



# DetECCIÓN Y CARACTERIZACIÓN DE COMUNIDADES DE MAGNETÓMETROS UTILIZANDO REDES COMPLEJAS.

Tesis  
entregada a la  
Universidad de Chile  
en cumplimiento parcial de los requisitos  
para optar al grado de  
Magíster en Ciencias con mención en Física  
Facultad de Ciencias

por

**Sebastián de la Maza Rodríguez**

Diciembre, 2022

Director de Tesis: **Dr. Víctor Muñoz Gálvez**

FACULTAD DE CIENCIAS  
UNIVERSIDAD DE CHILE

INFORME DE APROBACIÓN  
TESIS DE MAGÍSTER

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Magíster presentada por el candidato

**Sebastián de la Maza Rodríguez**

ha sido aprobada por la Comisión de Evaluación de la Tesis como requisito para optar al grado de Magíster en Ciencias con mención en Física, en el examen de Defensa de Tesis rendido el día 08/03/2023.

**Director de Tesis**

Dr. Víctor Muñoz Gálvez

\_\_\_\_\_

**Comisión de Evaluación de la Tesis**

Dra. Denisse Pastén Guzmán

\_\_\_\_\_

Dr. Pablo Moya Fuentes

\_\_\_\_\_

Dr. Alejandro Valdivia Hepp

\_\_\_\_\_

Dr. Rodrigo Soto Bertrán

\_\_\_\_\_



## BIOGRAFÍA

Flexible como un árbol en el viento,  
no tengo en vista ningún destino  
y utilizo todo aquello que la vida  
pone en mi camino.  
Como un viajero que no tiene planes fijos  
ni está preocupado por llegar.  
Como un artista que deja que su intuición  
le lleve donde quiera.

## AGRADECIMIENTOS

Gracias profe Víctor por haberme aceptado y ser mi profesor guía.

Gracias profe Carlos y profe Francisco por haberme ayudado a entrar al programa de Magíster.

...

Gracias!

### ■ Recursos materiales y financiamiento

El Departamento de Física de la Facultad de Ciencias de la Universidad de Chile posee los recursos computacionales necesarios para llevar a cabo este proyecto de tesis, acceso a bibliografía y licencias de código necesarias. Adicionalmente, el siguiente proyecto de investigación cubrirá parcialmente gastos de esta tesis: FONDECYT 1201967, “Time series and complex networks approaches for nonlinear processes in space plasmas”.

# Índice

<b>1. Introducción</b>	<b>1</b>
<b>2. Interacción Sol-Tierra</b>	<b>6</b>
2.1. Estructura del Sol . . . . .	7
2.2. Actividad solar . . . . .	7
2.2.1. Ciclo solar 23 . . . . .	8
2.3. Viento solar y campo magnético interplanetario (IMF) . . . . .	9
2.4. La magnetósfera de la Tierra . . . . .	11
<b>3. Campo magnético terrestre (CMT)</b>	<b>13</b>
3.1. Componentes del campo magnético terrestre . . . . .	15
3.2. Variaciones del campo magnético terrestre . . . . .	16
3.2.1. Variaciones del CMT producto de las tormentas solares . . . . .	17
3.3. Magnetómetros terrestres . . . . .	22

3.4. Mediciones satelitales del CMT . . . . .	23
3.4.1. Satélite Ørsted . . . . .	24
3.4.2. Satélite CHAMP . . . . .	24
<b>4. Redes Complejas</b>	<b>25</b>
4.1. Definición de red compleja . . . . .	27
4.2. Comunidades en redes complejas . . . . .	29
4.2.1. Método de detección de comunidades:	
Modularidad ( $M$ ) . . . . .	29
<b>5. Metodología</b>	<b>41</b>
5.1. Obtención de los datos de los magnetómetros . . . . .	43
5.2. Obtención de los datos satelitales . . . . .	45
5.3. Construcción de la red . . . . .	49
<b>6. Resultados</b>	<b>70</b>
6.1. Datos observados por los magnetómetros . . . . .	71
6.1.1. Número de comunidades y nodos . . . . .	72
6.1.2. Área de las comunidades . . . . .	80
6.1.3. Modularidad . . . . .	82
6.1.4. Entropía de Shannon . . . . .	83
6.2. Datos observados por los satélites . . . . .	85
6.2.1. Datos observados por el satélite CHAMP . . . . .	85
6.2.2. Datos observados por el satélite Ørsted . . . . .	95

<b>7. Conclusión</b>	<b>104</b>
<b>A. Apéndice</b>	<b>111</b>
<b>B. Apéndice</b>	<b>115</b>
<b>C. Apéndice</b>	<b>118</b>
C.1. Resultados con el índice Ap . . . . .	119
C.1.1. Número de comunidades . . . . .	119
C.1.2. Área de las comunidades . . . . .	123
C.1.3. Modularidad . . . . .	124
C.1.4. Entropía de Shannon . . . . .	125
C.1.5. Coeficiente de clustering . . . . .	126
C.2. Resultados con el índice AE . . . . .	127
C.2.1. Número de comunidades . . . . .	127
C.2.2. Área de las comunidades . . . . .	131
C.2.3. Modularidad . . . . .	132
C.2.4. Entropía de Shannon . . . . .	133
C.2.5. Coeficiente de clustering . . . . .	134



## RESUMEN

En esta tesis estudiamos las variaciones de la intensidad del campo magnético terrestre (CMT) utilizando técnicas de redes complejas. Específicamente se realizó un estudio de comunidades sobre una red compleja a partir de la serie de tiempo de campo magnético registrada por 59 magnetómetros distribuidos sobre la superficie terrestre, a lo largo del ciclo solar 23, donde los nodos representan puntos, sobre la superficie terrestre, y las conexiones representan el grado de similitud entre las series temporales observadas en dichos puntos.

En efecto, ha sido posible abstraer las variaciones espaciotemporales del CMT en lenguaje de redes complejas. Nuestros resultados muestran que existe una correlación entre la evolución de la estructura de comunidades y la actividad geomagnética. Además, estudiamos la dependencia de los resultados de los métodos usados para definir la similitud entre series de tiempo (y, por ende, para definir la conexión entre nodos), intentando establecer qué método es mejor para estudiar la variación de las comunidades. En este sentido encontramos que la elección de un método de similitud para este trabajo no muestra una diferencia significativa entre resultados. Pero la elección de un umbral de correlación si influye significativamente en la estructura de comunidades y su evolución con el ciclo solar. En particular, el umbral  $\theta = 0.8$  es el que presenta los mejores resultados, es decir, la variación de las métricas utilizando este umbral se correlacionan de mejor manera con la variación del índice *Dst*.

Nuestro trabajo sugiere que el análisis de las variaciones de la intensidad del CMT a través de estrategias de redes complejas puede ser útil para la comprensión de la dinámica del CMT a lo largo de un ciclo solar.

# Capítulo 1

## Introducción

En el tratamiento moderno de la mecánica estadística, la idea de redes complejas está recibiendo una atención significativa, surgiendo como un enfoque novedoso de la geociencia para obtener nuevos conocimientos sobre la naturaleza de los sistemas geofísicos.

Definimos una red compleja como un conjunto de nodos y enlaces, que representan relaciones entre ellos. Por ejemplo, en una red social, cada nodo puede representar a un individuo, y un enlace entre ellos puede indicar que ambas personas se conocen entre sí. Una red similar puede construirse a partir de los aeropuertos (nodos) y los diferentes vuelos que los conectan (enlaces). La elección sobre el significado de nodos y enlaces dependerá del sistema a estudiar, y de las propiedades que se desean estudiar del mismo.

En investigaciones basadas en datos sísmicos, Abe y Suzuki [1] construyeron redes crecientes definiendo una conexión entre dos terremotos sucesivos y encontraron que estas redes de terremotos eran libres de escala y de mundo pequeño. Suteanu [2] propuso un método basado en redes complejas para la evaluación de los terremotos

en cuanto a sus distribuciones de espacio, tiempo y magnitud y aplicó los resultados para el estudio de las variaciones temporales en los patrones de sismicidad volcánica.

Pastén et al. [3] estudiaron la universalidad de los exponentes críticos en redes complejas a partir de redes construidas por dos conjuntos de datos sísmicos (zona norte y zona central de Chile sin grandes terremotos).

Otra aplicación geofísica de redes complejas es la ciencia del clima. Peron et al. [4] construyeron una red de temperatura por correlación y consideraron los puntos de la cuadrícula global sobre América del Norte como nodos. Realizaron una detección de comunidades y se dieron cuenta de que en el lado este se forman varias comunidades y en el lado oeste solo una. Agarwal et al. [5] hicieron un estudio de comunidades en una red compleja de estaciones meteorológicas construida a partir de los datos de precipitaciones diarias de la India, para estudiar, a partir del número de conexiones inter e intracomunitarias, el papel de cada estación (es decir, de cada punto de observación).

Además, de manera similar a los sismos y los fenómenos climáticos, el campo magnético terrestre (CMT) también se distribuye espacialmente. La variación del CMT implica interacciones y flujos espaciales, siendo el Sol uno de los principales responsables, el cual produce una variación periódica en el CMT debido a su ciclo magnético, a través del forzamiento inducido por el viento solar.

Proponemos un estudio, a través de redes complejas, para estudiar las variaciones espaciales y temporales del CMT. Diversos estudios han hecho propuestas similares, mostrando que las redes complejas son una interesante herramienta para estudiar el campo magnético terrestre. Por ejemplo, Orr y Chapman, et al. [6] utilizaron técnicas

de la ciencia de redes para analizar 41 subtormentas usando datos de 137 magnetómetros ubicados en el hemisferio norte. Realizaron una detección de comunidades en la red y encontraron que el sistema pasa de una configuración de muchos pequeños grupos de magnetómetros no correlacionados antes del inicio de la subtormenta, a un gran sistema correlacionado espacialmente extendido durante la fase de expansión. Lu y Zhang, et al. [7] investigaron los procesos dinámicos en la ionósfera, construyendo una red compleja dirigida, basada en un gráfico probabilístico del contenido vertical total de electrones (VTEC) de 2012.

También se han utilizado redes complejas para estudiar el campo magnético solar. Por ejemplo, Najafi et al. [8] construyeron una red compleja de erupciones solares utilizando grafos de visibilidad junto con el método Abe y Suzuki [1]. Flández et al. [9] estudiaron la actividad solar, usando algoritmos de reconocimiento de imágenes para determinar la ubicación de las manchas solares, y a partir de ahí construir una red compleja, cuyas métricas se estudiaron en ventanas de tiempo a lo largo del ciclo solar 23, para así analizar su correlación con la actividad magnética solar.

Estas investigaciones son un excelente punto de partida para los estudios que intentan caracterizar, mediante redes complejas, la información colectiva de los datos de los magnetómetros terrestres, permitiendo contribuir y ampliar el conocimiento existente sobre la acción de los numerosos mecanismos físicos involucrados en la dinámica del CMT.

No sólo a nivel local, a escala de subtormentas, sino también a escala del planeta completo, el CMT presenta variaciones espaciotemporales, algunas de las cuales son intrínsecas, y otras externas, inducidas por el forzamiento del viento solar. En

particular, la actividad solar presenta un ciclo de 11 años aproximadamente, el cual impacta en la actividad geomagnética, aumentando intensidad y número de tormentas geomagnéticas. Estas variaciones en la actividad solar, a su vez, inducen variaciones espaciotemporales del campo magnético. Por tanto, basados en estudios previos de estructura de comunidades de redes complejas construidas a partir de mediciones de campo magnético [6, 7, 8], es interesante preguntarse si dicha estructura de comunidades presenta variaciones que pueden correlacionarse con el ciclo solar.

Un propósito adicional de este trabajo es estudiar la posible dependencia de los resultados (estructura de comunidades de la red compleja construida) con respecto al método de construcción de dicha red. En nuestro caso, nuestros datos corresponden a la serie de tiempo de campo magnético en distintos puntos del planeta. Para construir la red, seguiremos una estrategia similar a la desarrollada para construir redes complejas a partir de registros de lluvia, [5, 10] o registros de campo magnético [6], en donde los nodos corresponderán a los puntos donde se efectúan mediciones, y las conexiones están determinadas por el grado de similitud entre las series de tiempo respectivas. Sin embargo, esto precisamente abre el problema de cómo medir la similitud entre series de tiempo, lo cual determina las conexiones entre nodos, la topología de la red, y en definitiva su estructura de comunidades. Por ello, nos interesará también explorar cómo dependen los resultados obtenidos de estas elecciones.

Este trabajo está organizado de la siguiente manera. En el Cap. 2 revisaremos algunos conceptos generales respecto a la interacción entre el Sol y la Tierra. En el Cap. 3, mencionamos aspectos importantes sobre el campo magnético terrestre. En el Cap. 4 se da una introducción de la Teoría de Redes Complejas, definiendo algunos

conceptos esenciales para su descripción y caracterización. En el Cap. 5 se describe la metodología utilizada para la construcción de la red.

A continuación en el Cap. 6 se muestran los resultados obtenidos, caracterizando las comunidades de la red compleja a través de diversas métricas, y estudiando la evolución de dichas métricas a lo largo del ciclo solar 23. Para finalizar en el Cap. 7 se discuten los resultados y se derivan las conclusiones de este trabajo.

# Capítulo 2

## Interacción Sol-Tierra

La actividad magnética del Sol es un importante factor para entender diversos fenómenos en el sistema solar, debido a la interacción del material y la radiación que emite el Sol con los planetas, específicamente sus magnetósferas e ionósferas. En la Tierra, de hecho, el ecosistema existente es sumamente frágil, por lo que variaciones en la cantidad de radiación y partículas que recibe del Sol, pueden tener un efecto significativo en el medio ambiente [11].

El sistema Sol-Tierra puede considerarse como aquel constituido por cuatro componentes principales: el Sol, el medio interplanetario, la magnetósfera y la atmósfera terrestre [12]. A este sistema hay que agregar toda una serie de agentes externos como los rayos cósmicos (partículas de muy alta energía capaces de atravesar todo el medio interplanetario), el campo magnético terrestre, el medio interestelar a través del cual se mueve nuestro sistema solar, etc.

En general, la comprensión global del sistema constituido por la Tierra y sus alrededores en todas las escalas de tiempo es una cuestión sumamente compleja e interdisciplinaria. Cualquier cambio detectable en el sistema es, en última instancia,

resultado de la interacción, la retroalimentación o incluso la posible amplificación de muchos factores causantes, algunos de los cuales revisamos a continuación.

## 2.1. Estructura del Sol

El Sol es una gran esfera luminosa de gas, que emite distintos tipos de radiación: desde las de muy baja energía, como las ondas de radio, hasta las más penetrantes como los rayos gamma, así como partículas energéticas y plasma, este último en forma de haces (el viento solar) que llenan todo el Sistema Solar y se extienden hasta mucho más allá de sus límites, creando lo que se conoce como la heliósfera.

El Sol es una estrella enana de color amarillo, miembro de uno de los más numerosos tipos de estrellas, las del tipo espectral G2, que se mantiene unida por su propio campo gravitacional y presiones internas como la del plasma y la de radiación. Tiene un radio de 695980 km, es decir, 109 veces el radio de la Tierra, que tiene alrededor de 6371 km, y rota sobre su propio eje en aproximadamente 27 días. El interior está formado por tres capas: el núcleo, la zona radiativa y la zona convectiva, mientras que la atmósfera se divide en: fotosfera, cromósfera, zona de transición y corona solar [13].

## 2.2. Actividad solar

La actividad solar se refiere a variaciones temporales de largo y corto periodo que ocurren en la atmósfera solar a consecuencia del complicado patrón de campo magnético emergente del interior del Sol, el cual a su vez produce cambios en el plasma en el cual se encuentra inmersa la Tierra (viento solar), y por ende afecta al



campo geomagnético.

Una de las manifestaciones principales de la actividad solar son las manchas solares que presenta la superficie solar, al punto que dichas manchas se pueden usar para caracterizar la actividad magnética solar. En general, las manchas solares aparecen en ciclos de aproximadamente 11 años, en los que se registran periodos de gran actividad, con más manchas solares en comparación con otros periodos de menor actividad donde algunas o ninguna mancha aparece en la superficie solar [13].

### 2.2.1. Ciclo solar 23

El ciclo que nos interesa en el presente trabajo es el ciclo solar 23, que tuvo una duración de 12.3 años, comenzó en agosto de 1996 y finalizó en diciembre de 2008, alcanzando el máximo solar durante el 2001. El mínimo solar de 2008, comprendido dentro de este ciclo 23, fue anómalo, ya que duró más que cualquier mínimo anterior en los últimos 100 años. Además, se observó que la magnitud del campo magnético interplanetario fue muy baja. Debido a estos cambios observados, la intensidad de los rayos cósmicos fue bastante alta [14]. El uso del ciclo solar 23 nos permite comparar este trabajo con otros estudios similares que se han realizado sobre la complejidad de la dinámica solar [9, 15].

A continuación la Fig. 2.1 nos muestra la variación del número de manchas solares a lo largo del ciclo solar 23, en donde se puede observar que los años de alta actividad solar son de 1999 al 2004, en donde los máximos ocurren en los años 2001 y 2003. Y los años de baja actividad solar son entre los años 1996 y 1998, y entre 2005 y 2008.



Figura 2.1: Número de manchas solares correspondiente al Ciclo Solar 23. Imagen extraída de la Ref. [16].

### 2.3. Viento solar y campo magnético interplanetario (IMF)

Debido al gran desbalance de presiones entre la corona solar y el borde de la heliósfera (la heliósfera es la cavidad creada por la interacción del viento solar y el gas interestelar), el plasma en la zona coronal es expulsado (“viento solar”) en dirección radial respecto al Sol.

Debido a que el plasma de viento solar tiene una conductividad eléctrica muy alta, el campo magnético del Sol se desplaza con el viento, siguiéndolo en su expansión por el medio interplanetario, como si se encontrara “congelado” a él. Así pues, el campo magnético que llega a la Tierra junto con el viento solar, al que se conoce como campo magnético interplanetario, no es más que la extensión del campo magnético del Sol. El campo magnético general del Sol es de tipo dipolar. El viento solar expande este campo radialmente hacia fuera del Sol. Podría esperarse que se estableciera un campo cuyas líneas fueran también radiales. Pero como el Sol gira y al girar traslada las

líneas de campo emergentes del mismo, la geometría que finalmente resulta para las líneas del campo magnético interplanetario es la de una espiral (“espiral de Parker”). Esto se muestra en la Fig. 2.2, donde las flechas radiales muestran la dirección de flujo del viento solar. De esta manera, el viento solar transporta parte de la energía del Sol hacia el espacio, donde pequeñas perturbaciones en el flujo del viento solar tienen consecuencias apreciables en el entorno terrestre. Este plasma, constituido por partículas cargadas eléctricamente envuelve nuestro planeta y perturba su campo magnético.

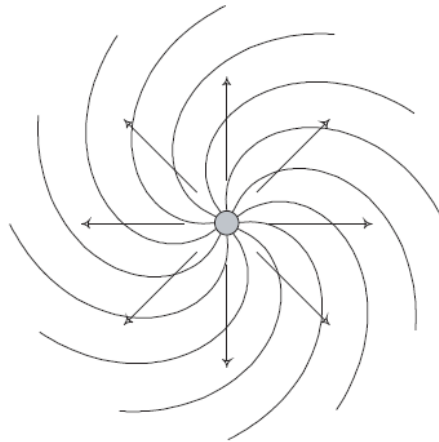


Figura 2.2: Configuración del campo magnético interplanetario en el plano del ecuador solar. Imagen extraída de la Ref. [17].

## 2.4. La magnetósfera de la Tierra

A su paso por el medio interplanetario ese plasma magnetizado que es el viento solar encuentra un obstáculo en el campo magnético de la Tierra, protegiendo a esta de su impacto directo. El viento solar no puede penetrar directamente este escudo, por lo que sufre una deflexión que deja una cavidad en forma de cometa (ver Fig. 2.3).

La cavidad del campo geomagnético o magnetosfera es una región de plasma que contiene partículas cargadas de un gran rango de energías, desde los cientos hasta los millones de electrón volts.

El campo geomagnético es empujado en la dirección del viento solar hasta que la presión de este iguale a la presión del campo geomagnético. La frontera donde estas dos presiones se balancean se denomina magnetopausa. La Fig. 2.3 ilustra las principales regiones de la magnetósfera.

Las regiones esquematizadas en la Fig. 2.3 son:

- Región Interplanetaria: Lugar donde las propiedades del medio no son afectadas por la Tierra y su campo magnético.
- La Magnetofunda: Es la región donde interactúan el viento solar con el campo geomagnético.
- Frente de choque: Separa la región interplanetaria de la magnetofunda.
- La Magnetopausa: El límite de la magnetósfera, separando el plasma unido a la Tierra del que viaja con el viento solar.

- Magnetósfera: El ambiente más cercano de la Tierra, dominado por el campo magnético de la Tierra.

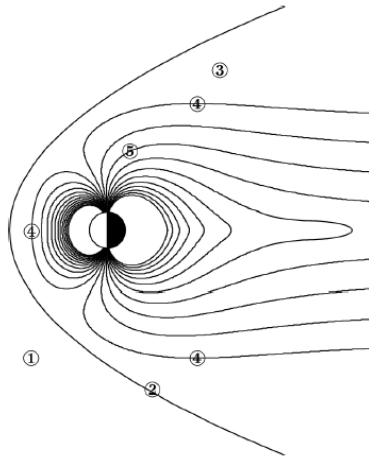


Figura 2.3: Sección transversal de la Magnetósfera terrestre. 1: Región interplanetaria, 2: Onda de Choque, 3: Magnetofunda, 4: Magnetopausa, 5: Magnetósfera. Imagen extraída de la Ref. [17].

La interacción del viento solar con la magnetósfera terrestre, se ve reflejada en las variaciones de la actividad geomagnética.

## Capítulo 3

# Campo magnético terrestre (CMT)

El campo magnético terrestre o campo geomagnético que se observa en un punto de la Tierra tiene dos contribuciones, una de origen interno y otra externa.

El campo de origen interno se explica por la llamada teoría del dínamo, según la cual el núcleo externo (de consistencia líquida) da vueltas lentamente (debido a la dinámica de fluido del hierro), con respecto al manto semisólido, generando de esta manera corrientes eléctricas que rodean al núcleo. Estas corrientes generan a su vez un campo magnético, parte del cual alcanza la superficie de la Tierra, dándonos el campo magnético que observamos, y otra parte interacciona con el núcleo líquido que está en movimiento, sosteniendo de esta manera la acción del dínamo. La configuración resultante es semejante a un dipolo magnético situado en el centro de la Tierra, como una barra magnética alineada aproximadamente sobre el eje de rotación [18] (ver Fig. 3.1).

Este campo no es constante ni uniforme, sino que presenta una variación lenta en el tiempo que se conoce como variación secular, entre otras [19, 20] (en la sección 3.2 se discutirán más en detalle las escalas de tiempo de las distintas variaciones del CMT).

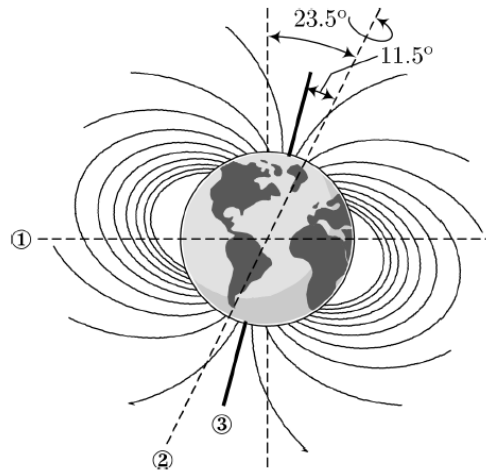


Figura 3.1: Campo geomagnético. En primera aproximación este puede ser considerado como el producido por un dipolo magnético. 1: Plano de la órbita de la Tierra, 2: Eje de rotación, 3: Eje magnético. Imagen extraída de la Ref. [17].

El campo de origen externo es debido principalmente a la interacción del viento solar con la ionósfera y la magnetósfera. Este presenta variaciones periódicas, siendo la más importante la variación diurna con un período de 24 horas [20].

El campo en la superficie tiene una magnitud que va desde unos 30000 nT cerca del ecuador hasta unos 60000 nT cerca de los polos [21].

### 3.1. Componentes del campo magnético terrestre

La magnitud del vector magnético  $\vec{B}$  en un punto  $P$  de la superficie terrestre está dada por la intensidad del campo  $F$  que puede proyectarse en dos ángulos, como se ilustra en la Fig. 3.2: la declinación magnética  $D$  o ángulo desde el norte geográfico hasta la proyección horizontal (positiva hacia el este) y la inclinación magnética  $I$  o ángulo desde la horizontal hasta la dirección del campo (positiva hacia abajo apuntando al centro de la Tierra). Estas tres magnitudes  $D$ ,  $I$  y  $F$  se denominan elementos magnéticos y pueden ser expresados en coordenadas esféricas. Y también se puede utilizar un sistema cartesiano, siendo las componentes  $X$ ,  $Y$  y  $Z$  correspondientes a las direcciones norte, este y nadir (en la dirección vertical con el sentido hacia abajo apuntando al centro de la Tierra) respectivamente [22] (ver Fig. 3.2).

La relación entre ambos sistemas es la siguiente:

$$X = F \cdot \cos(I) \cdot \cos(D), \quad (3.1)$$

$$Y = F \cdot \cos(I) \cdot \sin(D), \quad (3.2)$$

$$Z = F \cdot \sin(I). \quad (3.3)$$



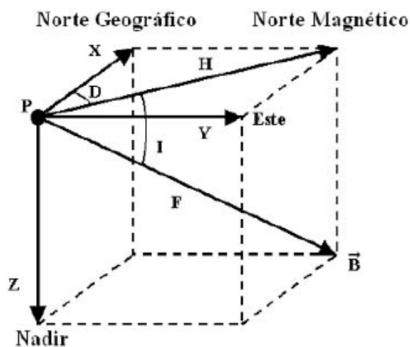


Figura 3.2: Relación entre las componentes cartesianas del campo magnético ( $X$ ,  $Y$ ,  $Z$ ) y los elementos magnéticos  $D$  (declinación),  $I$  (inclinación) y  $F$  (intensidad). El vector  $\vec{B}$  representa el campo magnético y  $H$  es su proyección sobre el plano horizontal. Imagen extraída de la Ref. [23].

### 3.2. Variaciones del campo magnético terrestre

El campo magnético terrestre no es estático, sino que varía con el tiempo en diferentes escalas que van desde los milisegundos hasta las decenas de millones de años. Las variaciones asociadas usualmente a los periodos más cortos están asociadas a causas de origen externo (corrientes en la ionósfera y la magnetósfera, tormentas geomagnéticas). Los campos asociados a estas fuentes son bastante más débiles que el campo interno. Con la excepción del ciclo solar, de aproximadamente 11 años, los periodos por encima de un año suelen estar asociados al campo principal, ya que las variaciones de origen interno que tuviesen periodos más cortos serían atenuadas en el manto inferior debido a su conductividad y no alcanzarían la superficie terrestre, ver por ejemplo, Merrill et al. [19].

Se pueden identificar las siguientes variaciones [24]:

- Variaciones seculares: Son variaciones que se extienden durante varios años. Se originan por cambios en el interior de la Tierra y son generalmente incrementos y decrementos suaves y lentos. Los valores de la variación secular de las componentes del campo van desde 10 nT a 150 nT por año.
- Variaciones periódicas: Se originan debido a la influencia de campos externos y en general equivalen a menos de 100 nT. Estas están relacionadas con la rotación de la Tierra y la influencia del Sol y la Luna.
- Tormentas magnéticas: Son perturbaciones súbitas del campo magnético, las cuales pueden durar horas o varios días y modifican el campo en más de 500 nT.

Dado que las tormentas solares son una de las principales causas de la variación del CMT y que el objetivo de esta tesis es explorar las variaciones de la intensidad del CMT considerando datos correspondientes al ciclo solar 23, hablaremos ahora sobre cómo inciden las tormentas solares en la variación del CMT.

### **3.2.1. Variaciones del CMT producto de las tormentas solares**

Las tormentas magnéticas son periodos de variación magnética intensa, usualmente debido a la actividad solar, donde las variaciones pueden ser del orden de cientos de nanoteslas, con duraciones de hasta 24 horas. El efecto puede ser mayormente observado en zonas alrededor de los polos magnéticos.

Son producto del aumento en la intensidad de la corriente de anillo que se traduce en la aparición de fuertes perturbaciones en el campo magnético de la Tierra,

conduciendo a una disminución de la intensidad de su componente horizontal, o sea, la corriente de anillo alrededor de la Tierra produce un campo magnético que está directamente opuesto al campo magnético de la Tierra.

Una tormenta típica consta de tres etapas que muestran efectos de compresión y expansión sobre la magnetósfera, aunque pueden existir tormentas que no presenten alguna de estas partes [20]. En la Fig. 3.3 se muestra un registro obtenido en Italia, de la variación de las componentes  $H$ ,  $D$  y  $Z$  de una tormenta sucedida el 15 de julio del año 2000 durante el máximo solar del ciclo solar 23. Esta es una de las tormentas más intensas que se han registrado, la cual se conoce como la tormenta solar del día de la Bastilla. Involucró una llamarada solar, un evento de partículas solares y una eyección de masa coronal que causó una severa tormenta geomagnética.

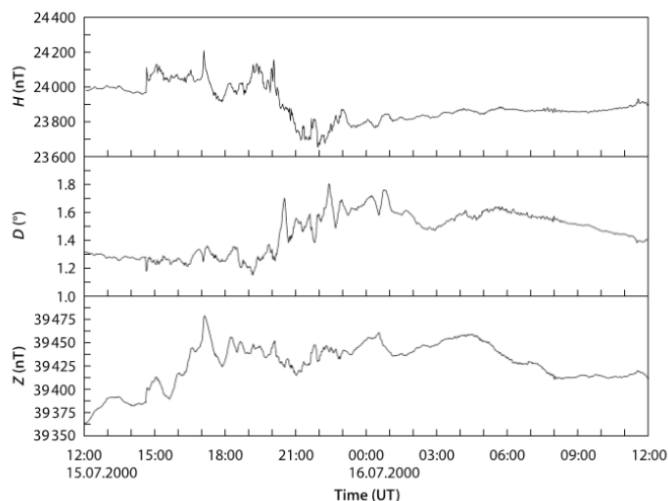


Figura 3.3: Tormenta geomagnética en los elementos magnéticos  $H$ ,  $D$  y  $Z$  (evento día de la Bastilla) del 15 de julio de 2000 en L'Aquila Italia. Imagen extraída de la Ref. [20].

Generalmente, pero no siempre, la tormenta tiene un comienzo repentino (*sudden storm commencement*, SSC) en la intensidad de la componente horizontal  $H$  del campo geomagnético. Inmediatamente después tenemos la fase inicial, asociada con un incremento en la componente horizontal que puede ser explicada por la compresión de la magnetósfera que produce el aumento de la presión del viento solar [25].

En la llamada fase principal se produce un aumento en la corriente de anillo, generada por el aumento en la concentración de partículas cargadas en el campo magnético terrestre y consecuentemente un campo magnético que se opone al terrestre, manifestándose como una caída brusca en la intensidad de la componente horizontal.

La fase final es la fase de recuperación, consistente en un incremento del campo hacia el nivel anterior al comienzo de la tormenta, que resulta de un decrecimiento de la concentración de partículas en la corriente de anillo [20].

Para cuantificar el efecto de la depresión en  $H$ , se calcula un índice horario  $Dst$  (*Disturbance Storm Time Index*) con las estaciones geomagnéticas cercanas al Ecuador, pero lo suficientemente lejos de los electrojets ecuatoriales. El índice  $Dst$  es, entonces, una manera de cuantificar la actividad magnética, a través de un promedio horario del campo magnético en el ecuador [20, 26, 27].

El índice se calcula generalmente a partir de los registros proporcionados por cuatro estaciones. Sus ubicaciones, como se mencionó, están alejadas de los electrojets y están distribuidas lo más uniformemente posible en longitud (ver Fig. 3.4). Los nombres de las estaciones son: Hermanus (HER) en el hemisferio sur, y Honolulu (HON), Kakioka (KAK) y San Juan (SJG) en el hemisferio norte [27].

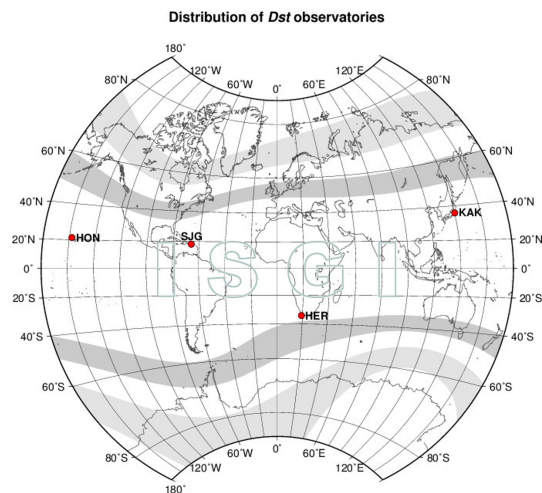


Figura 3.4: Ubicación de las cuatro estaciones geomagnéticas cercanas al Ecuador. Imagen extraída de la Ref. [27].

El cálculo del índice está determinado por la ecuación (3.4), donde  $\Delta H_i$  representa las variaciones de la componente horizontal (sin las contribuciones de la variación secular y la variación diurna) y  $\theta_i$  es la latitud geomagnética de las estaciones:

$$Dst(T) = \frac{1}{4} \frac{\sum_{i=1}^4 \Delta H_i}{\sum_{i=1}^4 \cos(\theta_i)}. \quad (3.4)$$

En la Fig. 3.5 se puede observar la variación del índice  $Dst$  entre los años 1996 y 2008.

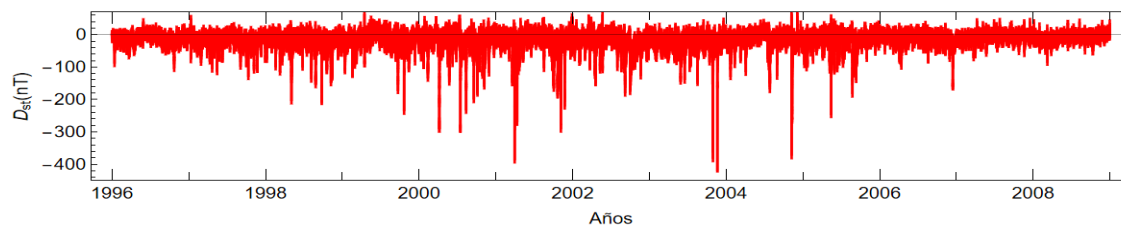


Figura 3.5: Variación del índice  $Dst$  correspondiente al Ciclo solar 23. Imagen adaptada de la Ref. [15].

En el presente trabajo nos enfocamos en las variaciones del CMT en escalas de tiempo de 1 año, que es el ancho de ventana escogido para la construcción de las redes complejas, como veremos en el Cap. 5. Para tener una estimación del tipo de variación que podemos tener en esta escala de tiempo, la Fig. 3.6 muestra el número de días con una tormenta geomagnética por año. Esperamos, entonces, en esta escala de tiempo, una variación suave, con un máximo en torno al período de máxima actividad solar.

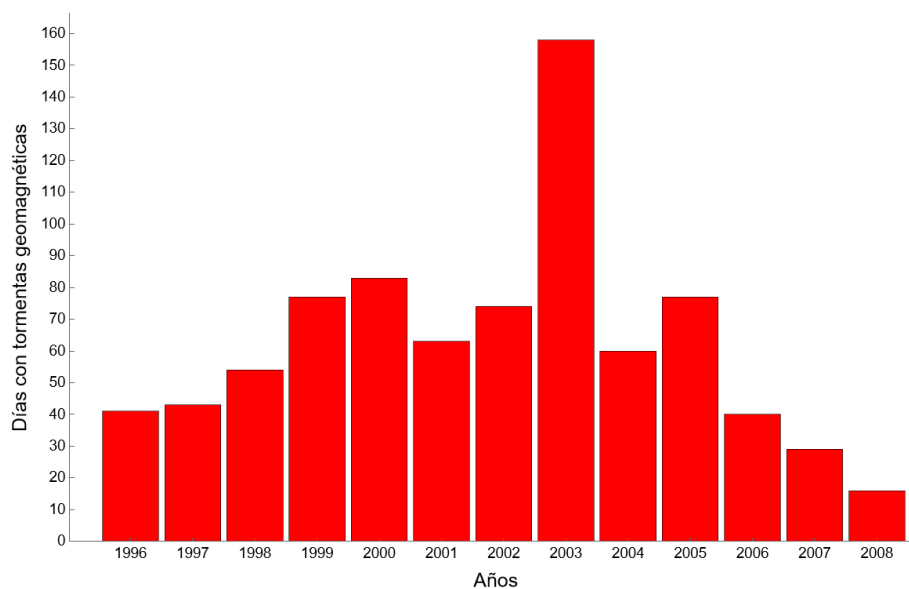


Figura 3.6: Número de días con una tormenta geomagnética por año. Imagen adaptada de la Ref. [16].

Además del índice  $Dst$ , existen otros índices para cuantificar la actividad geomagnética, como por ejemplo el índice  $Kp$ , el índice  $Ap$  y el índice  $AE$  [26, 27] (para más detalle revisar Apéndice A), que registran preferentemente actividad magnética cerca de los polos, o el índice  $SYM-H$ , que registra información análoga a  $Dst$ , pero con mayor resolución (1 minuto). Dado que nos interesan variaciones en escalas rela-

tivamente largas, de 1 año, el índice *Dst* resulta ser razonablemente adecuado para ilustrar las variaciones del campo magnético terrestre a lo largo del ciclo solar.

### 3.3. Magnetómetros terrestres

En pocas palabras, un magnetómetro es un instrumento que se utiliza para medir la magnitud y la dirección de un campo magnético [28]. Muchos satélites han sido equipados con magnetómetros in situ para el estudio y la observación del campo cambiante en el espacio. Pero antes de la era espacial, en los primeros años de la investigación magnetosférica, el principal instrumento de estudio era el magnetómetro terrestre [29]. Incluso hoy en día, una gran red de magnetómetros de estaciones terrestres recopila datos en todo el mundo, creando una base de datos bien distribuida en términos de resolución temporal y espacial.

Los magnetómetros tienen como objetivo registrar de forma continua y precisa los valores de las componentes del campo magnético terrestre y del campo total. Registran de forma continua y automática los elementos del campo geomagnético, requieren de una instalación en entornos controlados térmicamente y sobre una plataforma completamente estable. Uno de los magnetómetros más comunes para el uso en exploración geofísica es el magnetómetro de tipo fluxgate (magnetómetro de puerta de flujo), que es un magnetómetro vectorial, lo que significa que mide no solo la magnitud sino también la dirección del campo magnético de la Tierra. La saturación de materiales magnéticos es el principio sobre el que funcionan los magnetómetros fluxgate. El dispositivo consta de un núcleo ferromagnético rodeado por dos bobinas de alambre. La configuración es similar a la de un transformador. La corriente al-

terna pasa a través de una bobina, que se llama primaria. Esto produce un campo magnético alterno que genera una corriente eléctrica en la segunda bobina. Es posible medir la fase y la intensidad de la corriente alterna en la bobina secundaria. La exposición a variaciones en el campo magnético de la Tierra provoca variaciones en los campos magnéticos del núcleo, lo que conduce a cambios en la corriente que se mide en la bobina secundaria. El cambio también se puede medir para analizar la intensidad y orientación de las líneas de flujo. Este tipo de magnetómetros también se utiliza en sondas espaciales para analizar los campos magnéticos de los planetas, incluida la Tierra [30].

Los datos obtenidos en los magnetómetros se utilizan para determinar los índices de actividad geomagnética, los valores medios horarios, diarios, mensuales y anuales, así como la variación anual de los elementos geomagnéticos y confeccionar con ellos los Anuarios Geomagnéticos (para más detalles revisar [26, 31]).

### **3.4. Mediciones satelitales del CMT**

Como se menciona en la sección anterior muchos satélites han sido equipados con magnetómetros in situ para el estudio y la observación del CMT. Los magnetómetros se encuentran entre los instrumentos científicos más utilizados en satélites de exploración y observación. En este trabajo, dos satélites particulares serán de nuestro interés, por razones que explicaremos a continuación.



### 3.4.1. Satélite Ørsted

Ørsted es un satélite de investigación geomagnética de Dinamarca. La nave espacial fue lanzada en el año 1999, su órbita tiene una altura de 700 km. Los principales objetivos científicos de la nave espacial eran realizar mediciones muy precisas y sensibles del campo geomagnético y realizar un seguimiento global del entorno de partículas cargadas de alta energía en el entorno de la Tierra. Los datos se utilizan para mejorar los modelos geomagnéticos, para estudiar los fenómenos aurorales y para la correlación con las mediciones basadas en la Tierra para estudiar la relación entre el campo externo y el acoplamiento de energía del sistema viento solar-magnetósfera-ionósfera [32].

Después de más de veinte años en órbita, el satélite Ørsted todavía está en funcionamiento y continúa transmitiendo mediciones precisas del campo magnético de la Tierra. Para más detalles sobre los datos del satélite Ørsted, revisar [33].

### 3.4.2. Satélite CHAMP

Challenging Minisatellite Payload (CHAMP) fue un satélite alemán lanzado en el año 2000 y se utilizó para la investigación atmosférica e ionosférica, así como para otras aplicaciones geocientíficas, tenía una órbita de altura de 454 km. CHAMP permitió por primera vez generar simultáneamente mediciones de gravedad y campo magnético de alta precisión. Debido a su larga vida útil, no solo se pudo determinar la distribución espacial de estos campos, sino también su variabilidad temporal [34].

El final de la misión de CHAMP fue el año 2010, después de diez años. Para más detalles sobre los datos del satélite CHAMP, revisar [34].

# Capítulo 4

## Redes Complejas

Las redes complejas se han convertido en una herramienta muy útil para representar, estudiar y modelar fenómenos reales, permitiendo comprender de una manera novedosa las estructuras de datos y el análisis de algoritmos. La estructura de una red está íntimamente ligada a la dinámica del sistema que modela, por tanto, se puede extraer información valiosísima si se analiza adecuadamente.

Estudiar una red compleja implica determinar ciertas métricas, que permiten caracterizarla, revelar sus propiedades estructurales generales, o incluso comprender sus mecanismos de evolución [35, 36]. Cantidades tales como el grado, coeficiente de agrupamiento, distancia promedio entre nodos, o diversas medidas de centralidad se pueden calcular para cada nodo [10], estableciendo así propiedades locales que, promediadas sobre toda la red, permiten comprender diversos aspectos de la estructura de la red.

Sin embargo, hay propiedades de las redes que tienen que ver específicamente con su estructura mesoscópica. Por ejemplo, la red puede estar organizada en grupos de nodos, con una alta densidad de enlaces entre nodos del mismo grupo y una densidad

comparativamente baja de enlaces entre nodos de diferentes grupos. A estos grupos los llamamos comunidades o módulos [37, 38] que son el objetivo de este estudio: la estructura de comunidades.

Explorar las comunidades de la red es importante, ya que puede entregar una nueva perspectiva a la comprensión de la organización de la red, lo que puede ayudar a formular modelos realistas para la evolución de un sistema físico, y descubrir relaciones no triviales entre los nodos y entre subconjuntos de ellos, y que normalmente se pueden atribuir a la dinámica del sistema [39].

Como en este trabajo consideramos redes que cambian en el tiempo, es importante destacar que los nodos de las comunidades pueden ir cambiando y/o eliminando sus relaciones con los otros nodos de las comunidades, modificando la estructura de la red y por lo tanto, variando las estructuras de las comunidades. Esto nos permite responder preguntas como cuáles son las comunidades que perduran en el tiempo y cuáles no, estudiar factores internos o externos que llevan a la desaparición de una comunidad, quiénes son los nodos que conforman el core (centro) de una comunidad, y preguntas globales respecto a la evolución del sistema como un todo, etc.

Los estudios de la estructura de comunidades de redes complejas se han utilizado para estudiar distintos fenómenos, como por ejemplo sociales, biológicos [40] y geofísicos, tales como estudio de problemas en meteorología y variables climáticas [5, 41], y también para el caso de registros de subtormentas geomagnéticas en el hemisferio norte [6].

## 4.1. Definición de red compleja

Una red es una colección de puntos llamados vértices (o nodos) conectados por líneas llamadas aristas (o bordes) [42] (ver Fig. 4.1). Físicamente, los nodos pueden representar entidades que constituyen un determinado sistema (personas en una sociedad, moléculas en un sólido, servidores en Internet, etc.), y las conexiones pueden representar vínculos o interacciones entre ellas (relaciones de amistad, interacciones eléctricas, flujo de datos, etc.).

Si las aristas de una red tienen una dirección asignada, se dice que la red es dirigida. Si las aristas no tienen dirección, la red es no dirigida.

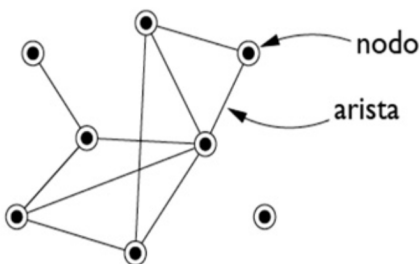


Figura 4.1: Ejemplo de red. Imagen extraída de la Ref. [43].

Utilizando una notación formal, una red  $R = (V, L)$  está constituida por un conjunto de nodos  $V = \{v_1, v_2, \dots, v_n\}$ , con  $n$  el total de nodos y un conjunto de aristas  $L = \{l_1, l_2, \dots, l_m\}$ , con  $m$  el total de aristas.

Una red se puede representar a través de una matriz de adyacencia  $\mathbf{A}$  con elementos  $A_{ij}$  tal que:  $A_{ij} = 1$  si los nodos  $v_i$  y  $v_j$  están conectados, y  $A_{ij} = 0$  si los nodos  $v_i$  y  $v_j$  no están conectados. Si la red es dirigida entonces su matriz de adyacencia  $\mathbf{A}$  no es simétrica, si la red es no dirigida entonces su matriz  $\mathbf{A}$  es simétrica.

En este trabajo, consideraremos redes no dirigidas, de modo que las aristas no tienen orientación, y si un nodo  $v_i$  está conectado a un nodo  $v_j$ , también el nodo  $v_j$  está conectado al nodo  $v_i$ . Este tipo de red se puede representar mediante una matriz de adyacencia  $\mathbf{A}$  simétrica.

Como ejemplo consideremos una red como la que se muestra en la Fig. 4.2.

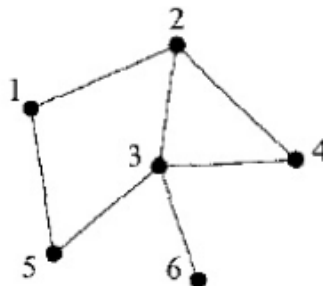


Figura 4.2: Ejemplo de red no dirigida. Imagen extraída de la Ref. [42].

Entonces nuestra matriz de adyacencia  $\mathbf{A}$  será de la forma:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \quad (4.1)$$

Notar que los elementos de la diagonal son 0 ya que no existen autoconexiones.

Si a cada arista de la red se le asigna un peso o valor numérico que mide la intensidad de la unión, entonces la red se llama red ponderada.

## 4.2. Comunidades en redes complejas

La estructura de comunidades, como se ha comentado anteriormente, es una propiedad de las redes complejas [40]. Una comunidad puede ser definida como un conjunto de nodos que están más densamente conectados entre ellos que con el resto de la red (ver Fig. 4.3). La importancia de este planteamiento radica en que se espera que los nodos que están contenidos dentro de una misma comunidad compartan atributos, características comunes o relaciones funcionales (ampliamente detallado en [37]).

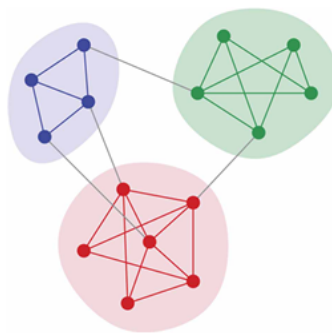


Figura 4.3: Ejemplo de red con estructura de comunidad. Los nodos de esta red están divididos en tres comunidades, donde la mayoría de las conexiones son intra-comunidad y sólo unas pocas entre comunidades. Imagen extraída de la Ref. [44].

### 4.2.1. Método de detección de comunidades: Modularidad ( $M$ )

Se ha encontrado que la elección del algoritmo de detección de comunidades tiene un pequeño impacto en las comunidades resultantes en los estudios de ciencia de datos geofísicos [45].

Utilizamos una detección basada en la modularidad, pues es una función que refleja qué tan buena es la partición de una red: mientras mayor sea el valor de  $M$  mejor es la partición. De este modo, el problema de la búsqueda de la estructura de comunidades se convierte en una estrategia de maximización de esta métrica.

La modularidad se define como un valor proporcional al número de aristas que caen dentro de una misma comunidad menos el número esperado en una red equivalente con aristas colocadas al azar. Los valores que se acercan al valor máximo,  $M = 1$ , indican una estructura de comunidad fuerte, es decir, valores altos de modularidad implican un mejor valor de asortatividad entre los nodos de una misma comunidad. Las particiones con valores  $M$  muy cercanos a cero, puede insinuar la existencia de comunidades con muy pocas aristas internas y muchas aristas entre comunidades (estructura multipartita). En la práctica, los valores de modularidad superiores a 0.3 sugieren la presencia de estructuras modulares [46, 47].

La modularidad ( $M$ ) se calcula como:

$$M = \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(C_i, C_j) , \quad (4.2)$$

donde  $A_{ij}$  son los elementos de la matriz de adyacencia  $\mathbf{A}$  de la red. Así,  $k_i = \sum_j A_{ij}$  es el grado del nodo  $v_i$ , y  $\frac{k_i k_j}{2m}$  es el valor esperado de aristas entre los nodos  $v_i$  y  $v_j$  si la red fuese aleatoria. De este modo, el término entre paréntesis es grande si dos nodos están mucho más conectados entre sí de lo que se espera para una red aleatoria. Además,  $C_i$  es la comunidad a la que se asigna el vértice  $v_i$ , y la función  $\delta(C_i, C_j)$  es 1 si  $C_i = C_j$  y 0 en caso contrario. Esto quiere decir que sólo aportan a la suma los nodos que pertenecen a la misma comunidad.

Entonces la modularidad  $M$  está definida de modo que es grande si los nodos dentro de una misma comunidad están muy conectados.

En este trabajo, emplearemos el método de Louvain, que maximiza la modularidad ( $M$ ) para encontrar la estructura comunitaria óptima en la red, es decir, particiona la red en grupos de nodos que no se superponen, maximizando el número de conexiones dentro de cada grupo y minimizando el número de conexiones entre grupos [48, 49]. Escogemos este método por sobre otros como por ejemplo el método de Clauset [50] o el método de Newman y Girvan [40, 51] por su alta velocidad de ejecución y por la calidad de los resultados que proporciona [48].

### **Algoritmo de Louvain**

El algoritmo de Louvain es un algoritmo eficiente de agrupamiento jerárquico para redes. Consiste en hacer que la modularidad de una partición en comunidades alcance el valor máximo a través de un proceso iterativo, en que los nodos se van asignando a diferentes comunidades en cada iteración [52].

Los pasos principales del algoritmo de Louvain son los siguientes:

- Paso 1: Configure cada nodo como una comunidad propia, es decir, la comunidad 1 está integrada sólo por el nodo 1, la comunidad 2 sólo por el nodo 2, y así sucesivamente.
- Paso 2: Encuentre todas las comunidades conectadas al nodo  $i$  y calcule el cambio de modularidad después de mover este nodo a cada comunidad vecina. Mueva el nodo  $i$  a la comunidad que maximice la modularidad.



- Paso 3: Iterar sobre todos los nodos y ejecutar el paso 2 hasta que no haya nodos para mover, obteniéndose así una primera capa de partición en comunidades.
- Paso 4: Combine cada comunidad en el paso 3 en un nuevo nodo, creando una nueva red donde los nodos son las comunidades de la fase anterior. Todos los enlaces entre nodos de la misma comunidad ahora están representados por auto-bucles en el nuevo nodo de comunidad y los enlaces de múltiples nodos en la misma comunidad a un nodo en una comunidad diferente están representados por bordes ponderados entre comunidades. Una vez que se crea la nueva red, regrese al paso 1 para fusionar los nuevos nodos (comunidades de la primera partición) entre sí, y construir una segunda partición a partir de dicha fusión. Iterar este proceso hasta que ya no se pueda seguir fusionando comunidades.

### Ejemplo

Consideremos una red con una matriz de adyacencia  $\mathbf{A}$ , con 16 nodos y 28 aristas como se muestra en la Fig. 4.4. Para determinar las comunidades de esta red, se debe calcular la modularidad  $M$  según la ecuación (4.2), maximizándola por medio del algoritmo de Louvain.

Lo primero es considerar todos los nodos como una comunidad propia. Luego se selecciona un nodo arbitrario, por ejemplo el nodo 0 y se calcula el valor de  $M$  para todos los posibles candidatos que podrían formar una comunidad con este nodo. En este caso, las nuevas comunidades posibles son  $\{0, 2\}$ ,  $\{0, 3\}$ ,  $\{0, 4\}$  ó  $\{0, 5\}$ .

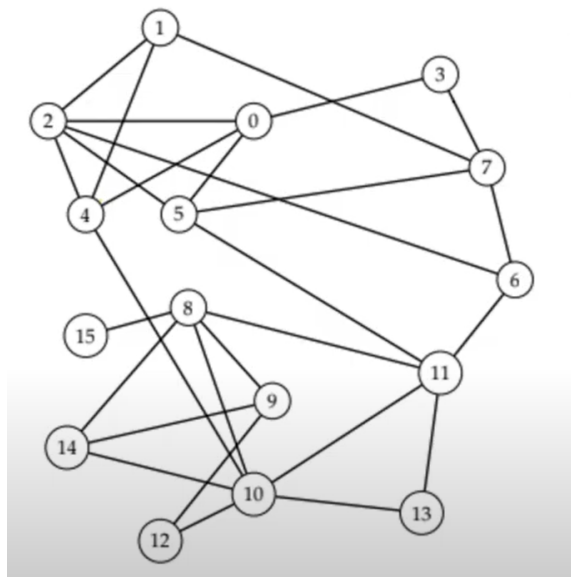


Figura 4.4: Red de ejemplo para explicar el algoritmo de Louvain. Considerar aquí cada nodo como una comunidad propia.

Entonces consideremos primero la posible comunidad  $\{0, 2\}$ . Para calcular el valor de  $M$  nos fijamos que la sumatoria de la ecuación (4.2) nos dice que tenemos que sumar para cada fila  $i$  de la matriz  $\mathbf{A}$  todos los elementos  $j$  de esa fila, pero el factor  $\delta(C_i, C_j)$  nos dice que solo hay que considerar los elementos  $A_{ij}$  que pertenezcan a la posible comunidad que se quiere formar ( $C_i = C_j$ ). En este caso serían los elementos  $A_{02}$  y  $A_{20}$ , por lo tanto podemos expandir la ecuación (4.2) de la siguiente forma:

$$M_{02} = \frac{1}{2m} \left( \left[ A_{02} - \frac{k_0 k_2}{2m} \right] + \left[ A_{20} - \frac{k_2 k_0}{2m} \right] \right) .$$

Considerando ahora que  $A_{02} = A_{20} = 1$ ,  $k_0 = 4$  y  $k_2 = 5$ , se obtiene:

$$M_{02} = \frac{1}{2m} \left[ 2 - \frac{4 \cdot 5}{m} \right] = 0.023 .$$

Ahora de la misma forma se determina el valor de  $M$  para los otros posibles candidatos del nodo 0, donde en este caso,  $k_3 = 2$ ,  $k_4 = 4$  y  $k_5 = 4$ , resultando:

$$M_{03} = \frac{1}{2m} \left[ 2 - \frac{4 \cdot 2}{m} \right] = 0.031 ,$$

$$M_{04} = \frac{1}{2m} \left[ 2 - \frac{4 \cdot 4}{m} \right] = 0.026 ,$$

$$M_{05} = \frac{1}{2m} \left[ 2 - \frac{4 \cdot 4}{m} \right] = 0.026 .$$

Observamos que el nodo 3 es la mejor opción del nodo 0 para formar una comunidad, entonces la comunidad seleccionada es  $\{0, 3\}$ , como se muestra en la Fig. 4.5.

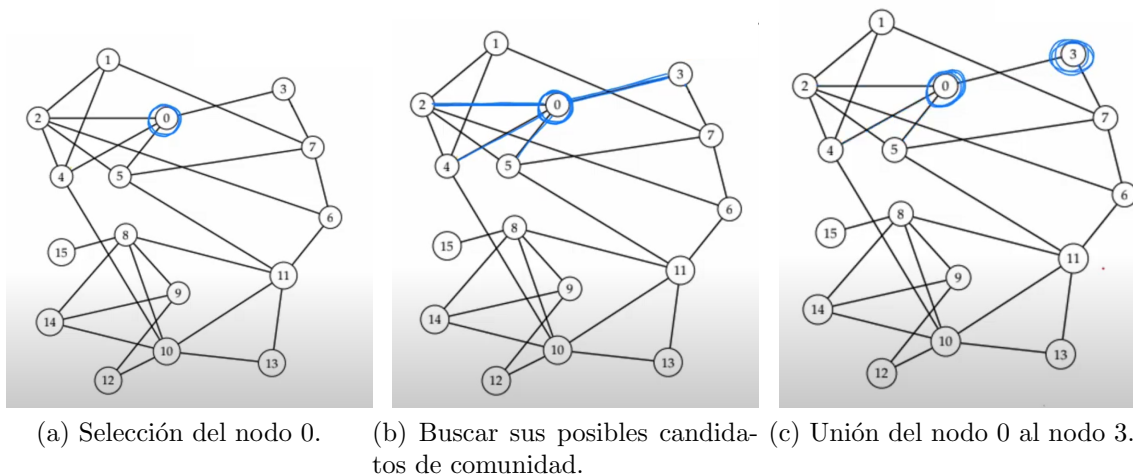


Figura 4.5: Ejemplo de cómo se forma la primera comunidad al partir seleccionando el nodo 0.

Ahora consideramos el nodo 1 que tiene tres comunidades potenciales  $\{1, 2\}$ ,  $\{1, 4\}$  y  $\{1, 7\}$ . Se repite el mismo procedimiento que se hizo para el nodo 0. En este caso, debemos usar que  $k_1 = 3$ ,  $k_2 = 5$ ,  $k_4 = 4$  y  $k_7 = 4$ .

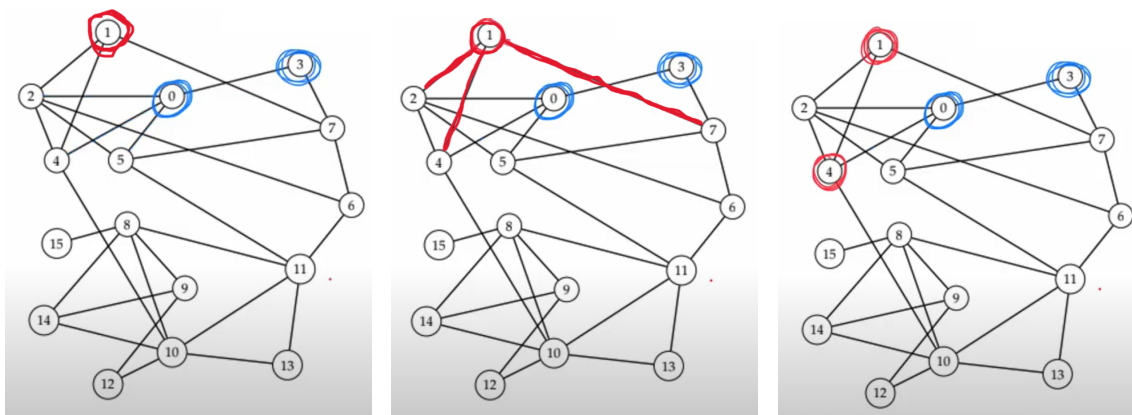
Entonces el valor de  $M$  para el nodo 1 y sus posibles comunidades es:

$$M_{12} = \frac{1}{2m} \left[ 2 - \frac{3 \cdot 5}{m} \right] = 0.026 ,$$

$$M_{14} = \frac{1}{2m} \left[ 2 - \frac{3 \cdot 4}{m} \right] = 0.028 ,$$

$$M_{17} = \frac{1}{2m} \left[ 2 - \frac{3 \cdot 4}{m} \right] = 0.028 .$$

Notar que  $M_{14} = M_{17} = 0.028$ , entonces aquí se puede seleccionar arbitrariamente una de las dos opciones. Seleccionaremos la comunidad  $\{1, 4\}$  (ver Fig. 4.6).



(a) Selección del nodo 1. (b) Buscar sus posibles candidatos de comunidad. (c) Unión del nodo 1 al nodo 4.

Figura 4.6: Ejemplo de cómo se forma la segunda comunidad seleccionando el nodo 1.

Ahora tomemos el nodo 2: Este nodo está conectado con los nodos 1, 0, 6, 5 y 4. Pero 1 y 4 son una comunidad, igual que 0 y 3. Entonces, en términos de comunidades, la comunidad  $\{2\}$  está conectada con 4 comunidades:  $\{0, 3\}$ ,  $\{1, 4\}$ ,  $\{5\}$  y  $\{6\}$ . Y, al calcular  $M$ , se deben considerar el grado de los nodos:  $k_2 = 5$ ,  $k_5 = 4$ ,  $k_6 = 3$ ,  $k_{\{1,4\}} = 7$  y  $k_{\{0,3\}} = 6$ . Similarmente con  $A_{ij}$ , que es el número total de aristas compartidas entre el nodo 2 y las comunidades con las que se conecta.

Entonces el valor de  $M$  para el nodo 2 y sus posibles comunidades es:

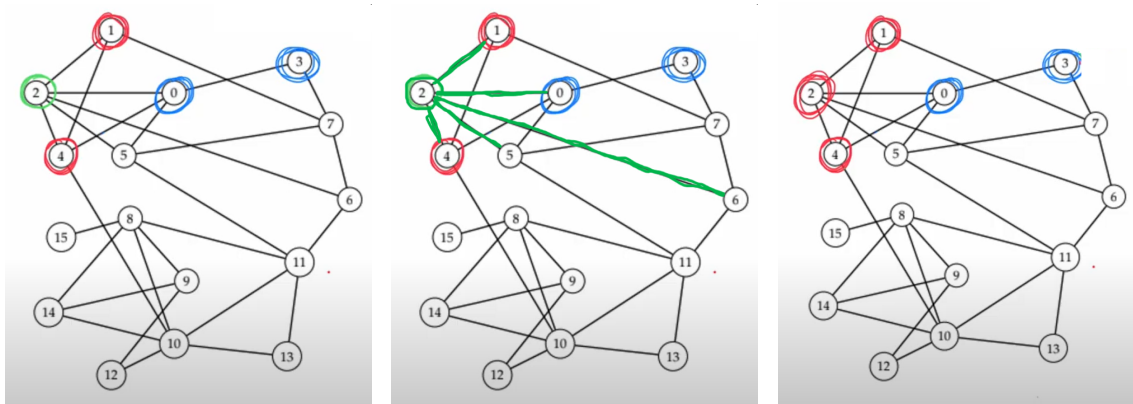
$$M_{25} = \frac{1}{2m} \left[ 2 - \frac{5 \cdot 4}{m} \right] = 0.023 ,$$

$$M_{26} = \frac{1}{2m} \left[ 2 - \frac{5 \cdot 3}{m} \right] = 0.026 ,$$

$$M_{2\{0,3\}} = \frac{1}{2m} \left[ 2 - \frac{5 \cdot 6}{m} \right] = 0.017 ,$$

$$M_{2\{1,4\}} = \frac{1}{2m} \left[ 4 - \frac{5 \cdot 7}{m} \right] = 0.049 .$$

Notamos que la comunidad  $\{1, 4\}$  es la mejor opción del nodo 2 para formar una comunidad (ver Fig. 4.7).



(a) Selección del nodo 2.

(b) Buscar sus posibles candida-  
tos de comunidad.

(c) Unión del nodo 2 a la comuni-  
dad  $\{1, 4\}$ .

Figura 4.7: Ejemplo de como se une el nodo 2 a la comunidad  $\{1, 4\}$ .

Ahora se repite este proceso desde el nodo 3 hasta el nodo 15, y así se finaliza la primera iteración. Al final de la primera iteración la red queda como se muestra en la Fig. 4.8.

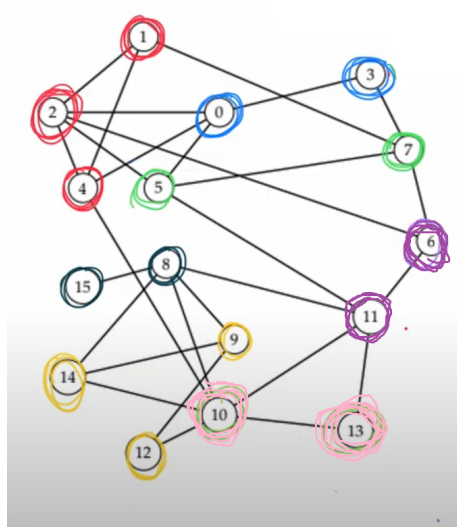


Figura 4.8: Comunidades formadas en la red de ejemplo al finalizar la primera iteración del algoritmo de Louvain.

Para poder evaluar los casos que no se consideraron en la primera iteración (unir el nodo 1 con el nodo 7 en vez de unir el nodo 1 con el nodo 4), a continuación el proceso se reinicia desde la situación inicial, pero tomando los otros caminos de decisión posibles. Si la estructura de comunidad cambia, teniendo una modularidad mayor que la obtenida en el camino anterior, entonces se toma la última estructura. Si no hay cambios, entonces se finaliza la primera fase del algoritmo.

Una vez que se finaliza la primera fase del algoritmo de Louvain se agrupan los nodos de cada comunidad, creando una nueva red donde los nodos son las comunidades de la fase anterior. Todos los enlaces entre nodos de la misma comunidad ahora están representados por auto-bucles en el nuevo nodo de comunidad y los enlaces entre nodos de comunidades distintas se representan por bordes ponderados, donde la ponderación es el número de aristas entre dichas comunidades.

En la Fig. 4.9 se puede observar la nueva red que se crea al combinar los nodos de la Fig. 4.8.

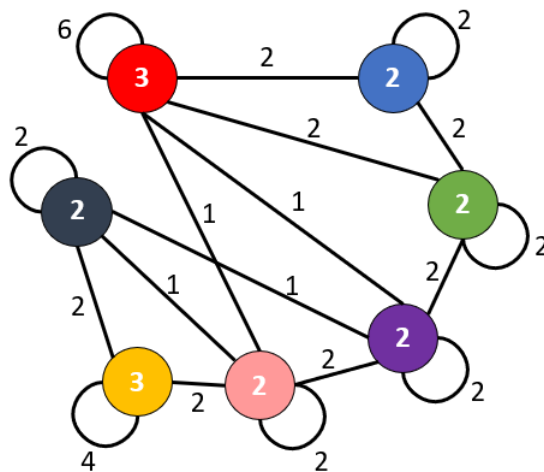


Figura 4.9: Combinación de los nodos de una misma comunidad en un nuevo nodo.

Por ejemplo, el nodo rojo de la red de la Fig. 4.9 representa la combinación de los 3 nodos de la comunidad roja de la red de la Fig. 4.8. El peso del auto-bucle de este nuevo nodo rojo representa la suma del grado de los 3 nodos de la comunidad roja de la red de la Fig: 4.8, donde el grado de cada uno de estos nodos se calcula considerando solo los enlaces dentro de la comunidad. Como el grado de los nodos 1,

2 y 4 de la red de la Fig. 4.8 es 2 cada uno, entonces el peso del auto-bucle es 6 para el nodo rojo en la nueva red de la Fig. 4.9. El peso de la arista entre los nodos en la nueva red de la Fig. 4.9 representa la cantidad de enlaces entre las comunidades de la red de la Fig. 4.8.

Una vez que se crea esta nueva red (Fig. 4.9), se regresa al comienzo, seleccionando un nodo (comunidad) a la vez, y calculando la modularidad que resulta al unirlo con cada uno de los nodos (comunidades) con los que se conecta. El procedimiento se aplica entonces recursivamente, hasta que ya no es posible aumentar la modularidad uniendo comunidades entre sí.

Al finalizar completamente el algoritmo, para esta red de ejemplo resultan dos comunidades que en la Fig. 4.10 se representan en colores verde y celeste. La línea roja en la Fig. 4.10 representa la partición de la red en ambas comunidades. Con esta estructura, se maximiza la modularidad  $M$ .

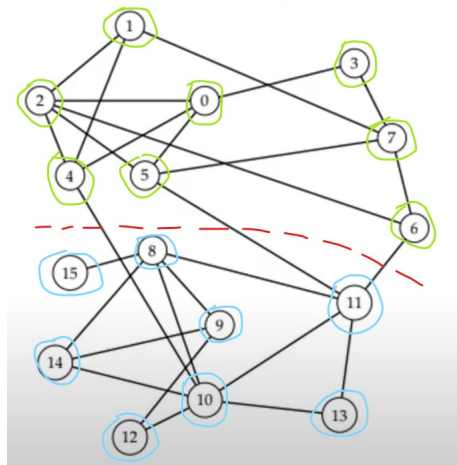


Figura 4.10: Estructura de comunidades de la red de ejemplo.



La suma de los grados de los nodos de la comunidad verde y celeste, respectivamente, son  $k_{\text{verde}} = 29$  y  $k_{\text{celeste}} = 27$ .

Asimismo, la suma de los grados de los nodos considerando solo las aristas internas de las comunidades verde y celeste, respectivamente son  $A_{\text{verde}} = 26$  y  $A_{\text{celeste}} = 24$ .

Entonces el valor de  $M$  para la estructura de comunidad óptima ( $M_{\text{óptima}}$ ) se obtiene de la suma de la modularidad de ambas comunidades:

$$M_{\text{óptima}} = M_{\text{verde}} + M_{\text{celeste}} ,$$

$$M_{\text{óptima}} = \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(C_{\text{verde}}, C_{\text{verde}}) + \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(C_{\text{celeste}}, C_{\text{celeste}}) ,$$

$$M_{\text{óptima}} = \frac{1}{2m} \left[ 26 - \frac{29 \cdot 29}{2m} \right] + \frac{1}{2m} \left[ 24 - \frac{27 \cdot 27}{2m} \right] ,$$

$$M_{\text{óptima}} = 0.4 .$$

Observemos que si se quisiera intentar unir a las comunidades verde y celeste en una sola, se obtendrá un valor negativo:

$$M = \frac{1}{2m} \left[ 6 - \frac{29 \cdot 27}{m} \right] = -0.39 ,$$

lo cual indica que ya no se puede seguir maximizando el valor de la modularidad.

# Capítulo 5

## Metodología

Realizamos un estudio de estructuras de comunidades sobre redes complejas construidas a partir de las series de tiempo de la intensidad de campo magnético registradas en diferentes puntos distribuidos sobre la superficie terrestre, a lo largo del ciclo solar 23. Trabajamos con dos redes diferentes, una de ellas se construye utilizando las series temporales que se obtienen a partir de las mediciones hechas por magnetómetros terrestres y la otra red se construye utilizando mediciones satelitales.

La construcción de una red compleja a partir de series temporales enfrenta dos desafíos principales: (1) cómo definir nodos y (2) cómo vincular nodos.

1. Los nodos de la red que se construye con los datos de magnetómetros terrestres son puntos sobre la superficie terrestre en donde se efectúan las mediciones, es decir, los nodos son la ubicación de los magnetómetros (ver Fig. 5.1). Y los nodos de la red que se construye a partir de los datos satelitales son puntos sobre la superficie terrestre que representan el promedio de los datos dentro de un área alrededor de este punto (ver Fig. 5.3). En las secciones 5.1 y 5.2 se explica con más detalles.

2. Las conexiones de los nodos (las aristas de la red) están dadas por el grado de similitud entre las series temporales observadas en dichos puntos. Luego dos nodos se consideran conectados solo si el valor de este grado de similitud es mayor o igual que algún umbral apropiado, lo que impone un nivel muy alto de confianza en esa interacción particular.

Existen diversos métodos para medir la similitud entre series temporales, como la correlación de Pearson [10, 53], la correlación de Spearman [10], correlación cruzada [54, 55], información mutua [53, 56], la sincronización de eventos [5], etc.

Para construir nuestras redes, se decidió seguir una estrategia similar a la utilizada en los trabajos para los registros de lluvia, [5, 10, 53] y para los de registro de campo magnético [6], en los cuales para medir similitudes temporales se utiliza la correlación de Pearson [10] y el método de sincronización de eventos [5]. Cada uno de ellos, a su vez, involucra la elección de ciertos umbrales de correlación, que determinan si dos nodos (sitios) están efectivamente conectados o no. Además de los dos criterios de similitud, consideraremos diversos umbrales de correlación, con el fin de establecer cómo depende la estructura de comunidades (y su evolución con el ciclo solar) de dichas elecciones.

## 5.1. Obtención de los datos de los magnetómetros

En primer lugar para construir la red compleja, consideramos las lecturas de campo magnético por hora, durante todo el ciclo solar 23, en 59 magnetómetros distribuidos sobre la superficie de la Tierra (ver Fig. 5.1, que son obtenidos del catálogo de datos del Centro Mundial de Datos (World Data Center, WDC) para Geomagnetismo, Kyoto [27]. Seleccionamos sólo aquellos magnetómetros del catálogo que cubren, al menos, el 90 % del ciclo solar 23. Esto nos permitirá estudiar la evolución de la estructura de comunidades a lo largo del ciclo solar realizando ventanas temporales móviles de datos de 1 año de duración, asegurando que en todas las ventanas hay datos obtenidos con todos los magnetómetros seleccionados.

Debido a que definiremos ventanas de 1 año de duración, por simplicidad consideraremos registros desde el comienzo del primer año calendario del ciclo (1 de enero de 1996) hasta el final del último año calendario del ciclo (31 de diciembre de 2008).

En la tabla 5.1 se muestran algunas características generales de los datos usados:

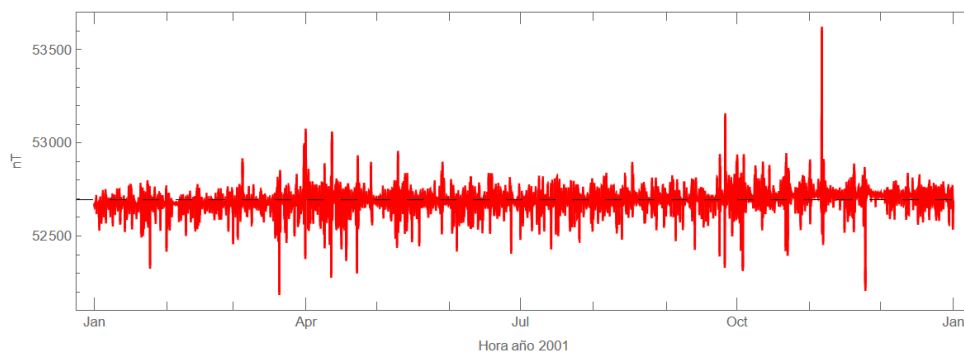
Total de magnetómetros	59
Total de datos por año de 1 magnetómetro	8760 (8784 año bisiesto)
Total de años del ciclo solar	13 (1996–2008)

Tabla 5.1: Informaciones generales sobre los datos.

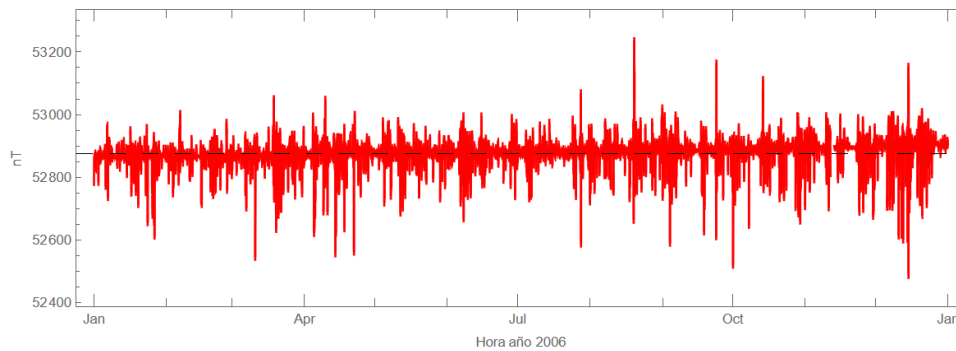
La Fig. 5.1 muestra la ubicación de las estaciones seleccionadas, y en la Fig. 5.2 se tiene un ejemplo de las series de tiempo registradas en dichos magnetómetros.



Figura 5.1: Ubicación de las estaciones geomagnéticas sobre la superficie terrestre.



(a) Serie de tiempo de la intensidad del campo magnético año 2001.



(b) Serie de tiempo de la intensidad del campo magnético año 2006.

Figura 5.2: Ejemplo de series de tiempo de la intensidad del campo magnético de los años (a) 2001 y (b) 2006 de la estación Abisko ubicada en latitud  $68^{\circ}\text{N}$  y longitud  $18^{\circ}\text{E}$ .

## 5.2. Obtención de los datos satelitales

Una dificultad que se ha encontrado es que las 59 estaciones seleccionadas, que tienen cobertura de datos suficiente para hacer un análisis a lo largo del ciclo solar completo, no están distribuidas uniformemente. Esto puede afectar los resultados, ya que dos estaciones muy cercanas pueden tener una mayor correlación sólo por su proximidad espacial. Por lo tanto, para complementar las conclusiones de los datos de los magnetómetros terrestres, realizamos un análisis similar al propuesto, pero ahora considerando datos desde el año 1999 al año 2005 del satélite Ørsted (tomados del Instituto Nacional del Espacio de la Universidad Técnica de Dinamarca, [https://www.space.dtu.dk/english/research/scientific\\_data\\_and\\_models/magnetic\\_satellites](https://www.space.dtu.dk/english/research/scientific_data_and_models/magnetic_satellites)), y datos desde el año 2000 al año 2008 del satélite CHAMP (tomados del Centro Alemán de Investigación de Geociencias, <https://isdg.gfz-potsdam.de/champ-isdg/access-to-the-champ-data/>).

Para ambos casos consideramos una red de 63 puntos distribuidos uniformemente con una separación de  $22.5^\circ$  en latitud y  $40^\circ$  en longitud como se observa en la Fig. 5.3. Cada uno de estos puntos representa el promedio de los datos dentro de un área alrededor de este punto con una altura de  $22.5^\circ$  de latitud y un ancho de  $40^\circ$  de longitud. Los satélites no miden permanentemente en cada posición, sino que cambian la posición de observación a medida que se mueven en su órbita, por lo que además del promedio espacial, se realiza un promedio sobre una cierta ventana temporal, adecuada para acumular suficientes observaciones en cada región rectangular. Dicha ventana temporal se escogió con una duración de 5 días, por razones que se argumentarán a continuación.

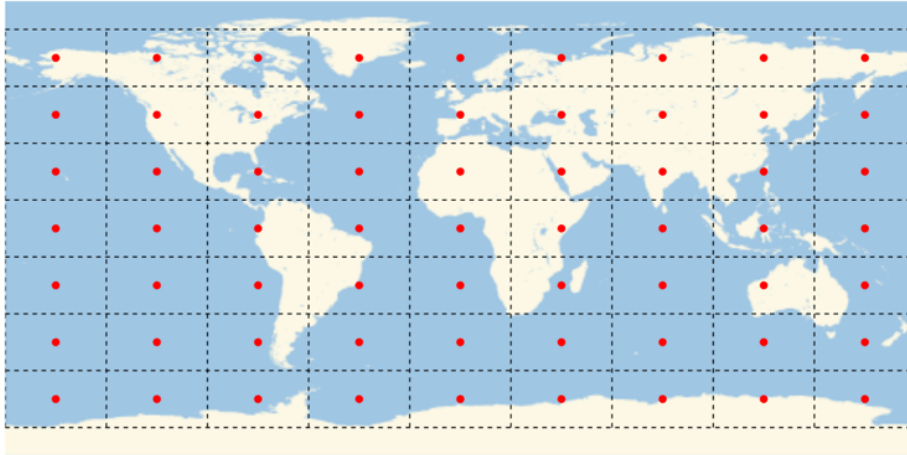


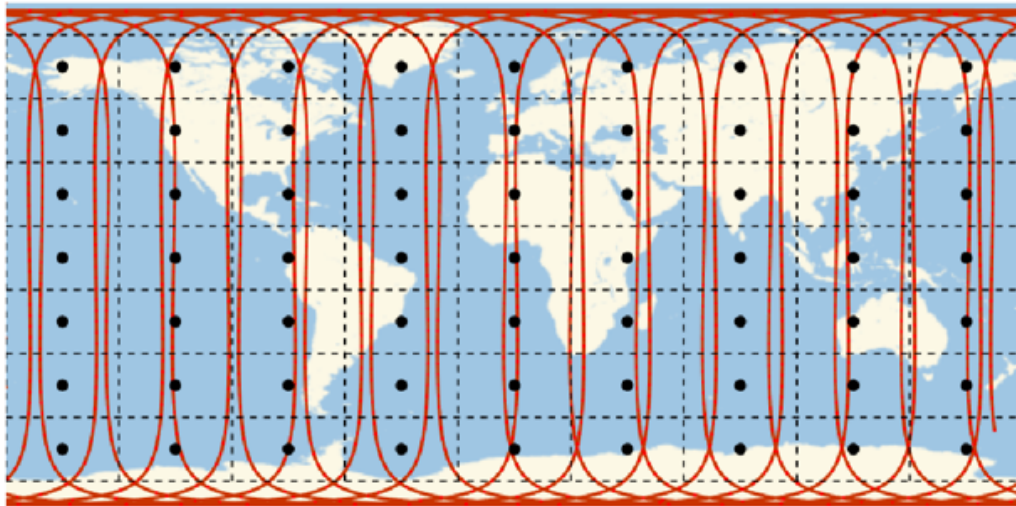
Figura 5.3: Distribución uniforme de puntos sobre la superficie terrestre. Cada punto es el promedio de cinco días de datos registrados dentro del área representada por las líneas segmentadas que rodean al punto.

Información básica sobre los datos satelitales se presenta en la tabla 5.2.

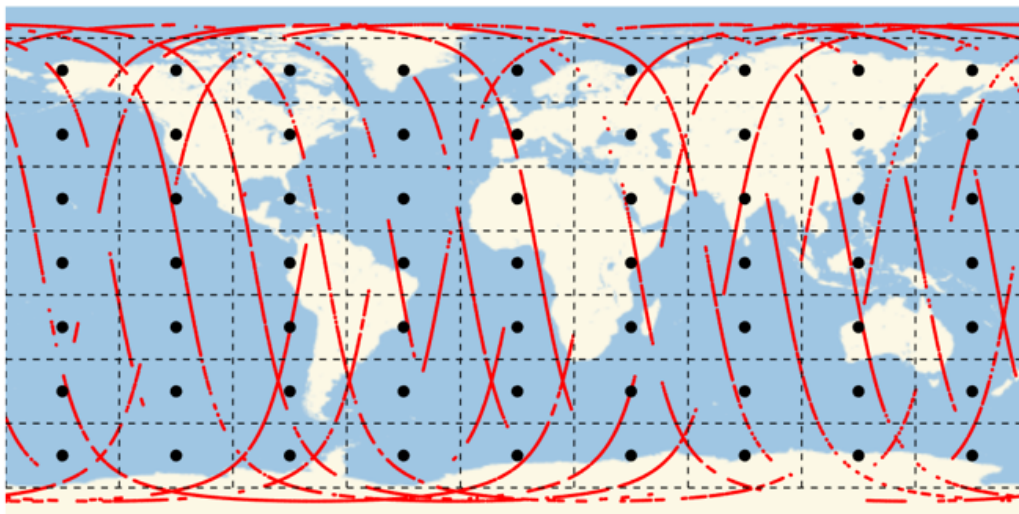
Total de puntos	63
Total de datos por año en 1 punto	72
Total de años del satélite CHAMP	9 (2000–2008)
Total de años del satélite Ørsted	7 (1999–2005)

Tabla 5.2: Datos satelitales.

La Fig. 5.4 muestra la proyección sobre la superficie terrestre de las trayectorias a lo largo de un día para ambos satélites, indicando la grilla y puntos representativos de ella mostrados en la Fig. 5.3.



(a) Trayectoria de un día del satélite CHAMP.

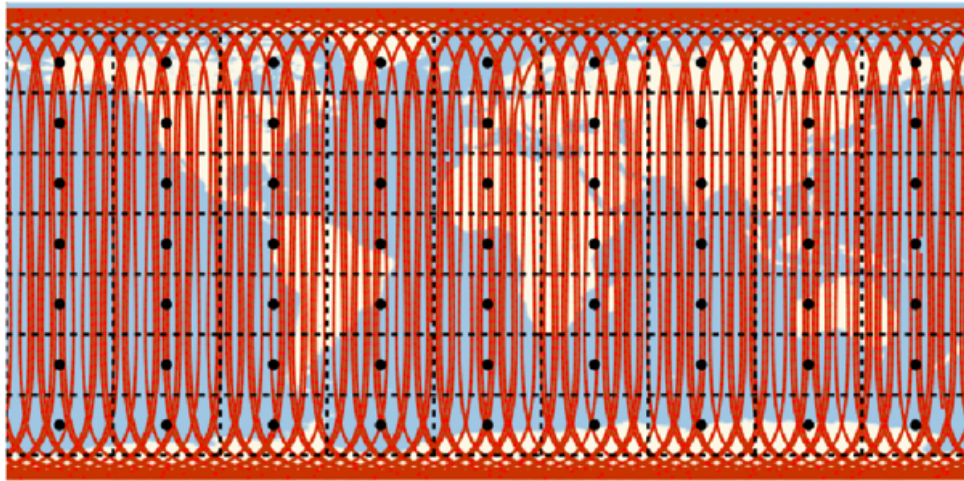


(b) Trayectoria de un día del satélite Ørsted.

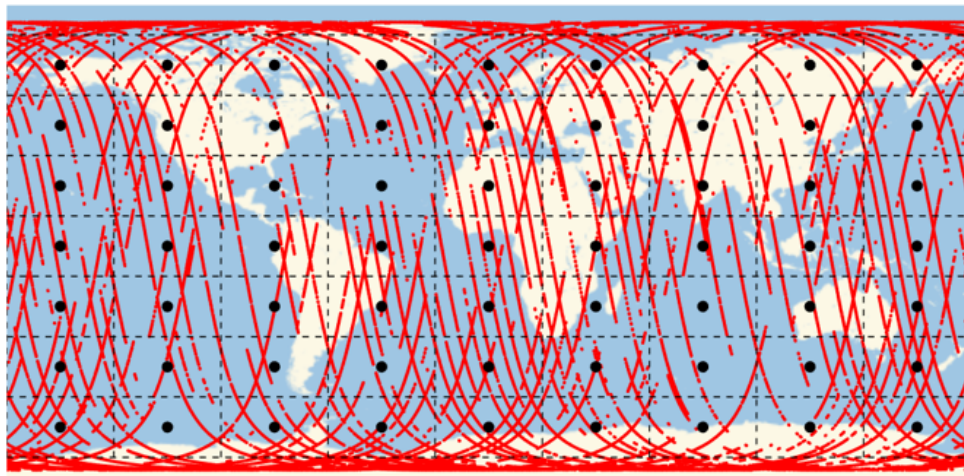
Figura 5.4: Trayectoria de un día de los satélites (a) CHAMP y (b) Ørsted.



Se puede apreciar que, considerando una rotación completa de la Tierra, no hay suficientes mediciones en cada celda de la grilla como para que el promedio sea representativo. Sin embargo, la situación cambia cuando se consideran 5 días, como se aprecia en la Fig. 5.5.



(a) Trayectoria de cinco días del satélite CHAMP.



(b) Trayectoria de cinco días del satélite Ørsted.

Figura 5.5: Trayectoria de cinco días de los satélites (a) CHAMP y (b) Ørsted.

De la Fig. 5.5 podemos observar que considerando 5 días de datos se logra cubrir mejor el área de cada celda, y los promedios sobre dicha ventana temporal serán representativos de la celda.

A partir de la información de las secciones anteriores y de la similitud entre las series de tiempo de cada magnetómetro, ya sea de los datos terrestres o de los datos satelitales, se construyen las redes complejas a estudiar.

### 5.3. Construcción de la red

Para construir nuestra red, como mencionamos anteriormente, vamos a considerar los nodos como los puntos donde se efectúan las mediciones. Para los datos de magnetómetros terrestres, los puntos corresponden a la ubicación de dichos magnetómetros (Fig. 5.1); para los datos satelitales, corresponde a los puntos de la grilla indicados en la Fig. 5.3, que a su vez representan los promedios sobre una cierta ventana temporal y espacial. La conexión entre dichos nodos, en ambos casos, es dada por la similitud entre sus respectivas series de tiempo de intensidad de campo magnético.

Sin pérdida de generalidad, a continuación se explica cómo se construye la red, pensando en series de tiempo correspondientes a lecturas de campo magnético tomadas con magnetómetros terrestres. La construcción de la red a partir de datos satelitales es totalmente análoga, cambiando las series de tiempo correspondientes.

Dado que la selección de un método de similitud involucra un grado de arbitrariedad, es importante saber cómo afecta esta selección a la red y cuál es más adecuado para la construcción de ella. Por ello, utilizamos dos métodos para construir las redes

de magnetómetros: la correlación de Pearson, el cual es un método común para definir la conexión entre nodos ya que evalúa la relación lineal entre dos variables [10], y la sincronización de eventos [57], debido a que se ha diseñado específicamente para cuantificar la sincronización de eventos extremos.

### 1. Correlación de Pearson (CP)

La correlación de Pearson (CP) entre dos series de tiempo  $x(t)$  y  $y(t)$  se define como:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (5.1)$$

donde  $n$  es el tamaño de la serie de tiempo;  $x_i$  e  $y_i$  son los puntos de las series individuales indexados por  $i$ ;  $\bar{x}$  e  $\bar{y}$  son la media de sus series de tiempo respectivas.

El valor de CP entre dos series de tiempo varía entre  $r_{xy} = -1$  y  $r_{xy} = 1$ , donde  $r_{xy} = 1$  indica una correlación positiva perfecta, es decir, existe una relación directamente proporcional entre las series de tiempo, y  $r_{xy} = -1$  indica una correlación negativa perfecta, es decir, existe una relación inversamente proporcional entre las series de tiempo.

### 2. Sincronización de eventos (SE)

La sincronización de eventos (SE) se ha diseñado específicamente para calcular relaciones no lineales entre series temporales, a partir de la identificación de eventos en cada serie de tiempo. Por ejemplo, si una serie tiene *peaks* bien definidos, y otra tiene también *peaks* simultáneos con los de la primera, podemos

afirmar que hay una correlación entre ambas. Lo mismo se puede decir si, después de cada *peak* en un serie de tiempo, sucede un *peak* en la otra, después de un cierto tiempo  $\Delta t$  fijo. En ese caso, también debería haber una correlación. Para series de tiempo reales, seguramente no habrá una coincidencia exacta de *peaks*, y si hay un retraso entre ambas series, tampoco será constante. Esto se deberá tener en cuenta para determinar la similitud entre series de tiempo, basada en los eventos extremos, como se apreciará en la definición del algoritmo presentado más abajo.

En nuestro caso, utilizamos el algoritmo propuesto por [58], modificado para considerar mínimos del campo magnético (que indican alta actividad magnética) en vez de máximos. Consideremos dos series de tiempo  $x(t)$  e  $y(t)$  que corresponden a la serie de tiempo de la intensidad del campo magnético medido en dos posiciones diferentes.

Primero se deben identificar los eventos extremos, que definiremos como todos los datos que están dentro del 5% de los valores más bajos de cada serie de tiempo. Dicho 5% permite un compromiso entre tener, por una parte, un umbral significativo para determinar que la intensidad del campo magnético está debilitado, y por otra parte, un número suficiente de eventos en cada serie de tiempo para tener una buena estadística (en este caso, el 5% del número de datos es aproximadamente 430) [5].

La Fig. 5.6 ilustra la ubicación de los eventos extremos determinados con el criterio anterior, para una de las series de tiempo de magnetómetros utilizadas en este estudio.

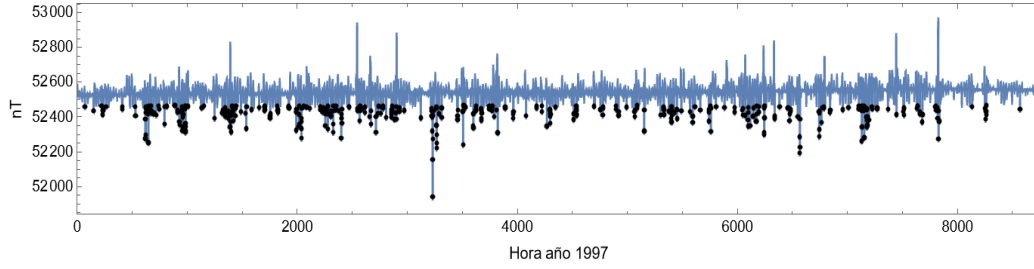


Figura 5.6: Ejemplo de serie de tiempo de la intensidad del campo magnético del año 1997 de la estación Abisko ubicada en latitud  $68^\circ\text{N}$  y longitud  $18^\circ\text{E}$ . Los puntos negros representan los eventos extremos de la serie.

Los eventos extremos ocurren en las series  $x(t)$  e  $y(t)$ , respectivamente, en ciertos tiempos  $t_l^x$  y  $t_m^y$ , donde  $l = 1, 2, 3, 4, \dots, S_x$ ,  $m = 1, 2, 3, 4, \dots, S_y$ , y  $S_x$ ,  $S_y$  son el número total de eventos extremos registrados por cada magnetómetro.

En los cálculos posteriores se debe excluir el primer y el último evento de cada serie, por lo tanto, establecemos  $l = 2, 3, 4, \dots, S_x - 1$  y  $m = 2, 3, 4, \dots, S_y - 1$ .

Como no es esperable que las dos series de tiempo tengan eventos realmente simultáneos, se buscará la coincidencia de eventos extremos con una tolerancia  $\tau_{lm}^{xy}$ , definida en base a la mínima separación entre eventos extremos en ambas series [59]:

$$\tau_{lm}^{xy} = \frac{1}{2} \min\{t_{l+1}^x - t_l^x, t_l^x - t_{l-1}^x, t_{m+1}^y - t_m^y, t_m^y - t_{m-1}^y\}. \quad (5.2)$$

A continuación se calcula  $C(x|y)$ , el número de veces que un evento extremo en la serie  $x(t)$  ocurre separado, a lo más, un tiempo  $\tau_{lm}^{xy}$  de un evento en la serie  $y(t)$ :

$$C(x|y) = \sum_{l=2}^{S_x-1} \sum_{m=2}^{S_y-1} J_{lm}^{xy}, \quad (5.3)$$

donde

$$J_{lm}^{xy} = \begin{cases} 1 & \text{si } 0 < t_l^x - t_m^y < \tau_{lm}^{xy} \\ \frac{1}{2} & \text{si } t_l^x = t_m^y \\ 0 & \text{otro caso} \end{cases}. \quad (5.4)$$

Notemos que, en la Ec. (5.3),  $l$  y  $m$  recorren todo el intervalo de enteros posibles, de modo que  $J_{lm}^{xy} = 1$  sólo cuando un evento extremo en una de las series ( $x(t)$ ) ocurre *antes* que un evento extremo en la otra serie ( $y(t)$ ). Si  $l$  y  $m$  se intercambian, los mismos eventos no contribuyen a la suma. El caso en que dos eventos extremos ocurren simultáneamente se trata por separado, ya que los eventos siguen coincidiendo si se intercambian sus índices  $l$  y  $m$ . Por esa razón, se define  $J_{lm}^{xy} = 1/2$  en este caso, de modo que, al final, ese par de eventos también contribuye con un término igual a 1 a la suma en la Ec. (5.3).

De forma análoga se calcula  $C(y|x)$ , que es el número de veces que un evento en la serie  $y(t)$  ocurre separado, a lo más, un tiempo  $\tau_{lm}^{xy}$  de un evento en la serie  $x(t)$ .

A partir de estas cantidades, obtenemos

$$Q_{xy} = \frac{C(x|y) + C(y|x)}{\sqrt{(S_x - 2)(S_y - 2)}} , \quad (5.5)$$

donde se ha normalizado, recordando que la cantidad de eventos extremos que se utilizan para calcular  $Q_{xy}$  son  $S_x - 2$  y  $S_y - 2$ , por haber excluido el primer y último evento extremo de cada serie  $x(t)$  e  $y(t)$ .

Por lo tanto,  $Q_{xy}$  es una medida de la sincronización de eventos entre  $x(t)$  e  $y(t)$ .  $Q_{xy}$  está normalizado de modo que  $0 \leq Q_{xy} \leq 1$ . Si  $Q_{xy} = 0$  indica ausencia de sincronización, y si  $Q_{xy} = 1$  corresponde a una sincronización perfecta entre las series de tiempo  $x(t)$  e  $y(t)$ . Observemos también que, al ser simétrico el numerador en la Ec. (5.5), esta cantidad no contiene información sobre la direccionalidad de la red, no siendo relevante si los eventos extremos en  $x(t)$  ocurren antes o después de los eventos extremos en  $y(t)$ .

Como ejemplo para entender la sincronización de eventos consideremos dos series  $x(t)$  e  $y(t)$  con 100 datos cada una.

a)  $x(t)$ : 924, 172, 105, 290, 752, 459, 145, 378, 904, 351, 511, 403, 368, 939, 647, 134, 61, 70, 905, 277, 155, 787, 313, 385, 551, 142, 125, 713, 218, 368, 715, 226, 473, 646, 558, 3, 662, 167, 336, 619, 538, 423, 940, 417, 982, 552, 476, 280, 761, 727, 812, 743, 502, 787, 323, 151, 667, 827, 592, 804, 490, 913, 724, 17, 263, 446, 869, 308, 24, 238, 278, 567, 357, 282, 725, 753, 617, 82, 26, 126, 164, 697, 544, 328, 44, 128, 286, 318, 431, 358, 713, 658, 866, 948, 662, 815, 615, 728, 360, 830

b)  $y(t)$ : 254, 521, 99, 660, 502, 162, 345, 660, 154, 231, 932, 926, 86, 999, 993, 592, 359, 773, 703, 793, 290, 912, 495, 667, 769, 483, 237, 443, 126, 349, 752, 910, 516, 120, 503, 701, 359, 334, 407, 742, 404, 336, 126, 310, 450, 932, 648, 92, 822, 428, 254, 3, 197, 863, 793, 24, 246, 831, 377, 986, 539, 551, 468, 29, 89, 449, 238, 552, 70, 39, 323, 378, 519, 275, 722, 308, 561, 423, 769, 956, 872, 286, 321, 601, 47, 245, 609, 483, 167, 132, 767, 65, 498, 181, 99, 368, 698, 647, 292, 307

Entonces, en cada serie de tiempo identificamos los eventos extremos, que son el 5% de los valores más bajos, por lo tanto, tenemos que el número de eventos extremos en cada serie es  $S_x = S_y = 5$ . Entonces,  $l = 1, 2, 3, 4, 5$  y  $m = 1, 2, 3, 4, 5$  son los índices que rotulan dichos eventos extremos en cada serie. Los eventos extremos en cada serie de tiempo son:

a) Para la serie  $x(t)$ : 3, 17, 24, 26, 44.

b) Para la serie  $y(t)$ : 3, 24, 29, 39, 47.

Tenemos que identificar los tiempos  $t_l^x$  y  $t_m^y$  de los eventos extremos en sus series de tiempo, que, en este caso, corresponden a la posición que ocupa el evento extremo en la serie de tiempo:

a)  $t_l^x = 36, 64, 69, 79, 85$ .

b)  $t_m^y = 52, 56, 64, 70, 85$ .

Por ejemplo el evento extremo de valor 17 de la serie de tiempo  $x(t)$  está en la posición 64.



Ahora se calcula la tolerancia  $\tau_{lm}^{xy}$  utilizando la ecuación (5.2), para el caso,  $l = m = 2$ :

$$\begin{aligned}\tau_{22}^{xy} &= \frac{1}{2} \min\{t_3^x - t_2^x, t_2^x - t_1^x, t_3^y - t_2^y, t_2^y - t_1^y\} \\ &= \frac{1}{2} \min\{69 - 64, 64 - 36, 64 - 56, 56 - 52\} = 2.\end{aligned}$$

Y así sucesivamente con todas las combinaciones posibles de  $l$  y  $m$ . Recordemos nuevamente, que se debe excluir el primer y el último evento de cada serie.

Por lo tanto, se obtiene una matriz como se muestra a continuación:

$$\tau_{lm}^{xy} = \begin{pmatrix} 2 & 2.5 & 2.5 \\ 2 & 2.5 & 2.5 \\ 2 & 3 & 3 \end{pmatrix}.$$

Luego utilizando la ecuación (5.4) calculamos la matriz  $J_{xy}$  y  $J_{yx}$ :

$$J_{xy} = \begin{pmatrix} 0 & 0.5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$J_{yx} = \begin{pmatrix} 0 & 0.5 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Usando la ecuación (5.3) se tiene que  $C(x|y) = 0.5$  y  $C(y|x) = 1.5$ , luego usando la ecuación (5.5) la medida de sincronización  $Q_{xy}$  es:

$$Q_{xy} = \frac{2}{\sqrt{(5-2)(5-2)}} = \frac{2}{3} \approx 0.6 .$$

Este valor de  $Q_{xy}$  para este ejemplo indica una sincronización entre las series de tiempo  $x(t)$  e  $y(t)$  de 0.6.

En este trabajo, estamos interesados en identificar correlaciones entre estaciones. Sin embargo, es ilustrativo ver cómo el método de sincronización de eventos permite, efectivamente, identificar correlaciones significativas entre dos series de tiempo. Para ello, tomaremos series de tiempo para dos estaciones particulares, una cerca del Ecuador, y otra cerca del polo norte, y las compararemos con la serie de tiempo del índice  $Dst$ , que estima el campo magnético ecuatorial.

En particular, consideraremos series de tiempo para los años 1997, 2001, 2005 y 2008, para dos estaciones distintas: la estación (a) se ubica cerca de una de las estaciones que se utilizan para calcular el índice  $Dst$  (representada por un rombo celeste en la Fig. 5.7) y la estación (b) se ubica en el hemisferio norte lejos de la línea del Ecuador (representada por un rombo negro en la misma figura).

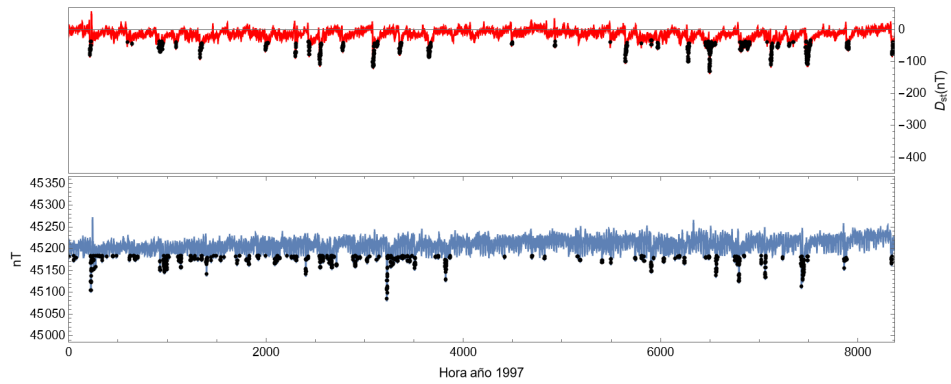
Para cada serie de tiempo, identificamos la posición de sus eventos extremos comparándolas con las posiciones del 5 % de los valores más bajos del índice *Dst*.

Los resultados se muestran en la Figs. 5.8–5.11.

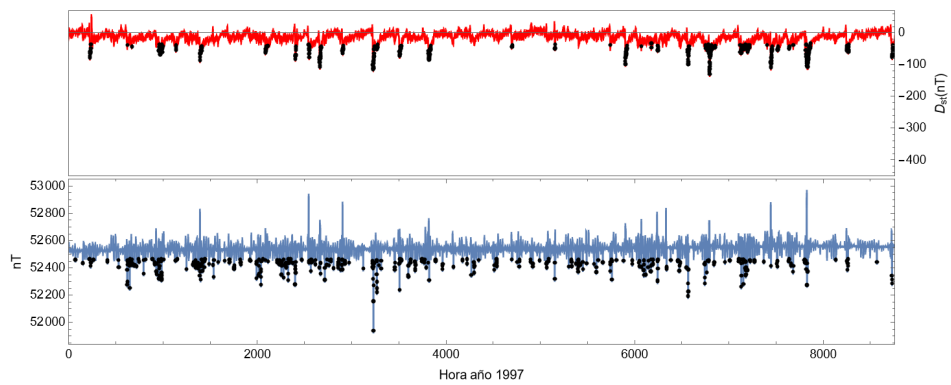


Figura 5.7: Ubicación de la estación (a) representada por un rombo celeste y la estación (b) representada por un rombo negro. Los puntos rojos representan las estaciones que se utilizan para calcular el índice *Dst*. La línea segmentada representa la línea del Ecuador.

■ Series de tiempo para el año 1997



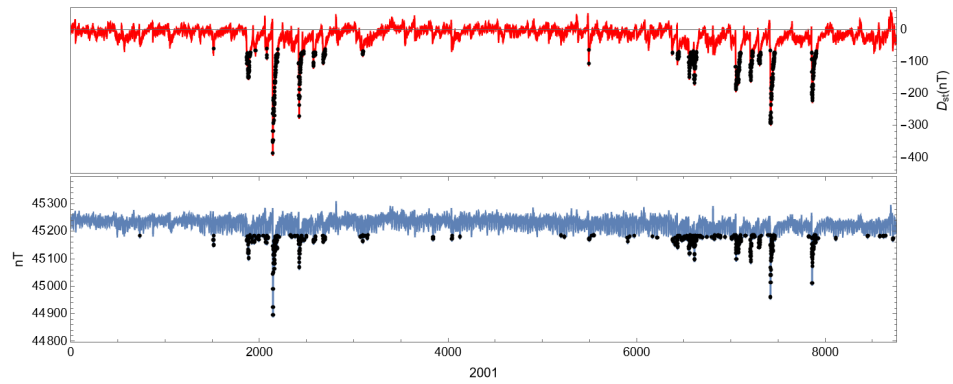
(a) Serie de tiempo de la intensidad del campo magnético año 1997 de la estación (a).



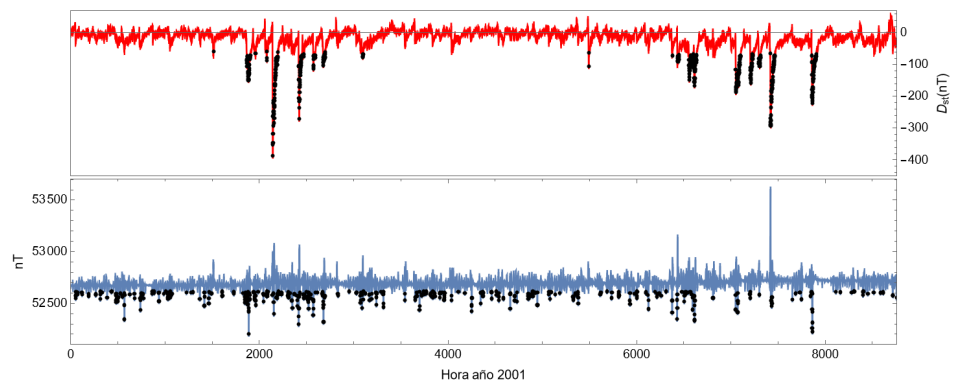
(b) Serie de tiempo de la intensidad del campo magnético año 1997 de la estación (b).

Figura 5.8: (Curva roja) Índice *Dst*. (Curva azul) Intensidad del campo magnético terrestre. (Puntos negros) Eventos extremos de cada serie de tiempo.

■ Series de tiempo para el año 2001



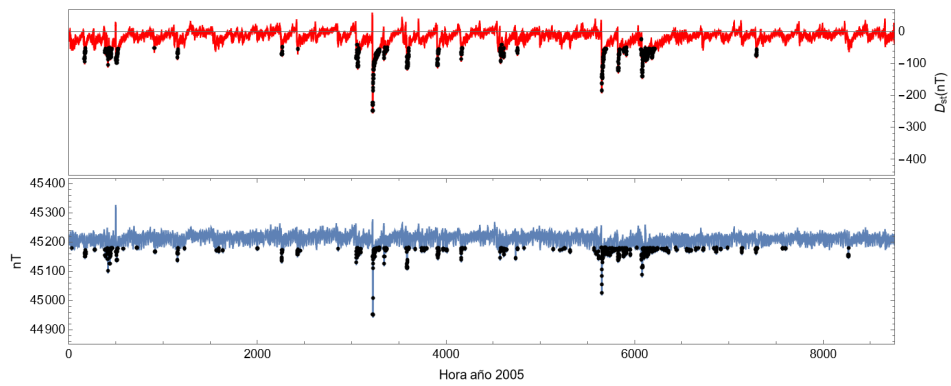
(a) Serie de tiempo de la intensidad del campo magnético año 2001 de la estación (a).



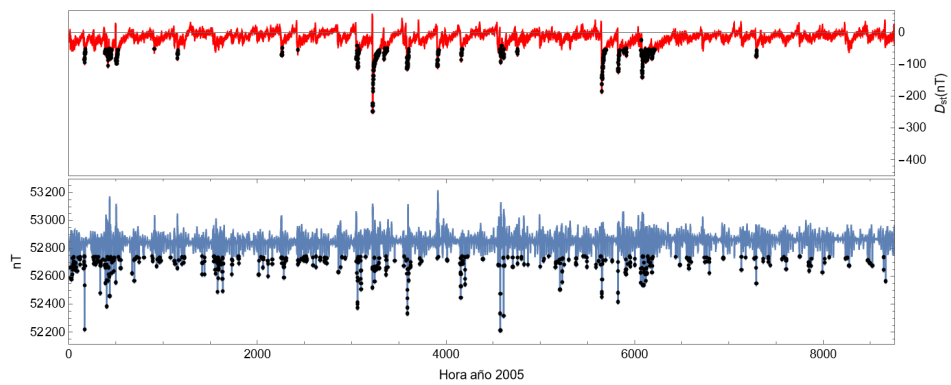
(b) Serie de tiempo de la intensidad del campo magnético año 2001 de la estación (b).

Figura 5.9: (Curva roja) Índice *Dst*. (Curva azul) Intensidad del campo magnético terrestre. (Puntos negros) Eventos extremos de cada serie de tiempo.

■ Series de tiempo para el año 2005



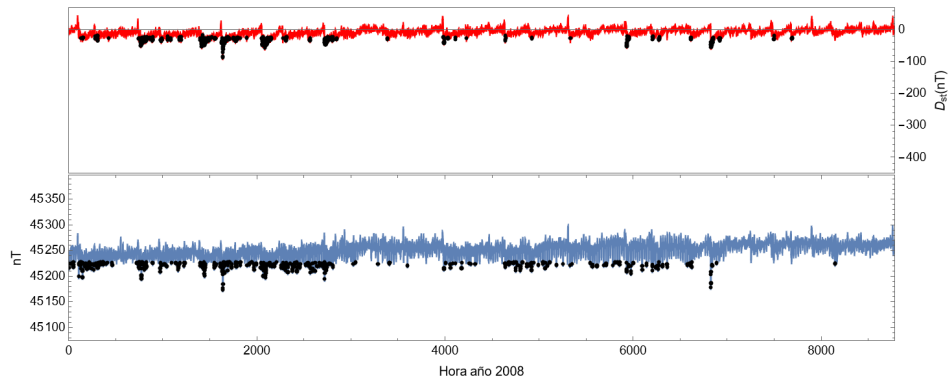
(a) Serie de tiempo de la intensidad del campo magnético año 2005 de la estación (a).



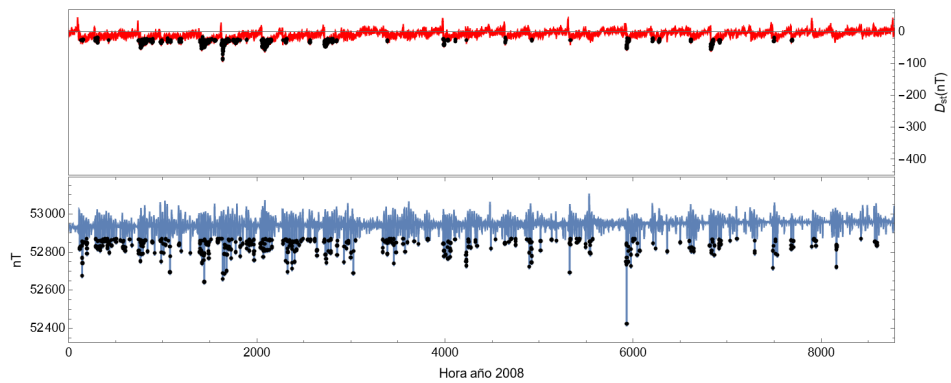
(b) Serie de tiempo de la intensidad del campo magnético año 2005 de la estación (b).

Figura 5.10: (Curva roja) Índice  $Dst$ . (Curva azul) Intensidad del campo magnético terrestre. (Puntos negros) Eventos extremos de cada serie de tiempo.

■ Series de tiempo para el año 2008



(a) Serie de tiempo de la intensidad del campo magnético año 2008 de la estación (a).



(b) Serie de tiempo de la intensidad del campo magnético año 2008 de la estación (b).

Figura 5.11: (Curva roja) Índice  $Dst$ . (Curva azul) Intensidad del campo magnético terrestre. (Puntos negros) Eventos extremos de cada serie de tiempo.

Las figuras muestran que hay una clara similitud entre las posiciones de los eventos extremos de las series de tiempo de la estación (a) ubicada cerca de la línea del Ecuador (rombo celeste ver Fig. 5.7) con la posición de los eventos extremos del índice  $Dst$ , lo cual es consistente con el hecho de que dicha estación está más cercana a la línea del Ecuador, y a una de las estaciones que se utilizan para determinar el índice  $Dst$ . Dicha similitud no se observa para la estación (b), en la cual los eventos extremos están más homogéneamente distribuidos, independiente de la ubicación de los eventos extremos en  $Dst$ .

Así, podemos apreciar que la identificación de eventos extremos permite establecer una similitud o correlación entre series de tiempo que es físicamente significativa.

Si calculáramos el valor de  $Q$  asociado a los pares de series de tiempo mostrados en el ejemplo anterior, dado por la Ec. (5.5), obtendríamos cierto número. Pero ahora corresponde determinar cuán alto debe ser  $Q$ , para afirmar que realmente las series están correlacionadas. Este umbral de correlación, que llamaremos  $\theta$ , también debe ser elegido en base a algún criterio, ya que determinará si dos nodos (sitios) están conectados o no. Lo mismo sucede con  $r_{xy}$ , el resultado de calcular la correlación de Pearson entre dos series de tiempo.

Como el valor de  $r_{xy}$  varía entre  $-1$  y  $1$ , y el valor de  $Q_{xy}$  varía entre  $0$  y  $1$  se debe definir un umbral de correlación  $\theta$  que determina si dos nodos (sitios) están conectados o no, con el fin de establecer cómo depende la estructura de comunidades (y su evolución con el ciclo solar) de dichas elecciones.



Entonces, si  $r_{xy}$  o  $Q_{xy}$ , según corresponda, es mayor que dicho umbral, se considerará que ambas series de tiempo son similares y, por lo tanto, ambos magnetómetros están conectados en la red [10]. Si  $\theta$  es demasiado pequeño, entonces aumentan las conexiones de la red resultante y si  $\theta$  es muy grande, entonces se reducen las conexiones de la red resultante.

Un proceso similar es el que se utiliza en la binarización de una imagen [9, 15], donde se utiliza un umbral de intensidad  $\alpha$  con la intención de convertir la imagen digital en una imagen en blanco y negro, de tal manera que se preserven las propiedades esenciales de la imagen, y así por ejemplo poder facilitar la identificación de las manchas solares en magnetogramas.

Con el objeto de no prejuzgar acerca del umbral, y examinar cómo incide su elección en las características de la red, y su eventual correlación con la actividad solar, consideramos 10 valores diferentes, a saber,  $\theta = 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9$  y  $0.95$ .

Esto para que, siendo una elección arbitraria, dicha arbitrariedad tenga un efecto controlado sobre el resultado final, y así poder observar cómo cambian nuestros resultados en función del umbral escogido.

Para construir la red de magnetómetros, se indica a los magnetómetros con índices  $i, j$ , entonces la matriz de adyacencia de la red es

$$A_{i,j} = \begin{cases} 1 & \text{si } Q_{i,j}, r_{i,j} \geq \theta_{i,j} \\ 0 & \text{otro caso} \end{cases}, \quad (5.6)$$

usándose  $r_{i,j}$  si se emplea el método CP, o  $Q_{i,j}$  si es el método SE.  $A_{i,j} = 1$  denota un vínculo entre el  $i$ -ésimo y el  $j$ -ésimo nodos, y 0 denota lo contrario.

La matriz de adyacencia (ecuación (5.6)) representa las conexiones en la red de magnetómetros. Como se mencionó en la sección 4.1 en este estudio trabajamos con una red no dirigida. Esto es consistente con el hecho de que la relación de similitud, definida por los métodos CP o SE, es simétrica: si  $x(t)$  está correlacionada en el sentido de Pearson o de eventos extremos con la serie  $y(t)$ , entonces  $y(t)$  también lo está con  $x(t)$ .

Veamos un ejemplo de cómo utilizar los umbrales. Consideremos 5 series de tiempo. Si se usa la sincronización de eventos, construimos una matriz  $Q$  tal que el elemento  $Q_{i,j}$  representa la magnitud de la sincronización entre las series de tiempo  $i$  y  $j$ . Digamos que esa matriz es la siguiente:

$$Q = \begin{pmatrix} 1 & 0.3 & 0.87 & 0.43 & 0.92 \\ 0.3 & 1 & 0.78 & 0.86 & 0.5 \\ 0.87 & 0.78 & 1 & 0.12 & 0.9 \\ 0.43 & 0.86 & 0.12 & 1 & 0.6 \\ 0.92 & 0.5 & 0.9 & 0.6 & 1 \end{pmatrix} .$$

Entonces, dos magnetómetros se conectan en la red si el valor de su sincronización es tal que  $Q_{i,j}$  está por encima del umbral  $\theta$ . Consideremos un valor  $\theta_{i,j} = 0.85$ , con el cual obtenemos la siguiente matriz de adyacencia  $A$ :

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} .$$

De este modo, la elección de umbral nos permite obtener  $A$ , que es la representación matricial de las conexiones entre magnetómetros.

La matriz  $A$  es simétrica, indicando que la red es no dirigida. Su diagonal representa el coeficiente de correlación de una serie consigo misma, que siempre será máximo, por lo tanto no entrega información relevante. Por lo tanto, construimos la red solo con las conexiones de la matriz triangular superior, como se ve en la Fig. 5.12. Esto es equivalente a decir que ignoramos las autoconexiones de todos los nodos.

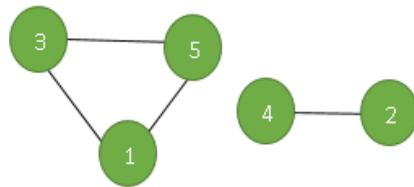


Figura 5.12: Ejemplo de una red construida utilizando la sincronización de eventos para 5 series temporales con un umbral  $\theta = 0.85$ .

Como indicamos en el Cap. 4, dependiendo de la estructura topológica de la red, los grupos de nodos (los magnetómetros) pueden agruparse formando comunidades [60]. Para determinar las comunidades dentro de cada red existen varios enfoques que tienen como objetivo estratificar los nodos en comunidades de una manera óptima (ver [37] para una revisión extensa).

Como se menciona en la sección 4.2, utilizaremos el método de Louvain para detectar las comunidades de la red, que en este caso significará agrupar los magnetómetros de acuerdo a sus correlaciones.

Como las redes de alta modularidad están densamente vinculadas dentro de las comunidades, pero escasamente vinculadas entre comunidades, entonces esperamos que bajos niveles de actividad geomagnética induzcan mayores correlaciones espa-

ciales entre puntos distantes, lo que se debería ver reflejado en un mayor número de nodos (magnetómetros) dentro de una comunidad, lo que implica que las comunidades que tengan más de dos estaciones sean más grandes.

Sin embargo, aunque podríamos esperar que el número de comunidades, número de nodos en ellas, su área, etc., sean cantidades correlacionadas, no deberíamos esperar una dependencia exacta, por lo tanto monitorearemos estas cantidades independientemente, para obtener una descripción lo más completa posible de la estructura de comunidades, y de su evolución a lo largo del ciclo solar.

Específicamente, obtenidas las comunidades en nuestra red, caracterizaremos el patrón resultante calculando las siguientes cantidades:

- a) número de comunidades de una sola estación, es decir, estaciones que no pertenecen a ninguna comunidad (estaciones aisladas);
- b) el número de comunidades de dos o más estaciones;
- c) el número total de comunidades (incluyendo las comunidades de dos o más estaciones y las de una sola estación);
- d) número total de nodos dentro de las comunidades (para calcular esta cantidad solo se consideran las comunidades de más de una estación);
- e) número promedio de nodos dentro de las comunidades (para calcular esta cantidad solo se consideran las comunidades de más de una estación);
- f) la comunidad con la mayor cantidad de nodos para cada año;
- g) área promedio cubierta por las comunidades.

En el caso del área, ésta será estimada como el área de la sección esférica comprendida entre los valores máximo y mínimo de latitud y longitud de la ubicación de los magnetómetros que componen dicha comunidad. Consistentemente, el área de las comunidades de una sola estación la consideramos igual a cero.

Otro modo de analizar los resultados a partir del estudio de comunidades propuesto es a través de la noción de “desorden”, por ejemplo utilizando la entropía de Shannon [61]:

$$S = \sum_{c \in C} -p_c \log_2(p_c) , \quad (5.7)$$

donde  $C = \{c_1, \dots, c_n\}$  es el conjunto de las comunidades de la red y  $p_c$  es la probabilidad de que un nodo pertenezca a la comunidad  $c$ . En este conjunto de comunidades se incluyen las comunidades de dos o más estaciones y las comunidades de estaciones aisladas. Así, la evolución de la estructura de comunidades podría reflejarse en un cambio de entropía.

Por ejemplo, si consideramos una red de 10 nodos, la cual tiene una comunidad de 5 nodos, una comunidad de 4 nodos y un nodo aislado (comunidad de un solo nodo) su entropía se calcula como:

$$S = \frac{-5}{10} \log_2 \left( \frac{5}{10} \right) + \frac{-4}{10} \log_2 \left( \frac{4}{10} \right) + \frac{-1}{10} \log_2 \left( \frac{1}{10} \right) ,$$

$$S = 1.36 .$$

También se considerará el valor de la modularidad óptima  $M$  obtenida, que también permite caracterizar la partición en comunidades resultante.

Se construye una red para cada año del ciclo solar 23 (entre 1996 y 2008), utilizando las dos medidas de similitud (SE y CP) y los umbrales ( $\theta_{i,j}$ ) mencionados anteriormente para cada medida.

Para analizar las variaciones de la intensidad del CMT graficamos estas métricas en función del tiempo y luego estos gráficos lo comparamos con el promedio anual de manchas solares (ver Fig. 2.1), con la variación del índice *Dst*, que corresponde a una estimación del campo magnético en el ecuador terrestre, y que es usualmente utilizado como un indicador de actividad geomagnética [26] (ver sección 3.2.1) y con el promedio anual del índice *Dst*.

Los gráficos de las cantidades calculadas en función del tiempo también se pueden comparar con la variación de otros índices geomagnéticos, tales como el índice AE y Ap (ver Apéndice A), estos resultados también permiten analizar las variaciones de la intensidad del CMT (ver Apéndice C).

Además de las métricas mencionadas anteriormente también se puede considerar calcular el coeficiente de clustering que entrega información sobre cuán agrupados están los nodos dentro de una red. Muestra la tendencia de que dos nodos vecinos de un nodo  $v_i$  estén conectados entre ellos [1]. En particular, el coeficiente de clustering para un nodo es la proporción de nodos vecinos a él, que están a su vez conectados entre sí. Esta cantidad que se calcula sobre la red, es independiente del método de detección de comunidades (para más detalle revisar Apéndice B).

# Capítulo 6

## Resultados

En este capítulo se muestran los resultados del estudio de comunidades que se obtienen utilizando el método de Louvain sobre las redes complejas no dirigidas, construidas mediante los algoritmos de similitud mencionados en la sección 5.3 sobre los registros de campo magnético del ciclo solar 23. Los nodos de la red corresponden a los puntos sobre la superficie terrestre (posiciones de los magnetómetros terrestres, o promedios sobre una determinada ventana espaciotemporal en el caso de datos satelitales).

Para estudiar las variaciones de las comunidades se consideran diversas métricas, relacionadas con el número de comunidades, su tamaño (en número de nodos o en área), la entropía de Shannon  $S$ , y la modularidad  $M$ . Estas métricas, enumeradas en detalle al final del Cap. 5, son comparadas con el promedio anual de manchas solares (Fig. 2.1), el índice  $Dst$  (Fig. 3.5) y el promedio anual del índice  $Dst$  a lo largo del ciclo solar 23, para intentar determinar la correlación entre estas métricas y la actividad magnética (en el Apéndice C se puede observar la comparación de estas métricas con los índices Ap y AE).

## 6.1. Datos observados por los magnetómetros

Como se mencionó en la sección 4.2 se utilizaron 10 umbrales de correlación  $\theta$  que determinan si dos nodos (sitios) están conectados o no. En esta sección mostraremos los resultados de las redes que se obtienen al aplicar los umbrales  $\theta = 0.7, 0.8$  y  $0.85$  en las matrices  $Q$  y  $r$  de los datos de los magnetómetros terrestres, los cuales son los que entregan resultados que muestran una mejor relación con la variación del índice  $Dst$  y con la variación del promedio anual de manchas solares a lo largo del ciclo solar.

En la Fig. 6.1 y Fig. 6.2 mostramos un ejemplo de las comunidades que se forman en 2 años distintos, uno de baja actividad solar (2005) y uno de alta actividad solar (2000), en base a CP y SE, respectivamente. Cada comunidad de magnetómetros es representada con rombos de colores distintos, salvo los magnetómetros aislados, en comunidades de un solo magnetómetro, que son indicados con triángulos naranjos. Para visualizar el tamaño espacial de las comunidades, se han agregado rectángulos que abarcan el rango de longitudes y latitudes de los magnetómetros correspondientes, rectángulo asociado al área cubierta por la comunidad, de acuerdo a lo indicado en la sección 5.3.

Observamos, para empezar, que los resultados sí dependen del método de similitud empleado para definir las conexiones, sin embargo, parece haber algunas características generales: en el año de baja actividad geomagnética, hay comunidades más grandes que en el año de alta actividad, lo cual podría ser consistente con nuestra sospecha inicial: la baja actividad implica un mayor “orden” en el campo magnético terrestre, reflejado en comunidades de mayor tamaño. Este tipo de análisis es el que



procuramos cuantificar de manera más sistemática a lo largo de este capítulo.

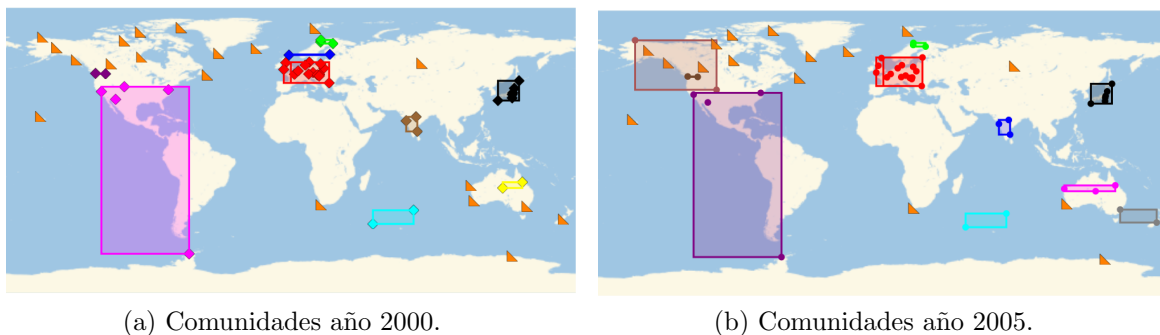


Figura 6.1: Comunidades de magnetómetros año (a) 2000 y (b) 2005, utilizando la correlación de Pearson para un umbral de 0.8. Los rombos de colores representan las distintas comunidades y los triángulos naranjos son magnetómetros en comunidades de una sola estación.

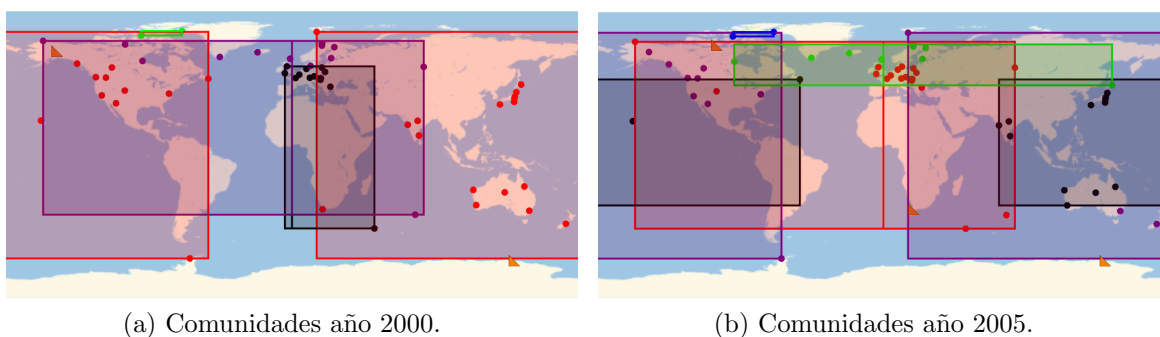


Figura 6.2: Comunidades de magnetómetros año (a) 2000 y (b) 2005, utilizando la sincronización de eventos para un umbral de 0.8. Los rombos de colores representan las distintas comunidades y los triángulos naranjos son magnetómetros en comunidades de una sola estación.

### 6.1.1. Número de comunidades y nodos

A continuación mostramos la variación en función del tiempo del número total de comunidades, del número de comunidades de una sola estación, del número de

comunidades de dos o más estaciones, del total de nodos dentro de las comunidades, del número promedio de nodos dentro de las comunidades y del máximo de nodos dentro de una comunidad para ambas medidas de similitud, CP y SE. Las figuras (Figs. 6.3–6.5) muestran los resultados para dichas métricas en las curvas azules, mientras que las curvas rojas representan el promedio anual de manchas solares (Fig. 2.1), el índice *Dst* (Fig. 3.5) y el promedio anual del índice *Dst*, permitiendo comparar las métricas con la actividad geomagnética.

### 1. Resultados para un umbral de $\theta = 0.85$ .

La Fig. 6.3 muestra los resultados para  $\theta = 0.85$ .

Se puede observar, en primer lugar, que ambas medidas de similitud, CP y SE, dan resultados parecidos, de modo que, al menos para este umbral, la elección de dicha medida de similitud no afecta fuertemente los resultados.

Luego, notamos que el número total de comunidades tiende a aumentar en períodos de alta actividad geomagnética, consistente con nuestra suposición de que en dichos períodos las mayores perturbaciones magnéticas se traducen en mayor fragmentación del campo. A su vez, el número de magnetómetros aislados aumenta cerca del máximo solar, lo cual es consistente con la observación anterior.

A pesar de lo anterior, el número de comunidades con más de una estación no presenta una correlación tan clara con la actividad geomagnética, sugiriendo que la tendencia en el número total de comunidades, se debe esencialmente al aumento de los magnetómetros aislados.

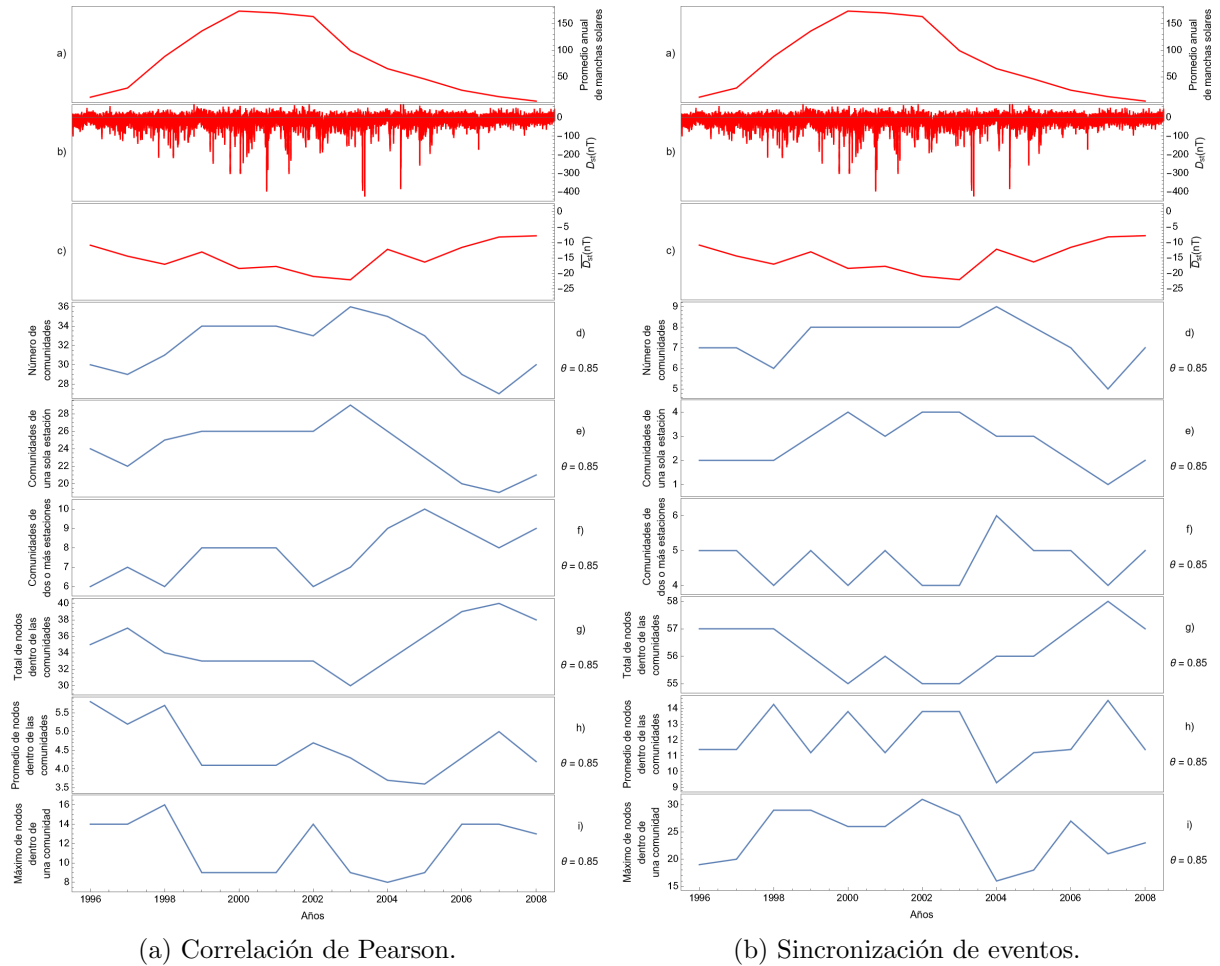


Figura 6.3: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice  $D_{st}$  y en **c**) promedio anual del índice  $D_{st}$ . (Curva azul) En **d**) número de comunidades, en **e**) comunidades de una sola estación, en **f**) comunidades de dos o más estaciones, en **g**) total de nodos dentro de las comunidades, en **h**) promedio de nodos dentro de las comunidades y en **i**) máximo de nodos dentro de una comunidad.

Como una manera de cuantificar el tamaño de las comunidades, hemos calculado también el número de estaciones dentro de comunidades de más de una estación, el promedio de estaciones por comunidad de más de una estación, y el máximo número de estaciones en las comunidades.

Se observa para estas métricas que ambas medidas de similitud arrojan resultados parecidos. De estas métricas, la que muestra una mejor correlación con la actividad geomagnética es el número total de estaciones en comunidades, que parece decrecer durante el máximo solar. En cambio, ni el promedio de estaciones por comunidad ni el máximo de estaciones por comunidad entregan una información parecida, aunque es interesante que, en el caso del método CP, se aprecia una disminución de la métrica en dos períodos que coinciden con dos períodos en que el índice *Dst* y el promedio anual de manchas solares muestran valores extremos (alrededor de los años 2001 y 2004).

## 2. Resultados para un umbral $\theta = 0.8$ .

La Fig. 6.4 es análoga a la anterior, Fig. 6.3, pero con un umbral menor,  $\theta = 0.8$ .

Los resultados son similares que para el mayor umbral: hay más comunidades, y es mayor también el número de magnetómetros aislados, cuando la actividad geomagnética es mayor, resultado que se aprecia para ambos métodos de similitud. Este aumento de los magnetómetros aislados es suficiente para explicar el mayor número de comunidades, ya que, como muestra la Fig. 6.4, no se aprecia una correlación clara entre el número de comunidades con más de una estación y el índice *Dst* y el promedio anual de manchas solares. Esto, al

menos, en promedio, ya que es interesante destacar el aumento del número de comunidades de dos o más estaciones en la Fig. 6.4 (a), con dos peaks en torno a los años 2001 y 2004, que muestran menores valores del  $Dst$  y valores altos para el promedio anual de manchas solares.

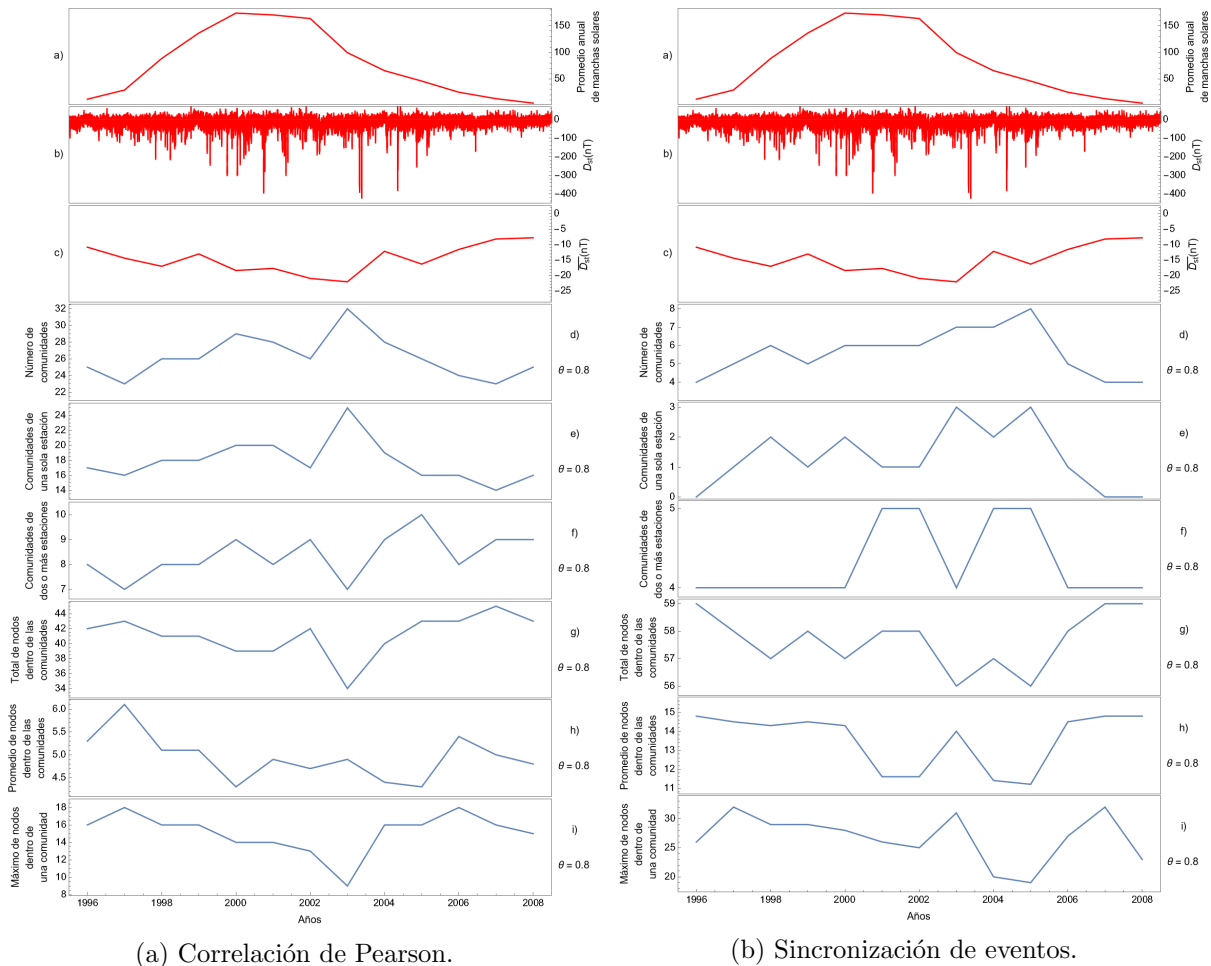


Figura 6.4: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice  $Dst$  y en **c**) promedio anual del índice  $Dst$ . (Curva azul) En **d**) número de comunidades, en **e**) comunidades de una sola estación, en **f**) comunidades de dos o más estaciones, en **g**) total de nodos dentro de las comunidades, en **h**) promedio de nodos dentro de las comunidades y en **i**) máximo de nodos dentro de una comunidad.

En cuanto al tamaño de las comunidades en términos del número de estaciones contenidas, el total de estaciones dentro de comunidades, el promedio de estaciones y el número máximo de estaciones.

Los resultados muestran que, para este umbral más bajo, estas métricas de tamaño sí exhiben una mejor correlación con la actividad geomagnética, a diferencia de lo que ocurría con  $\theta = 0.85$  (Figs. 6.3), en particular para el total de estaciones y el promedio de estaciones.

### 3. Resultados para un umbral $\theta = 0.7$ .

Por último, reducimos el umbral un poco más, a  $\theta = 0.7$  y obtenemos la Fig. 6.5.

A diferencia de los resultados anteriores, para este umbral las correlaciones, si existen, son mucho menos claras, y a veces incluso de signo opuesto entre sí al comparar ambos métodos de similitud (p. ej., ver el máximo de nodos dentro de una comunidad).

Esto sugiere que, aunque los métodos de similitud sean adecuados para determinar conexiones en la red de magnetómetros, la elección de un umbral demasiado bajo puede estropear las correlaciones. Algo similar sucede si se usa un umbral demasiado alto (el umbral máximo que hemos mostrado en estas figuras es  $\theta = 0.85$ , pero también hemos realizado los cálculos con umbrales mayores), ya que en ese caso son muy pocas las conexiones que se preservan en la red, haciendo que la estadística sea más débil, y las fluctuaciones sean mayores.

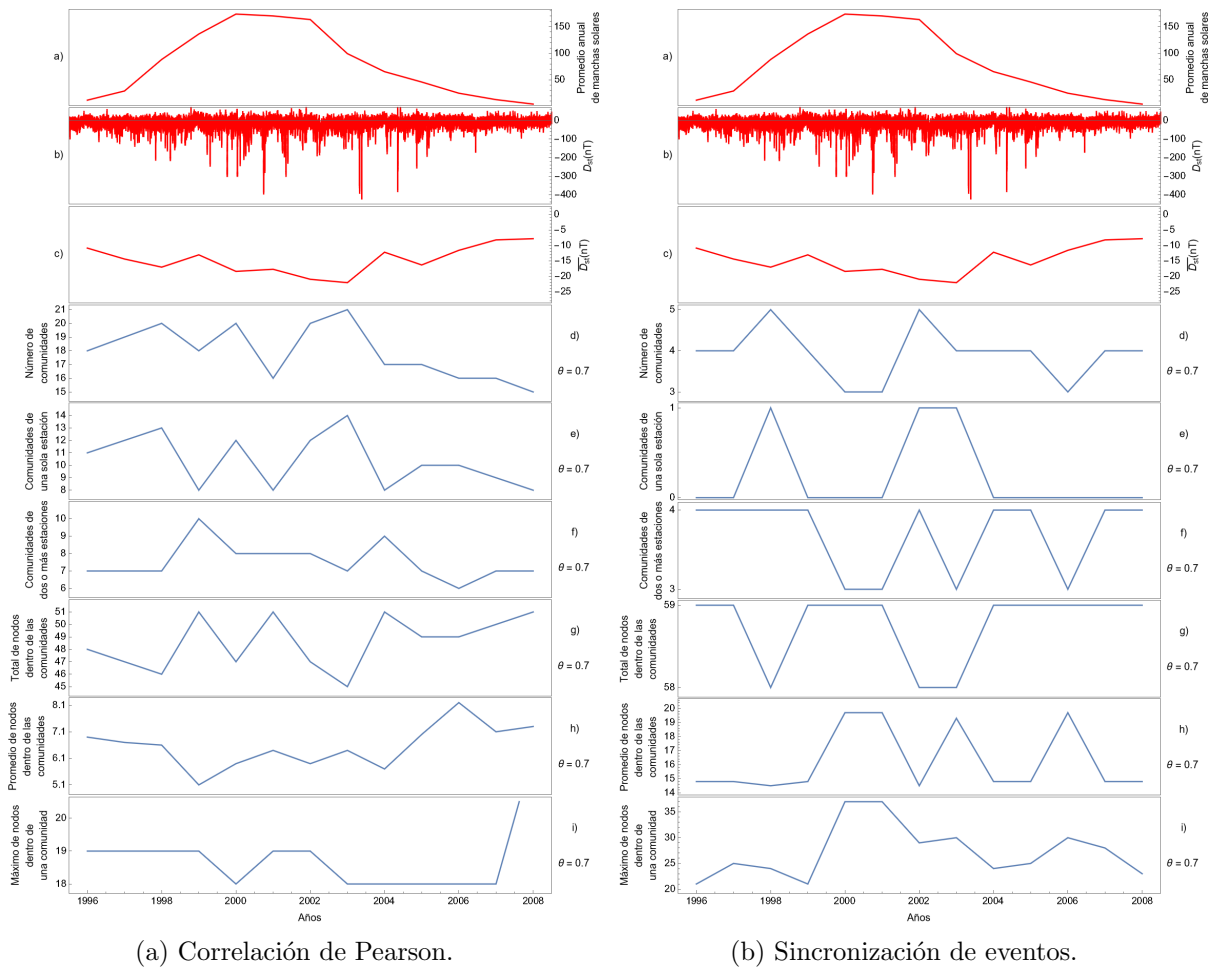


Figura 6.5: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

De estos tres umbrales, los que muestran una mejor relación de estas métricas con el promedio anual de manchas solares, el índice  $Dst$  y el promedio anual del índice  $Dst$  son  $\theta = 0.8$  y  $0.85$ . Con ellos, podemos observar que hay más comunidades, más

comunidades de una sola estación y menos estaciones en comunidades en los años de alta actividad magnética. A la inversa, en años de baja actividad, hay menos comunidades, menos estaciones aisladas y el número de estaciones en comunidades aumenta (Figs. 6.3 y 6.4). Estas observaciones son independientes del método de similitud usado, y son consistentes con la idea de que el estado del campo magnético terrestre es más ordenado en períodos de calma geomagnética, y más desordenado cuando las perturbaciones aumentan en número e intensidad.

Las otras métricas, como el número de comunidades de dos o más estaciones, el promedio de estaciones por comunidad, y el máximo número de estaciones en comunidades, no muestran correlaciones tan claras en general. Respecto al número de estaciones con dos o más estaciones, esto sugiere que el cambio en el número de comunidades se debe principalmente al número de estaciones aisladas.

También se puede observar una relación inversa entre las comunidades de dos o más estaciones con el promedio de estaciones dentro de una comunidad, cuando aumenta una disminuye la otra (Figs. 6.3 y 6.4).

Recordemos que los magnetómetros no están distribuidos homogéneamente en la superficie terrestre (Fig. 5.1). La densidad de magnetómetros aumenta en el hemisferio norte, específicamente en Europa y en el norte de América, y disminuye en el hemisferio sur. Esto induce un sesgo inherente en las medidas estudiadas. Sin embargo, es interesante que, a pesar de eso, las correlaciones no se alteran significativamente por el sesgo de densidad de nodos, estas parecen tener el signo correcto.



### 6.1.2. Área de las comunidades

A continuación mostramos los resultados obtenidos para otra medida del tamaño de las comunidades: la estimación del área promedio (Fig. 6.6).

La distribución no uniforme de los magnetómetros (Fig. 5.1) puede provocar que magnetómetros muy separados pueden quedar en la misma comunidad por un efecto estadístico, sin que eso signifique necesariamente que la comunidad tenga un área tan grande. Sin embargo, los resultados, a pesar de eso, son sugerentes.

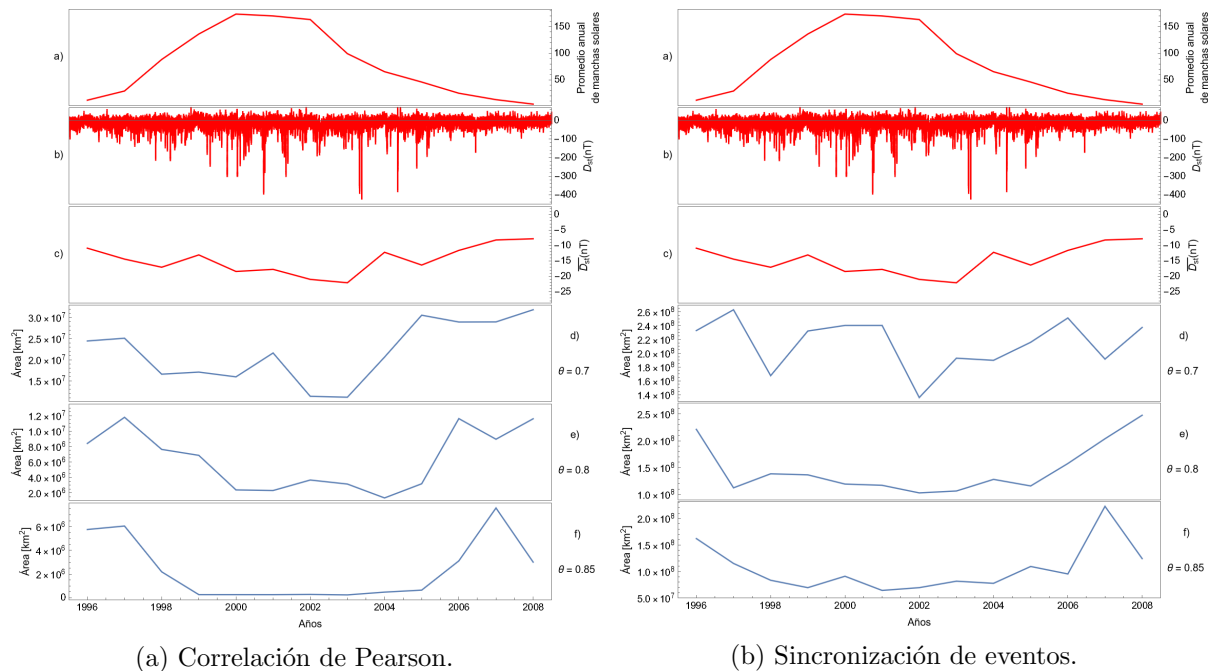


Figura 6.6: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) Área promedio de las comunidades en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

Para el umbral más alto,  $\theta = 0.85$ . Se observa una tendencia a que la extensión geográfica de las comunidades disminuya en períodos de alta actividad magnética, lo que es consistente con el mayor “desorden” de la configuración del campo, y con los resultados de la sección 6.1.1. Esto no parece depender del método de similitud empleado, aunque las variaciones son más claras cuando se usa la correlación de Pearson.

Para el umbral  $\theta = 0.8$ , los resultados son similares a los obtenidos con  $\theta = 0.85$ : el área cubierta por las comunidades disminuye cerca del máximo solar, observándose para ambos métodos de similitud, habiendo mayor claridad con el método CP.

El umbral  $\theta = 0.7$  muestra que, cuando se reduce aun más el umbral, la correlación observada anteriormente se sigue observando, aunque, en este caso, es sólo apreciable cuando se usa el método CP.

Es interesante notar que, a diferencia de lo presentado en la sección 6.1.1, sí se obtienen resultados razonables, consistentes con lo esperado, para  $\theta = 0.7$ , lo cual sugiere que el área de las comunidades puede ser una métrica menos sensible a la elección del umbral que las métricas basadas en el número de comunidades y de nodos.

### 6.1.3. Modularidad

A continuación mostramos la variación del valor de la modularidad  $M$  en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. 6.7).

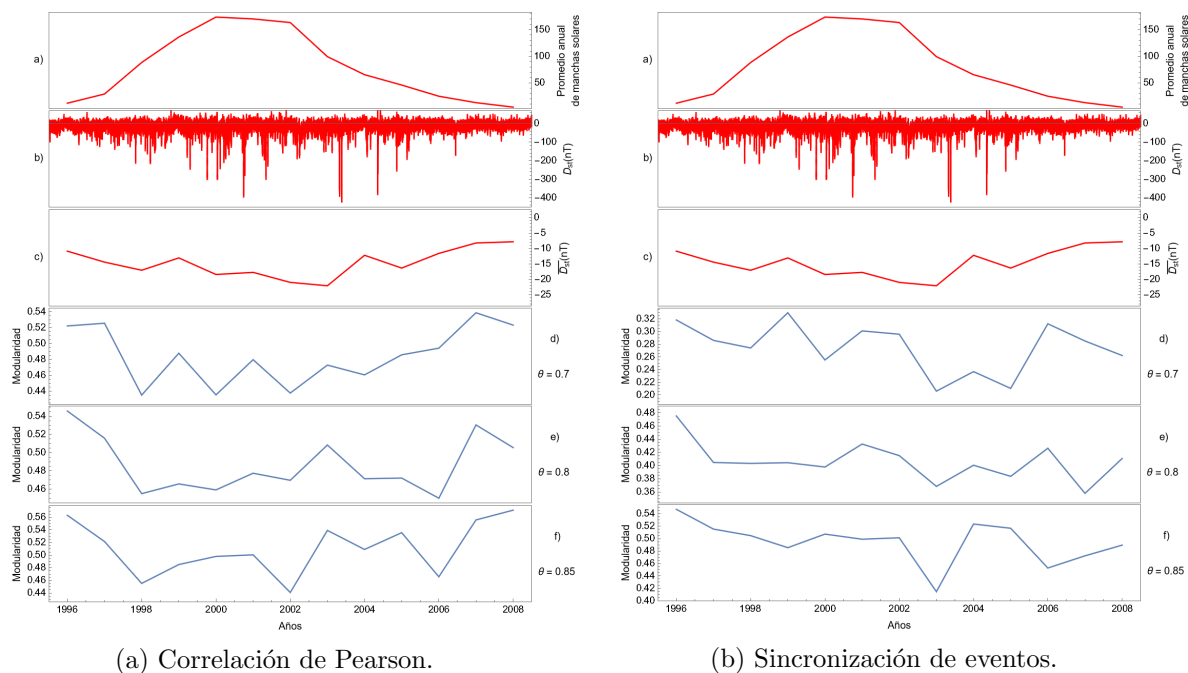


Figura 6.7: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice  $Dst$  y en **c**) promedio anual del índice  $Dst$ . (Curva azul) Modularidad en función del tiempo, utilizando en **d**) un umbral  $\theta = 0.7$ , en **e**) un umbral  $\theta = 0.8$  y en **f**) un umbral  $\theta = 0.85$ .

Para un umbral de  $\theta = 0.85$  se observa una tendencia de la modularidad a ser menor en torno al máximo del ciclo solar, aunque es una tendencia leve. Sin embargo, algo similar se observa con los otros umbrales,  $\theta = 0.8$  y  $\theta = 0.7$ , de modo que, aun cuando leve, parece ser un resultado robusto. Esto indica que, para esta red, las comunidades no sólo son más, y con menor área promedio en períodos de alta

actividad magnética, sino que las comunidades están más interconectadas entre sí.

Ambas medidas de similitud (CP y SE) entregan resultados similares, consistente con lo resultados anteriores. Sin embargo, a diferencia de lo observado antes, las correlaciones más claras ocurren para los umbrales menores,  $\theta = 0.8$  y  $\theta = 0.7$ .

### 6.1.4. Entropía de Shannon

La última manera de cuantificar la evolución de la estructura de comunidades será a través de la entropía de Shannon  $S$  (Fig. 6.8), basada en las probabilidades de que un magnetómetro pertenezca a una comunidad, ecuación (5.7).

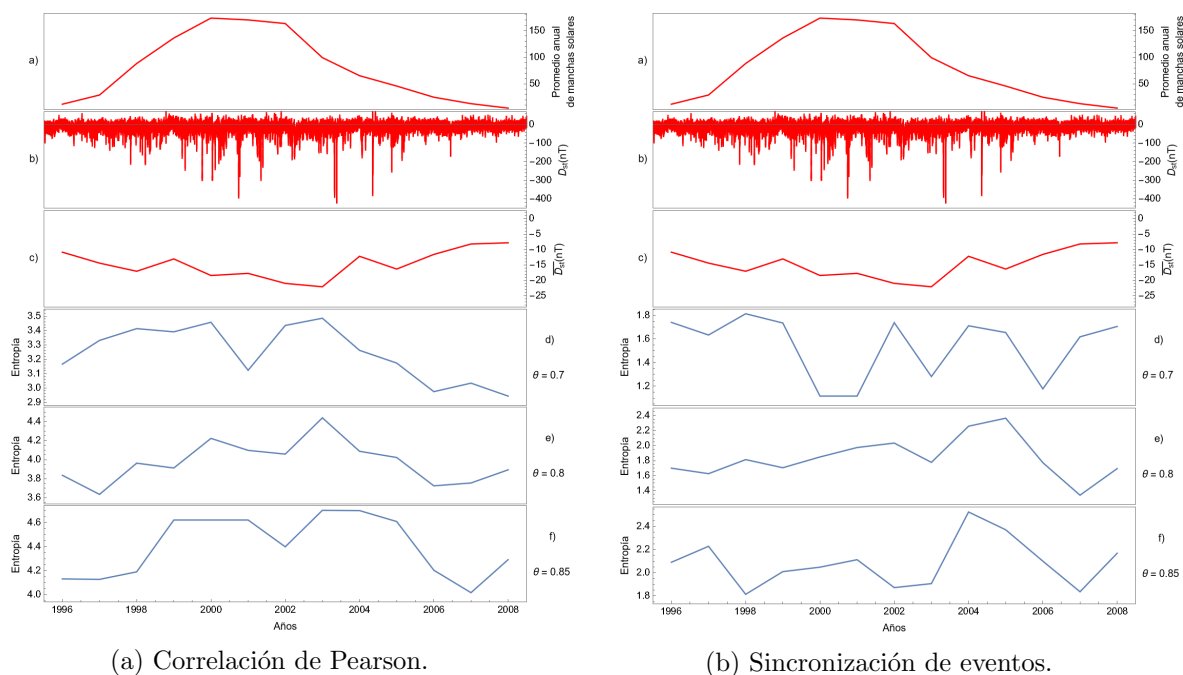


Figura 6.8: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice  $Dst$  y en **c**) promedio anual del índice  $Dst$ . (Curva azul) Entropía en función del tiempo, utilizando en **d**) un umbral  $\theta = 0.7$ , en **e**) un umbral  $\theta = 0.8$  y en **f**) un umbral  $\theta = 0.85$ .

En base a los resultados anteriores, esperaríamos una mayor entropía en períodos de mayor actividad geomagnética. Sin embargo, para los umbrales observados, esto no es claro, aunque se insinúa en algunos casos particulares. Por ejemplo, para todos los umbrales se observa un aumento de entropía cerca del año 2004, y en algunos casos también alrededor del año 2001. Las tendencias más claras de correlación con la actividad geomagnética parecen observarse en  $\theta = 0.8$ , con método SE y para  $\theta = 0.7$ , con método CP. Sin embargo, los resultados no son suficientemente consistentes para distintos umbrales y para ambos métodos, como se observó en las métricas anteriores.

## 6.2. Datos observados por los satélites

Como se observó en la sección 5.2, la falta de uniformidad en la ubicación de los satélites complica la interpretación de los resultados mostrados hasta el momento. Por ello, complementaremos con datos tomados sobre una grilla uniforme por medio de observaciones satelitales. En la sección 6.2.1 mostraremos los resultados obtenidos con el satélite CHAMP, y en la sección 6.2.2 los resultados con Ørsted.

A diferencia de lo sucedido con los datos de magnetómetros, el número de datos por año es mucho menor (ver tablas 5.1 y 5.2), el número de eventos extremos es muy pequeño, por lo tanto el método de sincronización de eventos no tiene una buena estadística. Por ello, solo utilizamos la correlación de Pearson (CP) para determinar las conexiones entre los nodos de las redes.

### 6.2.1. Datos observados por el satélite CHAMP

Recordemos que, en este caso, los datos sólo comprenden desde el año 2000 al 2008. En dicha ventana de tiempo, los mejores resultados para la correlación con el número promedio de manchas solares, índice *Dst* y el promedio anual del índice *Dst* se obtiene para los umbrales  $\theta = 0.8, 0.85$  y  $0.95$ , y son los resultados que mostraremos a continuación.

#### Número de comunidades y nodos

A continuación mostramos la variación en función del tiempo del número total de comunidades, del número de comunidades de una sola estación, del número de comunidades de dos o más estaciones, del total de nodos dentro de las comunidades,

del número promedio de nodos dentro de las comunidades y del máximo de nodos dentro de una comunidad.

En general, las figuras (Figs. 6.9–6.11) sugieren que las correlaciones, cuando existen, tienden a ser más claras a medida que el umbral  $\theta$  disminuye.

### 1. Resultados para un umbral de $\theta = 0.95$ .

Los resultados están contenidos en la Fig. 6.9.

Lo primero que observamos es que, en las diversas métricas calculadas, se ve una correlación débil con el ciclo solar como un todo. Sin embargo, al considerar las curvas en detalle, podemos observar algunas coincidencias entre el comportamiento de las métricas y la existencia de eventos más intensos en el índice *Dst* y en el promedio de manchas solares. Así, se ve una leve tendencia a que el número de comunidades y el número de comunidades de una estación aumenten cuando el índice *Dst* tiene sus menores valores y cuando aumenta el promedio de manchas solares. Esto es consistente con lo observado anteriormente para datos de magnetómetros (sección 6.1.1). Lo inverso sucede con el número de estaciones dentro de las comunidades, lo que también fue observado anteriormente para magnetómetros.

Es interesante notar que tanto el número de comunidades como el número de magnetómetros aislados muestran un claro aumento alrededor en el año 2001, coincidiendo con un período de alta actividad geomagnética. Por otra parte, todas las otras métricas tienen, en ese mismo año, una clara disminución. Esto es sugerente de que los resultados son consistentes con los resultados para mag-

netómetros, en la sección 6.1.1, pero el número de datos disponibles no permite tener una estadística suficiente para hacer afirmaciones más categóricas.

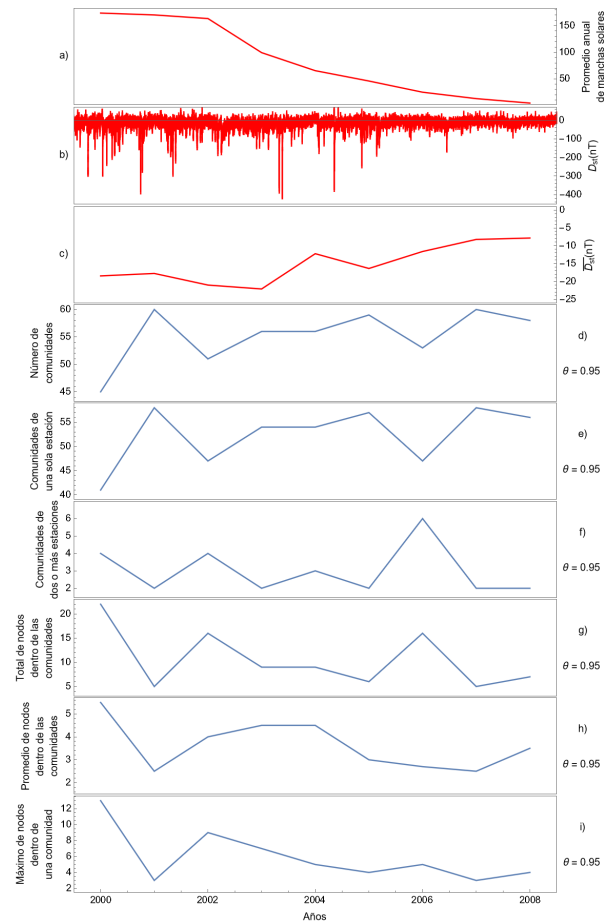


Figura 6.9: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.



## 2. Resultados para un umbral de $\theta = 0.85$ .

En este caso, los resultados son un poco más claros, como lo muestra la Fig. 6.10.

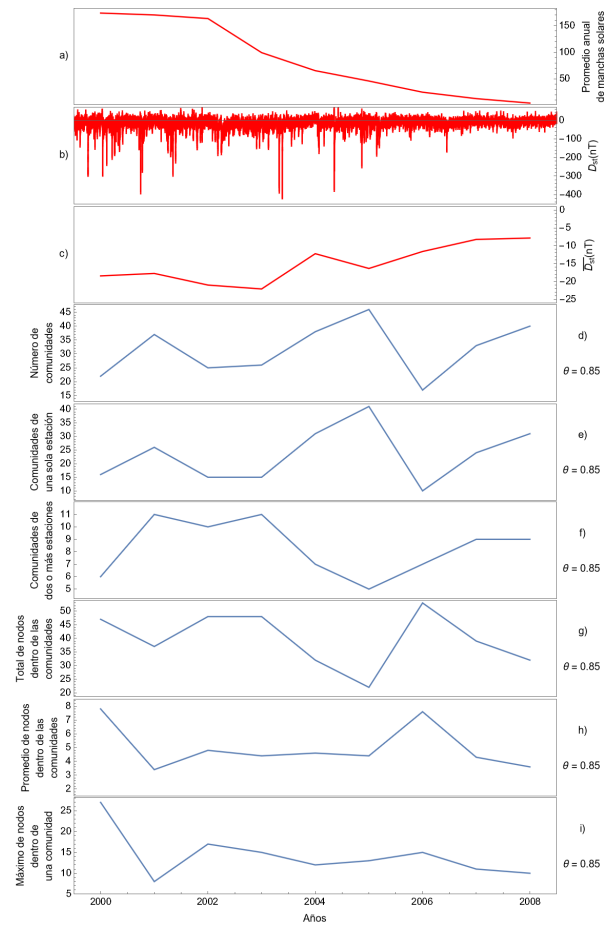


Figura 6.10: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

Una vez más, el mayor desorden de las configuraciones con mayor actividad geomagnética se aprecia en que hay más comunidades, y más comunidades de magnetómetros aislados, en los años en que el índice *Dst* y el promedio anual de manchas solares muestran sus *peaks* principales. Consistentemente, hay más estaciones en comunidades, y el promedio de estaciones dentro de cada comunidad disminuye en esos mismos años.

Las otras métricas no muestran un comportamiento claro, lo cual, en todo caso, también sucedió con los datos de magnetómetros (sección 6.1.1).

### 3. Resultados para un umbral de $\theta = 0.8$ .

Los resultados están contenidos en la Fig. 6.11.

Para este umbral es más claro que en los anteriores que, en años de mayor actividad en el índice *Dst* y en el promedio anual de manchas solares, hay más comunidades y hay más magnetómetros aislados, acompañado de una disminución de las comunidades formadas por más de una estación. Asimismo, la Fig. 6.11 muestra que el número de estaciones dentro de comunidades claramente disminuye en estos mismos años.

El promedio de estaciones en cada comunidad se mantiene relativamente bajo durante dichos años de mayor actividad.

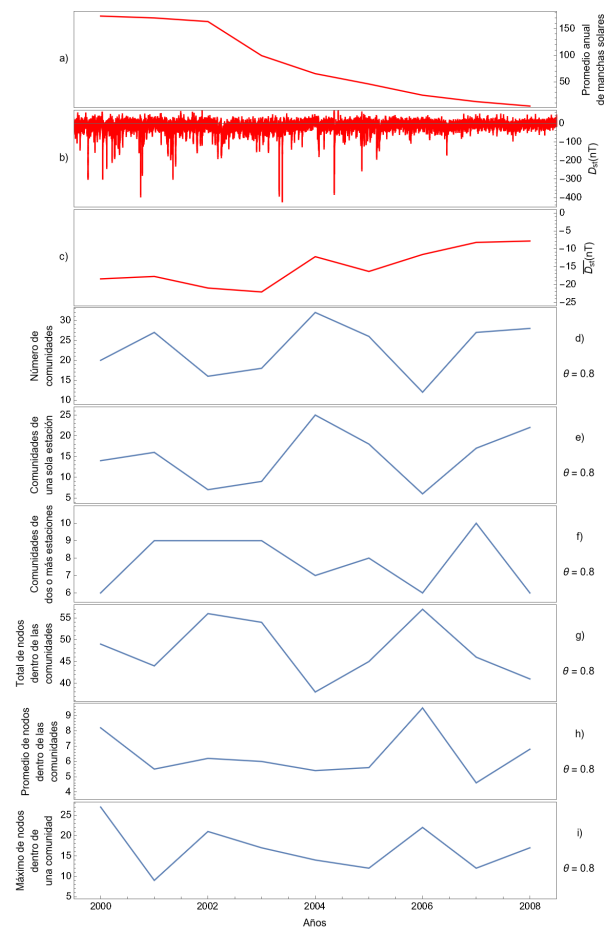


Figura 6.11: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

Podemos observar que estos resultados tienen un comportamiento similar al que se obtiene con las comunidades de las redes de magnetómetros terrestres (sección 6.1.1), o sea, para los años de alta actividad geomagnética podemos observar un aumento del número total de comunidades y del número de comunidades de una sola estación, y

una disminución del total de estaciones dentro de las comunidades, y para niveles de baja actividad geomagnética podemos observar que el número total de comunidades y el número de comunidades de una sola estación tienden a disminuir. Para estos mismos años el total de estaciones dentro de las comunidades aumenta.

De estos tres umbrales, los que muestran una mejor relación de estas métricas con el índice  $Dst$  y con el promedio anual de manchas solares son  $\theta = 0.85$  y  $0.8$ .

### Área de las comunidades

A continuación mostramos la variación del promedio de las áreas de las comunidades en función del tiempo (Fig. 6.12).

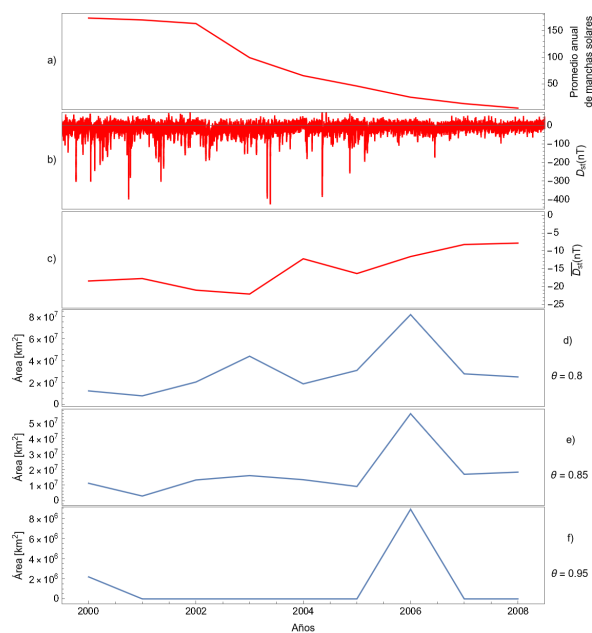


Figura 6.12: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice  $Dst$  y en **c**) promedio anual del índice  $Dst$ . (Curva azul) Área promedio de las comunidades en función del tiempo, utilizando en **d**) un umbral  $\theta = 0.8$ , en **e**) un umbral  $\theta = 0.85$  y en **f**) un umbral  $\theta = 0.95$ .

En general, las áreas se mantienen en un valor bajo en años de alta actividad magnética, con un claro aumento cuando dicha actividad alta termina, el año 2006. La tendencia del área a aumentar a medida que se acerca el mínimo solar es cada vez más clara a medida que el umbral disminuye ( $\theta = 0.8$ ). Estos resultados son consistentes con los obtenidos con magnetómetros (sección 6.1.2.)

## Modularidad

A continuación mostramos la variación del valor de la modularidad  $M$  en función del tiempo (Fig. 6.13).

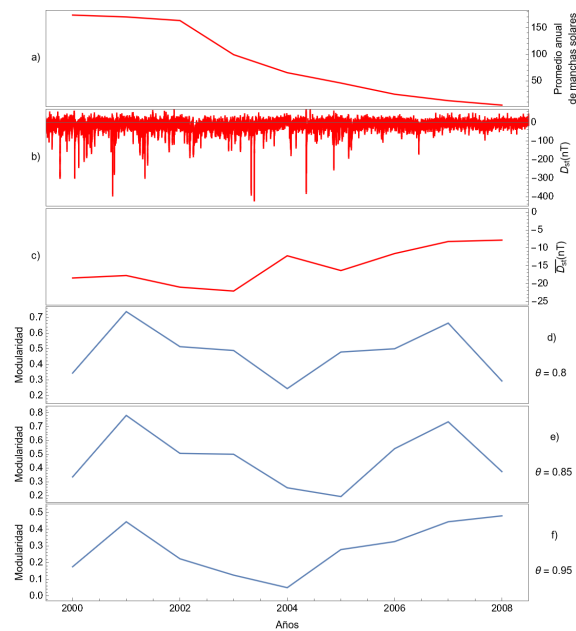


Figura 6.13: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) Modularidad en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.8$ , en **e)** un umbral  $\theta = 0.85$  y en **f)** un umbral  $\theta = 0.95$ .

Se observa que los valores de la modularidad de las comunidades de las redes construidas con las lecturas del campo magnético del satélite CHAMP tienen un comportamiento similar al de los valores de la modularidad para el caso de las comunidades de la red de magnetómetros terrestres (sección 6.1.3), es decir, los valores disminuyen en los años de alta actividad geomagnética y aumentan en los años de baja actividad.

Los resultados, en este caso, muestran que las curvas tienen una evolución muy similar para los tres umbrales.

### **Entropía de Shannon**

A continuación mostramos la variación del valor de la entropía de Shannon  $S$  de las comunidades en función del tiempo (Fig. 6.14).

En general, el valor de la entropía de las comunidades de las redes construidas con las lecturas de campo magnético del satélite CHAMP tiene un comportamiento contrario al obtenido con los magnetómetros (sección 6.1.4), es decir, los valores de  $S$  disminuyen en los años de alta actividad en el índice  $Dst$  y en el promedio anual de manchas solares.

Por tanto, no podemos extraer conclusiones claras para esta métrica.

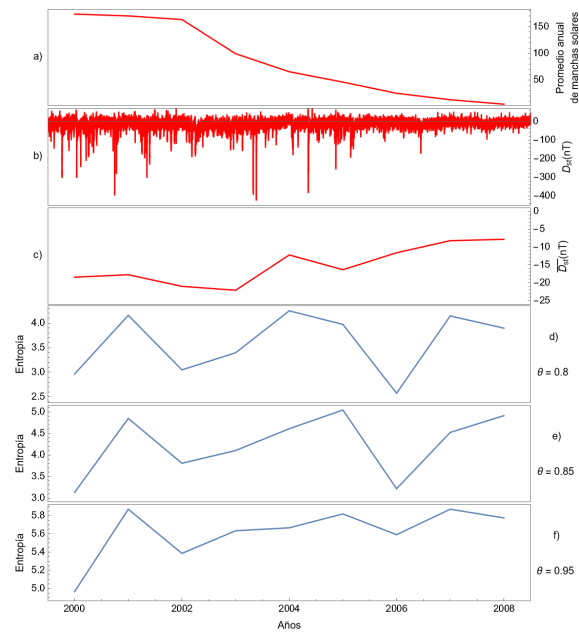


Figura 6.14: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice *Dst* y en **c)** promedio anual del índice *Dst*. (Curva azul) Entropía en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.8$ , en **e)** un umbral  $\theta = 0.85$  y en **f)** un umbral  $\theta = 0.95$ .

## 6.2.2. Datos observados por el satélite Ørsted

En esta sección mostraremos resultados análogos a los de la sección 6.2.1, pero para datos obtenidos por el satélite Ørsted. En este caso, los datos disponibles corresponden al intervalo entre los años 1999 y 2005. Las métricas usadas son las mismas que en dicha sección.

Los datos no comprenden todo el ciclo solar, lo que complica buscar correlaciones con la evolución completa de la actividad solar, pero sí es posible comparar con períodos de alta o baja actividad geomagnética, en escalas de tiempo menores.

### Número de comunidades y nodos

#### 1. Resultados para un umbral de $\theta = 0.95$ .

Los resultados se encuentran en la Fig. 6.15.

Para este umbral, que es bastante alto, el comportamiento de las comunidades no muestra una correlación muy clara con la actividad geomagnética. Por ejemplo, el número de comunidades aumenta en torno al año 2004, lo cual es consistente con los resultados de secciones anteriores, pero por otra parte hay menores valores en torno al año 2000, donde también hay una mayor actividad en el índice *Dst* y en el promedio anual de manchas solares.

El número de comunidades de más de una estación tiene también un claro máximo en el 2004, pero nada especial sucede en torno al año 2000.



Por último, el número de estaciones promedio por comunidad muestra comportamientos opuestos en ambos años.

Por tanto, no podemos extraer conclusiones claras para este umbral.

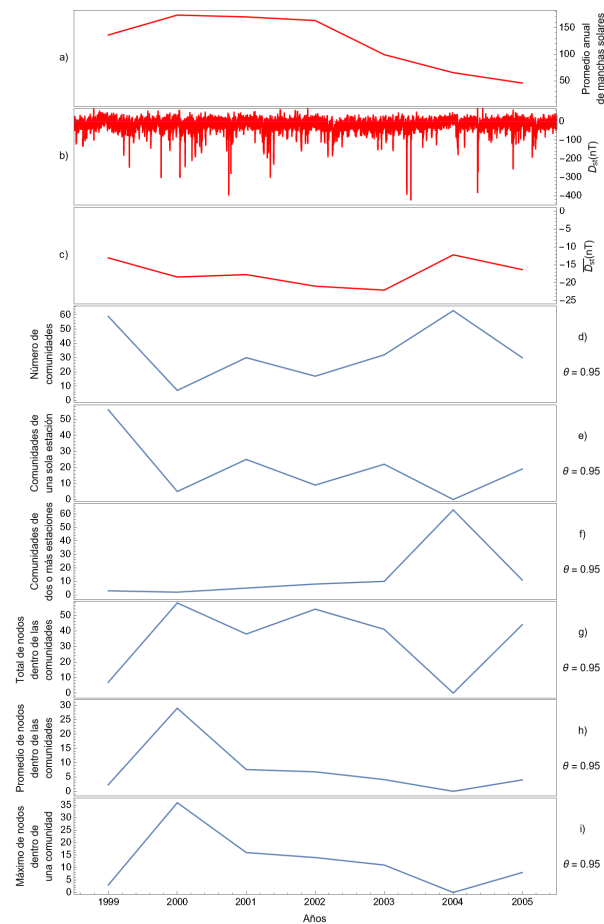


Figura 6.15: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

## 2. Resultados para un umbral de $\theta = 0.9$ .

Los resultados se encuentran en la Fig. 6.16.

Se observa que el número total de comunidades y el número de comunidades de una sola estación tienden a aumentar en los años 2001 y 2004, coincidiendo con un período de alta actividad en el índice *Dst* y en el promedio de manchas solares, lo cual es consistente con los resultados de secciones anteriores.

El número de comunidades de más de una estación muestra un comportamiento opuesto a los resultados de secciones anteriores.

Por último, el total de estaciones dentro de las comunidades, el promedio de estaciones en las comunidades y el máximo de estaciones dentro de una comunidad tienden a disminuir en los años 2001 y 2004, coincidiendo con un período de alta actividad en el índice *Dst* y en el promedio de manchas solares.

Por tanto, no podemos extraer conclusiones claras para este umbral.

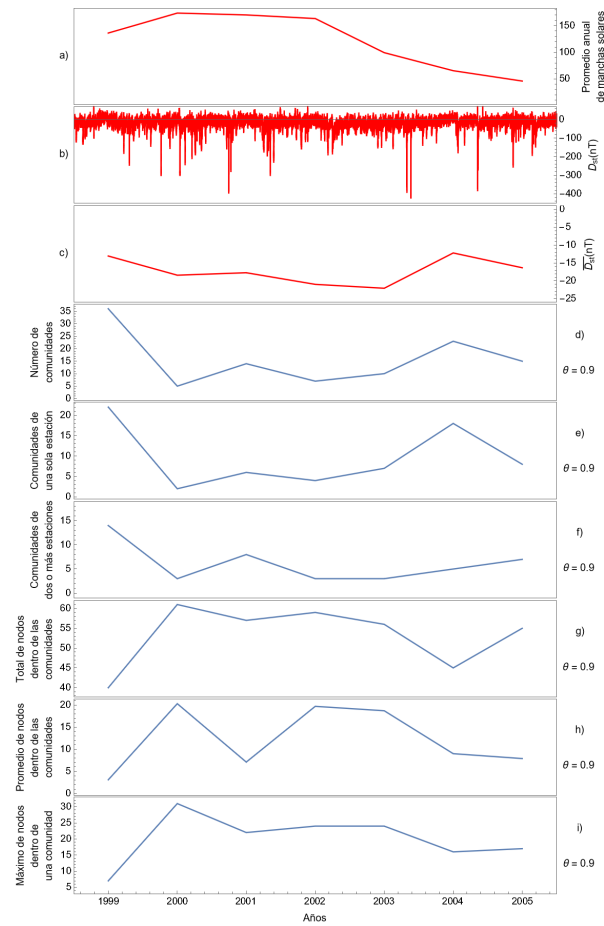


Figura 6.16: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

### 3. Resultados para un umbral de $\theta = 0.85$ .

Los resultados se encuentran en la Fig. 6.17.

Se observa, al igual que para los umbrales anteriores, un aumento en el número total de comunidades y el número de comunidades de una sola estación en los años 2001 y 2004, coincidiendo con un período de alta actividad en el índice *Dst* y en el promedio de manchas solares, lo cual es consistente con los resultados de secciones anteriores.

También se observa que el total de estaciones dentro de las comunidades, el promedio de estaciones en las comunidades y el máximo de estaciones dentro de una comunidad tienden a disminuir en los años 2001 y 2004, coincidiendo con un período de alta actividad magnética.

El número de comunidades de más de una estación tiende a mantenerse constante entre los años 2000 y 2005.

Por tanto, no podemos extraer conclusiones claras para este umbral.

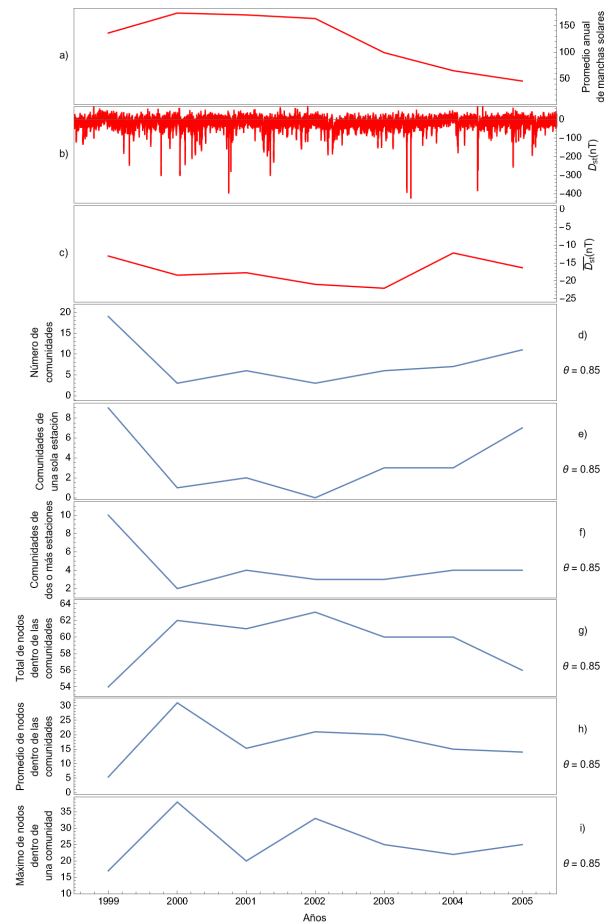


Figura 6.17: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

## Área de las comunidades

A continuación mostramos la variación del promedio de las áreas de las comunidades en función del tiempo (Fig. 6.18).

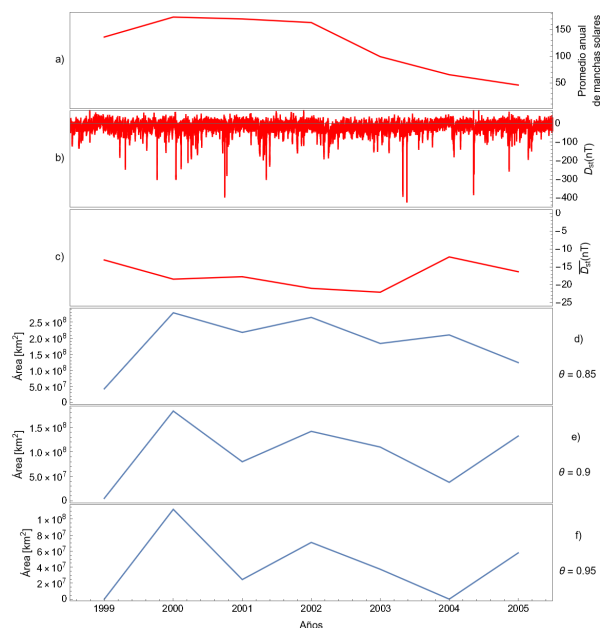


Figura 6.18: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice *Dst* y en **c)** promedio anual del índice *Dst*. (Curva azul) Área promedio de las comunidades en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.85$ , en **e)** un umbral  $\theta = 0.9$  y en **f)** un umbral  $\theta = 0.95$ .

Para los umbrales  $\theta = 0.95$  y  $0.9$ , el comportamiento del área promedio disminuye en torno a los años 2004 y 2001, lo cual es consistente con los resultados de secciones anteriores, pero por otra parte hay un crecimiento en el año 2000, donde también hay una mayor actividad magnética, siendo este un comportamiento opuesto a lo que se espera. Para el umbral  $\theta = 0.85$  el área promedio muestra un comportamiento opuesto, ya que mantiene valores altos en años de alta actividad magnética.

## Modularidad

A continuación mostramos la variación del valor de la modularidad  $M$  en función del tiempo (Fig. 6.19).

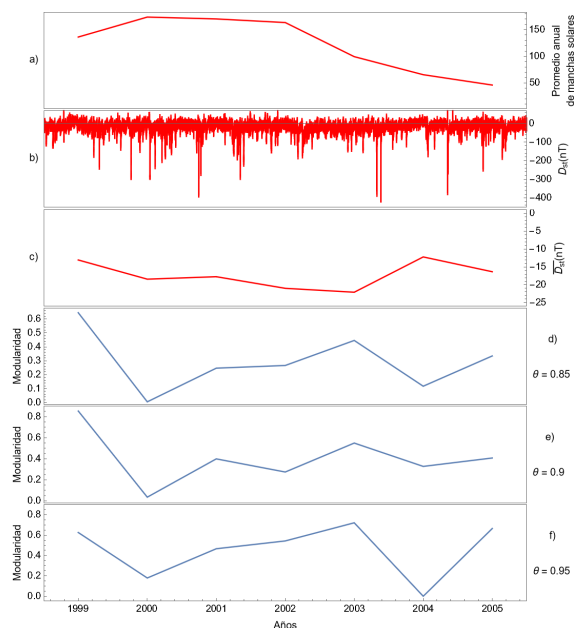


Figura 6.19: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice  $Dst$  y en **c**) promedio anual del índice  $Dst$ . (Curva azul) Modularidad en función del tiempo, utilizando en **d**) un umbral  $\theta = 0.85$ , en **e**) un umbral  $\theta = 0.9$  y en **f**) un umbral  $\theta = 0.95$ .

En general, el valor de la modularidad de las comunidades de las redes construidas con las lecturas de campo magnético del satélite  $\text{\O}rsted$  tienen un comportamiento contrario al de las secciones anteriores, es decir, los valores aumentan en los años de alta actividad en el índice  $Dst$  y en el promedio anual de manchas solares.

Por tanto, no podemos extraer conclusiones claras para esta métrica.

## Entropía de Shannon

A continuación mostramos la variación del valor de la entropía de Shannon  $S$  de las comunidades en función del tiempo (Fig. 6.20).

En general, el valor de la entropía de las comunidades de las redes construidas con las lecturas de campo magnético del satélite Ørsted tienen un comportamiento contrario al obtenido con los magnetómetros terrestres (sección 6.1.4), es decir, los valores de  $S$  disminuyen en los años de alta actividad magnética.

Por tanto, no podemos extraer conclusiones claras para esta métrica.

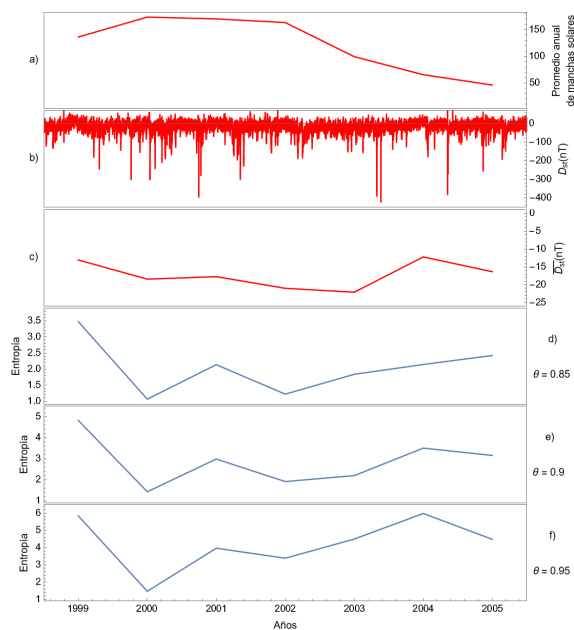


Figura 6.20: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) Entropía en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.85$ , en **e)** un umbral  $\theta = 0.9$  y en **f)** un umbral  $\theta = 0.95$ .



# Capítulo 7

## Conclusión

La ciencia de la complejidad se basa en el uso efectivo de datos de observación extensos para estudiar propiedades dinámicas y estadísticas. El campo magnético terrestre (CMT) es un ejemplo de un sistema abierto a gran escala en la naturaleza y los estudios de sus propiedades basados en datos han brindado aspectos destacados para comprender su complejidad.

En este trabajo se propone un nuevo enfoque, que consiste en estudiar la variación del CMT durante el ciclo solar 23 utilizando técnicas de redes complejas. La principal motivación para llevar a cabo esta tarea se relaciona con los avances de la ciencia de la complejidad, y diferentes estudios mencionados en el Cap. 1 referentes a la complejidad existente en la dinámica del CMT.

Hemos realizado una abstracción de las lecturas del CMT de diversos magnetómetros distribuidos sobre la superficie terrestre durante el ciclo solar 23 en lenguaje de redes complejas, al considerar que un cambio en la intensidad CMT debido al viento solar se puede manifestar en la estructura de la red de magnetómetros terrestres. Por esta razón, se ha aplicado el formalismo de redes complejas, específicamente

la detección de comunidades, para estudiar las variaciones espacio-temporales del CMT. Esto se traduce en calcular distintas métricas y estudiar si se correlacionan con la variación del índice *Dst* durante el ciclo solar 23 y, por tanto, pueden contener información acerca de la dinámica del CMT.

Se consideraron ventanas temporales de un año entre 1996 y 2008, lo que da un total de 13 años. Se construyeron redes para cada año donde los nodos (magnetómetros) representan puntos del espacio, y sus conexiones representan el grado de similitud (se utilizaron 2 medidas de similitud y 10 umbrales de correlación) entre las series temporales observadas en dichos puntos (el procedimiento se describe en el Cap. 5). Una vez construidas las redes, para cada una de ellas se determinó la estructura de comunidades utilizando el algoritmo de Louvain, el cual maximiza la modularidad, y luego se calcularon las métricas descritas en el Cap. 5.

Hemos encontrado que la elección de umbrales de correlación y de métodos de correlación, brindan algunos resultados interesantes sobre la red de magnetómetros terrestres.

La elección del umbral de correlación influye significativamente en la estructura de comunidades y su evolución con el ciclo solar 23. Esto se puede observar en cómo varían las métricas calculadas con los distintos umbrales y su relación con la variación del índice *Dst*. Si bien la elección del umbral  $\theta$  es arbitraria, se puede observar que el umbral  $\theta = 0.8$  es el que presenta mejores resultados: la variación de las métricas utilizando este umbral siguen la tendencia esperada, o sea, se correlacionan con la variación del índice *Dst*, lo que indica que sí existe una correlación entre la evolución de las comunidades detectadas y la actividad geomagnética. En efecto, niveles

mayores de actividad geomagnética inducen menores correlaciones espaciales entre puntos distantes. Esto se ve reflejado en que, en años de alta actividad geomagnética, aumenta el número total de comunidades y el número de comunidades de una sola estación, y disminuye el total de estaciones dentro de las comunidades. Para estos mismos años, el área promedio de las comunidades disminuye.

Por el contrario, en períodos de baja actividad geomagnética, hay una mayor correlación entre puntos distantes, lo que implica que hay mayor número de nodos (magnetómetros) en una comunidad, el número de comunidades disminuye, y su área aumenta.

Estas transiciones, entre un estado con muchos pequeños grupos de magnetómetros no correlacionados (alta actividad magnética) y pocos grupos mejor correlacionados entre sí (baja actividad magnética), se puede cuantificar a través de la medida de modularidad  $M$ . En efecto, en períodos de baja actividad geomagnética  $M$  es grande. Esto indica la presencia de una partición fuerte de la red en comunidades, donde las comunidades están densamente vinculadas dentro de ellas, pero con escasas vinculaciones entre comunidades. Por su parte, para niveles mayores de actividad geomagnética el valor de  $M$  es pequeño. Esto nos quiere decir que existen comunidades con muy pocas conexiones internas y muchas conexiones entre comunidades (estructura multipartita).

También cuantificamos estas variaciones a través de la entropía, como una medida del “desorden” en la estructura de comunidades.

Si el incremento de entropía  $S$  es positivo, esto sugiere que existe un mayor “desorden” en la red, o sea, existe un mayor número de nodos aislados que no per-

tenecen a una comunidad, y esto ocurre cuando hay un mayor nivel de actividad geomagnética. En cambio, cuando hay un menor nivel de actividad geomagnética el valor de la entropía  $S$  disminuye, es decir, los nodos son menos “desordenados”, por lo que aumenta la probabilidad de que un nodo sí pertenezca a una comunidad, aumentando el promedio de estaciones dentro de las comunidades y disminuyendo el número de comunidades de una sola estación (estaciones aisladas).

Con respecto a la elección del método de similitud entre series de tiempo, se debe tener en cuenta que la correlación de Pearson (CP) toma todos los puntos de datos existentes y cuantifica la magnitud de una relación funcional lineal entre dos series de tiempo, mientras que la sincronización de eventos (SE) solo considera los eventos extremos y cuantifica la magnitud de una relación no lineal entre dos series de tiempo.

Dado que la CP es una medida instantánea de la sincronía global, esta no proporciona información sobre el orden temporal entre las señales. Por otro lado, SE clasifica automáticamente dos eventos espacio-temporalmente distintos como sincronizados o no dependiendo solo de las distancias internas entre eventos. Esta sí revela información sobre el orden temporal entre las señales.

A pesar de lo anterior, en nuestro trabajo hemos encontrado que la elección de un método de similitud determinado no afecta significativamente los resultados.

Si bien el estudio de los magnetómetros terrestres sugiere que existe una relación entre la estructura de comunidades y la actividad geomagnética, las conclusiones pueden verse afectadas por la distribución de las estaciones (nodos de la red), la cual no es uniforme (ver Fig. 5.1). Por esto, para complementar los resultados, realizamos

un estudio similar pero con las lecturas de campo magnético de los satélites CHAMP y Ørsted, las cuales están distribuidas uniformemente sobre la superficie terrestre (ver Fig. 5.3).

Al igual que con los resultados obtenidos con los magnetómetros terrestres, los resultados obtenidos con los datos del satélite CHAMP indican que niveles mayores de actividad geomagnética inducen menos correlaciones entre distintos puntos y que niveles de baja actividad geomagnética inducen un mayor número de correlaciones entre los diversos puntos.

Aunque los datos del satélite CHAMP no cubren la totalidad del ciclo solar 23, se puede observar que dentro del rango de años, del 2000 al 2008, si existe una correlación entre la variación de las métricas calculadas sobre las comunidades de sus redes y el índice *Dst*, esta correlación se hace más clara a medida que el umbral  $\theta$  disminuye.

Por lo tanto, como existen tendencias similares entre ambos resultados, entonces los datos del satélite CHAMP sí corroboran los resultados obtenidos con los datos de los magnetómetros terrestres, indicando que estos resultados son útiles para describir la variación del campo magnético terrestre durante el ciclo solar 23, aun cuando la distribución de los magnetómetros terrestres no es uniforme.

Por otra parte, los resultados obtenidos con los datos del satélite Ørsted no presentan tendencias similares con los resultados de los magnetómetros terrestres y con los del satélite CHAMP. Esto se puede deber a la baja densidad de datos dentro de las áreas donde se realizan los promedios sobre una ventana temporal para definir los puntos de observación. Y también a que el período cubierto por el satélite Ørsted

corresponde a la zona del máximo solar, provocando que no se observe una clara variación entre años de alta actividad y baja actividad magnética. Lo anteriormente descrito incide en que las métricas calculadas no presenten correlaciones claras con el índice  $Dst$ , y por lo tanto los resultados obtenidos con los datos del satélite Ørsted son menos útiles.

En resumen en este trabajo :

1. Introdujimos un método novedoso para estudiar las variaciones del CMT, usando métodos de similitud entre series de tiempo de magnetómetros, y a partir de ahí construir las redes complejas.
2. El análisis de redes complejas y la detección de comunidades permite estudiar la evolución del CMT.
3. Los resultados obtenidos dependen del umbral de correlación que se utiliza para la construcción de la red.
4. No hay gran diferencia entre usar la correlación de Pearson y la sincronización de eventos.

En general, los resultados encontrados en este trabajo, al igual que en [9], sugieren que utilizar técnicas de redes complejas ofrece una perspectiva interesante en la investigación de la dinámica del CMT y del ciclo solar, en este caso a través de los estudios de comunidad.

Existen varias maneras en que estos resultados se pueden mejorar y extender. Por una parte, en este trabajo consideramos ventanas de 1 año, sin traslape entre ellas,

lo cual lleva a que la evolución de las métricas es calculada con poca resolución. Una versión más refinada de este trabajo debería considerar ventanas traslapadas, a fin de obtener curvas suavizadas, lo que permitiría apreciar mejor las variaciones con el ciclo solar.

Un segundo factor es que nuestro análisis ha estado restringido al ciclo solar 23, y además ha utilizado una manera particular de construcción de las redes complejas. Por ello, posibles mejoras en esta investigación pueden obtenerse analizando otros ciclos solares (ciclo 24), buscando tanto comportamientos universales o variaciones respecto a lo obtenido en este trabajo. Además, sería interesante considerar otros métodos para la construcción de la red, en particular otros métodos de similitud entre series de tiempo. Una posibilidad interesante es la correlación cruzada con retraso en el tiempo [55], que puede identificar la direccionalidad entre dos señales, como una relación líder-seguidor en la que el líder inicia una respuesta que el seguidor repite. Esto, a su vez, ofrece un modo natural de definir redes análogas a las que hemos trabajado, pero dirigidas.

Esperamos poder mejorar este trabajo en estas direcciones, y otras que emerjan como relevantes, en el futuro.

# Apéndice A

## Índices geomagnéticos

### Índice K

El índice K es una medición de tres horas de la variación del campo magnético de la Tierra. La medición se realiza con un magnetómetro. Esto indica la variación del flujo magnético en nanoTeslas. Esta lectura se convierte al índice K. La relación es directamente proporcional a una escala logarítmica.

El índice K se mide en 13 estaciones geomagnéticas ubicadas en diferentes lugares del mundo (ver Fig. A.3), resultando un valor particular para cada estación de medición.

Para obtener un índice K global, se calcula el promedio de los índices K de cada estación, obteniéndose así el llamado índice K<sub>p</sub> o K planetario [26, 27].

### Índice Ap

El índice Ap es análogo al índice K, pero convertido desde la escala logarítmica de K a una escala lineal, de modo que sus valores se extienden a un rango mucho más amplio. Para cada estación, se escala el índice K para dar un valor en una escala



lineal que se denomina índice «a». Esto se promedia durante el período de un día para dar el índice A. Al igual que el índice K, los valores resultantes se promedian sobre todas las estaciones para dar el índice Ap planetario. Los valores del índice Ap oscilan hasta 100 durante una tormenta y pueden subir hasta 400 en una tormenta geomagnética grave [26, 27]. En la Fig. A.2 se puede observar la variación del índice Ap entre los años 1996 y 2008.



Figura A.1: Ubicación de las 13 estaciones geomagnéticas que se utilizan para calcular el índice K.

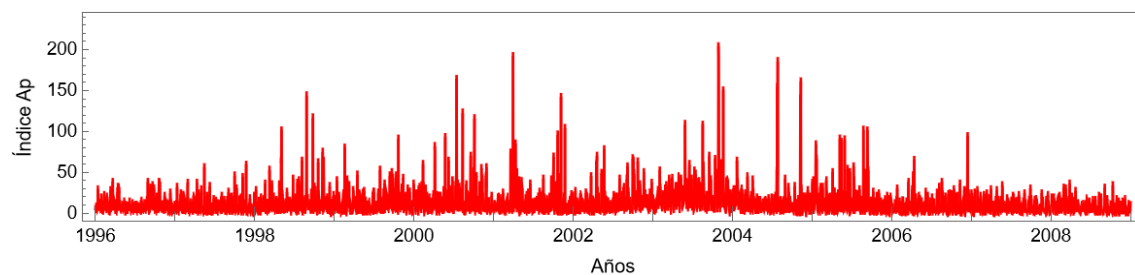


Figura A.2: Variación del índice Ap correspondiente al Ciclo solar 23.

## Índice AE

El índice Auroral Electrojet (AE) fue introducido originalmente en 1966 como una medida de la actividad global de electrojet en la zona auroral. El índice AE ahora se usa ampliamente para investigaciones en geomagnetismo, aeronomía y física solar-terrestre. El índice AE se deriva de las variaciones geomagnéticas en la componente horizontal ( $H$ ) registrado en 12 observatorios a lo largo de la zona auroral en el hemisferio norte (ver Fig. A.3) [26, 27].

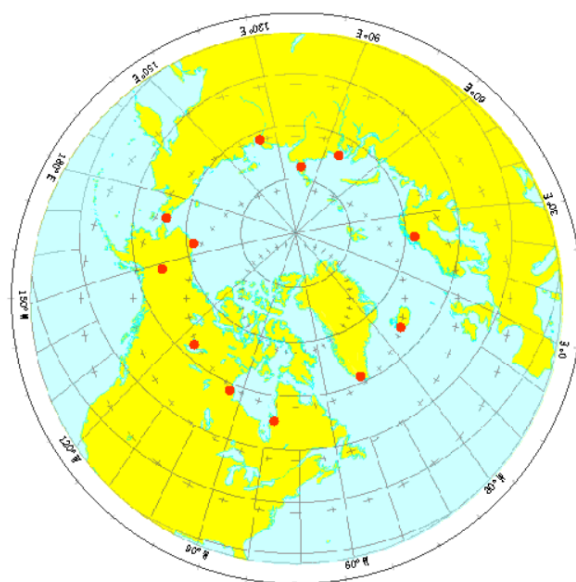


Figura A.3: Ubicación de las 12 estaciones geomagnéticas que se utilizan para calcular el índice AE. Imagen extraída de la Ref. [27]

En la Fig. A.4 se puede observar la variación del índice AE entre los años 1996 y 2008.

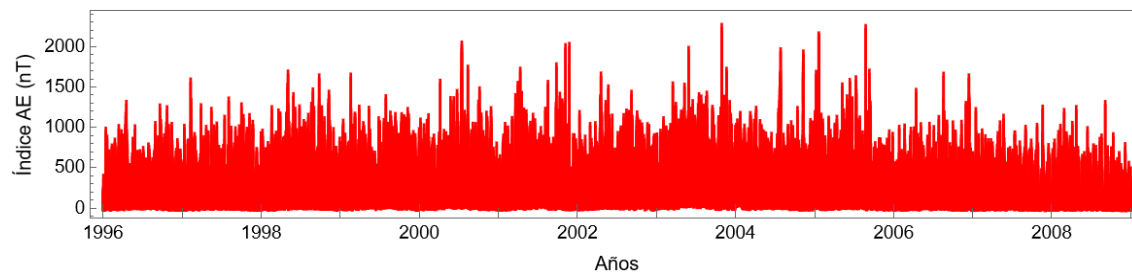


Figura A.4: Variación del índice AE correspondiente al Ciclo solar 23.

Para más información sobre los índices geomagnéticos, se pueden revisar los siguientes enlaces:

1. <https://kp.gfz-potsdam.de/en/about-kp>.
2. [https://www.ngdc.noaa.gov/stp/geomag/kp\\_ap.html](https://www.ngdc.noaa.gov/stp/geomag/kp_ap.html).
3. <https://www.ngdc.noaa.gov/stp/geomag/ae.html>.
4. <https://wdc.kugi.kyoto-u.ac.jp/kp/index.html>.

# Apéndice B

## Coeficiente de clustering

El coeficiente de clustering de un nodo  $v_i$  viene dado por:

$$c_i = \frac{2L_i}{k_i(k_i - 1)}, \quad (\text{B.1})$$

donde  $L_i$  corresponde al número de pares de vecinos de  $v_i$  que están conectados y  $k_i$  es el grado de  $v_i$ . Esta métrica representa la probabilidad de que pares de vecinos del nodo  $v_i$  estén conectados. Si se tiene un nodo  $v_i$  que tiene  $k_i$  vecinos, entonces a lo sumo pueden existir  $k_i(k_i - 1)/2$  conexiones entre ellos, lo que ocurre cuando cada vecino de  $v_i$  está conectado a todos los demás vecinos de  $v_i$ . Luego  $c_i$  es la fracción de estas conexiones permitidas que realmente existen [62]. El  $c_i$  es también una medida de densidad local. Si pensamos en una red social, entonces  $c_i$  representa la probabilidad de que un par de amigos de  $v_i$  sean amigos entre sí [42].

Es preciso notar que tanto  $L_i$  como  $k_i$  se pueden obtener a partir de la matriz de adyacencia, de modo que  $c_i$  también. Con esto se puede encontrar el coeficiente de clustering de la red, que es dado por:

$$C = \frac{1}{n} \sum_i c_i, \quad (\text{B.2})$$

el cual permite determinar en qué medida pares de vecinos de un nodo tomado al azar dentro de la red son vecinos entre sí.

A continuación mostramos la variación del valor del coeficiente de clustering  $C$  de la red en función del tiempo (Fig. B.1), para los distintos umbrales ( $\theta = 0.7, 0.8$  y  $0.85$ ), y para ambas medidas de similitud (en el Apéndice C se puede observar la comparación entre la variación del coeficiente de clustering  $C$  de la red en función del tiempo y la variación de los índices AE y Ap).

En base a los resultados anteriores (ver sección 6.1), esperaríamos un menor coeficiente de clustering en períodos de mayor actividad geomagnética, ya que el coeficiente de clustering es una medida de la proporción de vínculos débiles/fuertes en la red. Un coeficiente de clustering bajo es indicativo de una red compuesta por conexiones débiles. Sin embargo, para los umbrales observados, esto no es claro, aunque se insinúa en algunos casos particulares. Por ejemplo, para todos los umbrales se observa una disminución del coeficiente de clustering cerca del año 2003, y en algunos casos también alrededor del año 2001. Las tendencias más claras de correlación con la actividad geomagnética parecen observarse en  $\theta = 0.7$  y  $0.8$  (ver Fig. B.1), con método CP y para  $\theta = 0.7$  (ver Fig. B.1), con método SE. Sin embargo, los resultados no son suficientemente consistentes para distintos umbrales y para ambos métodos,

como se observó en las métricas anteriores (ver sección 6.1).

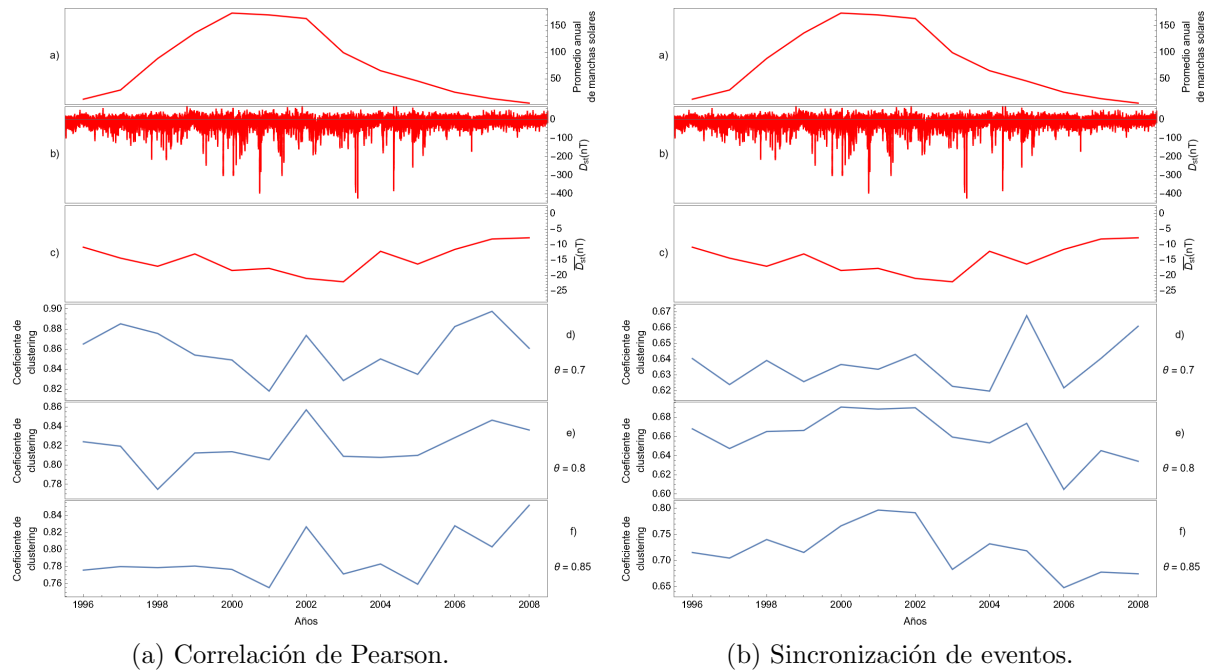


Figura B.1: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice  $Dst$  y en **c)** promedio anual del índice  $Dst$ . (Curva azul) Coeficiente de clustering en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

# Apéndice C

## Otros resultados

En esta sección mostraremos los resultados de las redes que se obtienen al aplicar los umbrales  $\theta = 0.7, 0.8$  y  $0.85$  en las matrices  $Q$  y  $r$  de los datos de los magnetómetros terrestres, los cuales son los que entregan resultados que muestran una mejor relación con la variación de los índices Ap y AE a lo largo del ciclo solar. Esto con la intención de mostrar que se puede estudiar y visualizar de otra manera la correlación entre las variaciones de las comunidades de magnetómetros y la actividad geomagnética a lo largo del ciclo solar 23 discutida en la sección 6.1.

Los índices Ap y AE, al igual que el índice  $Dst$  también son indicadores que se utilizan para estudiar la actividad geomagnética [26, 27].

En la sección C.1 mostramos la comparación entre la variación en función del tiempo de las métricas mencionadas al final del Cap. 5 con el promedio anual de manchas solares, con el índice Ap y con el promedio anual del índice Ap. Y de forma análoga en la sección C.2 utilizamos el índice AE para estudiar la variación de las métricas.

## C.1. Resultados con el índice Ap

En esta sección mostraremos resultados análogos a los de la sección 6.1, pero utilizando el índice Ap como indicador de actividad geomagnética.

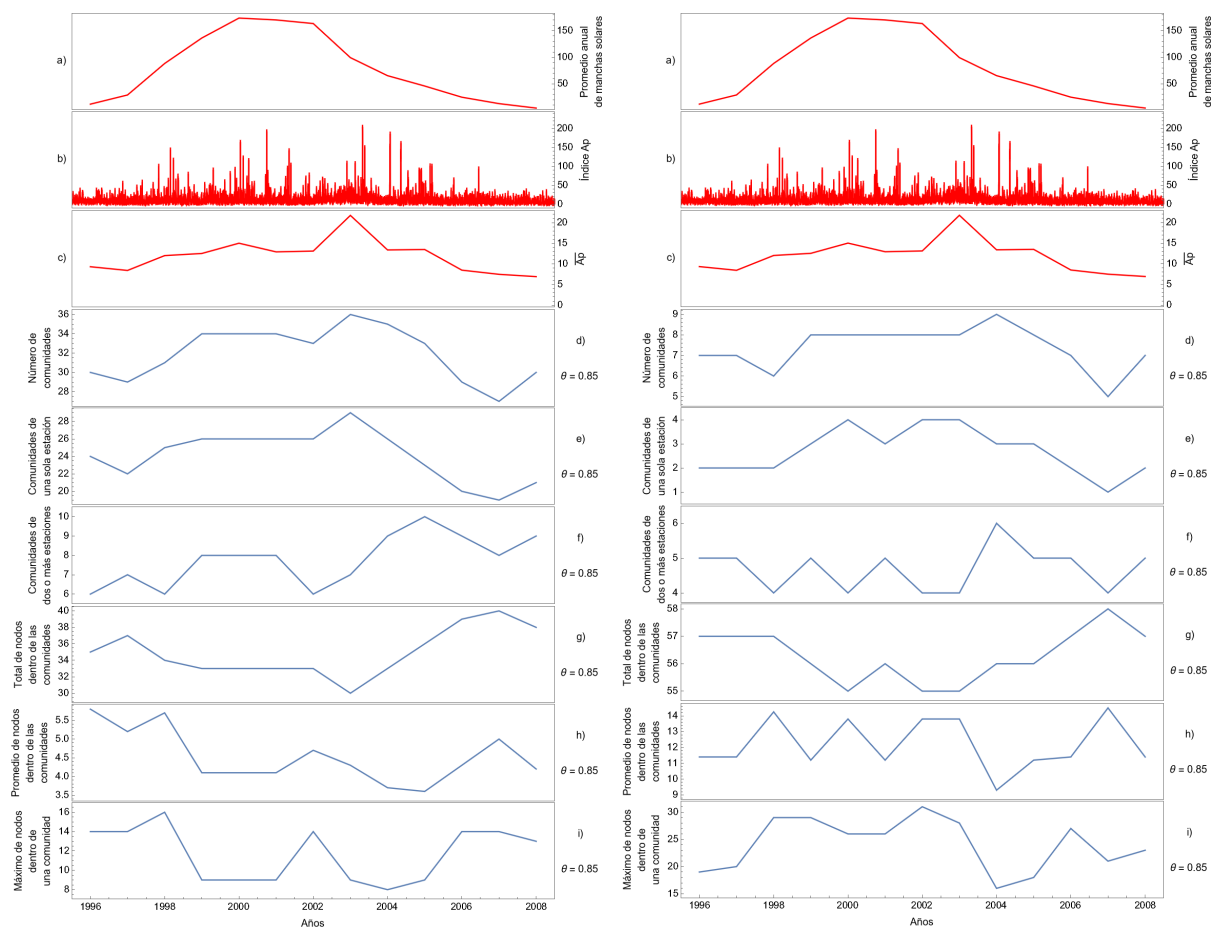
### C.1.1. Número de comunidades

A continuación mostramos la variación en función del tiempo del número total de comunidades, del número de comunidades de una sola estación, del número de comunidades de dos o más estaciones, del total de nodos dentro de las comunidades, del número promedio de nodos dentro de las comunidades y del máximo de nodos dentro de una comunidad. Las figuras (Figs. C.1–C.3) muestran los resultados para dichas métricas en las curvas azules, mientras que las curvas rojas representan el promedio anual de manchas solares (Fig. 2.1), el índice Ap (Fig. A.2) y el promedio anual del índice Ap, permitiendo comparar las métricas con la actividad geomagnética.



### 1. Resultados para un umbral de $\theta = 0.85$ .

La Fig. C.1 muestra los resultados para  $\theta = 0.85$ .



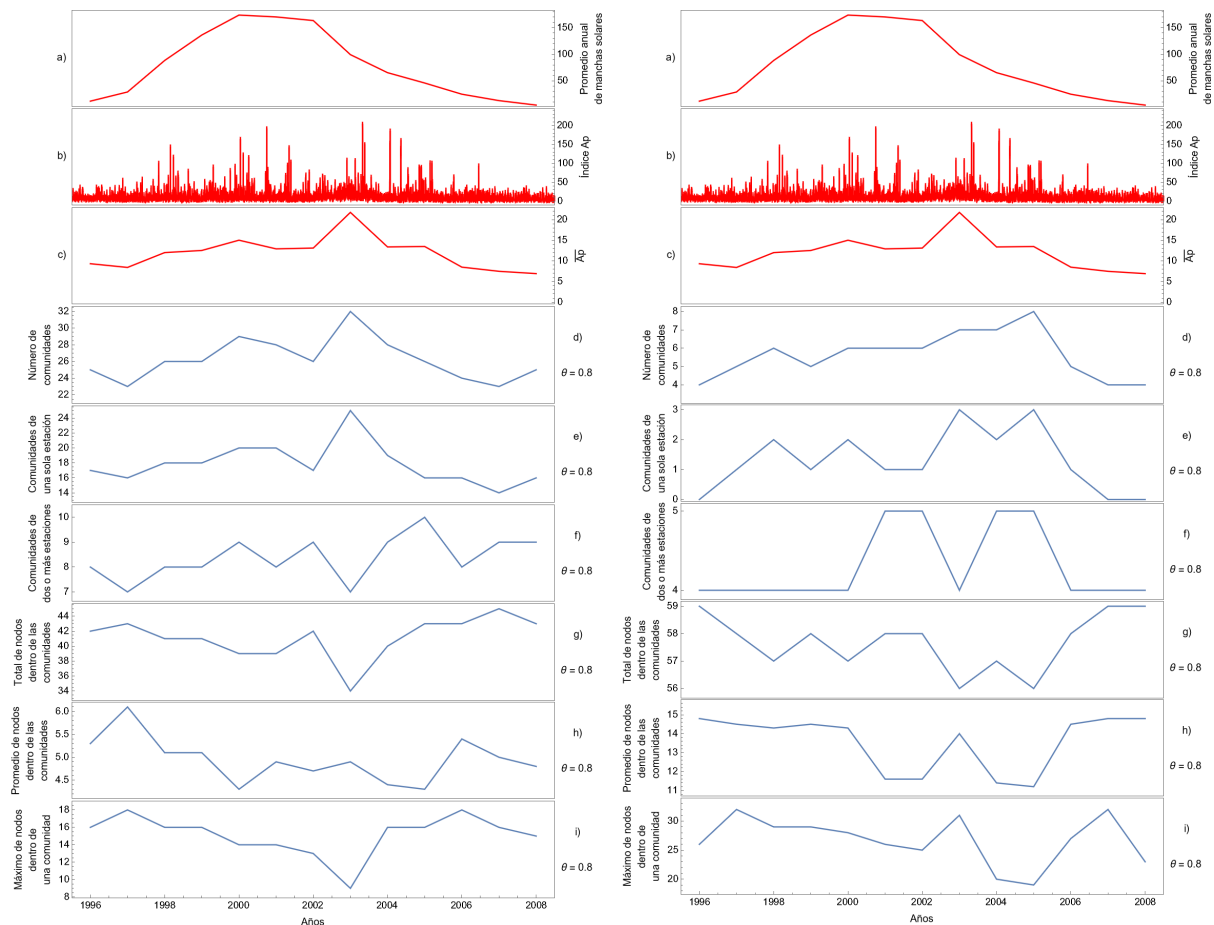
(a) Correlación de Pearson.

(b) Sincronización de eventos.

Figura C.1: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice Ap y en **c)** promedio anual del índice Ap. (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

## 2. Resultados para un umbral de $\theta = 0.8$ .

La Fig. C.2 muestra los resultados para  $\theta = 0.8$ .



(a) Correlación de Pearson.

(b) Sincronización de eventos.

Figura C.2: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice Ap y en **c**) promedio anual del índice Ap. (Curva azul) En **d**) número de comunidades, en **e**) comunidades de una sola estación, en **f**) comunidades de dos o más estaciones, en **g**) total de nodos dentro de las comunidades, en **h**) promedio de nodos dentro de las comunidades y en **i**) máximo de nodos dentro de una comunidad.

### 3. Resultados para un umbral de $\theta = 0.7$ .

La Fig. C.3 muestra los resultados para  $\theta = 0.7$ .

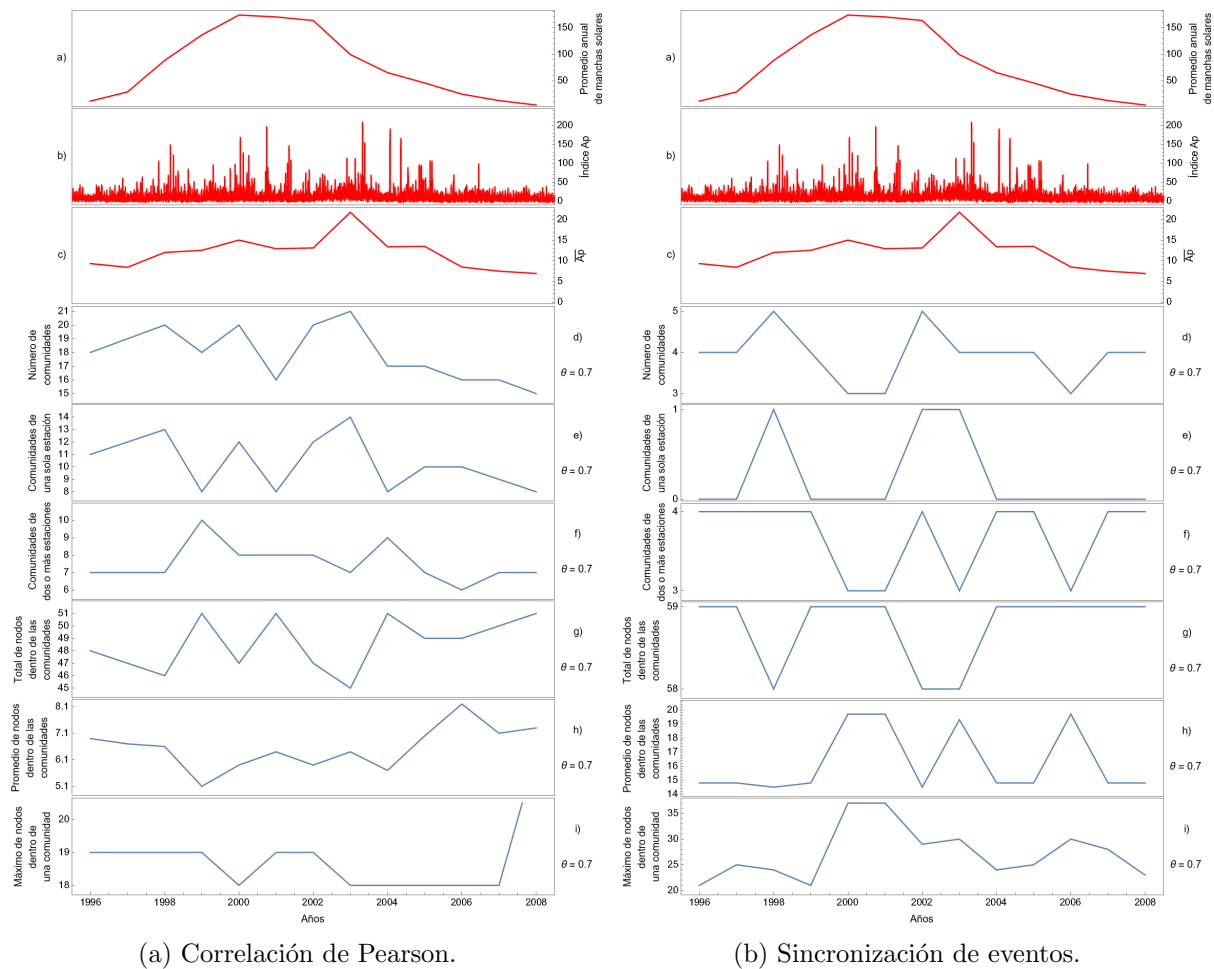


Figura C.3: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice Ap y en **c**) promedio anual del índice Ap. (Curva azul) En **d**) número de comunidades, en **e**) comunidades de una sola estación, en **f**) comunidades de dos o más estaciones, en **g**) total de nodos dentro de las comunidades, en **h**) promedio de nodos dentro de las comunidades y en **i**) máximo de nodos dentro de una comunidad.

### C.1.2. Área de las comunidades

A continuación mostramos los resultados obtenidos para otra medida del tamaño de las comunidades: la estimación del área promedio (Fig. C.4).

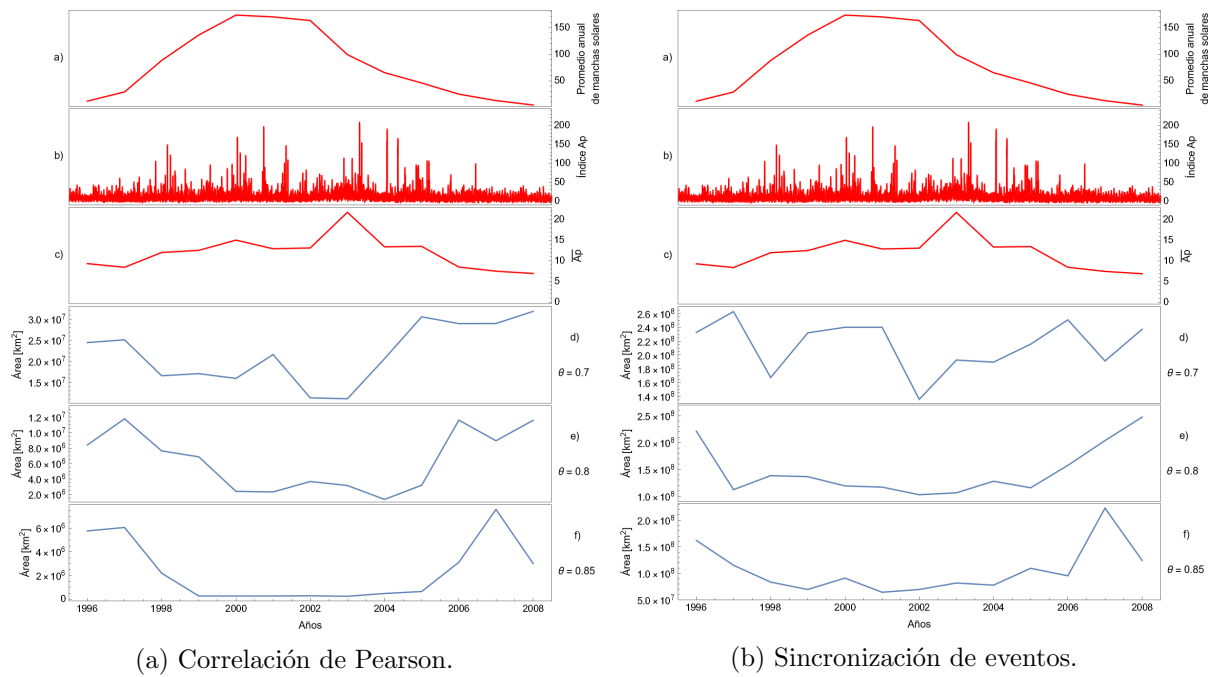


Figura C.4: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice Ap y en **c**) promedio anual del índice Ap. (Curva azul) Área promedio de las comunidades en función del tiempo, utilizando en **d**) un umbral  $\theta = 0.7$ , en **e**) un umbral  $\theta = 0.8$  y en **f**) un umbral  $\theta = 0.85$ .

### C.1.3. Modularidad

A continuación mostramos la variación del valor de la modularidad  $M$  en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. C.5).

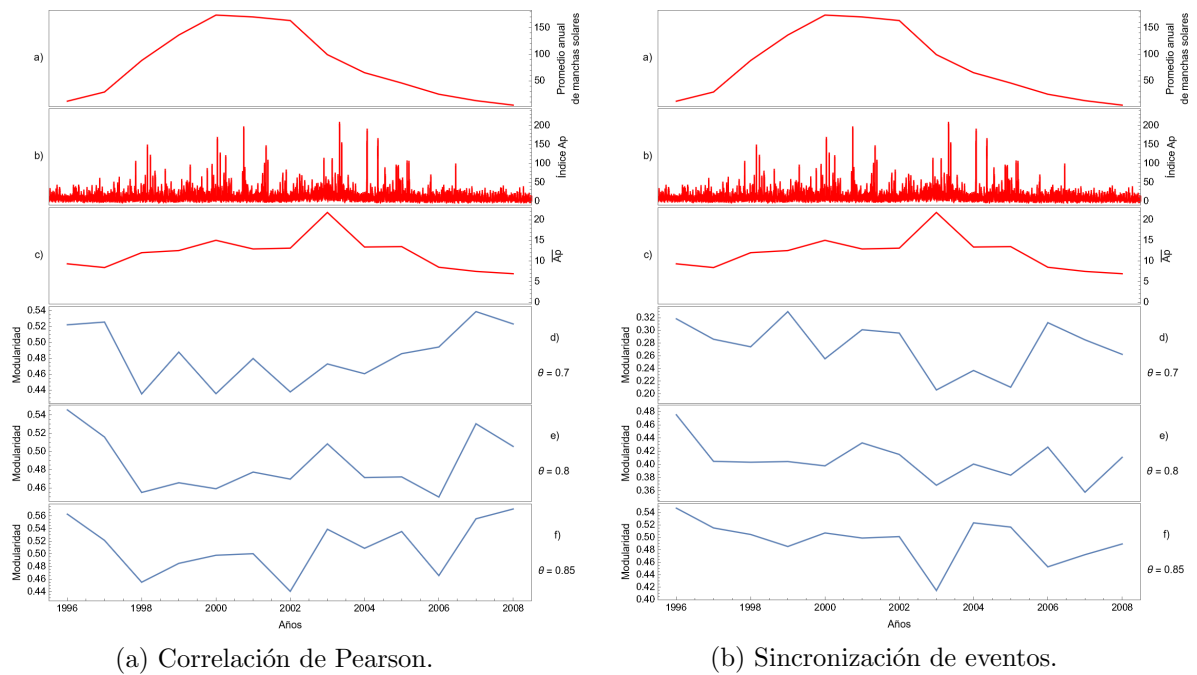


Figura C.5: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice Ap y en **c)** promedio anual del índice Ap. (Curva azul) Modularidad función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

### C.1.4. Entropía de Shannon

A continuación mostramos la variación del valor de la entropía de Shannon  $S$  de las comunidades en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. C.6).

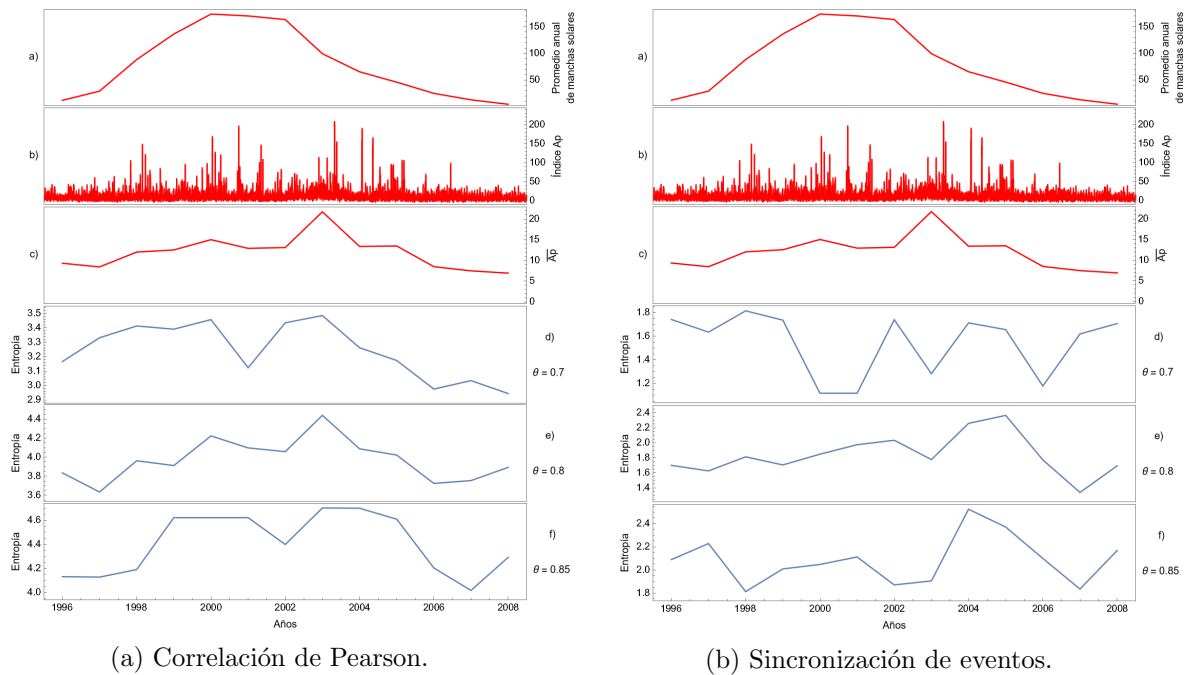


Figura C.6: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice Ap y en **c)** promedio anual del índice Ap. (Curva azul) Entropía en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

### C.1.5. Coeficiente de clustering

A continuación mostramos la variación del valor del coeficiente de clustering  $C$  de la red en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. C.7).

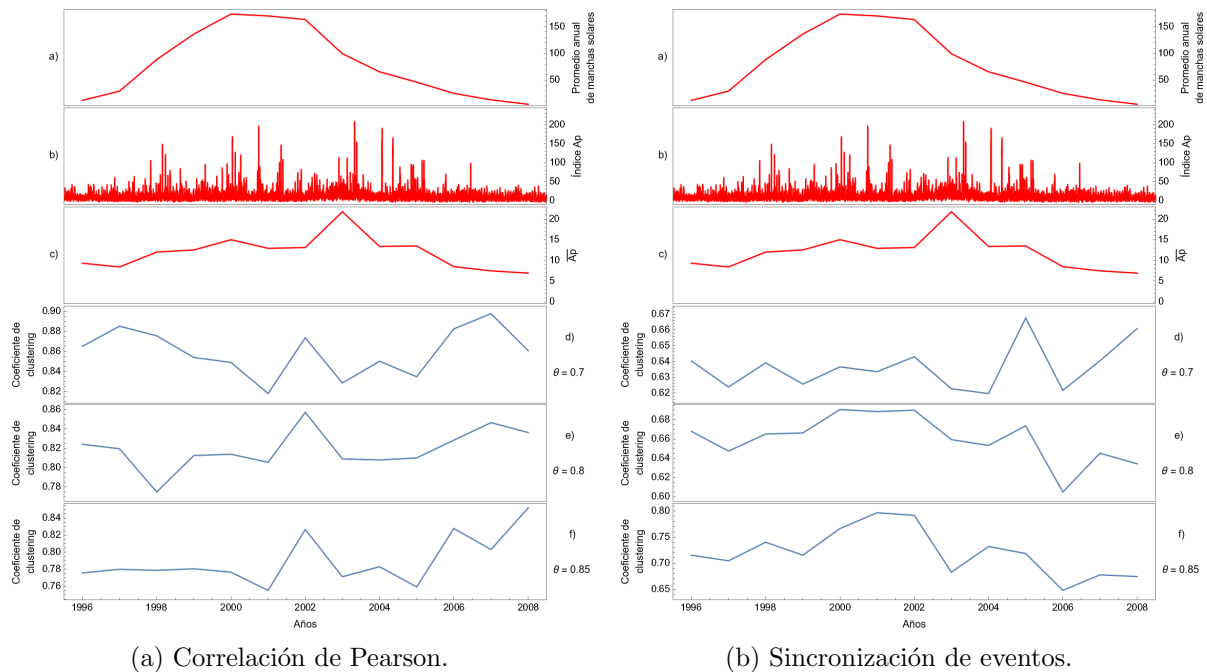


Figura C.7: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice Ap y en **c)** promedio anual del índice Ap. (Curva azul) Coeficiente de clustering en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

## C.2. Resultados con el índice AE

En esta sección mostraremos resultados análogos a los de la sección 6.1, pero utilizando el índice AE como indicador de actividad geomagnética.

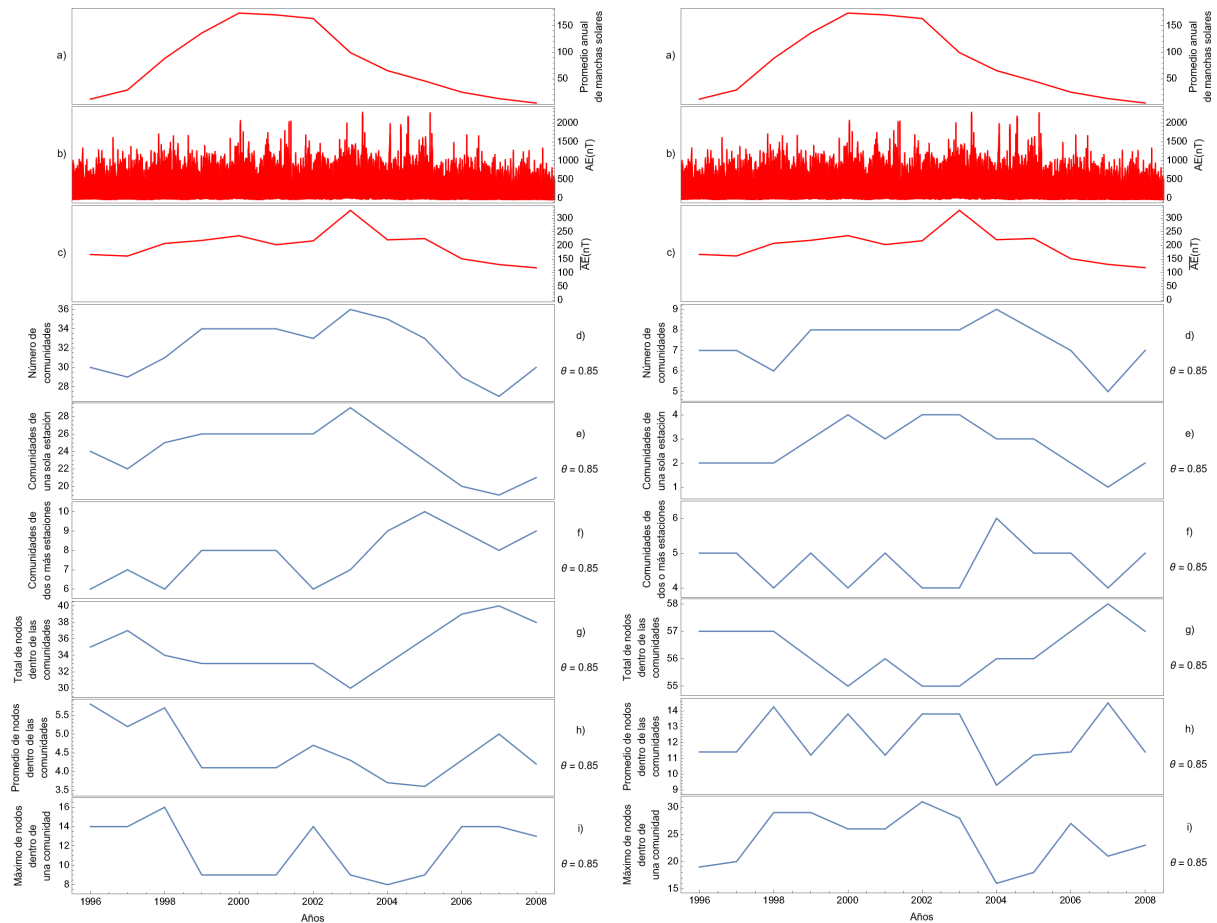
### C.2.1. Número de comunidades

A continuación mostramos la variación en función del tiempo del número total de comunidades, del número de comunidades de una sola estación, del número de comunidades de dos o más estaciones, del total de nodos dentro de las comunidades, del número promedio de nodos dentro de las comunidades y del máximo de nodos dentro de una comunidad. Las figuras (Figs. C.8–C.10) muestran los resultados para dichas métricas en las curvas azules, mientras que las curvas rojas representan el promedio anual de manchas solares (Fig. 2.1), el índice AE (Fig. A.4) y el promedio anual del índice AE, permitiendo comparar las métricas con la actividad geomagnética.



### 1. Resultados para un umbral de $\theta = 0.85$ .

La Fig. C.8 muestra los resultados para  $\theta = 0.85$ .



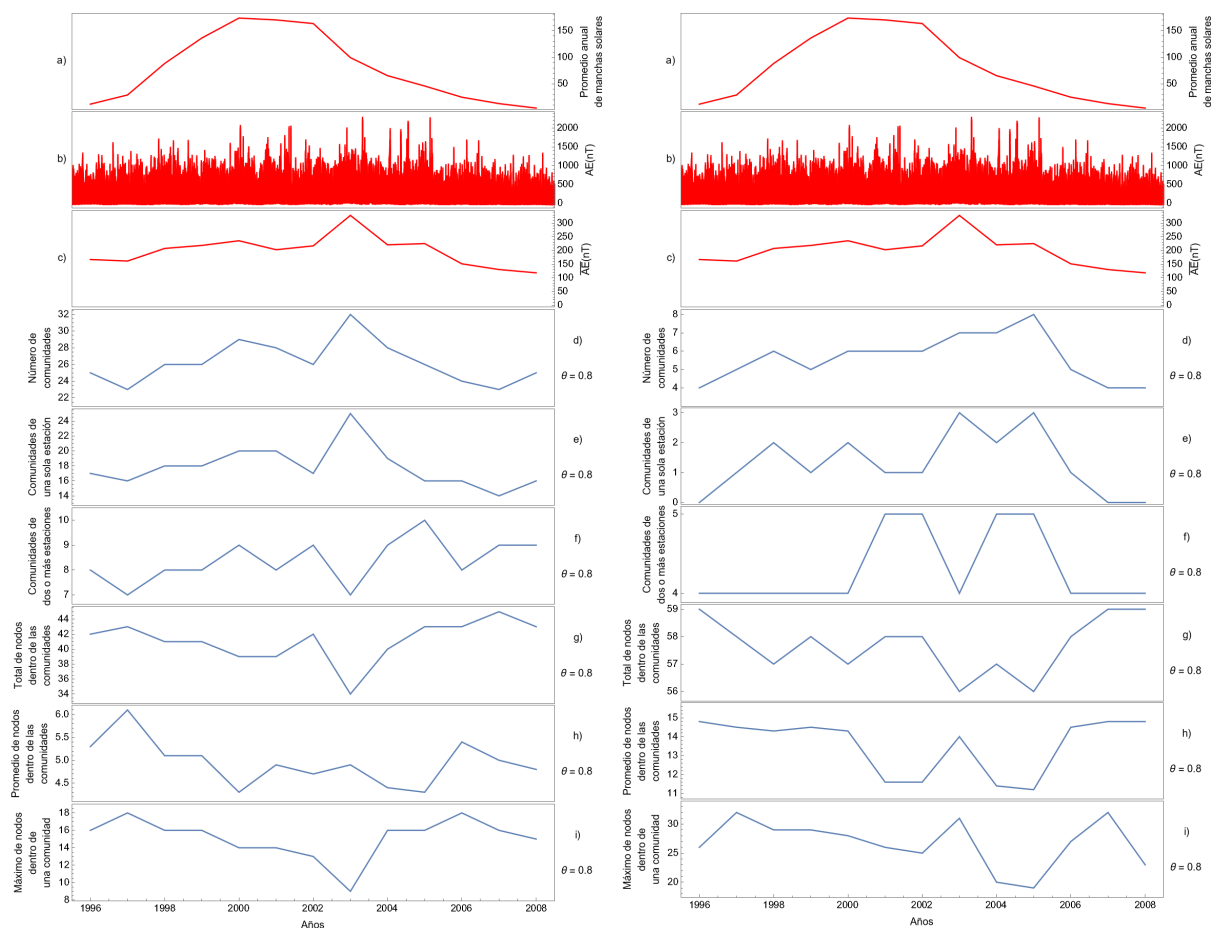
(a) Correlación de Pearson.

(b) Sincronización de eventos.

Figura C.8: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice AE y en **c)** promedio anual del índice AE. (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

## 2. Resultados para un umbral de $\theta = 0.8$ .

La Fig. C.9 muestra los resultados para  $\theta = 0.8$ .



(a) Correlación de Pearson.

(b) Sincronización de eventos.

Figura C.9: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice AE y en **c)** promedio anual del índice AE. (Curva azul) En **d)** número de comunidades, en **e)** comunidades de una sola estación, en **f)** comunidades de dos o más estaciones, en **g)** total de nodos dentro de las comunidades, en **h)** promedio de nodos dentro de las comunidades y en **i)** máximo de nodos dentro de una comunidad.

### 3. Resultados para un umbral de $\theta = 0.7$ .

La Fig. C.10 muestra los resultados para  $\theta = 0.7$ .

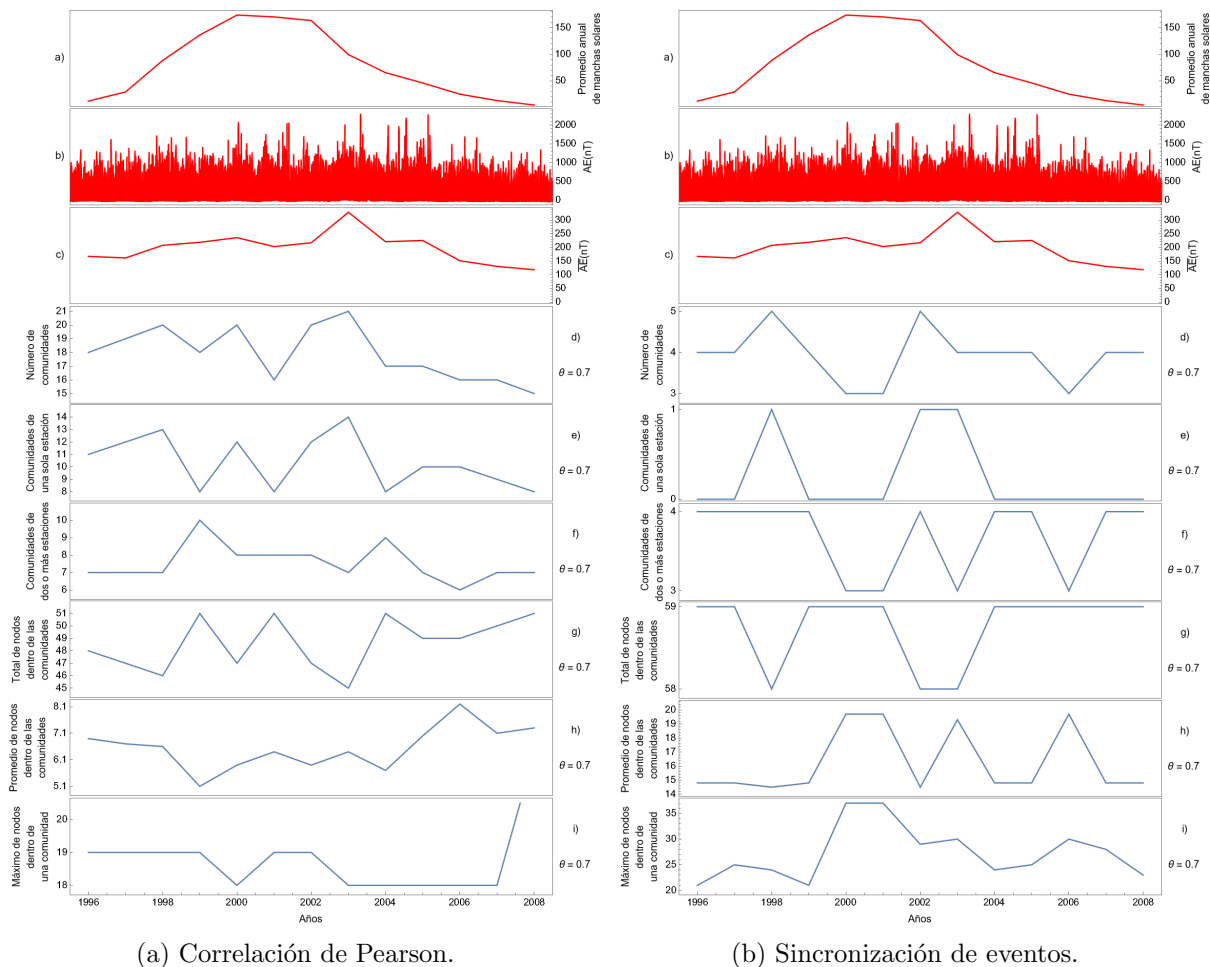


Figura C.10: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice AE y en **c**) promedio anual del índice AE. (Curva azul) En **d**) número de comunidades, en **e**) comunidades de una sola estación, en **f**) comunidades de dos o más estaciones, en **g**) total de nodos dentro de las comunidades, en **h**) promedio de nodos dentro de las comunidades y en **i**) máximo de nodos dentro de una comunidad.

## C.2.2. Área de las comunidades

A continuación mostramos los resultados obtenidos para otra medida del tamaño de las comunidades: la estimación del área promedio (Fig. C.11).

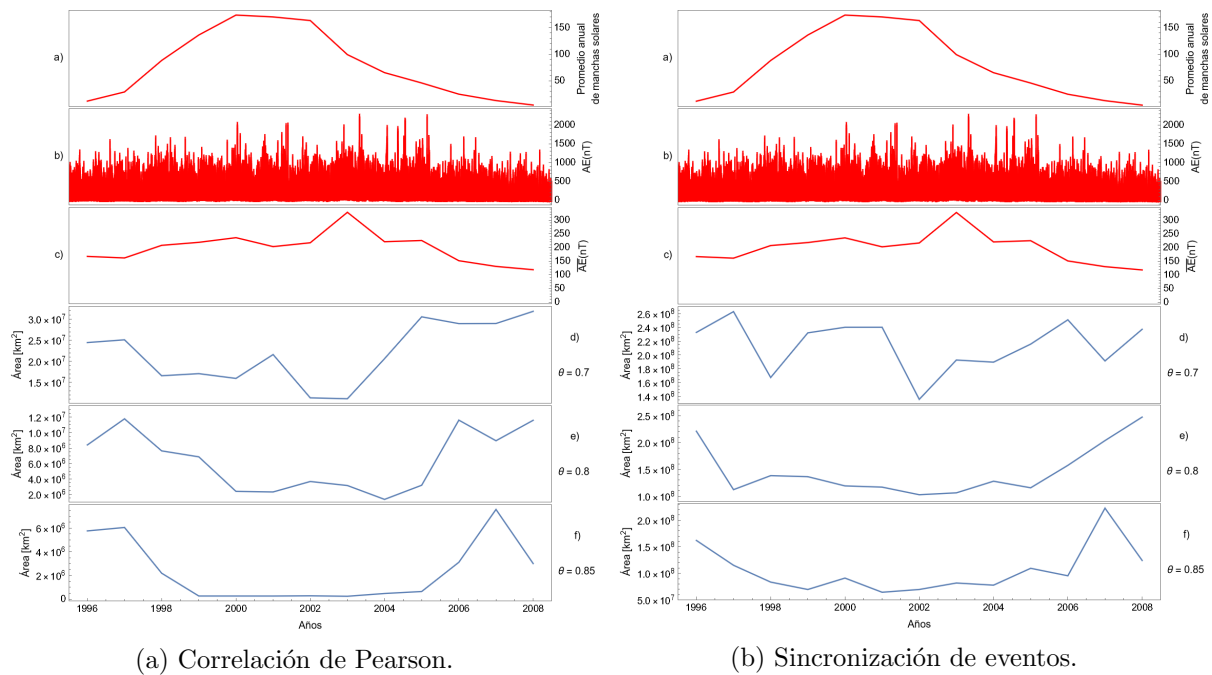


Figura C.11: (Curva roja) En **a**) promedio anual de manchas solares, en **b**) índice AE y en **c**) promedio anual del índice AE. (Curva azul) Área promedio de las comunidades en función del tiempo, utilizando en **d**) un umbral  $\theta = 0.7$ , en **e**) un umbral  $\theta = 0.8$  y en **f**) un umbral  $\theta = 0.85$ .

### C.2.3. Modularidad

A continuación mostramos la variación del valor de la modularidad  $M$  en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. C.12).

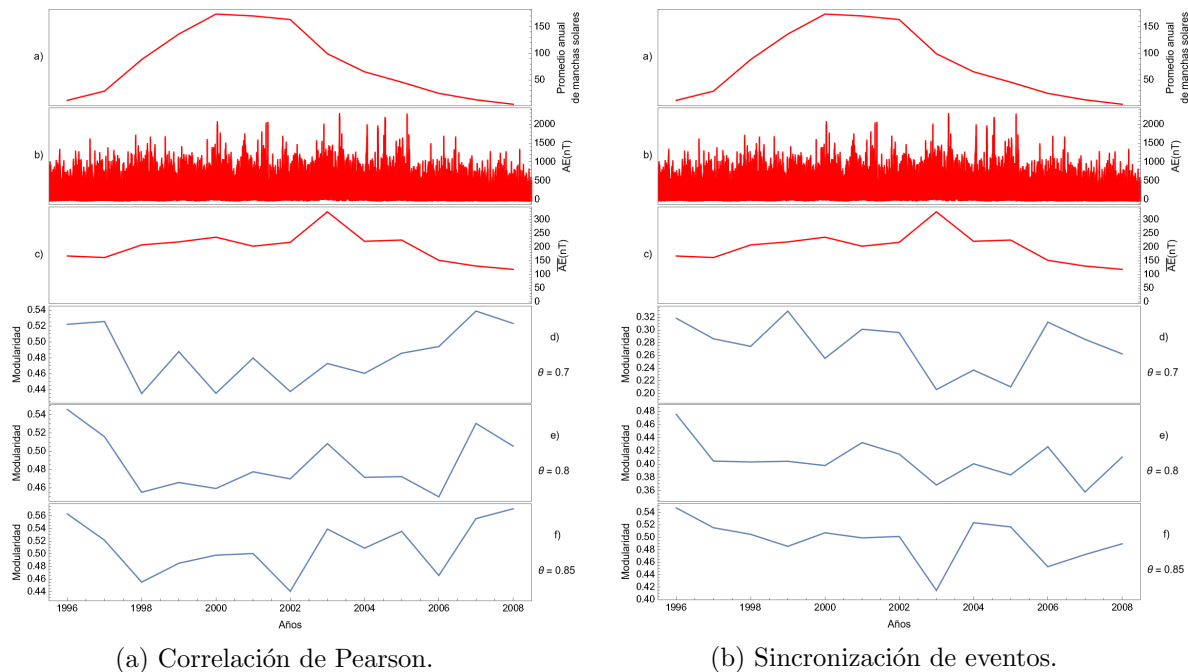


Figura C.12: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice AE y en **c)** promedio anual del índice AE. (Curva azul) Modularidad función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

### C.2.4. Entropía de Shannon

A continuación mostramos la variación del valor de la entropía de Shannon  $S$  de las comunidades en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. C.13).

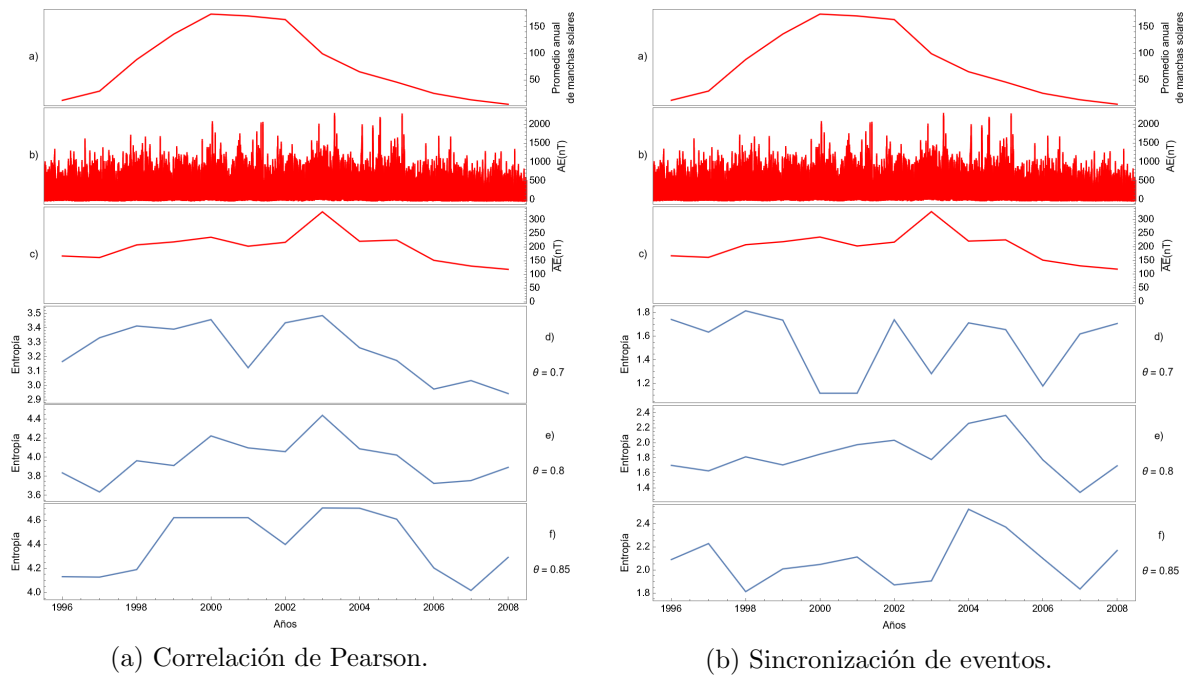


Figura C.13: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice AE y en **c)** promedio anual del índice AE. (Curva azul) Entropía en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

### C.2.5. Coeficiente de clustering

A continuación mostramos la variación del valor del coeficiente de clustering  $C$  de la red en función del tiempo, para los distintos umbrales, y para ambas medidas de similitud (Fig. C.14).

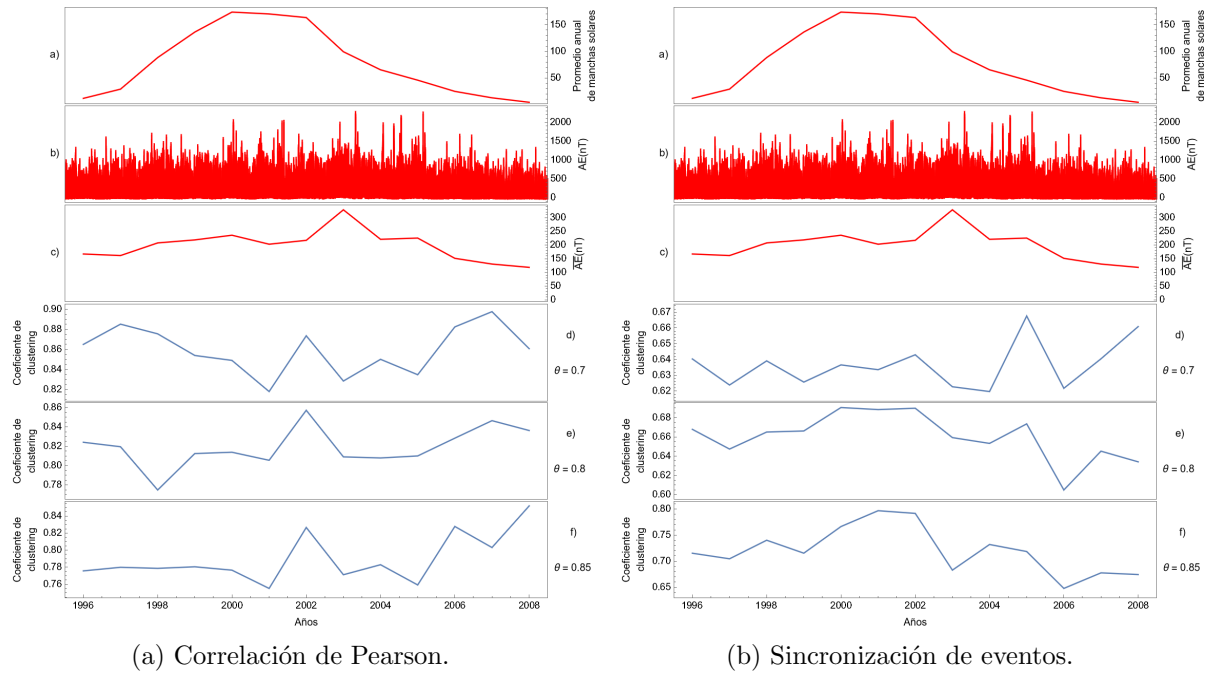


Figura C.14: (Curva roja) En **a)** promedio anual de manchas solares, en **b)** índice AE y en **c)** promedio anual del índice AE. (Curva azul) Coeficiente de clustering en función del tiempo, utilizando en **d)** un umbral  $\theta = 0.7$ , en **e)** un umbral  $\theta = 0.8$  y en **f)** un umbral  $\theta = 0.85$ .

# Bibliografía

- [1] S. Abe y N. Suzuki. “Complex network description of seismicity”. En: *Nonlinear Processes in Geophys* 13 (2006), págs. 145–150. DOI: <https://doi.org/10.5194/npg-13-145-2006>.
- [2] M. Suteanu. “Scale free properties in a network-based integrated approach to earthquake pattern analysis”. En: *Nonlinear Processes in Geophys* 21 (2014), págs. 427–452. DOI: <https://doi.org/10.5194/npg-21-427-2014>.
- [3] D. Pastén et al. “Non-universal critical exponents in earthquake complex networks”. En: *Physica A: Statistical Mechanics and its Applications* 491 (2018), págs. 445–452. DOI: <https://doi.org/10.1016/j.physa.2017.09.064>.
- [4] T. Peron et al. “Correlations between climate network and relief data”. En: *Nonlinear Processes in Geophys* 21 (2014), págs. 1127–1132. DOI: <https://doi.org/10.5194/npg-21-1127-2014>.
- [5] N. Agarwal et al. “Quantifying the roles of single stations within homogeneous regions using complex network analysis”. En: *Journal of Hydrology* 563 (2018), págs. 802–810. DOI: <https://doi.org/10.1016/j.jhydrol.2018.06.050>.
- [6] L. Orr, S. Chapman y J. Gjerloev. “Network community structure of substorms using SuperMAG magnetometers”. En: *Nature Communications* 12 (2021), pág. 1842. DOI: <https://doi.org/10.1038/s41467-021-22112-4>.
- [7] S. Lu et al. “Complex network description of the ionosphere”. En: *Nonlinear Processes in Geophys* 25 (2018), págs. 233–240. DOI: <https://doi.org/10.5194/npg-25-233-2018>.
- [8] A. Najafi et al. “Solar Flare Modified Complex Network”. En: *The Astrophysical Journal* 894 (2020). DOI: <https://doi.org/10.3847/1538-4357/ab8301>.
- [9] E. Flández y V. Muñoz. “Complex Network Study of Solar Magnetograms”. En: *Entropy* 24 (2022). DOI: <https://doi.org/10.3390/e24060753>.



- [10] Y. Xu et al. “Exploring the Clustering Property and Network Structure of a Large-Scale Basin’s Precipitation Network: A Complex Network Approach”. En: *Water* 12 (2020), pág. 1739. DOI: [10.3390/w12061739](https://doi.org/10.3390/w12061739).
- [11] P. Kumar y D. P. Singh. “Solar Cycle Variability and Global Climate Change”. En: *Journal of Earth Science Climatic Change* 10 (2019), pág. 514. DOI: [10.4172/2157-7617.1000514](https://doi.org/10.4172/2157-7617.1000514).
- [12] M. Moldwin. *An Introduction to Space Weather*. Cambridge University Press, 2008.
- [13] J.M. Rodríguez. *El Sol y la irradiancia. El Sol: conceptos básicos*. Editorial Académica Española, Chisinau, 2018.
- [14] C. Schrijver y G. Siscoe. *Heliophysics: Evolving Solar Activity and the Climates of Space and Earth*. Cambridge University Press, 2010.
- [15] M. Domínguez, V. Muñoz y J. A. Valdivia. “Temporal evolution of fractality in the Earth’s magnetosphere and the solar photosphere”. En: *Journal of Geophysical Research. Space Physics* 119 (2014). DOI: [10.1002/2013JA019433](https://doi.org/10.1002/2013JA019433).
- [16] Parsec VZW. URL: <https://www.spaceweatherlive.com/es/actividad-solar/ciclo-solar.html>.
- [17] M. Calcina. “Un modelo dinámico para el campo geomagnético”. En: *Revista Boliviana de Física* 15 (2009), págs. 44–62.
- [18] E. J. Tarbuck et al. *Ciencias de La Tierra: Una Introducción a La Geología Física*. Pearson Educación, 2005.
- [19] R. Merrill, P. McFadden y M. McElhinny. *The Magnetic Field of the Earth: Paleomagnetism, the Core, and the Deep Mantle*. Academic Press, San Diego, 1996.
- [20] R. Lanza y A. Meloni. *The Earth’s Magnetism. An Introduction for Geologist*. Alemania: Editorial Springer, 2006.
- [21] J. Jankowski y C. Sucksdorff. *Guide for Magnetic Measurements and Observatory Practice (International Association of Geomagnetism and Aeronomy)*. 1996.
- [22] W. Lowrie. *Fundamentals of Geophysics*. Vol. 91. 2nd. Cambridge: Cambridge University Press, 2007.
- [23] G. G. Hernández. “Análisis de espectrogramas de anomalías magnéticas para la detección de estructuras.” <https://repositorio.unam.mx/contenidos/3529133>. Tesis de maestría. Universidad Nacional Autónoma de México, 2010.

- [24] A. Cillis y S. J. Sciutto. *Geomagnetic Field and Air Shower Simulations*. Tech. rep., Departamento de Física, Universidad Nacional de La Plata, 2006.
- [25] Istituto Nazionale di Geofisica e Vulcanologia. URL: <http://roma2.rm.ingv.it/>.
- [26] National Centers for environmental information. URL: <https://www.ngdc.noaa.gov/geomag/indices/indices.html>.
- [27] Kyoto World Data Center for Geomagnetism. URL: <https://wdc.kugi.kyoto-u.ac.jp/index.html>.
- [28] J. L. Rasson. *Observatories, intermagnet*. In *Encyclopedia of Geomagnetism and Paleomagnetism*. Springer, 2007.
- [29] M. G. Kivelson y C. T. Russell. *Introduction to space physics*. Cambridge University Press, 1995.
- [30] G. Musmann. *Fluxgate Magnetometers for Space Research*. Books on Demand, 2010. ISBN: 9783839137024. URL: <https://books.google.cl/books?id=0xwYNpQ0bnUC>.
- [31] Instituto geográfico nacional. URL: <https://www.ign.es/web/ign/portal/inicio>.
- [32] NASA. URL: <https://nssdc.gsfc.nasa.gov/nmc/spacecraft/display.action?id=1999-008B>.
- [33] DTU. URL: [https://www.space.dtu.dk/english/research/scientific%5C\\_data%5C\\_and%5C\\_models/magnetic%5C\\_satellites](https://www.space.dtu.dk/english/research/scientific%5C_data%5C_and%5C_models/magnetic%5C_satellites).
- [34] GFZ. URL: <https://isdc.gfz-potsdam.de/champ-isdc/>.
- [35] S. Dorogovtsev y J. Mendes. “Evolution of networks”. En: *Advances in Physics* 51 (2002), págs. 1079–1187.
- [36] S. Boccaletti et al. “Complex networks: Structure and dynamics”. En: *Physics Reports* 424 (2006), págs. 175–308.
- [37] S. Fortunato. “Community detection in graphs”. En: *Physics Reports* 486 (2010), págs. 75–174.
- [38] M. Porter, J. Onnela y P. Mucha. “Communities in networks”. En: *Notices of the American Mathematical Society* 56 (2009), págs. 1082–1097.
- [39] A. Lancichinetti et al. “Characterizing the Community Structure of Complex Networks”. En: *Plos One* (2010). DOI: [10.1371/journal.pone.0011976](https://doi.org/10.1371/journal.pone.0011976).

- [40] M. Girvan y M. Newman. “Community structure in social and biological networks”. En: *Proceedings of the National Academy of Sciences of the United States of America* 99(12) (2002), págs. 7821–7826. DOI: [10.1073/pnas.122653799](https://doi.org/10.1073/pnas.122653799).
- [41] A. Tsonis et al. “Community structure and dynamics in climate networks”. En: *Climate Dynamics* 37 (2010), págs. 933–940. DOI: [10.1007/s00382-010-0874-3](https://doi.org/10.1007/s00382-010-0874-3).
- [42] M. Newman. *Networks: An Introduction*. Oxford University Press, 2010.
- [43] M. Newman. “The Structure and Function of Complex Networks”. En: *SIAM Review* 45 (2003), págs. 167–256. DOI: [10.1137/s003614450342480](https://doi.org/10.1137/s003614450342480).
- [44] M. Newman. “Communities, modules and large-scale structure in networks”. En: *Nature Physics* 8 (2012), págs. 25–31.
- [45] M. Halverson y S. Fleming. “Complex network theory, streamflow, and hydro-metric monitoring system design”. En: *Hydrology and Earth System Sciences* 19 (2015), págs. 3301–3318. DOI: <http://dx.doi.org/10.5194/hess-19-3301-2015>.
- [46] M. Newman. “Detecting community structure in networks”. En: *The European Physical Journal B: Condensed Matter and Complex Systems* 38 (2004), págs. 321–330. DOI: <http://dx.doi.org/10.1140/epjb/e2004-00124-y>.
- [47] M. Newman. “Modularity and community structure in networks”. En: *Proceedings of the National Academy of Sciences* 103(23) (2006), págs. 8577–8582.
- [48] V. Blondel et al. “Fast unfolding of communities in large networks”. En: *Journal of Statistical Mechanics: Theory and Experiment* (2008). DOI: <http://dx.doi.org/10.1088/17425468/2008/10/P10008>.
- [49] M. Rubinov y O. Sporns. “Weight-conserving characterization of complex functional brain networks”. En: *NeuroImage* 56 (2011), págs. 2068–2079. DOI: <http://dx.doi.org/10.1016/j.neuroimage.2011.03.069>.
- [50] A. Clauset, M. Newman y C. Moore. “Finding community structure in very large networks”. En: *Physical Review E* 70 (2004). DOI: [10.1103/physreve.70.066111](https://doi.org/10.1103/physreve.70.066111).
- [51] M. Newman y M. Girvan. “Finding and evaluating community structure in networks”. En: *Physical Review E* 69 (2004). DOI: [10.1103/physreve.69.026113](https://doi.org/10.1103/physreve.69.026113).

- [52] J. Zhang et al. “An Improved Louvain Algorithm for Community Detection”. En: *Hindawi, Mathematical Problems in Engineering* 2021 (2021). DOI: <https://doi.org/10.1155/2021/1485592>.
- [53] J. Donges et al. “Complex networks in climate dynamics: comparing linear and nonlinear network construction methods”. En: *The European Physical Journal Special Topics* 174 (2009), págs. 157–179. DOI: <http://dx.doi.org/10.1140/epjst/e2009-01098-2>.
- [54] K. Steinhäuser, N. V. Chawla y A. R. Ganguly. “An exploration of climate data using complex networks”. En: *ACM SIGKDD Explorations* 12 (2010), págs. 25–32.
- [55] J. Runge, V. Petoukhov y J. Kurths. “Quantifying the Strength and Delay of Climatic Interactions: The Ambiguities of Cross Correlation and a Novel Measure Based on Graphical Models”. En: *Journal of Climate* 27(2) (2014), págs. 720–739. DOI: <https://doi.org/10.1175/JCLI-D-13-00159.1>.
- [56] M. Paluš. “Linked by Dynamics: Wavelet-Based Mutual Information Rate as a Connectivity Measure and Scale-Specific Networks”. En: *Advances in Nonlinear Geosciences* (2018), págs. 427–463. DOI: <http://dx.doi.org/10.1007/978-3-319-58895-7-21>.
- [57] V. Stolbova et al. “Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka”. En: *Nonlinear Processes in Geophysics* 21 (2014), págs. 901–917. DOI: <http://dx.doi.org/10.5194/npg-21-901-2014>.
- [58] F. Conticello et al. “An event synchronization method to link heavy rainfall events and large-scale atmospheric circulation features”. En: *International Journal of Climatology* (2017). DOI: <http://dx.doi.org/10.1002/joc.5255>.
- [59] A. Rheinwalt et al. “Non-linear time series analysis of precipitation events using regional climate networks for Germany”. En: *Climate Dynamics* 46 (2016), págs. 1065–1074. DOI: <http://dx.doi.org/10.1007/s00382-015-2632-z>.
- [60] S. Jha et al. “Network theory and spatial rainfall connections: an interpretation”. En: *Journal of Hydrology* 527 (2015), págs. 13–19. DOI: <http://dx.doi.org/10.1016/j.jhydrol.2015.04.035>.
- [61] C. Shannon. “A Mathematical Theory of Communication”. En: *Bell System Technical Journal* 27 (1948), págs. 379–423.
- [62] D. Watts y S. Strogatz. “Collective dynamics of ‘small-world’ networks”. En: *Nature* 393(6684) (1998), págs. 440–442.