	SRS35	SRL09.mtz	0,1145	0,0965
24	SRS36	SRL09.mtz	0,1126	0,1032
25	SES37	SRL09.mtz	0,1142	0,0884

A pesar de que el error de generalización mínimo corresponde a la topología 7-4-1, al escoger la mejor red debe considerarse además la “calidad” del entrenamiento. En este ejemplo debe tenerse en cuenta que las RNA que fueron entrenadas con ruido más variado (2009 a 2011) tienen más *experiencia* y en consecuencia son más tolerables a cambios en las características del ruido. En este sentido la mejor topología de las redes entrenadas con mayor variedad de ruido está asociada al entrenamiento SRS34 cuya topología es 7 - 4 - 1, marcada con azul en la Tabla 5-3.

Siete coeficientes de predicción lineal con cuatro neuronas en la capa oculta son la mejor combinación para discriminar entre unidades de ruido y unidades de sismos.

Figura 5-5: Comportamiento de los errores mínimos global y de generalización de cada entrenamiento.



Finalmente, en la Figura 5-6 se muestra el comportamiento de los errores global y de generalización para el entrenamiento con los mejores resultados. De estas curvas se extrae el *clon* de red que se escoge en base al mínimo error de validación cruzada. Para el entrenamiento SRL34 se escogió el clon 710. En la gráfica se observa un comportamiento oscilante en el error de generalización, sin embargo después del clon 710 la tendencia es ascendente. El error global tiene un comportamiento claramente descendente en las primeras 100 iteraciones, luego tiende al descenso lentamente.

5.2.2. Clasificación VT vs LP

En este caso para definir una topología de RNA adecuada se entrenaron redes con 13 matrices de entrenamiento, todas con igual cantidad de elementos, 450 del tipo VT y 450 LP, de los cuales 50 elementos no se utilizaron en el proceso de entrenamiento y se destinaron para las pruebas de efectividad de las RNA. De los 850 elementos restantes, 170 (20 % del total disponible para los entrenamientos) se utilizaron para formar la matriz de validación. La Tabla 5-4 muestra las características con las que se generaron las matrices de entrenamiento.

Figura 5-6: Curvas de error global y de generalización para el entrenamiento SRS34.

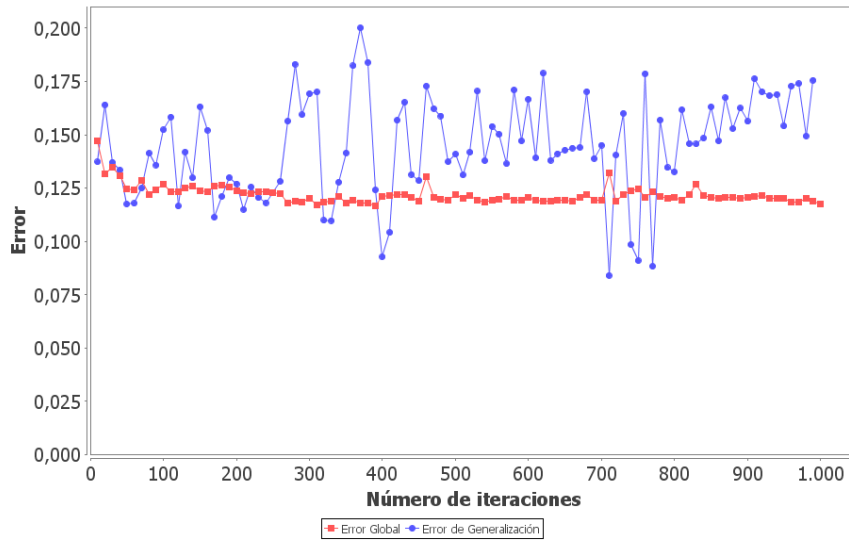


Tabla 5-4: Configuración de las matrices de entrenamiento para clasificación entre LP y VT.

Consecutivo	Nombre matriz	Duración segmento	Duración unidad	Número unidades	Orden LPC	Normalización LPC
1	LV01.mtz	20	3	7	5	[0,1]
2	LV02.mtz	20	3	7	7	[0,1]
3	LV02b.mtz	20	3	7	7	[-1,1]
4	LV02c.mtz	20	2	11	7	[-1,1]
5	LV02d.mtz	20	2	11	9	[-1,1]
6	LV02e.mtz	20	2	11	8	[-1,1]
7	LV03.mtz	20	3	7	9	[0,1]
8	LV04.mtz	25	3	9	5	[0,1]
9	LV05.mtz	25	3	9	7	[0,1]
10	LV06.mtz	25	3	9	9	[0,1]
11	LV07.mtz	30	3	11	5	[0,1]
12	LV08.mtz	30	3	11	7	[0,1]
13	LV09.mtz	30	3	11	9	[0,1]

Todas las trazas que conformaron las matrices de entrenamiento fueron normalizadas. Se incluyeron unidades de 2 segundos de duración y segmentos de traza que variaron entre 20 y 30 segundos.

La elección de la duración del segmento de traza se basó en un análisis de duraciones de los eventos tipo LP Y VT registrados históricamente en el OVSP. Se encontró que el 80 % de los sismos VT tienen duraciones menores que 21s y el 90 % menor que 30s (Figura 5-7). Para la sismicidad tipo LP el 90 % de los eventos tiene duración menor que 45s y el 72 % menor que 30s (Figura 5-8). Por tanto, al tomar segmentos no mayores que 30s se evita incluir información de ruido en un 90 % de los eventos tipo VT y se incluye la totalidad de la traza en el 72 % de los sismos tipo LP. La información espectral extraída de un segmento de entre 20 a 30 segundos es suficiente para lograr una adecuada representación en caso de que se evalúen sismos LP o VT con duraciones mayores.

En esta etapa, de clasificación entre LP y VT se entrenaron 31 RNA con las matrices construidas, se utilizó una capa oculta de 3 ó 5 neuronas en la mayoría de las topologías y se realizaron tres pruebas especiales con una capa oculta compuesta por 20 y 50 neuronas para contrastar con topologías más pequeñas. La Tabla 5-5 y la Figura 5-9 muestran los resultados de los errores global y de generalización para cada entrenamiento.

Figura 5-7: Diagrama de frecuencias que muestra la cantidad de sismos VT por duración, se ha tomado un valor por cada múltiplo de 3 segundos.

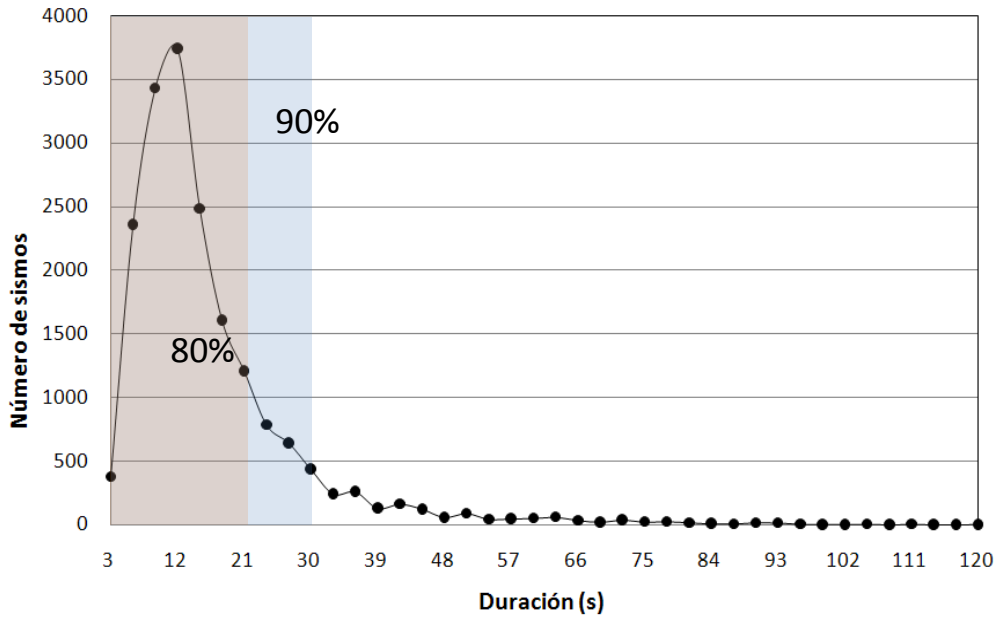


Figura 5-8: Diagrama de frecuencias que muestra la cantidad de sismos LP por duración, se ha tomado un valor por cada múltiplo de 3 segundos.

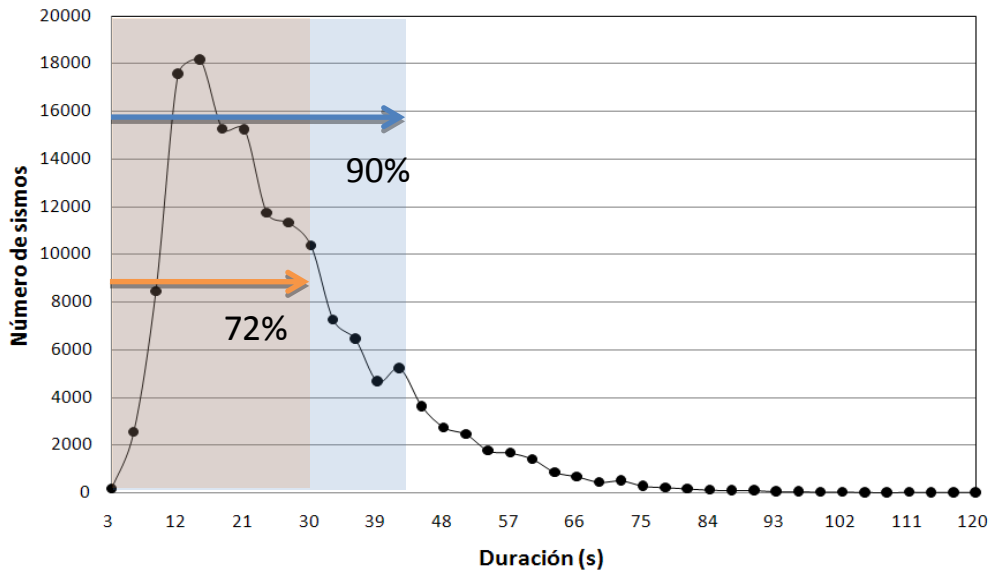
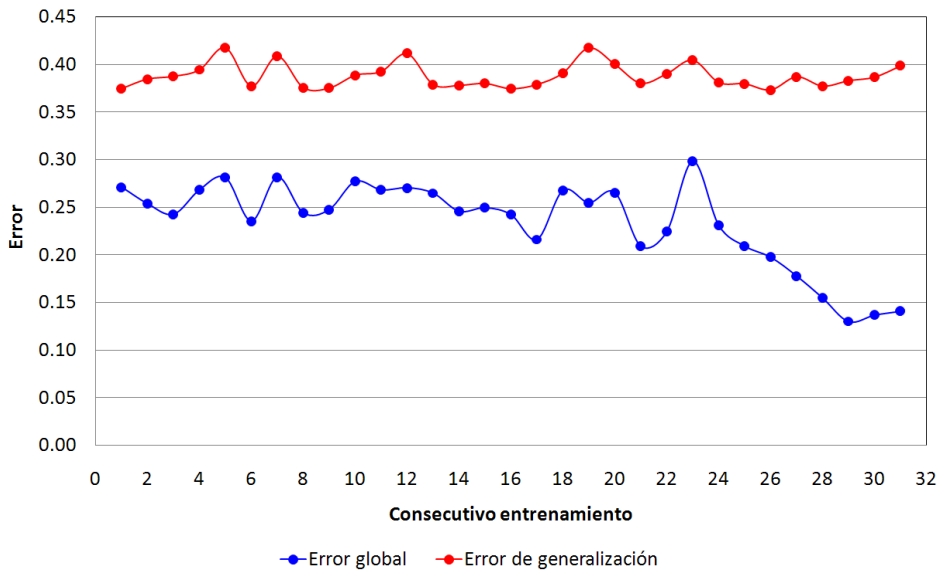


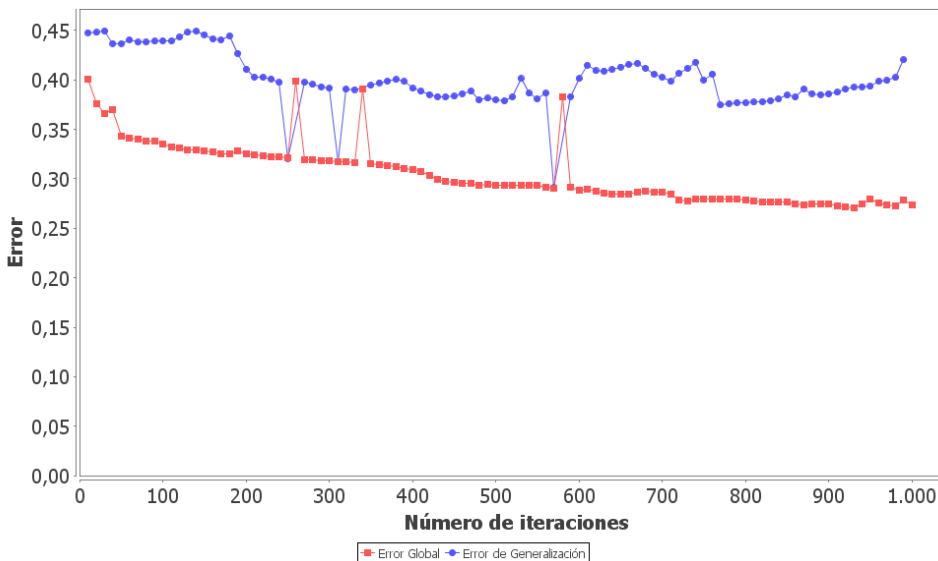
Tabla 5-5: Valores mínimos de los errores global y de generalización para cada entrenamiento, las filas en color naranja indican las mejores redes con base en el criterio de mínimo error de generalización.

Consecutivo	Nombre Entrenamiento	Matriz de entrenamiento utilizada	Error global	Error generalización
1	ELV01B	LV01.mtz	0,2711	0,3745
2	ELV02B	LV02.mtz	0,2537	0,3846
3	ELV03	LV03.mtz	0,2425	0,3875
4	ELV04	LV04.mtz	0,2681	0,3945
5	ELV05	LV05.mtz	0,2810	0,4179
6	ELV06	LV06.mtz	0,2354	0,3772
7	ELV07	LV07.mtz	0,2815	0,4090
8	ELV08	LV08.mtz	0,2440	0,3756
9	ELV09	LV09.mtz	0,2475	0,3751
10	ELV02Ba	LV02.mtz	0,2769	0,3887
11	ELV02Bb	LV02.mtz	0,2681	0,3926
12	ELV02Bc	LV02.mtz	0,2703	0,4117
13	ELV02Bd	LV02.mtz	0,2647	0,3789
14	ELV06a	LV06.mtz	0,2457	0,3779
15	ELV06b	LV06.mtz	0,2494	0,3799
16	ELV06c	LV06.mtz	0,2421	0,3744
17	ELV06d	LV06.mtz	0,2161	0,3789
18	ELV10	LV02.mtz	0,2676	0,3911
19	ELV10a	LV02.mtz	0,2547	0,4178
20	ELV10b	LV02.mtz	0,2653	0,4003
21	ELV11	LV02.mtz	0,2090	0,3803
22	ELV11a	LV02.mtz	0,2244	0,3903
23	ELV12	LV09.mtz	0,2981	0,4047
24	ELV13	LV02b.mtz	0,2310	0,3811
25	ELV14	LV02b.mtz	0,2091	0,3796
26	ELV14b	LV02b.mtz	0,1975	0,3732
27	ELV15	LV02c.mtz	0,1779	0,3871
28	ELV15a	LV02c.mtz	0,1549	0,3772
29	ELV16	LV02d.mtz	0,1301	0,3829
30	ELV16a	LV02d.mtz	0,1369	0,3867
31	ELV17	LV02e.mtz	0,1409	0,3986

Las RNA que mostraron mejores resultados tienen las siguientes topologías: 35 - 3-1 con un segmento de traza de 20s, 81 - 3-1 con un segmento de traza de 25s y 49 - 3-1 con un segmento de traza de 20s. En la Figura 5-9 se observa que el error de generalización no cambia considerablemente entre uno y otro entrenamiento. Se observa que para los seis últimos entrenamientos el error global tiende a descender, sin embargo eso no asegura que la red tendrá un mejor desempeño. Las RNA con 20 neuronas ocultas (consecutivo 21 y 22) y 50 neuronas ocultas (consecutivo 23) no muestran diferencias respecto a las de topología más pequeña.

Figura 5-9: Errores mínimos de generalización y error global en la etapa de entrenamiento.


Dos de las mejores RNA utilizan un segmento de traza de 20s y el orden de la predicción de cada unidad es de 5 ó 7 LPC; la duración de la unidad fue de 3s. Basados en el criterio de mínimo error de generalización, esta observación sugiere que las topologías más pequeñas arrojan mejores resultados. En la Figura 5-10 se muestra la variación del error del entrenamiento con topología 35 - 3 - 1. En la gráfica se observa que el error global es descendente hasta el final de las iteraciones, sin embargo el error de generalización varía mostrando algunos mínimos locales. Se observa que alrededor de los clones 250 y 600 el entrenamiento sacrifica error global por mejorar el error de generalización.

Figura 5-10: Variaciones del error global y de generalización con el número de iteraciones para el entrenamiento con RNA de topología 35 - 3 - 1.


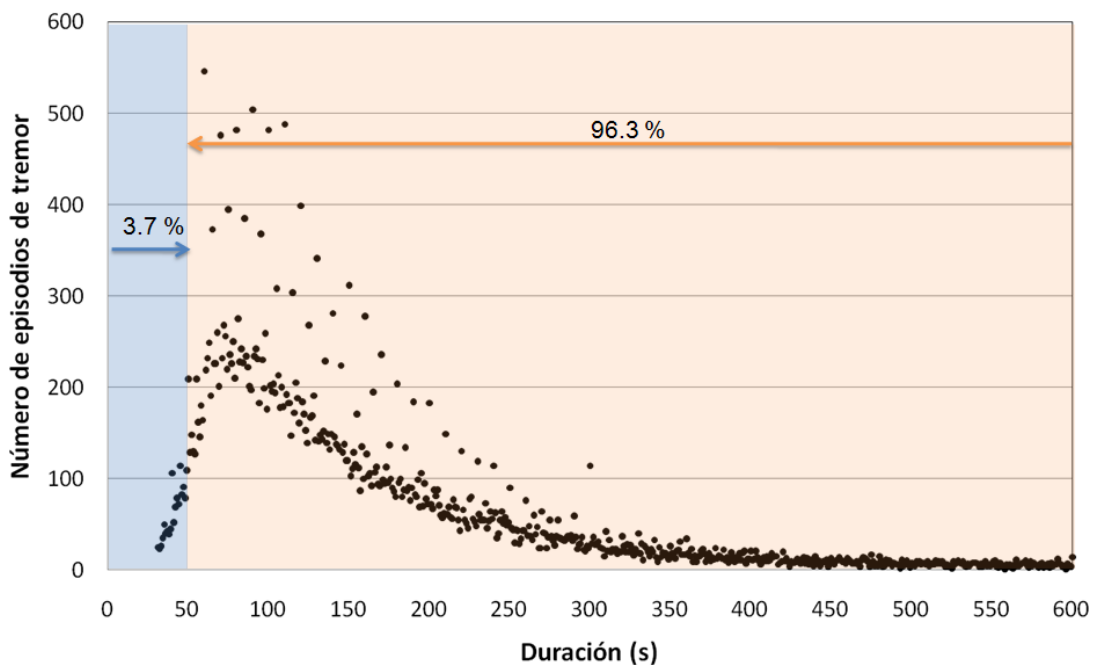
5.2.3. Clasificación de Tremor

El reconocimiento de los sismos TRE depende de la persistencia de la fuente, lo cual es medible en términos de la duración de los eventos. La Figura 5-11 muestra la distribución de la cantidad de estos eventos respecto a la duración. Se observa que el 3.7% de los eventos tiene una duración menor o igual a 51s, y el 96.3% de TRE tiene duración mayor o igual a 51s. Por tanto, dentro de un margen de error aceptable un episodio de tremor se reconoce cuando el número de unidades consecutivas clasificadas como segmentos de un sismo es mayor o igual a 17, que equivalen a 51 segundos de duración.

En el CD adjunto a este documento, en el directorio [Videos/TRE](#), se muestra el ejercicio de clasificación ejecutado en línea para 10 episodios de tremor ocurridos entre enero y marzo de 2011. Las RNA encargadas de clasificar entre unidades de ruido y unidades de sismos no interrumpen la detección de una señal continua asociada con un sismo de fuente persistente, acertando en el aviso de la presencia de sismos tipo TRE.

La evaluación cuantitativa del desempeño para eventos TRE se discute en la sección 5.3.1.

Figura 5-11: Diagrama de frecuencias que muestra la cantidad de sismos tipo TRE por duración.



5.3. Validación de resultados

5.3.1. Unidades de sismos vs. unidades de ruido y tremor

El proceso de validación de resultados para clasificación de unidades de ruido vs unidades de sismos es también útil para evaluar el desempeño de las RNA en la detección de TRE ya que

la distinción de este tipo de sismos está asociada al reconocimiento de unidades consecutivas de sismos.

Para el proceso de validación se utilizó 50 elementos de cada matriz de entrenamiento que no se incluyeron en el proceso de entrenamiento. De estos ejemplos 25 corresponden unidades de sismos y 25 a unidades de ruido. Para la validación se utiliza el modo fuera de línea con ayuda de la correspondiente interfaz del programa LAKiY. La manera directa de analizar la ejecución de las RNA es contando la cantidad de sus aciertos. La Tabla 5-6 muestra los resultados de la ejecución de las RNA asociadas con los 25 entrenamientos realizados. El campo *Promedio valor absoluto diferencias* es el promedio de la resta entre los objetivos y los valores calculados.

Tabla 5-6: Resultados de la ejecución fuera de línea de las RNA para unidades de señal y de ruido.

Nombre RNA	Consecutivo	Acierto %	Promedio valor abs. diferencias	Desv. Est.	Mínimo diferencia	Máximo diferencia
RNA01	1	100	8.41E-03	5.65E-03	6.48E-03	4.67E-02
RNA02	2	100	7.59E-03	3.03E-03	4.22E-03	1.62E-02
RNA03	3	100	1.44E-03	1.02E-02	4.22E-12	7.18E-02
RNA04	4	100	1.16E-02	3.27E-03	8.42E-03	2.02E-02
RNA05	5	100	4.38E-03	3.37E-03	1.42E-06	8.68E-03
RNA06	6	100	5.69E-05	6.66E-05	2.13E-10	3.37E-04
RNA07	7	100	3.15E-04	3.19E-04	3.04E-12	8.85E-04
RNA08	8	100	1.54E-02	6.95E-03	8.62E-03	2.71E-02
RNA09	9	100	2.41E-05	6.72E-05	8.43E-08	4.63E-04
RNA09b	10	100	8.62E-03	6.99E-03	4.13E-07	2.94E-02
RNA10	11	100	3.04E-03	1.28E-02	2.50E-10	7.94E-02
RNA11	12	98	1.72E-02	1.12E-01	2.66E-10	7.91E-01
RNA12	13	98	1.24E-02	8.59E-02	1.57E-07	6.08E-01
RNA26	14	100	7.03E-03	2.33E-03	6.99E-04	1.53E-02
RNA27	15	100	2.96E-03	3.32E-03	8.60E-04	1.44E-02
RNA28	16	100	3.37E-03	4.92E-03	1.77E-10	1.80E-02
RNA29	17	100	4.48E-03	1.62E-02	1.74E-04	1.12E-01
RNA30	18	100	3.13E-03	3.74E-03	3.78E-06	2.08E-02
RNA31	19	100	3.55E-03	8.35E-03	5.52E-07	5.86E-02
RNA32	20	100	6.94E-05	2.70E-04	9.81E-10	1.72E-03
RNA33	21	100	7.27E-03	4.99E-03	2.58E-05	1.48E-02
RNA34	22	100	7.80E-03	7.78E-03	3.48E-09	1.67E-02
RNA35	23	100	5.62E-03	1.48E-02	2.24E-08	1.02E-01
RNA36	24	100	7.12E-03	7.50E-03	9.34E-12	1.57E-02
RNA37	25	100	3.35E-04	4.35E-04	1.20E-12	1.11E-03

En general la cantidad de aciertos fue 50 de 50 ejemplos evaluados, excepto para las redes RNA11 y RNA12 en las que se registró un error, en la Figura 5-12 se muestra los resultados en forma gráfica que LAKiY generó para una de las redes ejecutadas fuera de línea (RNA09). En la ejecución de las redes se considera un acierto cuando el valor de salida supera el 0.5 en el caso de señal y está por debajo de este valor en caso de ruido.

En la Tabla 5-6 se han marcado con verde las RNA con menor valor promedio de las diferencias entre los objetivos y los valores calculados, se observa que ninguna de ellas coincide con las de menor error de generalización, marcadas de color naranja. Para analizar esto se graficaron las diferencias de los valores absolutos de los objetivos y los valores

calculados para cada unidad ejecutada por las redes RNA10 y RNA09, las mejores por: error de generalización y aciertos más cercanos a los objetivos respectivamente (Figura 5-13).

Figura 5-12: Resultados de la ejecución de la RNA09 con 50 ejemplos, 25 unidades de ruido y 25 unidades de sismos. Los puntos sobre las líneas representan las diferencias entre el objetivo y el valor calculado por la RNA. En la gráfica estas diferencias se miden alrededor de 0 para las unidades de ruido y alrededor de 1 para las unidades de sismos. En la tabla se pueden observar las diferencias exactas en la última columna para algunas unidades. Para este caso las curvas son líneas casi rectas indicando un buen desempeño de la red.

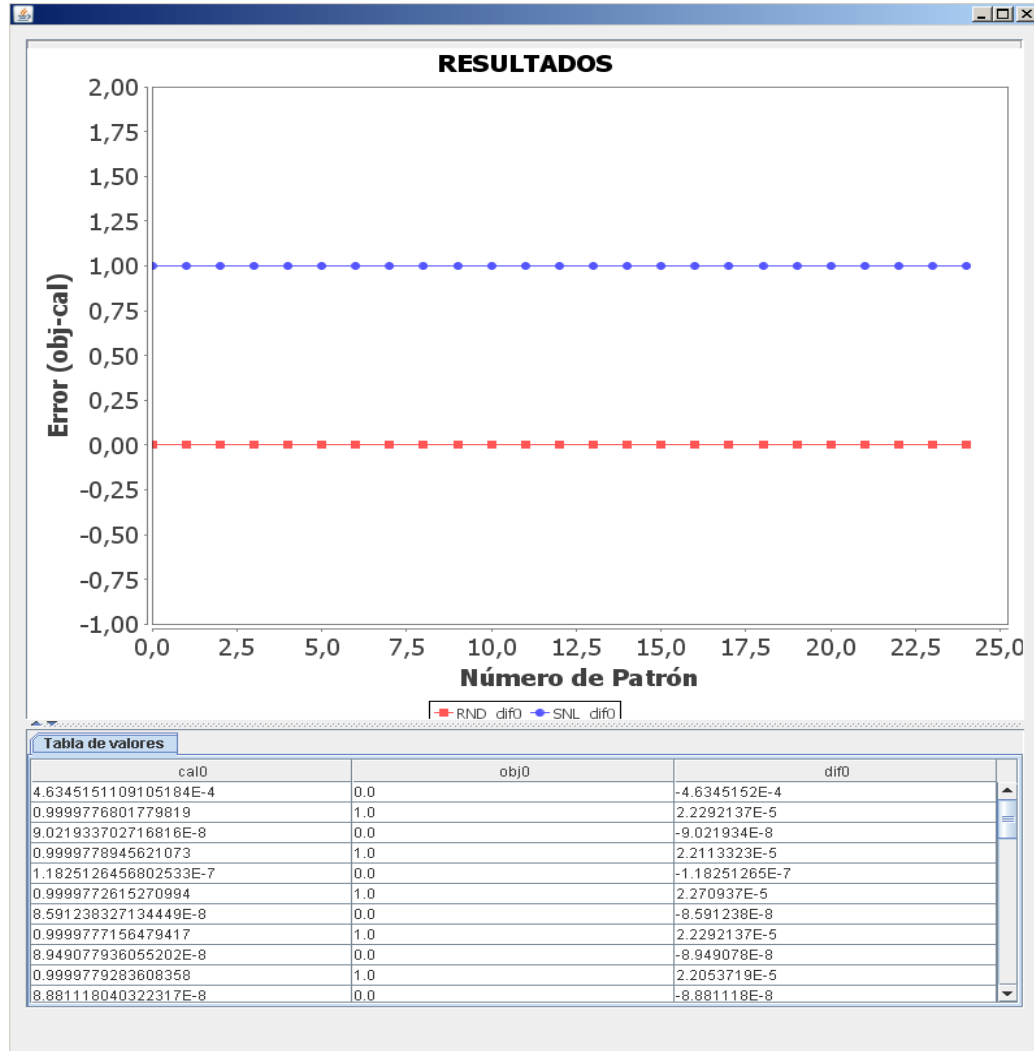
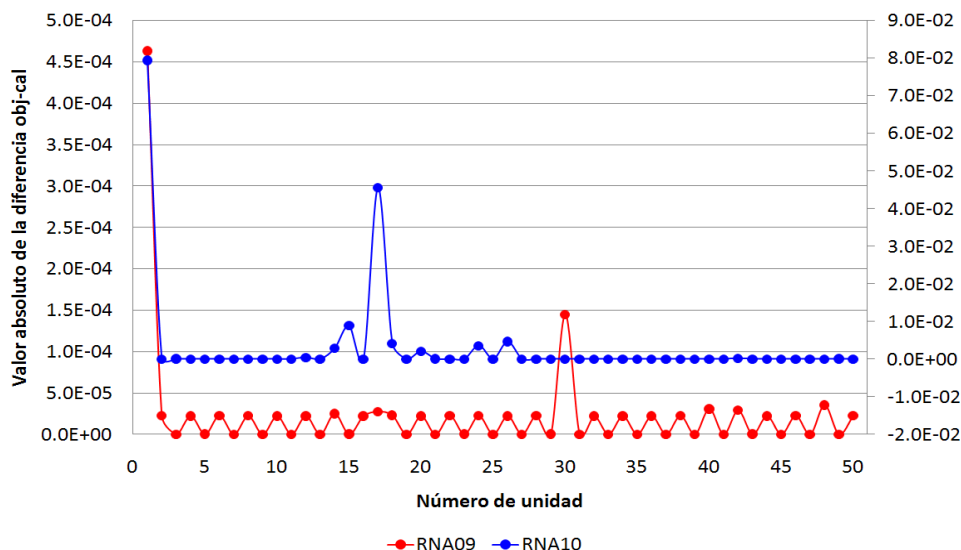


Figura 5-13: Valores absolutos de las diferencias obj-cal para las 50 unidades ejecutadas por las redes RNA09 (eje izquierdo) y RNA10 (eje derecho).



Se observa que las diferencias son bastante pequeñas, sin embargo, son mucho menores para la red RNA09 que fue elegida de acuerdo a la calidad de ejecución.

Para obtener un criterio más para la elección de la mejor RNA se evaluará su comportamiento ejecutando las redes en el modo *en línea* durante un mismo periodo de tiempo. La prueba consiste en ejecutar las redes RNA09, RNA10, RNA32 y RNA34 que corresponden a las mejores ya sea por menor error de generalización o por mejor calidad de ejecución fuera de línea. Se tomaron dos tramos de registro (Figura 5-14), cada tramo contiene un sismo. Entre los dos tramos suman 100 unidades, de las cuales 16 corresponden a sismo, las demás a ruido. Un resumen de los resultados se muestra en la Tabla 5-7.

Figura 5-14: Trazas utilizadas para ejecución en línea de las redes pre-seleccionadas, la marca azul corresponde al segmento de traza que contiene 7 unidades para el sismo 1 registrado el 18 de febrero de 2011 (arriba) y 16 para el sismo 2 registrado el 21 de febrero de 2011 (abajo).

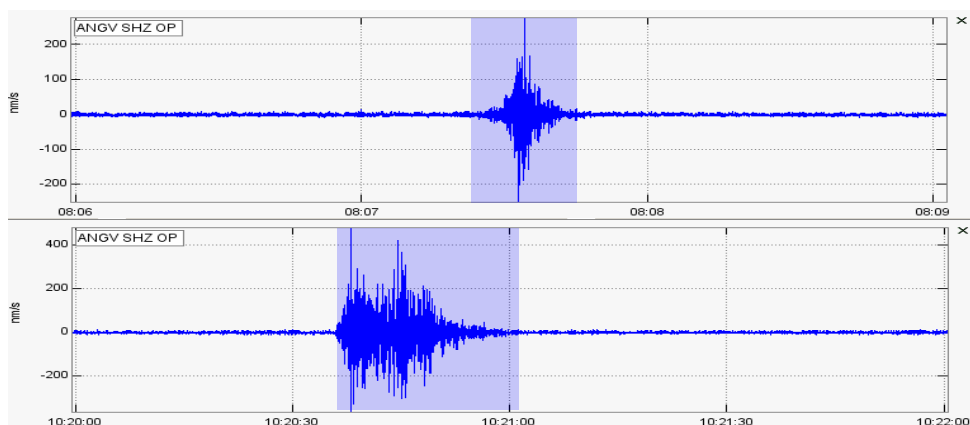


Tabla 5-7: Resumen de resultados de ejecución en línea para las cuatro RNA pre-seleccionadas, la columna verde corresponde a la red con mejor desempeño.

Tipo unidad	RNA09	RNA10	RNA32	RNA34
U. sismos	28	37	11	10
U. ruido	72	63	89	90
U. sismo 1	9	14	8	7
U. sismo 2	7	8	3	3
U. sismo falso	12	22	0	0
U. ruido falso	0	0	5	6

Se observa que la cantidad de unidades detectadas como sismo sobrepasa la correcta para las redes RNA09 y RNA10, mientras que para las redes RNA32 y RNA34 está por debajo del número correcto. Estas últimas redes, entrenadas con un conjunto de ruido más variado (que incluye ruido de 2011), clasificaron algunas unidades del inicio y final de los sismos como ruido. Por otro lado, las redes que fueron entrenadas con el conjunto de ruido más limitado en el tiempo, clasificaron como unidades de sismos algunos segmentos de traza en los que no hay sismo, sin embargo estas unidades clasificadas erróneamente no fueron continuas. La red RNA09 acertó con el número exacto de unidades de sismo.

En este experimento las redes con mejor desempeño fueron las que más aciertos obtuvieron en la ejecución fuera de línea, superando a las elegidas por menor error de generalización.

5.3.2. Clasificación LP vs VT

En este caso también se ejecutaron las RNA con elementos que no se incluyeron en las matrices de entrenamiento, un total de 50. Los resultados de la aplicación fuera de línea se muestran en la Tabla 5-8.

El rango de aciertos va desde 78 a 96% con un promedio del 86%. Se destaca que ninguna de las RNA con acierto mayor que el 90% (líneas color verde en la Tabla 5-8) coincide con una de las mejores elegidas por el criterio de mínimo error de generalización (líneas de color naranja en la Tabla 5-8). La RNA que alcanzó el 96% de acierto tiene una topología 77-3-1, fue entrenada con la matriz LV02c y contiene 20 segundos de traza normalizada, 11 unidades, 7 coeficientes de predicción para cada unidad y normalización del vector de coeficientes en el intervalo $[-1,1]$. En la Figura 5-15 se muestran los resultados de esta ejecución.

Lo novedoso de esta matriz es que sus unidades tienen 2 segundos de duración. Sin embargo las RNA entrenadas con unidades de 2 s y con más de 7 coeficientes de predicción por cada unidad no mostraron aciertos mayores que el 88%. Por otra parte las topologías con 20 y 50 neuronas en la capa oculta tampoco se destacaron en efectividad.

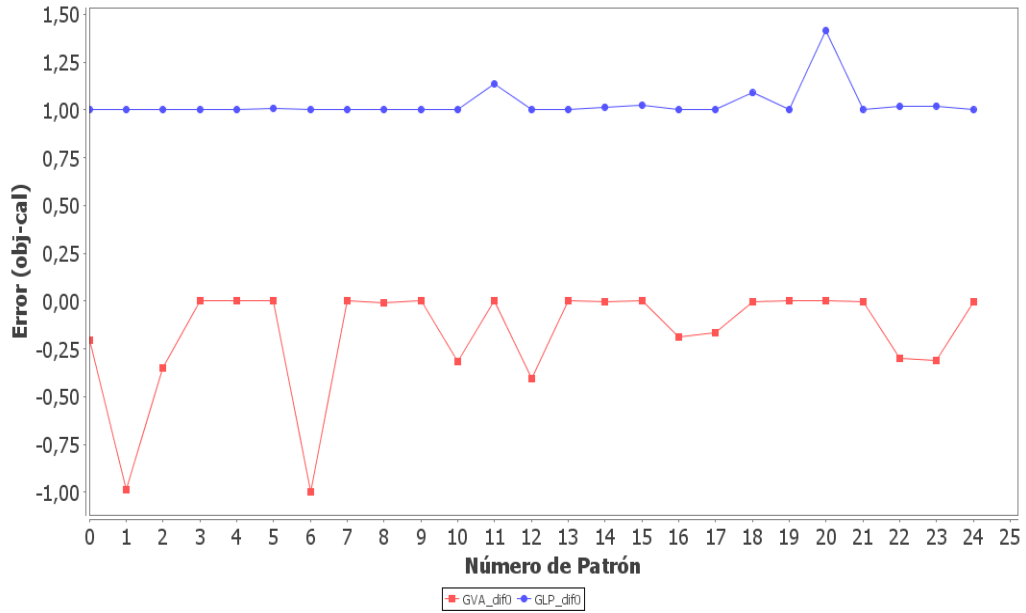
Lo común entre las mejores RNA elegidas por el criterio de mínimo error de generalización (RNA14b) y por el de mayor efectividad (RNA15) es que fueron entrenadas con 7 LPC por unidad, segmentos de traza de 20 segundos y normalización del vector de coeficientes en el intervalo $[-1,1]$. *De esta manera, los segmentos de traza más cortos, topologías pequeñas y 7 LPC arrojaron los mejores resultados.*

Los videos adjuntos en el CD, directorios [Videos/VT](#) y [Videos/LP](#) muestran el desempeño de estas redes cuando se ejecutan en el modo *en línea*.

Tabla 5-8: Resultados de la ejecución fuera de línea de las RNA entrenadas para la clasificación entre los tipos LP y VT.

Nombre RNA	Consecutivo	Aciertos %	Promedio	Desv.Est.	Mínimo	Máximo
RNALV01B	1	82	0.2078	0.3013	0.0039	0.9667
RNALV02B	2	88	0.1483	0.2331	0.0009	0.8279
RNALV03	3	86	0.1032	0.1738	0.0000	0.7937
RNALV04	4	74	0.2586	0.3271	0.0005	0.9292
RNALV05	5	78	0.0846	0.1482	0.0018	0.8587
RNALV06	6	88	0.1471	0.2303	0.0007	0.8404
RNALV07	7	82	0.1284	0.2103	0.0003	0.8747
RNALV08	8	82	0.1949	0.3019	0.0001	0.9073
RNALV09	9	82	0.1371	0.2463	0.0034	0.8755
RNALV02Ba	10	78	0.1325	0.2162	0.0004	0.7395
RNALV02Bb	11	92	0.2417	0.2819	0.0036	0.8762
RNALV02Bc	12	86	0.1747	0.2449	0.0015	0.8706
RNALV02Bd	13	88	0.0544	0.1194	0.0049	0.7677
RNALV06a	14	92	0.0756	0.1542	0.0002	0.8448
RNALV06b	15	86	0.1119	0.1915	0.0028	0.8294
RNALV06c	16	88	0.1093	0.1465	0.0012	0.5557
RNALV06d	17	84	0.0751	0.1560	0.0034	0.8835
RNALV10	18	84	0.1134	0.2118	0.0028	0.9010
RNALV10a	19	78	0.1365	0.2069	0.0001	0.8857
RNALV10b	20	84	0.1375	0.2104	0.0001	0.8965
RNALV11	21	92	0.0805	0.1718	0.0000	0.8956
RNALV11a	22	88	0.0676	0.1189	0.0000	0.6337
RNALV12	23	80	0.1046	0.1614	0.0000	0.6009
RNALV13	24	92	0.1418	0.2525	0.0001	0.9693
RNALV14	25	92	0.0455	0.0933	0.0000	0.3976
RNALV14b	26	88	0.0853	0.1566	0.0000	0.8463
RNALV15	27	96	0.1656	0.2768	0.0000	0.9960
RNALV15a	28	90	0.1351	0.2685	0.0000	0.9883
RNALV16	29	86	0.1369	0.3185	0.0000	0.9885
RNALV16a	30	88	0.1139	0.2922	0.0026	0.9734
RNALV17	31	86	0.1868	0.3589	0.0006	0.9965

Figura 5-15: Resultados de la ejecución de la RNA15 con 50 ejemplos, 25 LP y 25 VT. Los puntos sobre las líneas representan las diferencias entre el objetivo y el valor calculado por la RNA. En la gráfica estas diferencias se miden alrededor de 0 para el tipo VT y alrededor de 1 para los sismos LP. Para este caso los puntos rojos número 1 y 6 fueron los desaciertos de la red. Existen valores cuyas diferencias no sobrepasan ± 0.5 , por tanto no se consideran errores.



6. Discusión

Las aplicaciones que utilizan redes neuronales artificiales, especialmente en las que se aplica el entrenamiento supervisado, están sujetas a aprender únicamente lo que el supervisor (el usuario) les enseña. De aquí una de las debilidades del método es que se debe conseguir la correcta elección de los ejemplos de entrenamiento, validación y prueba. La subjetividad en la elección de las señales es uno de los factores más influyente sobre los resultados de las RNA. Lo ideal es contar con un grupo de temáticos quienes se encarguen de estudiar los ejemplos utilizados en la fase de aprendizaje y desechar para esta fase aquellos ejemplos en los que no hubo consenso. Incluso después de la fase de aprendizaje se podría utilizar los eventos en los que el grupo no estuvo de acuerdo para evaluar la generalización de la red.

Los resultados obtenidos para la ejecución en línea no fueron mejores que los de validación fuera de línea, esto se debe a que para el segundo caso se separó un número de ejemplos de prueba que hacían parte del mismo periodo del que se conformó la matriz de entrenamiento, en cambio, para la ejecución en línea se utilizó una traza de un periodo diferente, en el que las características de ruido pudieron ser diferentes, pese a esto las RNA muestran una buena capacidad de generalización.

El método aplicado en este estudio puede ser mejorado utilizando más de una técnica de análisis espectral, inclusive el mismo dominio del tiempo, por ejemplo, parametrizando las envolventes de las formas de onda. En la medida en que se tenga más características que diferencien un conjunto de otro, la discriminación se realizará con mayor precisión.

En la fase de detección o clasificación sismo-ruido, el método utilizado funciona bastante bien, logrando cerca del 100 % de acierto en las pruebas realizadas, sin importar cuales fueran los parámetros utilizados, los cuales se ajustaron en varias configuraciones.

La ejecución en línea de esta fase mostró resultados interesantes, en varias ocasiones la detección de un sismo iniciaba segundos antes de que se observe claramente el arribo de las primeras ondas del mismo. Con esto no sugerimos que el sistema predice sismos, sino que el método encuentra componentes espectrales correspondientes a unidades de sismo, en lugares de la traza donde no es posible su reconocimiento a simple vista. Este resultado forma parte de la capacidad de generalización que tienen las RNA.

Con base en estos resultados, en caso en que se desee re-entrenar redes especializadas en este tipo de clasificación, ya no será necesario buscar minuciosamente una configuración de red óptima, las pruebas realizadas en este trabajo mostraron que con una topología de red sencilla se logran excelentes resultados.

Lo anterior también ofrece una gran confiabilidad en cuanto a la detección de sismos de fuente sostenida en el tiempo, como son los episodios de tremor volcánico. Para este tipo de eventos no es necesario ejecutar un clasificador especializado ya que su principal característica es la duración del sismo. El número de segmentos continuos de señal clasificados como sismo daría un muy buen estimativo de la duración del tremor.

Una posible extensión del sistema, sería el cálculo de la magnitud de duración para el caso de los sismos tipo VT ya que el clasificador sismo-ruido nos sugeriría de una manera confiable la duración del sismo.

La clasificación entre los tipos LP y VT tiene una gama más amplia en lo que a resultados se refiere, el cambio en los parámetros evaluados afecta el nivel de acierto en la clasificación. Contar con el mayor número de ejemplos en la fase de entrenamiento, construir una arquitectura de red con una capa oculta de 3 neuronas e inicializar los pesos varias veces, nos da la posibilidad de encontrar mejores resultados. El porcentaje de acierto obtenido en este estudio, que en promedio está en 87% y que para algunas redes alcanza el 92% y hasta el 96%, es lo suficientemente confiable como para utilizar los resultados en la interpretación parcial de la actividad volcánica.

Los resultados obtenidos indican que un sistema de este tipo puede ser una herramienta de ayuda en el procesamiento primario de sismicidad volcánica, sin embargo, es necesario dejar claros algunos aspectos con la finalidad de no generar falsas expectativas acerca de la ejecución del sistema:

- El sistema en ningún momento puede ser ejecutado sin la supervisión de los resultados por parte de un temático. Si bien es cierto que la detección y clasificación en línea agilizan el proceso de interpretación preliminar del fenómeno con cierto nivel de confiabilidad, no se obtendrá un 100% de aciertos en la clasificación, por ende se hace necesaria una supervisión rutinaria de la herramienta.
- En cuanto a la clasificación sismo-ruido es importante re-entrenar las redes utilizadas con cierta periodicidad, escogiendo matrices de entrenamiento con información lo más cercana al presente y de especial manera cuando las condiciones de ruido de fondo cambien notoriamente, por ejemplo, en épocas de lluvias o vientos fuertes. Esto aseguraría la correcta detección.
- Existen tipos de señales, no volcánicas, como sismos regionales, avalanchas, tormentas eléctricas, picos de voltaje y otros, que el sistema no es capaz de distinguir. Sobre los segmentos de este tipo de señales se les asignará sus respectivos vectores LPC y el sistema intentará discernir si corresponden a señal o ruido y probablemente estas señales serán asociadas, **erróneamente**, con un alguno de los tipos de sismicidad volcánica. Debe ser claro que en el presente trabajo no se incluyó este tipo de señales, se deja su tratamiento para un trabajo posterior. Esta es otra razón para realizar una continua supervisión en la ejecución de la herramienta.

7. Conclusiones y recomendaciones

7.1. Conclusiones

La metodología aplicada para detección y clasificación de sismos volcánicos proporciona un grado de confianza bastante aceptable dentro de las condiciones necesarias para su aplicación. La detección de segmentos de traza que contienen características de señales sísmicas son reconocidas con facilidad, el nivel de aciertos fue cercano al 100% para varias de las RNA seleccionadas.

Una vez probado el proceso de distinción entre unidades de sismos y unidades de ruido, la detección de episodios de tremor fue posible, el aprovechamiento de la característica de duración de este tipo de eventos facilitó su reconocimiento.

La discriminación entre eventos de los tipos LP y VT, en promedio, para el total de redes entrenadas alcanzó el 87%, sin embargo, las mejores redes de este grupo superaron el 92%, alcanzando el 96% en el mejor de los casos. La diversidad en los resultados para este tipo de redes fue consecuencia del número de pruebas realizadas con diferentes parámetros durante los entrenamientos.

La representación de los registros con 7 coeficientes de predicción lineal por cada unidad y las topologías pequeñas con 3 neuronas en la capa oculta mostraron los mejores resultados.

A un nivel operativo, el programa LAKiY mostró un desempeño adecuado, agilizando los procesos de conformación de matrices de entrenamiento, prueba y validación, haciendo versátil el proceso de entrenamiento por cuanto la mayor parte de parámetros ajustables pueden cambiarse desde su interfaz.

En cuanto a la ejecución de las redes entrenadas, la posibilidad de aplicar el método en línea y fuera de línea permite estudiar el comportamiento de las redes antes de incluirlas como parte del proceso primario.

La posibilidad de conectarse a servidores de ondas para procesar información de periodos pasados es útil para reevaluar la metodología a medida que en esta se incluyan nuevas técnicas de representación y clasificación.

La base de datos implementada facilitó la organización de los datos de entrada y salida en las diferentes fases del proceso de detección y clasificación.

7.2. Recomendaciones

Con miras a futuros trabajos relacionados con la metodología aplicada, se sugiere explorar otras técnicas de representación de las señales; si bien es cierto, la representación a través de

LPC arroja buenos resultados, estos podrían ser mejorados acudiendo a la combinación de técnicas que puedan extraer características en diferentes dominios, inclusive en el dominio del tiempo.

Existen variantes del algoritmo de retropropagación del error para la fase de aprendizaje, es posible que una de esas variantes mejore los resultados. Aprovechando la modularidad del sistema construido para este trabajo, se recomienda codificar variantes del algoritmo de retropropagación, realizar las respectivas pruebas y comparar.

Se sugiere estudiar la detección y clasificación para cada una de las estaciones de la red sísmica, el programa está diseñado para ser aplicado en varias estaciones al tiempo, para esto es necesario extraer los respectivos conjuntos de entrenamiento, validación y prueba.

Se recomienda estudiar el problema de la clasificación de señales no volcánicas con sismos volcánicos y del entrenamiento de redes especializadas en discernir señales que usualmente causan controversia en la clasificación, por ejemplo, en épocas en que los vientos se incrementan el nivel base de la señal aumenta visiblemente, debido a esto puede generarse confusión con el apareamiento de una señal de tremor volcánico poco energético, pero importante para la interpretación del fenómeno. Si el grupo corrobora que la procedencia de las señales es no volcánica (en este caso generada por viento), sería posible extraer los conjuntos de entrenamiento, validación y prueba para construir redes que distingan estos dos tipos de señales.

A. Anexo: Guía para el uso de LAKiY versión 1.0

Instalación

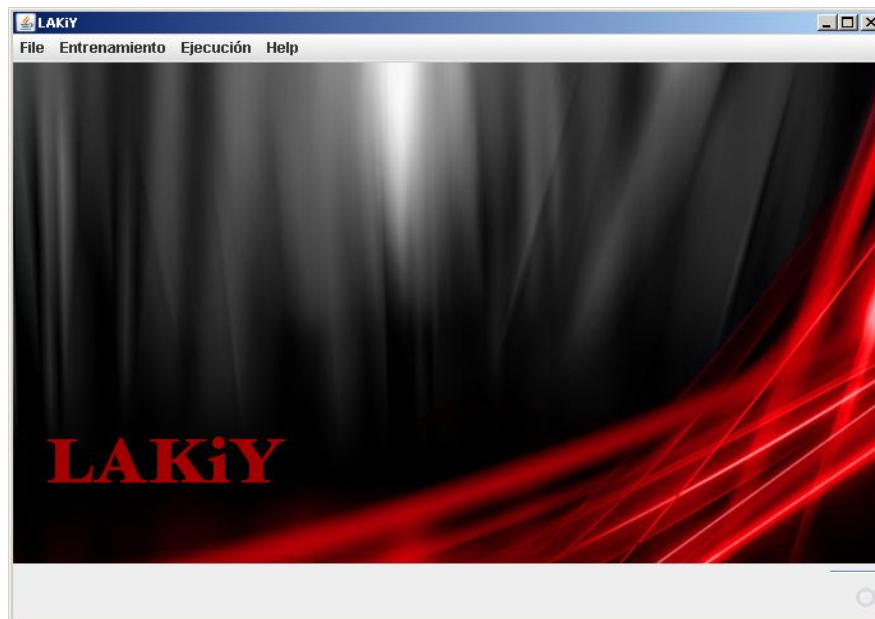
Para iniciar el proceso de instalación de LAKiY 1.0 es necesario contar con el gestor de base de datos PostgreSQL versión 8.1.4 o posterior y la JRE de Java en su versión 6.0 o posterior. La ejecución del programa LAKiY inicia cuando se ejecuta el archivo *ReconocimientoSismicidadVolcanica.Jar*, ubicado en el directorio *ReconocimientoSismicidadVolcanica/ bin*. Junto a este directorio existe una estructura de cuatro carpetas tituladas: *Senales*, *Matrices Entrenamiento*, *Entrenamientos* y *Resultados Ejecucion*. De la primera carpeta el programa toma los registros para entrenamiento y en las demás el programa organiza la información de los resultados. Es necesario editar el archivo *scr/configuración/bases.xml* digitando una IP de uso local y la clave de la base.

Se debe aclarar que para la ejecución en línea ya sea en tiempo real o tiempo pasado es necesario que el sistema de adquisición sea el Earthworm versión 6.4 o posterior. Por otro lado, los archivos de los cuales se extrae los ejemplos para entrenamiento y prueba deben estar en formato *suds* y para su preparación es necesario contar con el programa PSW (Raigoza, 2008).

El menú principal

En la Figura A-1 se muestra el panel inicial con su respectivo menú principal. En este podemos observar las pestañas de *Entrenamiento* y *Ejecución*.

Figura A-1: Panel inicial y menú principal del programa LAKiY.



Dentro de la pestaña *Entrenamiento* (Figura A-2) encontramos las opciones: *Extraer Señales*, que abre el programa PSW donde se selecciona y extrae los registros en el formato necesario; *Generar Matriz Entrenamiento*, donde se crean las matrices de entrenamiento validación y prueba y *Entrenar Red Neuronal*, donde se configuran los parámetros de entrenamiento de las RNA.

Figura A-2: Contenido de la pestaña Entrenamiento del menú principal.



En la pestaña *Ejecución* (Figura A-3) del menú principal se tiene las opciones *En línea* y *Fuera de línea*. De la primera de estas se despliegan dos opciones más, *Tiempo Real* y *Tiempo Pasado*.

Figura A-3: Contenido de la pestaña Ejecución del menú principal.



Generar matrices de entrenamiento

La opción *Generar Matriz Entrenamiento* despliega una ventana que permite la parametrización de una matriz de entrenamiento (Figura A-4). La ventana tiene tres paneles, el de configuración de la matriz de entrenamiento, el de objetivos y el de ubicación de los archivos de entrada y salida. Dentro del panel de configuración se encuentra una pestaña que despliega el nombre de las matrices ya generadas, si se selecciona una de estas los campos en todos los paneles se actualizan con los valores con que se configuro la respectiva matriz. Cuando se crea una nueva matriz es necesario llenar manualmente todos los campos con excepción del campo *Número de unidades* que se llena automáticamente dependiendo de los valores incluidos en los campos *Duración de unidad*, *Traslapo* y *Duración traza*. La opción *Contar n segundos antes máxima amplitud* permite cortar las trazas n segundos antes de encontrar su máxima amplitud, esto sirve especialmente para recortar registros de sismos tipo LP o VT cuando se desee entrenar sismos sin iniciar exactamente con el arribo de onda P. Si esta opción no se utiliza basta con dejarla en blanco.

Figura A-4: Ventana de configuración de parámetros para generar matrices de entrenamiento.

Generar matrices de entrenamiento

Configuración patrones entrenamiento

Matrices generadas: LV09

Nombre matriz: LV09

Estación: ANGV

Tasa de muestreo: 100 Hz

Tasa de remuestreo: 100 Hz

Duración de unidad: 3 s

Número de unidades: 11

Duración traza: 30 s

Traslapo: 15 %

Normalizar traza:

Orden LPC: 9

Normalizar LPC: [0,1]

Cortar n segundos antes máxima amplitud:

Objetivos

Extensión	Objetivo
<input type="text"/>	<input type="text"/>
GVA	0

Adicionar Eliminar

LAKY

Ubicación registros: C:/ReconocimientoSismicidad/Senales/LV Examinar

Ubicación matriz: C:/ReconocimientoSismicidad/MatricesEntrenamiento Examinar

Aceptar Modificar Eliminar Cancelar

Para incluir los objetivos es necesario especificar la extensión del archivo que contiene la señal y su respectivo valor objetivo, seguido a esto se hace clic en el botón *Adicionar*. Este procedimiento debe repetirse para cada tipo de señal. En las pestañas desplegadas ubicadas dentro del panel de objetivos se almacenan las extensiones y los objetivos asignados a cada tipo. El botón *Eliminar* borra la extensión y el respectivo objetivo seleccionado. En el panel inferior se deben seleccionar los directorios donde se encuentran las señales, es necesario que en el mismo directorio se encuentre todos los archivos de los diferentes tipos. En este panel también se selecciona el directorio donde se guardará la matriz generada. Los botones *Aceptar*, *Modificar*, *Eliminar* y *Cancelar*, que se encuentran por fuera de los paneles, se encargan respectivamente de: iniciar la generación de una nueva matriz de entrenamiento, modificar la configuración de una matriz existente, eliminar una matriz existente y cancelar la operación.

Entrenar RNA

La ventana de entrenamiento (Figura A-5) contiene cinco paneles en cuyos campos se consigna la información para configurar un entrenamiento. En el panel *General* se tiene una pestaña que despliega el nombre de los entrenamientos realizados, cuando se escoge uno de estos, los campos de los demás paneles se actualizan automáticamente, mostrando los valores con que se configuró dicho entrenamiento. En este panel también se introduce el nombre de un nuevo entrenamiento y en la pestaña inferior del panel se hace la elección de la matriz con que se desea entrenar. El siguiente panel (inferior) muestra la topología de la RNA que se entrenará, el único campo a llenar es *Capa oculta* que hace referencia al número de neuronas en la capa oculta ya que los campos *Capa de entrada* y *Capa de salida* obtienen sus valores de la configuración de la matriz de entrenamiento. El siguiente panel inferior contiene los parámetros de aprendizaje: *Modo*, *Tasa* y *Momentum*, el campo de *Modo* puede tener los valores 0, 1 o 2 dependiendo del tipo de aprendizaje que se desee utilizar que puede ser *on-line*, *batch* o *gradient*, respectivamente. En el panel de *Patrones de entrenamiento* se selecciona el archivo que contiene la matriz de entrenamiento, el número de patrones de entrenamiento, el número de iteraciones y la ubicación de los archivos de salida que consisten en un archivo con los errores global y de generalización y los clones de red generados.

Figura A-5: Ventana de configuración de los parámetros de entrenamiento.

The screenshot shows the 'Entrenar RNA' window with the following configuration details:

Panel	Field	Value
General	Entrenamientos realizados	ELV15a
	Nombre entrenamiento	ELV15a
	Matriz de entrenamiento	LV02c
Arquitectura - Perceptrón Multicapa	Capa de entrada	77
	Capa oculta	3
	Capa de salida	1
Parámetros de aprendizaje	Modo	1
	Tasa	0.5
	Momentum	0.5
Patrones de entrenamiento	Ubicación patrones de entrenamiento	ismicidad/MatricesEntrenamiento/LV02c.mtz
	Número de patrones de entrenamiento	680
	Número de iteraciones	1000
	Ubicación archivos de salida	ReconocimientoSismicidad/Entrenamientos
Patrones de validación	Ubicación patrones de validación	oSismicidad/MatricesEntrenamiento/LV02c.mtz
	Línea de inicio en archivo de patrones	681
	Número de patrones de validación	170
	Número de iteraciones por validación	10

Buttons: Modificar, Eliminar, Entrenar

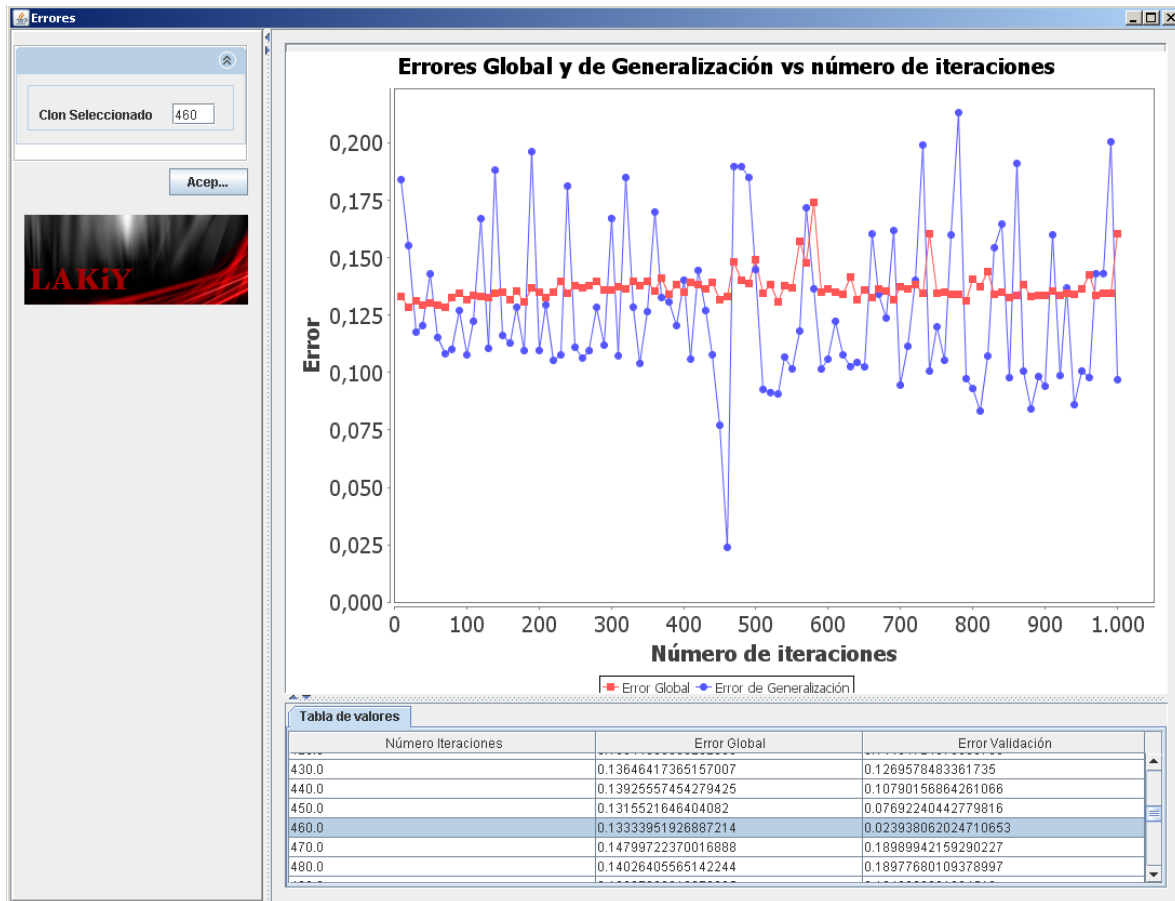
Logo: LAKiY

Finalmente en el panel *Patrones de validación* se ubica el archivo que contiene la matriz de validación que inclusive puede estar en el mismo archivo que contiene los ejemplos de entrenamiento, de ser así se especifica la línea de inicio en que inician los patrones de validación dentro de la matriz. Se consigna también el número de patrones de validación y el

número de iteraciones que deben transcurrir para generar periódicamente los clones de red. Los botones *Modificar* y *Eliminar* modifican la configuración de un entrenamiento existente y lo eliminan respectivamente; el botón *Entrenar* inicia el proceso de entrenamiento.

Cuando el entrenamiento ha terminado se despliega información gráfica y tabular de los resultados del entrenamiento (Figura A-6). En el panel izquierdo se introduce el número de iteración que generó el clon de red que se desea seleccionar. En la tabla del panel inferior se puede detallar cuál es el número de iteración al que corresponde un clon con determinado valor en sus errores. El botón *Aceptar* almacena el clon seleccionado.

Figura A-6: Información de salida de entrenamiento.



Ejecución

Existen tres modos de ejecución: fuera de línea, en línea tiempo real y en línea tiempo pasado. Para tener acceso a cualquiera de estos modos se acude a la pestaña *Ejecución* en el menú principal.

Modo Fuera de línea

Este tipo de ejecución requiere un archivo con los patrones que se le presentarán a la red, estos inclusive pueden hacer parte de la matriz de entrenamiento o validación, también pueden o no contener los objetivos en caso de que se conozca el tipo al que pertenece cada

patrón. Dentro del panel de la ventana *Ejecución fuera de línea* (Figura A-7) se tiene una pestaña con el listado de los clones seleccionados en cada entrenamiento, el nombre de estos clones es igual al nombre del entrenamiento ya que por cada entrenamiento se elige un solo clon. Se debe ubicar el archivo que contiene los patrones que se presentarán a la red, en el siguiente campo se escribe la cantidad de patrones a evaluar y en el siguiente se escribe la línea de inicio de los patrones en caso de que se encuentren dentro de una matriz de entrenamiento o validación. Se elige una ubicación de salida y se responde si los patrones tienen asociado objetivos o no.

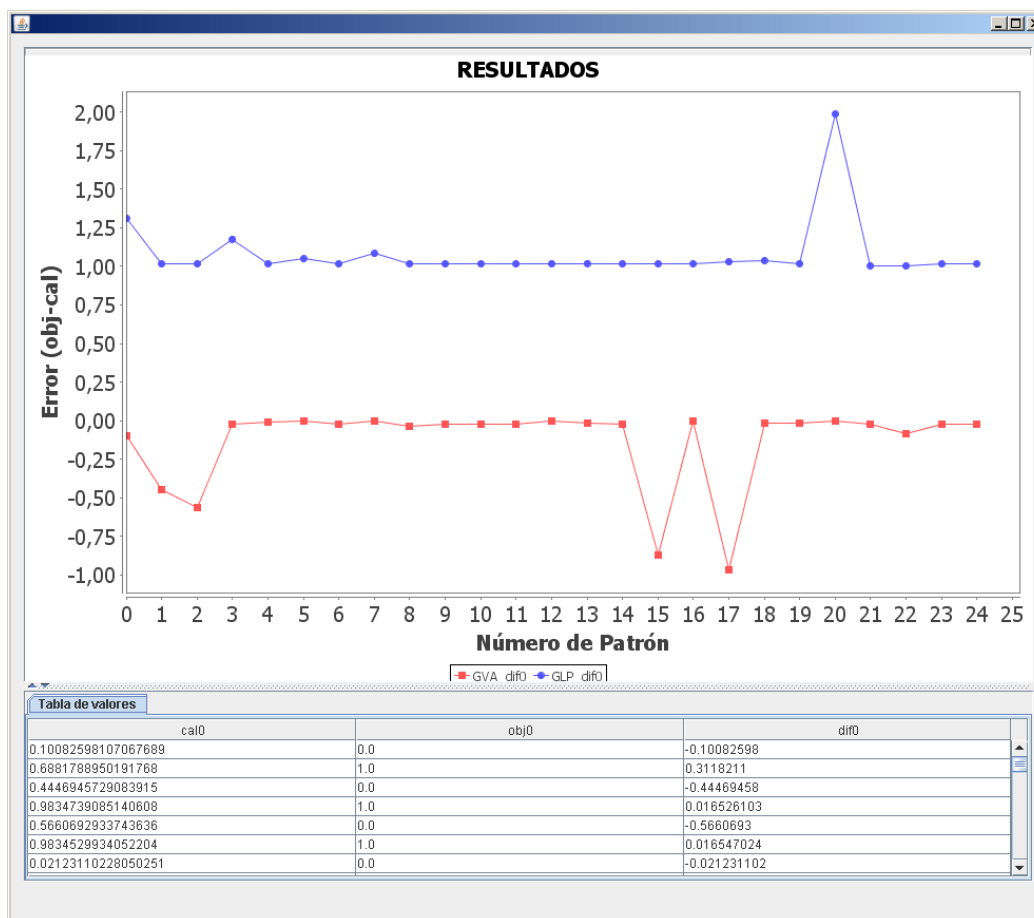
Figura A-7: Ventana de ejecución en el modo fuera de línea.

Seleccione los siguientes parámetros

Identificador de red	ELV10a
Ubicación archivo patrones	Examinar
Número de patrones	50
Línea de inicio en archivo de patrones	720
Ubicación archivo salida	Examinar
¿ El archivo contiene los objetivos ?	<input checked="" type="checkbox"/>

Aceptar

Una vez que se acepta la ejecución y esta se lleva a cabo, se despliega una ventana con información gráfica y tabular de la operación (Figura A-8). En la tabla se presentan columnas con los valores objetivos, valor calculado y diferencia entre el objetivo y el calculado para cada tipo de conjunto. En la figura se grafican las diferencias entre el objetivo y el calculado para cada conjunto. Las líneas base de estas gráficas están separadas una unidad para hacer más fácil su visualización.

Figura A-8: Información de salida del modo ejecución fuera de línea.


En línea - tiempo real y tiempo pasado

Este modo puede operar para detección o para detección y clasificación. La ventana de configuración contiene tres paneles (Figura A-9). En el primero (de arriba a abajo) se tiene la información general de cada set de clasificación, en la primera pestaña se despliega el nombre de los sets existentes, al elegir uno de estos la información en los campos se actualiza automáticamente mostrando los valores con que el set fue configurado. En este panel se debe llenar también el nombre de un nuevo set, la ruta de la información de salida y responder si se desea clasificar o no. Una vez se ha consignado esta información se presiona el botón *Agregar*, si se desea eliminar o modificar un set se utilizan los respectivos botones de este panel. En el panel *Configuración para detección* se ingresan la dirección IP del servidor de ondas (Waveserver o Winston), el puerto de donde se extrae los datos, la estación sismológica que se va a utilizar y la RNA que se ejecutará, esta última está ligada a un entrenamiento y lleva su mismo nombre. En la tabla de la derecha se muestran los valores con que se configuró la detección para cada estación sismológica.

Figura A-9: Ventana de configuración del modo en línea tiempo real.

Detección y clasificación en línea

Conjuntos

Conjuntos de clasificación existentes: ANGV-03-www

Nombre nuevo conjunto de clasificación: ANGV-03-www

Ubicación archivo resultados: ntoSismicidad/ResultadosEjecucion Examinar

Clasificar por tipo de sismo? Modificar Agregar Eliminar

Configuración para detección

IP servidor de ondas: 192.168.9.233

Puerto: 16099

verificar conexión

Estación: ANGV

Escoja una RNA: SRS32

IP	PUER...	ESTAC...	IDENTIFICA...
192.16...	16099	ANGV	113

Agregar Eliminar

Configuración para clasificación

IP servidor de ondas: 192.168.9.233

Puerto: 16099

verificar conexión

Estación: ANGV

Escoja una RNA: ELV15

IP	PUER...	ESTAC...	IDENTIFICA...
192.16...	16099	ANGV	152

Agregar Eliminar

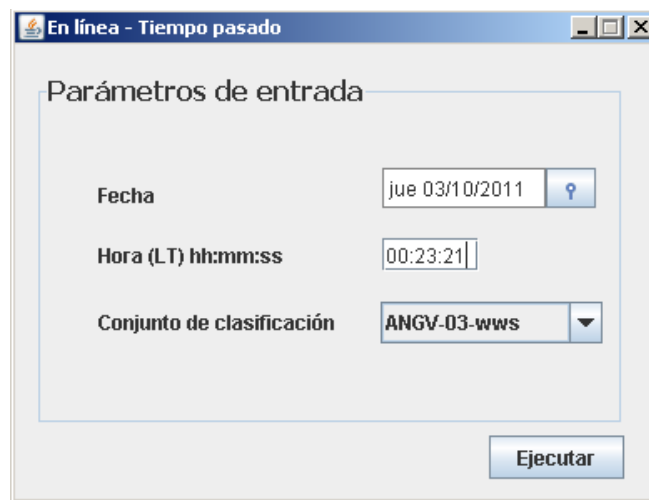
Ejecutar

En el panel de configuración para clasificación se ingresan los valores en campos iguales que en el panel anterior pero con información para clasificación. Se puede hacer modificaciones a un set pre-existente en la configuración para detección o clasificación individualmente. El

botón *Verificar conexión* indica si se estableció o no conexión con el servidor de ondas. Una vez configurado un set el botón *Ejecutar* inicia la operación.

Los sets creados en esta ventana sirven como insumo de entrada al modo de ejecución en *línea tiempo pasado*, su ventana correspondiente (Figura A-10) también pide información de la fecha y hora local en que se desea ejecutar la RNA. La operación emula el tiempo real empezando desde la fecha y hora con que se configuró la ejecución y utilizando para esto el set escogido en la pestaña inferior que despliega el nombre de los sets existentes. Si se desea utilizar este modo con un set diferente a los existentes es necesario crearlo en el modo *en línea tiempo real* para luego utilizarlo en el modo *en línea tiempo pasado*.

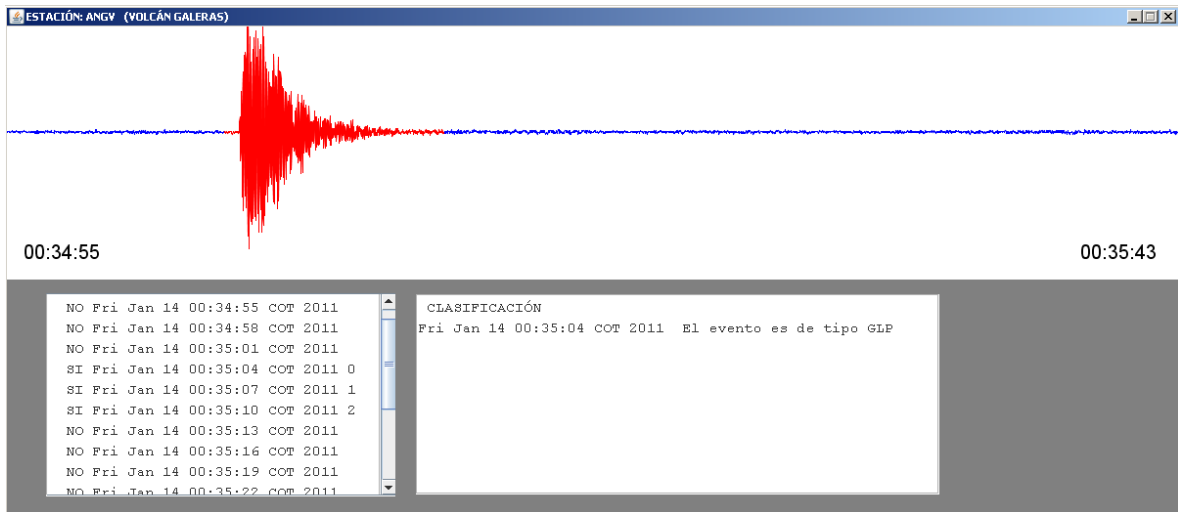
Figura A-10: Ventana del modo en línea tiempo pasado.



The screenshot shows a software window titled "En línea - Tiempo pasado". Inside the window, there is a section labeled "Parámetros de entrada" (Input Parameters). This section contains three input fields: "Fecha" (Date) with the value "jue 03/10/2011" and a calendar icon; "Hora (LT) hh:mm:ss" (Local Time) with the value "00:23:21"; and "Conjunto de clasificación" (Classification Set) with a dropdown menu showing "ANGV-03-www". At the bottom right of the window, there is a button labeled "Ejecutar" (Execute).

Para ambos casos, *en línea tiempo real* y *en línea tiempo pasado*, la ejecución de un set despliega información gráfica y textual de la operación (Figura A-11). Se muestra una gráfica de la traza a medida que el servidor de ondas responde las peticiones del programa. La traza con color azul indica que existe ruido, cuando se detecta la presencia de unidades de sismo la traza se pinta de color rojo. En los paneles de texto se consigna información de la fecha y hora de la detección y de la clasificación. Para la detección se muestra *NO* o *SI* dependiendo de si en ese lapso de tiempo se detectó o no una unidad de sismo, también muestra un contador de unidades de sismo continuas. En el caso de clasificación aparece una frase que menciona el tipo de sismo que presentó así como información de su fecha y hora. En el directorio de salida, seleccionado durante la configuración del set, se genera un archivo de texto llamado *resultadosEnLinea.acs* que contiene información de la detección y clasificación.

Figura A-11: Información de salida del modo ejecución en línea.



B.Anexo: Estructura de la base de datos

Con los siguientes diagramas no se pretende estructurar un diagrama entidad-relación. Sin embargo, se utilizan algunas de sus convenciones para que la estructura de la base de datos sea entendible. El nombre de la base de datos es *clasiauto* y consta de 7 tablas.

Convenciones:

Rectángulo. Representa una tabla.

Ovalo. Representa un atributo de la tabla.

Subrayado rojo. Representa el atributo que contiene la llave primaria de la tabla.

Figura B-1: Diagrama de la tabla Proyecto.

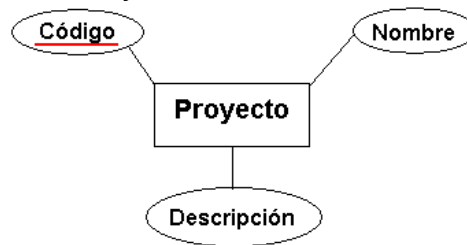


Figura B-2: Diagrama de la tabla Matriz entrenamiento.

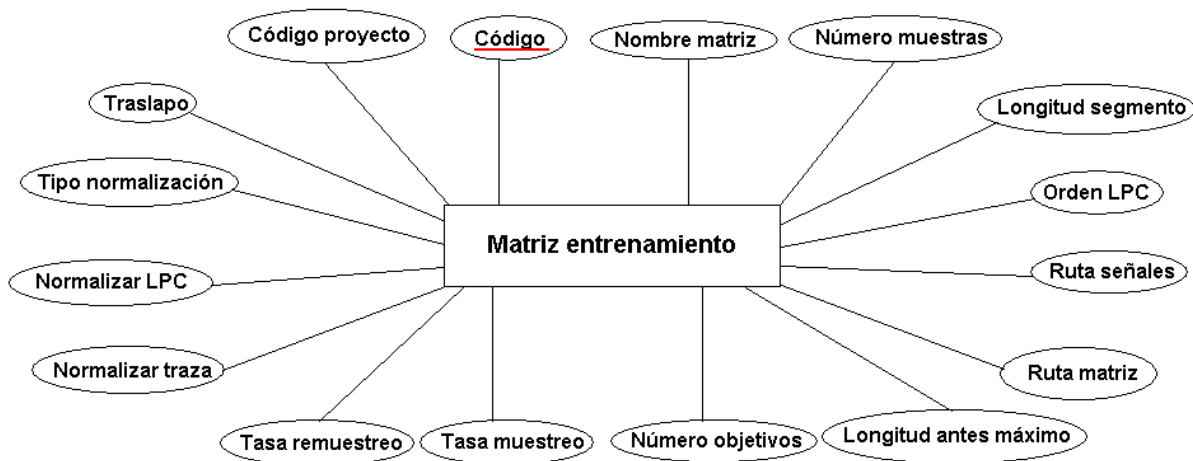


Figura B-3: Diagrama de la tabla Objetivos matriz.

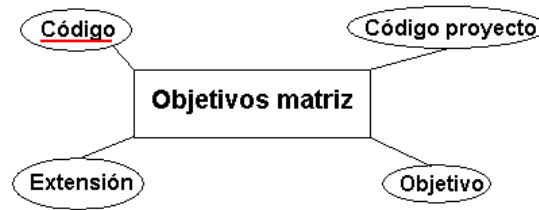


Figura B-4: Diagrama de la tabla Entrenamiento red.

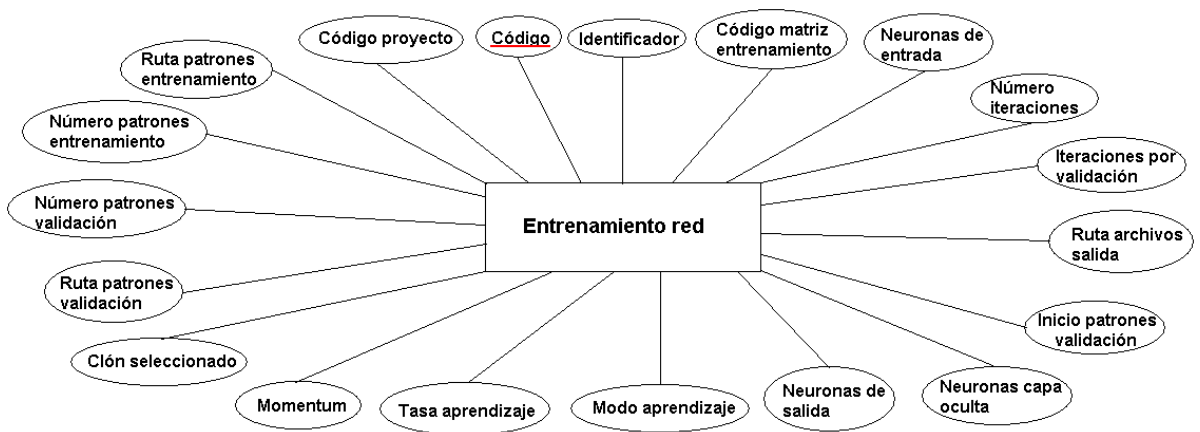


Figura B-5: Diagrama de la tabla Configuración estación.

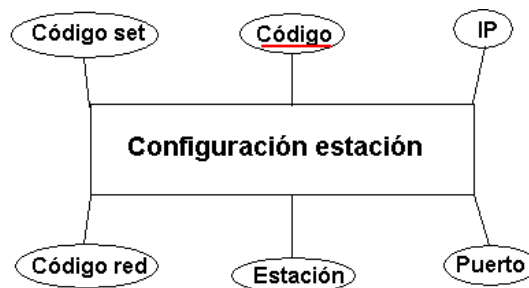
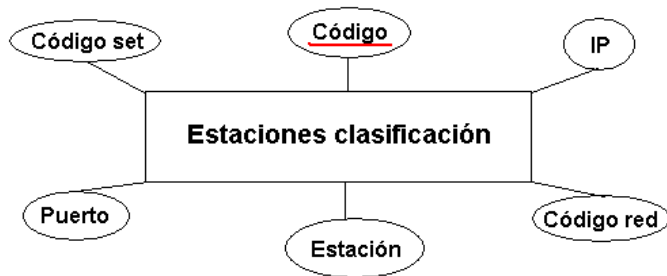
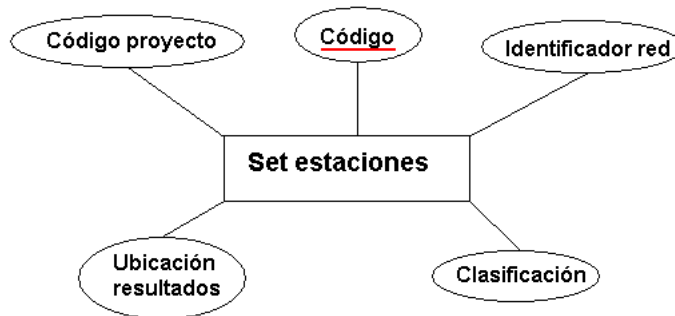


Figura B-6: Diagrama de la tabla Estaciones clasificación.**Figura B-7:** Diagrama de la tabla Set estaciones.

Archivos adjuntos

1. Carpeta con videos de clasificación LP y VT en el modo en línea.
2. Carpeta con videos de reconocimiento en línea de tremor volcánico.
3. Ejecutable, base de datos con información generada en este trabajo y código fuente del programa LAKiY.

Bibliografía

BENBRAHIM, Mohammed. DAOUDI, Adil. BENJELLOUN, Khalid. IBENBRAHIM, Aomar. Discrimination of seismic signals using artificial neuronal networks; *Enformatika v4* 2005 ISSN 1305-5313; Morocco, Unite Estates.

BISHOP,C., (1995) *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK.

CADENA Oscar. Arcos Darío. Detección automática de registros sísmicos asociados al comportamiento del volcán Galeras haciendo uso de Redes Neuronales Artificiales. Universidad de Nariño, 2008.

CHOUET, B., 1981. Ground motion in the near field of a fluid-driven crack and its interpretation in the study of shallow volcanic tremor. *Journal of Geophysical Research*, 86: 5985-6016.

CHOUET, B., 1985. Excitation of a buried magmatic pipe: a seismic source model for volcanic tremor. *Journal of Geophysical Research*, 90: 1881-1893.

CHOUET, B., 1986. Dynamics of a fluid-driven crack in three dimensions by defined difference method. *Journal of Geophysical Research*, 91: 13967-13992.

CHOUET, B., 1988. Resonance of a fluid-driven crack: radiation properties and implications for the source of long-period events and harmonic tremor. *Journal of Geophysical Research*, 93: 4375-4400.

CHOUET, B., 1992. A seismic model for the source of long-period events and harmonic tremor. In: P. Gasparini, R. Scarpa and K. Aki (Editors), *Volcanic Seismology*. Springer-Verlag, Berlin, pp. 133-156.

CHOUET, B., 1996. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature*, 380: 309-316.

DEL BRIO, Bonifacio. SANZ,Alfredo. *Redes Neuronales y Sistemas Difusos*; Universidad de Zaragoza. 2002.

DEL PEZZO, Edoardo. ESPOSITO, Anna. GIUDICE, Flora. MARINARO, Maria.MARTINI, Marcello. SCARPETTA, Silvia. Discrimination of earthquakes and underwater explosions using neuronal networks. 2003.

GIL-CRUZ, Fernando. Observations of two special kinds of tremor at Galeras volcano, Colombia (1989-1991). INGEOMINAS. 1999.

GOMEZ Diego. LAVERDE Carlos. NARVAEZ Lourdes. ORTEGA Adriana. SILVA Betty. TORRES Jesús. TORRES Roberto. Catálogo de señales sísmicas volcánicas de Colombia. INGEOMINAS. 2004.

GOMEZ, D. M., TORRES, R. A., 1997. Unusual low-frequency volcanic seismic events slowly decaying coda waves observed at Galeras and other volcanoes. *J. Volcanol. Geoth. Res.*

GOMEZ, D. M., TORRES, R. A., SEIDL, D., HELLEG, M., RADEMECHER, H., 1999. Tornillos seismic events at Galeras volcano, Colombia: A summary and new information from broad band three-component measurements. *Ann. Geofis.*

ISLAM, Tamana. Interpolation of Linear Prediction Coefficients for speech coding; McGill University. Montreal Canada. 2000.

KUBOTERA, A., 1974. Volcanic tremors at Aso volcano. In: L. Civetta, P. Gasparini, G. Luongo and A. Rapolla (Editors), *Physical Volcanology*. Elsevier, New York, pp. 29-48.

KUBOTERA, A., and Yoshikawa, K., 1963. Prediction of volcanic eruption at Aso and Sakurazima and some related geophysical problems, *Bulletin of Volcanology*, 26, 297-317.

LATTER, J. H., 1979. Volcanological observations at Tongariro National Park, 2. Types and classification of volcanic earthquakes, 1976-1978. Department of Scientific and Industrial Research, Geophysics Division, Wellington, New Zealand, 150, 60 p.

MALONE, S.D., 1983. Volcanic earthquakes: Examples from Mount St. Helens, in *Earthquake, Theory and Interpretation*, edited by H. Kanamori and E. Boschi, Soc. Itala. Di Fisica, pp. 436-455, Bologna, Italy, 1983.

MINAKAMI, T., 1974. Seismology of volcanoes in Japan. In: L. Civetta, P. Gasparini, G. Luongo and A. Rapolla (Editors), *Physical Volcanology*, Elsevier Scientific Publishing Company. The Netherlands, pp. 1-27.

MINAKAMI, T., 1960. Fundamental research for predicting volcanic eruptions, Part I, *Bulletin of Earthquake Research Ins.*, 38, 497-544.

PROAKIS, John. MANOLAKIS, Dimitris. Tratamiento digital de señales. 1998.

RUGELES, Alvaro. CADENA Oscar. Discriminación entre registro sísmicos tipo LP y VT producidos por el volcán Galeras, Utilizando Redes Neuronales Artificiales. Universidad de Nariño, 2006.

SAKURABA, Ataru. OIKAWA, Jun. IMANISHI, Yuichi. Free oscillations of a fluid sphere in an infinite elastic medium and long-period volcanic earthquakes. *Earth plane space*, 2002.