

## CAPÍTULO 5

# APLICACIONES Y RESULTADOS

---

### *5.1 Introducción*

El estudio por medio de wavelets permite describir señales no estacionarias proponiendo un enfoque más amplio que el análisis de Fourier; un GLE se describe como un evento de rápida transitoriedad lo cual es ideal para el análisis por wavelets. El contenido en frecuencias de los RCG antes y después de un evento altamente energético, puede correlacionarse con frecuencias de otros índices de actividad solar, proporcionando información sobre las conexiones entre diferentes capas solares. En el presente capítulo proponemos una clasificación de los GLE en función de su espectro de potencia wavelet.

En el mismo capítulo realizamos la descripción del espectro precursor, con características que consideramos no son comunes en los espectros de RCG lejanos a un evento. Inicialmente suponemos que una atmósfera lejana a un evento puede considerarse estable, de tal forma, que al acercarse la fecha de ocurrencia la atmosfera comienza a modificarse; estas alteraciones se registran como perturbaciones en la señal de RCG y son las que describimos en los espectros de potencia wavelet. El análisis cualitativo realizado por wavelets lo corroboramos con la técnica de ACP; donde el objetivo es probar que un espectro que precede un evento altamente energético posee características distintas a un espectro lejano a un evento.

En cada apartado presentamos los resultados obtenidos dejando la discusión y conclusiones para el capítulo final. El código del análisis wavelet es una librería modificada de *Matlab*; para el estudio por ACP, el algoritmo es una versión modificada del algoritmo original proporcionado por el *Dr. Mauricio Orozco*<sup>1</sup> y el *Dr. Carlos Ortiz Alemán*<sup>2</sup>.

---

<sup>1</sup> Dr. Mauricio Gabriel Orozco del Castillo, Investigador del Instituto Mexicano del Petróleo.

<sup>2</sup> Dr. Carlos Ortiz Alemán, Investigador del Instituto Mexicano del Petróleo.

## 5.2 Clasificación de GLE en Función del Contenido Espectral

La duración de un GLE es relativamente corta (de unas cuantas horas) en comparación con los periodos representativos de los RCG (años, meses, días, horas); la actual clasificación se encuentra en función del porcentaje de incremento en el contador de partículas de acuerdo al promedio de oscilación en un intervalo de tiempo observado (denominado en inglés como background); con lo anterior tenemos una idea de la magnitud de los eventos.

El comportamiento de los RCG antes y después de un GLE puede brindar información del medio circundante; en publicaciones recientes se han encontrado relaciones entre diferentes capas solares por el análisis de las periodicidades presentes en esta señal (Pérez-Peraza *et al.*, 2009).

Proponemos una clasificación en función del contenido espectral que no solo toma en cuenta el incremento de partículas; ahora clasificamos los GLE de acuerdo a los periodos que los describen.

El análisis por medio de wavelets permite identificar como impacta un GLE sobre la señal de RCG; un incremento instantáneo en el número de partículas es descrito con un espectro de potencia característico; los valores que registra la potencia espectral no son de interés en el estudio ya que la descripción se encuentra enfocada en la *distribución temporal* de los periodos que representan a un GLE. En todos los espectros de potencia wavelet presentados a continuación, las zonas en rojo pertenecen a las regiones de mayor potencia y las regiones en azul corresponden al extremo opuesto de menor potencia. La escala de los espectros de potencia wavelet se presentan en días, para la comodidad del lector mostramos en la siguiente tabla la equivalencia en minutos y horas para una mejor percepción de los periodos en los ejes:

PERIODO DÍAS	EQUIVALENCIA	PERIODO DÍAS	EQUIVALENCIA
<b>4</b>	4 DÍAS	0.125	3 HORAS
<b>2</b>	2 DÍAS	0.0625	1.5 HORAS
<b>1</b>	1 DÍA	0.03125	45 MINUTOS
<b>0.5</b>	12 HORAS	0.015625	22.5 MINUTOS
<b>0.25</b>	6 HORAS	0.0078125	11.25 MINUTOS

Tabla 5.1 Periodos en días con sus respectivas equivalencias.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

Un GLE se observa como una dilatación súbita en la intensidad de la señal de RCG; en los eventos de mayor incremento puede apreciarse como una discontinuidad sobre la señal; al realizar el espectro wavelet se toma en cuenta la distancia de esta discontinuidad (incremento), y además, se caracteriza la evolución en tiempo hacia su estabilidad oscilatoria.

Considerando el decaimiento temporal, un evento puede acceder a distintos periodos. Existen casos en los cuales la señal de RCG vuelve a su periodo promedio de oscilación horas después de ocurrido el evento, mientras que en algunos otros (generalmente los de mayor intensidad) puede tardar varios días.

Puesto que la duración de los eventos es de unas cuantas horas, consideramos que en una ventana de aproximadamente cuatro semanas podemos capturar las características que guarda la señal de RCG, analizando quince días antes y quince después de ocurrido el evento; dado esto, el tamaño de las ventanas utilizadas es cercano a 30 días, centrando el evento en el día 15; el tamaño no influye en el comportamiento característico de los espectros; por comodidad la unidad de tiempo correspondiente es de un día. Debido a la falta de información respecto a los primeros 20 eventos, utilizamos datos con un muestreo horario; para al resto de los evento accedemos a muestras cada 5 minutos. En todos los gráficos siguientes el eje de las abscisas representa la evolución temporal (excepto el EGW) y el eje de las ordenadas a los periodos o número de partículas dependiendo el caso.

El registro de un GLE en la señal de RCG presenta una combinación entre el incremento súbito en el contador de partículas y el tiempo de estabilización; esto tiene por consecuencia una composición de periodos que describen la relación en el plano *tiempo-frecuencia*. Cuando no existe registro de evento, la señal de RCG guarda una oscilación promedio el cual contiene periodos característicos. Cuando se presenta un GLE, la señal de RCG se deforma saliéndose de su comportamiento común; analizamos los espectros de potencia wavelet de esta deformación en todo los eventos, y de acuerdo a la relación *incremento-estabilización* logramos dividir los 70 registros en tres grupos que denominamos Grupo A, Grupo B y Grupo C. A continuación presento las características que describen cada grupo, así como los eventos que contienen.

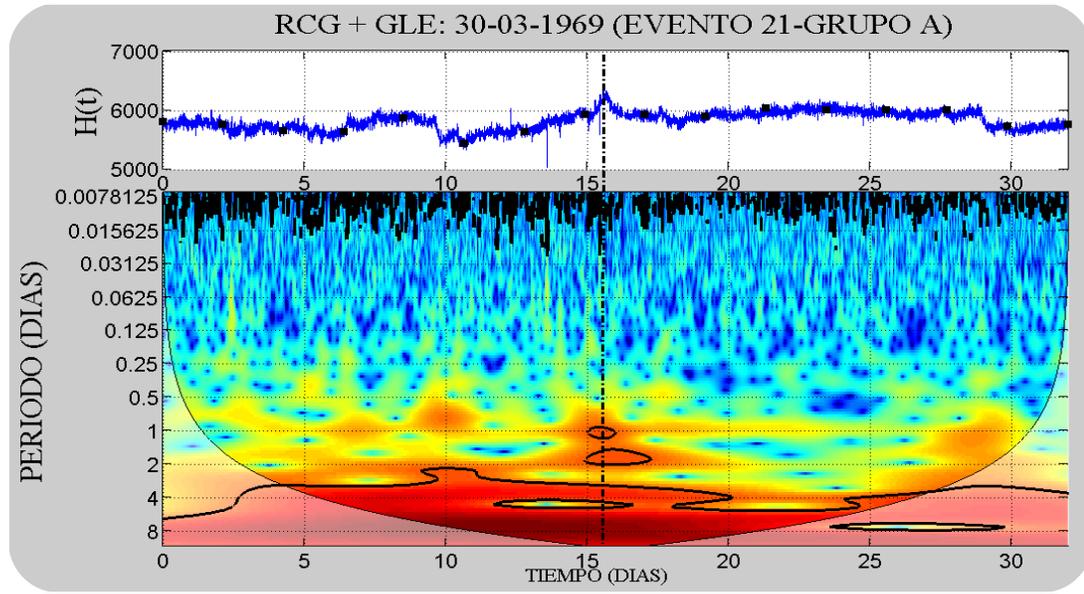


Figura 5.1 Espectro wavelet del evento 21 con un incremento reportado del 6%; el evento está señalado con la línea negra punteada

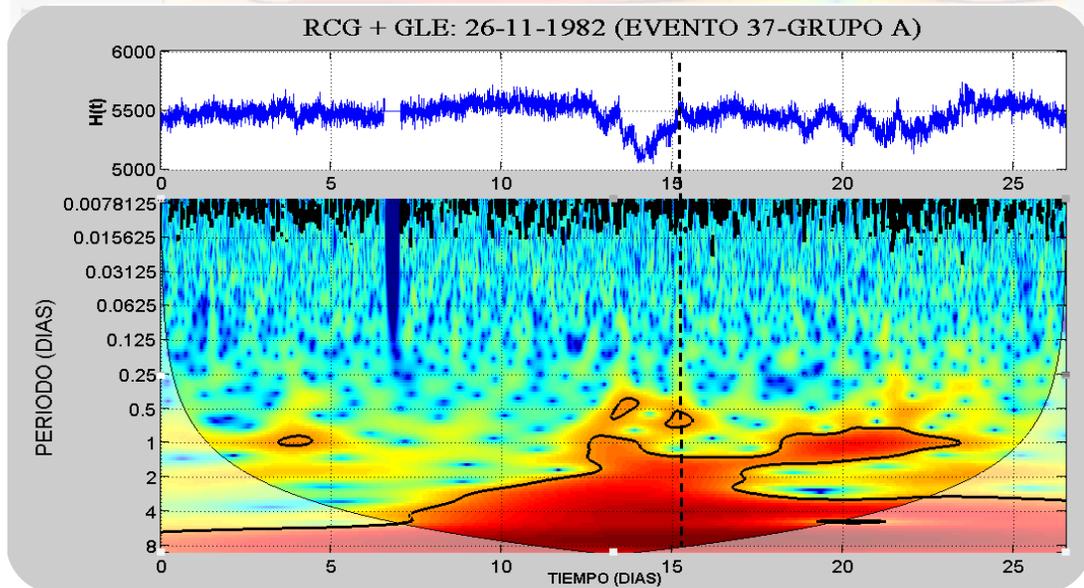


Figura 5.2 Espectro wavelet del evento 37 con un incremento reportado de 4%, el evento está señalado con la línea negra punteada.

La tabla B.1 (anexo B) incluimos 24 eventos; el mayor incremento está dado por el evento 12 y 21 con un 6%, y el mínimo se debe a los eventos 15, 20 y 34 con 1%. Este grupo contiene aproximadamente la tercera parte de todos los GLE registrados, describiendo eventos de muy baja intensidad que apenas son detectables en la señal de RCG; su espectro se manifiesta como un

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

tenue disturbio sobre los periodos de RCG; a continuación muestro los espectros wavelet de 2 eventos que caracterizan al denominado Grupo A.

Los eventos presentados en las dos figuras anteriores, son característicos del Grupo A, la señal  $H(t)$  es el número de partículas registrados en el contador de la estación de Oulu, el muestreo correspondiente es de 5 minutos. La escala de tiempo del panel superior es la misma que la descrita por su espectro.

En el Grupo A la potencia espectral de los GLE es muy pequeña que apenas logra perturbar la señal de rayos cósmicos, es por esto que en el Espectro Wavelet el evento no es evidente (figuras 5.1 y 5.2) como lo es en los casos de los grupos B y C.

La base de periodos altos es muy similar a los Grupos B y C, lo que identifica el Grupo A es el periodo mínimo al que logra acceder; el periodo menor que describe la perturbación por causa del GLE es de 12 horas, estos eventos son difíciles de detectar debido a que la señal de rayos cósmicos puede tener perturbaciones que se describen con un espectro similar.

El Grupo B (tabla B.2 anexo B) contiene 23 eventos; observamos el mayor incremento en los registros 16 y 17 con 17%, el incremento mínimo se debe a los eventos 25, 53, 64 y 65 con 5%.

La tabla B.2 contiene aproximadamente la tercera parte de todos los eventos; los describen un mayor número de periodos con un rango de acceso que puede alcanzar los 0.125 días (3 horas), a diferencia de la tabla B.1 el Grupo B abarca periodos menores de 12 horas.

El espectro de los eventos 36 y 55 (figuras 5.3 y 5.4 respectivamente) representa el comportamiento característico en los eventos del Grupo B; estos se encuentran sustentados por un contenido espectral más amplio que los del Grupo A; los periodos van desde 4 días hasta 0.125 días, pero aun el rango de valores abarcados es menor que los incluidos el Grupo C.

En las siguientes figuras logramos observar el espectro de dos señales, estas describen un abanico donde se concentran potencias altas señalada con una línea punteada;  $H(t)$  representa el número de partículas en el tiempo, con una muestra cada 5 minutos. Las unidades de tiempo del panel superior son las mismas a las de su espectro representativo.

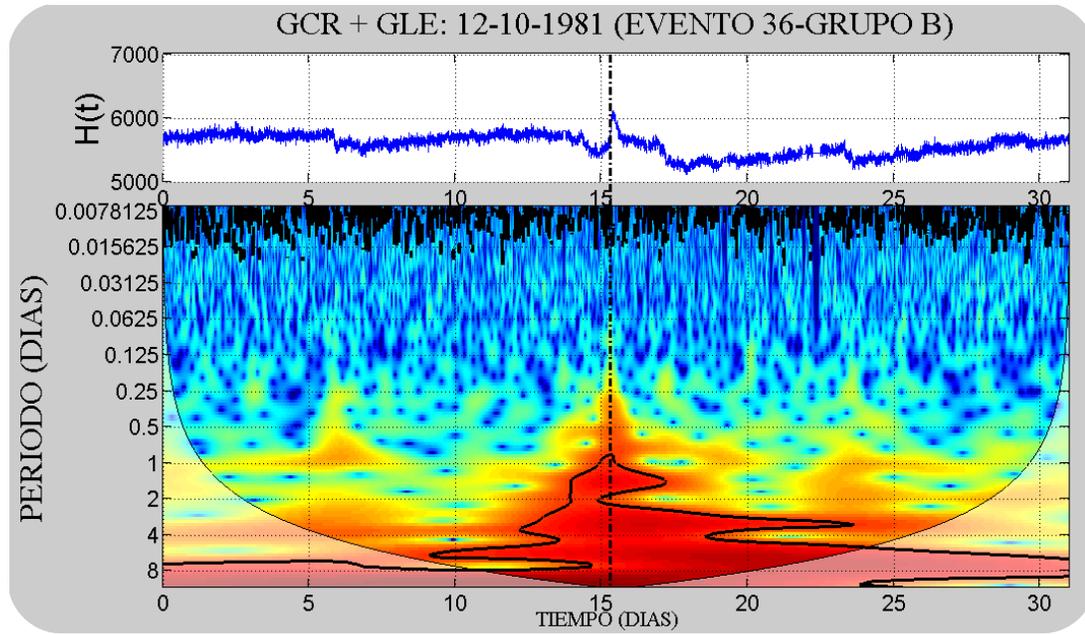


Figura 5.3 Evento 36 con un incremento reportado de 11%, el evento se encuentra resaltado por la línea negra punteada

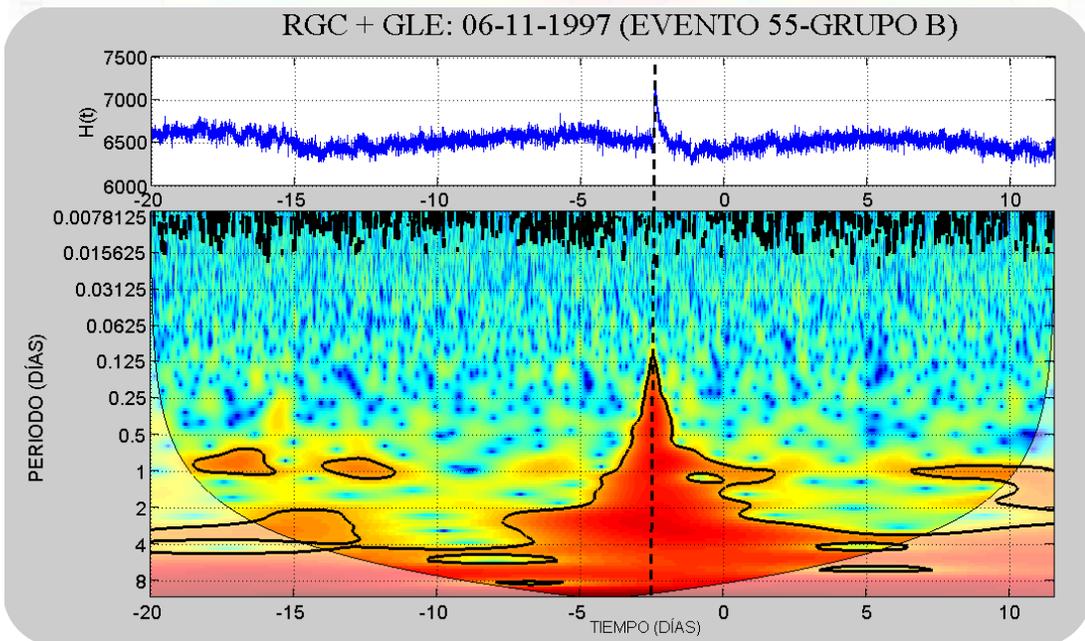


Figura 5.4 Evento 55 con un incremento reportado del 11%, el evento se resalta con una línea negra punteada.

La potencia espectral que aportan los eventos de esta clasificación logra perturbar el espectro wavelet de la señal de rayos cósmicos, representando un comportamiento específico. A partir de este grupo los periodos describen un comportamiento definido que solo pertenece a un GLE.

Finalmente los eventos de mayor potencia se encuentran en el Grupo C (tabla B.3 anexo B).

En la tabla B.3 (23 eventos), el evento más pequeño tiene apenas 12% en incremento (evento 41 y 61), mientras que el mayor sobrepasa 4000% (evento 5). La perturbación generada por el GLE es muy clara tanto en la señal temporal como en el espectro wavelet.

Los eventos de la tabla B.3 acceden a periodos menores a 0.125 días y pueden llegar a abarcar 0.007825 días (11 minutos, esto con un muestreo 5 minutos); el evento altera totalmente el espectro wavelet de la señal de rayos cósmicos centrándose la mayor cantidad de potencia en el momento de la perturbación. En las figuras 5.5 y 5.6 observamos el espectro característico del Grupo C (evento 69 y 42 respectivamente), en ambos espectros logramos detectar una atenuación en los periodos fuera del abanico descrito por el GLE.  $H(t)$  representa el número de partículas detectadas por el contador, con un muestreo de 5 minutos; las unidades de tiempo del panel superior son similares a las de su espectro correspondiente.

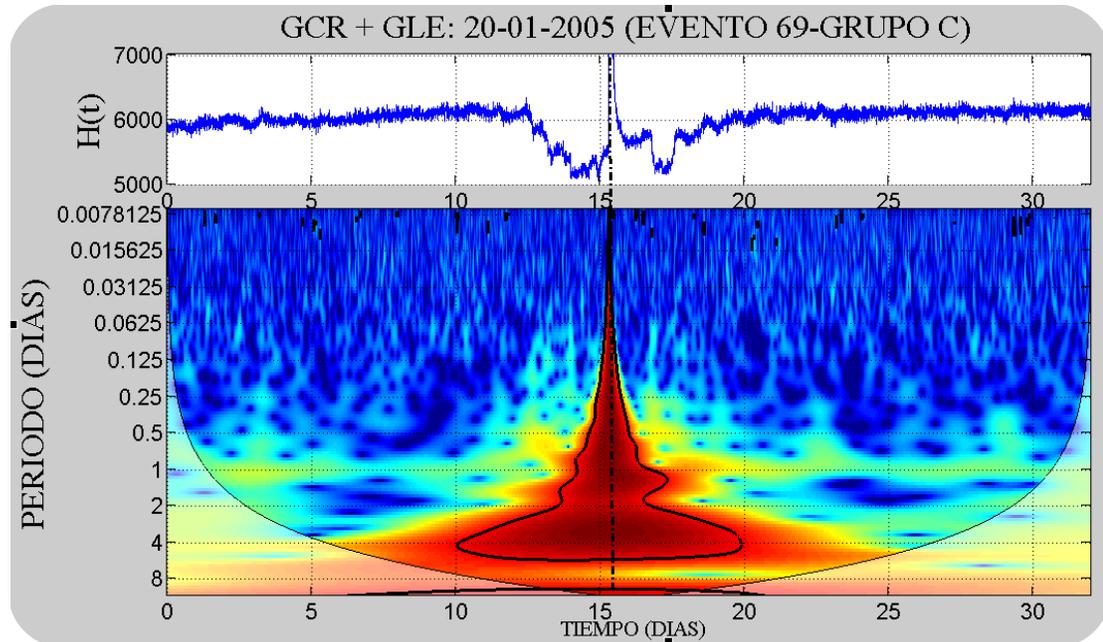


Figura 5.5 Evento con un incremento reportado de 269%, el evento es señalado con una línea negra punteada aunque es evidente en ambos dominios.

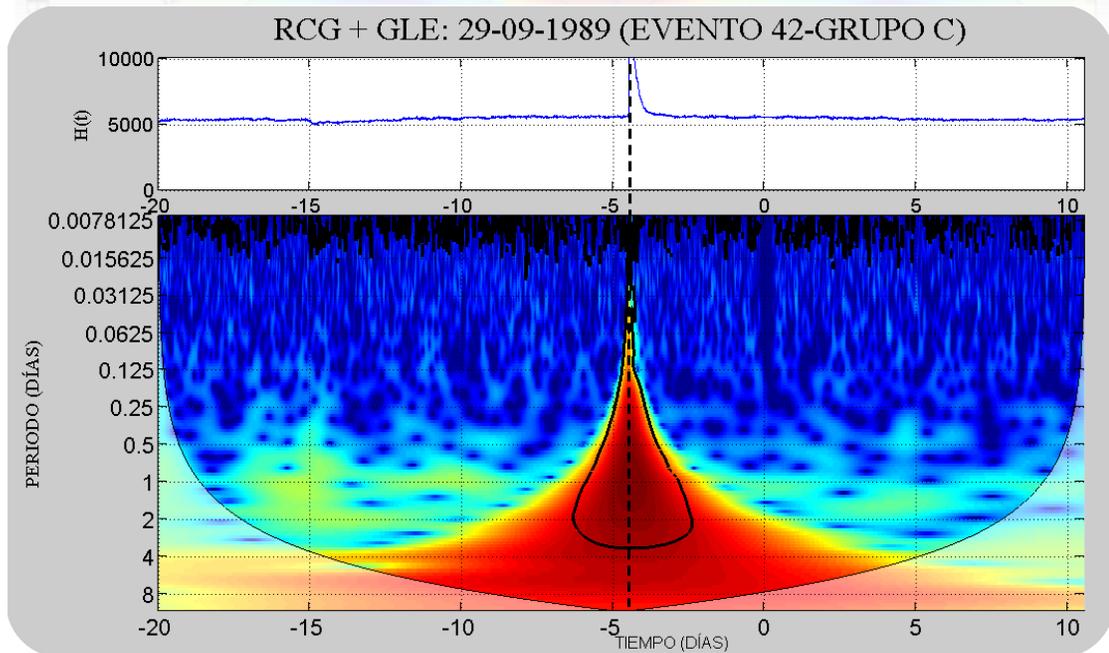


Figura 5.6 Evento con un 174% de incremento, el evento está señalado con una línea negra punteada aunque es evidente en ambos dominios.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

Los límites de los porcentajes no son exactos, basándonos del análisis anterior proponemos lo siguiente:

<b>INCREMENTO %</b>		
LÍMITE INFERIOR	LÍMITE SUPERIOR	GRUPO
1	<5	A
5	6	A ò B
>6	<12	B
12	17	B ó C
>17		C

Tabla 5.2 Límites asociados a los diferentes grupos.

<b>INCREMENTO %</b>		
GRUPO	LÍMITE INFERIOR	LÍMITE SUPERIOR
<b>A</b>	(<1	6]
<b>B</b>	[5	17]
<b>C</b>	[12	>17)

Tabla 5.3 Rango en porcentaje de los distintos grupos.

<b>PERIODOS CONTENIDOS DURANTE GLE [DÍAS]</b>										
GRUPO	4	2	1	0.5	0.25	0.125	0.0625	0.0312	0.01562	0.00781
<b>A</b>	X	X	X	X						
<b>B</b>	X	X	X	X	X	X				
<b>C</b>	X	X	X	X	X	X	X	X	X	X

Tabla 5.4 Periodos de acceso de los diferentes grupos.

Presentamos en las tablas 5.2, 5.3 y 5.4 los periodos asociados de cada grupo con los incrementos reportados en la literatura. Observamos en la tabla 5.2 que eventos con un mismo incremento pueden tener un contenido espectral distinto (representado por el Grupo A, B o C); estos resultados pueden brindar información sobre el origen de los eventos; en la tabla 5.3 asociamos

cada grupo al rango de porcentaje que describe un espectro similar, por último la tabla 5.4 presenta la gama de periodos que caracteriza a cada grupo.

### 5.2.1 Características por Grupo

#### Características del Grupo A

Contiene eventos menores o iguales a 6%, estos no alcanzan a perturbar el espectro wavelet, los periodos que describen al GLE son iguales en intensidad que los periodos característicos definidos por los rayos cósmicos, por lo tanto en estos casos la contribución en potencia espectral es similar. El grupo tiene periodos que van de 4 días hasta mayores a 0.5 días. 24 de los 70 eventos analizados pertenecen a este grupo.

#### Características del Grupo B

Contiene eventos con un incremento mayor o igual al 5% y eventos menores o iguales al 17%; el grupo logra perturbar al espectro wavelet generando periodos superiores a 0.125 días, aunque el rango de periodos abarcados es mayor que en el Grupo A, la contribución en potencia aun sigue siendo similar a la contribución energética de los periodos característicos de los RCG, 23 eventos de los 70 analizados pertenecen a este grupo.

#### Características del Grupo C

El grupo C es el que aporta mayor potencia espectral a la señal de rayos cósmicos, contiene eventos con un incremento mayor o igual al 12%; la característica esencial del GLE es generar una gran gama de periodos que pueden ir desde los 4 días hasta los 0.007813 días. El rango de periodos que describe al evento es de mayor intensidad que los periodos característicos de los RGC. La tabla contiene algunos eventos con incrementos relativamente pequeños (en comparación con el máximo reportado); la característica peculiar que hace que se encuentren dentro del grupo es el corto tiempo que tarda la señal de RGC en estabilizarse; un incremento súbito y una rápida estabilización hace que eventos reportados con un porcentaje pequeño (17%), accedan a periodos similares a los eventos con un porcentaje mucho mayor (200%); existen

eventos en los que la señal de RCG tarda días para volver al rango promedio de oscilación después de un GLE.

Lo anterior se propone porque necesitamos describir los GLE de acuerdo a su impacto sobre la señal de rayos cósmicos galácticos, de esta manera sabemos cómo afecta el evento altamente energético al espectro wavelet de la señal y al mismo tiempo, el modo en que se estabiliza la señal después de ocurrido el evento. Es por esto que clasificamos dentro un mismo grupo un evento de 50% en incremento que un evento de más de 200%.

Un segundo objetivo de esta clasificación, es proporcionar información sobre el origen de los GLE, los periodos pueden relacionarse con otro tipo de señales y detectar el origen entre diferentes regiones del sol (Pérez-Peraza *et al.*, 2009).

### 5.3 Análisis Wavelet por Ventanas Orientado a la Detección de Precursores GLE.

Una gran cantidad de fenómenos detectados en la naturaleza se conducen de manera sinusoidal, el comportamiento general de los rayos cósmicos no exenta esta tendencia ondulatoria (figura 5.7 datos captados por la estación de Oulu con una muestra por día).

Esta señal se encuentra modulada con frecuencias características brindando información sobre la física solar, así entonces, la señal captada por los monitores de neutrones es afectada en gran parte por procesos que se llevan a cabo dentro de la estrella; los RCG contienen información del estado del clima espacial y al gestarse un GLE la atmosfera solar comienza a cambiar, postrando características particulares sobre la señal de RCG, características que en una atmosfera lejos de un evento no se muestran o por lo menos no son comunes. Cabe aclarar que lo RCG no son de origen solar; provienes de otras fuentes del universo bañando a la heliósfera y el espacio interestelar isotrópicamente, en su camino franquean el sol, cuya fenomenología produce diversas modulaciones.

Los GLE son considerados de naturaleza estocástica, al superponer los eventos con la señal de rayos cósmicos (figura 5.8) observamos una asociación directa con el ciclo de actividad solar; los eventos parecen tener mayor frecuencia respecto a los mínimos de un ciclo promedio de 11 años, mínimo perteneciente a los RCG. En la figura 5.8 presentamos el espectro wavelet de la señal de RCG (1964-2010, con una muestra por día); en la misma figura superponemos los eventos como pulsos (líneas negras verticales); en el espectro observamos cómo se distribuye la potencia con periodos que se intensifican y decaen pareciendo seguir la ocurrencia de GLE.

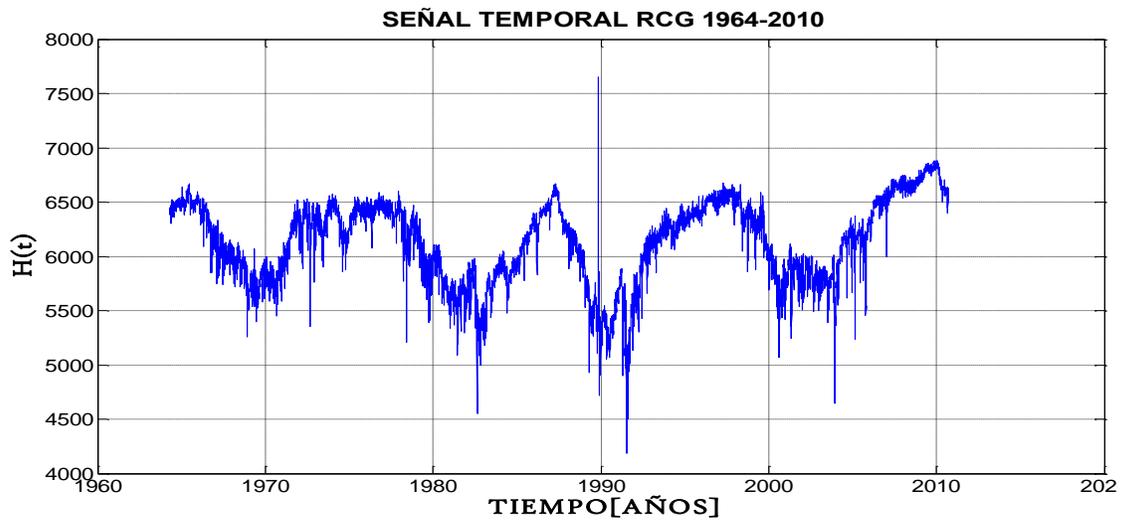


Figura 5.7 Señal temporal de RCG, muestreo diario, estación Oulu.

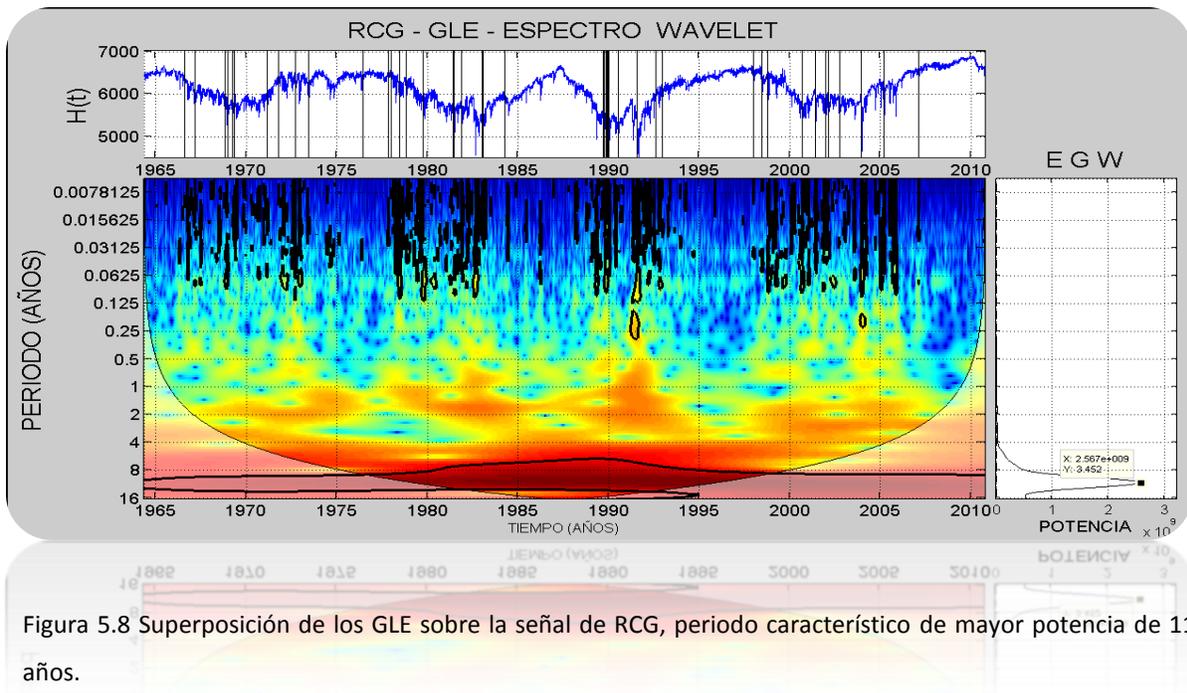


Figura 5.8 Superposición de los GLE sobre la señal de RCG, periodo característico de mayor potencia de 11 años.

En el EGW de la figura 5.8 distinguimos el periodo de mayor potencia, la lectura tomada en el máximo se encuentran en potencia base 2, por lo tanto la lectura del periodo máximo asociado es:

$$2^{3.452} = 10.9434 \text{ años.}$$

Del espectro wavelet de la serie temporal de 1964 hasta 2010 observamos los periodos reportados en publicaciones anteriores; el ciclo predominante es de aproximadamente 11 años; para detectar los periodos de menor potencia utilizamos el filtro pasa altas quitando el mayor aporte.

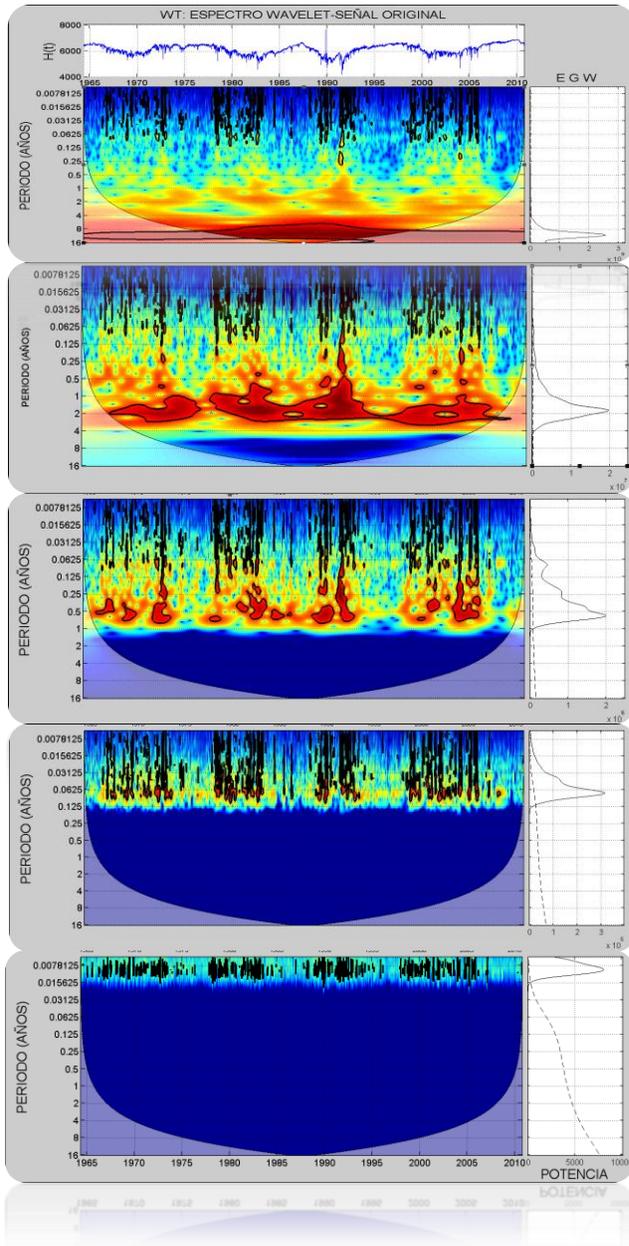


Figura 5.9 Espectros con filtros de Daubechies orden 8, 6, 4 y 1.

Realizando el proceso de filtrado detectamos los periodos con mayor contribución de potencia; con este proceso tenemos una idea general de los periodos característicos asociados a nuestra estación de estudio; en los espectros mostrados a la izquierda es posible distinguir que la intensidad de los periodos incrementan durante la ocurrencia de GLE, como la duración de los eventos es de unas cuantas horas los periodos observados describen directamente el comportamiento de los RCG, como el muestreo de la señal es de un día no podremos detectar la variación diurna que también es un periodo característico importante.

Los espectros en disposición descendente a excepción del primero contienen un filtro de orden 8, 6, 4 y 1 respectivamente, todos muestran el aporte de energía por encima del ruido rojo representado por la línea punteada en EGW, lo que significa una contribución con una significancia aceptable.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

En la siguiente tabla ordenamos los periodos de mayor a menor contribución de potencia, la lectura es tomada directamente del EGW en el proceso de filtrado (figura 5.9, panel derecho); estos periodos serán base de análisis para detectar el precursor, más adelante crearemos ventanas de distinto tamaño para el estudio a detalle y su caracterización temporal (evolución).

Los periodo están clasificados como mediano, corto y ultra corto plazo (Pérez-Peraza *et al.*, 2009); estos últimos los utilizaremos para el estudio de ACP.

PERIODOS MEDIANO PLAZO[años]	Promedio
10.94	11 años
5.16	5 años
1.72	1.7 años
1.22	1.2 años
0.58	7 meses
0.24	3 meses

Tabla 5.5 Periodos de Mediano Plazo

PERIODOS CORTO PLAZO [años]	Promedio
0.07615	27 días
0.038	14 días
0.01797	7 días
0.008987	3.5 días

Tabla 5.6 Periodos de corto plazo.

Enfocándonos al estudio de series temporales anuales, en las cuales buscamos diferencias entre años anómalos (con ocurrencia de GLE) y años normales (si ocurrencia de GLE) presentamos cuatro series del análisis total. Los periodos de estudio comprenden algunos del grupo de mediano plazo pero están directamente enfocados a caracterizar el comportamiento de las periodicidades en el grupo de corto plazo. Es así que realizamos una inspección de los periodos que anticipan los eventos.

Del análisis espectral de series anuales de RCG en la figura 5.10, las dos series contienen GLE mientras que en la figura 5.11 pertenece a series de años en los cuales no existe presencia de eventos. Tomando la lectura del Espectro Global Wavelet (panel derecho, EGW gráfico 5.10 y 5.11) el mayor aporte energético viene dado por periodos superiores a 16 días.

Los espectros con y sin evento están sustentados por una base de bajas frecuencias que se extienden sobre todo el intervalo de tiempo, además, el comportamiento de los periodos altos tiene la misma tendencia antes, durante y después de ocurrido el evento, por tanto pasan

desapercibido la anomalía en la serie de RCG; con este resultado nos dirigimos al análisis de periodos de corto plazo

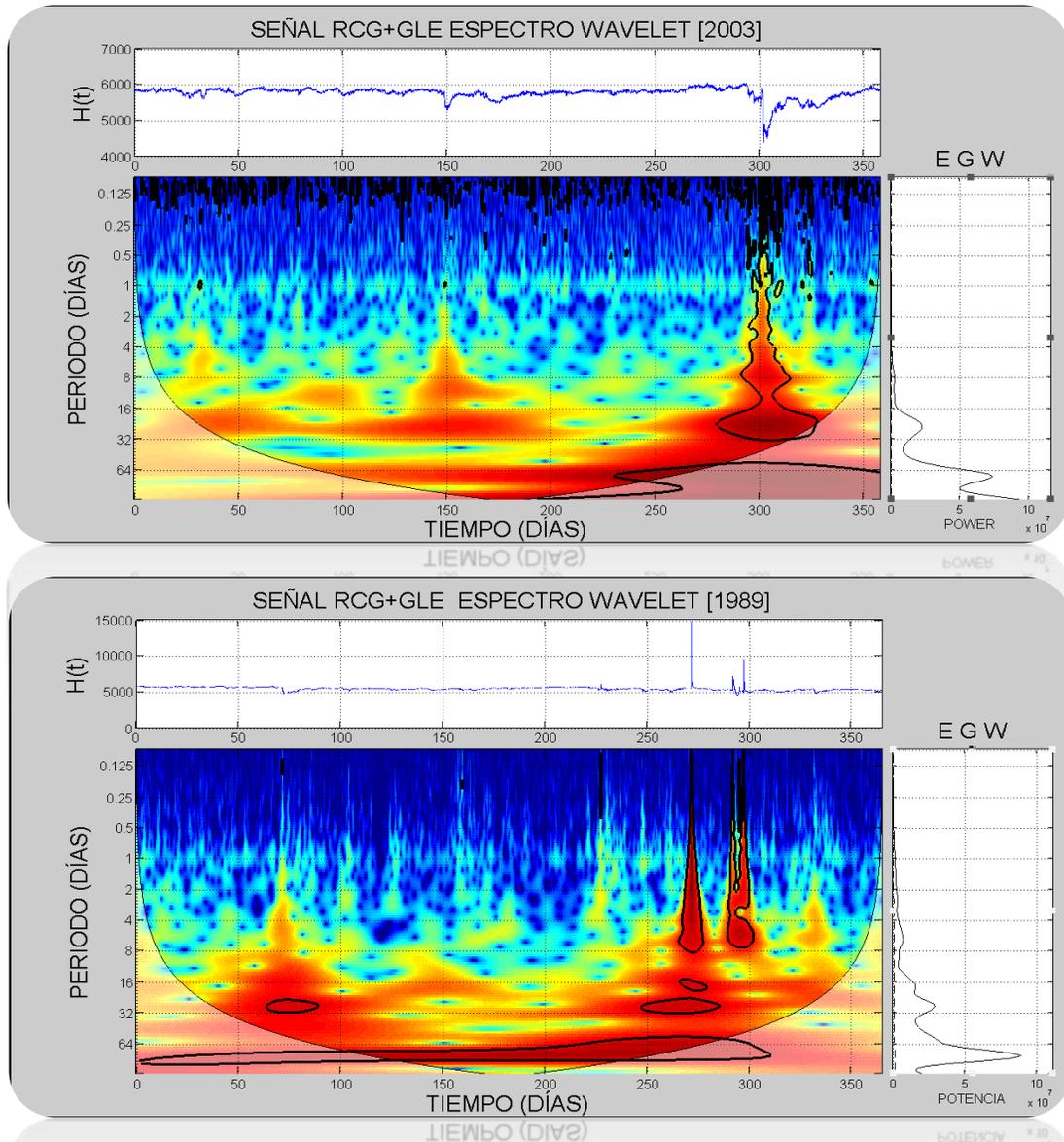


Figura 5.10 Espectro anual de series con presencia de GLE, con una muestra por hora, correspondientes al año 2003 y 1989 respectivamente.

De las figuras 5.10 y 5.11 observamos la distribución temporal de las frecuencias bajas; los espectros expuestos tienen en común la base de periodos largos, es así que distinguimos una gama de periodos similares mayores a 16 días; ya sea en un año con evento o sin evento la base se mantiene describiendo el comportamiento general de los RCG.

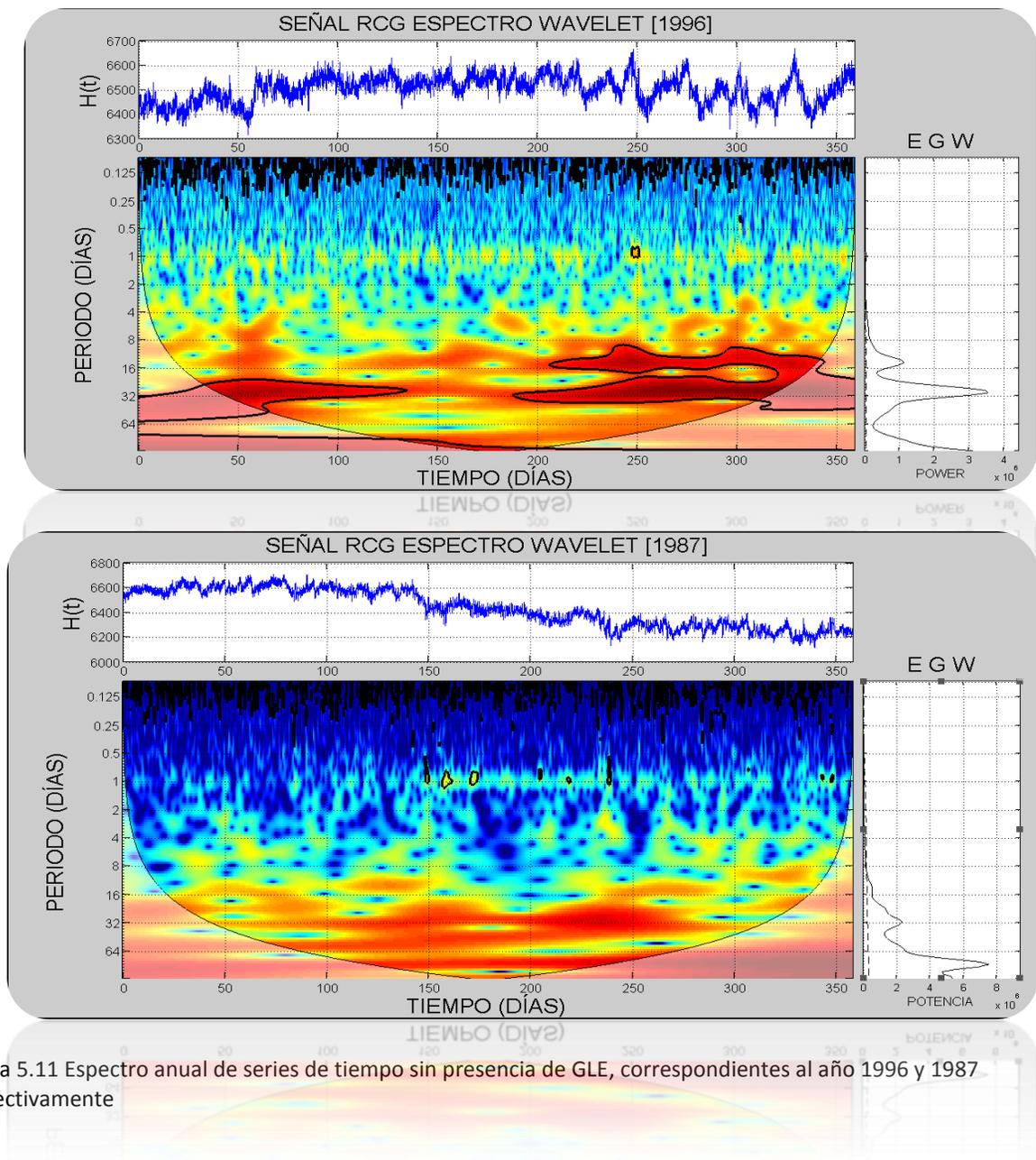


Figura 5.11 Espectro anual de series de tiempo sin presencia de GLE, correspondientes al año 1996 y 1987 respectivamente

Respecto al análisis de los periodos dominantes anuales (leídos en el espectro global EGW), en cada uno de los años sin evento y de los años en los cuales tenemos registros de GLE, no encontramos alguna tendencia en su comportamiento, dado esto limitamos nuestra búsqueda a periodos relativamente cortos; ahora nos enfocamos a ciclos menores a diez días, del análisis anterior elaboramos la tabla C.1 del anexo C.

Los resultados en la tabla C.1 ayudan a reducir el tamaño en la ventana, ya que el contenido de periodos mayores a 10 días es similar en ventanas que anticipan un GLE a ventanas de ruido espacial (señal RCG).

Consideramos que las periodicidades de interés se encuentran en el rango de corto plazo, abarcando días o posiblemente horas; para caracterizar las altas frecuencias reducimos el tamaño temporal de las ventanas.

Una observación más que podemos distinguir es que durante años con mayor actividad solar la potencia de los RCG aumenta, los espectros anuales con anomalía tienen una potencia con un valor de  $10^7$ , mientras que los espectros en años en los que no se tiene evento la potencia alcanza una magnitud de  $10^6$ , sin embargo el interés de estudio no se halla en el espectro global sino en la evolución de los periodos que contiene la señal; lo importante es determinar el desarrollo característico de la señal de RCG antes de un evento altamente energético.

De acuerdo a todo el análisis anterior y a la clasificación de los eventos, tenemos una gran gama de periodos, que si los analizamos como un solo conjunto es muy difícil utilizarlos como objetivos de pronóstico; es así que nos volvemos a eventos individuales para caracterizar la evolución antes de la fecha de ocurrencia del GLE.

Las ventanas de estudio son de 30, 15, 5 y 1 días, éstas nos permiten observar a detalle la evolución temporal de periodos menores a 10 días.

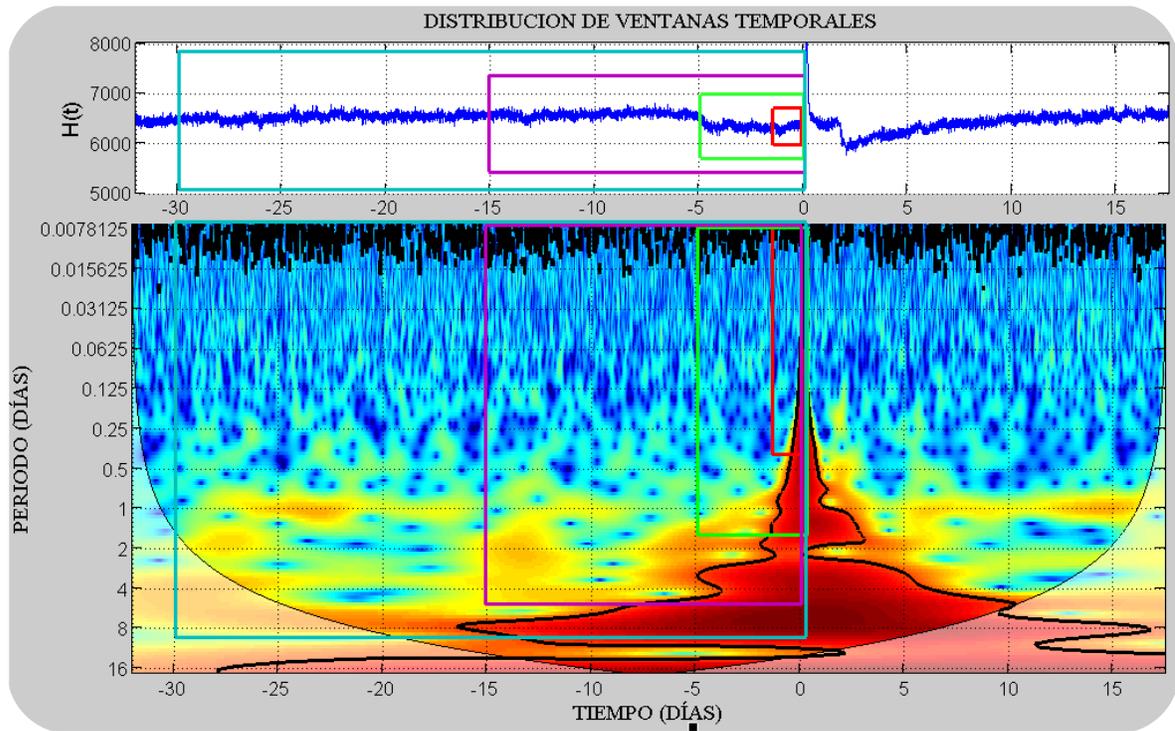


Figura 5.12 Distribución de ventanas temporales para el análisis de cada evento.

El GLE está ubicado en el día cero, por lo tanto la señal que lo anticipa se muestra en una escala de tiempo negativa, así ubicamos las ventanas aproximadamente 60 días antes del evento y observamos la evolución de los periodos antes de llegar al día cero.

En la figura 5.12 mostramos la distribución de las ventanas utilizadas, éstas avanzan en tiempo cada hora, actualizándose 24 veces en un día; podemos ver su evolución al realizar una película con el conjunto de espectros para cada evento. Las ventanas tienen asociadas un rango limitado de acceso a los periodos, por esto, para una ventana de  $N$  días accedemos a periodos máximos de  $N/3$ .

Como la señal de RCG está constituida por un amplio rango de periodos y la intensidad de color (en este caso la potencia) de los periodo es una función exponencial, realizamos un filtrado de bajas frecuencias (filtro pasa altas) para quitar el periodo de mayor aporte energético y de este modo, resaltar el comportamiento de periodos con menores potencias.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

La figura 5.13 presenta ventanas de distintos tamaños (30, 15, 5 y 1 días en lectura occidental) sin filtro, todas anticipan en una hora al GLE, cuando reducimos la ventana los periodos de corta duración se visualizan con un mayor detalle.

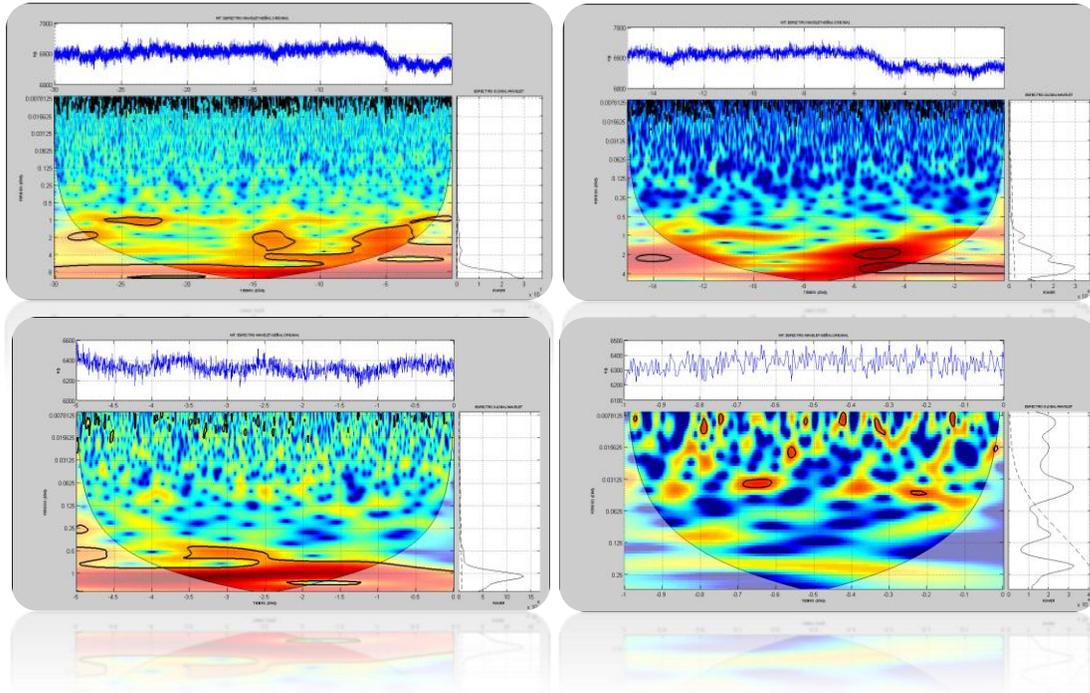


Figura 5.13 Ventanas de distintos tamaños con una distribución de frecuencias con distintas resoluciones; para eventos ocurridos después de 1969 el muestreo en la señal de RCG es de 5 minutos.

En la figura 5.14 observamos las ventanas anteriores con filtro pasa altas (orden 9), resaltando los periodos de menor contribución de potencia; en el caso de la ventana de 1 día, el filtro no cambia el espectro porque la mayor potencia es aportada por los periodos de plazo ultracorto (esto lo observamos en el EGW).

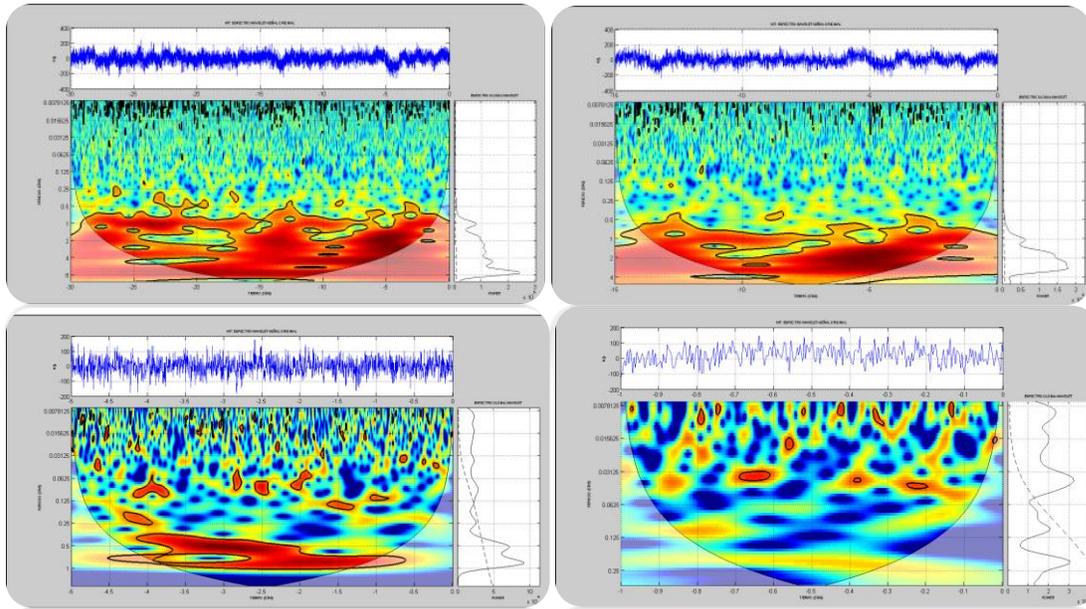


Figura 5.14 Espectros filtrados para resaltar periodos de menor potencia, las ventanas anticipan a un mismo evento en una hora.

Al caracterizar la evolución la primer ventana se encuentra a 60 días del evento, el paso temporal que le damos es de 1 hora, así entonces tenemos una ventana de tamaño  $N$  días que avanza en tiempo cada hora con un muestreo en la señal de 5 minutos, esto es, que la ventana al avanzar en tiempo se actualiza con 12 nuevos datos y desecha los últimos doce elementos de la serie.

Ya en este punto la hipótesis a probar es que los periodos fundamentales de la señal de RCG es la clave para detectar las firmas espectrales; esperamos que la intensidad de los periodos incrementen, decaigan o desaparezcan totalmente antes de un evento altamente energético, o de igual manera, que periodos no existentes en la señal comiencen a gestarse.

Al efectuar las distintas ventanas consideramos favorable una de 15 días, esta entrega los detalles suficientes para describir la evolución temporal del contenido espectral (periodos de 4, 2, 1 y 0.5 días), donde consideramos existe el precursor; esto porque no encontramos alguna característica en periodos mayores a 5 días, de esta manera discriminamos ventanas mayores a quince días (eliminando periodos superiores a 5 días).

Los periodos anteriores se encuentran catalogados como periodos característicos, por lo tanto podemos basarnos en su comportamiento para anticipar un GLE. La ventana de 15 días se justifica porque brinda información sobre los periodos requeridos, el interés se reduce a un rango de 5-0.5 días.

Al tener una resolución de 5 minutos es posible acceder a periodos mínimos de 10 minutos, el problema que surge es que los periodos menores a 12 horas se tornan turbulentos a nuestra percepción, de tal forma que no es posible detectar un patrón característico para describirlos; al observar la evolución del espectro wavelet de estos periodos pequeños (periodos de plazo ultra corto) parecen intensificarse, disminuir y desaparecer sin ningún patrón reconocible, por lo que con este primer método de caracterización, los periodos menores a 12 horas no es posible incluirlos dentro del estudio. Con esto podríamos justificar que el muestreo de la señal pudiese ser de 6 horas y optimizar el proceso de cómputo, pero trabajamos con 5 minutos por el tratamiento que realizamos por Componentes Principales, en cual hacemos que la computadora con ayuda de un método matemático encuentre las características que no son perceptibles a nuestros ojos.

### 5.3.1 Características

Los resultados de la caracterización no muestran solo un patrón antes de un GLE, esto resulta complicado debido al desarrollo temporal de los periodos; esta parte es un símil con la interpretación sísmica, en la cual para la definición de estructuras no existe una regla general, de manera que tenemos la necesidad de personal especializado en donde la interpretación está en función de la experiencia adquirida por la cantidad de imágenes analizadas.

En la figura 5.15 ilustramos un caso de cada grupo (A, B, C respectivamente) observando un comportamiento peculiar en la evolución de los periodos de los RCG; antes de la fecha de evento, existe una perturbación en la señal (línea punteada) que es reflejada en el espectro como un rango de periodo continuos [4-0.25] días, dichas perturbaciones presentan una tendencia a la deriva de alta a baja frecuencia, esta característica se presentan de 50 a 20 días antes del evento y de vez en cuando una segunda perturbación de 1 a 4 días antes del día cero.

La intensidad de los periodos en el momento de la perturbación se torna máxima, y a medida que se acerca la fecha del evento (día cero) comienza a atenuarse. Puesto que este patrón peculiar no

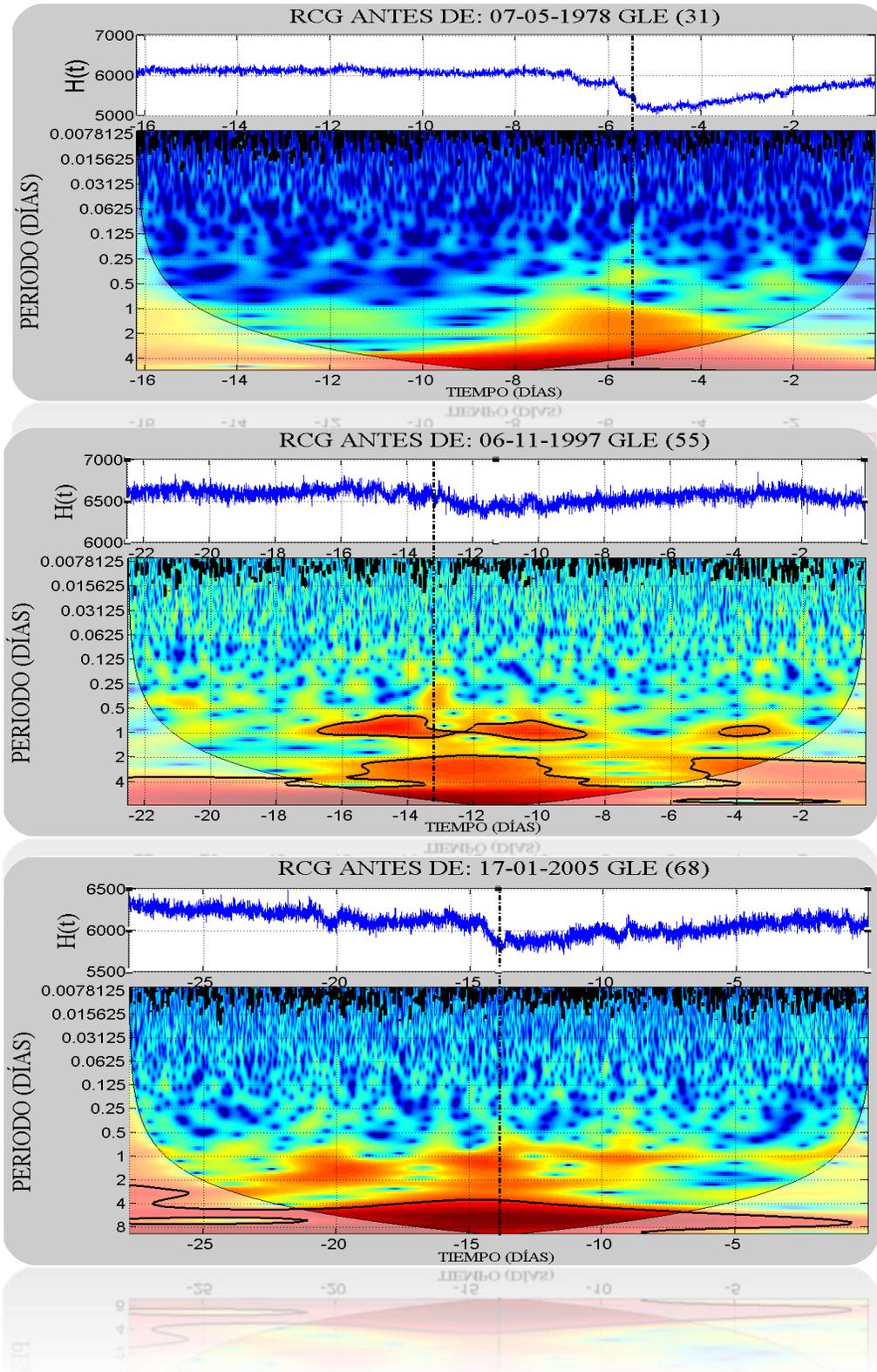
es común fuera de los periodos de ocurrencia de GLR, el resultado puede ser utilizado como herramienta de pronóstico.

Existe otro tipo de comportamiento que podemos caracterizar, es algo similar a la descripción anterior; para este caso el periodo de 1 día se mantiene intensificado días antes del evento formando una franja en tiempo muy notable que puede decaer o no días antes (u horas) de la fecha de ocurrencia, esto es, que se mantiene el comportamiento descrito anteriormente, donde existe la perturbación, solo que no todo el rango de periodos deriva a bajas frecuencias, si no que, se mantiene el periodo de 1 día por más tiempo antes del evento altamente energético, esta segunda característica la podemos observar en la figura 5.16.

El análisis anterior es totalmente cualitativo, lo que necesitamos es un argumento matemático que sustente nuestra teoría de que el espectro wavelet que anticipa a un GLE unos instantes antes es distinto a un espectro que se encuentra muy lejos de un evento; si probamos que los espectros son diferentes probamos al mismo tiempo que el precursor existe.

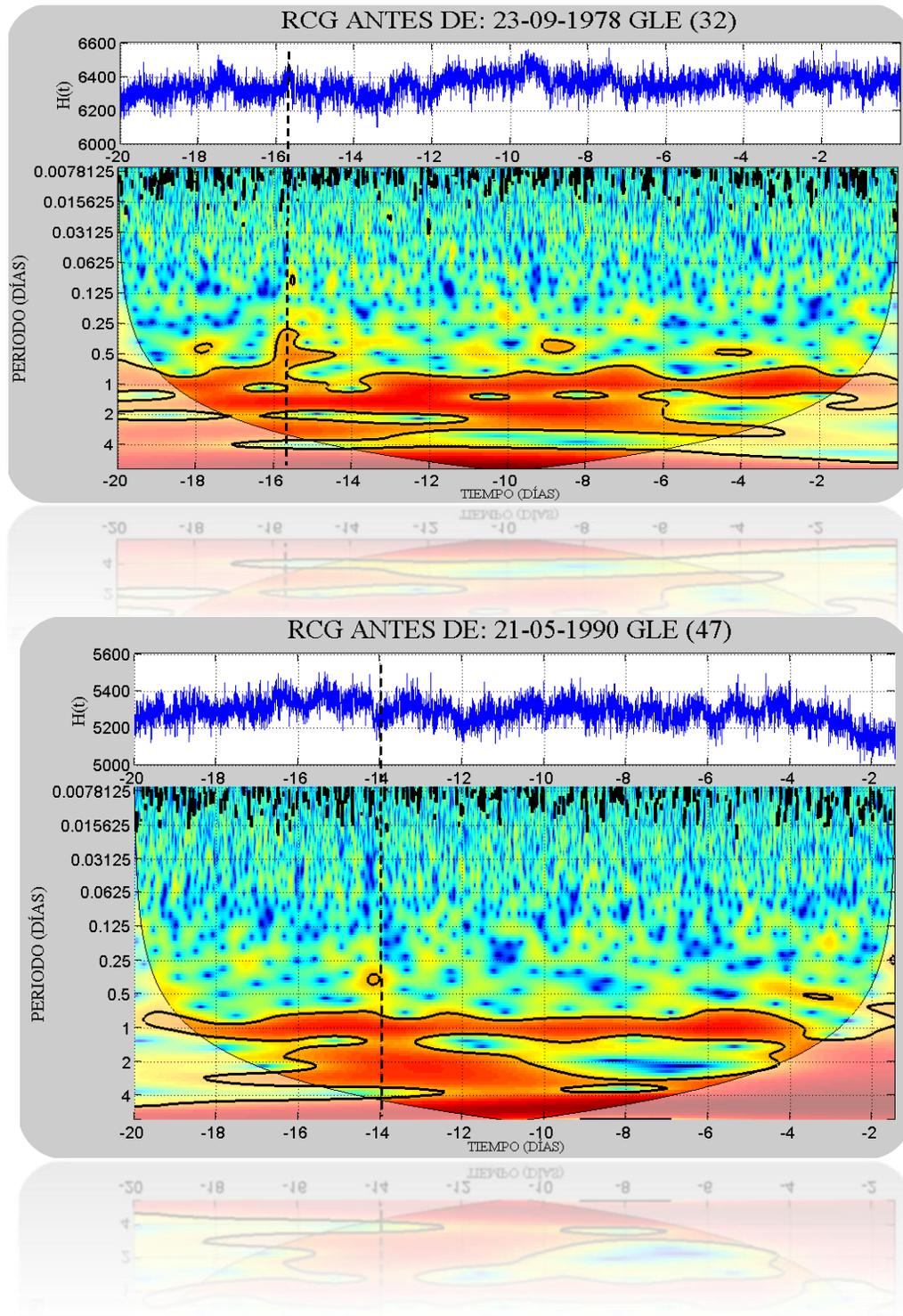
### 5.3.2 Comportamiento 1

Figura 5.15 Perturbaciones registradas en la señal de RCG reflejadas en los espectros wavelet.



### 5.3.2 Comportamiento 2

Figura 5.16 Comportamiento característico de los periodos antes de un GLE.



En todos los casos una perturbación es registrada en la señal de RCG antes de un GLE, de manera contraria no todas las perturbaciones anticipan a un GLE, de esta forma necesitamos trabajar para detallar las características sobre la señal de RCG que precede a un evento altamente energético.

Para esto proponemos el Análisis de Componentes Principales (PCA de sus siglas en inglés), este método toma las matrices de datos que generan los espectros wavelet y compara los espectros precursores con espectros no precursores, los detalles se describen a continuación.

## 5.4 Análisis de Componentes Principales de la Matriz Generada por Wavelets

Este tipo de análisis es utilizado para el reconocimiento de imágenes, una de sus aplicaciones se encuentra en la detección de rostros, enfocándose a características definidas tratando de simular lo que realiza el ojo humano. Aun cuando los rostros son similares en gran parte, el análisis de componentes principales es capaz de identificar los detalles específicos que distinguen uno de otro y los hace diferentes.

Las características que observamos en los espectros de potencia wavelet en gran parte se encuentran en función de nuestra capacidad de percepción; el método científico establece que un experimento debe de ser reproducible bajo las mismas condiciones, para dejar de lado el análisis cualitativo de los espectros wavelet sometemos las imágenes (espectros) a un estudio por sus componentes principales, de esta forma, todos podemos observar los resultados desde el mismo punto de vista, el punto de vista cuantitativo siendo este es totalmente reproducible.

La técnica es implementada para aprobar la premisa de que el precursor existe; suponemos que una ventana de quince días que precede al evento en una hora contiene información que podemos dirigir como precursor, así entonces, ventanas lejanas a un evento mostrarán características distintas a las ventanas que llamaremos precursoras o ventanas pre-evento.

El algoritmo es una adaptación de la versión original proporcionada por el *Dr. Mauricio Orozco* y el *Dr. Carlos Ortiz*. El método se encuentra inspirado en sus artículos publicados, orientados al reconocimiento de cuerpos salinos es secciones sísmicas.

El análisis de componentes principales permite observar características a detalle, por este motivo utilizaremos señales con un muestreo de 5 minutos, los eventos que cumplen esta particularidad comienzan en el evento 18.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

Trabajamos directamente sobre la matriz de datos obtenida por el análisis de ondículas; la señal temporal de 15 días contiene 4320 elementos, el tamaño de la matriz es 115x4320 definida en el dominio complejo, las 115 hileras son por un factor de 12 y por la escala máxima obtenida a partir del tamaño en la ventana. Presentamos las pruebas realizadas para tres matrices diferentes: real, valor absoluto y utilizando un filtro pasa altas.

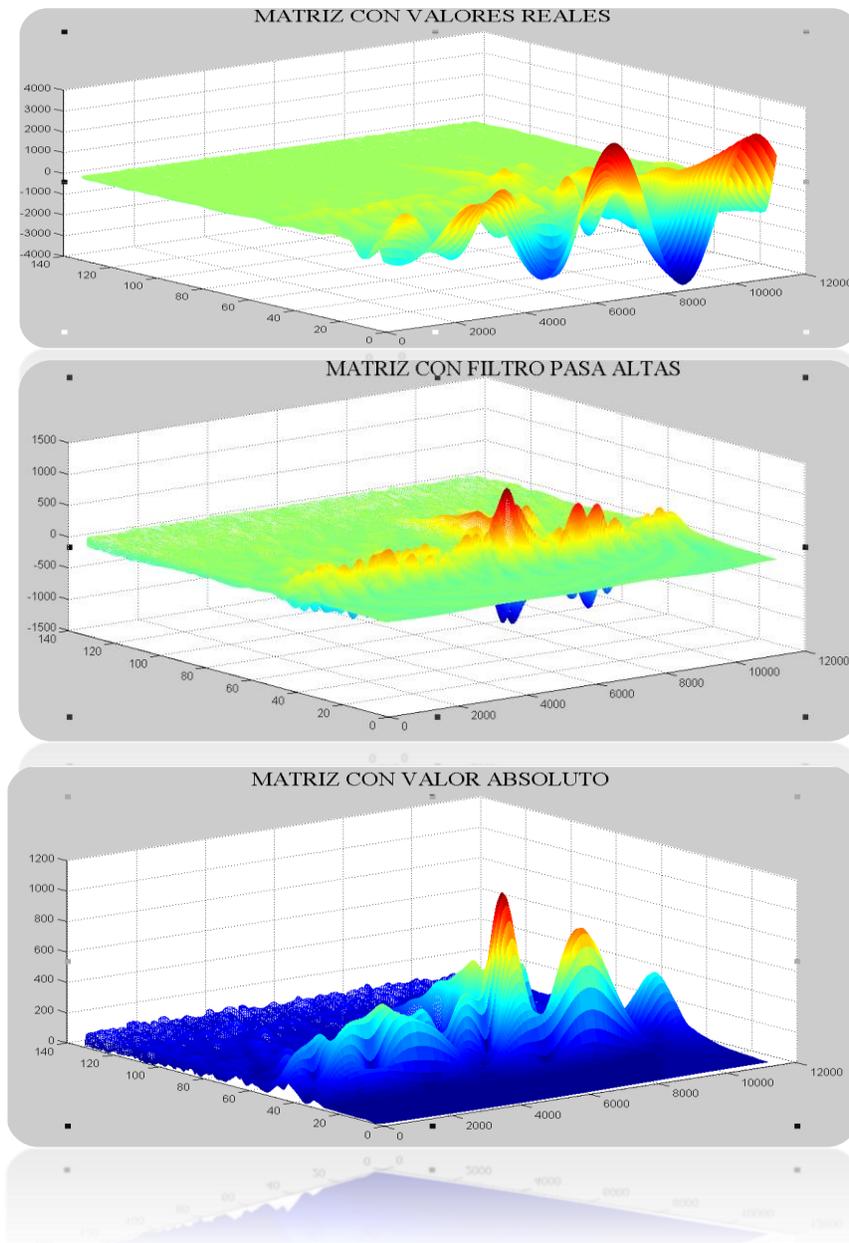


Figura 5.17 Matriz generada por el proceso de ondículas: real, filtrada y con valor absoluto.

El objetivo de realizar las pruebas con cada una de las matrices es seleccionar alguna que resalte las características del precursor; podremos evaluar la mejor matriz en función de la distancia de sus espacios vectoriales respecto a la base de entrenamiento, que se reflejará en un valor del estadístico  $Z$  más alto (respecto a una base entrenada con matrices no precursoras).

Tenemos acceso a 53 ventanas (una por cada evento); recortamos la ventana precursora unas horas antes del GLE y suponemos que todas éstas forman un conjunto con la característica esencial de preceder a un evento altamente energético.

Generamos un segundo conjunto de matrices que llamaremos matrices de ruido espacial; para asegurar que se hallan lejos de un evento las obtenemos de ventanas en años en los que no existe presencia de GLE. En total formamos un conjunto de 900 ventanas de ruido espacial que generan 900 matrices no precursoras y podemos comparar con el conjunto de matrices precursoras.

$$A = \{\text{MATRICES PRECURSORAS, 53 ELEMENTOS}\}$$

$$B = \{\text{MATRICES NO PRECURSORAS, 900 ELEMENTOS}\}$$

$$C = \{N \text{ MATRICES DE ENTRENAMIENTO} \in B\}$$

El algoritmo obtiene los valores y vectores característicos de un conjunto de matrices  $C$  que representa los elementos de entrenamiento; la base extraída del conjunto  $C$  es comparada con los vectores característicos del conjunto  $A$  y un subconjunto aleatorio proveniente de  $B$ . Esta comparación en realidad es la distancia de los espacios vectoriales; una distancia pequeña se refiere a la similitud de los eigenvectores, esto es, que los conjuntos son similares y pueden describirse con la misma base vectorial, de manera contraria si la distancia es grande, el conjunto al proyectarse en el espacio de entrenamiento estará lejos de ser reconstruido con la base vectorial entrenada.

La base de entrenamiento pertenece a un subconjunto de  $N$  elementos aleatorios contenidos en  $B$ . Elegimos una base de entrenamiento del conjunto  $B$  porque tenemos un número mayor de elementos de comparación; de este modo podemos tomar un conjunto de 40 matrices contenidas en  $B$  de manera aleatoria, y acceder a 860 matrices distintas para comparar la base entrenada con el mismo conjunto  $B$ ; de la misma forma accedemos a 52 matrices diferentes al comparar la base

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

entrenada con el conjunto  $A$ . De otra manera, si entrenamos la base con 40 elementos del conjunto  $A$  tenemos 900 matrices distintas en el conjunto  $B$  y solo 12 diferentes pertenecientes al conjunto  $A$ . Ya creada la base podemos acceder a  $N$  eigenvectores, discriminamos los aportes menores a 0.1%.

Al generar un subconjunto de matrices aleatorio  $f$  (contenida en  $B$ ) del cual realizamos un entrenamiento con todos los elementos, los vectores característicos serán capaces de reconstruir todos y cada uno de los elemento de  $f$  (si comparamos la distancia de las matrices proyectadas en el espacio de entrenamiento el resultado es cero)

Mostramos un resultado del algoritmo diseñado (de entrada solo necesitamos las matrices de entrenamiento):

$f =$  [Matrices de entrenamiento, 38 elementos]

En éste punto tenemos acceso a 38 valores característicos, el número siempre será igual o menor a la cantidad de matrices usadas en el entrenamiento, esto para mantener la independencia lineal (capítulo 4); accediendo a los valores característicos cuantificamos el aporte que realiza cada uno para la reconstrucción de un elemento de  $f$ .

Tabla 5.7 Distribución en porcentaje de los eigenvectores.

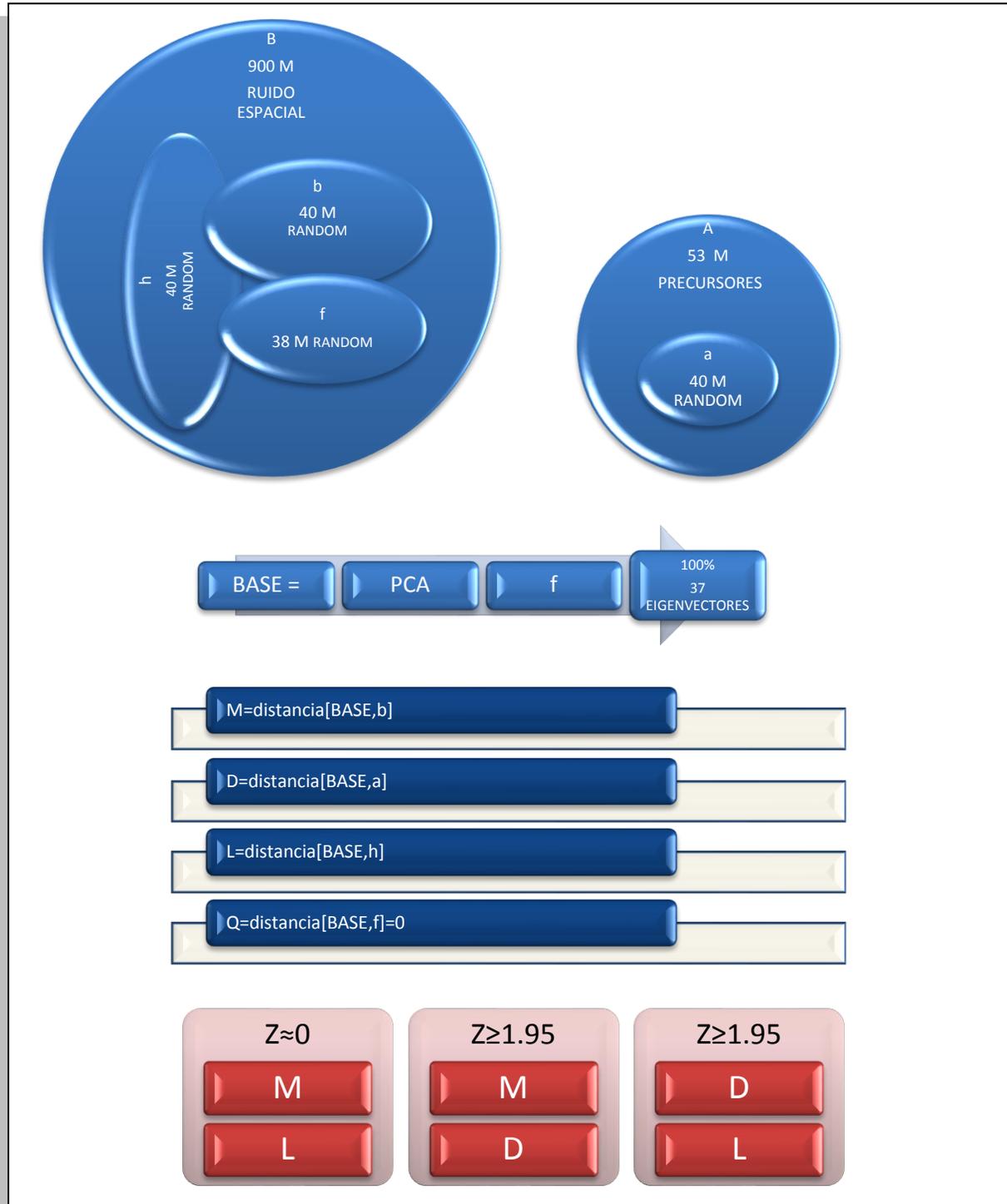
#	%	de	%												
	aporte		acumulado												
1	26.4159677		26.41596774	11	2.29788126		83.34585728	21	0.64267331		95.10888878	31	0.23766436		98.95744315
2	15.4919071		41.90787485	12	1.90073397		85.24659124	22	0.58833748		95.69722626	32	0.22258332		99.18002648
3	9.43536576		51.34324061	13	1.73536327		86.98195451	23	0.57212161		96.26934787	33	0.18649091		99.36651739
4	6.62192482		57.96516542	14	1.45556981		88.43752432	24	0.50343662		96.7727845	34	0.17493899		99.54145638
5	5.78159465		63.74676008	15	1.37055685		89.80808117	25	0.40347193		97.17625643	35	0.17081481		99.71227119
6	4.13122504		67.87798511	16	1.1938537		91.00193487	26	0.35986034		97.53611677	36	0.14557399		99.85784517
7	3.81860735		71.69659247	17	0.97952946		91.98146433	27	0.33327237		97.86938913	37	0.14215483		100
8	3.60968081		75.30627328	18	0.88340465		92.86486898	28	0.32205027		98.1914394	38	7.53E-16		100
9	3.00362834		78.30990161	19	0.81829618		93.68316517	29	0.27029259		98.461732				
10	2.7380744		81.04797602	20	0.78305031		94.46621547	30	0.25804679		98.71977879				



CAPÍTULO 5 APLICACIONES Y RESULTADOS

aleatorio de *A*; la *BASE* es entrenada con *f*; 37 eigenvectores logran caracterizar el 100% de los elementos pertenecientes a *f*.

Cuadro 5.1 Ejemplificación de la distribución de los conjuntos.



## CAPÍTULO 5 APLICACIONES Y RESULTADOS

Los vectores  $M$ ,  $D$ ,  $L$  y  $Q$  representan las distancias de cada elemento de los subconjuntos hacia la BASE de entrenamiento, en una expresión estadística representan la covarianza de cada matriz contenida en un subconjunto hacia la BASE vectorial entrenada; para el caso de  $M$ ,  $D$  y  $L$  los vectores contienen 40 números reales (cada uno representa la covarianza de las matrices hacia la BASE); se espera que las distancias en  $M$  y  $L$  sean menores que los contenidos en  $D$ , ya que los valores de  $M$  y  $L$  representan matrices pertenecientes al mismo conjunto de entrenamiento ( $b, h, f \in B$ ), de manera similar esperamos que los valores en  $D$  sean mayores que los anteriores, porque la premisa es que " $a$ " pertenece a un conjunto distinto lejos de ser representado por la BASE vectorial entrenada ( $D$  representa un conjunto que inicialmente consideramos diferente [ $a \in A$ ]).  $Q$  representa un vector nulo de 38 elementos, porque podemos representar cada elemento de  $f$  con la BASE de entrenamiento.

El estadístico  $Z$  lo utilizamos para tener un elemento matemático que garantice que los conjuntos son diferentes, así entonces la hipótesis a probar es la siguiente:

*"Las matrices del conjunto A son distintas a las contenidas en el conjunto B"*

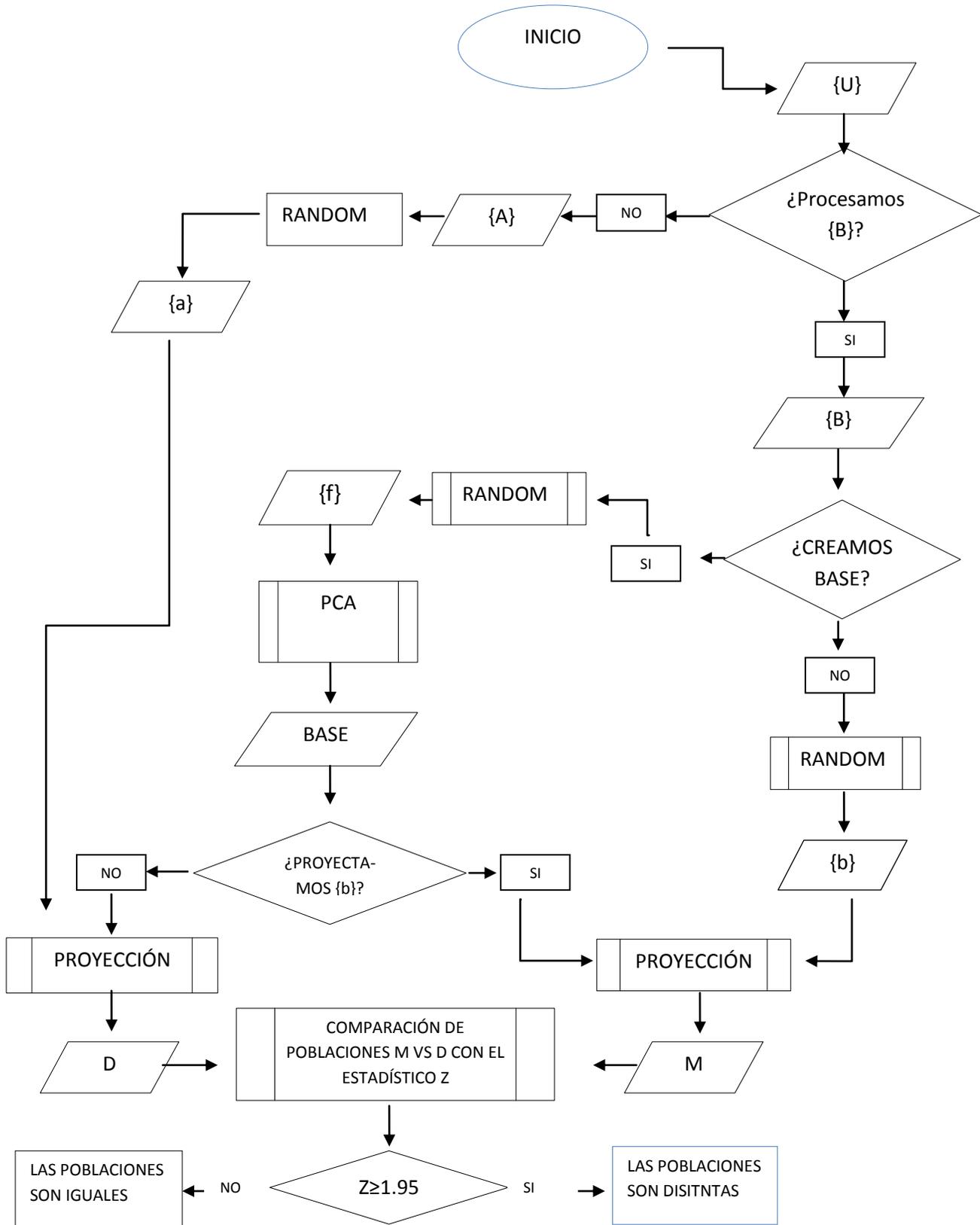
*Si  $Z \geq 1.95$  la hipótesis se acepta*

*Si  $Z \ll 1.95$  la hipótesis se rechaza*

Los recuadros en rojo muestran el valor de  $Z$  (al comparar las distancias) que se esperan en los experimentos; si comparamos los valores de  $M$  y  $L$  obtendremos un resultado  $Z$  cercano a cero debido a que las distancias son similares por representar a elementos de un mismo conjunto; para los casos siguientes al comparar  $M-D$  y  $D-L$  esperamos el valor  $Z \geq 1.95$  dado que confrontamos distancias que pertenecen a conjuntos diferentes.

El estadístico  $Z$  es útil para comprobar hipótesis de conjuntos con un número de elementos diferente, en las pruebas realizadas elegimos conjuntos representativos de 40 elementos para cumplir la condición de muestras grandes (mayores a 30 elementos).

En el siguiente diagrama explicamos el proceso del algoritmo con una base entrenada con  $N$  elementos contenidos en el conjunto B (ruido espacial):



$U = \{A, B, \setminus A \cap B = 0\}$

$A = \{\text{MATRICES PRECURSORAS, 53 ELEMENTOS}\}$

$a = \{n \text{ MATRICES ALEATORIAS } \in A, n \in \mathbb{N}\}$

$B = \{\text{MATRICES DE RUIDO ESPACIAL, 900 ELEMENTOS}\}$

$b = \{n \text{ MATRICES ALEATORIAS } \in B, n \in \mathbb{N}\}$

$f = \{N \text{ MATRICES ALEATORIAS } \in B, N \in \mathbb{N}\}$

$BASE = \{\sum EIGENVECTORES \geq 99.99\%, \#EIGENVECTORES \leq N\}$

$M = \{n \text{ DISTANCIAS LEJANAS DEL ESPACIO B}\}$

$D = \{n \text{ DISTANCIAS CERCANAS DEL ESPACIO B}\}$

### 5.4.1 Resultados del Algoritmo ACP

Presentamos los resultados del análisis con distintos números de entrenamiento, utilizando matrices de diferentes características, manteniendo constante la ventana de 15 días y una comparación (para evaluar el estadístico Z) de 40 matrices  $\in A$  vs 40 matrices  $\in B$ .

<b>MATRIZ GENERADA POR ONDÍCULAS (datos con filtro pasa altas)</b>		
<b># DE ENTRENAMIENTO</b>	<b># DE EIGENVECTORES <math>\approx 100\%</math></b>	<b><math>\langle Z \rangle</math></b>
20	[17-18]	6.2342
40	[35-38]	6.3335
180	[152-167]	6.1754
<b>MATRIZ GENERADA POR ONDÍCULAS (sin filtro)</b>		
<b># DE ENTRENAMIENTO</b>	<b># DE EIGENVECTORES <math>\approx 100\%</math></b>	<b><math>\langle Z \rangle</math></b>
20	[17-18]	6.2352
40	[35-38]	6.4476
180	[154-165]	6.5259
<b>MATRIZ GENERADA POR ONDÍCULAS (valor absoluto)</b>		
<b># DE ENTRENAMIENTO</b>	<b># DE EIGENVECTORES <math>\approx 100\%</math></b>	<b><math>\langle Z \rangle</math></b>
20	[17-18]	6.4307
40	[35-38]	6.2778
180	[152-160]	6.3309

Tabla 5.8 Resultados de la comparación de los conjuntos de matrices precursoras y matrices de ruido espacial.

La tabla 5.8 resulta de las pruebas realizadas, el número de matrices de entrenamiento propuestos son [20 40 180], 180 forma la quinta parte de la población de entrenamiento; el estadístico Z muestra la diferencia entre las poblaciones y en todo los casos aprueba la hipótesis de una manera contundente.

La diferencia en el trabajo de una matriz real, con valor absoluto o filtrada no influye de manera drástica en los resultados; adoptamos la matriz real para el análisis posterior.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

La matriz filtrada no influye de manera radical en los resultados porque solo filtramos un rango de periodos pequeño ( $>4$  días) para resaltar ligeramente los periodos de plazo ultra corto. El algoritmo logra identificar características similares en una matriz que trabaja con la parte real o con el valor absoluto de la matriz compleja.

En el análisis wavelet por ventanas describimos que la ventana de 15 días es el tamaño en el cual logramos caracterizar espectros distintos entre ventanas precursoras y ventanas lejanas a un evento; realizamos pruebas variando el tamaño en la ventana de 1 hasta 50 días, los resultados se presentan en la figura 5.19.

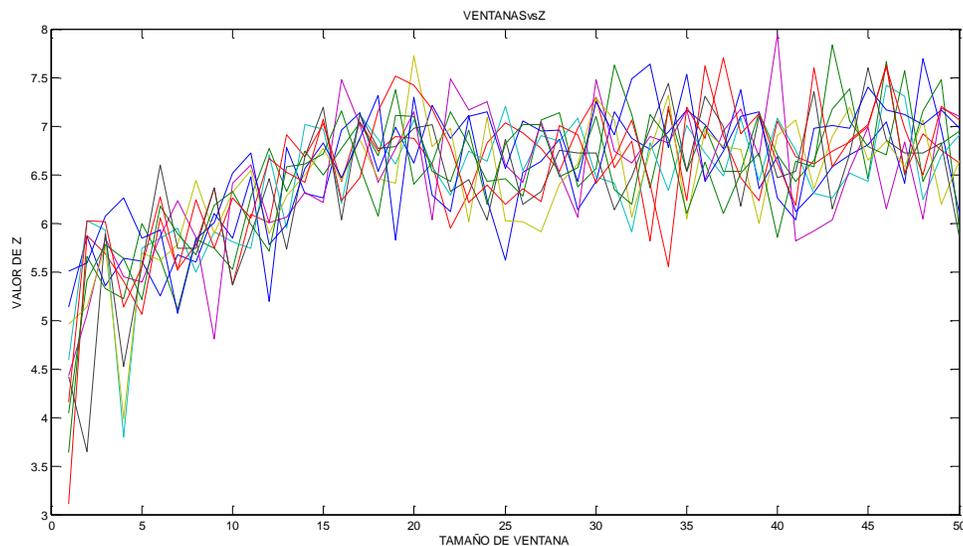


Figura 5.19 Pruebas con ventanas de distinto tamaño.

Del la figura 5.19 notamos que la hipótesis se aprueba en todos los tamaños de ventana; en estos casos mantuvimos constante el número de elementos de entrenamiento (40 matrices reales) realizando 10 pruebas sobre cada ventana. Podemos observar que existe una pendiente muy marcada de la ventana de tamaño 1 hasta la 14, y de acuerdo a nuestras observaciones por el análisis wavelet el sistema considera de igual manera que la ventana de 15 días contiene mejores características como precursor; de la figura anterior las distancias más grandes comienzan en la ventana de 15 días, esto se interpreta como una mayor diferencia entre los conjuntos en comparación.

## CAPÍTULO 5 APLICACIONES Y RESULTADOS

Una prueba más que realizamos es el número óptimo de matrices de entrenamiento; para éste caso mantenemos constante el tamaño en la ventana de 15 días y entrenamos la base con subconjuntos aleatorios de 10 hasta 180 elementos (con un intervalo de 10) contenidos en el conjunto  $B$ ; cada prueba se realizó 10 veces bajo las mismas condiciones y los resultados se muestran en la figura 5.20.

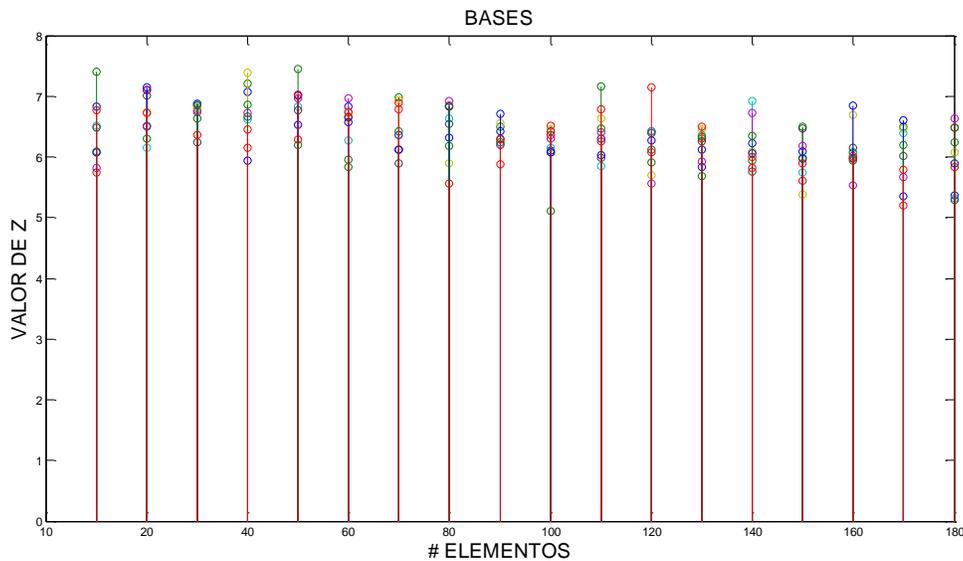


Figura 5.20 Pruebas con números de elementos de entrenamiento distintos.

Al analizar la figura 5.20 apreciamos que el número de elementos de entrenamiento no es un parámetro que influye de manera drástica en los resultados; parece ser que el sistema logra capturar los detalles de los espectros con una base de entrenamiento de 10 elementos o de 180. Este resultado refleja que las matrices de entrenamiento muestran una homogeneidad. Para acelerar el proceso de cómputo recomendamos una base de entrenamiento con 40 elementos.

Finalmente en el capítulo siguiente presentamos las conclusiones y recomendaciones de este trabajo de tesis.