



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL**
Unidad Mérida
DEPARTAMENTO DE FÍSICA APLICADA

**Análisis de series de tiempo usando el algoritmo de
visibilidad natural: Alcances y limitaciones en señales de
temperatura y ritmo cardiaco**

TESIS

que presenta

José Alberto Rosales Pérez

Para obtener el grado de

Doctor en Ciencias

en

Fisicoquímica

Director de Tesis:

Dr. Rodrigo Huerta Quintanilla

Mérida, Yucatán

Agosto 2020

Agradecimientos

A mis padres, hermana y mi familia en general por todo su apoyo, así como a mi Novia Lupita por su enorme paciencia y consideración.

A mis amigos dentro y fuera del CINVESTAV los cuales mediante su buen ánimo, consejos, risas y paciencia hicieron muy ameno este periodo.

A mi asesor de Tesis, Dr. Rodrigo Huerta, por sus consejos, su ánimo y apoyo constante, pero sobre todo por ser figura determinante en mi proceso de formación como investigador.

Al M. en C. Efraín Canto Lugo por su apoyo a lo largo del doctorado y específicamente en la revisión de tesis y la elaboración del mapa de Yucatán.

Al Dr. David Valdés quien aportó los datos con los que se realizó la tesis, así como sus explicaciones sobre el tema climático.

A mis sinodales Dr. Gabriel Pérez, Dr. Francisco Larios, Dr. Jesús Siqueiros y Dr. Manuel Rodríguez, por sus valiosas aportaciones hacia este trabajo así como su tiempo y dedicación.

Al CONACYT por el aporte económico que me brindó durante los estudios de Doctorado.

Al CINVESTAV Unidad Mérida en general por su apoyo durante el Doctorado. en especial mención a Zhirnay Rodriguez por su gran disponibilidad siempre, su apoyo y paciencia.

Al Dr. Román Castro, Coordinador Académico del Departamento, por su apoyo, esfuerzo y su gran dedicación para hacer posible los procesos de titulación en estas circunstancias tan especiales que vivimos ahora. Así también agradezco al personal que apoya al Dr. Román en esta ardua tarea.

Índice

Índice de figuras	5
Índice de tablas	7
Abreviaturas	8
Resumen	9
Abstract	10
1. Motivación	11
2. Análisis de la dinámica de temperatura usando DFA y ApEnt	14
2.1. Introducción	14
2.2. Metodología	17
2.2.1. DFA	17
2.2.2. ApEnt	19
2.2.3. Datos	20
2.2.4. Análisis de Datos usando DFA y ApEnt	20
2.3. Resultados	21
3. Análisis de la dinámica de temperatura usando redes	25
3.1. Introducción	25
3.2. Metodología	26
3.2.1. Algoritmo de Visibilidad Natural	26
3.2.2. Caracterización de las redes	27
3.2.3. Análisis de datos	28
3.3. Resultados	28
3.4. Discusión	30
3.4.1. Invariancia en el tiempo	30
3.4.2. Diferencia entre dinámicas	31
4. Señales fisiológicas y redes	34
4.1. Introducción	34
4.2. Metodología	36
4.3. Resultados	37
4.4. Discusión	39

5. Termodinámica y redes en el contexto del algoritmo de visibilidad	42
5.1. Introducción	42
5.2. Modelo Teórico	44
5.3. Cálculo de temperaturas	45
5.4. Discusión	46
6. Discusión e integración de resultados	50
6.1. ¿En qué casos se puede usar LEN y Ent?	50
6.2. ¿Qué mide el análisis de series de tiempo usando redes?	51
6.3. Complejidad, Termodinámica y NVA	52
7. Conclusiones y Trabajos futuros	53
8. Referencias	55

Índice de figuras

1.	Distribución de grado y medidas de las redes obtenidas para las ciudades de Mérida y Chelem de las series de tiempo tomadas en 2012.	11
2.	Detalles de las series de tiempo de Mérida y Chelem.	12
3.	Clasificación de las diferentes técnicas que existen para cuantificar la complejidad de una serie de tiempo. Imagen tomada de [9].	15
4.	Diagrama donde se muestran las dos grandes familias fgn y fbm, el rango del exponente de escalamiento de DFA y la forma de cada serie de tiempo. Figura tomada del artículo [15].	18
5.	Ubicación geográfica de las ciudades donde se obtuvieron las series de tiempo de temperatura.	21
6.	Detalle de la serie de tiempo de Mérida y Progreso. El periodo comprende de Mayo 1 a Mayo 5 de 2015.	23
7.	Representación gráfica del algoritmo de visibilidad natural. Se muestra la serie de tiempo, su red y su distribución de grado.	61
8.	Distribuciones de grado promedio para cada una de las ciudades analizadas . .	62
9.	Ejemplo gráfico de cómo afecta la variación de la serie de tiempo en LEN. Derecha, una serie de tiempo donde todos los elementos se ven entre sí lo que genera una red completamente conectada, izquierda, una serie de tiempo donde los valores 2 y 4 aumentan, disminuyendo la visibilidad de los nodos produciendo una red menos conectada, y por lo tanto el valor de LEN disminuye.	63
10.	Detalle de la variabilidad en un segmento de la serie de tiempo de Progreso .	64
11.	Serie de tiempo antes (izq.) y durante (der) la sesión de meditación, para datos normales (arriba) y suavizados (abajo).	65
12.	Gráfica donde se muestra la forma de la entropía (Ecuación 25) con respecto al número de enlaces. Tanto la entropía como el número de enlaces se normalizaron a la entropía máxima S_0 y el número de enlaces máximo m_0	66
13.	Gráfica de S vs E construida a partir del modelo descrito en el Capítulo 5. Se normaliza a la S máxima (S_0) y E máxima (E_0).	66
14.	Comparación del modelo teórico vs el experimental para N=50,100 y 500. Mientras más grande N, el error entre las curvas es menor.	67

- 15. Explicación visual sobre el concepto de temperatura negativa en redes. En la parte superior se muestra un sistema termodinámico con una energía máxima acotada, de izquierda a derecha se muestran 3 diferentes sistemas con un aumento de energía. En la parte media se muestran 3 redes de 5 nodos ejemplificando los extremos donde no hay enlaces y donde todos los nodos tienen enlaces (Red completa). En la parte inferior se muestra el diagrama de Entropía vs Energía correspondientes a los sistemas descritos anteriormente. 68
- 16. Representación gráfica del experimento de los dados, el cual consiste en lo siguiente: Se tienen dos dados los cuales al sumarse los puntos obtenidos por las caras nos genera una cantidad X. Nosotros queremos saber qué número X es el mejor de forma tal que sea muy probable que al lanzar los dados obtengamos ese puntaje. Se observa que si escogemos el 7, tenemos 6 formas distintas en las cuales podemos obtener dicha suma y sería más probable ganar que si escogemos el 12 o el 2, los cuales sólo hay una forma en la que nos salgan estos puntajes. Tomado de [68]. 69
- 17. Clasificación propuesta por Tang et al. [9] donde se incluye la propuesta de dónde podrían colocarse ΔEnt y LEN (resaltado en color rojo). 70

Índice de tablas

1.	Caracterización general de las ciudades estudiadas	22
2.	Reporte de los valores de significancia para cada par de ciudades contiguas y su distancia de separación	22
3.	Promedio de las medidas de las redes para cada ciudad.	29
4.	Reporte de los valores de significancia para cada par de ciudades contiguas. .	30
5.	Resultados de la caracterización de las series de tiempo para cada paciente antes (pre) y durante (med) la sesión de meditación.	37
6.	Promedio de las medidas de las redes para cada paciente usando redes.	38
7.	Resultados de la caracterización de las series de tiempo suavizadas para cada paciente antes (pre) y durante (med) la sesión de meditación.	39
8.	Promedio de las medidas de las redes para las series de tiempo suavizadas. . .	40
9.	Promedios calculados a partir del modelo teórico para las redes reales del Clima.	46
10.	Reporte de los valores de significancia para cada par de ciudades contiguas. .	46
11.	Promedio de los cálculos del número de enlaces (m), entropía (S) y temperatura (T) y el resultado de la prueba t para los pacientes antes (Pre) y durante (Med) la sesión de meditación.	47

Abreviaturas

NVA..... Natural Visibility Algorithm

DFA..... Detrended Fluctuation Analysis

ApEnt..... Approximate entropy

LE..... Laplacian Energy

LEN..... Laplacian Energy per Node

Δ Ent..... Diferencia de Entropías

HRV..... Heart Rate Variability

Resumen

En el presente escrito se analizan los alcances y limitaciones del algoritmo de visibilidad natural, el cual es un método que proyecta una serie de tiempo en una red. En este estudio se analiza el comportamiento de la Energía Laplaciana y la Entropía de Shannon calculadas a partir de las redes generadas de distintas dinámicas. Como primer análisis se trabajaron las series de tiempo de temperatura de 4 ciudades de la Península de Yucatán. Los resultados se compararon con técnicas ya bien conocidas como Detrended Fluctuation Analysis y Approximate Entropy diseñadas para analizar la autosimilaridad y el desorden estructural de la serie de tiempo respectivamente. Se encontró que el análisis de redes pudo identificar entre las distintas ciudades analizadas mientras que ApEnt y DFA no pudieron. Se observó que la dinámica de las ciudades costeras es más aleatoria, mientras más nos alejamos de la costa, las señales se hacen más regulares. El análisis de la técnica se extendió a las series de tiempo de la variabilidad del ritmo cardiaco en 8 personas durante una sesión de meditación guiada y se comparó con la señal producida cuando hacen sus actividades normales. Se encontró que el enfoque de redes no pudo diferenciar entre las dinámicas mientras que ApEnt y DFA sí pudieron. No fue hasta que se hizo un suavizado a la serie de tiempo que las redes pudieron diferenciar entre las dinámicas. Finalmente se hizo una discusión sobre la posibilidad de tomar la Entropía y la Energía Laplaciana como variables termodinámicas a partir de la construcción de un modelo de mecánica estadística aplicada a redes. Uno de los hallazgos más interesantes es que es posible encontrar redes con temperaturas negativas. En general, del estudio de esta tesis se concluye que el uso del algoritmo y la cuantificación de Entropía y Energía Laplaciana tienen la capacidad de medir la variación estructural de una serie de tiempo, la cual debe de poseer una componente periódica importante y un nivel de ruido bajo para que puedan funcionar sin problema. Se recomienda el uso de técnicas alternativas si la serie de tiempo no cumple con estas condiciones. Una de las grandes preguntas que quedan abiertas para un trabajo futuro es mostrar si estos cuantificadores pueden ser usados como medidas de complejidad. Finalmente, recomendamos el uso de esta técnica no como un sustituto de alguna otra sino más bien como un instrumento complementario, ya que cada técnica está enfocada en extraer información distinta de la dinámica que está bajo estudio.

Abstract

In this paper, the scope and limitations of the natural visibility algorithm are analyzed. This method projects a time series on a network. The objective of this work is to analyze the behavior of Laplacian Energy and Shannon Entropy calculated from the networks generated from different dynamics. As a first analysis, we work on the temperature time series of 4 cities of the Yucatan Peninsula with different geographic characteristics. The results were compared with well-known techniques such as Detrended Fluctuation Analysis and Approximate Entropy which analyzes the self-similarity and the structural dynamics of the time series respectively. With the network analysis we could identify between the different cities analyzed while apEnt and DFA could not. The dynamics of the coastal cities were found to be more random, while further away from the coast the signals become more regular. The analysis of the visibility algorithm was extended to the time series of heart rate variability in 8 people during a guided meditation session and compared to the signal produced when doing their normal activities. It was found that the network approach could not differentiate between the dynamics while ApEnt and DFA could. When the time series were smoothed, the networks could differentiate between the dynamics. Finally, a discussion was made on whether it was possible to take Entropy and Laplacian Energy as thermodynamic variables from the construction of a statistical mechanics model applied to networks. One of the most interesting findings is that within networks it is possible to find networks with negative temperatures. In general, from the study of this thesis it is concluded that the use of the algorithm and the quantification of Laplacian Entropy and Energy measure the structural variation of a time series, which must have an important periodic component and a low noise level in order to assure a better performance. The use of alternative techniques is recommended if the time series does not meet these conditions. One of the big questions that remains open for future work is to show if these quantifiers can be used as measures of complexity. Finally, we recommend the use of this technique not as a substitute for any other, but rather as a complementary instrument, since each technique is focused on extracting different information from the dynamics under study.

1. Motivación

El presente trabajo de doctorado es una continuación a la investigación que se comenzó en la maestría, cuyo objetivo estuvo enfocado a una exploración general del uso de redes como una herramienta alternativa para el análisis de series de tiempo.

El análisis se aplicó para la serie de tiempo de temperatura de varias ciudades de Yucatán, siendo las más representativas Chelem, una ciudad costera, y Mérida, una ciudad alejada de cuerpos de agua. Lo que se quería analizar era qué tipo de redes se obtenían de cada ciudad y si era posible asociar los patrones observados en las redes a las diferencias entre las dinámicas de cada ciudad cuyas características geográficas son notablemente diferentes. Ese análisis se aplicó para el mes de Febrero a Agosto de 2012.

En el estudio se encontró que la forma de las distribuciones de grado para cada ciudad generaban patrones distintos como se muestra en la Figura 1.

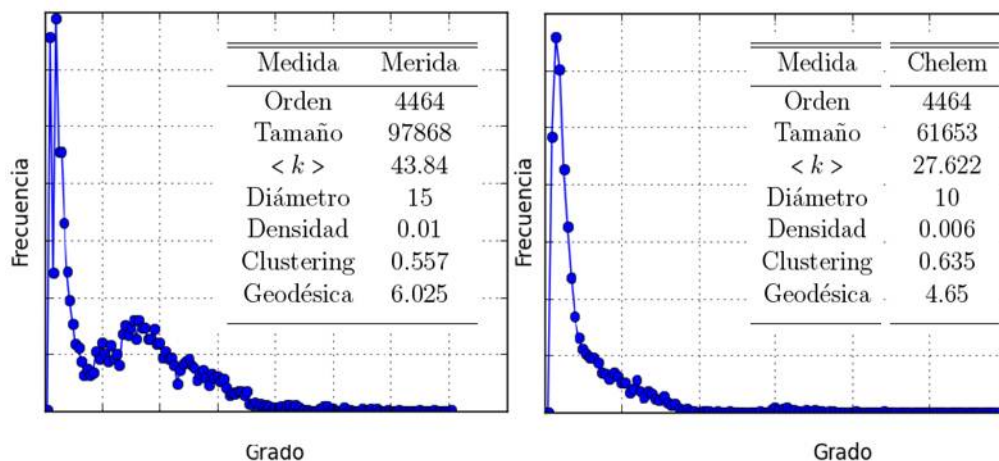


Figura 1: Distribución de grado y medidas de las redes obtenidas para las ciudades de Mérida y Chelem de las series de tiempo tomadas en 2012.

Se concluyó que las diferencias tanto en la forma de la distribución de grado como en las medidas de las redes obedecían a que las series de tiempo de temperatura tienen una estructura diferente. Esta diferencia radica en el contexto geográfico de cada ciudad, Mérida tiene un perfil de calentamiento que depende del sol y la nubosidad, mientras que Chelem está sujeta, además, a las brisas marinas que se generan dada la dinámica de calentamiento que existe por

ser zona costera. Otros factores como la cantidad de lluvia que se genera, cómo impactan los frentes fríos y la dinámica de viento en cada zona, generan formas muy diferentes en las series de tiempo de temperatura de cada ciudad (Figura 2).

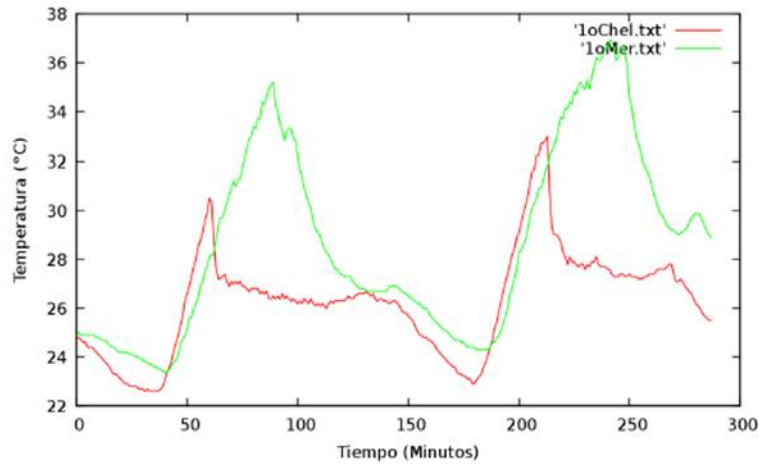


Figura 2: Detalles de las series de tiempo de Mérida y Chelem.

Con este trabajo se corroboró lo reportado por Veleva et al.[1] sobre la forma de una serie de tiempo. En el estudio se analizaron las series de tiempo de temperatura y humedad relativa desde 1992 hasta 1995. En el análisis se encuentra que la forma de la serie de tiempo de temperatura de Mérida y Progreso (ciudad muy cercana a Chelem) depende de las lluvias ocurridas en cada zona así como del efecto que tiene el mar sobre el amortiguamiento de la temperatura en las ciudades costeras. Aunque en su trabajo generalizan que para ambas ciudades la forma de la serie de tiempo puede modelarse con una línea y una parábola, los factores de ajuste serán distintos para cada ciudad.

A partir de estos antecedentes, la primer cuestión pendiente a resolver es si las diferencias encontradas entre las ciudades puedan ser cuantificadas usando redes y que éstas sean estadísticamente significativas. Otro factor a tomar en cuenta es el usar técnicas conocidas en el análisis de series de tiempo, para verificar, tanto los resultados obtenidos con las redes, como la interpretación de lo que sucede en el sistema climático. Finalmente se podría integrar todo este conocimiento para poderlo aplicar a otros sistemas además del clima.

En este trabajo de doctorado se optó por profundizar más en la técnica del análisis de series de tiempo usando redes con el objetivo de generar conclusiones mas sólidas sobre sus alcan-

ces y limitaciones, aumentando la ventana de tiempo de análisis, comparando los resultados usando técnicas más establecidas y ampliando la aplicación del algoritmo a otros sistemas. La presentación del trabajo está distribuida de la siguiente forma:

En el Capítulo 1 se describen las series de tiempo de temperatura usadas para este estudio en torno a los resultados obtenidos usando Approximate Entropy y Detrended Fluctuation Analysis. Una vez que se ha caracterizado el sistema a estudiar, en el Capítulo 2 se reportan los resultados obtenidos al aplicar el análisis usando redes. Se hace una discusión sobre el significado de cada medida de la red en la dinámica analizada dependiendo de la zona climática en la que está ubicada. En el Capítulo 3 se presentan los resultados del análisis usando redes a la serie de tiempo de variabilidad de frecuencia cardiaca de varios individuos antes y durante una sesión de meditación. En el capítulo 4 se genera un modelo usando mecánica estadística para discutir si la Entropía y la energía Laplaciana pueden ser consideradas variables termodinámicas. En el Capítulo 5 se presenta una discusión general a partir de los hallazgos comentados en cada uno de los capítulos, contestando las siguientes preguntas ¿Qué estamos midiendo a las series de tiempo cuando aplicamos la técnica y caracterizamos las redes usando la Energía Laplaciana y la Entropía? ¿En qué casos se puede usar el análisis de series de tiempo usando redes? ¿Es posible medir la complejidad de una dinámica a partir del análisis usando redes mediante el uso de termodinámica? Finalmente en el Capítulo 6 se muestran las conclusiones y algunas propuestas de líneas de investigación.

2. Análisis de la dinámica de temperatura usando DFA y ApEnt

2.1. Introducción

Entre los fenómenos que han llamado la atención a la humanidad desde tiempos muy remotos es la predictibilidad del Clima. El conocer las condiciones del tiempo nos permite desde saber cómo vestirnos ese día hasta aplicaciones más cruciales como en la agricultura. En estos ejemplos podemos ver que la temporalidad a la que necesitamos el pronóstico tiene ventanas de tiempo distintas. Para el primero nos interesan saber el pronóstico a escala de tiempo corta (horas) mientras que para el segundo la escala es mayor (1 año). En inglés, el primer fenómeno es denominado “weather” mientras que el segundo se denomina “climate”. Dado que no hay una palabra en español para diferenciarlos, a lo largo del capítulo se hará uso de estos anglicismos.

Es importante esta diferenciación ya que el “weather” no se puede pronosticar más allá de unos pocos días, los aspectos del clima pueden ser predecibles durante años, décadas y quizás más tiempo [2]. Sin embargo, dado el contexto del cambio climático que estamos viviendo el día de hoy, el poder establecer qué impacto tiene el calentamiento global en las condiciones climáticas de corto (weather) y largo alcance (climate) y cómo impactará económicamente ha sido un tema central entre los grupos de investigación que se enfocan en esta área [3].

Un enfoque para la caracterización del clima es mediante las herramientas que abarca el análisis de sistemas complejos. Dichos sistemas se caracterizan por ser un conjunto de componentes que interactúan de forma no lineal. Esta interacción genera comportamientos emergentes, como la capacidad de adaptarse y evolucionar con el tiempo [4; 5]. El clima de la Tierra es un sistema complejo ya que comprende diversos procesos dinámicos que operan en todas las escalas espacio-temporales [6].

Rosen [7] define un sistema complejo como aquel que no puede descomponerse en un conjunto de partes, es decir, el resultado del sistema no es la suma directa de las partes que lo integran. Entonces ¿Cómo estudiar un sistema como el clima? Una estrategia es el analizar el comportamiento con respecto al tiempo de las principales variables de respuesta del sistema. Con este enfoque se busca medir la complejidad de un sistema a partir de la complejidad de una señal.

Para caracterizar una serie de tiempo se han propuesto varias técnicas [8] las cuales se pueden clasificar en 3 enfoques [9] (Ver Figura 3): autosimilitud, caracterización del espacio

fase y el estado de desorden.

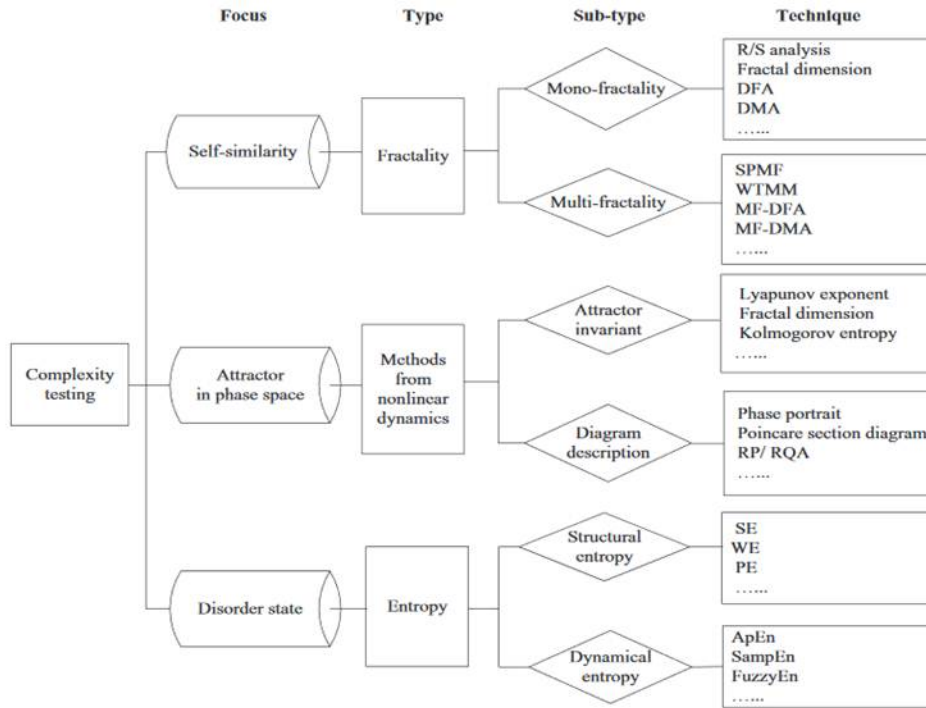


Figura 3: Clasificación de las diferentes técnicas que existen para cuantificar la complejidad de una serie de tiempo. Imagen tomada de [9].

En este trabajo se decidió analizar la dinámica de la temperatura ya que es una característica que captura todos los fenómenos circundantes en el sistema: la variación en la temperatura ambiente en una zona geográfica es una combinación de un comportamiento periódico (ordenado) en combinación con uno aleatorio, los cuales están en constante evolución y adaptación generando correlaciones de largo alcance (fractalidad temporal).

El comportamiento ordenado de la temperatura está modulado por la radiación solar el cual es el principal factor que aleja del equilibrio al sistema. Las fluctuaciones en la temperatura estarán inducidas por la diversidad y distribución de la vegetación en el área, los cuerpos de agua circundantes, los gases en la atmósfera y la actividad humana [10].

Para hacer el análisis, nos centramos en el enfoque de autosimilaridad de la serie de tiempo usando Detrended Fluctuation Analysis (DFA) y el análisis de estados desordenados usando

Approximate Entropy (ApEnt) ya que son técnicas que han sido aplicadas a diferentes regiones mostrando un buen desempeño a la hora de caracterizar la dinámica de distintas ciudades.

DFA es una técnica que cuantifica qué tan fractal es una estructura analizando las fluctuaciones del sistema en diferentes ventanas temporales (Ver sección 2.2.1 para mayor detalle). Bunde et al. [11] aplicó esta técnica para el máximo de temperatura diaria en diferentes lugares: Tucson EUA (97 años de datos colectados), Melbourne Australia (136 Años), Luling EUA (90 Años), Seúl Corea (86 Años), Praga Rep. Checa(218Años) y Kazán Rusia (96 años). Se encontró que existen 2 tipos de comportamiento: a larga escala todas las ciudades se comportaron de forma semejante muy cerca de ser una serie de tiempo aleatoria, mientras que a una escala mas corta (menos de 10 días) el comportamiento era variable, ya que en esta escala la geografía comienza a jugar un papel importante.

Zheng [12] hace un análisis de la temperatura anual media en Beijing (133 años), encontrando que las series de tiempo tienen correlaciones de largo alcance, sin embargo, encontró además que la dinámica tiene 4 escalas distintas. Las diferencias encontradas entre los datos de Bunde et al. y Zheng radican en la preparación de los datos, mientras que el primero tomó sólo los máximos diarios, el segundo hizo un promedio, por lo cual se pierde cierta información de la dinámica al no exponer los datos sin tratamiento a la técnica.

En cuanto al enfoque de estados desordenados uno de los estudios que relacionan la entropía de una serie de tiempo de temperatura con la geografía de donde fue medida es en el trabajo de Shuangcheng et al. [13] llevado a cabo en las provincias de Yunnan, Sichuan, Quizhou y Quangxi en China. El resultado arroja que cada zona climática tiene una entropía distinta. Los valores con mayor entropía correspondían a las regiones en donde existe una interacción entre distintas masas de aire provocada por la orografía del lugar, mientras que las regiones hacia el sur (más planas), sólo fluye una sola masa de aire, obteniendo valores de entropía menores.

El objetivo de este capítulo es el caracterizar las series de tiempo de temperatura de distintas ciudades de la Península de Yucatán usando DFA y ApEnt para entender cada ciudad en torno a su auto-afinidad y su grado de desorden dependiendo de su ubicación geográfica. Posteriormente nos enfocaremos en analizar qué tan efectiva es cada técnica para diferenciar la dinámica de una ciudad con respecto a otra dada su zona climática usando pruebas estadísticas.

2.2. Metodología

2.2.1. DFA

A continuación, se describirá el algoritmo DFA y posteriormente se analizará el significado del resultado en la fenomenología del sistema. Se tomó la notación del artículo original [14]. Para aplicar DFA a una serie de tiempo $B = B(i), i = 1, 2, \dots, N$, el procedimiento es el siguiente:

1. Integración de la señal usando la ecuación:

$$y(k) = \sum_{i=1}^k (B(i) - \langle B \rangle) \quad (1)$$

donde $\langle B \rangle$ es el promedio de la serie de tiempo B .

2. Se dividen los datos en segmentos de tamaño n y se busca una aproximación lineal del segmento (línea de tendencia) mediante el método de mínimos cuadrados. Este procedimiento se hace para cada segmento por separado.
3. Se calcula la fluctuación promedio $F(n)$ de la señal alrededor de la línea de tendencia para el tamaño de segmento n usando la siguiente ecuación:

$$F(\Delta n) = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(k) - y_{\Delta n}(k))^2} \quad (2)$$

Este cálculo se va a repetir para los tamaños de ventana que se hayan considerado.

4. Como el objetivo es caracterizar el comportamiento que tiene $F(\Delta n)$ mientras aumentamos n se graficará $\log(F(n))$ vs $\log(\Delta n)$ y mediante el método de mínimos cuadrados calculamos la pendiente (α) de esa recta.

La pendiente que se calcula en la etapa final de la metodología se llama exponente de escala, exponente de autocorrelación o parámetro de autosimilaridad. Tal como lo sugieren los nombres anteriormente citados, lo que mide el DFA es qué tanta similitud hay en las diferentes ventanas temporales.

El valor del exponente nos permite delimitar entre dos grandes familias (ver Figura 4), el Ruido fraccional Gaussiano (fgn por sus siglas en inglés), y el Movimiento fraccional Browniano (fbm), las cuales provienen los famosos procesos estocásticos Ruido Blanco y Movimiento Browniano [15].

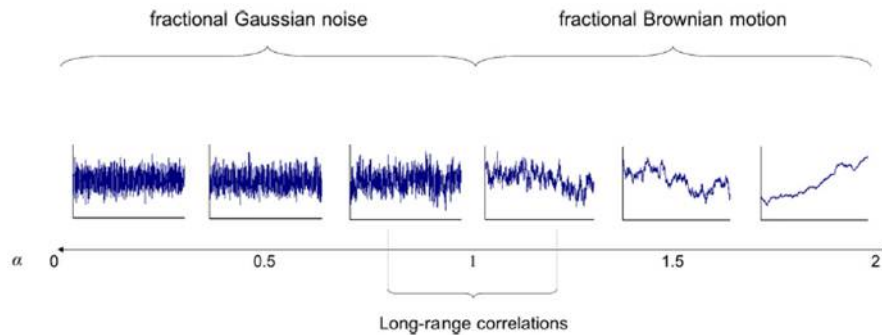


Figura 4: Diagrama donde se muestran las dos grandes familias fgn y fbm, el rango del exponente de escalamiento de DFA y la forma de cada serie de tiempo. Figura tomada del artículo [15].

El fbm es una extensión del movimiento Browniano el cual permite medir la correlación o memoria de lo que pasó anteriormente. Es decir, un fbm persistente quiere decir que, si en los datos anteriores hubo una tendencia a aumentar la magnitud, es más probable que para el futuro haya un incremento y en el diagrama serían todas las series de tiempo con $\alpha > 1,5$. Cuando el movimiento es antipersistente o también llamado compensatorio, si la tendencia histórica presenta un incremento, lo más probable es que haya un decremento en el futuro. Por esta razón parece que el sistema tiene “memoria” porque tiene que “recordar” qué tipo de tendencia había en su historia. El movimiento antipersistente del fbm tiene el rango de $1 < \alpha < 1,5$. Cuando $\alpha = 1,5$, tenemos un movimiento Browniano clásico sin memoria [16].

El ruido fgn también comprende procesos persistentes ($0,5 < \alpha < 1$) y antipersistentes ($\alpha < 0,5$) sólo que la construcción es bajo el proceso de ruido blanco. Cuando $\alpha = 0,5$, se habla de que es un ruido blanco puro sin memoria alguna. Cabe mencionar que mientras más nos alejamos de los procesos puros, es decir del $\alpha = 0,5$ y $\alpha = 1,5$, el sistema tiene que recordar más de su historia [16].

Cuando el resultado del análisis es $\alpha = 1$ la memoria del sistema se maximiza ya que se encuentra en la posición donde la memoria es máxima para el movimiento persistente del proceso fgn y el máximo del proceso antipersistente del proceso fbm. Experimentalmente se ha encontrado que los sistemas complejos producen series de tiempo alrededor de $\alpha = 1$, por lo que se tiene la conjetura de que si un sistema se aleja de $\alpha = 1$, entonces está perdiendo su complejidad y puede colapsar al hacerse muy aleatorio o muy ordenado [17]. Un ejemplo de esto son las señales de corazón y la movilidad motriz [18; 19].

2.2.2. ApEnt

El algoritmo de Entropía Aproximada (ApEnt por su acrónimo en inglés), es un análisis estadístico que cuantifica qué tan regulares son las fluctuaciones en una serie de tiempo. Esto es importante dado que una dinámica con patrones repetitivos a lo largo de la dinámica es altamente predecible.

ApEnt refleja qué tan probable es que patrones similares no serán seguidos de observaciones similares. Una serie de tiempo con una estructura muy repetitiva tendrá un ApEnt bajo, mientras que una serie de tiempo con fluctuaciones sin patrón recurrente obtendrá un coeficiente de ApEnt alto. Tiene una interpretación directa con el orden/desorden de un sistema, mientras más entrópico, más desordenado [20].

Para una serie de tiempo $y = \{y(i)\}, i = 1, 2, \dots, N$ el algoritmo se construye de la siguiente forma [21]:

1. Comenzamos definiendo dos vectores de la siguiente forma:

$$x(i) = [y(i), y(i+1), \dots, y(i+m+1)] \quad (3)$$

$$x(j) = [y(j), y(j+1), \dots, y(j+m+1)] \quad (4)$$

donde $0 < m < N$

2. Posteriormente se calcula el número de vectores similares $n_m^i(r)$, esto se establece como el número de vectores $j \leq N - m - 1$ tal que $d[x(i), x(j)] < r$, donde:

$$d[x(i), x(j)] = \max_{k=1,2,\dots,m} |y(i+k-1) - y(j+k-1)| \quad (5)$$

3. Con estos resultados se hace el siguiente cálculo:

$$C_m^i(r) = \frac{n_m^i(r)}{N - m - 1} \quad (6)$$

4. Ahora se procede a calcular el siguiente factor:

$$\Phi^m(r) = \frac{1}{N - m - 2} \sum_{i=1}^{N-m+1} \log C_m^i \quad (7)$$

5. Finalmente se calcula ApEnt como sigue:

$$ApEnt = \Phi^m(r) - \Phi^{m+1}(r) \quad (8)$$

ApEnt depende de los parámetros m (embedding dimensión) y r (scale factor). Delgado et.al [20] recomiendan m entre 2 o 3 y r entre 0.1 y 0.25 veces la desviación estándar de los datos.

2.2.3. Datos

Los datos se obtuvieron mediante las mediciones de una estación meteorológica (DAVIS Vantage Pro 2 Plus) instalada en cada una de las ciudades estudiadas: Mérida, Sierra Papacal, Progreso y Sisal (Figura 5) ubicadas en el estado de Yucatán, México, para el año 2015 y 2016.

Las mediciones se realizaron cada 2 segundos y el promedio de todas las lecturas obtenidas en un período de 10 minutos se registró en la base de datos del instrumento. Por lo tanto, para cada día, obtendremos 144 mediciones, para evitar diferencias en el número de nodos dado que tenemos meses de 28, 30 y 31 días, todas las series de tiempo se consideraron sólo 28 días para tener redes con nodos de 4032 en todas.

2.2.4. Análisis de Datos usando DFA y ApEnt

El análisis de datos se realizó mensualmente, es decir, para la serie de tiempo de cada mes se le calculó el coeficiente de escalamiento DFA y ApEnt ($m = 2, r = 0,20\sigma$), se reporta el promedio por cada ciudad.

Para definir si existe una diferencia estadísticamente significativa entre las mediciones de ciudades contiguas, se realizó una prueba t para poblaciones de medias independientes con un 95 % de nivel de confianza. Es decir, que se va a calcular la probabilidad de que los resultados arrojados por Mérida y Por Sierra Papacal provengan de la misma distribución para determinar si son la misma o diferentes poblaciones, con un 95 % de certeza en que el resultado es correcto. Las ciudades a comparar fueron Mérida-Sierra Papacal, Sierra Papacal-Progreso, Progreso-Sisal (ver Figura 5).

Se optó por una prueba t y no por un ANOVA (la cual es una prueba estadística que también está enfocada a determinar diferencias entre poblaciones) ya que ésta se enfoca en determinar si las poblaciones sometidas a la prueba son distintas o no, pero no dice cuáles. Si se aplica un

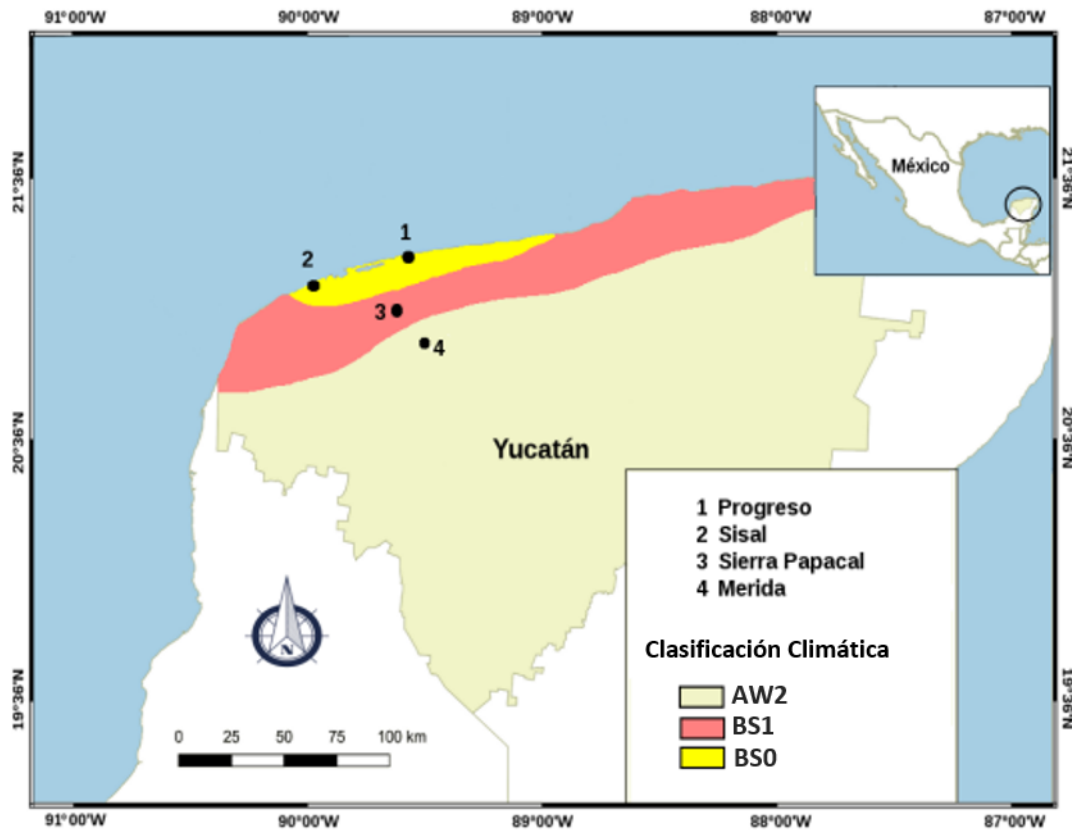


Figura 5: Ubicación geográfica de las ciudades donde se obtuvieron las series de tiempo de temperatura.

ANOVA, adicionalmente habría que aplicar las pruebas de rango post hoc, las cuales permiten determinar qué medias difieren, mientras que la prueba t te dice directamente si dos poblaciones son distintas o no.

2.3. Resultados

Estos resultados fueron publicados en PLoS One [22]. En la Tabla 1 se muestran algunas generalidades de las series de tiempo así como de las poblaciones que se están estudiando y el resultado de DFA y ApEnt por ciudad.

Los datos nos indican un patrón muy especial cuando se analiza la temperatura, la canti-

Tabla 1: Caracterización general de las ciudades estudiadas

	Mérida	Progreso	SierraP	Sisal
Temperatura Media (C)	27.2	25.9	26.7	26.6
Zona Climática*	Aw2	BS0	BS1	BS0
Lluvia Acumulada (mm)	958	466	679	574
ApEnt	0.2434	0.3114	0.2631	0.2931
DFA	1.14	1.10	1.12	1.09

*Simbología de la clasificación de Koppen modificada por García: AW2 Cálido Sub-Húmedo, BS0 Árido Muy cálido, BS1 Semi-Árido muy cálido [23].

dad de lluvia acumulada mensual y la zona climática a la que pertenecen. Encontramos que mientras más nos acercamos a la costa, la temperatura media anual es menor dado el efecto amortiguador que tiene el mar. Así también encontramos que hay menos precipitación anual acumulada en las ciudades costeras dado que en el Golfo de México permanece una columna de alta presión que hace que no genere las condiciones para las lluvias, es por esta razón que Progreso y Sisal se clasifica en una zona climática más seca que Mérida.

En la Figura 6 se muestra un fragmento de las series de tiempo de Mérida y Progreso, donde se aprecia de forma clara la diferencia entre el comportamiento de la temperatura. Mérida se observa más regular que Progreso.

Para poder analizar el comportamiento de DFA y ApEnt, analizaremos primero las pruebas estadísticas para saber entre qué ciudades podemos hacer una comparación válida presentada en la Tabla 2.

Tabla 2: Reporte de los valores de significancia para cada par de ciudades contiguas y su distancia de separación

	ApEnt	DFA	Distancia (Km)
Mérida-SierraP	0.051012	0.272336	15
SierraP-Progreso	0.00823*	0.165988	42
Progreso-Sisal	0.207246	0.37817	15
Mérida-Progreso	0.000897*	0.073192	27

*Valores con significancia estadística del 95 % de confianza.

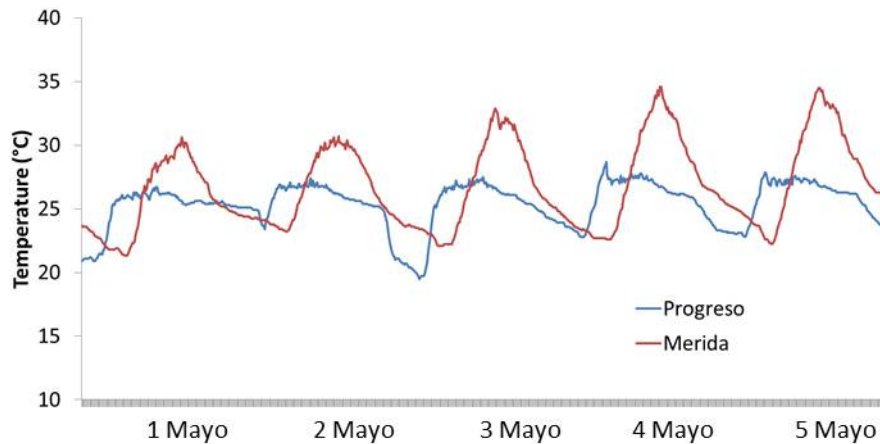


Figura 6: Detalle de la serie de tiempo de Mérida y Progreso. El periodo comprende de Mayo 1 a Mayo 5 de 2015.

Se observa que DFA no puede distinguir entre ningún par de ciudades dado que no pasaron las pruebas estadísticas. Lo único que se puede decir hasta el momento es que son muy cercanos a un $\alpha = 1$ que es donde se encuentran las señales de los sistemas complejos. Estrictamente hablando, las señales de las ciudades analizadas de Yucatán pertenecen a un proceso fmb anti-persistente con correlaciones de largo alcance.

Cuando se analizan los resultados estadísticos de ApEnt (Tabla 2), encontramos que puede diferenciar perfectamente entre Mérida-Progreso y Sierra Papacal-Progreso. Por lo que podemos hacer un análisis válido sólo entre las ciudades costeras vs no costeras.

Discusión

DFA busca la auto-similitud de las fluctuaciones en diferentes escalas de tiempo, en otras palabras qué tan fractal es una serie de tiempo. ApEnt mide qué tan similares son las estructuras de segmentos fijos de la serie de tiempo, pero de forma estadística.

Como se mencionó anteriormente, la técnica de DFA no distingue entre ciudades, por lo que se puede decir que no hay una diferencia en el grado de fractalidad en la serie de tiempo de las ciudades estudiadas. Para este estudio se puede decir que la geografía del lugar no impacta de forma estadísticamente significativa en la autoafinidad de la dinámica, la cual se mantiene muy cerca de $\alpha = 1$. Es probable que la forma en que DFA mide la auto-afinidad de una dinámica no sea tan sensible a pequeños cambios. Hay que recordar que las distancias

entre las poblaciones son menores a 30km.

Para ApEnt encontramos que puede distinguir entre ciudades no costeras y costeras. Para las primeras la magnitud de ApEnt es menor, lo que significa que son series de tiempo más regulares. Los valores más altos pertenecen a las ciudades costeras, seguido de Sierra Papacal y Mérida. Esto es de esperarse ya que en las zonas costeras existe la introducción de aire húmedo de menor temperatura llamadas brisas marinas que hacen que se generen más fluctuaciones cuya magnitud y frecuencia de ocurrencia no es regular. Este efecto es prácticamente nulo en Mérida y Sierra Papacal, por lo que su calentamiento y enfriamiento carece de fluctuaciones, lo que lo hace más regular.

Encontramos que ApEnt es una técnica más sensible a pequeños cambios que el análisis de DFA. Sin embargo, este hecho no implica que una técnica sea mejor que otra, sino más bien nos da diferentes puntos para hacer una caracterización integral de la dinámica.

En este estudio se podría concluir que las dinámicas tienen correlaciones de largo alcance, en donde la dinámica de las ciudades costeras es más aleatoria que las no costeras. En otras palabras, si se hace un pronóstico de la temperatura, se podría tener mayor certeza en todas las ciudades por la gran correlación que hay en los datos, pero tendríamos menos errores al tratar de predecir a Mérida y Sierra Papacal que las ciudades costeras.

Por otra parte, si el estudio estuviese enfocado en mostrar qué técnica puede clasificar cada ciudad por el tipo de clima al que pertenecen, DFA es una técnica que no es la mejor opción, mientras que ApEnt tiene un mejor desempeño al poder diferenciar ciudades costeras de las no costeras, pero falla en diferenciar Mérida y Sierra Papacal, las cuales están en distintas zonas climáticas.

3. Análisis de la dinámica de temperatura usando redes

3.1. Introducción

Una forma alternativa de analizar series temporales es mapeando su estructura en una red. Dependiendo de qué datos se toman de la serie de tiempo para transformarlo a una red, los algoritmos que hasta el momento se han propuesto se pueden dividir en dos grupos: aquellos que mapean estados dinámicos similares al extraer información del espacio de fase [24; 25; 26] y aquellos que mapean la estructura de la serie temporal [27; 28].

El análisis de series de tiempo usando redes es un método para caracterizar la dinámica de un sistema de manera integral. Lacasa et al. [27] mostró que, con el algoritmo de visibilidad natural, las características estructurales de un sistema dinámico se trasladan a la estructura de la red.

Debido a la simplicidad del algoritmo, su bajo costo computacional y la información estructural que se puede obtener de la dinámica [29], el algoritmo de visibilidad ha sido muy popular en diversas aplicaciones, por ejemplo, sismos [30; 31] y finanzas [32; 33].

Donner et al. [34] analizaron los alcances del algoritmo de visibilidad transformando la serie de tiempo diaria del nivel relativo de la marea desde 1916 a 2005 obtenidas de Ratan (Suecia) y Estocolmo. En el estudio se detecta el efecto de “El niño” alrededor de 1982-1983 el cual se cataloga como uno de los más fuertes. También se alcanza a observar el efecto que tiene en la marea la temporada de huracanes.

En el mismo estudio también se analiza la variación de las anomalías medidas en las estalagmitas de la cueva Dongge en el sureste de China. Este tipo de datos han demostrado su excelente potencial como fuente de información sobre la variabilidad de las condiciones paleoambientales, particularmente relacionado con cambios de precipitación. En el estudio se reporta que la distribución de grado es muy similar a la encontrada para el trabajo de Elsner et al. [35]. Los autores mencionan que es posible que exista una fuerte influencia del mecanismo de construcción de la red que podría restringir los posibles resultados del análisis realizado.

Elsner et al. [35] analizó la ocurrencia de huracanes en las costas de EUA. Se reporta que la distribución de grado es consistente con un proceso aleatorio de Poisson. Los hubs pertenecen a los años de alta ocurrencia de huracanes.

Aunque no pertenece a un estudio sobre el clima, el análisis de Zou et al. [36] presenta

importantes hallazgos para series de tiempo con cierta periodicidad integrada. En el artículo los autores analizaron las series de tiempo de manchas solares (International Sunspot Number ISN) diarias (59473 datos) y mensuales(1944 datos). Se pudieron identificar los ciclos de la series de tiempo a partir de las comunidades que se forman en la red. También se reporta una distribución de grado con colas pesadas, donde la variación a largo plazo de los ciclos se refleja en los hubs (nodos con muchas conexiones) de la red.

En general, el análisis de series de tiempo ha mostrado que en la práctica se pueden encontrar características importantes de las dinámicas como por ejemplo identificar eventos que han perturbado la dinámica normal de la serie de tiempo. En este capítulo se presenta la definición del algoritmo de transformación y cómo se caracterizaron cuantitativamente las redes obtenidas por cada ciudad. Se hizo un análisis estadístico para determinar qué cuantificadores de la red pueden distinguir entre ciudades contiguas.

3.2. Metodología

3.2.1. Algoritmo de Visibilidad Natural

El algoritmo de visibilidad es una técnica desarrollada por Lucas Lacasa et al. [27], la cual se usa para transformar series de tiempo en redes. El proceso de transformación comienza asignando a cada valor del eje horizontal (t) de la serie de tiempo como un nodo. La conexión entre nodos se construye de la siguiente forma: Dados dos valores (t_a, y_a) y (t_b, y_b) de la serie de tiempo, se desea analizar si éstos están conectados o no. Si para todo valor (t_c, y_c) comprendido entre los dos anteriores se cumple que:

$$\frac{y_b - y_c}{t_b - t_c} > \frac{y_b - y_a}{t_b - t_a} \quad (9)$$

existe un enlace entre y_a y y_b .

De forma gráfica (Figura 7) si se traza una recta entre dos impulsos de la serie y no existe nada que cruce la recta que trazamos, entonces se coloca un enlace. Los nodos contiguos siempre estarán conectados, y los nodos cuyo valor del eje vertical sean más altos, generalmente tendrán mayor visibilidad, es decir más vecinos y por lo tanto mayor grado (hubs).

Con este procedimiento, cualquier serie de tiempo puede ser convertida en una red y a partir de sus medidas se pueden analizar diferentes aspectos que se relacionen con la dinámica

del fenómeno que se está estudiando. Es importante mencionar que la estructura de una serie es conservada durante la transformación, esto es, series periódicas se convierten en redes regulares, series aleatorias en redes aleatorias y series fractales en redes libres de escala [27].

3.2.2. Caracterización de las redes

Una vez que se ha generado la red, la caracterizamos con algunas métricas que se presentarán en esta sección [37; 38]. Supongamos que G es una red con m enlaces y N nodos, donde el número de nodos vinculados con el nodo i se denomina grado de nodo i , denotado por k_i , y el grado medio es el número promedio de vecinos en todos los nodos de la red. Para caracterizar la conectividad de la red, se traza un histograma llamado distribución de grado. Este diagrama muestra la probabilidad $p(k)$ de tener un nodo con grado k , como se muestra en la Figura 7.

La asortatividad es una medida de la correlación de la red. Si los nodos más conectados prefieren estar vinculados a los nodos más conectados, la red está correlacionada y su valor es cercano a 1. Por otro lado, si los nodos menos conectados prefieren vincularse con los nodos más conectados, la red es anticorrelacionada, y su valor es cercano a -1. Un valor de cero significa que no hay ninguna preferencia.

Se han sugerido otras medidas, como la Entropía de Shannon (Ent) y la Energía Laplaciana (LE). El primero se define como [39],

$$Ent(G) = - \sum_k p(k) \ln(p(k)) \quad (10)$$

El concepto de entropía suele estar siempre referenciado a un punto para poder hacer una interpretación. En este caso, lo referenciamos al valor de la entropía obtenida de la red generada por una serie de tiempo aleatoria. Para obtener este resultado, tomamos la serie de tiempo a analizar, por ejemplo Enero de 2015 de Mérida, la transformamos y obtenemos la entropía denominada $Ent(G_{org})$. Para obtener el valor de referencia, tomamos la misma serie de tiempo y aleatorizamos la posición en la que se encuentran los datos conservando su magnitud y aplicamos la transformación para poder obtener el valor de entropía $Ent(G_{rand})$, este procedimiento se hace 100 veces y se reporta el promedio. A partir de estas cantidades, la “entropía referenciada” ($\Delta Ent(G)$) se define de la siguiente manera:

$$\Delta Ent(G) = |Ent(G_{org}) - Ent(G_{rand})| \quad (11)$$

El cálculo LE utilizado para este estudio fue definido por Kragujevac [40], [41]

$$LE(G) = 2m + \sum_{i=1}^N k_i^2 \quad (12)$$

donde k_i es el grado del i -ésimo nodo y m el número de enlaces de la red G . La energía Laplaciana por nodo (LEN) se define como:

$$LEN(G) = \frac{LE}{N} \quad (13)$$

3.2.3. Análisis de datos

El análisis de datos se realizó mensualmente, es decir, la serie de tiempo de cada mes se transformó en una red usando el algoritmo de visibilidad natural. Posteriormente se obtuvo para cada red su distribución de grado, grado medio, distancia geodésica, clustering, asortatividad, Ent, Δ Ent y LEN.

Se realizó una prueba t para poblaciones de medias independientes para determinar si las variables que estamos midiendo son diferentes, estadísticamente hablando, entre ciudades vecinas, por ejemplo, Mérida y Sierra Papacal, Sierra Papacal y Progreso, etc. En las pruebas buscamos que el valor de p sea mejor que 0.5 para garantizar un 95 % de confianza de los resultados.

3.3. Resultados

Estos resultados fueron publicados en PLoS One [22]. En la Figura 8 encontramos el promedio de las distribuciones de grado para cada una de las ciudades. Encontrando un patrón muy similar reportado en el trabajo de maestría. Colas pesadas en todos los casos debido a los nodos que corresponden a las temperaturas más altas, las cuales pueden ver a más nodos y por lo tanto las más conectadas.

En la Tabla 3 se muestran los resultados cuantitativos obtenidos de las redes para cada ciudad utilizando NVA. Se encontró mayor conectividad media y Energía en la red de Mérida, seguido de Sierra Papacal, Sisal y finalmente Progreso.

Para el Clustering se encontró que las ciudades costeras tienen más triadas que las ciudades no costeras. De igual forma encontramos que la distancia geodésica es mayor para las ciudades costeras que para las ciudades no costeras.

Finalmente, al analizar la asortatividad de las redes no se encuentra un patrón definido. Para la Energía y la entropía encontramos un patrón similar al grado medio.

Tabla 3: Promedio de las medidas de las redes para cada ciudad.

	Mérida	Progreso	SierraP	Sisal
Mean Degree	45.1	29.9	39.7	32.2
Assortativity	-0.1145	-0.0725	-0.1193	-0.1287
Ent(G)	1.99	1.86	1.95	1.88
Δ Ent(G)	0.9244	0.7702	0.8882	0.7984
LEN(G)	2919	1713	2397	1936

Cabe resaltar que encontramos las mismas tendencias que en el estudio presentado en el trabajo de maestría, donde tenemos un coeficiente de clustering arriba de 0.5 y geodésicas pequeñas, lo que da pie a ser redes muy conectadas entre sí.

En cuanto a la asortatividad hay una ligera tendencia a ser redes anticorrelacionadas. Sin embargo, dado que los valores están muy cercanos a cero, se podría decir que no existe correlación, es decir no hay preferencia de conexión entre los nodos.

Una de las grandes preguntas que se pretende responder es sobre si las medidas de las redes de dos ciudades contiguas son estadísticamente diferentes o no, para ello se realizó una prueba t para poblaciones de medias independientes y se hizo para dos ciudades que fueran contiguas.

Para determinar si dos ciudades son diferentes, buscamos un valor $p < 0,05$ en la prueba t de medias independientes, que equivale a un 95 % de confianza. Cada ciudad tendrá un grupo de mediciones independientes y normalmente distribuidas (Grado medio, asortatividad, LEN, Ent y Δ Ent), una para cada mes. La distancia geodésica y el clustering no fueron distribuciones normales, por lo que no se les puede aplicar este tipo de pruebas.

La Tabla 4 muestra los resultados de la prueba t . El grado medio, LEN, Ent y Δ Ent son medidas que pueden distinguir entre los pares Mérida-Sierra Papacal y Sierra Papacal-Progreso. Es decir que se está diferenciando entre poblaciones de distintas zonas climáticas, Sub húmedo, semi-árido y árido.

Tabla 4: Reporte de los valores de significancia para cada par de ciudades contiguas.

	Grado Medio	Asortatividad	Ent	ΔEnt	LEN
Merida-SierraP	0.000024*	0.396552	0.000066*	0.000173*	0.000037*
Sierra-Progreso	0.00001*	0.01132*	0.00001*	0.00001*	0.00001*
Progreso-Sisal	0.057084	0.001293*	0.104444	0.149133	0.080656
Merida-Progreso	0.00001*	0.00001*	0.00001*	0.00001*	0.00001*

*Valores con significancia estadística del 95 % de confianza.

Como se había esperado, los valores de Mérida y Progreso se distinguen sin problemas usando cualquiera de las medidas usadas para caracterizar la red, por lo que podemos decir que tenemos un instrumento de medición que puede distinguir entre dinámicas las cuales se encuentran en distintas zonas climáticas.

La asortatividad fue una medida que no diferencia entre Mérida y Sierra Papacal, es decir, puede diferenciar una ciudad que está en un clima subhúmedo y árido, pero no llega a hacer la distinción entre un clima semiárido y subhúmedo. Sin embargo, encontramos que puede diferenciar entre los demás pares, incluso entre Progreso y Sisal, donde todas las demás medidas no pasaron la prueba estadística.

3.4. Discusión

3.4.1. Invariancia en el tiempo

Cuando comparamos el comportamiento de las distribuciones de grado, encontramos que recuperamos los mismos patrones en el análisis que se hizo en 2012 y el análisis de 2015-2016.

Mérida tiene una campana mientras que las ciudades costeras no la tienen o es muy cercana al eje y. Esto quiere decir que las formas de la distribución de grado son una característica que no depende del tiempo en el cual se tomaron, sino que la geografía determina la dinámica de la temperatura.

Aunque necesitamos más datos en el año 2012 para hacer una comparación más robusta, podemos darnos una idea de que las distribuciones de grado pueden ser un tipo de huella digital de cómo es la dinámica de cada ciudad. Cuando se analizan las medidas de las redes,

encontramos que la ciudad costera posee un grado medio menor que la ciudad no costera. Las magnitudes son muy semejantes, 45 de Mérida 2015-2016 vs 43 de Mérida 2012.

Esta posible invariancia en el tiempo que estamos observando en la forma de las distribuciones de grado y las medidas de las redes se debe a que la forma en la que se comporta la dinámica no cambia dado que las condiciones geográficas desde 2012 a 2015 no han cambiado de forma tal que se vea reflejada en la serie de tiempo y pueda ser detectada en las redes. Sin embargo, esto se debe de comprobar con un análisis más robusto usando una ventana de tiempo más amplia.

3.4.2. Diferencia entre dinámicas

Como resultado principal y en respuesta a una de las preguntas que motivaron este trabajo es que el análisis de redes tiene la capacidad de distinguir entre 2 dinámicas distintas. En este caso, el grado medio, la entropía y la Energía Laplaciana fueron capaces de distinguir con un 95 % de confianza entre las ciudades que pertenecían a diferentes regiones climáticas.

El grado medio y la energía están relacionados dado que ambas dependen del conteo del número de vecinos, por esta razón encontramos que ambas tienen la misma tendencia. Sin embargo, el grado medio puede que no sea una medida válida o representativa de la vecindad de los nodos debido a las distribuciones que estamos encontrando. Por lo que LEN puede ser una mejor medida sobre la conectividad de los nodos.

En las Figuras 2 y 6 se observa que la serie de tiempo de Progreso tiene más variación que Mérida. En la tabla 4 encontramos que la serie de tiempo de Progreso produce redes con una Energía más baja que Mérida. Por lo que se podría decir que una serie de tiempo periódica, mientras más fluctuaciones tenga, generarán redes menos energéticas. Esto se explica en el siguiente ejemplo:

Supongamos que tenemos una serie de tiempo que posee 5 datos y que al transformarla obtenemos una red completa (Figura 9 derecha) con un LEN de 20, pero cuando los datos 2 y 4 aumentan (Figura 9 izquierda), el enlace entre estos dos nodos ya no se da y entonces la Energía Laplaciana cae a 16.8. Las variaciones en la serie de tiempo hacen que haya menos visibilidad entre los nodos y por lo tanto varios enlaces ya no pueden darse, disminuyendo el número de vecinos del nodo y por lo tanto su energía.

La entropía mide qué tan regular es una red. Si la entropía es cero, quiere decir que todos

los nodos tienen el mismo número de vecinos, si la entropía crece, quiere decir que aumenta la variedad de número de vecinos en la distribución de grado. Visto desde la interpretación de la teoría de la información de Shannon, la entropía mide qué tanta incertidumbre existe en el sistema usando la forma de la distribución de probabilidad, en nuestro caso, la distribución de grado, el cual es la probabilidad de obtener un nodo con vecindad k . Una red regular, con una distribución de grado parecida a una Delta de Dirac tendrá $Ent=0$ ya que no importa qué nodo escojamos, va a tener la misma vecindad. Cuando el número de vecinos por nodo comienza a crecer en variedad, entonces es cuando la incertidumbre crece y la entropía aumenta así como se muestra en el ejemplo de la Figura 9.

La distribución de grado de Mérida se observan dos máximos, ($k = 2$ y $k = 45$), mientras que Progreso sólo tiene uno el cual se ubica en $k = 2$. En la red de Mérida hay más incertidumbre que en la de Progreso y por lo tanto, la entropía es mayor.

El primer máximo de la distribución de grado proviene de los datos que se encuentran “atrapados” por fluctuaciones, así como se muestra en la Figura 10. Estos nodos no pueden ver más que a sus vecinos inmediatos en la serie de tiempo. Esto también explica por qué estas redes tienen coeficientes de clustering tan altos (Progreso del 0.62 Mérida 0.56). Se observa que la probabilidad de encontrar estos nodos “atrapados” es mayor en Progreso (0.04) que en Mérida (0.02), lo cual refleja una mayor fluctuación en Progreso originada por las brisas marinas (Figura 10).

Debido a que Mérida no presenta estas estructuras tan aleatorias, existe una “conectividad típica” entre los datos que pertenecen al descenso y ascenso en las curvas de temperatura, lugar donde ocurren la mayoría de los enlaces. Por otra parte, los ascensos y descensos de las curvas en Progreso no son uniformes y por lo tanto no tenemos una visibilidad típica (Figura 6).

Al principio parece contra intuitivo que una serie de tiempo regular como Mérida genere redes con más incertidumbre que una serie de tiempo como Progreso que tiene más irregularidades. Una vez que se hace el análisis anterior, se clarifica la interpretación. Sin embargo, es preciso generar un coeficiente que se pueda interpretar de forma más directa.

Para esto se propone el cálculo de ΔEnt , el cual es el valor absoluto de la diferencia de entropías entre una serie de tiempo y su respectiva aleatorización. De esta forma podemos ver qué tan lejos está la serie de tiempo original de su contraparte aleatoria y entonces poder tener una interpretación directa sobre los datos.

En Tabla 3 encontramos que Mérida es quien posee el valor de ΔEnt mayor seguido de

Sierra Papacal y finalmente Progreso y Sisal, es decir que Mérida está más alejado de ser una dinámica aleatoria y Progreso es el más cercano, por lo que seguimos corroborando los hallazgos que estamos reportando en los diferentes análisis presentados. Otro dato importante es que seguimos manteniendo la diferenciación validada por la prueba t de medias independientes entre ciudades contiguas para ΔEnt .

En este estudio encontramos que, en un ejercicio de caracterización de series de tiempo, el análisis usando NVA nos ayuda a ver cómo la aleatoriedad de las brisas marinas genera una aleatorización en las series de tiempo de temperatura generando redes menos energéticas y con menor entropía. Se observa cómo este efecto se va perdiendo mientras vamos tierra adentro hasta que es casi Nula como en el caso de Mérida. Para la caracterización es mejor usar ΔEnt para la asociación con la dinámica.

Si el ejercicio que se quiere hacer es de clasificación, el estudio muestra que usando NVA junto con LEN, Ent ó ΔEnt , se puede hacer sin problema.

4. Señales fisiológicas y redes

4.1. Introducción

En los capítulos anteriores se mostró que el análisis de series de tiempo usando el algoritmo de transformación NVA y caracterizando las redes usando Entropía de Shannon y LEN se obtuvo un buen desempeño superando incluso a técnicas bien establecidas como ApEnt y DFA. Además de la descripción de la dinámica de las series de tiempo, la técnica fue capaz de diferenciar entre ciudades ubicadas en distintas zonas geográficas a pesar de que la distancia entre ellas era alrededor de 15km.

Dada la eficacia mostrada por esta técnica y sus cuantificadores, nos dimos a la tarea de expandir su aplicación hacia el área médica, donde el objetivo principal del análisis de series de tiempo es poder predecir si una persona tiene tendencia a padecer cierta enfermedad degenerativa (Parkinson, Cáncer, Alzheimer) o tener un evento catastrófico (derrame cerebral, paro cardíaco, ataques epilépticos) y cuándo sucederá este evento con tan sólo analizar las señales fisiológicas del paciente.

En diversos estudios llevados a cabo por Peng et al. [42] y Goldberger et al. [43] encontraron que las series de tiempo provenientes de procesos fisiológicos de personas sanas, exhiben propiedades caóticas y fractales en su estructura. Cuando una persona envejece o se enferma, estas propiedades se pierden. Este fenómeno se ha encontrado en señales en el comportamiento de la motricidad del paciente [44; 45] en la variabilidad del ritmo cardíaco (HRV por sus siglas en inglés) [46; 47] y en la dinámica de las señales cerebrales (Electroencefalograma) [16].

Lipstiz y Goldberger [48] proponen que una persona saludable y joven es un sistema con complejidad máxima, ya que es el estado en el cual la capacidad de responder a los cambios en su entorno es mayor, los infantes apenas están desarrollando estas capacidades mientras que los adultos mayores ven disminuidas dichas facultades con el paso del tiempo. Esta constante adaptación al medio genera que las señales fisiológicas tengan correlaciones de largo alcance y fractalidad. Un sistema que pierde esta habilidad de adaptarse, comienza a generar señales con cierto grado de aleatoriedad o regularidad.

Como se mencionó en el Capítulo 1, DFA es una técnica con el enfoque de autoafinidad que calcula las correlaciones que existen en diferentes ventanas de tiempo por lo que es una técnica muy usada en el área.

Dentro de los algoritmos no lineales que se han aplicado a señales fisiológicas para su

caracterización, se encuentra el algoritmo de visibilidad. La variable de respuesta es la fractalidad de la red, la cual se puede calcular a partir de la pendiente que se forma al graficar $\log P(k)$ vs $\log 1/k$ obtenida de la distribución de grado de la red. Este valor es llamado PSVG (Power of Scale-freeness of Visibility Graph por sus siglas en inglés) y es una variante del método originalmente desarrollado por Lacasa et al. [49]. Este procedimiento se ha aplicado con éxito en el análisis fractal en procesos de producción multipartícula [50; 51; 52; 53].

Para las aplicaciones médicas, Nilanjana et al. [54] analizaron 12 horas de una señal obtenida por electrocardiograma (ECG) de 23 pacientes que experimentaron ataques al corazón repentinos y los comparó con 18 pacientes sanos. Para cada paciente se dividió la serie de tiempo en ventanas iguales de 1 minuto y mapeó cada una a red de forma tal que para cada minuto se calculó PSVG. Posteriormente se hizo un promedio de 30 datos (30 minutos de ECG) para cada paciente, es decir, en un lapso de 12 horas, se obtuvieron 24 PSVG's. Se encontró que los pacientes sanos obtenían un PSVG mayor que los que tuvieron el ataque. Es decir que los pacientes sanos generan una red con mayor fractalidad.

Bhaduri et al. [55] aplicó el mismo método usando PSVG para analizar las diferencias entre la variabilidad cardiaca de una persona cuando medita y cuando se encuentra en sus actividades cotidianas, los datos se encuentran en el banco de datos de Physionet [56]. En el artículo no se detalla un tratamiento previo a las series de tiempo ni estandarización en la longitud. Se reportó que el valor de PSVG era mayor en el momento en que la persona meditaba.

El mismo set de datos fue utilizado por Nasrolahzadeh et al. [57] aplicando la transformación a redes, pero se usó un coeficiente llamado Índice de Complejidad de Red (GIC por sus siglas en inglés) el cual fue propuesto por Kim et al. [58]. Este índice se construye de la siguiente manera: si λ_{max} es el eigenvalor máximo de la matriz de adyacencia de una red con N nodos, entonces GIC se calcula como sigue:

$$GIC_{\lambda_{max}} = 4c(1 - c) \quad (14)$$

donde

$$c = \frac{\lambda_{max} - 2\cos\left(\frac{\pi}{n+1}\right)}{n - 1 - 2\cos\left(\frac{\pi}{n+1}\right)} \quad (15)$$

El tratamiento de los datos para este estudio fue el siguiente: se fragmentaron las series de tiempo en ventanas no sobrelapadas de 500 puntos que fueron mapeados a una red y posteriormente se obtuvo el GIC. También se hizo el cálculo para las series de tiempo completas.

Se encontró que, para el tratamiento de las series de tiempo fragmentadas en 500 puntos, el GIC era mayor en el momento de meditación que cuando la persona estaba en otra actividad. Sin embargo, cuando se hace el análisis de toda la serie de tiempo completa, no se encuentra una diferencia estadísticamente significativa.

En este capítulo se usó el mismo conjunto de datos de las personas dentro y fuera de meditación con el objetivo de comparar cómo es el performance de LEN y Ent con los hallazgos reportados por Nasrolahzadeh et al. [57] y Bhaduri et al. [55].

4.2. Metodología

Los datos se tomaron del banco de Physionet [56] basado en el artículo de Goldberger et al. [59] donde se midió la frecuencia cardíaca de 8 meditadores de Chi, 5 mujeres y 3 hombres (rango de edad 26-35, con una edad media de 29 años) usando un Holter durante aproximadamente 10 h mientras realizaban sus actividades cotidianas. Aproximadamente a las 5 h de grabación, cada uno de ellos practicó 1 h de meditación. Durante estas sesiones, los meditadores de Chi se sentaron en silencio, escuchando la guía grabada del maestro. Los meditadores recibieron instrucciones de respirar espontáneamente mientras visualizaban la apertura y el cierre de un loto perfecto en el estómago. La sesión de meditación duró aproximadamente 1 h.

Dado que el método que se propone es sensible a diferencias en las longitudes de las series de tiempo, se decidió estandarizar el tamaño a 1000 datos los cuales representan un total de 16.6 minutos (Figura 11). Para la serie de tiempo de meditación se tomó a partir del minuto 30 hasta el minuto 45 ya que es el tiempo en el que se observó mayor estabilización en el periodo de meditación de todos los participantes. Para la serie de tiempo fuera de meditación se tomaron 15 minutos donde las series de tiempo fueran los más similares posible.

El suavizado de la serie de tiempo se hizo con el comando *smooth* en el lenguaje de programación *R*. Este comando está basado en el algoritmo *Running Medians* de Tukey [60].

Para el análisis de diferencias, se utilizó la prueba *t* pero en este caso, a diferencia del clima, se usó la de medias dependientes ya que estamos hablando de la misma persona en diferentes situaciones.

4.3. Resultados

En la Tabla 5 se reporta por paciente la media de la variabilidad cardiaca $\langle HVA \rangle$, ApEnt y DFA. Encontramos que antes de meditar, los pacientes tienen un ritmo cardiaco más elevado que cuando están meditando con excepción del paciente 8. Este fenómeno se da gracias a que el sistema respiratorio está acoplado fuertemente al corazón, si hay más requerimiento de oxígeno en el cuerpo, la frecuencia respiratoria aumenta y por lo tanto el bombeo de sangre debe aumentar también ya que es el que transporta el O_2 que ingresa y el CO_2 que se expulsa, Sin embargo, cuando se analiza si las diferencias son significativas, encontramos que no se pueden distinguir entre las dinámicas.

Cuando se analiza la regularidad de los patrones en la estructura de las series de tiempo usando ApEnt encontramos que en meditación son menos regulares que antes de meditar. Dichas diferencias fueron significativas.

Sobre la correlación entre ventanas de tiempo, el análisis de DFA muestra que antes de meditar la respiración se comporta como un proceso fbm antipersistente excepto para el Paciente 1 el cual genera una señal fgn. Por otra parte, las correlaciones de largo alcance se pierden en meditación. Las diferencias antes y durante meditación resultaron ser significativas.

Tabla 5: Resultados de la caracterización de las series de tiempo para cada paciente antes (pre) y durante (med) la sesión de meditación.

ID Paciente	$\langle HVA \rangle$		ApEnt		DFA	
	Pre	Med	Pre	Med	Pre	Med
1	69.01	74.46	1.286	1.081	0.818	0.666
2	94.49	82.14	0.5562	1.076	1.148	0.583
3	95.58	76.86	1.0742	1.166	1.162	0.965
4	99.44	73.20	0.6559	1.088	1.131	0.543
5	79.84	66.96	0.8870	1.183	1.223	0.642
6	80.85	67.22	0.9502	1.388	1.074	0.716
7	89.01	77.77	0.6217	1.127	1.268	1.084
8	81.22	100.2	1.2430	1.272	1.198	0.803
Promedio	86.18	77.35	0.9093	1.178	1.128	0.7505
t-test	0.1248		0.0275*		0.0007*	

*Valores con significancia estadística del 95 % de confianza.

En la Tabla 6 se muestran los resultados obtenidos usando las redes en donde se reporta que ninguna de las medidas fue capaz de distinguir las dinámicas. Para LEN encontramos que todos los pacientes aumentan la energía de la red cuando están en meditación excepto los pacientes 6 y 7.

El valor de Ent los valores aumentan en meditación exceptuando los pacientes 4 y 6. Para los resultados obtenidos para ΔEnt encontramos que los pacientes 6 y 7 no hay cambio en sus estados, mientras que los demás generan un cambio muy pequeño. Como era de esperarse, no se encontraron diferencias estadísticas entre las poblaciones.

Tabla 6: Promedio de las medidas de las redes para cada paciente usando redes.

ID Paciente	LEN		Ent		ΔEnt	
	Pre	Med	Pre	Med	Pre	Med
1	135.26	205.81	1.264	1.353	0.178	0.309
2	139.98	202.29	1.243	1.357	0.209	0.306
3	150.80	189.41	1.234	1.369	0.306	0.294
4	170.68	179.71	1.331	1.280	0.294	0.233
5	156.42	182.39	1.268	1.335	0.233	0.239
6	164.13	121.53	1.301	1.242	0.179	0.179
7	154.28	153.23	1.192	1.233	0.144	0.144
8	160.53	183.43	1.272	1.366	0.289	0.290
Promedio	153.97	177.21	1.264	1.317	0.1964	0.2496
t-test	0.01118		0.07145		0.07923	

*Valores con significancia estadística del 95 % de confianza.

Cuando se analizan los resultados de las series de tiempo suavizadas (Tabla 7), encontramos que $\langle HRV \rangle$ y DFA generan valores muy similares antes y después del suavizado. Sin embargo, para el cálculo de ApEnt, los valores cambian de forma notable, disminuyendo su magnitud después del suavizado. Esto es de esperarse ya que con este tratamiento, la serie de tiempo deja de tener tanto ruido interno y las estructuras se vuelven más "lisas" de forma tal que existirán más patrones similares. Seguimos encontrando que las diferencias estadísticas se preservan según lo discutido en la tabla anterior.

Cuando se hace el análisis usando redes para la serie de tiempo suavizada (Tabla 8), los valores cambian de forma notable al punto de que LEN, Ent y ΔEnt sí pueden distinguir entre dinámicas según los resultados de la prueba t .

Las redes que pertenecen a las series de tiempo durante meditación tienen en promedio una LEN de 261 vs 433 fuera de meditación, lo que nos dice que la serie de tiempo en meditación tiene más variabilidad. En cuanto a la entropía se muestra que ΔEnt es mayor antes de la meditación que durante lo que significa mayor variación en la serie de tiempo de Meditación. Con estos resultados se alcanza a reproducir la interpretación encontrada por DFA y ApEnt antes y después del suavizado. Es decir, en meditación hay más aleatoriedad que fuera de meditación.

Tabla 7: Resultados de la caracterización de las series de tiempo suavizadas para cada paciente antes (pre) y durante (med) la sesión de meditación.

ID Paciente	$\langle HVA \rangle$		ApEnt		DFA	
	Pre	Med	Pre	Med	Pre	Med
1	68.53	74.44	0.771	0.835	0.899	0.683
2	94.24	82.15	0.194	0.843	1.211	0.595
3	95.42	76.92	0.390	0.860	1.240	1.007
4	99.48	73.16	0.330	0.913	1.182	0.553
5	79.66	66.96	0.387	0.892	1.313	0.692
6	80.67	67.22	0.581	0.861	1.139	0.786
7	88.52	78.69	0.207	0.502	1.338	1.143
8	81.01	100.2	0.726	0.982	1.249	0.838
Promedio	85.94	77.45	0.448	0.836	1.196	0.787
t-test	0.1396		0.0083*		0.0005*	

*Valores con significancia estadística del 95 % de confianza.

4.4. Discusión

El primer punto a tocar es que en este ejercicio no se pudo diferenciar entre las dinámicas dentro y fuera de meditación para cada paciente usando redes, mientras que ApEnt y DFA sin problemas detectan las diferencias. Hasta que se suavizan las series de tiempo es que los cálculos de LEN y Ent son capaces de diferenciar entre estos dos estados.

En este experimento numérico evidenciamos una limitante para el análisis de series de tiempo usando el algoritmo de transformación NVA y los cuantificadores LEN y Ent. Esta

Tabla 8: Promedio de las medidas de las redes para las series de tiempo suavizadas.

ID Paciente	LEN		Ent		Δ Ent	
	Pre	Med	Pre	Med	Pre	Med
1	252.08	239.71	1.392	1.334	0.299	0.300
2	404.03	256.05	1.472	1.328	0.450	0.280
3	527.28	306.97	1.504	1.442	0.464	0.364
4	477.35	204.33	1.524	1.254	0.456	0.217
5	498.30	225.35	1.509	1.332	0.398	0.238
6	346.95	172.16	1.457	1.276	0.395	0.211
7	626.86	577.84	1.550	1.553	0.463	0.468
8	338.57	268.12	1.441	1.388	0.345	0.329
Promedio	433.93	281.30	1.481	1.363	0.409	0.301
t-test	0.00366*		0.0076*		0.00452*	

*Valores con significancia estadística del 95 % de confianza.

limitante radica en el nivel de ruido de la serie la cual evita que se puedan detectar los cambios en la estructura de los datos. Cuando se hace el suavizado, se limpia la serie de tiempo, haciendo posible su cuantificación mediante el uso de redes.

Por otra parte, Según lo mostrado en la introducción de este capítulo, Nasrolahzadeh et al. [57] y Bhaduri et al. [55] trabajaron con el mismo conjunto de datos obteniendo una diferencia entre estados. En el trabajo de Nasrolahzadeh et al. [57] se utilizó GIC para cuantificar las diferencias en complejidad transformando la serie de tiempo completa y en otro experimento lo hizo por ventanas, encontrando diferencias estadísticas sólo en el segundo ejercicio. Por lo que es posible que, si se hace un análisis en ventanas pudiésemos encontrar diferencias estadísticas entre los diferentes estados de los meditadores. Sin embargo, no se encuentran elementos teóricos como para establecer qué tamaño de ventanas usar, si las ventanas deben sobrelaparse o no, si no se sobrelapan como en el análisis de Nasrolahzadeh et al. [57] ¿qué pasa con la visibilidad entre los elementos que no tienen oportunidad de interactuar? Es posible que las diferencias encontradas se deban más a un efecto del tratamiento de los datos y no de la dinámica en sí. Para corroborar que estamos capturando sólo la fenomenología de la dinámica, el análisis se debió acompañar de los resultados de técnicas alternativas.

Sin embargo, Bhaduri et al. [55] encuentran diferencias entre dinámicas al introducir la serie de tiempo completa y sin un pre-tratamiento. Sus datos revelan que las redes tienen un carácter fractal mayor cuando provienen de una serie de tiempo en meditación que cuando

no lo está. Es interesante este resultado ya que nos da indicios sobre la propuesta inicial de Lacasa et al. cuando menciona que la transformación conserva toda la información de la serie de tiempo y la deposita en la estructura de la red, el reto es encontrar qué método aplicar.

Si bien es cierto que en el análisis presentado en este trabajo se mostró que tras el suavizado de los datos, los valores numéricos de $\langle HVA \rangle$ y DFA no varían mucho y que la interpretación sobre cuál serie de tiempo es más aleatoria, se mantuvo, es mejor operar con los datos tal cual se obtuvieron del sistema para evitar agregar o quitar elementos que nos puedan llevar a una conclusión errónea. Hemos de recordar que en estos ejemplos sabemos a priori cómo es el sistema, si estuviésemos sin esta información como pasa en la vida real de un Médico, habría más incertidumbre de qué pre-tratamiento hacer a los datos para encontrar diferencias, o ventanas de qué longitud usar y si se van a sobrelapar o no. Por lo que es más recomendable dejar la serie de tiempo tal como es y aplicar diferentes técnicas para tener la fotografía completa de la dinámica, la experimentación con el suavizado sería un recurso complementario.

5. Termodinámica y redes en el contexto del algoritmo de visibilidad

5.1. Introducción

A lo largo de este trabajo se ha hablado de la Energía (LEN) y Entropía(Ent) de las redes y su interpretación en cuanto qué tan aleatoria es la dinámica. Sin embargo, estos términos sugieren de forma natural que se puede hacer una analogía con la termodinámica, lo que permitiría implementar conceptos como transición de fase y criticalidad auto-organizada [61; 62; 63] para poder anticiparnos a eventos catastróficas a partir de las series de tiempo [64; 65], pero desde el enfoque de redes.

¿Podemos relacionar las cantidades calculadas Ent y LEN con un sentido termodinámico? Para contestar la pregunta comenzaremos a analizar los límites de cada variable. Con respecto a la energía, su valor máximo corresponde a una serie de tiempo que pueda generar una red completa. Es posible que una serie de datos que tenga la forma de una “U”, como la función x^2 , pueda generar estas redes. El mínimo de energía sería la red generada por la serie de tiempo completamente aleatoria.

La entropía generada por la red completa sería cero dado que todos los nodos tendrían el mismo grado k . La energía de la red proveniente de la serie de tiempo aleatoria sería alrededor de uno dado los resultados numéricos que se han obtenido hasta ahora. El máximo de entropía sería aquella red que tuviese una distribución de grado tipo uniforme, donde todos los grados de la red tengan el mismo número de nodos, es decir la misma probabilidad de ser elegidos al azar. Dilucidar las características que posee una serie de tiempo para generar este tipo de redes es una tarea no trivial y la cual se deja como un trabajo futuro.

Otro enfoque que podemos abordar para analizar las relaciones entre las variables calculadas en este trabajo y la termodinámica es a partir de la mecánica estadística aplicada a redes. Cimini et al. [66] y Park et al. [67] proponen un marco teórico llamado ERG (Exponential Random Graphs por sus siglas en inglés) donde aplican los conceptos de la mecánica estadística a la teoría de redes.

La propuesta consiste en modelar un sistema de redes G a través de un conjunto Ω de redes con el mismo número N de nodos. El modelo está especificado por $P(G)$, la probabilidad de encontrar una red $G \in \Omega$. La mejor opción de distribución de probabilidad es el que maximiza la entropía de Gibbs:

$$S = - \sum P(G) \ln P(G) \quad (16)$$

Aplicando la técnica de los multiplicadores de Lagrange teniendo como restricciones:

$$\sum_G P(G) = 1 \quad (17)$$

$$\sum_G P(G) x_i(G) = \langle x_i \rangle \quad (18)$$

donde $\langle x_i \rangle$ es el valor medio de x_i el cual se puede interpretar como un observable en G , por ejemplo el número de enlaces o el grado medio. La probabilidad de encontrar una red G viene dada por:

$$P(G) = \frac{e^{-H(G)}}{Z} \quad (19)$$

donde $H(G)$ es el Hamiltoniano de la red definido de la siguiente forma:

$$H(G) = \sum_i \theta_i x_i \quad (20)$$

θ_i se puede pensar como un parámetro acoplado al observable x_i el cual puede ser expresado en términos de variables que ya conocemos o se pueden calcular. De esta forma, la función de partición queda de la siguiente manera:

$$Z = \sum_G e^{-H(G)} \quad (21)$$

Para poder contestar la pregunta que se planteó al principio de este capítulo sobre el significado de los cálculos de Ent y LEN se plantean las siguientes restricciones para poder simplificar los cálculos y llegar al objetivo de esta sección.

5.2. Modelo Teórico

Para un sistema de N nodos distinguibles y m enlaces, el número de grafos Ω que podemos obtener (multiplicidad) se calcula de la siguiente manera:

$$\Omega = \frac{M!}{m!(M-m)!} \quad (22)$$

donde M representa el cálculo de las diferentes formas en las que podemos acomodar los m enlaces en los N nodos, visto de otra manera, sería contar el número de formas en las que podemos tomar 2 nodos sin reemplazo de una caja que contiene N nodos. El cálculo es de la siguiente forma:

$$M = \frac{N!}{2!(N-2)!} \quad (23)$$

Una vez que tenemos la multiplicidad de estados podemos calcular la entropía usando la ecuación:

$$S = \ln \Omega \quad (24)$$

aplicando la aproximación de Stirling obtenemos la siguiente expresión la cual relaciona la entropía del sistema con el número de enlaces que tiene:

$$S(G(N, m)) = M \ln M - m \ln m - (M - m) \ln(M - m) \quad (25)$$

Con estas ecuaciones generamos un modelo con límites superiores e inferiores definidos para la entropía, cuya forma se muestra en la Figura 12

En las derivaciones de los modelos clásicos como el Sólido de Einstein o la ecuación de Sakur-Tetrode, el término de multiplicidad se construye de forma tal que la ecuación de la entropía queda naturalmente expresada como función de la energía. En nuestro caso, la entropía resulta una función de m . Sin embargo, La definición de energía de la Ecuación 12 está en función de m y de la suma del cuadrado del grado de los nodos. Este segundo término se puede aproximar de la siguiente forma:

$$\sum_{i=1}^N k_i^2 \approx \beta m^2 + \gamma m \quad (26)$$

finalmente la ecuación de energía con respecto a los enlaces tiene la forma:

$$E(G(N, m)) = \beta m^2 + \gamma m \quad (27)$$

donde $\beta = \frac{4}{N}$ y $\gamma = \frac{12}{N}$

En la gráfica de la Figura 13 se muestra el comportamiento teórico de S vs E usando la aproximación de la Ecuación 27. Encontramos que dada la forma cuadrática de E se pierde la simetría que habíamos reportado para S en la Figura 12.

Para saber qué tan bien funciona la aproximación para la energía, se tomaron redes con distintas N y se calculó S y E a partir de m y también se calculó LE a partir de la definición original (Ecuación 12). En la Figura 14 se muestra el comportamiento de las variables, en donde encontramos que la aproximación funciona muy bien al tener poca desviación con respecto al real. Se observa que mientras mayor N mejor es la aproximación.

Ya que se tiene el comportamiento de la Entropía del sistema con respecto a la Energía, podemos aplicar la definición clásica de temperatura:

$$\frac{1}{T} = \frac{\delta S}{\delta U} \quad (28)$$

Lo que nos dice esta ecuación es que el inverso de la temperatura es la pendiente de la curva de la Figura 13. Cuando analizamos de izquierda a derecha, encontramos que existe una pendiente positiva, luego se hace cero en el máximo y la pendiente cambia de signo, señalando la existencia de temperaturas negativas.

5.3. Cálculo de temperaturas

Ahora que hemos validado las ecuaciones derivadas a partir de la multiplicidad, procedemos a calcular este nuevo término de entropía y analizar las estructuras en las redes de las series de tiempo de Clima (Tablas 9 y 10) y Meditación (Tabla 11):

Los resultados muestran el mismo patrón que hemos discutido en los capítulos anteriores, en donde la entropía es mayor en Mérida que en Progreso. En la Tabla 10 encontramos que todos los pares son significativamente diferentes excepto Sisal y Progreso, como se reportó en los capítulos anteriores. La entropía máxima para una red con N=4032 es 1397 por nodo y

Tabla 9: Promedios calculados a partir del modelo teórico para las redes reales del Clima.

	Merida	Progreso	SierraP	Sisal
m	90860	60459	80120	64990
S/N	124	88	111	94
LEN(G)	2919	1713	2397	1936
T	23.5	19.46	21.59	20.60

LEN es $4,1 \times 10^6$, lo cual indica que es una red poco conectada con respecto a su nivel máximo y que estamos en la zona de temperaturas positivas.

En relación con las series de tiempo, encontramos que mientras más variabilidad en los datos, obtenemos redes con temperaturas más bajas que aquellas que tienen una dinámica más regular.

Tabla 10: Reporte de los valores de significancia para cada par de ciudades contiguas.

	m	S	T
Merida-SierraP	0.000049*	0.000047*	0.000037*
Sierra-Progreso	0.00001*	0.00001*	0.00001*
Progreso-Sisal	0.114168	0.111399*	0.110886
Merida-Progreso	0.00001*	0.00001*	0.00001*

*Valores con significancia estadística del 95 % de confianza.

Para los datos suavizados de Meditación se muestra en la Tabla 11 en donde encontramos la misma tendencia reportada en los resultados del Capítulo 5, encontrando diferencias significativas ente los estados pre y med. Mientras más variabilidad, tenemos una entropía mayor y una temperatura menor, así como en el caso del clima. La entropía máxima para una red con $N=1000$ es 346 por nodo y la LEN es $6,2 \times 10^4$, lo cual indica que es una red poco conectada con respecto a su nivel máximo y que estamos en la zona de temperaturas positivas.

5.4. Discusión

La experiencia que se tiene día a día, es que si a un sistema comenzamos aumentamos su energía agregando alguna flama, la temperatura comenzaría a subir sin límite alguno. Sin embargo, cuando tenemos definido un límite superior para la energía, es posible tener una

Tabla 11: Promedio de los cálculos del número de enlaces (m), entropía (S) y temperatura (T) y el resultado de la prueba t para los pacientes antes (Pre) y durante (Med) la sesión de meditación.

ID Paciente	m		S		T	
	Pre	Med	Pre	Med	Pre	Med
Promedio	7805	6856	40171	36158	10.65	7.57
t-test	0.02232*		0.02302*		0.00918*	

*Valores con significancia estadística del 95 % de confianza.

disminución de la entropía del sistema y por lo tanto la existencia de temperaturas negativas.

Para explicar este punto pensemos en un sistema de bosones con N partículas. En el límite inferior de temperatura encontramos que todas las partículas tienen la misma energía, observaríamos que se “acomodan” en el estado base (Figura 15, Sistema Termodinámico). Cuando se introduce energía al sistema, las partículas comienzan a alcanzar otros niveles superiores, haciendo que la entropía aumente. Si el sistema no tiene un límite superior, las partículas seguirán subiendo de niveles de energía. Sin embargo, cuando existe un límite superior, las partículas comienzan a ocupar el nivel máximo de energía disponible del sistema hasta que todas llegan a ese nivel, generando un sistema con menos entropía. El comportamiento de la entropía con respecto a la energía es una parábola, y como se mencionó anteriormente, la pendiente de esa curva nos indica el signo de la Temperatura, resultando en la posibilidad de tener temperaturas negativas.

Cuando se hace la analogía con las redes, podemos pensar en un sistema donde sólo existen nodos sin ningún enlace, lo que generaría una entropía y energías mínimas (Figura 15 Redes). Posteriormente, comenzamos a agregar enlaces de forma aleatoria sin permitir enlaces múltiples ni loops. El cálculo de energía de la red comenzaría a aumentar así como el de su entropía. Sin embargo, encontraremos un punto en el que la energía sigue aumentando pero la entropía comienza a descender hasta que llega a cero. Este punto es cuando la red es completa. Como se ha mencionado a lo largo de este capítulo, la gráfica de la entropía con respecto a la energía sigue siendo una parábola y como en el ejemplo anterior encontrando temperaturas negativas.

Hasta ahora este es un hecho meramente teórico el cual tiene mucho campo de exploración para el área de redes. ¿Qué significa una red con temperatura negativa? hasta el momento se puede decir que es un indicio de que es una red con una conectividad muy alta entre los nodos, de forma tal que cierta proporción de nodos está conectado con todos los demás. Sin embargo,

falta profundizar más las propiedades de estas redes de forma tal que se pueda caracterizar estos umbrales.

Por otra parte, creemos que lo más importante a resaltar de lo desarrollado en este capítulo, es que se puede formular la siguiente conjetura: en el límite donde la pendiente de la curva es cero, la complejidad de la red es máxima. Habría que diseñar más experimentos y generar más teoría para poder comprobar esto, sin embargo, creo que esto puede ser cierto ya que la complejidad es un comportamiento que ha mostrado ser un equilibrio homeostático del sistema entre el orden y la aleatoriedad.

Otro de los temas centrales para discutir es la diferencia entre Ent calculado a partir de la Ecuación 10 y S calculada a partir de la Ecuación 25. El primero nos indica qué tanta variedad de nodos con k vecinos existen en la red. Una red con entropía cero indica que todos los nodos tienen el mismo k mientras que la máxima entropía sería una red donde todos los grados existentes posean la misma cantidad de nodos (una función uniforme en la distribución de grado).

La segunda definición nos indica cuántas redes se pueden formar con m enlaces en un grafo de N elementos. El máximo de entropía se centra en $M/2$ (donde M se calcula con la Ecuación 23).

Si se analiza a partir del modelo teórico desarrollado, Ent está calculando la entropía distribuida en un micro estado en particular, mientras que S calcula la entropía del conjunto de micro estados correspondiente al macro estado $G(m,N)$. Si tomamos como analogía el experimento de la probabilidad de encontrar un par de dados que al sumarlos me resulte una cantidad X de puntos (Figura 16), Ent está midiendo la entropía de un par de dados mientras que S mide la de todos los pares (configuraciones) posibles que suman X cantidad de puntos.

Desde el punto de vista del análisis de series de tiempo, ambas definiciones funcionan bien para caracterizar e identificar diferentes dinámicas.

En cuanto a la energía se podría decir que encontramos una buena analogía con la termodinámica estadística. La aproximación que se propuso para poder calcular la energía a partir del número de enlaces funcionó muy bien, sin embargo, se necesitan más experimentos que exploren las diferentes posibilidades y generar más resultados para una discusión más robusta.

Otro punto interesante es que al tener un sistema con energías acotadas en el mínimo y el máximo, es la posibilidad de tener redes con temperaturas negativas. Esto puede abrir todo un campo de investigación con respecto a la termodinámica de redes, el análisis de la capacidad

calorífica y el símil con sistemas físicos que presentan temperaturas negativas [69].

6. Discusión e integración de resultados

En este Capítulo integraremos los hallazgos discutidos en los capítulos anteriores. Los puntos para discutir en este apartado se resumen en las preguntas ¿En qué casos se puede usar LEN y Ent? ¿Qué estamos midiendo a las series de tiempo usando LEN y Ent? y finalmente ¿Las variables trabajadas a lo largo de este documento pueden ser usados como cuantificadores de complejidad de una dinámica? Sabemos que existen muchas preguntas más que resolver, sin embargo, las propuestas en este capítulo nos ayudarán a tener más claro los alcances y limitaciones del análisis de redes, el cual fue el objetivo de este trabajo.

6.1. ¿En qué casos se puede usar LEN y Ent?

En este trabajo se ha mostrado que el análisis de series de tiempo usando redes permite caracterizar las series de tiempo de temperatura de distintas ciudades mediante el cálculo de LEN y Ent mientras que DFA y ApEnt no distinguieron entre dinámicas.

Por otra parte, se mostró también que la técnica tiene sus limitaciones al no poder identificar entre las series de tiempo dentro y fuera de meditación. Por el contrario, DFA y ApEnt detectaron las diferencias sin problema.

Sin embargo, en el estudio de Bhaduri et al. [55] muestra que con la misma transformación, pero cuantificando la fractalidad de la red usando PSVG, se es capaz de encontrar diferencias consistentes entre ambos estados. Por lo que se concluyó en ese capítulo que la red sí captura la información de la serie de tiempo, sólo que hemos de saber cómo caracterizar la red para obtener la información que deseamos.

La primera pregunta que se va a abordar es ¿Por qué LEN y Ent tuvieron mejor desempeño en los datos de temperatura que en los de frecuencia cardíaca? Si observamos las series de tiempo de ambos sistemas, podemos notar que la diferencia radica en la componente sinusoidal y el tamaño de las fluctuaciones. LEN y Ent funcionan mejor para series de tiempo periódicas con fluctuaciones cortas. Esto se corroboró con el suavizado de la variabilidad cardíaca, encontramos que LEN y Ent pudieron diferenciar entre dinámicas.

Es posible que se pregunte el por qué no se utilizó PSVG en los datos del clima. Para poder aplicar esta técnica, es necesario que la distribución de grado tenga un comportamiento lineal cuando se hace la gráfica $\log(1/k)$ vs $\log P(k)$. Como se muestra en la Figura 8, la forma de la gráfica no cumple dicha característica.

Para responder concretamente la pregunta de este apartado, encontramos que para series de tiempo estacionales y con pocas fluctuaciones usar LEN y Ent, mientras que las series de tiempo contienen muchas fluctuaciones y no tienen estacionalidad ni tendencia, usar PSVG puede ser mejor opción.

6.2. ¿Qué mide el análisis de series de tiempo usando redes?

Para responder esta pregunta, utilizaremos la clasificación de Tang et al. [9] (Ver Figura 3). ApEnt se encuentra catalogado dentro del enfoque de estados desordenados en entropía dinámica y el DFA en el enfoque de autosimilitud en monofractalidad. ¿En dónde se podría clasificar el análisis de redes?

En este trabajo se propone que el análisis de redes se puede clasificar dependiendo de qué cuantificador se esté usando para caracterizar la red. Si se utiliza PSVG, se estaría catalogando como autosimilitud, sin embargo, se necesitarían hacer más estudios para saber si es posible colocarlo en mono o multifractal.

Si se usa Δ Ent, entonces estaríamos pensando en colocarlo en el enfoque de estados desordenados, pero ¿sería entropía estructural o dinámica? Para definirlo discutiremos el significado de dichas opciones. Tang et al. [9] define como entropía estructural al cálculo que se hace a partir del espectro de potencias de una serie de tiempo cuando se mapean los datos al espacio de frecuencias. Lo que se busca medir es qué tan concentrada está la distribución a partir de cálculos de entropía. Por otra parte, la entropía dinámica se refiere a la similitud que existe entre patrones internos en la serie de tiempo (como ApEnt).

Δ Ent no entra en la clasificación propuesta por Tang ya que no estamos pasando al espacio de frecuencias ni estamos comparando patrones internos. Nosotros proponemos incorporar un tercer subtipo llamado “Referenced Network Entropy” dado que estamos obteniendo la entropía de la red cuando transformamos la serie de tiempo y se hace referencia (se compara) con la entropía de la red originada de la serie de tiempo aleatorizada.

Finalmente tenemos LEN, el cual es una medición que nos indica qué tan aleatoria es una serie de tiempo, por lo que podríamos colocarla en el enfoque de estado desordenado. Dado que no puede estar en la categoría de entropía, sugerimos abrir otro tipo al enfoque estado desordenado que se llame Energía (Ver Figura 17).

En resumen, lo que mide las redes de una dinámica es su aleatoriedad visto desde la en-

tropía referenciada y por la energía de la red. Aunque hacen falta más estudios, PSVG estaría midiendo la fractalidad de una serie de tiempo.

6.3. Complejidad, Termodinámica y NVA

En el Capítulo 5 se habló acerca de si las cantidades calculadas en los Capítulos 3 y 4 podrían equipararse al marco teórico de la termodinámica. Se mostró que LEN sí era una cantidad equiparable a la energía, así como la cantidad S, mientras que Ent es una cantidad que pertenece al micro estado.

A pesar de que ambas definiciones de entropía se derivaron de diferente manera, la interpretación de la dinámica, tanto de los datos de temperatura, como los de la variabilidad del ritmo cardiaco, mostraron ser siempre consistentes. Así también se mostró que se podía diferenciar estadísticamente las dinámicas entre sí.

¿Alguna de estas cantidades puede cuantificar la complejidad de una dinámica? En el artículo de Tang et al. [9] se explica el enfoque de estados desordenados y menciona a la termodinámica como un símil natural entre los cálculos hechos usando la entropía y el orden del sistema y que los métodos enumerados en su clasificación pueden medir la complejidad de una serie de tiempo. Sin embargo, este concepto sigue siendo un gran debate en la literatura, ya que no existe ni si quiera una definición aceptada por toda la comunidad científica de lo que es la complejidad de un sistema. A lo más que se ha llegado a aceptar son algunas de las características que comparten. Por lo tanto asegurar que se puede cuantificar, es una aseveración que es imposible mostrar plenamente.

En la discusión del Capítulo 5 sobre la pendiente de la curva S vs LEN, se menciona que puede ser un probable cuantificador de complejidad del sistema, algo similar a la escala de DFA, donde un $\alpha = 1$ indicaba una serie de tiempo como una memoria de muy largo alcance. Sin embargo, como se mencionó en la discusión de ese capítulo, habría que hacer una investigación más profunda sobre este diagrama y las características de las redes obtenidas.

Lo propuesto en este trabajo dista mucho de establecer cuantificadores de la complejidad de un sistema, lo único que se puede decir hasta el momento es que ayudan a cuantificar qué tan alejadas están las dinámicas de la aleatoriedad. Para poder proponerlos como cuantificadores que describan la complejidad de una dinámica haría falta un análisis más profundo.

7. Conclusiones y Trabajos futuros

A lo largo de este trabajo se habló de los resultados que arroja el analizar una serie de tiempo utilizando el algoritmo de visibilidad natural. Para la serie de tiempo de temperatura se encontró que las medidas como LEN y Ent ayudaron a identificar las distintas ciudades a pesar de que la distancia entre ellas era de 15Km y en el caso de las ciudades extremo de 30Km. Cuando se hizo la comparación con técnicas no lineales como DFA y ApEnt, se mostró que el análisis de redes tenía mejor desempeño si la tarea era de clasificación. Se encontró que la dinámica es más aleatoria mientras más cerca estamos de la costa. Esta aleatoriedad se traduce en redes con mayor LEN y menor Ent.

El estudio se amplió a las series de tiempo de la variabilidad del ritmo cardiaco en personas mientras toman una sesión de meditación guiada y mientras están en sus actividades cotidianas. Se encontró que DFA y ApEnt clasificaron a los pacientes de manera correcta, mientras que el análisis de redes no tuvo esa capacidad. No fue hasta que se hizo un suavizado de las series de tiempo que el análisis de las series de tiempo produjo frutos. Para este sistema se hizo relevante que las dinámicas más aleatorias surgían en el proceso de meditación, generando redes con un LEN menor y un Ent menor que cuando estaban fuera de meditación.

Con estos hallazgos se concluyó que el algoritmo de transformación funciona bien para series de tiempo con una componente fuerte de periodicidad, mientras que se recomienda el uso de otras técnicas cuando las series de tiempo tienen mucho ruido (otro trabajo futuro es el establecer cuanto es mucho o poco ruido). Aunque se observó que al suavizar las series de tiempo el análisis de redes tuvo un buen desempeño equiparable a DFA y ApEnt, se recomienda que la técnica sea más como una confirmación de lo que puedan encontrar DFA y ApEnt.

También se pudo concluir que lo que cuantifican las medidas de la red es la variabilidad estructural de una serie de tiempo. Esta cuantificación puede ser apartir de la entropía o energía.

Existen aún muchas preguntas sin responder las cuales pueden traducirse a trabajos futuros, como por ejemplo el comportamiento de las variables termodinámicas en series de tiempo con distintos coeficientes de Hurst lo que derivaría en responder la pregunta ¿Es posible cuantificar la memoria de una serie de tiempo a partir de las variables termodinámicas de la red que genera usando NVA?

Otra pregunta que quedó abierta es sobre la complejidad de un sistema y su cuantificación vía redes. ¿Es posible anticiparse a un evento catastrófico como un sismo o un ataque cardiaco? ¿Se podría ver como una transición de fase? ¿Cómo sería la LEN, Ent S y T de un sistema

simple vs uno complejo?

Un hecho importante es el descubrimiento de que una red puede tener temperaturas negativas. ¿Qué tipo de redes reales están en esta zona? Si se enlaza con las series de tiempo ¿Qué tipo de series de tiempo pueden generar redes con temperaturas negativas? Así también el trabajar sobre el significado de la pendiente de la curva Entropía-Energía y su relación con la complejidad de la estructura de la red

Para terminar este escrito, concluimos que el análisis de series de tiempo usando el algoritmo de visibilidad natural es una técnica que puede usarse como complemento con otras más. Como se ha discutido a lo largo de este documento, distintos análisis extraen una información específica de la dinámica, cuando se juntan todas como complemento, se puede obtener una fotografía enriquecida de todo el sistema, lo cual puede ayudar a una caracterización más profunda y facilitar a generar instrumentos para poder predecir su comportamiento.

8. Referencias

- [1] Veleva L., Pérez G., Acosta M. Statistical analysis of the temperature-humidity complex and time of wetness of a tropical climate in the Yucatan Peninsula in Mexico. *Atmospheric Environment*, 31(5), 773–776.
- [2] Keller C.F. Climate, modeling, and predictability. *Physica D*, vol. 1999; 133, pp. 296–308
- [3] Gerten D. et al Asynchronous exposure to global warming: freshwater resources and terrestrial ecosystems. *Environ. Res. Lett.* 8 (3)
- [4] Mitchell Melanie *Complexity, A guided Tour*. Oxford University Press 2009, Nueva York
- [5] Amaral L.A.N, Ottino J.M. Complex systems and networks: challenges and opportunities for chemical and biological engineers. *Environ. Res. Lett.* 8 (3)
- [6] Rind D. *Complexity and Climate Science* 1999;284,105-107
- [7] Rosen R. *Complexity and System Descriptions Facets of Systems Science* 1999;7, pp 477-482.
- [8] Bradley E., Holger Kantz H. Nonlinear time-series analysis revisited. *Chaos*. 2015;25.
- [9] Tang L., Lv H., Yang F. Yu L. Complexity testing techniques for time series data: A comprehensive literature review. *Chaos, Solitons, Fractals* 2015; 81; 117–135.
- [10] Shuangcheng L, Qiaofu Z, Shaohong W, Erfu D. Measurement of Climate Complexity Using Sample Entropy. *Int. J. Climatol.* 2006;26: 2131–2139.
- [11] Bunde A., Havlin S., Koscielny-Bunde E., and Schellnhuber H.J. The science of disasters. Chapter 5: Atmospheric Persistence Analysis: Novel Approaches and Applications Springer. 2002, 171-189
- [12] Zheng Z. Characteristics of Long-Term Climate Change in Beijing with Detrended Fluctuation Analysis Sixth International Conference on Natural Computation 2010, Yantai, 1084-1088
- [13] Shuangcheng L.I., Qiaofu Z., Shaohong W., Erfu D Measurement of climate complexity using Sample Entropy *Int. J. Climatol.*,2006;26,2131–2139.
- [14] Bryce RM, Sprague KB. Revisiting detrended fluctuation analysis.. *Sci Rep*. 2012;2:315.

- [15] Marmelat V, Torre K, Delignières D. Relative roughness: an index for testing the suitability of the monofractal model. *Front Physiol.* 2012;3:208.
- [16] Hardstone R, Poil SS, Schiavone G, et al. Detrended fluctuation analysis: a scale-free view on neuronal oscillations. *Front Physiol.* 2012;3:450.
- [17] Colley D., Dean R. Origins of 1/f noise in human music performance from short-range autocorrelations related to rhythmic structures . *PLOS.* 2019.
- [18] Irmischer M, Houtman SJ, Mansvelder HD, Tremmel M, Ott U, Linkenkaer-Hansen K. Controlling the Temporal Structure of Brain Oscillations by Focused Attention Meditation . *Hum Brain Mapp.* 2018;39(4):1825-1838..
- [19] Goldberger AL, Amaral LA, Hausdorff JM, Ivanov PCh, Peng CK, Stanley HE. Fractal dynamics in physiology: alterations with disease and aging . *Proc Natl Acad Sci U S A.* 2002;99 Suppl 1(Suppl 1):2466-2472.
- [20] Delgado-Bonal, A., Marshak, A. Approximate Entropy and Sample Entropy: A Comprehensive Tutorial. *Entropy* 2019, 21(6), 541.
- [21] Pincus, S. Approximate entropy (ApEn) as a complexity measure. *Chaos.* 1995;5(1):110-117.
- [22] Rosales-Pérez J.A., Canto-Lugo E., Valdés-Lozano D., Huerta-Quintanilla R. Temperature time series analysis at Yucatan using natural and horizontal visibility algorithms. *PLoS ONE.* 2019, 14(12)
- [23] Garcia E. Modificaciones al Sistema de Clasificación Climática de Köpen. Instituto de Geografía Univeridad Nacional Autónoma de México, 1998
- [24] Zhang J, Small M. Complex network from pseudoperiodic time series: Topology versus dynamics. *Phys Rev Lett.* 2006; 96: 238701.
- [25] Xu XK, Zhang J, Small M. Superfamily phenomena and motifs of networks induced from time series. *Proc Natl Acad Sci USA.* 2008;105:601-605,
- [26] Donner RV, Small M, Donges JF, Marwan, N Y. Zou, Xiang R, Kurths J. Recurrence-based time series analysis by means of complex network methods. *International Int J Bifurcat Chaos.* 2011;21:1019–1046 .
- [27] Lacasa L, Luque B, Ballesteros F, Luque J, Nuño JC. From time series to complex networks: the visibility graph. *PNAS.* 2008;105: 4972-4975.

- [28] Campanharo ASLO, Irmak M, Dean R, Ramos FM, Nunes LA. Duality between Time Series and Networks. PLoS ONE. 2011;6:pp 1-12.
- [29] Liu J, Li Q. Planar Visibility Graph Network Algorithm For Two Dimensional Timeseries. 29th Chinese Control And Decision Conference (CCDC). 2017. doi: 10.1109/CCDC.2017.7978727
- [30] Telesca L, Lovallo M, Aggarwal SK, Khan PK. Precursory signatures in the visibility graph analysis of seismicity: An application to the Kachchh (Western India) seismicity. J. Phys. Chem. Earth. 2015;85:195-200.
- [31] Telesca L, Lovallo M, Toth L. Visibility graph analysis of 2002–2011 Pannonian seismicity. Physica A. 2014;416:219-224.
- [32] Zhuang E, Small M, Feng G. Time series analysis of the developed financial markets' integration using visibility graphs. Physica A. 2014; 410:483–495.
- [33] Yu L. Visibility graph network analysis of gold price time series. Physica A. 2013;392:3374-3384.
- [34] Donner RV, Donges JF. Visibility graph analysis of geophysical time series: Potentials and possible pitfalls. Acta Geophys. 2012;60:589-623.
- [35] Elsner JB, Jagger TH, Fogarty EA. Visibility network of United States hurricanes. Geophys Res Lett, 2009; 36.
- [36] Zou Y, Small M, Liu Z, Kurths J. Complex network approach to characterize the statistical features of the sunspot series. New J. Phys. 2014;16.
- [37] Costa LF, Rodrigues FA, Travieso G, Villas PR. Characterization of complex networks: A survey of measurements. Adv Phys. 2007;56:167-242.
- [38] Vega-Redondo F. Complex Social Networks. 1st ed. Cambridge University Press, 2007.
- [39] Luque B, Lacasa L, Ballesteros FJ, Robledo A. Analytical properties of horizontal visibility graphs in the Feigenbaum scenario. Chaos. 2012; doi: 10.1063/1.3676686
- [40] Kragujevac ML. On the Laplacian Energy of a graph. Czech Math J. 2006;56:1207–1213.
- [41] Stevanovic D, Stankovic I, Milosevic M. More on the relation between energy and Laplacian Energy of graphs. Commun. Math. Comput. Chem., 2009;61:395-401.

- [42] Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Time Series. *Chaos*. 1995 5(1):82-87
- [43] Ary L. Goldberger, Luis A. N. Amaral, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, C.-K. Peng, H. Eugene Stanley
Fractal dynamics in physiology: Alterations with disease and aging . *Proceedings of the National Academy of Sciences* Feb 2002, 99 (suppl 1) 2466-2472;
- [44] Hausdorff JM, Peng CK, Ladin Z, Wei JY, Goldberger AL. Is walking a random walk? Evidence for long-range correlations in stride interval of human gait.. *J Appl Physiol*. 1995;78(1):349-358.
- [45] Hausdorff JM, Mitchell SL, Firtion R, et al. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease . *J Appl Physiol* (1985). 1997;82(1):262-269.
- [46] Glass L., Mickey M.C. From clocks to chaos the Rythms of life . Princento University Press 1988.
- [47] Denton T.A., Diamond G. A., Helfant R.H., Khan S., Karagueuzian H. Fascinating rhythm: a prime on chaos theory and its applications to cardiology. *Am. Heart J*. 120, 1419-1440.
- [48] Lipsitz L.A., Goldberger A.L. Loss of complexity and aging. Potential applications of fractals and chaos theory to senescence. *JAMA*. 1992;267, 1806-1809
- [49] Lacasa L., Luque B., Luque J., Nuno J.C. The visibility graph: Anew method for estimating the Hurst exponent of fractionalBrownian motion. *EPL Europhysics Lett* 86: 30001
- [50] Bhaduri A. Ghosh D. Multiplicity fluctuation and phase transition in high-energy collision — A chaos-based study with complex network perspective. , *Int J Modern Phys A*, 3 (2016) 1650185.
- [51] Bhaduri A. and Ghosh D. Pion Fluctuation in High Energy Collisions - A Chaos-based Quantitative Estimation with Visibility Graph Technique. *Acta Phys Pol B*, 48 (2017) 741
- [52] Bhaduri S., Bhaduri A. and Ghosh D A new approach of chaos and complex network method to study fluctuation and phase transition in nuclear collision at high energy . *Eur Phys J A*, 53 (2017) 135.

- [53] Bhaduri S. and Ghosh D, Fractal study of pion void probability distribution in ultra-relativistic nuclear collision and its target dependence . *Mod Phys Lett A*, 31 (2016) 1650158.
- [54] Nilanjana P., Anirban B., Susmita B. and Dipak G. Non-Invasive Alarm Generation for Sudden Cardiac Arrest: A Pilot Study with Visibility Graph Technique. *Transl Biomed.* 2016, 7:3.
- [55] Bhaduri A. and Ghosh D. Quantitative Assessment of Heart Rate Dynamics during Meditation: An ECG Based Study with Multi-Fractality and Visibility Graph. *Front. Physiol.* 2016 7:44.
- [56] Peng C, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals . *Circulation [Online]*. 101 (23), pp. e215–e220..
- [57] Nasrolahzadeh M, Mohammadpoory Z, Haddadnia J. Analysis of heart rate signals during meditation using visibility graph complexity . *Cogn Neurodyn.* 2019;13(1):45-52.
- [58] Kim J, Wilhelm T. What is a complex graph? . *Phys A Stat Mech Appl.* 2008; 387:2637–2652.
- [59] Goldberger, A. et al. Exaggerated heart rate oscillations during two meditation techniques . *International Journal of Cardiology.* 1999; 70 (2),101-107
- [60] Tukey, J. W. *Exploratory data analysis.* Reading Massachusetts: Addison-Wesley 1977
- [61] Solé R., Manrubia S., Luque B, Delgado J., Bascompte J. Phase transitions and complex systems: Simple, nonlinear models capture complex systems at the edge of chaos . *Complexity* , 1996;1(4), 13-26
- [62] Bak P. *How Nature Works: The Science of Self-Organised Criticality..* Copernicus Press 1996
- [63] Christensen K., Moloney N. *Complexity and criticality.* Imperial College Press advanced physics texts Vol.1
- [64] Scheffer M., Bascompte J., Brock W., Brovkin V., Carpenter S., Dakos V., et al. Early-warning signals for critical transitions. *Nature.* 2009;461(7260), 53–59.
- [65] Scheffer M., Carpenter S., Lenton T., Bascompte J., Brock W., Dakos V., et al. Anticipating Critical Transitions. *Science.* 2012;338(6105) e344–348

- [66] Cimini, G., Squartini, T., Saracco, F. et al. The statistical physics of real-world networks. *Nat Rev Phys* 2019. 1, 58–71.
- [67] Park J., Newman E. J. Statistical mechanics of networks. *PHYSICAL REVIEW E*. 2004;70, 66171:1-13
- [68] Nave C. R., HyperPhysics: Entropy as a Measure of the Multiplicity of a System. Página de Internet: <http://hyperphysics.phy-astr.gsu.edu/hbase/Therm/entrop2.html>
- [69] Braun, S., Ronzheimer J. P., Schreiber M., Hodgman S. S., Rom T., Bloch I., Schneider U. Negative Absolute Temperature for Motional Degrees of Freedom. *Science* 2013; 339(6115), 52-55

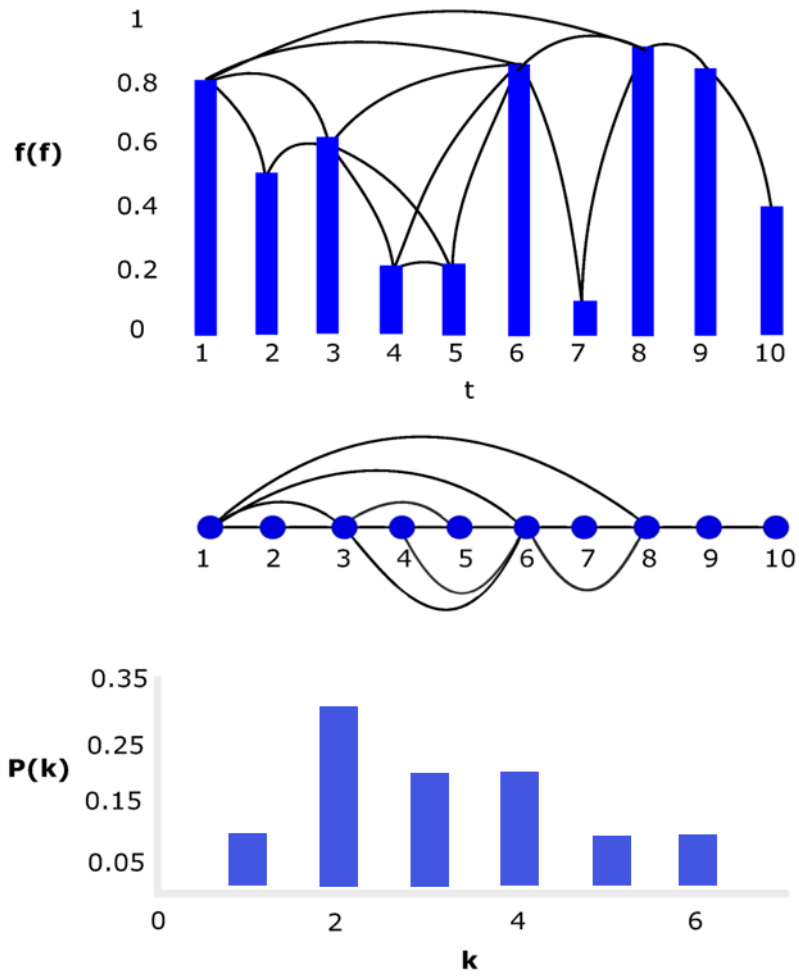


Figura 7: Representación gráfica del algoritmo de visibilidad natural. Se muestra la serie de tiempo, su red y su distribución de grado.

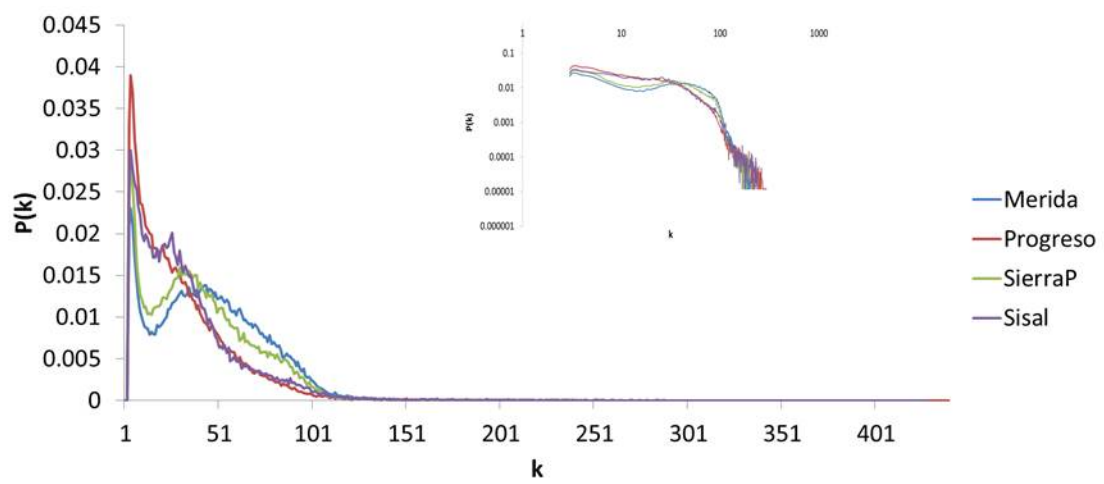


Figura 8: Distribuciones de grado promedio para cada una de las ciudades analizadas .

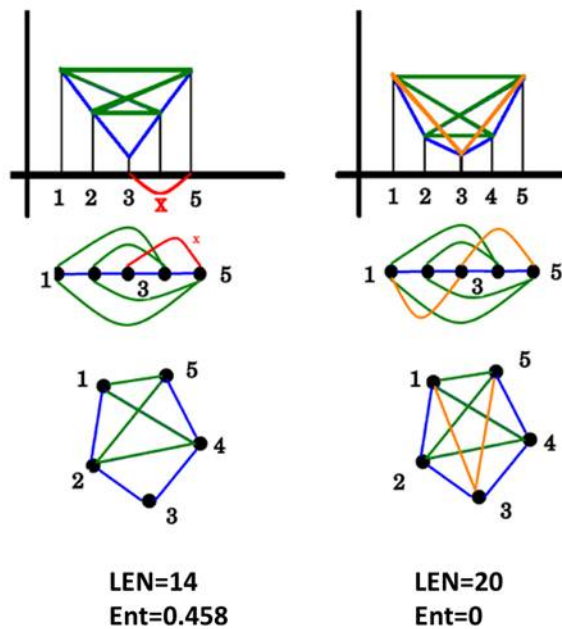


Figura 9: Ejemplo gráfico de cómo afecta la variación de la serie de tiempo en LEN. De-
 recha, una serie de tiempo donde todos los elementos se ven entre sí lo que genera una red
 completamente conectada, izquierda, una serie de tiempo donde los valores 2 y 4 aumentan,
 disminuyendo la visibilidad de los nodos produciendo una red menos conectada, y por lo tanto
 el valor de LEN disminuye.

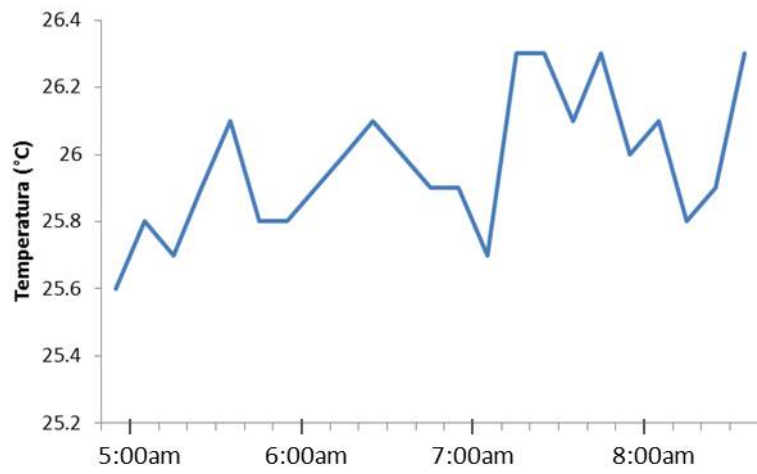


Figura 10: Detalle de la variabilidad en un segmento de la serie de tiempo de Progreso

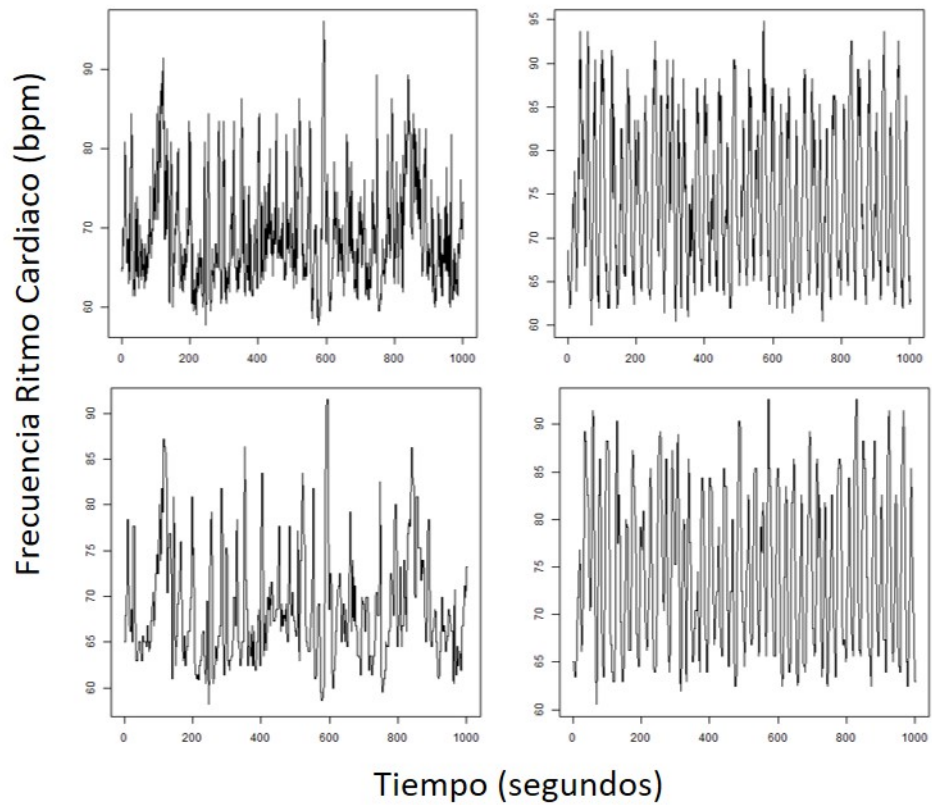


Figura 11: Series de tiempo antes (izq.) y durante (der) la sesión de meditación, para datos normales (arriba) y suavizados (abajo).

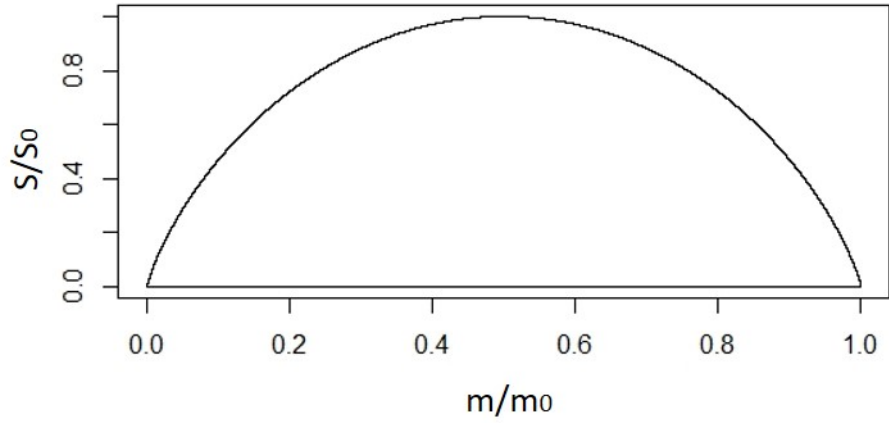


Figura 12: Gráfica donde se muestra la forma de la entropía (Ecuación 25) con respecto al número de enlaces. Tanto la entropía como el número de enlaces se normalizaron a la entropía máxima S_0 y el número de enlaces máximo m_0

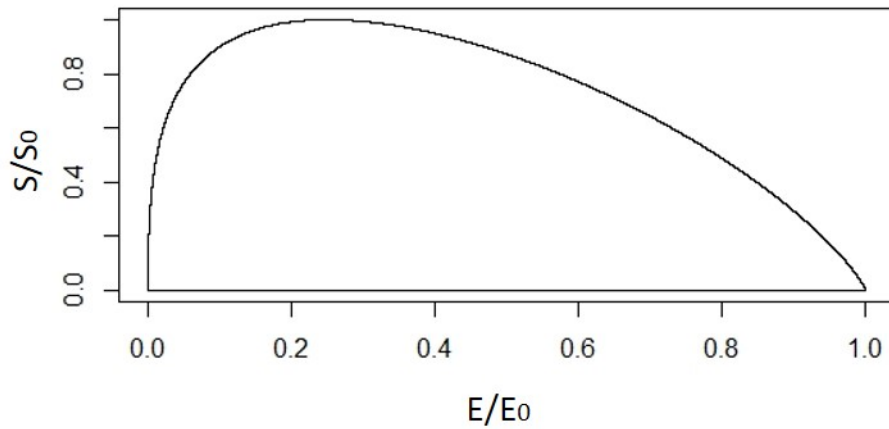


Figura 13: Gráfica de S vs E construida a partir del modelo descrito en el Capítulo 5. Se normaliza a la S máxima (S_0) y E máxima (E_0).

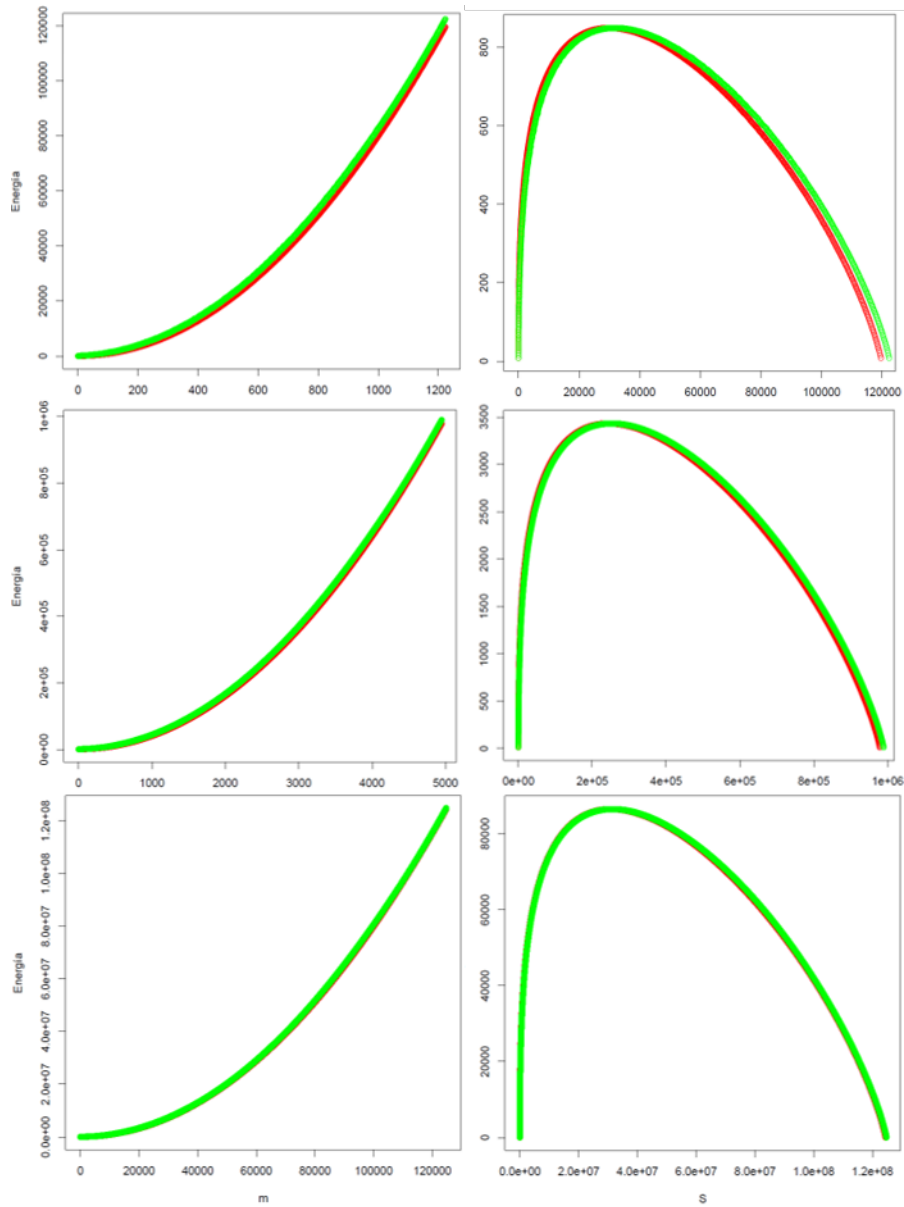


Figura 14: Comparación del modelo teórico vs el experimental para $N=50,100$ y 500 . Mientras más grande N , el error entre las curvas es menor.

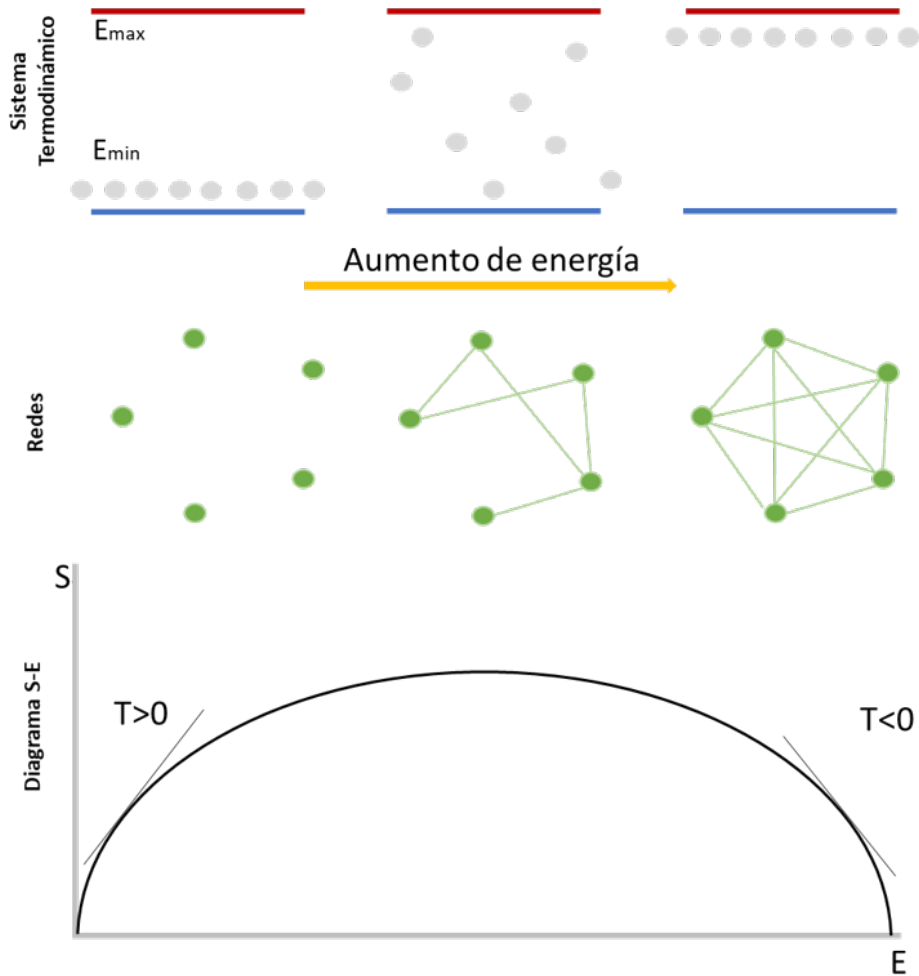


Figura 15: Explicación visual sobre el concepto de temperatura negativa en redes. En la parte superior se muestra un sistema termodinámico con una energía máxima acotada, de izquierda a derecha se muestran 3 diferentes sistemas con un aumento de energía. En la parte media se muestran 3 redes de 5 nodos ejemplificando los extremos donde no hay enlaces y donde todos los nodos tienen enlaces (Red completa). En la parte inferior se muestra el diagrama de Entropía vs Energía correspondientes a los sistemas descritos anteriormente.

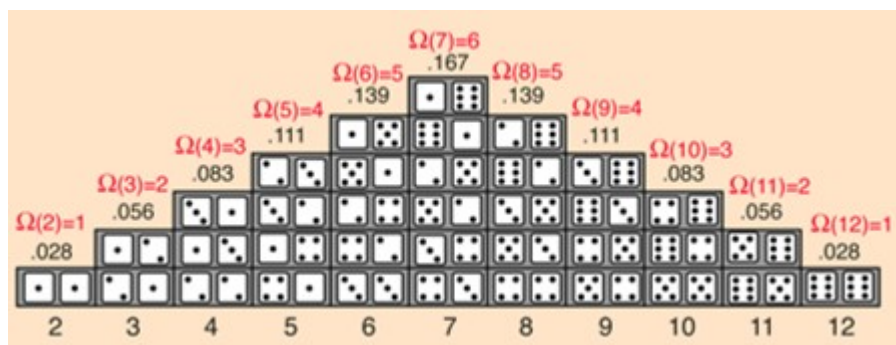


Figura 16: Representación gráfica del experimento de los dados, el cual consiste en lo siguiente: Se tienen dos dados los cuales al sumarse los puntos obtenidos por las caras nos genera una cantidad X. Nosotros queremos saber qué número X es el mejor de forma tal que sea muy probable que al lanzar los dados obtengamos ese puntaje. Se observa que si escogemos el 7, tenemos 6 formas distintas en las cuales podemos obtener dicha suma y sería más probable ganar que si escogemos el 12 o el 2, los cuales sólo hay una forma en la que nos salgan estos puntajes. Tomado de [68].

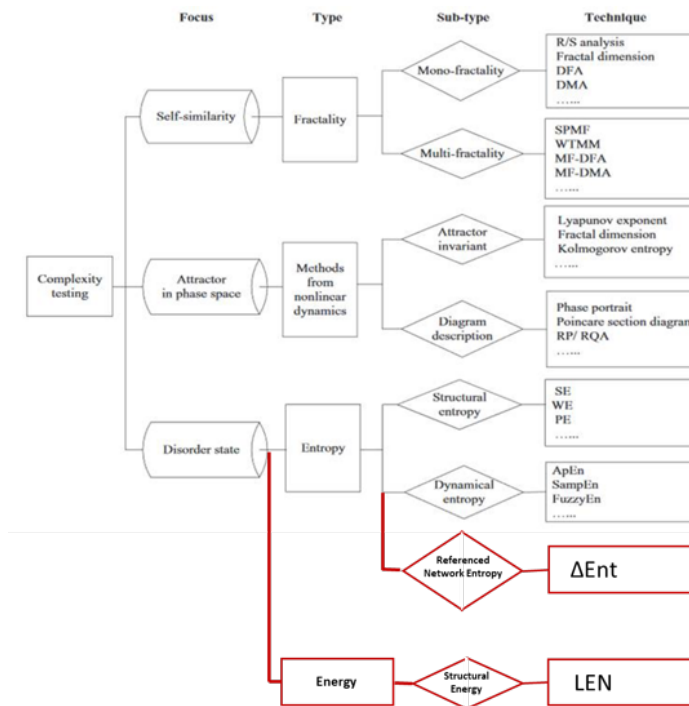


Figura 17: Clasificación propuesta por Tang et al. [9] donde se incluye la propuesta de dónde podrían colocarse ΔEnt y LEN (resaltado en color rojo).