



XVIII Congreso Peruano de Geología

IDENTIFICACIÓN DE PATRONES DE RECONOCIMIENTO PARA CLASIFICACIÓN AUTOMÁTICA DE EVENTOS SÍSMICOS DEL VOLCÁN MISTI (PERÚ)

Riky Centeno ¹, y Orlando Macedo ¹

¹ Observatorio Vulcanológico del Sur, Instituto Geofísico del Perú (IGP-OVS), Urb. La Marina B-19 Cayma, Arequipa, Perú (riky.centeno@igp.gob.pe)

1. Introducción

La principal tarea del Observatorio Vulcanológico del Sur (OVS) es analizar minuciosamente la actividad sísmica asociada a los volcanes Misti, Ubinas, Sabancaya y Ticsani, debido a que hoy en día, el método sísmico ha demostrado ser una herramienta eficaz en la comprensión de procesos volcánicos internos y el pronóstico de erupciones (Aki y Richards, 1980; McNutt, 1996; Chouet, 2003). Este análisis tradicionalmente requiere primero la detección y posteriormente la clasificación manual de los eventos sismo-volcánicos. Sin embargo, el correcto etiquetado de los eventos muchas veces se ve afectado por diferentes factores tales como: la baja relación señal-ruido de los eventos, o por un incremento notable de la sismicidad (centenares de sismos) justo antes de una erupción y se requieren de decisiones rápidas (Ibáñez et al., 2009). Recientemente, los modernos observatorios, tratan de complementar el trabajo manual con modernos sistemas de reconocimiento automático de eventos sismo-volcánicos, esto ayuda en gran medida a reducir el volumen de trabajo en el análisis de los datos, y permite concentrarse en actividades más complejas como la elaboración de reportes técnicos, implementar de modelos que expliquen el comportamiento volcánico, etc. Según el catálogo de sismos del volcán Misti, existen tres tipos de eventos principales: Volcano-Tectónicos (VT), Largo Periodo (LP) y Tremor (TRE). El objetivo de este trabajo es seleccionar parámetros de identificación de señales y diseñar un algoritmo de clasificación, asistido por computador, en base a máquinas de soporte vectorial (SVM), que automatice la clasificación de los eventos del

volcán Misti, y que la metodología empleada, sea de utilidad para los sismos de los volcanes Ubinas, Sabancaya y Ticsani.

2. Selección de datos

Para este estudio se emplearon datos de tres estaciones sísmicas, de la red sísmica de monitoreo permanente del volcán Misti (Fig. 1), estas son: la MISC, es de periodo corto y utiliza un sensor Lennartz LE-3D Lite de 1 segundo; la MISD y MISE, ambas de banda ancha y cada una tiene un sensor Guralp GMG-3SPEC de 30 segundos. Todas de tres componentes (Z, N, E), pero solo fue necesario utilizar la componente vertical (Z), debido a que proporciona una mejor relación señal-ruido en la mayoría de eventos. La base de datos contiene una selección de 100 eventos volcánicos por clase: VT, LP y TRE, así como otros 100 eventos de origen no volcánicos, denominados OTR. Todos estos eventos fueron registrados en al menos dos de las tres estaciones mencionadas, durante los meses de enero a diciembre del año 2015 (Fig. 2).

3. Procesamiento

El método adoptado en este estudio se dividió en dos etapas: pre-procesamiento y clasificación. En la etapa de pre-procesamiento, a la señal sísmica se le aplica un filtro pasa-banda butterworth de 4to orden entre las frecuencias 0.5-24Hz, debido a que la energía sísmica de los eventos volcánicos está contenida en ese rango de frecuencias. Luego "uno a uno" a cada evento de la base de datos, se selecciona el máximo de características o patrones típicos de cada tipo de evento, y en este estudio fueron seleccionadas 6 características.

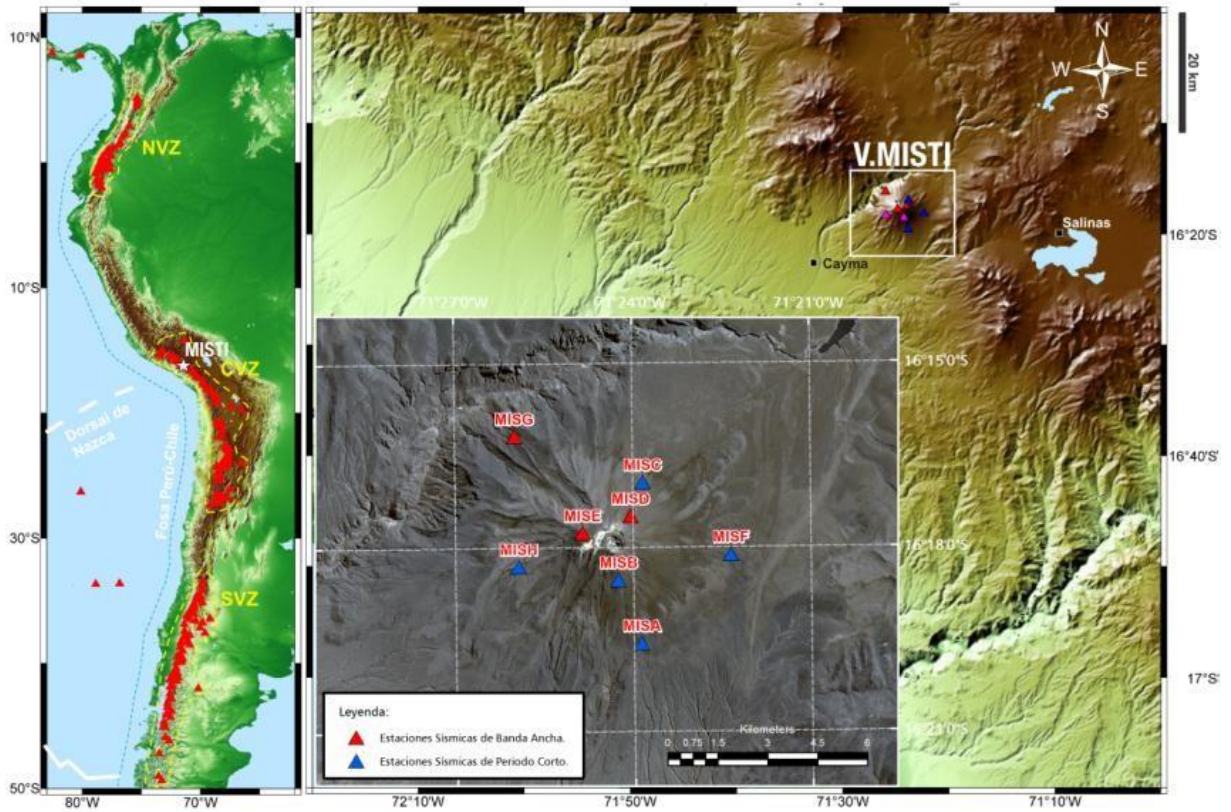


Figura 1.- Ubicación de las estaciones sísmicas utilizadas en este estudio, así como la red sísmica de monitoreo volcánico del Misti.

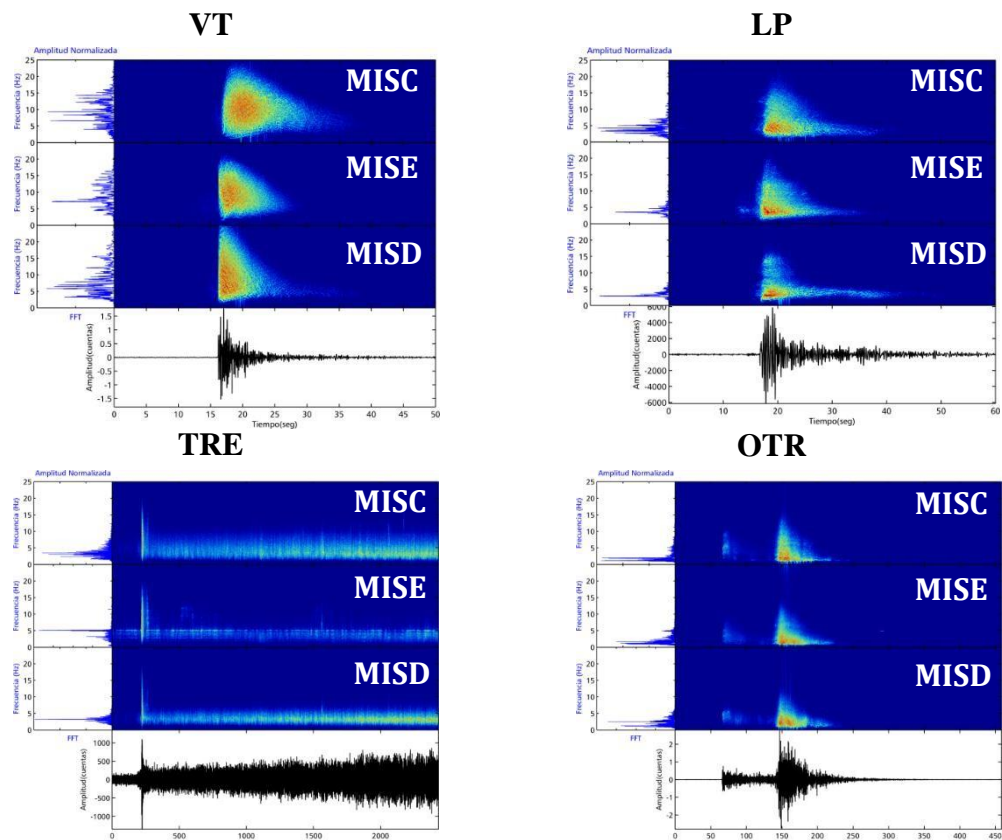


Figura 2.- Representación espectral y en el tiempo de los cuatro tipos de eventos (VT, LP, TRE y OTR), registrados en la componente vertical (Z) de las estaciones MISC, MISE y MISD.

En la Tabla 1, se describe cada uno de los elementos seleccionados, 2 de estos parámetros fueron extraídos de la representación en el dominio del tiempo; y los cuatro restantes, se obtuvieron a partir del espectro de potencia de Fourier, construido con ventanas de longitud variable que cubrían la duración completa del registro. Seguidamente, se forma una matriz o vector de n características \times m estaciones ($n=6$; $m=3$), y al final se obtiene un vector de 18 características por evento. Este vector es clave en el éxito del modelo de clasificación seleccionado.

Tabla 1.- Descripción general de las características utilizadas para clasificar los eventos volcánicos.

Dominio	Característica	Descripción
Tiempo	Dt	Duración del evento (inicio-fin).
Frecuencia	F5	Frecuencia pico principal del evento total.
Frecuencia	Es	Frecuencia pico máxima al 97% de la energía espectral máxima.
Tiempo	Ev	Índice asimétrico de la envolvente (amplitud normalizada al cuadrado) de la señal, en los primeros 35 segundos.
Frecuencia	Kt	Medida de la kurtosis espectral.
Frecuencia	As	Medida del índice de asimetría del espectro de Fourier.

En la etapa de clasificación, se diseñó un modelo de clasificación en base a máquinas de soporte vectorial (SVM), que es uno de los métodos más modernos utilizados en modelos de clasificación y una herramienta muy útil en sistemas de aprendizaje automático. En general un modelo de clasificación automática requiere básicamente tres cosas: una tabla o matriz de aprendizaje, en la cual están las variables predictoras (características) y la variable a predecir (tipo de evento); una tabla de decisión, en la cual el usuario va a validar la construcción del modelo, esta tabla debe contener las variables predictoras y la variable a predecir en el mismo orden de la anterior; y por último un modelo de clasificación generado a partir de la tabla de aprendizaje, que en este caso sería una máquina de soporte vectorial. Este modelo estadístico se construye a partir la tabla de decisión, y en esta misma se evalúa el modelo de clasificación diseñado, es decir, se compara el resultado del modelo con el valor verdadero de la variable a predecir (tipo de evento). Si ambos coinciden, el modelo es aceptable, y se le aplica a nuevos individuos o eventos volcánicos futuros.

La idea principal del algoritmo de las máquinas de soporte vectorial (SVM), es encontrar un hiperplano que separe dos clases o grupos con el mayor margen posible, en principio este algoritmo solo funciona para dos clases, pero si se aplica iterativamente puede aplicarse a más de dos clases. El problema clasificación consiste en observar en qué lado del hiperplano, queda el vector que represente a un nuevo individuo. Entonces el problema matemático será encontrar estas separaciones mediante una simple función lineal $g(x)=w^t x+b$, donde: $g(x)$ representa a la clase, x a las variables predictoras, w y b son las variables a calcular que determinan el plano de separación. Mientras esta función lineal discriminante

tenga un mayor margen de separación, se reducirá en gran medida la posibilidad de mezcla entre las clases.

Para la aplicación de las máquinas de soporte vectorial, así como el procesamiento de las señales (lectura, filtrado, reconocimiento de patrones, etc.) se utilizó el software Matlab.

4. Discusión de los resultados

En muchos trabajos de clasificación automática de señales sismo-volcánicas, se han enfocado en la aplicación de redes neuronales artificiales (ANN), modelos ocultos de Markov (HMM) y en algunos casos bayesianas, pero todos ellos basados únicamente en señales provenientes de una sola estación de registro (Chu et al., 2010). Muchas veces diferenciar un evento VT y un LP resulta difícil, debido a que las señales en su trayecto (fuente-estación) sufren alteraciones en su forma y firma espectral (efectos de sitio), por lo que un VT fácilmente puede ser confundido por un LP. En este estudio se propone utilizar información proveniente de varias estaciones, para mejorar la estructura de nuestro modelo de clasificación utilizando máquinas de soporte vectorial (SVM).

Para determinar la precisión del algoritmo de clasificación aplicado a los eventos sismo-volcánicos (VT, LP, TRE y OTR) del Misti, según la metodología propuesta en este estudio, se analiza primero la matriz de confusión (Tabla 2), y los resultados muestran que los eventos VT están correctamente clasificados en un 99% de los casos, los eventos LP en un 97%, los tremores en un 91% y la categoría OTR un 90%. En general, las máquinas de soporte vectorial, clasifican correctamente las señales en promedio un 94%. Del análisis de los resultados, muestran que hay una buena concordancia entre el resultado del modelo aplicado a la tabla de decisión construida, con el valor manual (real). La clasificación de los eventos VT y LP fue la mejor, debido a dos de sus principales características: su duración y energía espectral. Otro buen resultado fue la clasificación de los eventos TRE y OTR con un 91 y 90% de efectividad.

Tabla 2.- Matriz de confusión, como resultado de la clasificación de las máquinas vectoriales de soporte.

Clase	Pronostico				Total	
	VT	LP	TRE	OTR		
Real	VT	97	0	0	1	98
	LP	0	91	3	2	96
	TRE	0	3	83	6	92
	OTR	1	0	9	84	93
%	VT	99	0	0	1	100
	LP	0	97	2	1	100
	TRE	0	2	91	7	100
	OTR	1	0	9	90	100

5. Conclusiones

En este trabajo de investigación, se ha desarrollado con éxito un algoritmo asistido por computador que utilizando máquinas de soporte vectorial (SVM), se identificó seis patrones de reconocimiento automático para los tres tipos

de eventos sismo-volcánicos del Misti: Volcano Tectónicos (VT), Largo Periodo (LP) y Tremor (TRE). Además de un cuarto tipo que no tiene ninguna relación con la actividad volcánica del Misti denominado OTR. Este algoritmo en lo posterior será implementado y complementara el monitoreo permanente que viene realizando el OVS para los volcanes Ubinas, Sabancaya y Ticsani.

Agradecimientos

Un agradecimiento especial al director del Observatorio Vulcanológico del Sur, Dr. Orlando Macedo Sánchez, por haberme permitido trabajar estos últimos 5 años en el monitoreo sísmico de volcanes como: el Ubinas, Misti y Sabancaya. Desarrollando mi talento en el arte del procesamiento de señales sismo-volcánicas.

Referencias

Benson P., Vinciguerra S., Meredith P. and Young P. 2010. Spatio-temporal evolution of volcano seismicity: A laboratory study, *Earth and Planetary Sciences Letters*, p. 297, 315-323.

Centeno, R. 2012. Análisis de la actividad sísmica del volcán Misti entre octubre 2005-diciembre 2008 y su dinámica interna actual. Tesis de Grado, Universidad nacional de San Agustín, p. 73-99.

Chu. A., Orozco M., Londoño M. 2010. Combinación fija de clasificadores para la discriminación de señales sísmicas volcánicas, *Boletín de Ciencias de la Tierra* número 27, Noviembre de 2009 - Junio de 2010, Medellín, ISSN-01203630.pp.37-48.

Cortés, G., García, L., Alvarez, I., Benítez, C., De la Torre, A. and Ibáñez, J.M. 2014. Parallel System Architecture (PSA): An efficient approach for automatic recognition of volcano-seismic events. *Journal of Volcanology and Geothermal Research*, Volume 271, 1 February 2014, p 1-10.

Curilem, M., Vergara, J., Martin, C. S., Fuentealba, G., Cardona, C., Huenupan, F. and Yoma, N. B. 2014, Pattern recognition applied to seismic signals of the Llaima volcano (Chile): An analysis of the events' features. *Journal of Volcanology and Geothermal Research*, 282, 134-147.10.1016/j.jvolgeores.2014.06.004

Orozco M., Acosta C., Londoño M., 2012, The Automated Identification of Volcanic Earthquakes: Concepts, Applications and Challenges, *Earthquake Research and Analysis - Seismology, Seismotectonic and Earthquake Geology*, Dr Sebastiano D'Amico (Ed.), InTech, DOI: 10.5772/27508.pp.377-379

Scarpetta, S., Giudicepietro, F., Ezin, E. C., Petrosino, S., Del Pezzo, E., Martini, M. and Marinaro, M., 2005. Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks. *Bulletin of the Seismological Society of America*, vol. 95, no. 1, pp.185-196.