



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

**“ANÁLISIS DE SERIES DE TIEMPO DE VARIABLES
ATMOSFÉRICAS APLICANDO LA TEORÍA DE
SISTEMAS DINÁMICOS NO LINEALES Y UN
ENFOQUE DE MINERÍA DE DATOS:
ESTADO DE TLAXCALA”**

TESIS QUE PARA OBTENER EL TÍTULO DE
INGENIERO GEOFÍSICO

PRESENTA: DANIEL PEÑA MACIEL

DIRECTOR DE TESIS: DR. TOMÁS MORALES ACOLTZI

MÉXICO D.F. ENERO DE 2012.





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

I. Agradecimientos

A mi querida Universidad Nacional Autónoma de México, mi alma máter, porque ha sido y será siempre una fuente inagotable de nuevas experiencias, emociones y conocimiento en mi vida.

A la Facultad de Ingeniería y mis profesores, que me guiaron en la adquisición de las bases de la Ingeniería y la Geofísica.

Al Comité de Titulación: Dr. Ernesto Caetano Dos Santos, M. en C. Gerardo Cifuentes Nava, M. en C. David Escobedo Zenil y el M. en C. Mauricio Nava Flores, porque con sus valiosas observaciones y sugerencias este trabajo se vio enriquecido, pero más que nada por su apoyo profesional durante la duración de mis estudios.

Al Dr. Tomás Morales Acoltzi, investigador del Centro de Ciencias de la Atmósfera, por su amistad, el apoyo brindado para concluir esta etapa y por mostrarme nuevos horizontes para explorar.

II. Dedicatorias

A mi padre, José Daniel, porque forjaste mi carácter, mis principios y me guiaste sin soltarme en los años que definen a una persona.

A mi madre, Reyna, por tu amor inconmensurable e incondicional, y tu apoyo en todo momento.

A mis hermanos, Jahzeel, Luis y Alejandro, porque crecimos juntos y hemos reído, llorado, amado y disfrutado a la par; además de que han sido un motor en mi vida.

A Vero, porque eres el *yin* de mi vida que tanto estuve esperando y el futuro que aguarda.

A la gran y numerosa familia Maciel Sánchez, porque todos y cada uno han contribuido y apoyado en mi existir.

A la familia Peña Hernández, porque mi existir, siempre me han extendido la mano.

A todos y cada uno de mis amigos: de la preparatoria, la universidad y demás lugares que no puedo enumerar por completo, que han estado conmigo en las altas y en las bajas.

A mi familia del *Tae Kwon Do* y las demás Artes Marciales, en especial, *Sabom-Nim* Oscar Hernández Zarza y *Sensei* René Galván Barrios, ya que momento en el entrenamiento ha contribuido en mi búsqueda del equilibrio mental, físico y espiritual.

Índice

Capítulo 1. Introducción y Planteamiento del Problema.....	1
1.1. Preliminares.....	1
1.2. Trabajos relacionados concernientes al Estado de Tlaxcala.....	2
1.3. Objetivos.....	3
1.4. Problemática e hipótesis	4
Capítulo 2. Marco teórico	6
2.1. Generalidades sobre sistemas	6
2.1.1. Modelo: Representación de un sistema.....	8
2.2. Análisis de Series de Tiempo	11
2.3. Generalidades sobre Geometría fractal.....	13
2.3.1. Dimensiones no fractales: la dimensión topológica y la dimensión euclidiana	15
2.3.2. Relación de la geometría fractal y el concepto de dimensión con el Análisis No Lineal de Series de Tiempo	21
2.4. Teoría de Sistemas Dinámicos No Lineales (TSDNL)	23
2.4.1. Antecedentes	23
2.4.2. Conceptos fundamentales de la TSDNL	27
2.4.2.1. Sistema dinámico.....	27
2.4.2.2. Complejidad.....	27
2.4.2.3. Linealidad y No Linealidad	28
2.4.2.4. Determinismo y aleatoriedad	30
2.4.2.5. Sistemas conservativos y sistemas no conservativos o disipativos.....	31
2.4.2.6. Caos y Teoría del Caos.....	32
2.5. Análisis lineal de señales	35
2.5.1. Espectro de potencia: caracterización preeliminar de un sistema	38
2.6. Análisis No lineal de Series de Tiempo (ANST).....	39
2.6.1. Espacio de estado y espacio de fase.....	41

2.6.2. Atractor	44
2.6.2.1. Tipos de atractores	45
2.6.3. Reconstrucción del atractor	47
2.6.3.1 Estimación de los parámetros de reconstrucción	48
2.6.3.1.1. Obtención del corrimiento óptimo.....	49
2.6.3.1.2. Obtención de la dimensión de inserción (contenedora).....	50
2.6.3.1.3. Estimación de la Dimensión de Correlación.....	51
2.6.4. Minería de datos (MD).....	53
Capítulo 3. Métodos y Herramientas.....	55
3.1. Base de Datos (BD).....	55
3.2. Zona de estudio	55
3.3. Métodos y Herramientas computacionales	56
Capítulo 4. Experimentos numéricos	59
4.1. Experimentos con ST teóricas.....	60
4.1.1. Experimentos ST Lorenz: Variable “x”.....	61
4.1.2. Experimentos ST Lorenz: Variable “y”.....	64
4.1.3. Experimentos ST Lorenz: Variable “z”.....	67
4.2. Experimentos con ST observadas	70
4.2.1. Experimentos Estación Atlanga P1: 1960-1970.....	71
4.2.2. Experimentos Estación Atlanga P1: 1960-1970; Temperatura mínima.....	74
4.2.3. Experimentos Estación Atlanga P2: 1973-1987.....	77
4.2.4. Experimentos Estación Atlanga P2: 1973-1987; Temperatura mínima.....	80
4.2.5. Experimentos Estación Tlaxco P1: 1960-1970	83
4.2.6. Experimentos Estación Tlaxco P1: 1960-1970; Temperatura mínima.....	86
4.2.7. Experimentos Estación Tlaxco P2: 1973-1987	89
4.2.8. Experimentos Estación Tlaxco P2: 1973-1987; Temperatura mínima.....	92
Capítulo 5. Análisis y Discusión de Resultados	95

5.1. Resultados.....	95
5.1.1. ST teóricas: Modelo de Lorenz.....	95
5.1.2. ST observadas diarias.....	96
5.2. Discusión de Resultados.....	97
5.2.1. ST teóricas: Modelo de Lorenz.....	97
5.2.2. ST observadas: Estación Atlanga	98
5.2.3. ST observadas: Estación Tlaxco.....	100
5.3. Interpretación global	101
Capítulo 6. Conclusiones y Trabajos a Futuro	103
6.1. Conclusiones.....	103
6.2. Aportaciones	104
6.3. Trabajos a Futuro.....	105
Bibliografía y Referencias.....	106
Anexo 1. Modelo de Lorenz	110

Capítulo 1. Introducción y Planteamiento del Problema

1.1. Preliminares

Cada vez que el hombre busca entender y describir los procesos que tienen lugar a su alrededor, debe recurrir a conceptos y herramientas que le permitan realizar un análisis de esa cierta porción de su entorno, en otras palabras, estudia un sistema. Existen varias clasificaciones referentes a los sistemas, sin embargo, ésta depende generalmente de alguna característica particular que de éstos resalte. En general, la mayoría de los sistemas reales son no lineales, sin embargo, frecuentemente se utiliza como primera aproximación a la dinámica de un sistema particular, un modelo lineal. Este tipo de modelos son preferibles desde el punto de vista científico, porque el análisis matemático que implican es menos complejo que para el caso de los sistemas no lineales; por ejemplo, no es posible obtener soluciones analíticas para las ecuaciones diferenciales no lineales usadas para describir la evolución temporal de los sistemas en cuestión. En el presente trabajo, el interés se vuelve hacia los sistemas atmosféricos, los cuales en esencia, pertenecen a dos clasificaciones: dinámicos y no lineales, el primer término se refiere a aquellos que evolucionan en el tiempo, la segunda clasificación alude a un comportamiento que no puede ser representado por un modelo lineal. Así, se pretende estudiarlos mediante métodos alternativos que se adecuan a los dos términos presentados.

En la actualidad, dada la complejidad de los sistemas naturales, ya se han aplicado ampliamente y en varias áreas del conocimiento a nivel mundial, nuevas herramientas para estudiarlos, por ejemplo, para cuantificar el impacto de la variabilidad que tiene el sistema irregular denominado **ENSO** (*El Niño-Southern Oscillation*, El Niño-La Oscilación del Sur), en la cantidad de lluvia registrada en la región N-O de los Andes argentinos (Marwan et al. 2003), para estudiar las propiedades de los electrocardiogramas (Richter & Schreiber, 1998), así como para buscar o al menos cuantificar la presencia de caos determinístico en los mercados bursátiles (Schreiber, 1999). Algunas de dichas herramientas y las cuales toman parte en el presente trabajo son: la Teoría de Sistemas Dinámicos no Lineales (**TSDNL**) conocida como la Teoría del Caos (**TC**) y la Geometría Fractal (**GF**)

1.2. Trabajos relacionados concernientes al Estado de Tlaxcala

Específicamente para el estado de Tlaxcala, se han realizado varios estudios utilizando distintas metodologías, con el fin de mejorar el entendimiento de los procesos atmosféricos y así poder actuar con base en el conocimiento obtenido; de ellos, los principales son los siguientes:

Meneses (1993) aplicó la Teoría de Sistemas Dinámicos Disipativos y la Teoría del Caos a series de tiempo de temperatura diaria de las estaciones Atlanga, Tlaxco y Apizaco para determinar: si las series eran o no caóticas, el número de variables para describir determinísticamente la evolución del clima para cada estación y el periodo válido de predicción impuesto por la propiedad caótica.

Waelbrock et al.(1994), desarrollaron un algoritmo de pronóstico de precipitación diaria, utilizando el método de reconstrucción local en un espacio de fase, con datos del observatorio de Tlaxcala Capital, con un enfoque no lineal. Waelbrock et al.(1995), realizaron otro estudio de las propiedades dinámicas de la precipitación del observatorio de Tlaxcala Capital a escala diaria, con énfasis en el límite de predictibilidad, asociado a métodos determinísticos.

Por su parte, Hernández (1996) utilizó la variable horas de calor obtenida como la diferencia de la temperatura máxima y la temperatura mínima; el objetivo principal de su trabajo fue demostrar que a partir de los registros de dichas temperaturas es posible determinar la dinámica (puntos fijos, puntos periódicos, etc.) del tiempo para las estaciones de Atlanga y Tlaxco y determinar el límite de predictibilidad de la variable denominada horas de calor.

Magaña et al.(1999), realizaron un estudio del impacto de la variabilidad climática sobre la precipitación a escala mensual con énfasis en el **EL NIÑO** y **LA NIÑA**. Gay-García et al. en 2004, realizaron un estudio de pronóstico de precipitación mensual con aplicación a actividades agrícolas, aplicando el método empírico de "Ensamble de Análogos observados".

El trabajo de Casillas-Lavín (2004) consistió en comprender y analizar algunos de los métodos que se están explotando en la TSD, para proporcionar información sobre la física del tiempo/clima, los cuales permiten extraer información cualitativa sobre los sistemas dinámicos a partir de series de datos experimentales. Utilizó en particular dos series de tiempo de precipitación mensual: una correspondiente a la población de Apizaco (Ene 1942- Dic 2002) y la segunda correspondiente a la población de Balboa, Panamá (Ene 1906-Dic 2000); además, construyó una red neuronal para reproducir la dinámica del atractor y poder efectuar predicciones.

Por su parte, Chamizo (2008) analizó la base de precipitación diaria de 12 estaciones climatológicas a diferentes escalas temporales y utilizó la plataforma del modelo numérico de mesoescala MM5 para demostrar su potencial para pronosticar eventos extremos.

1.3. Objetivos

A pesar de todos los esfuerzos mencionados, aún falta comprender muchos aspectos de los sistemas atmosféricos que siguen incluso sin poder ser descritos, por lo que deben utilizarse todas las herramientas disponibles para lograrlo. Así, en algunas áreas aún no se ha tenido una difusión más amplia de las herramientas de análisis no lineal de series de tiempo, y sobre todo, no han sido aplicadas en algún estudio práctico, es decir que **poco o nada** se ha **experimentado** con ellas usando **datos reales**.

La atmósfera siendo compleja e irregular, es supuesta impredecible. Lo interesante “es aclarar si la irregularidad de la atmósfera es completamente al azar o existe un orden atrás de su irregularidad” (Meneses, 1993); como se mencionó antes, nuevas herramientas para el análisis de sistemas ya están en uso, al respecto, por ejemplo, la **TC** permite diferenciar sistemas cuyo comportamiento presenta algún tipo de orden con estructura compleja o incluso irregular, de aquellos que en efecto, muestran cierto comportamiento irregular,

pero que es debido a alguna entrada aleatoria externa al sistema. Teniendo en cuenta lo mencionado anteriormente, pueden contarse entre los objetivos de este trabajo los siguientes:

- Realizar un acercamiento inicial a los métodos y las herramientas elementales utilizadas en el Análisis No Lineal de Series de Tiempo (**ANST**), exponiendo de forma concreta las bases teóricas en las que se fundamenta: Teoría de Sistemas Dinámicos o Teoría del Caos y Geometría Fractal; y por ende contribuir a su difusión. Todo lo cual conlleva a la caracterización de la dinámica de series de tiempo desde un enfoque no lineal.
- Aplicar un enfoque de Minería de Datos (MD) para la explicación de la dinámica de los sistemas en estudio: que los datos nos describan o proporcionen indicios del por qué son como son. De manera que se genere conocimiento nuevo y útil sobre la evolución de los sistemas atmosféricos de las regiones bajo estudio.

1.4. Problemática e hipótesis

Por lo general, la importancia de los métodos desarrollados en los campos de la ciencia resalta más, cuando éstos tienen aplicaciones prácticas en el análisis de datos provenientes de situaciones de la vida real.

Cuando se estudia algún sistema específico, es posible contar con mediciones de varias variables, de acuerdo a la teoría, **cada variable** observada debe tener la **misma dinámica** registrada por lo que, generalmente, se asume que **cualquier variable** del **espacio** de **estado** puede usarse para reconstruir el espacio de fase, en términos de las respectivas derivadas, lo que implica que las **propiedades dinámicas** de **cada variable** deben ser las **mismas**, es decir, el tiempo de corrimiento, la dimensión de correlación, etc. Considerando dicha situación, en la presente tesis se utilizan series de tiempo tanto teóricas (soluciones de las ecuaciones de Lorenz) como observadas (provenientes de estaciones climatológicas) para **comprobar** si esto se cumple, y en su caso,

dar alguna explicación pertinente al respecto. En el caso de ST observadas, éstas corresponden a las variables **precipitación (Pcp)**, **temperatura máxima (TempMax)** y **temperatura mínima (TempMin) diarias**.

De esta forma, el contenido del presente trabajo queda distribuido de la siguiente forma: en el capítulo 2 se explican cuestiones pertinentes y concretas respecto a la Teoría de Sistemas Dinámicos (**TSD**), la Geometría Fractal (**GF**) y la Teoría del Caos (**TC**), las cuales constituyen la base de los métodos no lineales de análisis de series de tiempo. El siguiente capítulo (3) expone lo referente a la base de datos observados utilizada, la zona a la que corresponden dichos registros y los métodos y las herramientas computacionales usadas para efectuar el procesamiento de los datos, así como el esquema correspondiente a dicho tratamiento de las observaciones. El capítulo 4 describe los experimentos numéricos concebidos para probar las cuestiones planteadas en el presente capítulo 1, de igual forma contiene las gráficas correspondientes. Dentro del capítulo 5 se plasmaron los resultados obtenidos mediante los experimentos numéricos efectuados y también se realiza la discusión de los mismos, con base en la teoría expuesta en el capítulo 2 y lo resuelto por los métodos aplicados. Por último, el capítulo 6 comprende las conclusiones a las que se llegaron mediante el presente análisis, así como las sugerencias y perspectivas sobre trabajos futuros relacionados directamente con los conceptos tratados en este trabajo.

Capítulo 2. Marco teórico

2.1. Generalidades sobre sistemas

Un sistema es un conjunto organizado de elementos (físicos o abstractos) interdependientes y que interactúan (se encuentran conectados) entre sí, con el fin de alcanzar un objetivo común, conformando así un todo unitario y complejo, además de que se encuentra delimitado por fronteras ya sea de tipo físico o abstractas (modificado de García-Carrillo 2003 y Madrigal-Loyola).

Los elementos básicos por los que se encuentra formado un sistema son tres:

Entradas: son los ingresos del sistema, que pueden ser recursos materiales, recursos humanos o información.

Procesos: es aquello que transforma una entrada en salida.

Salidas: son los resultados que se obtienen al procesar las entradas como consecuencia del funcionamiento o diseño del sistema.

El análisis de sistemas implica el estudio de la respuesta(o salida) de éstos a entradas conocidas, su síntesis se realiza especificando las salidas deseadas para entradas particulares y estudiando cuál es el más adecuado (identificación de sistemas).

Ahora bien, el estudio de los sistemas en las ciencias se refiere al proceso de observar y caracterizar el comportamiento de una cierta porción del universo seleccionada, de la cual quiere saberse en la mayoría de los casos su comportamiento futuro. Específicamente, en las ciencias de la Tierra, los sistemas estudiados son en realidad subsistemas del planeta mismo, así, entre éstos podemos contar los océanos, los volcanes, la atmósfera, etc. Existen varias clasificaciones referentes a los sistemas, sin embargo, ésta depende generalmente de alguna característica o variable particular que de éstos interese (Ver Tabla 1). Por ejemplo, en las ciencias de la Tierra muchos de los sistemas estudiados pueden denominarse “naturales” (característica), específicamente en el caso de que no hayan sido concebidos ni implementados por el hombre, aunque sí influenciados en cierta medida por él.

Variable o característica	Nombre que recibe el sistema	Definición
Tiempo	Continuo	Las variables tienen un valor para todos y cada uno de los instantes de tiempo (los últimos se representan con números reales).
Tiempo	Discreto	Las variables tienen un valor sólo para ciertos instantes de tiempo (los últimos se representan con números enteros).
Relación entrada-salida	Causal	La salida es debida a un estímulo de entrada.
Relación entrada-salida	No causal	La salida es independiente de la entrada.
Capacidad de almacenar energía	Con memoria	La salida se debe a entradas anteriores y a la entrada actual.
Capacidad de almacenar energía	Sin memoria	La salida se debe sólo a la entrada actual.
Inversión	Invertible	Si ocurre que para distintas entradas, se obtengan distintas salidas.
Invariabilidad	Invariante en el tiempo	Si al aplicar un estímulo se obtiene una salida determinada, y esa misma salida se obtiene al aplicar el mismo estímulo en tiempos distintos a la primera aplicación. Es decir, si los parámetros del modelo que representa al sistema no son dependientes del tiempo, el último se llama invariante en el tiempo. De lo contrario, se denomina variante en el tiempo.
Estabilidad	Estable	Si ante una entrada perfectamente acotada, la salida también es acotada. En otras palabras, si el sistema regresa a un estado lo más parecido al que poseía antes de haber interactuado con dicha entrada.
Relación causa-efecto	Lineal	Si la transformación cumple con las propiedades de ser aditiva y homogénea, es decir, con el principio de superposición. Es decir, el sistema puede analizarse dividiéndolo en subsistemas (componentes más simples). Esto significa que las

		causas y los efectos de los eventos que experimenta el sistema son proporcionales.
Relación causa-efecto	No lineal	No cumplen con el principio de superposición. Las causas y los efectos de los eventos que experimenta el sistema no son proporcionales.
Estado del sistema	Dinámico	Interesa analizar la evolución de las variables que representan al sistema conforme transcurre el tiempo.
Estado del sistema	Estático	Aquel sistema del que básicamente no interesa la evolución temporal de las variables que lo describen.

Tabla 1. Clasificación de sistemas.

2.1.1. Modelo: Representación de un sistema

La visión genérica de los modelos es que un modelo es una representación de la realidad. Mientras que la realidad puede ser arbitrariamente compleja, el modelo, por necesidad debe ser más simple. Además, el modelo se construye no directamente de la realidad, sino de las limitadas observaciones que de ella se tienen. Éste es precisamente el marco de trabajo proporcionado por el teorema de inserción de Takens: uno puede observar la realidad y construir un modelo de ella. Pero el modelo es sólo tan bueno como las observaciones (Small, 2005).

De acuerdo a Small (2005), en general, pueden describirse dos tipos de modelos:

Ontológicos: aquellos derivados de la lógica de hechos conocidos (o supuestos) sobre la realidad, están basados en reglas presuntamente verdaderas. Dichas reglas, son deducidas de las observaciones físicas de la realidad y los modelos pueden confirmarse o refutarse con observaciones adicionales. Sin embargo, los modelos no “reaccionan” directamente a las observaciones de la realidad.

Fenomenológicos: aquellos que originalmente son establecidos en el dominio de la estadística, pero que han ido encontrando una aplicación más amplia (con la llegada de las computadoras modernas), se construyen específicamente para imitar las observaciones. Se caracterizan por una falta de comprensión profunda de la realidad y de la aproximación pragmática (llamada coloquialmente fuerza bruta) de forzar que el modelo se ajuste a los datos observados. A partir de una clase general de modelos repleta de un gran conjunto de parámetros, uno escoge valores del parámetro que especifica el miembro de esa clase de modelo que mejor refleja nuestras observaciones de la realidad. Ya que el modelo es construido a partir de observaciones de la realidad, las mismas no pueden usarse para validarlo. No obstante, la consistencia del modelo elegido con nuevas observaciones puede utilizarse para validarlo. Cuando se carece de una adecuada comprensión de la realidad, es mejor construir un modelo fenomenológico. **Usualmente, en el análisis de series de tiempo, se lidia con modelos de este último tipo.**

En otras palabras, un modelo es una representación de un sistema. Así, un determinado sistema puede ser representado alternativamente de varias formas. Dicha representación dependerá de los objetivos del modelo (Heij et al., 2007) y del sistema particular que éste represente.

Muy frecuentemente, la forma más común de representar un sistema es describirlo por medio de un conjunto de ecuaciones matemáticas (constituyendo éste un modelo matemático), es decir, las leyes del sistema, cada una de las cuales expresa una relación entre las variables, o bien, la forma en la que interactúan las últimas en la evolución de dicho sistema.

Dependiendo de si el sistema está en función o no de una o más variables independientes, las ecuaciones que conforman un modelo matemático pueden pertenecer a dos diferentes clasificaciones: las ecuaciones diferenciales ordinarias y las ecuaciones diferenciales parciales. Las primeras se usan cuando se tiene una sola variable independiente, mientras que se recurre al segundo tipo cuando se ven involucradas dos o más variables independientes (Figura 1).

$$a) \frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}$$

$$b) m \frac{d^2 x}{dt^2} + c \frac{dx}{dt} + kx = 0$$

Figura 1. a) Ecuación diferencial parcial de calor en una dimensión, describe la variación de la temperatura “u” respecto al tiempo “t” y la posición “x”, en una barra calentada que se extiende a lo largo de un eje coordenado nombrado como “x”. b) Ecuación diferencial ordinaria de segundo orden homogénea que describe las vibraciones libres de una masa conectada a un resorte.

Además de las dos clasificaciones de ecuaciones diferenciales antes mencionadas, debe señalarse una distinción adicional respecto a las ecuaciones de carácter ordinario, ya que pueden resaltarse además, dos “subtipos” de especial interés: las ecuaciones diferenciales ordinarias y las ecuaciones ordinarias en diferencias. Básicamente, ambos describen la razón a la que cambia cierta variable respecto a una variable independiente. Ahora bien, a pesar de que en el presente trabajo se utilizan series de tiempo, no sobra aclarar que éste último es la única variable independiente presente, por lo que se estará tratando con ecuaciones diferenciales ordinarias y ecuaciones ordinarias en diferencias.

Respecto al párrafo anterior, debe aclararse que las primeras expresan la variación o evolución de las variables de interés considerando todos y cada uno de los instantes de tiempo transcurridos (tiempo continuo). Debido a que en la realidad no es posible tener datos de cada instante de tiempo para cada variable, se considera para cada una de éstas una serie de valores (de una cierta variable) que han sido registrados secuencialmente en el tiempo, con un cierto periodo de muestreo, es decir, a intervalos regulares o iguales de tiempo; esto se conoce como serie de tiempo discreto. Ahora bien, si dicha serie se expresa de forma que la ecuación resultante describa la evolución de la variable de interés con respecto al tiempo discreto, se tendrá una ecuación en diferencias (Figura 2).

$$a) \dot{x}(t) = \mathbf{F}(x(t))$$

$$b) x_{n+1} = \mathbf{f}(x_n)$$

Figura 2. a) Esta ecuación representa un conjunto de ecuaciones diferenciales ordinarias, implicando así tiempo continuo. b) La segunda ecuación hace referencia al tiempo discreto $t=N\Delta t$, y representa una ecuación en diferencias o “mapeo” en un espacio dimensional.

Así, la estrecha relación que guardan los conceptos de sistema y modelo puede expresarse en los términos siguientes:

Resumiendo y de acuerdo a Hernández (1996), el estudio de un sistema ocurre en 3 etapas básicas:

1) Elaboración del modelo conceptual del sistema: identificación de las variables que describen el estado del sistema en cualquier instante de tiempo.

2) Estudio del comportamiento de las variables identificadas para encontrar reglas o fórmulas que expresen la variación instantánea de las primeras (sus derivadas respecto al tiempo), así, de tener éxito en esta etapa, se obtendrá un modelo matemático que irá refinándose al seguir efectuando nuevos experimentos con él, llegando así al punto en que éste describa y simule apropiadamente al sistema bajo estudio.

3) Esta última etapa consiste en predecir la evolución del sistema, si se tiene éxito en las 3 etapas el modelo obtenido será determinista, sin embargo, no siempre es posible pasar de una etapa a otra, por lo que el modelaje **a partir de series de tiempo observadas** es una **herramienta** a la que se **recurre** con **frecuencia**.

Para el presente trabajo la “materia prima” con la que se cuenta son precisamente las series de tiempo observadas de precipitación, temperatura máxima y temperatura mínima. También, cabe mencionar que los sistemas de interés para el presente, pertenecen en esencia, a dos clasificaciones: dinámicos y no lineales, de manera muy simple, el primer término se refiere a aquellos que evolucionan en el tiempo, es decir, que aquello que cambia es el estado del sistema. La segunda clasificación alude a un comportamiento que no puede ser representado por un modelo lineal.

2.2. Análisis de Series de Tiempo

Desde cierto punto de vista, un sistema es un procesador de señales (una señal puede considerarse como la información relevante obtenida sobre un

proceso), así, las mediciones registradas, pueden considerarse como una señal de salida proporcionada por éste.

Una serie de tiempo (**ST**) es una **serie de mediciones efectuadas secuencialmente** en el **tiempo**, éstas pueden considerarse como el resultado de los distintos estados (evolución) que experimenta un cierto sistema. Generalmente, cada medición o muestra se registra usando un determinado intervalo de tiempo, generalmente regular. A dicho intervalo de tiempo usado para registrar los valores de alguna variable se le denomina intervalo de muestreo.

El análisis de **ST** pretende estudiar el comportamiento temporal de una o varias variables y sus métodos investigan secuencias de datos como función del tiempo.

Según Chatfield (Modificado de, 1995) existen varios objetivos posibles al analizar las **ST**:

a) Descripción.- Permite obtener medidas descriptivas de las propiedades principales de la series. Por ejemplo para detectar patrones, valores extremos o puntos de bifurcación (aquellos donde cambia abruptamente el comportamiento de la serie). El patrón temporal de una secuencia de eventos puede ser aleatorio, agrupado, cíclico o caótico. El análisis de series de tiempo provee varias herramientas con las cuales detectar dichos patrones.

b) Explicación.- Al registrar dos o más variables de un sistema es posible usar la variación en una serie de tiempo para explicar la variación de otra serie. Lo que puede llevar a un entendimiento mayor sobre el mecanismo que generó una determinada serie y así establecer ciertas propiedades del sistema.

c) Predicción.- El entender el proceso subyacente (mecanismo) que produjo los datos observados permite predecir futuros valores de una variable a partir de una serie temporal.

d) Control.- Cuando se genera una serie temporal que mide la "calidad" de un proceso de manufacturación, el objetivo del análisis puede ser el control del proceso.

2.3. Generalidades sobre Geometría fractal

Debido a que todo lo que nos rodea se encuentra en constante cambio, las técnicas, métodos y demás herramientas que son utilizadas para el estudio de nuestro entorno también deben por ende, ir perfeccionándose conforme transcurre el tiempo, y hasta dar paso a aquellas que mejor expliquen el funcionamiento de éste.

A este respecto, recientemente, la geometría fractal ha crecido como una vasta área del conocimiento, ocasionando que casi todas las ramas de la ciencia y la ingeniería se beneficien de las nuevas ideas que ésta aporta. La **geometría fractal** se **ocupa** de las **propiedades de los objetos fractales**, conocidos simple y usualmente como fractales. Los últimos pueden encontrarse en la naturaleza o generarse mediante una “receta matemática”.

Por otra parte, Benoit Mandelbrot es considerado como el padre de la geometría fractal, él acuñó el término “fractal”, el cual hace referencia al adjetivo latín *fractus* y al verbo latino *frangere* (romper, crear fragmentos irregulares). Mandelbrot puntualizó el hecho de que muy frecuentemente es imposible describir la naturaleza utilizando solamente geometría euclidiana, es decir, líneas rectas, círculos, cubos, etc. Él propuso que tanto los fractales como la geometría fractal podían ser usados para describir objetos reales, tales como árboles, relámpagos, los meandros de los ríos, las líneas de costa, etc.

Las nubes no son esferas, las montañas no son conos, las costas no son círculos, y las cortezas de los árboles no son lisas, ni los relámpagos viajan en una línea recta (Mandelbrot, 1982).

Existen varias definiciones de “**fractal**”, de las cuales, la más simple es: un fractal es un **objeto** que se muestra **autosimilar** bajo diferentes **cambios** de **escala**, teniendo simetría a través de ésta, con cada pequeña parte del objeto replicando la estructura del todo.

Independientemente del método de aproximación al concepto de fractal que se utilice, existe un concepto central que es de suma importancia y debe de considerarse: **dimensión**, el cual en pocas palabras, se refiere a la **capacidad de un objeto para ocupar espacio**. Los fractales poseen su propia dimensión, conocida como dimensión fractal, la cual usualmente (pero no siempre) es no entera(**fraccionaria**) y **mayor** que su **dimensión topológica**(D_T), además de

ser **menor** que su **dimensión euclidiana**(D_E). Existen varias definiciones de dimensión fractal, por ejemplo, dimensión: de similaridad, divisor, de Hausdorff, de conteo de cajas, de correlación, de información, punto a punto (puntual), puntual media, de Lyapunov, etc. De las cuales, las últimas 7 son útiles para caracterizar la estructura de atractores extraños asociados a dinámica caótica. Así, puede decirse que un **fractal** es un **objeto geométrico** que **posee autosimilitud y dimensión fraccionaria**. Ahora bien, debe mencionarse que la autosimilitud puede ser de dos tipos:

Exacta: Cada porción de un objeto tiene **exactamente** las mismas características del objeto completo. Ejemplo de un fractal con autosimilitud exacta (Figura 3).

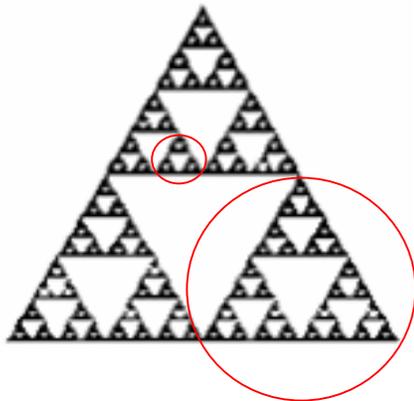


Figura 3. Fractal regular: Triángulo de Sierpinsky.

Estadística: Cada región de un objeto conserva, de manera estadísticamente similar, sus características globales.

Aquellos **fractales** que poseen **similitud exacta** se denominan **regulares**, en el entendido de que están compuestos por copias idénticas de ellos mismos rotadas y de menor escala.

Mientras que existe otro grupo de fractales que presentan el elemento aleatorio o estadístico, es decir, que cada pequeña parte del fractal posee las **mismas propiedades estadísticas** que el fractal completo. Dichos fractales son útiles en la descripción de las propiedades de diversos objetos y procesos naturales y se denominan **aleatorios**.

2.3.1. Dimensiones no fractales: la dimensión topológica y la dimensión euclidiana

Para entender el concepto de dimensión fractal, es conveniente recordar las dos definiciones de dimensión más comunes y quizá más comprensibles, ambas son dimensiones de tipo entero:

Dimensión euclidiana (D_E).- Simplemente es el número de coordenadas requeridas para especificar todos los puntos de un objeto (en el espacio).

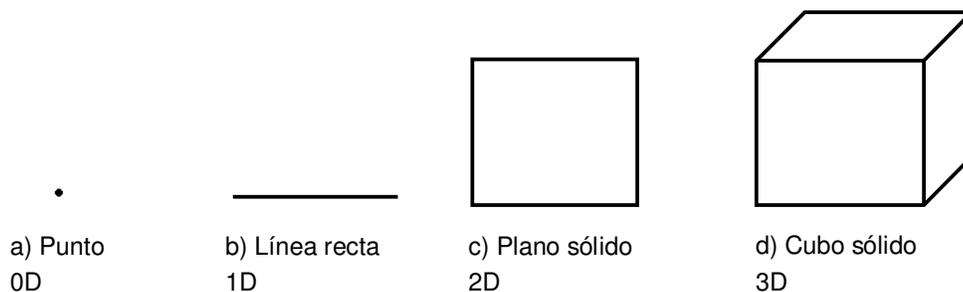


Figura 4. Dimensiones enteras comunes.

Dimensión topológica (D_T).- Ésta es más complicada que la dimensión euclidiana. La rama de las matemáticas denominada Topología considera el contorno y la forma de un objeto desde un punto de vista esencialmente cualitativo. Esta rama se ocupa de estudiar las maneras en las que objetos que poseen cierta forma, pueden ser distorsionados en otra, sin perder sus características esenciales. Así, las líneas rectas pueden ser transformadas en curvas suaves, o ser torcidas para obtener curvas quebradas (Figura 5), siendo cada una de las construcciones mencionadas topológicamente equivalentes. Es importante mencionar que dentro de las transformaciones (“distorsiones”) aplicadas a los objetos está permitido doblar, estirar, encoger, retorcer, etc., pero siempre que se haga **sin romper ni separar** lo que estaba **unido** (la transformación debe ser continua), **ni pegar** lo que estaba **separado** (la inversa también debe ser continua).

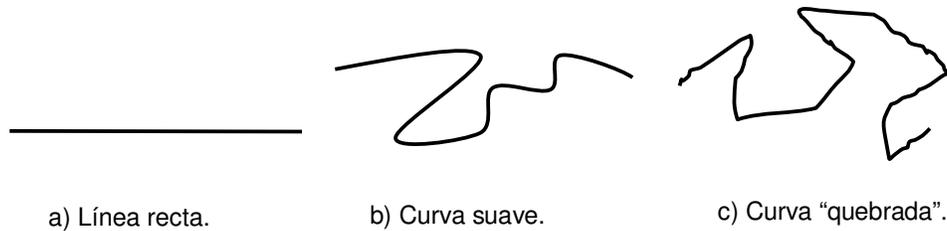


Figura 5. Curvas topológicamente equivalentes. Modificado de Addison(1997).

Ciertas características son invariantes bajo las transformaciones apropiadas (denominadas homeomorfismos por los topólogos). Por ejemplo, los hoyos en los objetos permanecen siéndolo, independientemente del estiramiento y la torsión a los que los segundos sean sometidos en su transformación de una forma a otra. Todas las superficies de dos orificios en la Figura 7, aunque bastante diferentes en forma, son topológicamente equivalentes, dado que cada una puede ser estirada y moldeada en las otras.

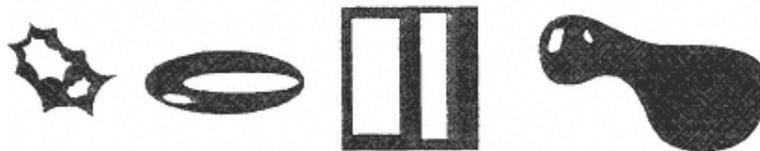


Figura 7. Superficies de dos orificios topológicamente equivalentes. Tomado de Addison(1997).

La **dimensión topológica** de un **objeto no cambia bajo la transformación de éste**. Dicha dimensión se deduce a partir de la capacidad para cubrir el objeto con discos de radio pequeño. El segmento de línea puede ser cubierto usando muchos discos que se intersecan varias veces entre ellos mismos (Figura 8a). No obstante, es posible refinar este cubrimiento usando discos que se intersecten sólo una vez entre pares de discos adyacentes. Incluso cuando la línea se contorsiona, es posible encontrar discos lo suficientemente pequeños para cubrirla de forma que las intersecciones entre éstos ocurran sólo entre pares adyacentes (Figura 8b). El segmento dentro de cada disco puede ser cubierto en sí usando discos más pequeños que requieran intersectarse sólo

en pares (Figura 8c). De forma similar, una superficie puede ser cubierta usando esferas de radio pequeño con un número mínimo de intersecciones que requiera grupos de tres esferas intersectándose (Figura 8d). La definición de dimensión topológica resulta de esta observación. El cubrimiento de un objeto por elementos (discos o esferas) de radio pequeño requiere intersecciones entre un mínimo de D_T+1 grupos de elementos.

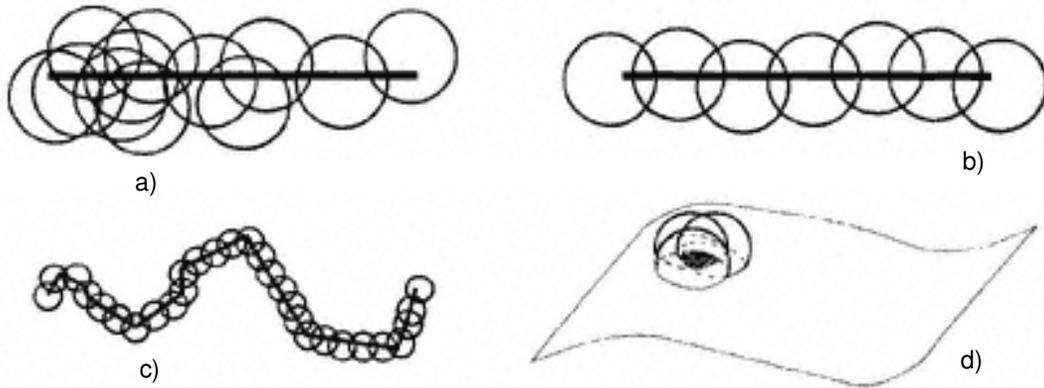
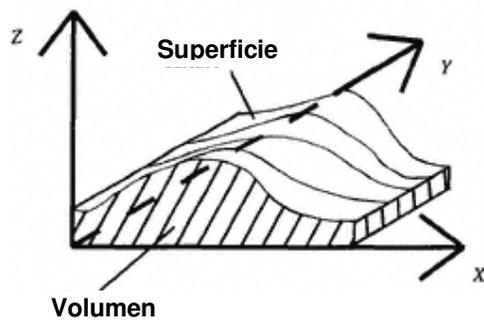
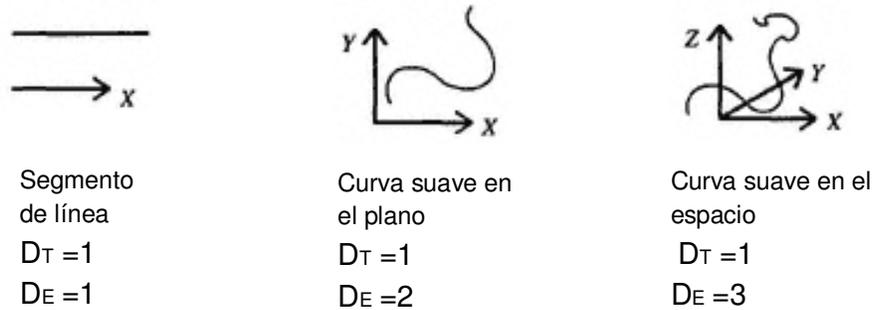
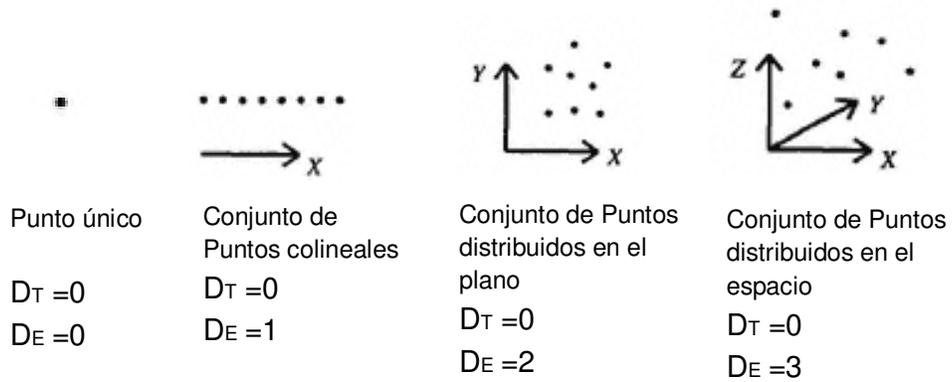


Figura 8. Cubrimiento de objetos con discos y esferas para develar su dimensión topológica. a) Segmento de línea cubierto por discos. b) Segmento de línea cubierto por discos que se intersecan sólo en pares. c) Línea quebrada cubierta por discos que intersecan sólo en pares. d) Superficie cubierta por esferas, la región de intersección se muestra sombreada.

La utilidad de tener una gran variedad de definiciones del concepto de dimensión es que puede seleccionarse algún tipo en particular para caracterizar objetos fractales específicos. Esto dependiendo si el fractal es regular o aleatorio. En el primer caso, la **dimensión de similitud** es ideal para caracterizar fractales con **autosimilitud exacta**, mientras que no es posible usar ésta cuando se desea estudiar fractales que poseen autosimilitud estadística.

La figura siguiente muestra un conjunto completo de formas comunes con sus respectivas dimensiones euclidiana y topológica.



Sección de un objeto sólido

$D_T(\text{superficie}) = 2$	$D_T(\text{volumen}) = 3$
$D_E(\text{superficie}) = 3$	$D_E(\text{volumen}) = 3$

Figura 9. Elementos geométricos en diferentes dimensiones

Un ejemplo ilustrativo de un fractal aleatorio es una línea de costa real, ya que si se seleccionan aleatoriamente segmentos de ésta, se encuentra que los últimos poseen las mismas propiedades estadísticas sobre todas las escalas de aumento, por lo tanto, las líneas de costa son curvas fractales aleatorias. Al respecto, se menciona también que una frontera fractal es una curva que no se cruza y que revela más estructura conforme se amplía la misma.

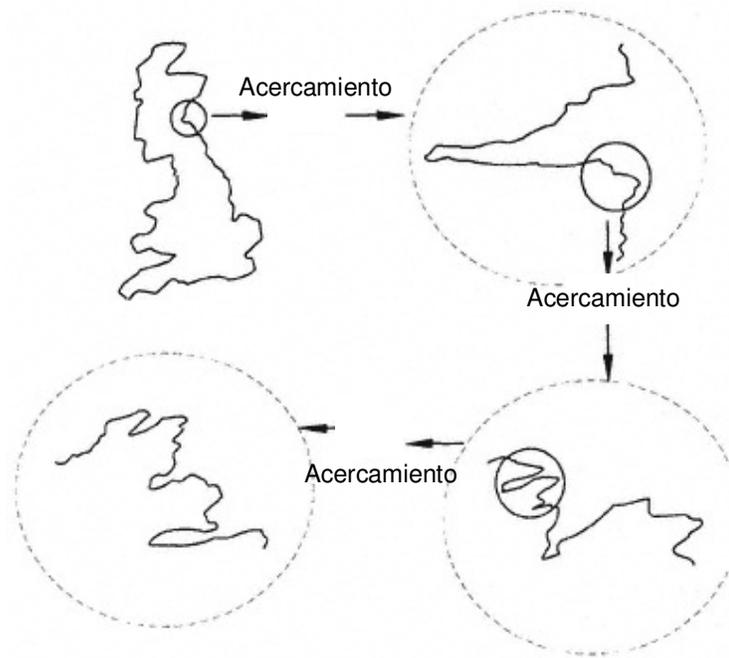


Figura 10. Acercamiento a la línea de costa para revelar más estructura. Modificado de Addison (1997).

Vale la pena aclarar que incluso los mejores ejemplos de fractales naturales, incluyendo las líneas de costa, poseen autosimilitud sobre un rango finito de escalas. No obstante, dicho rango es suficientemente grande para permitir que la geometría fractal juegue un papel importante en su caracterización.

Uno de los conceptos de dimensión más importantes para clasificar fractales es la dimensión de Hausdorff, de hecho Mandelbrot sugirió que un fractal puede definirse como un objeto que tiene una dimensión de Hausdorff que excede su dimensión topológica, sin embargo, esta aproximación al concepto de dimensión no es útil para un ingeniero o un científico que desea calcularla para un objeto fractal representado por datos reales. En sustitución puede usarse

una aproximación muy cercanamente relacionada llamada dimensión de conteo de cajas (box counting dimension), la cual básicamente es análoga a la dimensión de similitud usada para fractales regulares, porque consiste esencialmente en cubrir el objeto estudiado (por ejemplo una **ST**) con elementos o “cajas” de lado δ Figura 11. El número de cajas N requerido para cubrir el objeto está relacionado con δ mediante su dimensión de conteo de cajas (D_B). Por lo que la expresión para calcular la dimensión D_B de un objeto con un hipervolumen (es decir, longitud, área, volumen o hipervolumen fractal) V^* ¹ que no sea igual a la unidad está dado por

$$D_B = \frac{\log(N) - \log(V^*)}{\log(1/\delta)} \quad 2$$

Donde N es el número de elementos o hipercubos de lado δ requeridos para cubrir el objeto, es decir, $N = V^* / \delta^{D_B}$. Reacomodando la ecuación anterior se obtiene:

$$\log(N) = D_B \log(1/\delta) + \log(V^*)$$

La cual es una expresión que tiene la forma de la ecuación de una línea recta, en la que el gradiente (la pendiente) D_B , de la línea, corresponde a la dimensión de conteo de cajas del objeto. Dicha expresión es conveniente para calcular la D_B de una gran variedad de objetos fractales al graficar $\log(N)$ contra $\log(1/\delta)$.

No obstante, en este trabajo no se aplicará ninguno de los conceptos de dimensión antes revisados a las **ST**, en su lugar, se utilizará la dimensión de correlación, ya que es una opción alternativa de dimensión muy popular, la cual frecuentemente es más fácil de calcular que la dimensión de conteo de cajas, hablando en términos de su aplicabilidad a datos reales (Kostelich, 1997).

¹Hipervolumen de n-dimensiones: es una medida que generaliza el concepto de volumen a espacios de dimensión superior a tres. Lo mismo aplica para el concepto de hipercubo: generalización de un cubo a espacios de dimensión superior a tres.

² Es posible seleccionar cualquier base para el logaritmo en cualquiera de las ecuaciones anteriores, ya que se obtienen los mismos resultados.

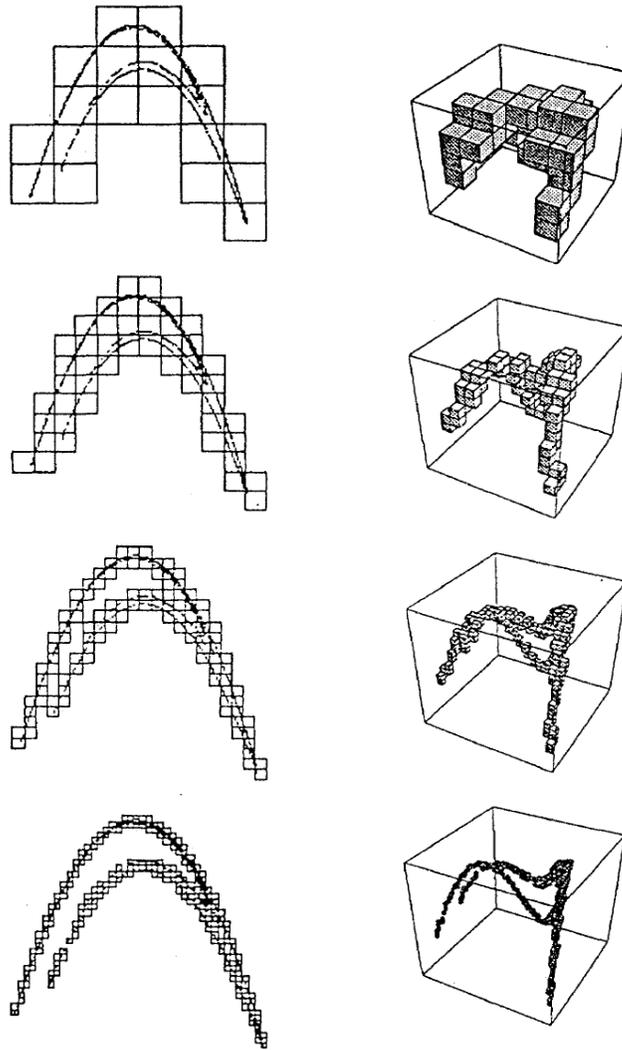


Figura 11. Dimensión por conteo de cajas para el atractor de Hénon en las dimensiones de inserción (*embedding*) $m=2,3$ y a diferentes resoluciones espaciales. Tomado de Abarbanel et al., 1993.

2.3.2. Relación de la geometría fractal y el concepto de dimensión con el Análisis No Lineal de Series de Tiempo

A primera vista, puede parecer que los fractales no tienen relación directa con una serie de datos de tiempo, y ésta a su vez con el concepto de dimensión, sin embargo ésta no es visible en primera instancia. Es cierto que los fractales no contienen pistas sobre la dinámica temporal determinista y por lo tanto no están relacionados directamente con las series de tiempo o el caos; pero existen dos conceptos que logran enlazar a los tres mencionados al inicio: la **iteración** y el **atractor extraño**.

Por un lado, y en palabras de Devaney (1992), el término iteración se refiere a repetir un proceso una y otra vez, en Dinámica no lineal, dicho proceso es la aplicación repetida de una función.

El **comportamiento no lineal** de los sistemas básicamente se debe a los **procesos de retroalimentación** que tienen lugar entre **todos** y cada uno de sus **componentes**; teniendo esto en cuenta, el proceso de **iteración** tiene lugar en el momento en el que el **resultado** de un **proceso** (una interacción entre dos partes del sistema por ejemplo) es **utilizado nuevamente** para iniciar el **mismo proceso** una y otra vez, es decir, la salida de un proceso, se utiliza como entrada para otro proceso, lo cual puede describirse mediante una ecuación iterativa, en otras palabras, considerando la iteración de una función, esto sería equivalente a efectuar la operación matemática denominada composición de una función consigo misma. La **iteración** es una **característica matemática común** de los **atractores extraños** y el **vínculo** entre la **Teoría del Caos** y la **geometría fractal**.

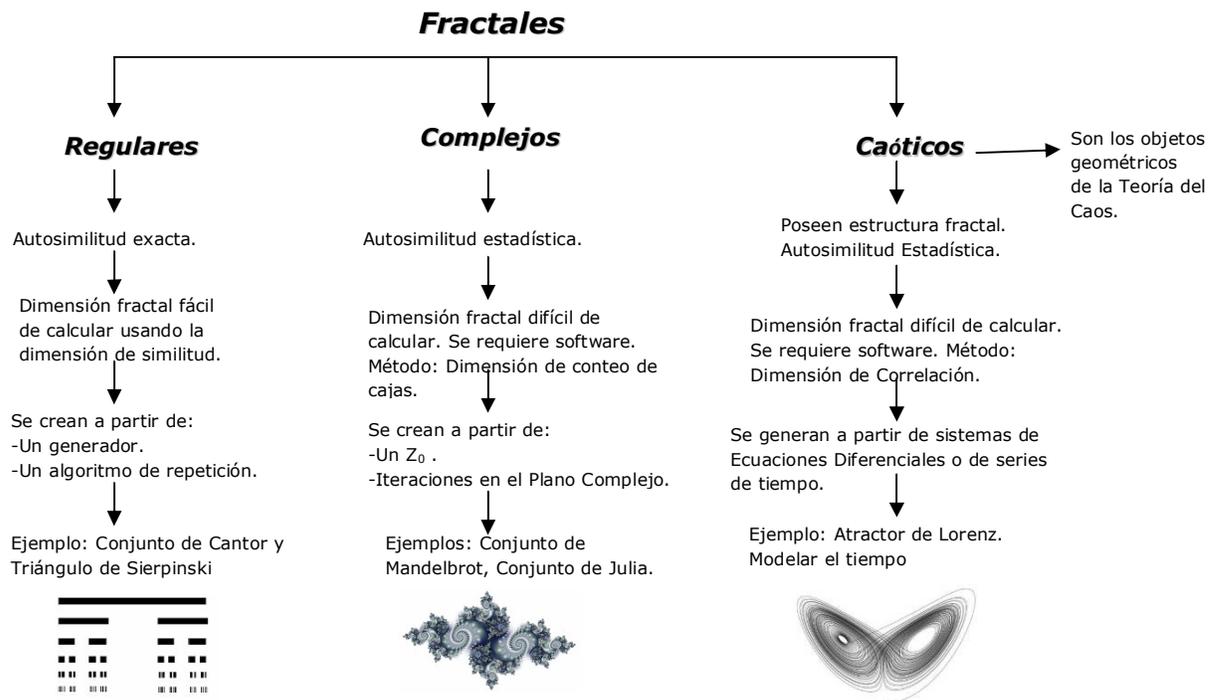
El término atractor extraño se refiere a la representación geométrica de la evolución de un sistema en el tiempo, trazada en un espacio coordinado ortogonal, además, sucede que dichos objetos presentan propiedades fractales, tales como la autosimilitud estadística, por lo que es de suponerse que también puede estimarse su dimensión, la cual será fraccionaria si representa a un sistema no lineal; la **dimensión** de un **atractor caótico proporciona una medida del número de grados de libertad generalizados** que son **relevantes** para la **dinámica del sistema** (Kostelich,1997). Más adelante en este capítulo se abordará con mayor detalle dicho término, por el momento es suficiente contar con la explicación proporcionada.

Ahora bien la **relación** que guardan los **tres conceptos** nombrados al principio es la siguiente: una **serie de tiempo** de alguna variable **describe** el comportamiento de un **sistema**, es posible recrear un **atractor** a partir de dicha serie, el cual proporcionará una descripción gráfica de la evolución del sistema, para **recrear** dicho **atractor** es necesario **estimar** la **dimensión mínima** del espacio geométrico que lo contendrá, así como la dimensión misma del objeto en cuestión. Por lo que como se observa, para analizar la serie temporal es

necesario reconstruir su atractor, pero para efectuar dicha reconstrucción se requiere estimar su dimensión.

Por último, se muestra una síntesis sobre los tipos de fractales y sus características más importantes.

Figura 12. Clasificación conceptual de estructuras en la naturaleza



2.4. Teoría de Sistemas Dinámicos No Lineales (TSDNL)

2.4.1. Antecedentes

En esencia, el término presentado como subtítulo del presente apartado, hace referencia a un dominio científico que los **matemáticos** prefieren llamar “**Teoría de Sistemas Dinámicos**”, mientras otros, incluyendo la mayoría de los **físicos**, aún siguen hablando de “**Teoría del Caos Determinístico**” o simplemente “**Caos**” (Aubin & Dahan-Dalmedico, 2002).

Podemos darnos una idea del impacto que ha tenido dicho dominio científico, al abrir nuevas herramientas para la investigación de los sistemas naturales; citando a algunos autores experimentados en dicho campo:

Según Cambel(1993), la Teoría de Caos(**TC**) sirve como el paradigma más poderoso y general para el estudio de los sistemas complejos.

En palabras de Aubin & Dahan-Dalmedico(2002), James Gleick fue más lejos al escribir en un *best seller* mundial:

La ciencia del siglo XX será recordada por tres cosas: la relatividad, la mecánica cuántica y el caos. El Caos se ha convertido en la tercera gran revolución del siglo en las ciencias físicas. Al igual que las dos primeras revoluciones, el caos ha cortado los principios de la física de Newton. Como lo expresó un físico: “La Relatividad eliminó la ilusión Newtoniana del espacio y tiempo absolutos; la teoría cuántica eliminó el sueño Newtoniano de un proceso de medición controlable; y el caos eliminó la fantasía Laplaciana de la predictibilidad determinística.

Si bien no existe un punto estándar de inicio para explicar la Teoría del Caos (**TC**), sí es posible afirmar que posee una sólida base matemática en la que se fundamenta. No obstante, es necesario aclarar que las raíces y los conceptos de los que proviene se extienden mucho más allá de sólo las Matemáticas. Para poder comprender los orígenes de esta teoría es necesario ser capaz de tratar al mismo tiempo con campos de conocimiento tan variados como mecánica de fluidos, partes de la ingeniería y dinámica de población, por mencionar algunos.

Por un lado, la **TSD** se refiere a la rama de las Matemáticas cuyo objetivo es entender los procesos en movimiento, en consecuencia, también busca predecir el comportamiento del sistema a partir de su movimiento o su cambio en el tiempo. Así, dicha rama considera a las ecuaciones diferenciales como la principal técnica para describir procesos que evolucionan continuamente en el tiempo (Devaney,1992).

Por otro lado, debido a la innovación, la variedad de herramientas, conceptos y métodos desplegados por el matemático, científico y filósofo francés Henri Poincaré durante el siglo XIX, no hay duda alguna de que su trabajo es el

punto de origen de la **TC** y también la piedra angular sobre la que fue construida.

En su obra científica, Poincaré logró articular cuatro temas de especial e importante interés:

- a) la teoría cualitativa de ecuaciones diferenciales;
- b) el estudio de la estabilidad global de los conjuntos de trayectorias (en lugar de centrarse en las soluciones individuales);
- c) la noción de bifurcación junto con el estudio de familias de sistemas dinámicos dependientes de un parámetro, y
- d) la introducción de conceptos probabilísticos en la dinámica, con respecto a la teoría ergódica y la exclusión de casos excepcionales.

En el curso del siglo siguiente (siglo XX), cada uno de estos cuatro temas se movilizó en forma conjunta o por separado. Asociados con los conceptos y métodos fundamentales, Así, los últimos pudieron establecer el esquema del dominio.

En su serie de memorias *Sur les courbes définies par une equation différentielle*, publicadas entre 1881 y 1886, Poincaré forjó los elementos de un análisis geométrico cualitativo, haciendo posible, conocer el aspecto general de las soluciones cuando las ecuaciones diferenciales no pueden resolverse, es decir, conocer sus diagramas (retratos)³ de fase y los resultados globales de su estado.

La **TC** es una teoría marcada por abruptas rupturas temporales y por transferencias de métodos y herramientas conceptuales. Por esto, puede afirmarse que su establecimiento formal emergió de forma aproximadamente independiente, de las obras individuales que sobresalieron en diversos campos científicos. La razón de esto es que los resultados obtenidos a partir de dichas obras permanecieron estrictamente confinados a los ámbitos disciplinarios y profesionales que las produjeron, algunos ejemplos son: las bases de la teoría hiperbólica (rama de las matemáticas), el modelo de ecuaciones de Lorenz de baja dimensionalidad con trayectorias complejas y la importancia que tienen

³El diagrama de fase es el conjunto de curvas solución trazadas en el espacio de fase.

los elementos de una clasificación global cualitativa de la dinámica para la modelación de los fenómenos naturales, ilustrada de forma sobresaliente por Ruelle⁴.

Por esto, y de acuerdo a su evolución histórica, es posible afirmar que la **TC** es un dominio que se constituyó en la década de 1970, sobre la base de una confrontación triple:

- 1) la teoría matemática de sistemas dinámicos;
- 2) el estudio de los fenómenos no lineales, desordenados y turbulentos en el mundo natural y/o tecnológico, y
- 3) la tecnología, es decir, las computadoras.

Por esta razón, se dice que no es una teoría dentro de la matemática pura, sino una teoría **matemática** en **constante interacción** con un gran **número** de **disciplinas**, pertenecientes a: las **ciencias físicas** (física, mecánica de fluidos), las **ciencias** de la **ingeniería** (regulación automática), e incluso a las **ciencias** de la **vida** (dinámica de la población, epidemiología).

Entre las varias aplicaciones de la **TC**, se encuentra la habilidad de proveer ideas sobre los sistemas no lineales, los cuales involucran aspectos o conceptos como aleatoriedad, determinismo, complejidad, dinámica, estados, orden, organización, condiciones iniciales, interacción, retroalimentación, equilibrio, estabilidad, divergencia, convergencia, etc.

Hasta este punto, se han usado términos como caos, sistemas dinámicos, determinismo, etc., por lo que a continuación se abordan dichos conceptos fundamentales de forma breve pero concreta, para entender a lo que se refiere dicha teoría. Cabe mencionar que una vez revisados dichos fundamentos, será posible entender la razón por la que el presente trabajo maneja el término de Teoría de Sistemas Dinámicos No Lineales (**TSDNL**).

⁴Lo que marca la pauta para afirmar que al igual que ha sucedido con otras disciplinas, para entender su origen y tener un panorama más general sobre todas las contribuciones y aquellos que las hicieron, es necesario consultar trabajos especializados al respecto, ya que dicho contexto histórico escapa al alcance del presente trabajo de tesis, se remite al lector al trabajo de Aubin & Dahan-Dalmedico (2002), el cual ofrece interesantes referencias.

2.4.2. Conceptos fundamentales de la TSDNL

2.4.2.1. Sistema dinámico

Aquel sistema cuyo comportamiento o estado **evoluciona** conforme **transcurre** el **tiempo**, en otras palabras, puede considerarse que posee un estado inicial y conforme transcurre el tiempo, alcanzará un estado final, como resultado de los cambios de estado que ha experimentado. Dichos sistemas son objeto de ser **representados** por **ecuaciones diferenciales** y en **diferencias** (siendo ambas operadores matemáticos).

La **TC** aplicada a sistemas dinámicos permite determinar el grado de: orden, desorden y predictibilidad de estos sistemas.

2.4.2.2. Complejidad

El adjetivo “complejo”, se refiere a que se compone de diversos elementos⁵. Como una regla, entre **mayor** sea el **número** de **partes** o componentes del **sistema**, es más factible que ocurra la **complejidad**; considerando que los componentes de un sistema: coexisten grandes y pequeños de forma cooperativa, están ligados y se afectan mutuamente de forma sinérgica, presentan paradojas tales como la presencia de eventos rápidos y lentos, además de posiblemente repentinos, formas regulares e irregulares, etc., todo lo que en consecuencia da lugar a una amplia gama de estados finales, partiendo desde varios o un solo estado inicial.

Cambel (1993) señaló que de acuerdo con Pagels, los sistemas complejos poseen un lugar a lo largo de un espectro que se extiende desde el orden perfecto hasta la completa aleatoriedad. En realidad, la existencia de cualquiera de dichas condiciones extremas es difícil de alcanzar, ya que los sistemas reales exhiben ambas características.

Dado que la complejidad es tan penetrante, es necesario contar con buenas bases teóricas para entenderla, de acuerdo a Ludwig von Helmholtz: **no hay nada más práctico que una buena teoría.**

⁵De acuerdo al diccionario de la Real Academia Española.

2.4.2.3. Linealidad y No Linealidad

Muchas clases de **SDs** pueden distinguirse dependiendo de las propiedades del operador que describe su evolución; si dicho **operador** obedece a la **propiedad de superposición**, es decir, es **lineal**, el **sistema** correspondiente es **lineal**, de otra forma, el sistema es no lineal. Si el operador que describe la evolución del sistema **depende implícitamente del tiempo**, el sistema correspondiente es **autónomo**, es decir, no contiene fuerzas externas aditivas o multiplicativas que dependen explícitamente del tiempo, en otro caso, se trata de un sistema **no autónomo**.

En los sistemas no lineales, la magnitud de un cierta **salida (y)**, está controlada por la magnitud de una **entrada (x)**, de acuerdo a ecuaciones simples, cuya forma genérica es **$y=mx+b$** . Es posible citar variados ejemplos de dicha afirmación, uno de ellos es la ley de Ohm: $V=IR$, donde V es el voltaje en un circuito, y variará de forma lineal con la corriente I , siempre y cuando la resistencia R se mantenga constante. Dos características centrales de los **sistemas lineales** son la **proporcionalidad** y la **superposición**.

La **proporcionalidad** significa que la **relación** que guarda la **salida** con la **entrada**(el comportamiento del sistema), puede representarse mediante una **línea recta**, en otras palabras y de manera breve, esto quiere decir que la **dinámica** intrínseca del **sistema** está gobernada por el paradigma de que **causas similares ocasionan efectos semejantes**.

Por otro lado, la **superposición** se refiere al hecho de que el **comportamiento** de los **sistemas lineales** al estar compuesto por **múltiples componentes** puede ser completamente **entendido** y **predicho** al **separar** dichas componentes y averiguar sus relaciones entrada-salida de manera individual. La **salida total** será simplemente la **suma** de dichas **partes constituyentes**. Es decir, las componentes de un sistema lineal literalmente se suman, no existiendo sorpresas o comportamientos anómalos. Dicha característica quiere decir que es posible describir y estudiar a dichos sistemas aplicando un enfoque reduccionista⁶.

⁶Reduccionismo: Partir sucesivamente al sistema en componentes (subsistemas) más pequeños para entender la estructura fundamental y su funcionamiento interno.

En contraste, incluso los **sistemas no lineales** más simples **violan** los **principios** de **proporcionalidad** y **superposición**. Así, para los **sistemas no lineales**, la **proporcionalidad** no se cumple: **pequeños cambios**(causas) pueden tener **efectos dramáticos** e **inesperados**.

Siendo además una complicación adicional el hecho de que, al estar **compuestos** por **múltiples subunidades** (o subsistemas), **no** es posible **entenderlos** al **analizar** dichas **componentes** de **forma individual**. Esta estrategia reduccionista falla debido a que los **componentes**, conforman una **red no lineal** de **interacción mutua**, en otras palabras, están **acoplados** (Goldberger et al.). La **no linealidad** de la conducta de estos sistemas está **determinada** por los **procesos** de **retroalimentación** que **aparecen** en las **interacciones** entre sus **elementos constitutivos**, donde el **resultado** de un **proceso** es **utilizado nuevamente** para **iniciar** el **mismo proceso**, marcando una **iteración**(característica matemática común de los atractores extraños y el vínculo entre la teoría del caos y la fractal) (Pichín et al., 2004).

Así, como puede apreciarse, la descripción de los sistemas no lineales se inclina hacia la aplicación de un enfoque holista⁷.

Ahora bien, hablando en términos de los modelos que representan a los sistemas físicos es prácticamente obligatorio mencionar a la principal herramienta de modelación matemática de dichos sistemas: los sistemas de ecuaciones diferenciales. Al respecto de los últimos sistemas mencionados, se puede decir que una de las situaciones de mayor apremio que conllevan a determinar qué tan complejos pueden ser, depende del sistema físico en estudio y el nivel de detalle con el lo describan, esto influye directamente en su propia resolución, ya que mientras existen soluciones explícitas generales para la mayor parte de los sistemas de ecuaciones diferenciales que describen sistemas lineales, éste no es el caso para los problemas matemáticos no lineales. Debido a esta situación, en el pasado, la tendencia era lidiar con los problemas matemáticos no lineales considerándolos como "aberraciones" e ignorarlos o bien, "linealizándolos", siempre que esto fuera posible; al

⁷Holismo: Entender el sistema en su totalidad, como un todo.

considerar que los términos no lineales y/o de orden superior son muy pequeños y así poder despreciarlos; o de igual forma, simplificar en exceso el modelo matemático para que su solución fuese fácilmente alcanzada.

De todo esto, debe resaltarse el hecho de que usualmente la complejidad implica no linealidad, desafortunadamente, no existen soluciones generales explícitas para las ecuaciones diferenciales parciales, por lo que es necesario buscar formas alternativas de lidiar con ellos, dos posibles opciones son el análisis numérico y otra, precisamente es la **TSDNL**.

2.4.2.4. Determinismo y aleatoriedad

Puede afirmarse que un **sistema determinista** es aquel cuya **evolución** está **completamente determinada** por las **condiciones iniciales**⁸.

Así, al hablar de sistemas deterministas es muy razonable o hasta obvio hacer la afirmación de que los datos que describen a dichos sistemas, pueden obtenerse a partir de fórmulas matemáticas explícitas. En otras palabras, es posible **describir** la **evolución** de un **sistema determinista mediante** una **regla** (ecuación matemática), ya que ésta indica la forma en la que el sistema cambia desde un estado a otro al transcurrir el tiempo. Así, pueden obtenerse datos que describan al sistema partir de cada estado que éste experimente.

Desde el punto de vista científico, el determinismo estricto debe descartarse, ya que las mediciones se ven afectadas por la presencia del observador. Incluso las denominadas mediciones no invasivas afectan el sistema al menos de forma microscópica, esto entre muchos otros factores que ejercen cierta influencia en las mediciones.

Por otro lado, el concepto de aleatoriedad se refiere a todo proceso cuyo resultado no es previsible, de igual forma, se usa para expresar una aparente carencia de propósito, causa, u orden.

⁸Conjunto de condiciones o características que posee un sistema en el instante específico de tiempo en el que se comienza a estudiar su evolución.

Para un **sistema** puramente **aleatorio** o **estocástico**, no debe de existir relación causal entre; una observación en un momento presente $t(n)$, y/o una observación pasada considerando un incremento de tiempo $t(n-1)$ y/o una observación futura con incremento de tiempo $t(n+1)$. Aunque un conjunto de valores parezca aleatorio, no significa que los datos que lo conforman posean esa naturaleza, deben realizarse análisis para establecer la naturaleza de los datos. Lo que se denomina aleatorio, en realidad es pseudoaleatorio, ya que a pesar de que no es posible establecer una regla con la que pueden predecirse los valores de estados futuros, sí se puede establecer con cierto grado de certidumbre un intervalo dentro del cual se encontrarán dichos valores que describen los estados futuros.

Los **sistemas complejos**(los cuales corresponden a la realidad) **no son completamente aleatorios** o **completamente deterministas**, sino que **exhiben ambas características**.

2.4.2.5. Sistemas conservativos y sistemas no conservativos o disipativos

La contracción del volumen en el espacio de fase en los sistemas mecánicos corresponde a una pérdida de energía como resultado de la disipación. Un crecimiento del volumen en el espacio de fase implica un suministro de energía para el sistema, el cual puede denominarse disipación negativa. Por lo tanto, aquellos Sistemas Dinámicos en los que la energía o el volumen en el espacio de fase varían se denominan sistemas disipativos.

A los **sistemas conservativos** también se les llama hamiltonianos, son aquellos que no experimentan pérdidas de energía. Para un sistema conservativo, el volumen y la energía en el espacio de fase se conservan conforme transcurre el tiempo.

Los **sistemas disipativos o no conservativos** son aquellos que se presentan en la mayoría de las situaciones de la vida real y que experimentan pérdidas de energía. Para su subsistencia deben recibir energía e información. Usualmente, para estos sistemas, su volumen en el espacio de fase se contrae.

Entre una gran variedad de clases de sistemas dinámicos, un lugar especial lo ocupan aquellos sistemas que manifiestan oscilaciones, es decir, procesos que

muestran repetición total o parcial. Los sistemas oscilatorios, al igual que los Sistemas Dinámicos en general, se dividen en lineales y no lineales, en agrupados y distribuidos, conservativos y disipativos, y autónomos y no autónomos. Una clase especial incluye los denominados sistemas autosostenidos. Los sistemas disipativos no lineales en los que pueden aparecer oscilaciones que no decaen y que pueden sostenerse sin la presencia de una fuerza externa se denominan autosostenidos, de igual forma las oscilaciones que tienen lugar en dichos sistemas se denominan oscilaciones autosostenidas. La energía perdida en forma de disipación en un sistema autosostenido es compensada desde una fuente externa. Una peculiaridad de las oscilaciones autosostenidas es que sus características (amplitud, frecuencia, forma de onda, etc.) no dependen de las propiedades de una fuente de poder y permanecen al menos pequeñas bajo variación de condiciones iniciales.

2.4.2.6. Caos y Teoría del Caos

A veces, los términos caos y Teoría del Caos son confundidos uno con el otro, y se usan de manera indistinta, lo que puede provocar malinterpretaciones.

Para Cambel (1993):

La **TC** es una colección de técnicas matemáticas, numéricas y geométricas que permiten lidiar con problemas de naturaleza no lineal, para los cuales no existen soluciones generales explícitas; y enfatiza que las ecuaciones caóticas no lineales no son generales, y tienen soluciones únicas para condiciones iniciales dadas. En oposición a la **TC**, el **Caos** es una **condición**. Para que éste tenga lugar, deben presentarse condiciones específicas. No todos los problemas dinámicos no lineales son caóticos, pero hasta el momento se cree que todos los problemas caóticos son no lineales. **La esencia del caos radica en la observación de que las ecuaciones deterministas que describen un sistema pueden poseer soluciones impredecibles.**

Desde un punto de vista físico, el término **Caos** hace referencia a comportamiento irregular, además e impredecible.

Las aproximaciones más destacadas en la **TC** son:

- a. La propuesta por A. M. Lyapunov, la cual pone énfasis en el entendimiento de la estabilidad de los problemas y cuya naturaleza es más cualitativa;
- b. la aproximación de Lotka-Volterra, la cual se fundamenta en la ecuación integro-diferencial no lineal;
- c. aquella que hace uso de atractores extraños y,
- d. la geometría fractal.

Dichas aproximaciones son aplicables sin la necesidad de representar al sistema mediante ecuaciones no lineales. De éstas, las que se tratan en esta tesis son las dos últimas, las dos restantes son de interés en áreas como Sistemas de control, entre muchas otras.

Al igual que para muchos otros términos en la ciencia, no existe una definición estándar de **Caos**, sin embargo, sí hay una convención sobre las características típicas que presenta el **Caos**, de acuerdo con Munakata & Ditto, (1995), éstas son:

- **No linealidad**: si es lineal, no puede ser caótico.
- **Determinismo**: Es posible describir cada estado futuro posible que el sistema puede tener, es decir, posee reglas deterministas (más que probabilísticas) que lo describen por completo de manera precisa.
- **Sensibilidad a las condiciones iniciales**: Pequeños cambios en su estado inicial pueden llevar un comportamiento radicalmente diferente en su estado final. Este “efecto mariposa” permite la posibilidad de que incluso la más ligera perturbación producida por el aleteo de una mariposa puede afectar dramáticamente si días después predominará un cielo con nubes o soleado.
- **Irregularidad** sostenida en el comportamiento del sistema: Orden oculto, incluyendo un gran o infinito número de patrones (o movimientos) periódicos inestables. Dicho orden oculto conforma la infraestructura de los sistemas caóticos irregulares, en otras palabras, “orden en desorden”.
- **Predicción** (pero no control) a **largo plazo** en general **imposible**, debido a la sensibilidad a las condiciones iniciales, las cuales pueden conocerse sólo con cierto grado finito de precisión.

De esta forma, con todos los antecedentes expuestos a lo largo de este apartado, es posible introducir el término: **sistema caótico**. Por lo que de acuerdo a lo anterior, puede decirse que un **sistema** es **caótico** si cumple con las condiciones de ser **no lineal, determinista, sensible** a las **condiciones iniciales, irregular e impredecible a largo plazo**. Por lo general, la mayoría de autores considera que un sistema es caótico si es sensible a las condiciones iniciales, ya que las otras características se encuentran un tanto “implícitas” en ésta.

Así, puede observarse que complejidad e irregularidad son casi sinónimos de no predictibilidad, lo que nos remite directamente a un sistema caótico.

La **TC** nos permite ver orden en procesos que fueron pensados completamente azarosos (Meneses, 1993). Así, es posible diferenciar sistemas cuyo comportamiento presenta algún tipo de orden con estructura compleja o incluso irregular, de aquellos que en efecto, presentan cierto comportamiento irregular, pero que es debido a alguna entrada aleatoria externa al sistema. Lo último se pensaba debido a que las ecuaciones lineales pueden llevar sólo a soluciones que oscilan periódicamente o crecen exponencialmente, así, todo el comportamiento irregular del sistema era atribuido precisamente a alguna entrada aleatoria externa a éste.

Se considera que hasta este punto se han establecido los conceptos teóricos necesarios para entender la importancia, y el por qué dicha teoría hasta el momento no ha sido acogida por completo en algunas áreas del conocimiento, precisamente la razón es que está conformada por un gran número de conceptos y herramientas con distintos orígenes. Esta situación ha ocasionado que en muchos trabajos no exista homogeneidad en los términos para referirse a la “**TC aplicada a sistemas dinámicos**”, esto se puede notar fácilmente al consultar la diversa bibliografía existente sobre el tema.

En la opinión del autor de la presente tesis al respecto de esta inconsistencia; es que debería por lo menos aclararse que el término “caótico” implica no linealidad, determinismo, no predictibilidad, irregularidad y sensibilidad a las condiciones iniciales, ya que cada uno de dichos conceptos juega un papel importante en la evolución de los sistemas naturales. Por otro lado, si el

nombre de dicho dominio científico fuera **Teoría de Sistemas Dinámicos Disipativos Deterministas No Lineales**, dicha denominación resultaría demasiado redundante, dado que siendo estrictos, el término **Sistemas Caóticos**, contendría a todos los demás, siendo así el más apropiado para denominar a este campo del conocimiento.

Cabe mencionar que en la actualidad dicho dominio es considerado como parte de las denominadas **ciencias de la complejidad**, dado que éstas engloban otros campos como; la teoría general de sistemas, la teoría de la información, etc., teniendo en común que consideran al sistema en su totalidad y no como la suma de subsistemas fundamentales. No obstante, tomando en cuenta las cuestiones planteadas en el presente trabajo, se utilizará el término **Teoría de Sistemas Dinámicos No Lineales (TSDNL)** para hacer alusión al dominio científico que incluye cada uno de los conceptos abordados en este apartado.

2.5. Análisis lineal de señales

Durante mucho tiempo, los sistemas estudiados en las diferentes áreas del conocimiento humano han sido objeto de varias simplificaciones y suposiciones, tanto físicas como matemáticas respecto a su funcionamiento, debido principalmente a que gran parte de su comportamiento y de las estructuras que los componen no han podido ser caracterizadas y representadas de manera adecuada, dada la complejidad del sistema o debido a limitaciones en equipos, costos, o simplemente porque no es posible una adecuada observación y registro de las características del sistema de interés (algunos sistemas implican riesgos para aquellos investigadores que los estudian, por ejemplo los volcanes, mediciones realizadas en el océano, etc.).

Los modelos lineales han sido muy utilizados en un sinnúmero de campos, con el objeto de obtener un primer acercamiento o aproximación para representar a diversos sistemas. Muchos de ellos representan de forma bastante cercana la manera en la que funciona o evoluciona cierto sistema, a pesar de que consideran simplificaciones e idealizaciones (suposiciones) que en la vida real son difíciles de encontrar, o que sólo bajo ciertas circunstancias muy particulares son completamente válidas, sin embargo, muchos de ellos

cumplen excepcionalmente con los propósitos para los cuales han sido concebidos.

De acuerdo a lo planteado anteriormente, un sistema puede considerarse representado mediante una señal, y ésta dividirse en componentes más simples, las cuales a su vez pueden procesarse individualmente, y los resultados pueden reunirse. Así, de acuerdo a Smith (1999), esta aproximación posee la poderosa característica de dividir un sólo problema complejo en varios problemas más simples y fáciles de resolver. Incluso la mayoría de técnicas concernientes al Procesamiento Digital de Señales están basadas en esta estrategia, llamada superposición, que puede resumirse en la frase “divide y vencerás”.

El análisis tradicional de señales está basado en lo que se conoce como análisis de Fourier, el cual es un conjunto de técnicas matemáticas, cuyo fundamento yace en efectuar la descomposición de una señal en sinusoides. El objetivo de la descomposición es manejar algo con lo que sea más fácil tratar que la señal original. Para efectuar dicha descomposición se realiza un cambio del dominio del tiempo al dominio de la frecuencia.

El dominio de la frecuencia contiene exactamente la misma información que el dominio del tiempo, sólo que en diferente forma. Si se conoce un dominio, puede calcularse el otro. Dada la señal en el dominio del tiempo, el proceso de cálculo del dominio de la frecuencia se denomina, descomposición, análisis, o simplemente Transformada de Fourier (**TF**). Si se conoce el dominio de la frecuencia, el cálculo del dominio del tiempo se denomina síntesis, o Transformada Inversa de Fourier (**TIF**).

La **TF** permite evaluar la **contribución** que tiene cada **frecuencia** en una **ST**, en otras palabras, permite buscar las frecuencias dominantes de una señal.

Como es sabido, las señales que pueden obtenerse como resultado de la evolución de un proceso real son discretas y en muchos casos aperiódicas, por lo que el análisis espectral⁹ de la serie temporal debe efectuarse aplicando la Transformada Discreta de Fourier (**TDF**), así como la correspondiente

⁹El término espectral se refiere al análisis del espectro de la señal. Un espectro es un conjunto de frecuencias.

Transformada Inversa Discreta de Fourier (**TIDF**), la cuales están dadas por las ecuaciones:

Par Transformado Discreto de Fourier en forma exponencial

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-\frac{i2\pi kn}{N}} \quad \forall k = 0, 1, 2, 3, \dots, N-1$$

Transformada Discreta de Fourier (**TDF**)

$$x(n) = \sum_{k=0}^{N-1} X(k) e^{\frac{i2\pi kn}{N}} \quad \forall n = 0, 1, 2, 3, \dots, N-1$$

Transformada Inversa Discreta de Fourier (**TIDF**)

Al graficar la **TF** de una serie de tiempo, es posible diferenciar entre señales periódicas, casi periódicas o no periódicas. Se dice que la **TF** es lineal, ya que cumple las propiedades matemáticas de homogeneidad (proporcionalidad) y aditividad (superposición). Por esta razón puede afirmarse que el análisis de una serie de tiempo efectuado mediante la **TF** es lineal, ya que es posible aplicar un enfoque reduccionista, al dividir la **TF** de la señal, manejar sus componentes de forma individual y volver a reunirlos.

A partir de dicha transformación se obtienen los denominados: espectro de amplitudes y el espectro de fases. Dado que la **TF** está conformada por una parte real y una imaginaria, dichos espectros se calculan mediante las siguientes relaciones:

Espectros de frecuencia de una señal

$$|X(k)| = \sqrt{[\text{Re}\{X(k)\}]^2 + [\text{Im}\{X(k)\}]^2} \quad \text{Espectro de amplitudes}$$

$$\phi(k) = \tan^{-1} \left[\frac{\text{Im}\{X(k)\}}{\text{Re}\{X(k)\}} \right] \quad \text{Espectro de fases}$$

2.5.1. Espectro de potencia: caracterización pre-eliminar de un sistema

A menudo, en lugar de un espectro de frecuencia, se utiliza un espectro de potencia. Dicho espectro se estima a partir del espectro de amplitudes, ya que es simplemente el cuadrado de la magnitud, es decir, del espectro de amplitudes.

$$|X(k)|^2 = \left| \sqrt{[\operatorname{Re}\{X(k)\}]^2 + [\operatorname{Im}\{X(k)\}]^2} \right|^2 \quad \text{Espectro de potencia}$$

Es recomendable que dicho gráfico sea semilogarítmico (que sólo el eje vertical sea logarítmico) para que en caso de presentarse, el espectro de potencia muestre con mayor claridad las componentes espectrales de banda ancha intercaladas con las componentes de alta frecuencia.

Sobre la relación del espectro de potencia con los sistemas dinámicos, Hernández (1996) afirma que:

Aplicando los conocimientos acerca de la **TF** es posible distinguir entre sistemas dinámicos que poseen atractores regulares y atractores extraños.

Si se desea analizar y clasificar un proceso dinámico, el espectro de potencia es particularmente apropiado para reconocer a los diferentes tipos de atractores. De acuerdo a su espectro de potencia es posible reconocer a los siguientes atractores:

Atractor puntual: Su espectro de potencia consiste en un simple impulso.

Atractor de ciclo límite: Su espectro de potencia corresponde a impulsos equidistantes.

Atractor toro: El espectro de potencia obtenido es una combinación lineal entera de dos impulsos básicos.

Por otro lado, si se desea distinguir procesos caóticos deterministas de procesos con ruido blanco, el espectro de potencia de éste último es una constante.

En el caso de un proceso caótico, su espectro de potencia consiste en una banda ancha y continua de frecuencias, posiblemente acompañada de picos característicos simples.

En los sistemas deterministas, la existencia de un espectro de banda ancha es una marca característica de la presencia de caos.

Es necesario aclarar, que el análisis del espectro de potencia tiene un alcance limitado para una serie temporal caótica por dos razones. La primera es que las series de tiempo caóticas usualmente, al poseer espectros de banda ancha, presentan la situación de que varias o muchas frecuencias pueden ser relevantes para la dinámica del sistema. La segunda es que un espectro de banda ancha, puede corresponder a una serie temporal estocástica o a ruido.

Esto es sabido porque, como apunta Kostelich (1997), es posible construir una serie temporal aleatoria que tenga el mismo espectro de potencia que una serie determinista; al generar un nuevo conjunto de coeficientes complejos que tengan las mismas magnitudes pero fases aleatorias. Así la correspondiente serie posee el mismo espectro de potencia que la serie original, pero no es determinista.

En las señales correspondientes a sistemas dinámicos reales, es más difícil diferenciar entre las componentes de banda ancha asociadas a ruido y las componentes de banda ancha asociadas a comportamiento caótico. Por lo que en general los espectros de frecuencia no son suficientes para confirmar la presencia de caos en una señal experimental, la cual inevitablemente contiene ruido. Por lo tanto, es necesario contar con otras técnicas que puedan diferenciar entre aquellas series que son aleatorias(o ruidosas) y aquellas que son caóticas; esto puede hacerse ya sea midiendo la divergencia exponencial de trayectorias cercanas sobre el atractor de la serie o investigando su estructura fractal, lo cual corresponde a métodos usados en el **Análisis No Lineal de Series de Tiempo**.

2.6. Análisis No lineal de Series de Tiempo (ANST)

A través de las secciones anteriores de este capítulo, se han proporcionado los conceptos básicos para entender lo que se abordará en el presente apartado. Así, se inicia la explicación teórica de los métodos usados en el **ANST**, resumiendo de forma puntual las características principales de ambos análisis

de series de tiempo: lineal y no lineal, esto mediante la siguiente tabla modificada, propuesta por Abarbanel et al. (1993):

Procesamiento lineal de señales		Procesamiento no lineal de señales
Reducción de ruido; detección Separación del ruido de banda ancha de la de la señal de estrecha banda usando diferentes características espectrales. Si se conoce el sistema, se diseña un filtro adecuado en el dominio de la frecuencia.	Separación de la señal	Reducción de ruido; detección Separación de la señal de banda ancha del ruido de banda ancha usando la naturaleza determinista de la señal Si se conoce el sistema o se dispone de una serie temporal, se diseña un filtro adecuado en el dominio del tiempo. Se usan la dinámica o la distribución invariante y probabilidades de transición de Markov.
Transformadas de Fourier Se usan los métodos del espacio de Fourier para convertir ecuaciones diferenciales o relaciones de recurrencia a formas algebraicas. $x(n)$ Es observada. $X(k) = \sum x(n) \exp[i2\pi k]$ Es usada.	Reconstrucción del espacio	Reconstrucción del espacio de fase Se usan las variables retrasadas en tiempo. Conforman las coordenadas del espacio de fase, en m dimensiones: $y(n) = [x(n), x(n + \tau), \dots, x(n + (m - 1)\tau)]$ τ y m se determinan mediante los métodos llamados información mutua y falsos vecinos más cercanos, respectivamente.
Picos espectrales puntiagudos. Frecuencias resonantes del sistema Cantidades independientes de las condiciones iniciales.	Clasificación de la señal	Invariantes de las órbitas(trayectorias), Exponentes de Lyapunov; Varias dimensiones fractales; invariantes topológicos; Vínculos de los números de órbitas periódicas inestables. Cantidades independientes de las condiciones iniciales.
$x(n+1) = \sum c_j x(n-j)$ Encontrar los parámetros c_j consistentes con los clasificadores invariantes(picos espectrales).	Elaboración de modelos, predicción	$y(n) \rightarrow y(n+1)$ Evolución temporal $y(n+1) = F[y(n), a_1, a_2, \dots, a_j]$ Encontrar parámetros consistentes con clasificadores invariantes(exponentes de Lyapunov, dimensiones). Encontrar dimensiones dinámicas a partir de los datos.

Tabla 2. Conexión entre los requisitos del procesamiento de señales lineal y no lineal. Los objetivos son los mismos en cada caso; las técnicas son muy diferentes.

Al respecto de la definición del **ANST**, Small (2005), plantea la siguiente:

El **Análisis No lineal de Series de Tiempo** es el estudio de datos de series de tiempo con técnicas computacionales sensibles a la no linealidad en los datos.

Y acerca del nacimiento de éste, Lauterborn et al. (1997) escribió:

En los modelos matemáticos de los sistemas dinámicos, la dinámica se visualiza en el espacio de estado, cuyas dimensiones (enteras) están dadas por el número de variables dependientes del modelo. En los experimentos, usualmente sólo es medida una variable como función del tiempo, y el espacio de estado no se conoce. Por lo que surge la pregunta: ¿cómo puede llegarse al atractor que pueda caracterizar al sistema? Este vacío entre los conceptos teóricos y los valores observados fue llenado cuando Packard *et al.* publicaron su artículo “Geometría a partir de una serie de tiempo”; y Takens proporcionó una justificación matemática de dicha aproximación. Takens probó que, a partir solamente de una serie de tiempo, es posible (re)construir un nuevo atractor, el cual es difeomorfo equivalente al atractor (generalmente desconocido) en el espacio de estado original del sistema experimental. Esto fue el nacimiento del **Análisis No lineal de Series de Tiempo**.

Para entender lo afirmado por Lauterborn en el párrafo anterior, es necesario aclarar varios conceptos.

2.6.1. Espacio de estado y espacio de fase

Al tratar de comprender el comportamiento de un sistema es necesario conocer las variables (dependientes e independientes) que determinan dicho comportamiento, además, de las relaciones existentes entre éstas, con el fin de describir cómo evoluciona el sistema de un estado a otro. Además, en el mundo real, los sistemas evolucionan en un espacio físico, en el que intercambian masa y energía con su entorno, por lo que al estudiarlos se hace necesario utilizar un marco de referencia para establecer el cambio que han sufrido a partir de su estado inicial conforme pasa el tiempo, es decir su movimiento respecto al punto considerado como el origen en el que se inició su estudio. Entonces, teniendo esto en cuenta, se debe utilizar un marco de referencia mediante el cual pueda seguirse el movimiento del sistema, lo que es posible considerando un espacio puramente matemático, lo que proporciona la pauta para introducir los siguientes conceptos.

Espacio de estado: es un **espacio** matemático (vectorial) que se obtiene al **considerar** como **coordenadas** a **todas** las **variables** que describen el **estado** instantáneo de un **sistema**, es decir su evolución; cada punto de este espacio constituye un estado del sistema y viceversa. Dichas coordenadas conforman así, un sistema coordinado ortogonal, en el que puede observarse de forma gráfica la forma en la que el sistema evoluciona conforme transcurre el tiempo. En consecuencia, la dimensión de un espacio de estado es el mínimo número de coordenadas necesarias para especificar la localización de un punto de la evolución de cualquier sistema dinámico.

Ahora bien, es posible simplificar el estudio del sistema, usando para esto, sólo **una** de las **variables** que describen su comportamiento en **función** del **tiempo**, esto se efectúa utilizando al espacio de fase, definido a continuación:

El espacio de fase es un **espacio matemático**(vectorial) que se **obtiene** al **considerar** como **coordenadas** a **una** sola de las **variables** que describen el estado instantáneo de un sistema y a **sus** respectivas **derivadas**; cada punto de este espacio constituye un estado del sistema y viceversa. Dichas coordenadas conforman así, un sistema coordinado ortogonal, en el que puede observarse de forma gráfica la forma en la que el sistema evoluciona conforme transcurre el tiempo.

La **justificación física** del **uso** del **espacio** de **fase** es que el registro de **una** sola **variable dependiente** del **tiempo**; es el **resultado** de la **interacción** o interrelación de las restantes **variables** del espacio de estado que determinan la evolución del sistema dinámico. De esta manera, no se pierde la información de la dinámica del sistema, pues está incluida en cualquier variable observada. Si recordamos, esto es una parte elemental de la premisa básica del comportamiento no lineal, como se estableció en su momento: los **componentes de un sistema no lineal** conforman una **red de interacción mutua**, en otras palabras, están **acoplados, por lo que la retroalimentación existente entre estos elementos debe de reflejarse en todas y cada una de las variables que caracterizan al sistema.**

De igual forma, existe una justificación matemática para efectuar la sustitución del espacio de estado por el espacio de fase.

Se parte de la suposición de que el espacio de estado de un sistema está representado por N diferentes variables (las cuales pueden considerarse como coordenadas), todas distintas; denotadas como x y dependientes del tiempo ($i=1,2,\dots,N$), de tal forma que el estado inicial del sistema en un tiempo t_0 , estará descrito por el vector $\mathbf{x}_0 = (x_1(t_0), x_2(t_0), x_3(t_0), \dots, x_N(t_0))$ en el espacio de estado. Posterior a esto, se describe la evolución del sistema respecto al tiempo, en otras palabras, se emplea un sistema de ecuaciones diferenciales ordinarias, el cual indica cómo varía cada proceso (cada variable) que interviene en la evolución del sistema estudiado.

Después, dichas ecuaciones pueden transformarse convenientemente a un conjunto de N ecuaciones diferenciales ordinarias de primer orden de la forma:

$$\begin{aligned}\dot{\mathbf{x}}_1 &= f_1(x_1, x_2, x_3, \dots, x_N) \\ \dot{\mathbf{x}}_2 &= f_2(x_1, x_2, x_3, \dots, x_N) \\ \dot{\mathbf{x}}_3 &= f_3(x_1, x_2, x_3, \dots, x_N) \\ &\vdots \\ \dot{\mathbf{x}}_N &= f_N(x_1, x_2, x_3, \dots, x_N)\end{aligned}$$

Donde el punto indica la primera derivada con respecto al tiempo, y las variables en negrita denotan vectores. Como se observa, cada ecuación del sistema puede obtenerse sustituyendo los valores $i=1,\dots,N$ en la ecuación que tiene la forma general:

$$\dot{\mathbf{x}}_i = f_i(x_1, x_2, x_3, \dots, x_N)$$

De esta forma, las coordenadas del espacio de estado son $(x_1, x_2, x_3, \dots, x_N)$, todas diferentes. Por lo tanto, el sistema de ecuaciones diferenciales ordinarias puede ser sustituido o reducido a una ecuación diferencial altamente no lineal de N-ésimo orden al efectuar la derivación sucesiva de una de las ecuaciones que describen al sistema. Por ejemplo, diferenciando la primera ecuación:

$$\dot{\mathbf{x}}_1^{(N)} = f_1(x_1, \dot{x}_1, \ddot{x}_1, \ddot{\ddot{x}}_1, \dots, x_1^{(N-1)})$$

Por lo que es posible reemplazar las coordenadas del espacio de estado por:

$$(x_1, \dot{x}_1, \ddot{x}_1, \dots, x_1^{(N-1)}) \quad \text{o} \quad (\dot{x}_1, \ddot{x}_1, \ddot{\ddot{x}}_1, \dots, x_1^{(N)})$$

Con lo que no se pierde generalidad en la dinámica del sistema y el espacio de fase queda descrito sólo por una sola variable y sus derivadas.

Habiendo establecido los conceptos de espacio de estado y espacio de fase, es posible hablar de la estructura que se genera en dichos espacios coordinados: el atractor.

2.6.2. Atractor

El término **atractor** hace referencia a un conjunto cerrado, conformado por trayectorias que se encuentran confinadas (acotadas) en un determinado espacio de estado o de fase, dichas trayectorias constituyen o representan el conjunto de estados del sistema, es decir su evolución temporal a partir de una condición inicial específica. De manera sencilla, este término se usa debido a que dicho conjunto literalmente atrae a las trayectorias. La dimensión del atractor, representa el **número de grados de libertad activados en la dinámica del sistema**, o equivalentemente, el **número de ecuaciones necesarias para modelar la dinámica del sistema**, representado por la **ST**. Esto es la clave central de la Teoría de Sistemas Dinámicos No Lineales (Galca, Maab and Pfister, 1998; Grassberger and Procaccia, 1983).

Los atractores tienen las siguientes propiedades (Strogatz 2000):

- El atractor es invariante, cualquier trayectoria que comience dentro del atractor, permanecerá en él en todo momento (siempre y cuando no se perturbe el sistema).
- Al atractor convergen las trayectorias desde un conjunto abierto de condiciones iniciales, al cual se le llama la “cuenca de atracción”.
- El atractor es el mínimo conjunto que cumpla las dos condiciones anteriores, es decir, no contiene subconjuntos propios con sus mismas propiedades.

2.6.2.1. Tipos de atractores

Atractor puntual: describe a sistemas dinámicos que evolucionan hacia un solo estado final, es decir, hacia un solo punto; no importando desde qué estado inicial partan, eventualmente alcanzarán dicho punto. Este atractor representa que en el espacio físico, dichos sistemas alcanzan un estado de reposo, un ejemplo físico es una cuenca o sumidero, hacia el que se dirige un flujo.

Atractor de ciclo límite (llamado también péndulo, de equilibrio o periódico): describe a un sistema que evoluciona oscilando de forma regular entre dos puntos o lugares. Las representaciones típicas son las representaciones gráficas de un péndulo y un sistema masa-resorte-amortiguador.

Atractor toro: describe a un sistema que evoluciona de una forma compleja y predecible. Los estados presentados por el sistema se repiten de forma exacta o aproximada en el transcurso del tiempo. Un ejemplo de lo primero es un laberinto, en el que sólo hay un camino que lleva a la salida y que eventualmente conduce de vuelta al inicio para volver a empezar; una vez que la solución está definida a lo largo del laberinto, cada nueva vez se repiten las rutas previas para completar el recorrido. La otra variante, llamada “autosimilar”, es parecida a una pieza larga de alambre que se envuelve alrededor de una superficie con forma de “dona”; conforme el cable se extiende a lo largo de dicha superficie, describe una trayectoria característica; ahora bien, cada trayectoria sucesiva sólo varía en que es un poco mayor que el ancho del cable alrededor de la dona. Para la mayoría de propósitos, puede considerarse que este tipo de atractor se repite exactamente, ya que las pequeñas diferencias tienen consecuencias prácticas menores en el sistema.

Atractor extraño: Describe un sistema que evoluciona de formas complejas pero impredecibles, pero que al mismo tiempo se organiza en un orden emergente. Lo hace estableciendo un patrón de funcionamiento que está acotado, autosimilar, pero nunca repitiéndose exactamente. Un atractor extraño es aquel cuya dimensión es un valor no entero (dimensión fractal). Para un sistema caótico, el correspondiente atractor puede ser fractal.

Un atractor fractal (extraño) no representa una condición suficiente ni necesaria que indique la presencia de caos. A pesar de esto, la inserción de series de tiempo caóticas a menudo producirá atractores con una dimensión de correlación fraccionaria, por el contrario, sistemas con atractores fractales frecuentemente son caóticos. Así, la dimensión de correlación fraccionaria es un indicador útil pero de ninguna forma implica la obligatoria presencia de caos.

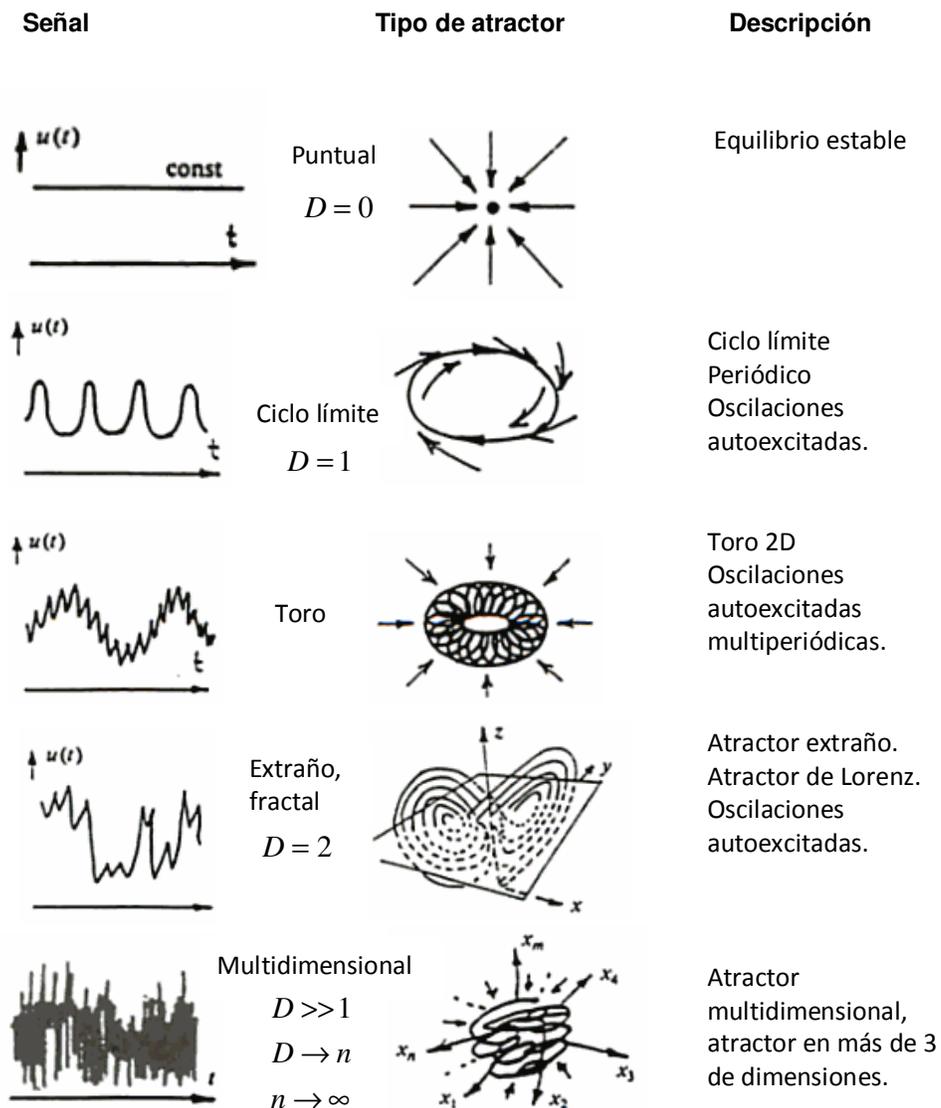


Figura 13. Donde, D denota dimensión. Modificado de Abarbanel et al. (1993).

2.6.3. Reconstrucción del atractor

La base natural para formular los algoritmos de análisis no lineal de series de tiempo a partir de la teoría del caos es un espacio de fase multidimensional, más que el dominio del tiempo o el dominio de la frecuencia (Hegger R. & Kantz H., 1999).

En la sección anterior se habló de que el espacio de estado puede sustituirse por el de fase, lo cual implica que se conozca el modelo del sistema estudiado, sin embargo, por lo general, los elementos obtenidos a partir de los procesos de registro de los sistemas naturales, son las **ST** de variables observadas, por lo que si se cuenta con el registro de una sola variable observada, no se conocen ni el espacio de estado(o de fase) ni el atractor que representa a las variables o a sus derivadas. No obstante, una serie de tiempo unidimensional proveniente de una variable observada, puede expandirse a un espacio de dimensión superior, en el que tiene lugar la dinámica del proceso generador que subyace a ésta.

Esto se realiza usando por lo general, una técnica denominada inserción de retraso coordinado (*delayed coordinate embedding*), la cual recrea un retrato del espacio de fase del sistema dinámico estudiado a partir de una serie escalar temporal. Para expandir una señal unidimensional a un espacio de fase m-dimensional, se sustituye cada observación de la señal original $\{x_i\}_{i=1}^N$, con el vector

$$y_i = \{x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}\}$$

Donde: $i = 1, 2, 3, \dots, N + (m-1)\tau$

$i \rightarrow$ Índice de tiempo (o posición de la muestra).

$m \rightarrow$ Dimensión de inserción o de *embedding*. Cantidad de vectores de inserción ejes coordenados del espacio de inserción.

$\tau \rightarrow$ Tiempo de corrimiento (*lag* o *delay*) o retraso.

Como resultado se obtiene una serie de vectores:

$$\mathbf{Y} = \{y_1, y_2, y_3, \dots, y_{N+(m-1)\tau}\}$$

Donde:

$N \rightarrow$ Es la extensión de la serie original.

El número de componentes de cada vector de inserción es $N + (m - 1)\tau$.

La idea de dicha reconstrucción es capturar los estados originales del sistema en cada instante de tiempo en que se tiene una observación de la salida del sistema.

Cada estado desconocido s_t en el instante(o tiempo) t es aproximado por un vector de coordenadas retrasadas:

$$Y_t = \{x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(m-1)\tau}\}$$

En otras palabras, esto es: supóngase que un cierto sistema se encuentra evolucionando en el espacio físico, de igual manera, las correspondientes variables que lo describen, evolucionan en un cierto espacio de estado desconocido, si se tuviera conocimiento de alguna regla que relacionara a las variables, podría transformarse el sistema a una ecuación diferencial de mayor orden y así obtener el espacio de fase para representar dicho sistema, sin embargo, en la práctica sólo se cuenta con alguna(s) variable(s) observada(s), en este punto, teniendo los registros de esa sola variable se reconstruye el espacio de fase al introducir un retraso o corrimiento, lo cual equivale a la operación “derivación”, sólo que para el caso discreto.

2.6.3.1 Estimación de los parámetros de reconstrucción

La reconstrucción usando coordenadas retrasadas se basa en la estimación de dos parámetros: la dimensión de inserción m y el tiempo de corrimiento τ .

Por lo general, la estimación de uno, requiere el conocimiento, por lo que su estimación puede parecer redundante.

El teorema de Takens relaciona la dimensión D_2 del atractor con la dimensión del espacio que contiene a dicha estructura geométrica, así, establece que:

$$m = 2D_2 + 1$$

Es decir, que la dimensión del espacio contenedor (de inserción) del atractor debe ser al menos dos veces más uno, la dimensión de éste último.

2.6.3.1.1. Obtención del corrimiento óptimo

Básicamente existen dos criterios más importantes para estimar el corrimiento óptimo:

La **función de autocorrelación** estima la relación lineal entre las observaciones de una ST separadas por un corrimiento (retraso) de τ unidades de tiempo.

$$C(\tau) = \frac{\sum_{i=1}^{N-\tau} (x'_i)(x'_{i+\tau})}{\sum_{i=1}^{N-\tau} (x'_i)^2}$$

Donde $x'_i = x_i - \bar{x}_i$

Es decir, la componente fluctuante de la ST, es separada de la componente temporal media. El corrimiento elegido corresponde generalmente al valor del primer mínimo de la función de autocorrelación.

Por otro lado, el otro criterio para estimar el corrimiento óptimo es utilizar la **función de información mutua media**, la cual mide la dependencia general de dos variables. Ésta se define como

$$M(\tau) = \sum_{i=1}^{N+(m-1)\tau} P(x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}) \left[\log \frac{P(x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau})}{P(x_i)P(x_{i+\tau})P(x_{i+2\tau}) \dots P(x_{i+(m-1)\tau})} \right]$$

Donde $P(x_i)$ es la probabilidad de ocurrencia de la variable x_i , correspondiente a la serie de tiempo, y $P(x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau})$ es la probabilidad adjunta de ocurrencia de la coordenada $| Y_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau})$ del atractor. M es una medida de la dependencia estadística de las variables reconstruidas. Si las coordenadas son estadísticamente independientes, entonces $M = 0$, lo cual correspondería a un proceso puramente aleatorio, como el ruido blanco. En contraste, una completa dependencia resultaría en $M = \infty$. Una elección adecuada del corrimiento implica que la información mutua sea mínima, para que el atractor se extienda lo más posible, por lo que usualmente se usa el corrimiento correspondiente al primer mínimo de la función de información mutua.

2.6.3.1.2. Obtención de la dimensión de inserción (contenedora)

Al igual que la mayoría de los métodos de **ANST**, la obtención de este parámetro se basa en la búsqueda “vecinos cercanos”, es decir, que se considera cada punto (estado) de la **ST** original como referencia, y se estima la cantidad de puntos a su alrededor que son “vecinos cercanos”, considerando un cierto radio de vecindad, es decir aquellos puntos que se encuentran dentro de dicho radio. Posteriormente se aumenta la dimensión (la cantidad de ejes coordenados) en uno y se vuelve a efectuar el mismo procedimiento para cada estado de la **ST**. Mientras que en una dimensión cada punto de referencia puede poseer una cantidad de vecinos cercanos, al aumentar la dimensión, dicha proporción disminuye o aumenta, esto debido a la evolución temporal del sistema, ahora bien, dada la dinámica de éste, existen vecinos que ya no aparecen de una dimensión a otra, éstos se denominan falsos vecinos más cercanos, cuando se alcanza una dimensión contenedora adecuada, la fracción de falsos vecinos más cercanos a hasta cero. Así, inspeccionando todos los datos de la serie y verificando con cuál dimensión se eliminan todos los falsos vecinos más cercanos, se removerán secuencialmente los traslapes de las orbitas de dimensión más baja. Cuando no se detecten más falsos vecinos cercanos, se habrá identificado el valor en el cual el atractor está completamente desplegado.

Así, en cada ejecución para cada punto en la **ST**, se busca su vecino más cercano en un espacio m -dimensional. Se calcula la distancia,

$$R_i = \frac{|x_{i+1} - x_{j+1}|}{\|\bar{x}_i - \bar{x}_j\|}$$

Si R_i excede un valor umbral R_t , se marca que dicho punto tiene un falso vecino más cercano. El criterio de que la dimensión de inserción es suficiente para contener al atractor es que la fracción de puntos para los que $R_i > R_t$ es cero o al menos suficientemente pequeña. De acuerdo a Small (2005), típicamente $10 \leq R_t \leq 30$, por conveniencia, $R_t = 15$ puede ser un buen valor inicial para el umbral de distancia.

2.6.3.1.3. Estimación de la Dimensión de Correlación

La **dimensión** del **atractor** es una medida de sus propiedades geométricas de escalamiento (su “fractalidad”) y ha sido llamada “la **propiedad** más **básica** de un **atractor**”.

El cálculo de la dimensión de conteo de cajas en la práctica demanda recursos computacionales excesivos, siendo ésta una de las razones por las que para calcular la dimensión de un atractor, se usa preferiblemente la dimensión de correlación (D_c), cuyo algoritmo computacional es eficiente y relativamente rápido. Para definirla, primero es necesario considerar el concepto de suma de correlación C_r :

$$C_r = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1; j \neq i}^N \theta(r - |X_i - X_j|)$$
$$\theta(X) = \begin{cases} \theta(X) = 0; & x \leq 0 \\ \theta(X) = 1; & x > 0 \end{cases}$$

Donde θ es la función de Heaviside, r es el radio de una hiperesfera¹⁰ centrada en cada punto muestreado sobre la trayectoria del atractor, X_i , $i=1,2,3,\dots,N$. Aquí, X_i , denota el vector multidimensional que corresponde a la i -ésima coordenada del espacio de fase del atractor, por ejemplo, para un atractor en un espacio de fase 3D con coordenadas x, y, z , $X_i = (x_i, y_i, z_i)$.

Es recomendable hacer un poco de énfasis en la interpretación física de la ecuación anterior. De acuerdo al esquema presentado en la Figura 14, se tiene una hiperesfera centrada en uno de los puntos que definen la trayectoria X_i del atractor. De hecho, el papel que cumple la función de Heaviside es la de contabilizar el número de otros puntos sobre el atractor que se encuentran dentro de la esfera de radio r . Dicha función es igual a uno si el valor dentro de los paréntesis es positivo, o cero si dicho valor es negativo. Los puntos X_i

¹⁰ Hiperesfera o n-esfera: generalización de una esfera a más de 3 dimensiones.

corresponden a aquellos puntos sobre la trayectoria de referencia, mientras que los puntos X_j son otros puntos sobre el atractor en la vecindad de X_i .

$|X_i - X_j|$ es la distancia de separación entre los puntos. La función $\theta(r - |X_i - X_j|)$ proporciona un valor de uno si la distancia entre los dos puntos es menor que el radio de la hiperesfera, en otro caso, proporciona un valor de cero.

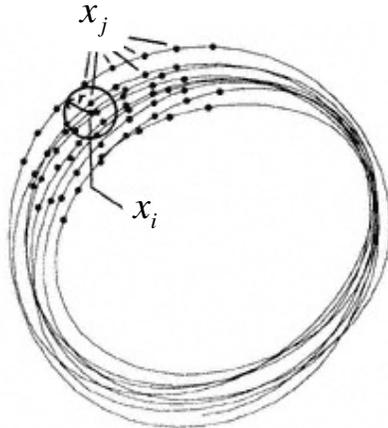


Figura 14. Determinación de la dimensión del atractor. Esfera de prueba sobre el atractor.

El cálculo de la suma de correlación (C_r) involucra darle seguimiento a la trayectoria de referencia, detenerse en cada punto discreto sobre ésta y contar el número de puntos sobre el atractor localizados dentro de una hiperesfera de radio r .

Así, la suma acumulada de todos los puntos contabilizados se divide por $N(N-1)$ para obtener la suma de correlación (C_r). El máximo valor de C_r es uno, y es alcanzado cuando el radio de la esfera de prueba es mayor que el diámetro mayor del atractor y todos los puntos se han contabilizado. El mínimo valor de C_r es $2/(N(N-1))$, y se presenta cuando sólo los dos puntos más cercanos sobre el atractor son contabilizados. Cabe advertir que la suma de correlación contabiliza el vecino más cercano dos veces (proporcionando los dos en el numerador) conforme la esfera de prueba visita ambos puntos en su trayecto alrededor del atractor.

La suma de correlación se escala con el radio de la hiperesfera de acuerdo a una ley de potencia de la forma

$$C_r \propto r^{D_c}$$

Donde el exponente D_c es la dimensión de correlación. Por lo tanto, al examinar el atractor mediante el método expuesto en el párrafo anterior para

una variedad de valores de r (diferentes hipersferas), D_c se obtiene a partir de la pendiente correspondiente a la región de escalamiento de una gráfica $\log(r)$ - $(\log C_r)$, como se muestra en la Figura 15.

La región de escalamiento es la principal región lineal de la curva de correlación, limitada por dos regiones que asemejan curvas no lineales, donde los efectos de escala grandes y pequeños sobre el atractor influyen la pendiente. En general, una buena estimación de la D_c requiere un gran número de datos N , sin embargo, es prudente resaltar que el tiempo de cómputo requerido para realizar los cálculos de dicho invariante se incrementa en N^2 .

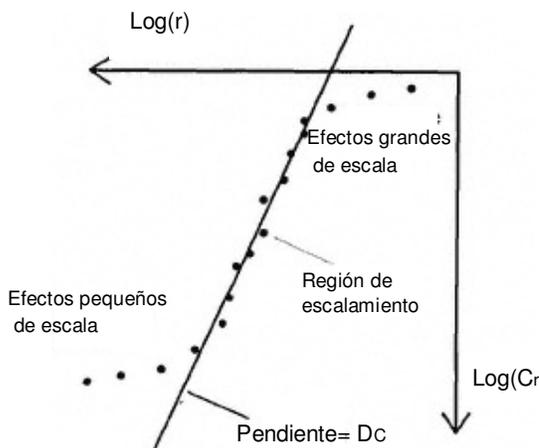


Figura 15. Determinación de la dimensión del atractor. Gráfica $\log(r)$ - $(\log C_r)$.

De esta forma, la dimensión de correlación se define como

Donde: es la correlación del conjunto de datos, dada por

$$D_C = \lim_{r \rightarrow 0} \frac{\ln(C_r)}{\ln(r)} \quad \text{por} \quad C_r = \lim_{N \rightarrow \infty} \frac{1}{N^2}$$

$$D_C = D_2$$

2.6.4. Minería de datos (MD)

La **MD** se define como el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos (Witten & Frank, 2000).

En sí la **MD** es sólo una etapa de un proceso de mayor magnitud y que representa todo un esquema de análisis de **BD**, conocido como “Descubrimiento de Conocimiento en Bases de Datos”, en inglés: Knowledge Discovery in Databases(KDD). De acuerdo a como se muestra en la figura 16:



Figura 16. Tomado de Hernández-Orallo et al, 2004.

Hernández-Orallo et al. (2004), mencionan que de acuerdo a Fallad et al., el KDD (*Knowledge Discovery in Databases*) es un proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos.

Por otra parte, la **MD** es un proceso que invierte la dinámica del método científico en el siguiente sentido: En el método científico, primero se formula la hipótesis y luego se diseña el experimento para coleccionar los datos que confirmen o refuten la hipótesis. Si esto se hace con la formalidad adecuada (cuidando cuáles son las variables controladas y cuáles experimentales), se obtiene un nuevo conocimiento (Morales-Acoltzi et al., 2011-a).

En la **MD**, se **coleccionan** los **datos** y se espera **extraer hipótesis** de ellos. En otras palabras, se busca que los datos nos describan o indiquen por qué son como son. Validar esa hipótesis inspirada por los datos en los datos mismos, será numéricamente significativa, pero experimentalmente inválida. De ahí que la **MD** debe presentar un **enfoque exploratorio**, y **no confirmador**; por lo que el resultado de tal exploración deberá ser interesante y su calidad no debe ser afectada por mayores volúmenes de datos o por ruido en los datos. En este sentido, los algoritmos de descubrimiento de información deben ser altamente robustos. Pero no todos los datos son necesarios para la extracción de características por lo que lo que se debe de definir el cubo de información o el espacio de búsqueda para efectuar las investigaciones.

Capítulo 3. Métodos y Herramientas

3.1. Base de Datos (BD)

La **BD** utilizada para el presente estudio comprende 3 variables: Precipitación (**Pcp**), Temperatura Máxima (**TempMax**) y Temperatura Mínima (**TempMin**) diarias, considerando dos periodos: 1960-1970 y 1973-1987, para dos estaciones climatológicas ubicadas en el Estado de Tlaxcala: Atlangatepec (o solamente Atlanga) y Tlaxco; es producto de un estudio realizado por Morales-Acoltzi T. et al. (2011) denominado: "Variabilidad y Cambio Climático de la lluvia en Tres Sitios en el Estado de Tlaxcala ". Dichos autores, aplicaron análisis de calidad a los datos mediante el software libre RCLimDex(1.0), el cual se usa comúnmente para calcular índices de extremos climáticos y es desarrollado y mantenido por Xuebin Zhang y Feng Yang del Departamento de Investigación Climática del Servicio Meteorológico de Canadá (Santos, 2004). Lo último asegura que se ha disminuido el grado de incertidumbre de las mediciones presente en los datos.

3.2. Zona de estudio

El Estado de Tlaxcala es el más pequeño de México en superficie, con 3914 [km²], lo cual representa 0.2% del territorio nacional y su altitud media es de 2,230 [msnm].

El clima predominante en Tlaxcala (Figura 1) es templado subhúmedo, con lluvias en verano, aunque por estar situado en el eje neovolcánico, su topografía es montañosa y por lo tanto presenta una gran cantidad de microclimas. La temperatura media anual estatal es de 27[°C]. La lluvia media anual estatal es de 705 [mm], con lluvias medias anuales mayores en el centro y sur (de 600 a 1,200 [mm]). Las lluvias medias mensuales se concentran de junio a septiembre (70% del total anual) (Morales-Acoltzi, 2011; Meneses, 1993). En cuanto a las estaciones climatológicas, es importante mencionar que éstas se encuentran separadas una distancia de 10 [km] y presentan una diferencia de alturas respecto al nivel medio del mar de 40 [m] aproximadamente. Además, en la estación Atlanga existe un cuerpo de agua (una presa) permanente que influye notablemente en los cambios

meteorológicos de los alrededores, mientras la estación Tlaxco se encuentra sobre un terreno plano, ligeramente inclinado hacia la presa.



Figura 1. Zona de estudio.

3.3. Métodos y Herramientas computacionales

Con el fin de estimar los parámetros para efectuar la reconstrucción del atractor, existen métodos teóricos y algoritmos prácticos, como se mostró de manera teórica en el capítulo 2, en este trabajo se consideran los métodos de **ANST** "estándar".

Al contrario que sus contrapartes lineales, los métodos utilizados en el **ANST** tienen relativamente poco tiempo de haber alcanzado su madurez, por lo que aún se encuentran en una fase tanto "experimental" como de desarrollo y refinamiento, a pesar de la sólida base matemática que los fundamenta. Así, dentro de este contexto, se tienen dos opciones para implementar de forma práctica los métodos de **ANST**.

En primera instancia, surgió un paquete de programas denominado **TISEAN**, el cual es el resultado de un grupo interdisciplinario de investigadores en varios años. Constituye la herramienta ideal para el análisis no lineal de series de tiempo. Desarrollado por el Instituto de Física y Teoría Química de la Universidad de Frankfurt, el Instituto Max Planck de Física Compleja, el

Departamento de Física de la Universidad de Wuppertal, Alemania, por Rainer Hegger, Holger Kantz, Thomas Schreiber y Andreas Schmitz (Kantz & Schreiber, 1997). **TISEAN** es un software libre, cuyo proyecto ha crecido incluso por fuera de dicho grupo de investigadores y está disponible desde febrero de 1998. Los programas fuente están disponibles en código C y Fortran, la teoría en la que se basan y la explicación de su funcionamiento puede consultarse en la página. Además, debido a que es un software libre, es posible descargar tanto los códigos fuente como los archivos ejecutables. Sobre su funcionamiento es preciso señalar que no posee interfaz gráfica alguna, en su lugar, se debe utilizar la consola de comandos (shell) de Unix o bien la de MS-DOS; por lo que tanto las entradas como las salidas (los resultados), en la mayoría de los casos son archivos de tipo texto que deben de tratarse o procesarse con otras aplicaciones, con el fin de ejecutar el análisis correspondiente.

Al respecto de esta situación, Morales-Acoltzi et al. (2011), desarrollaron un sistema en el software **MATLAB**, con el cual efectuaron experimentos numéricos, acoplando algunas rutinas del software **TISEAN** (Hegger, Kantz & Schreiber, 1999), con énfasis en el cálculo de la dimensión del atractor. En lo posterior, dicho sistema será referido con el nombre "TISEAN_M".

En la presente tesis se tuvo acceso al código fuente de dicho sistema computacional, así, éste se modificó a manera de dejar intacta su simpleza, se corrigieron algunos errores y sobre todo, se acoplaron otras rutinas de **TISEAN** con **MATLAB**, por lo que fue necesario aprender primero el funcionamiento de dicho sistema para después mejorarlo y complementarlo.

Las rutinas de **TISEAN** acopladas en el sistema original son: "lorenz", "henon", "ikedada", "mutual" y "d2". De las cuales sólo se consideró la penúltima; adicionalmente a ésta, se acoplaron: "corr", "false_nearest" y "delay".

Por otro lado, existe también **VRA**, el cual es un software escrito en C++ por Eugene Kononov, corre en Windows 95, 98, 2000 y NT, y tiene como función efectuar el análisis topológico, cualitativo y cuantitativo y la predicción no paramétrica de **ST** no lineales y caóticas. Éste último es fácil de usar, ya que posee una interfaz gráfica.

Por consiguiente, el procedimiento seguido para tratar los datos tiene el esquema siguiente:

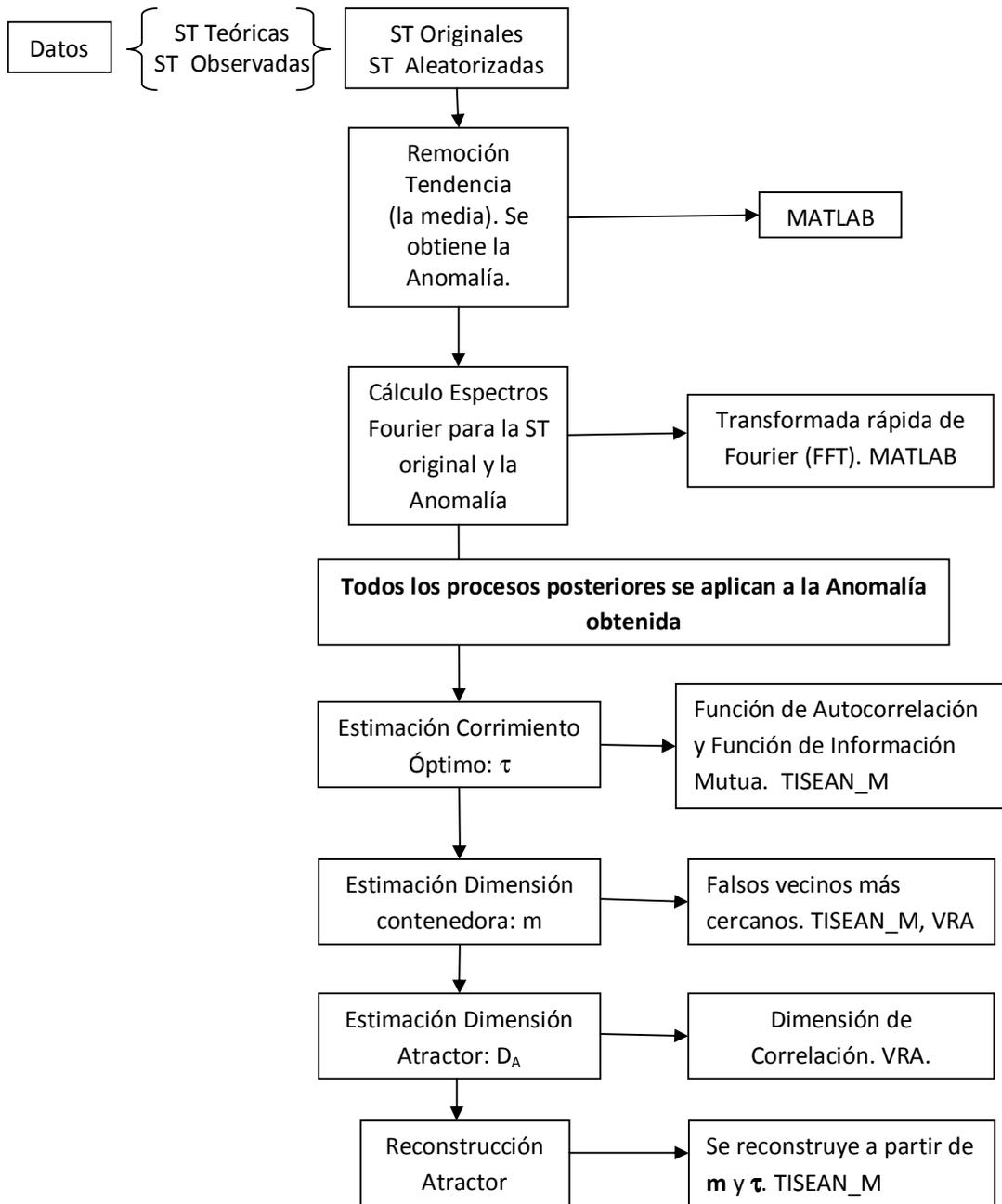


Figura 2. Secuencia del tratamiento de los datos mediante los diferentes métodos. Se señala el método aplicado con el respectivo programa computacional utilizado para su implementación práctica.

Capítulo 4. Experimentos numéricos

Se efectuaron varios experimentos numéricos; tanto con series teóricas-deterministas, como con las series experimentales descritas en el esquema planteado en el capítulo 3:

Lo primero es graficar la ST original contra el número total de muestras. Posteriormente se elimina la tendencia (se remueve la media aritmética de los datos), es decir, se obtiene la anomalía y se calculan los espectros de Fourier con el objeto de evaluar cuáles son las frecuencias que más contribuyen a la varianza, recordando que éstas son frecuencias medias. Así, se obtienen los espectros de amplitudes y los respectivos espectros de potencia, tanto para la ST original como para la anomalía, efectuando un cambio de escala en el eje vertical con el objeto de apreciar mejor los valores de las amplitudes, obteniendo de esta forma, gráficos semilogarítmicos. Los posteriores procesos son aplicados a la anomalía.

El siguiente paso es la estimación del corrimiento óptimo, denominado τ . Para esto se emplea la función de autocorrelación, la cual opera en el dominio del tiempo y mide correlaciones lineales. Se acostumbra utilizar el primer valor del corrimiento que corresponda al primer cero de la función de autocorrelación. Por otro lado, también se calcula la función de información mutua media, la cual opera en términos de probabilidades y mide la dependencia general (no lineal) de dos variables. De acuerdo a la teoría, se usa el primer mínimo de dicha función. Cabe mencionar que de coincidir los valores del corrimiento obtenido mediante ambas funciones, no sería necesario utilizar el criterio proporcionado por la función de información mutua media, ya que la dependencia descrita sería de tipo lineal y el uso de dicha herramienta sobraría. Después, se estima la dimensión que contendrá al atractor de la ST, mediante la búsqueda de falsos vecinos más cercanos, de acuerdo a la teoría. Posterior a esto, se calcula la Dimensión de Correlación, lo que corresponde a estimar la dimensión del atractor. Una vez estimadas la dimensión de inserción m y el corrimiento óptimo τ , se procede a calcular las componentes de los m-vectores de inserción del respectivo espacio de fase y así, reconstruir el

atractor y representarlo gráficamente en tres dimensiones. Así, la dinámica del sistema queda resumida mediante su representación geométrica.

Ahora bien, para mostrar que está presente el determinismo en las series, se realizó la aleatorización de todas las series teóricas; en el caso de las ST observadas, sólo se aleatorizó las correspondientes a la temperatura mínima, y de igual manera se les aplicó el mismo procedimiento planteado, obteniendo lo siguiente:

4.1. Experimentos con ST teóricas

Las ST seleccionadas para efectuar la primer parte de experimentos, corresponden a las soluciones del sistema de ecuaciones diferenciales de Lorenz, las cuales constituyen un modelo teórico para estudiar el atractor que éstas generan en el espacio de estado:

$$\begin{aligned}\dot{x} &= -\sigma(x - y); \sigma > 0 \\ \dot{y} &= -xz + rx - y; r > 0 \\ \dot{z} &= xy - bz; b > 0\end{aligned}$$

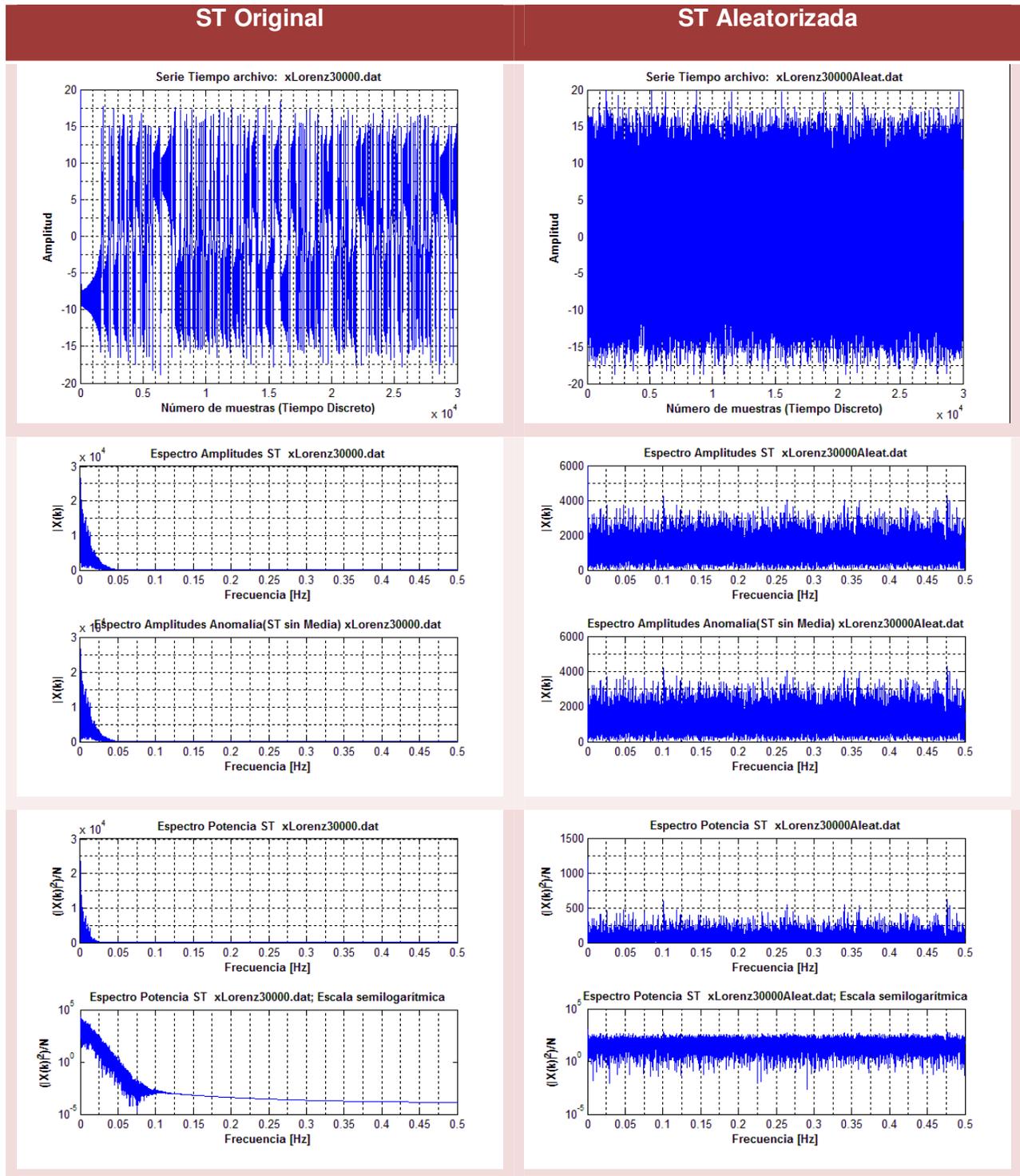
Donde:

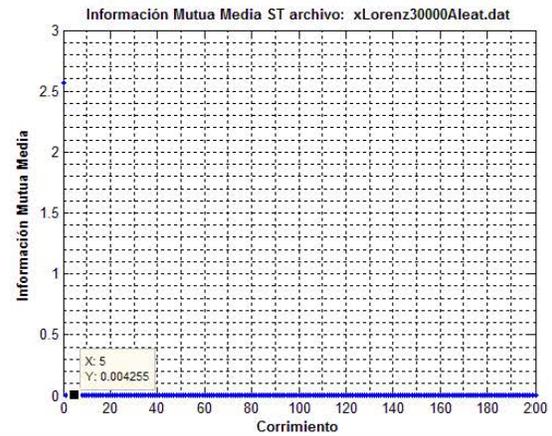
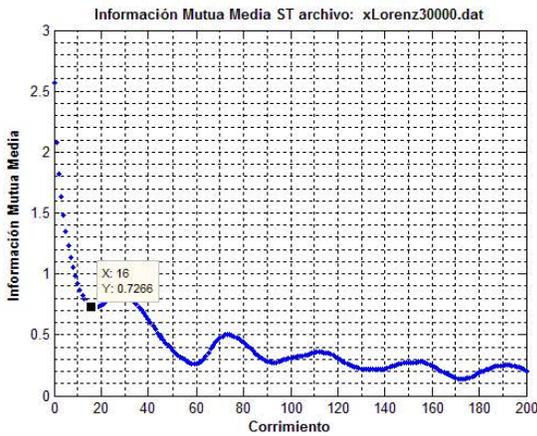
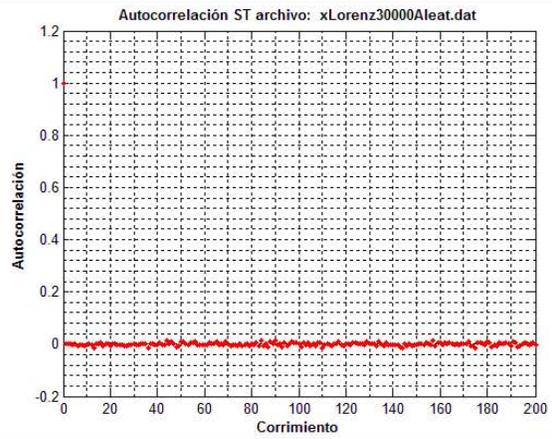
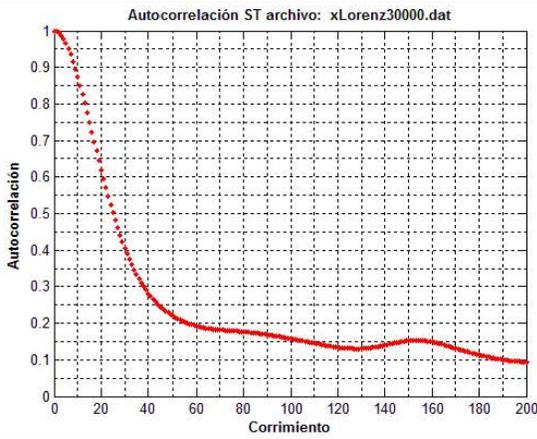
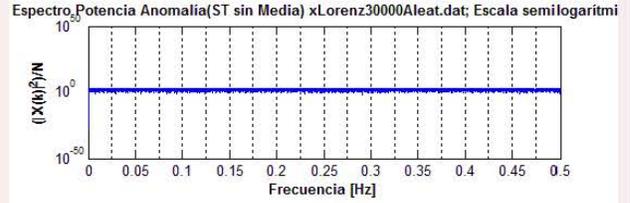
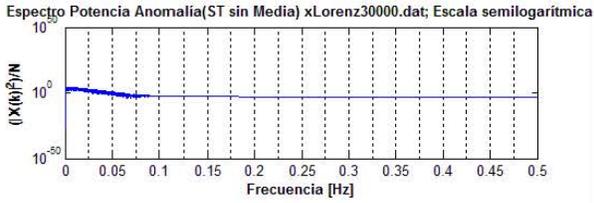
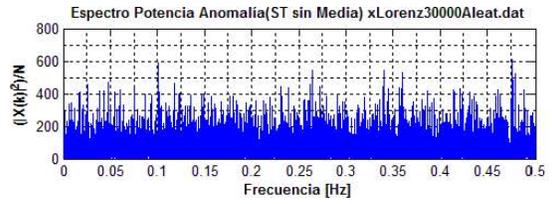
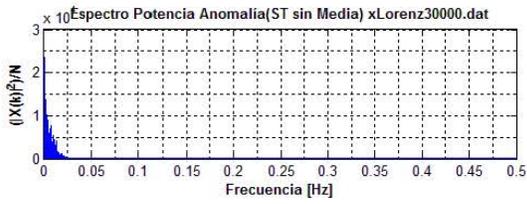
$\sigma, r, b \rightarrow$ Son escalares y el punto denota la primera derivada de las respectivas variables. Ver Anexo 1. Para cada variable se tienen 30000 datos. De acuerdo a la teoría, puede pasarse del espacio de estado al espacio de fase considerando una sola variable y sus derivadas, así, puede usarse sólo una variable de las anteriores para reconstruir el atractor en el espacio de fase, y dado que todas las variables deben registrar la misma dinámica, deberían obtenerse los mismos resultados al usar cualquiera de las tres:

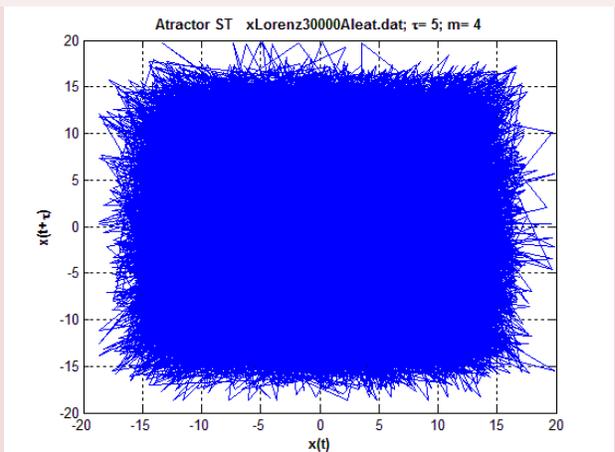
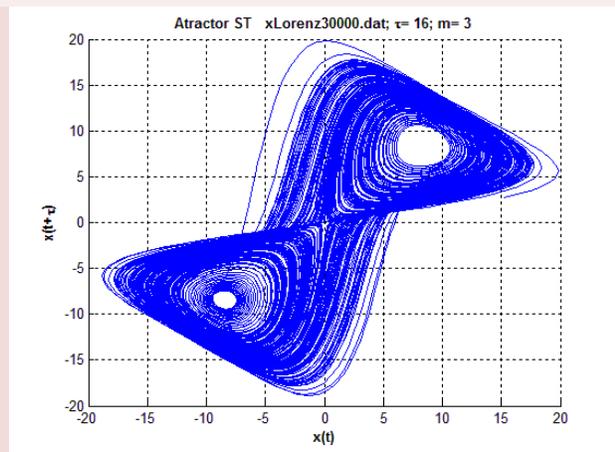
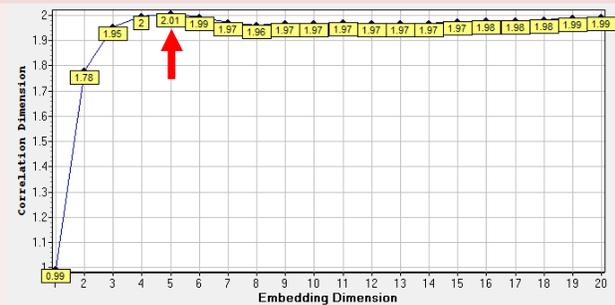
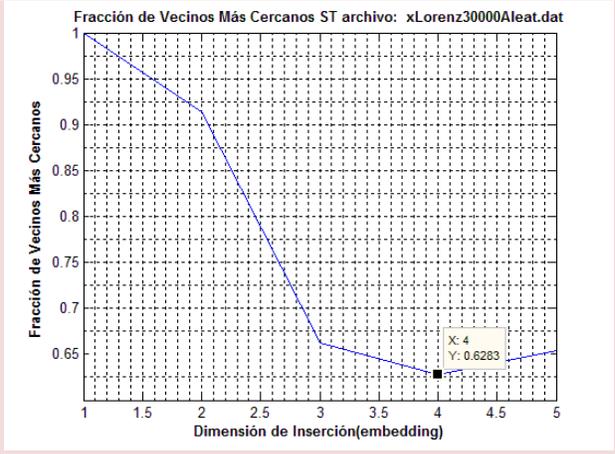
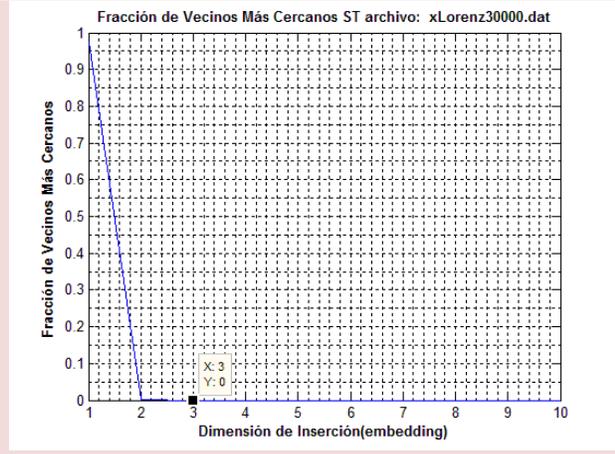
$$\left. \begin{array}{l} (x, y, z) \\ \text{Espacio} \\ \text{de estado} \end{array} \right\} \left\{ \begin{array}{l} \ddot{x} = f(x, \dot{x}, \ddot{x}) \\ \ddot{y} = f(y, \dot{y}, \ddot{y}) \\ \ddot{z} = f(z, \dot{z}, \ddot{z}) \end{array} \right\} \begin{array}{l} \text{Espacios} \\ \text{de fase} \end{array}$$

En otras palabras, las propiedades dinámicas: τ (corrimiento óptimo) y la dimensión del atractor (D_2), entre otras, deberían de ser iguales. También, como parte de los experimentos, cada una de las ST originales fue aleatorizada para probar el hecho de que provienen de un sistema determinista de ecuaciones, y que el comportamiento descrito por los datos se ve afectado por el orden de éstos. Así, de acuerdo al esquema de tratamiento de las ST propuesto en el capítulo 3, se presentan los experimentos realizados para el modelo de Lorenz:

4.1.1. Experimentos ST Lorenz: Variable “x”

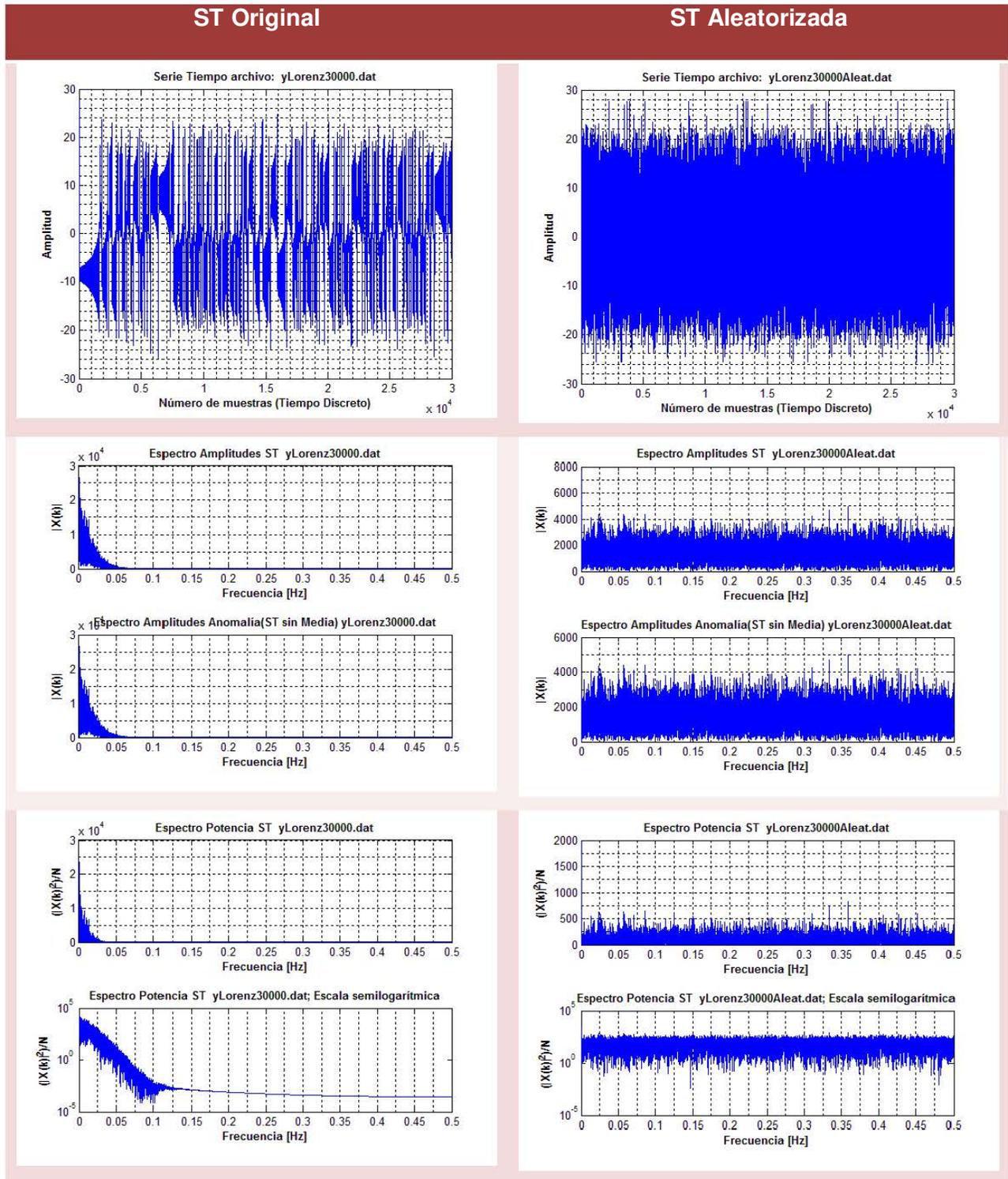


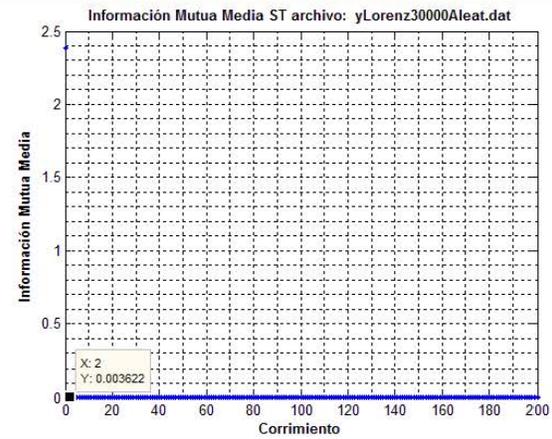
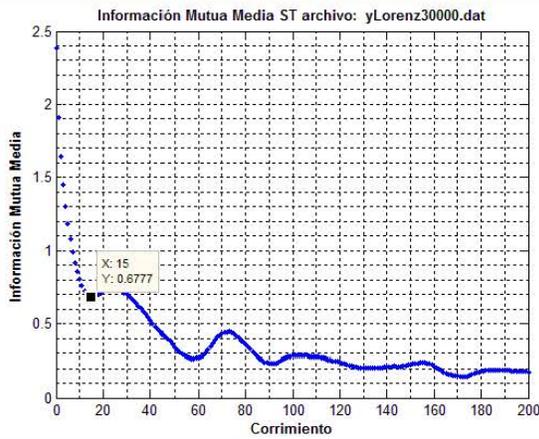
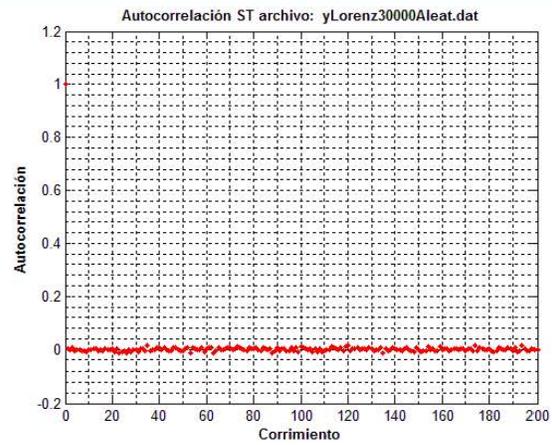
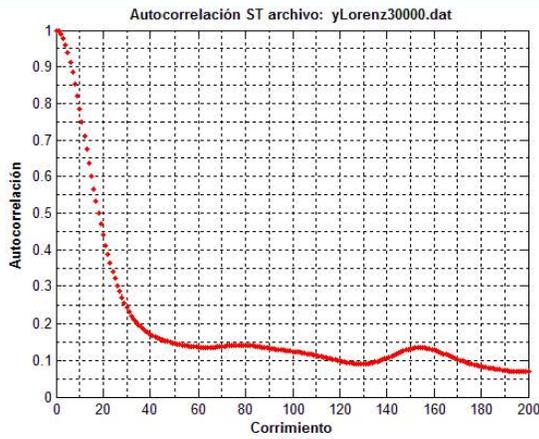
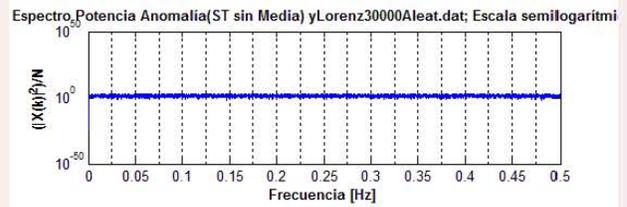
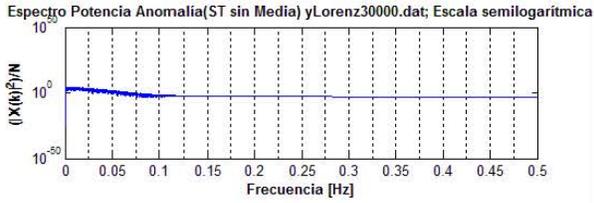
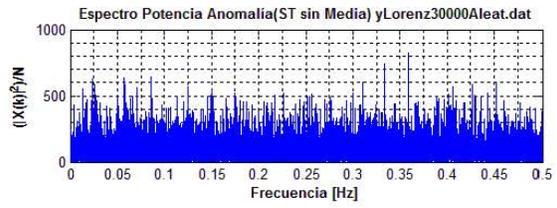
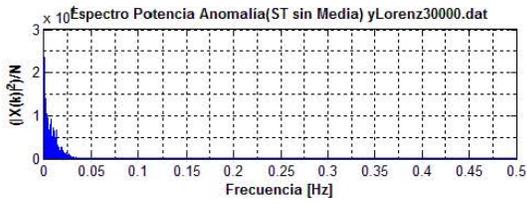


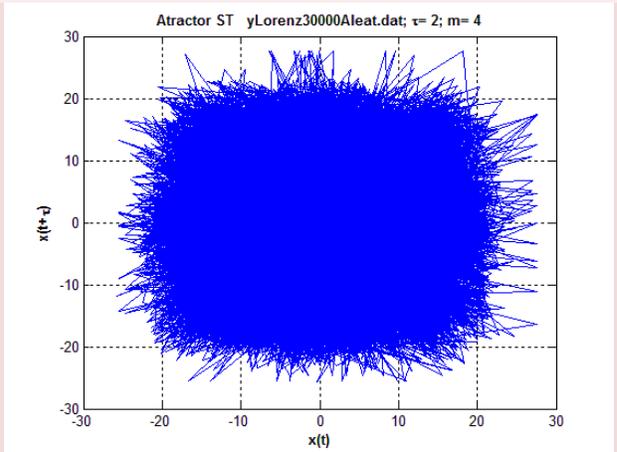
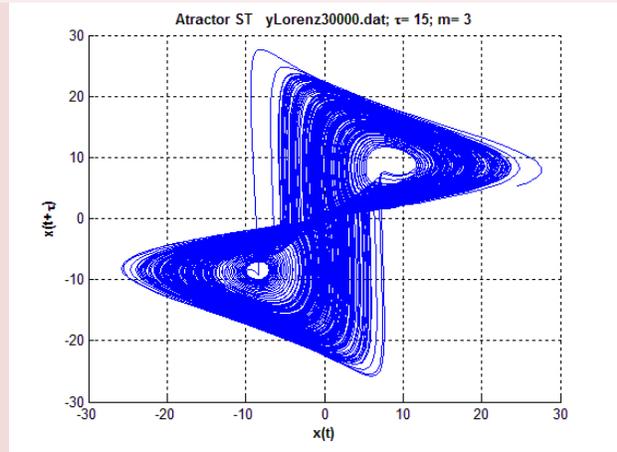
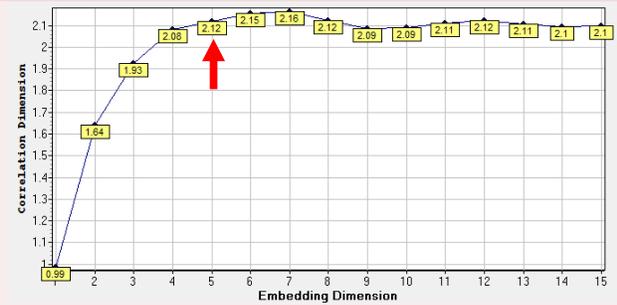
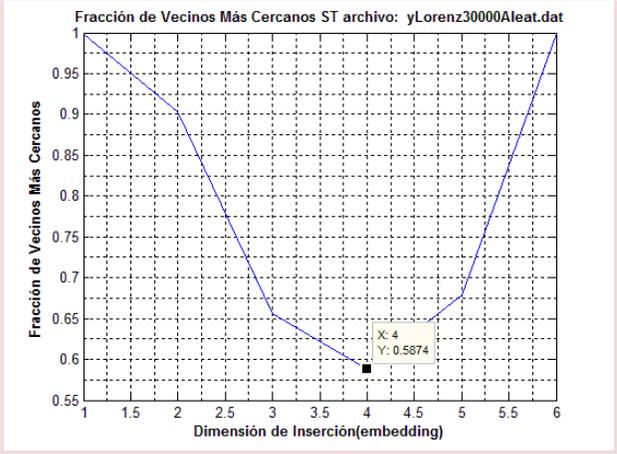
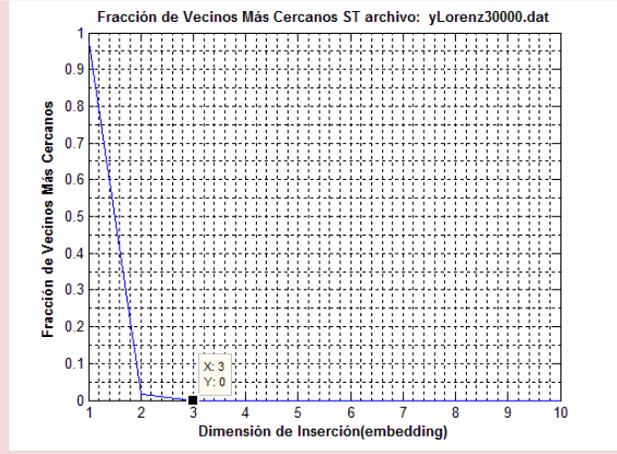


Experimento numérico para comparar la ST teórica correspondiente a la variable “x” del sistema de ecuaciones de Lorenz, con la misma ST pero aleatorizada, usando muestreo con reemplazo. Esto con el fin de exponer los dos casos extremos y demostrar la existencia de una estructura definida en los datos, determinismo.

4.1.2. Experimentos ST Lorenz: Variable “y”

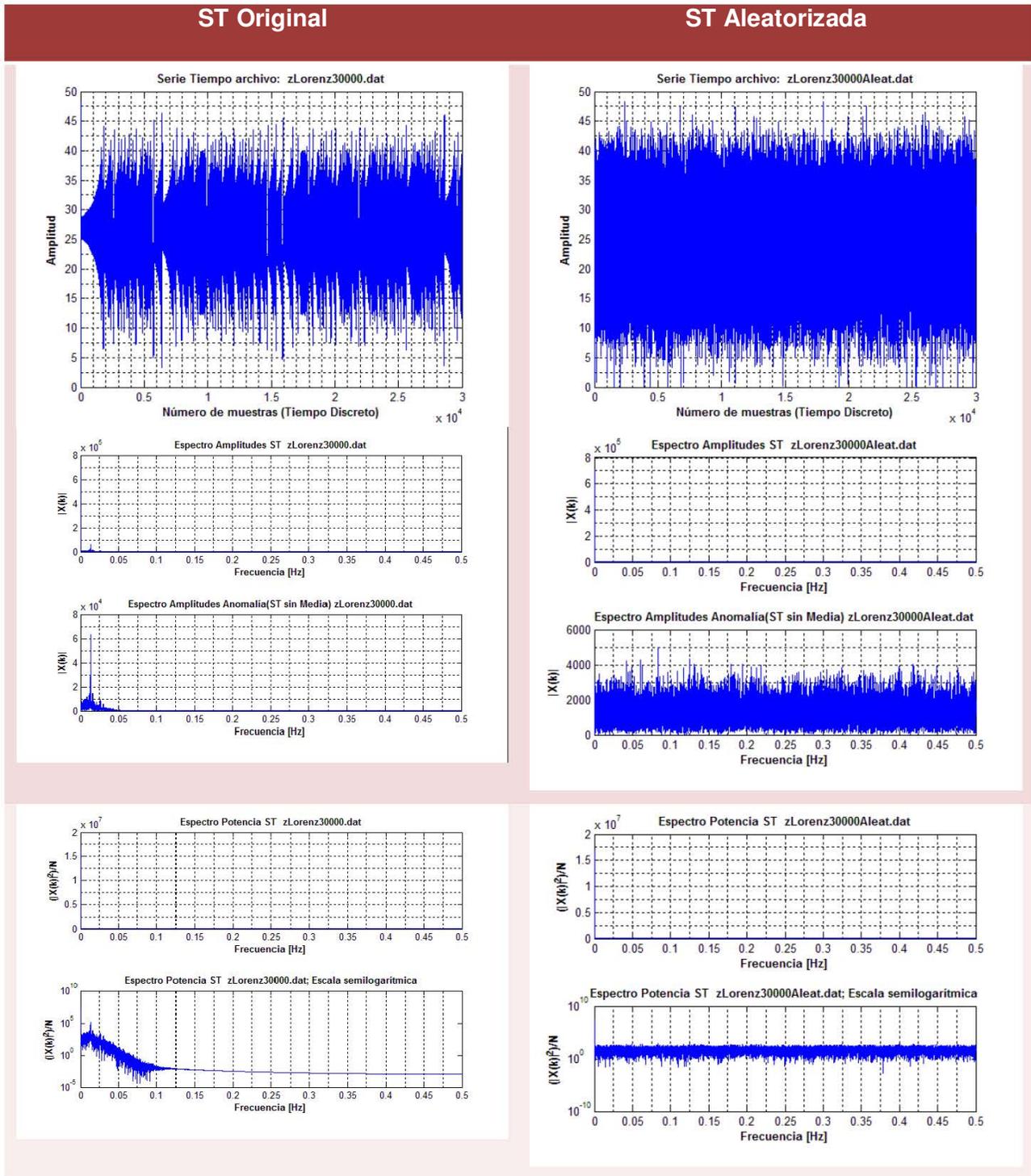


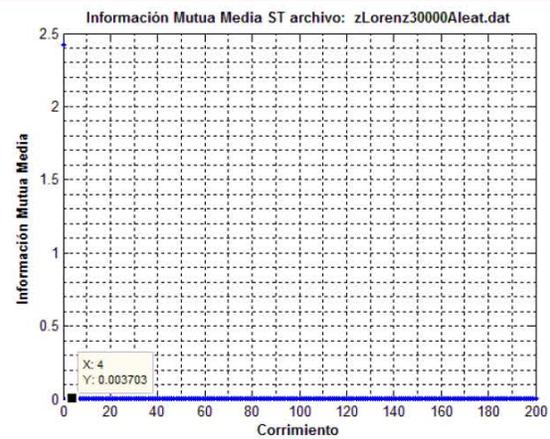
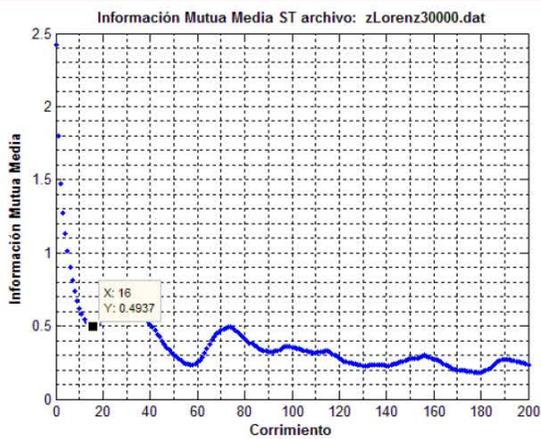
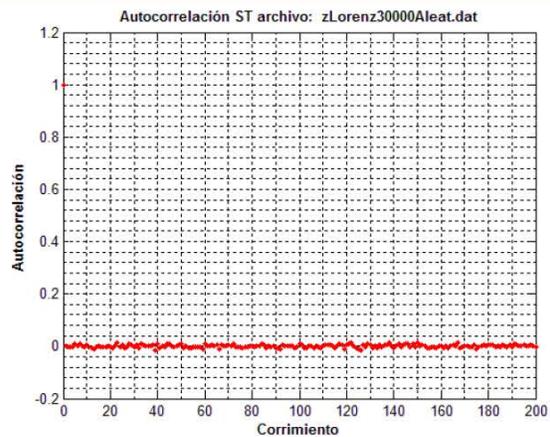
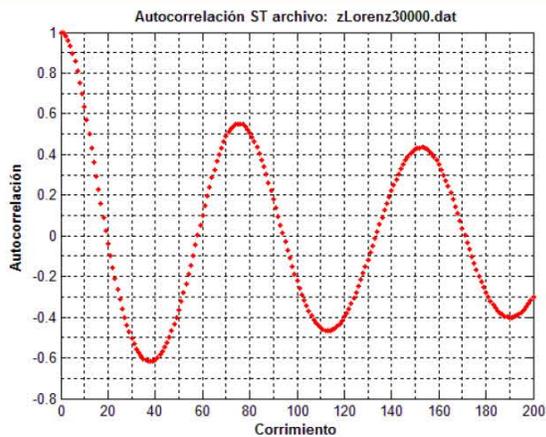
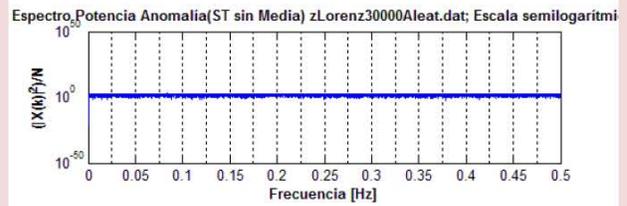
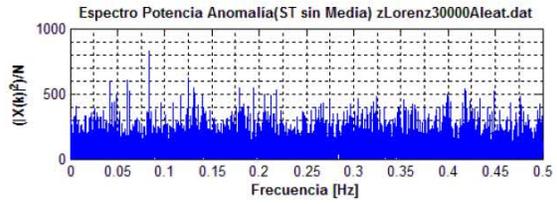
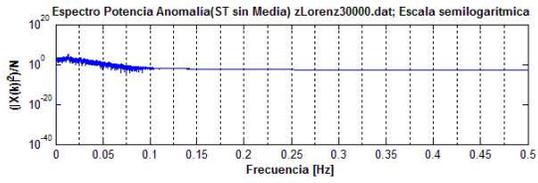
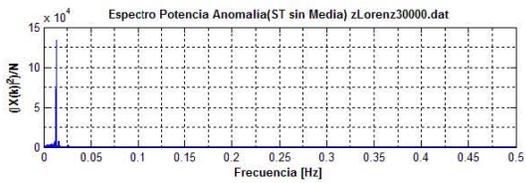


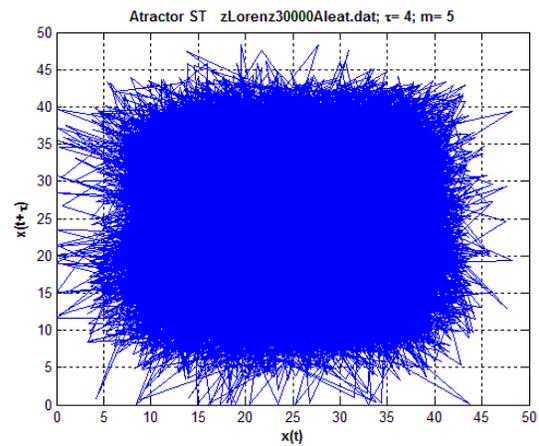
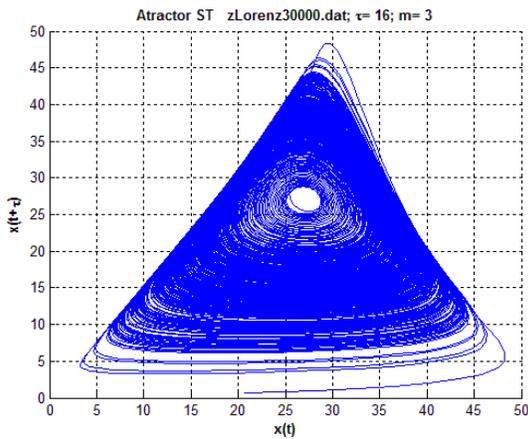
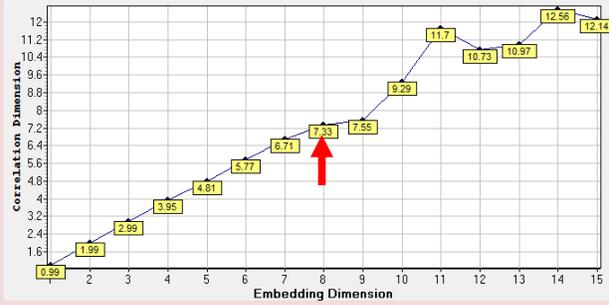
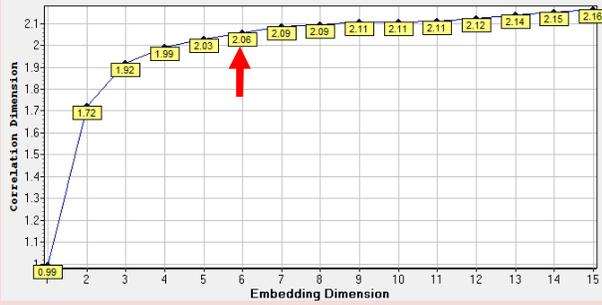
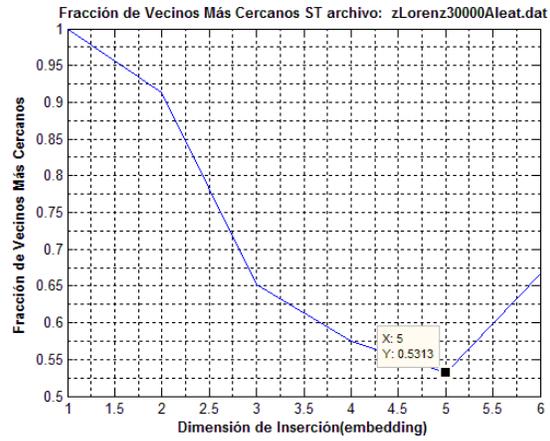
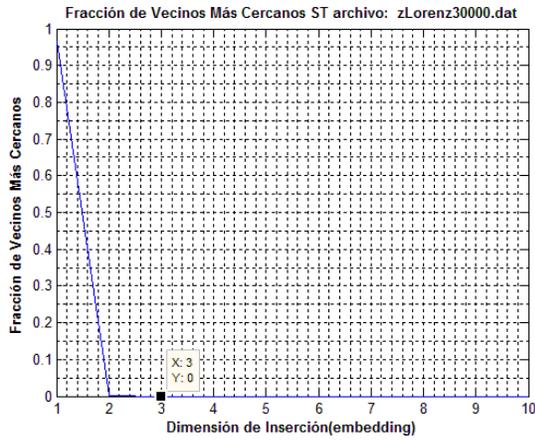


Experimento numérico para comparar la ST teórica correspondiente a la variable “y” del sistema de ecuaciones de Lorenz, con la misma ST pero aleatorizada, usando muestreo con reemplazo. Esto con el fin de exponer los dos casos extremos y demostrar la existencia de una estructura definida en los datos (determinismo).

4.1.3. Experimentos ST Lorenz: Variable “z”







Experimento numérico para comparar la ST teórica de la variable “z” del sistema de ecuaciones de Lorenz, con la misma ST pero aleatorizada, usando muestreo con reemplazo. Lo anterior busca exponer los dos casos extremos y demostrar la existencia determinismo en los datos.

4.2. Experimentos con ST observadas

Los experimentos concernientes a datos observados, abarcan dos periodos de tiempo, así las ST de precipitación, temperatura máxima y temperatura mínima con las que se cuenta, se descompusieron de la siguiente forma:

Denominación	Periodo comprendido	Número de datos
P1	1960-1970	4018
P2	1973-1987	5478

Estudios realizados por Morales-Acoltzi et al. (2011) señalan que en el estado de Tlaxcala, del año 1970 al año 1993 aproximadamente, se presentó un incremento en el número de días que contribuyeron a la lluvia anual (Figura 1), por lo que la descomposición de las ST planteada, corresponde a dos regímenes diferentes de condiciones atmosféricas diferenciables: el comienzo del aumento gradual (suave) en la cantidad de lluvia (P1), y el periodo en el que la cantidad de precipitación se mantuvo entre 600 y 900 [mm/año], denominado P2.

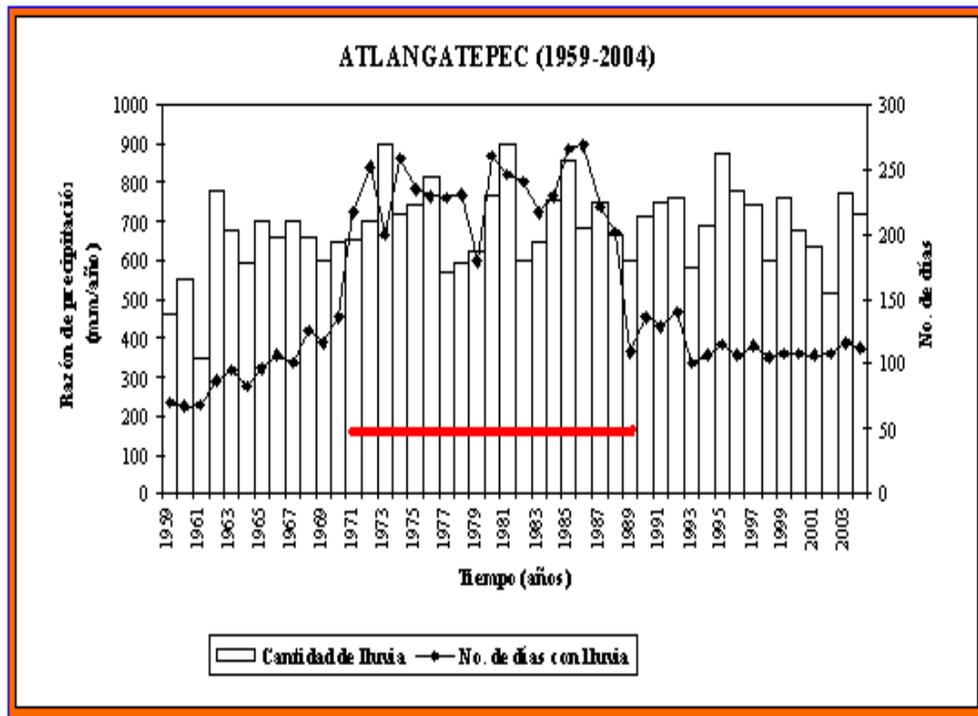
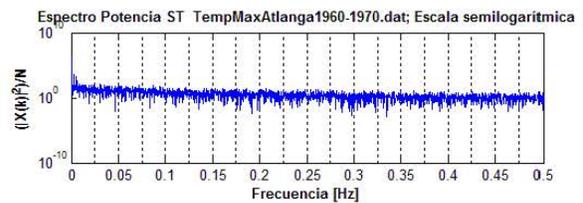
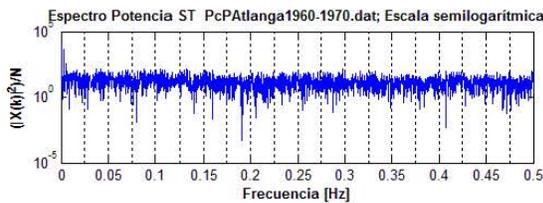
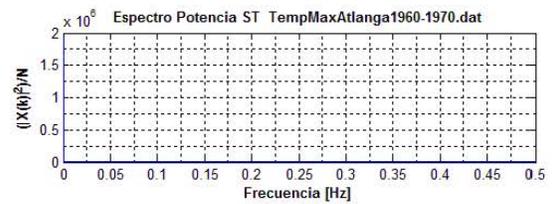
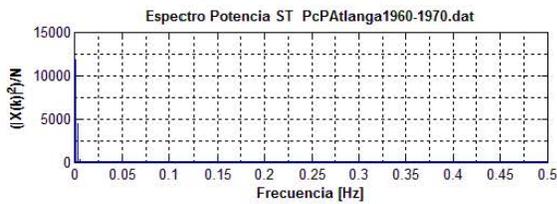
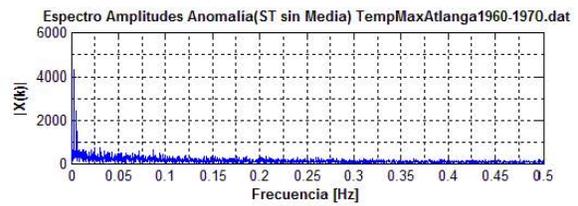
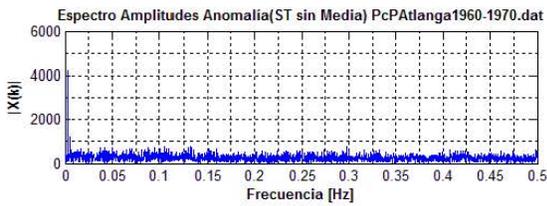
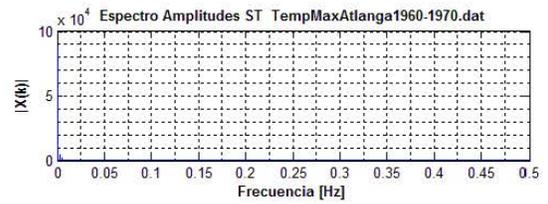
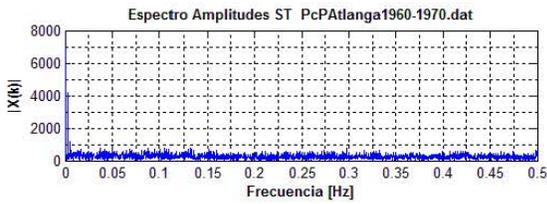
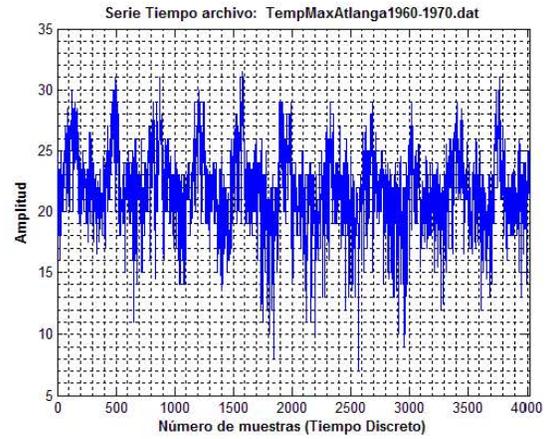
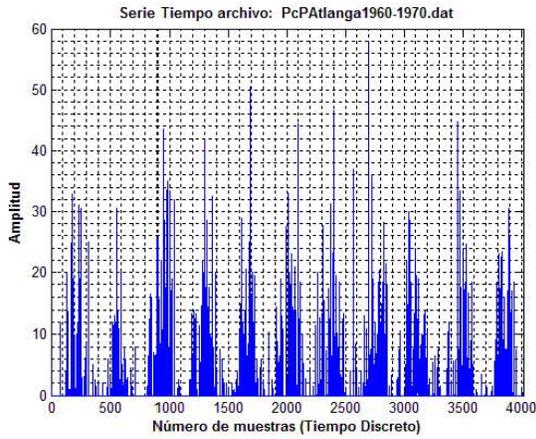


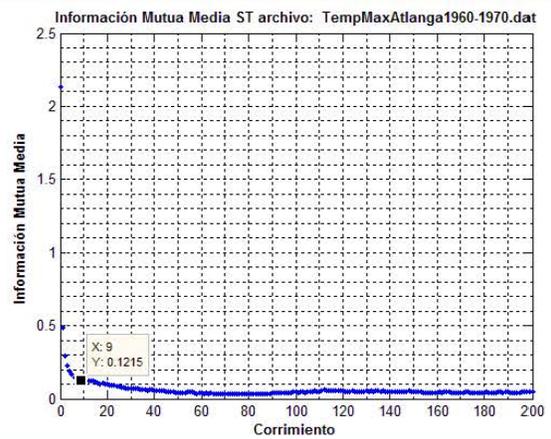
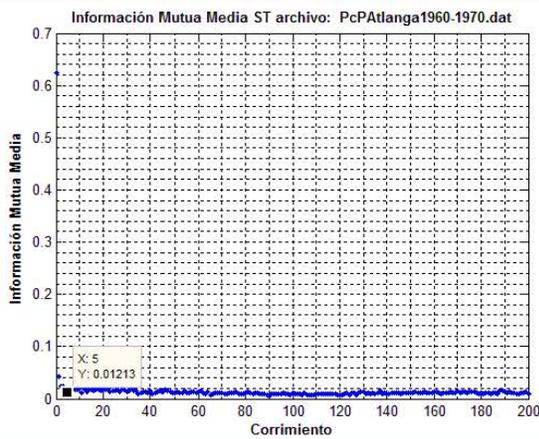
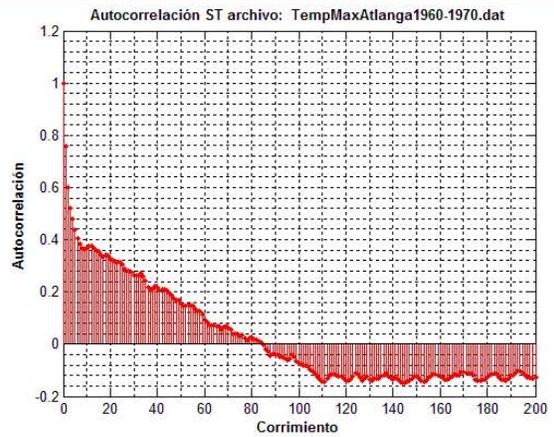
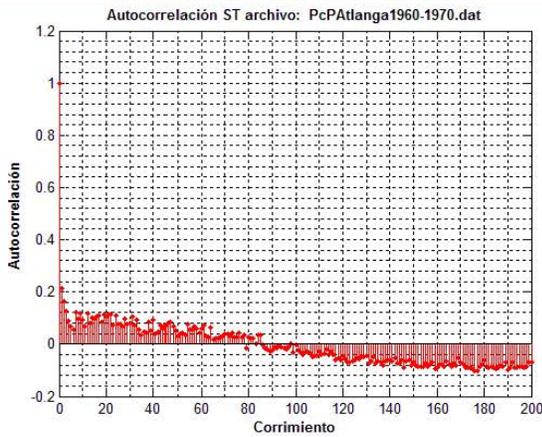
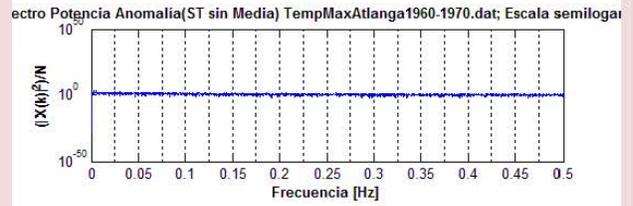
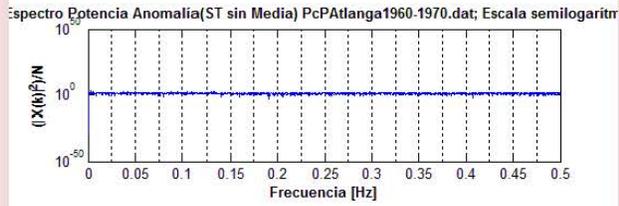
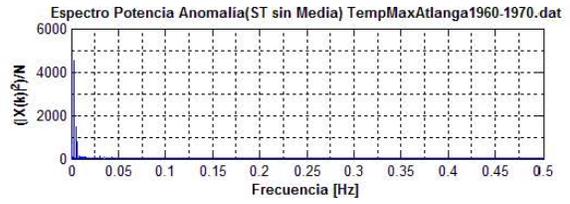
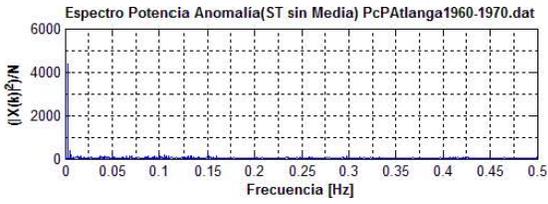
Figura 1. Precipitación anual Vs. Número de días que contribuyeron a la lluvia anual para la región de Atlangatepec (Tomado de Morales Acoltzi et al., 2011).

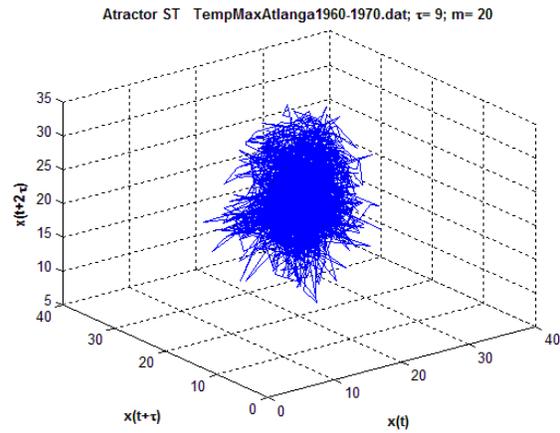
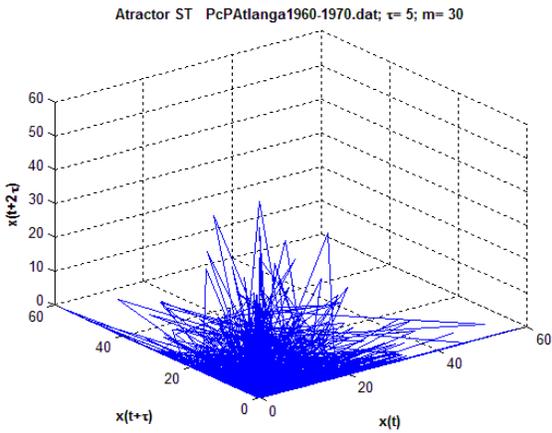
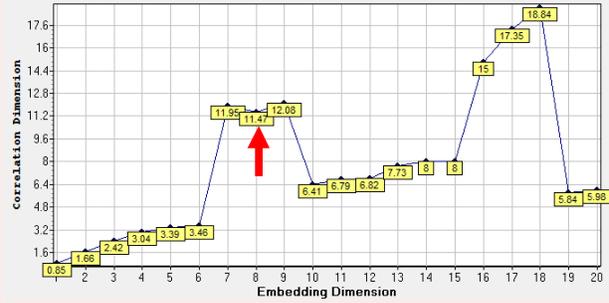
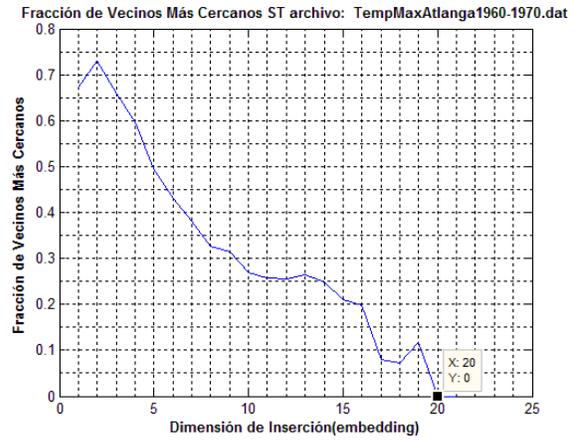
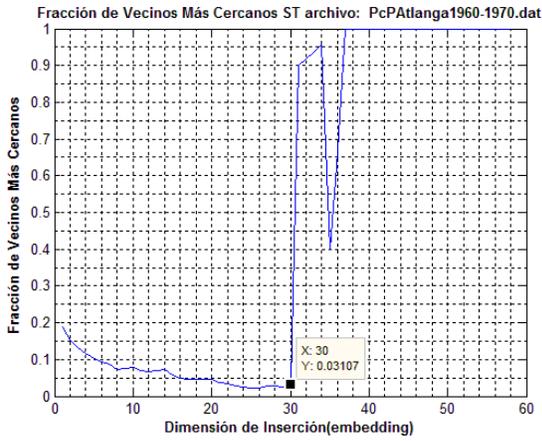
4.2.1. Experimentos Estación Atlanga P1: 1960-1970

ST Precipitación

ST Temperatura Máxima

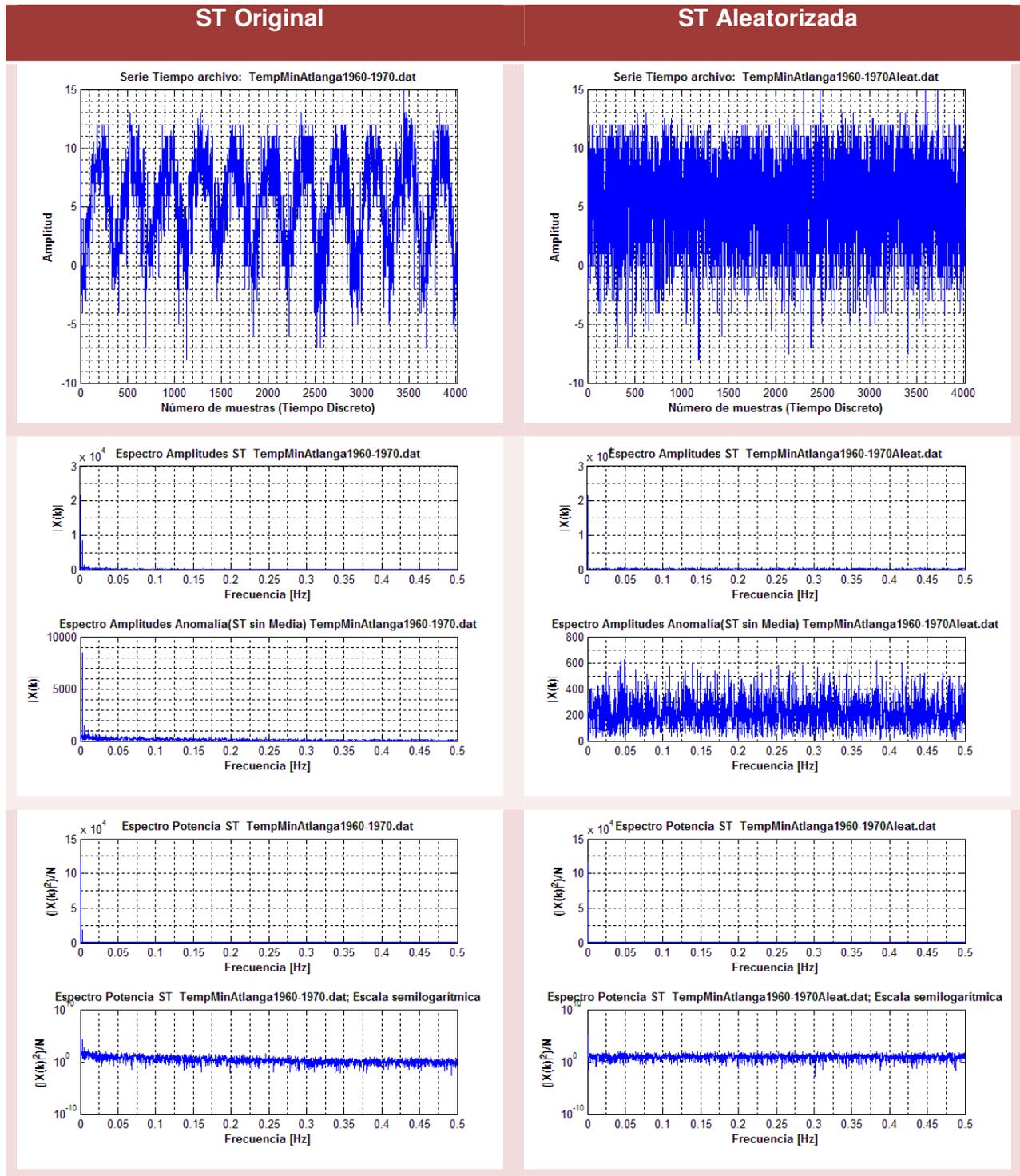


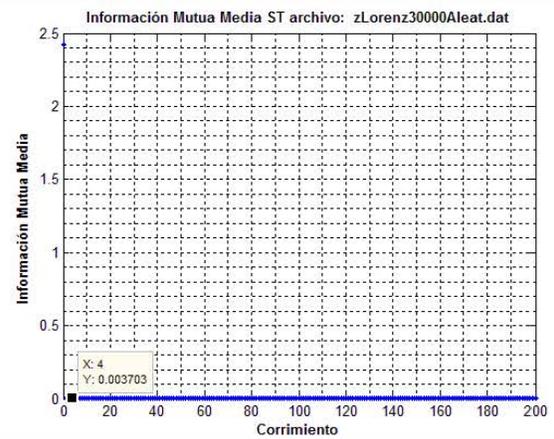
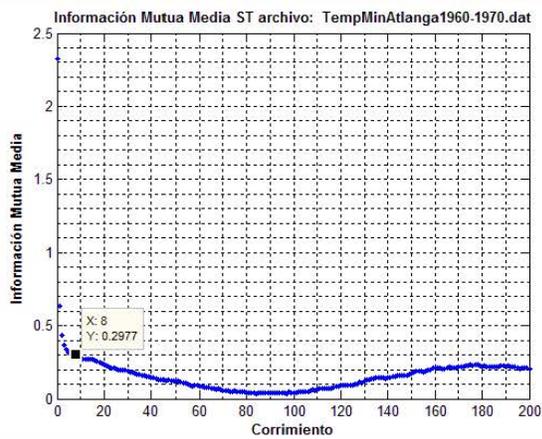
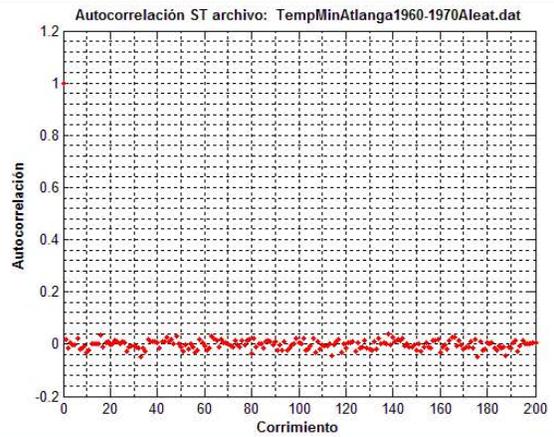
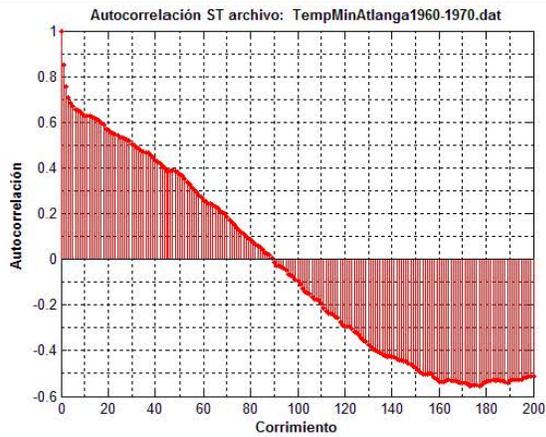
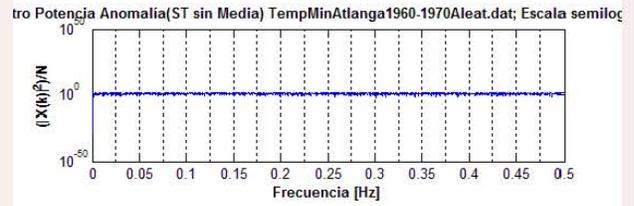
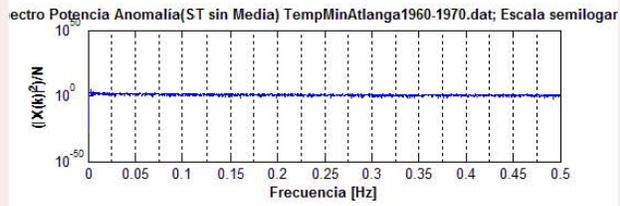
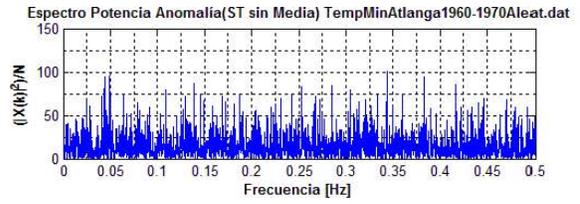
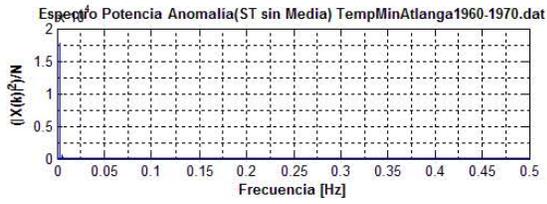


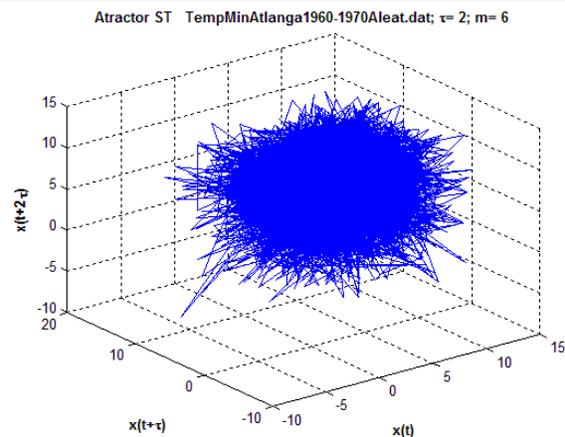
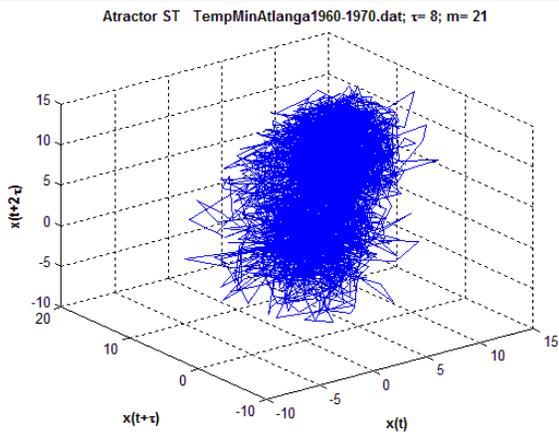
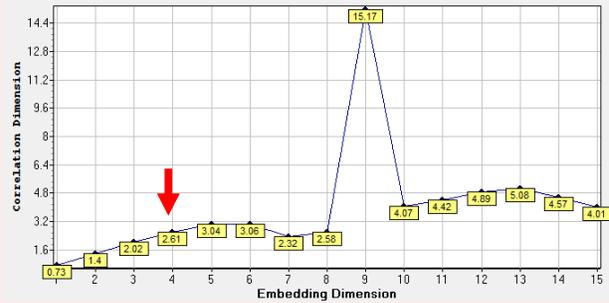
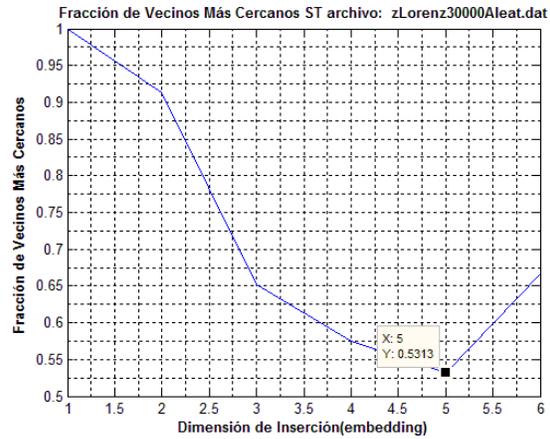
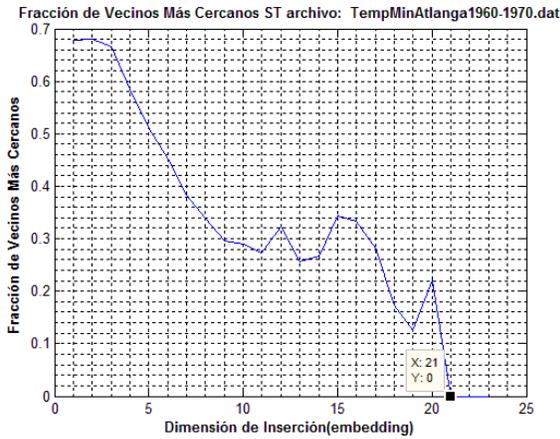


Experimento numérico para comparar el comportamiento de anomalías de las variables observadas precipitación y temperatura máxima y estimar los valores de los parámetros correspondientes a su dinámica.

4.2.2. Experimentos Estación Atlanga P1: 1960-1970; Temperatura mínima

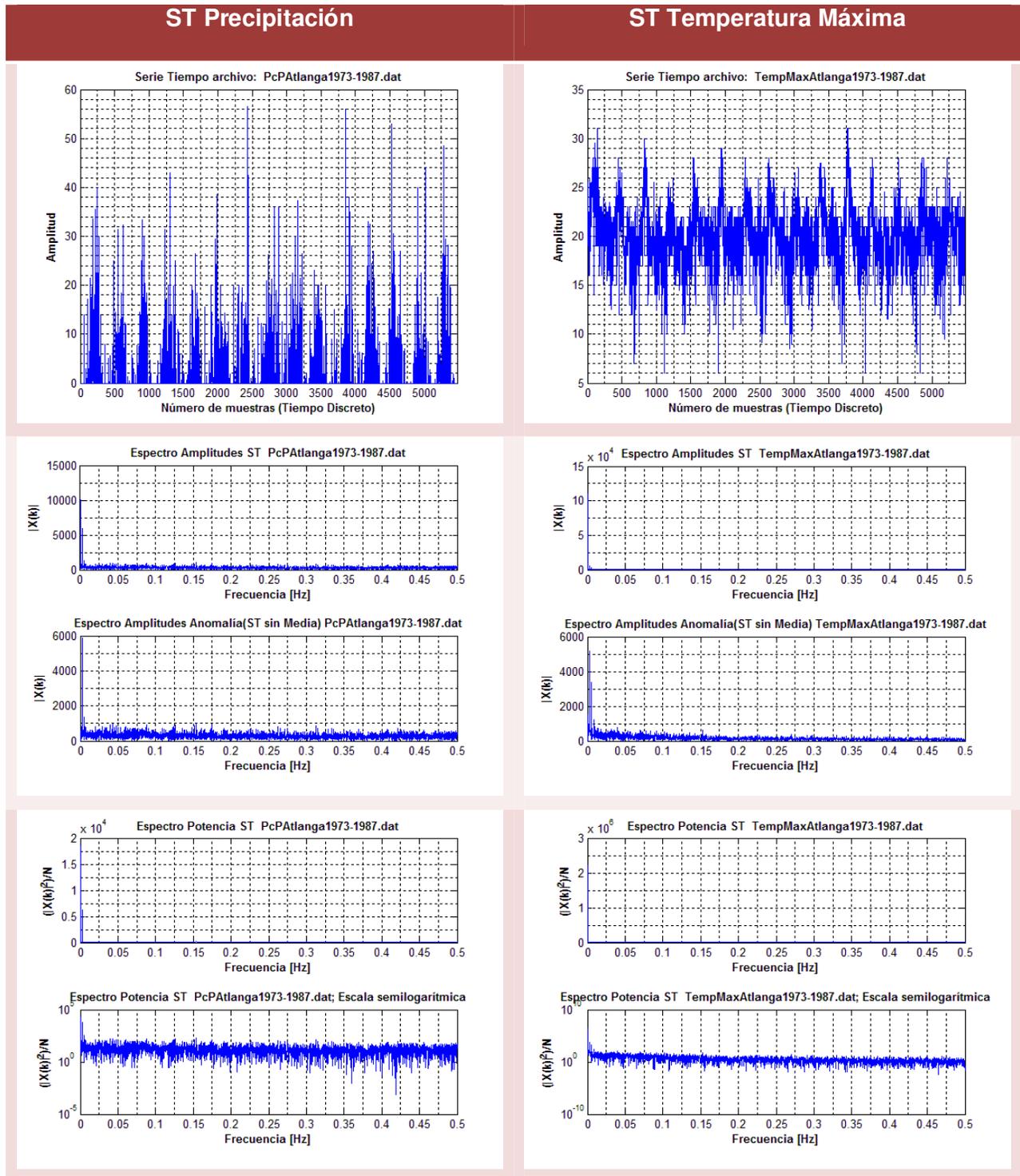


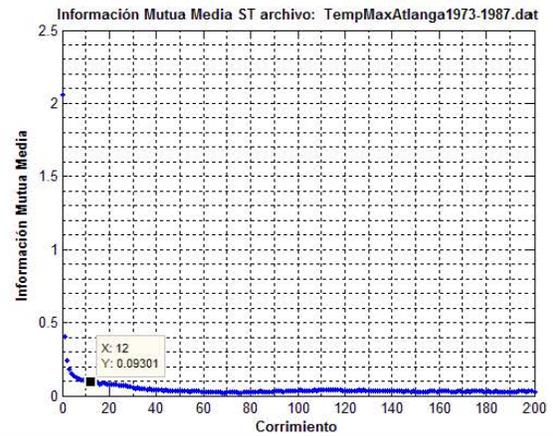
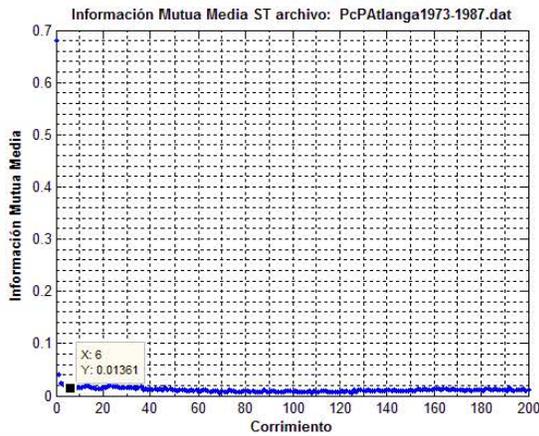
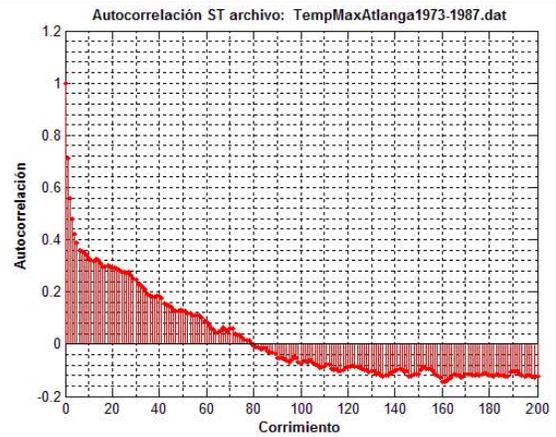
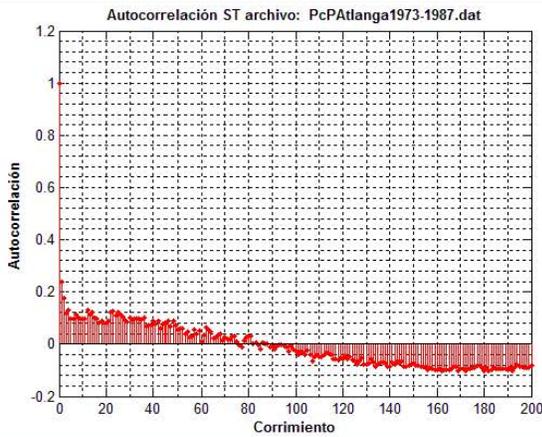
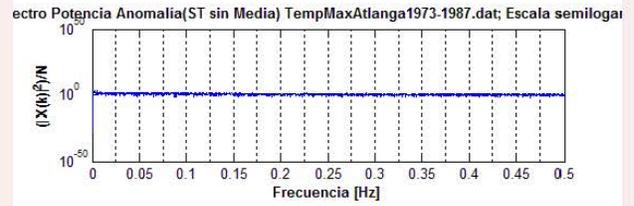
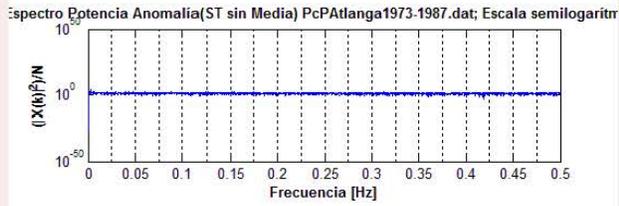
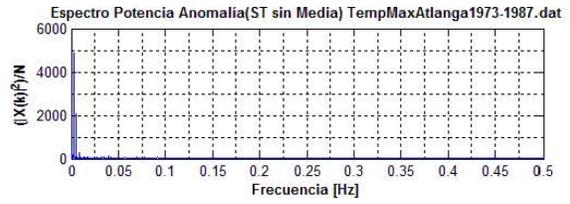
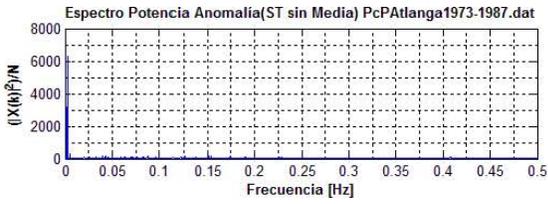


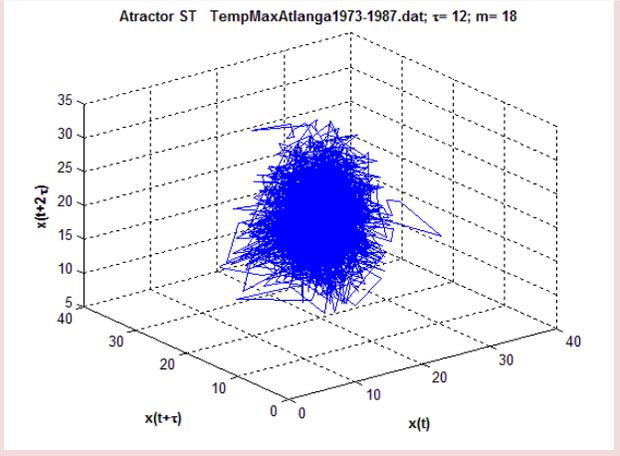
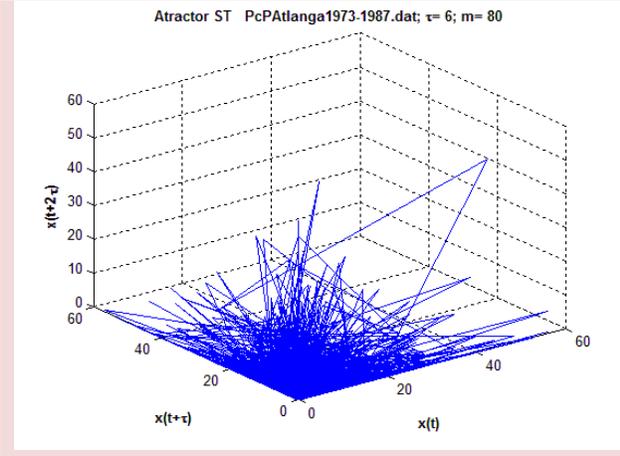
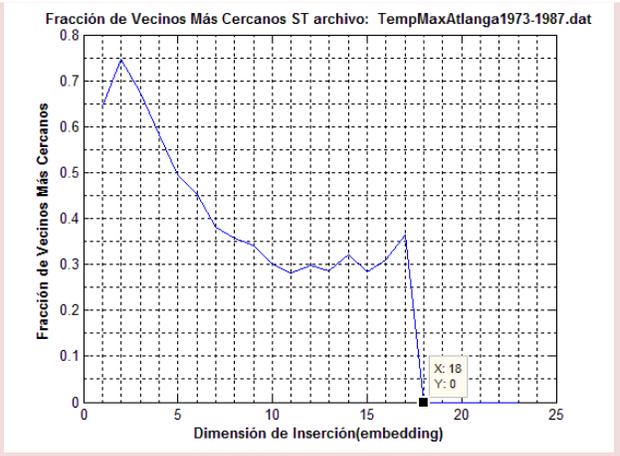
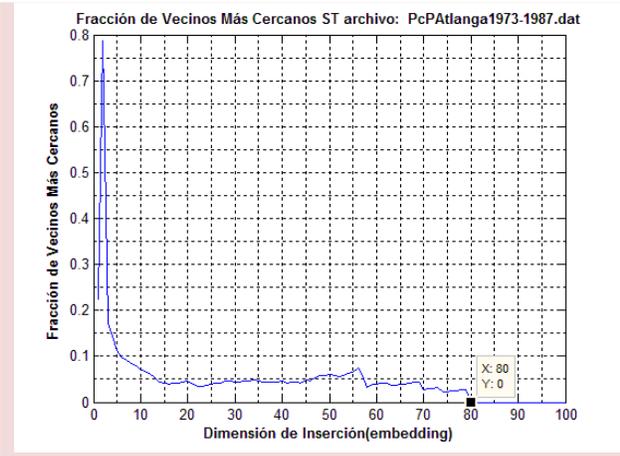


Experimento numérico para comparar el comportamiento de la anomalía de la variable observada temperatura mínima y su contraparte aleatoria y, exponer la presencia de una estructura interna en los datos, determinismo. Además, estimar los valores de los parámetros correspondientes a su dinámica.

4.2.3. Experimentos Estación Atlanga P2: 1973-1987

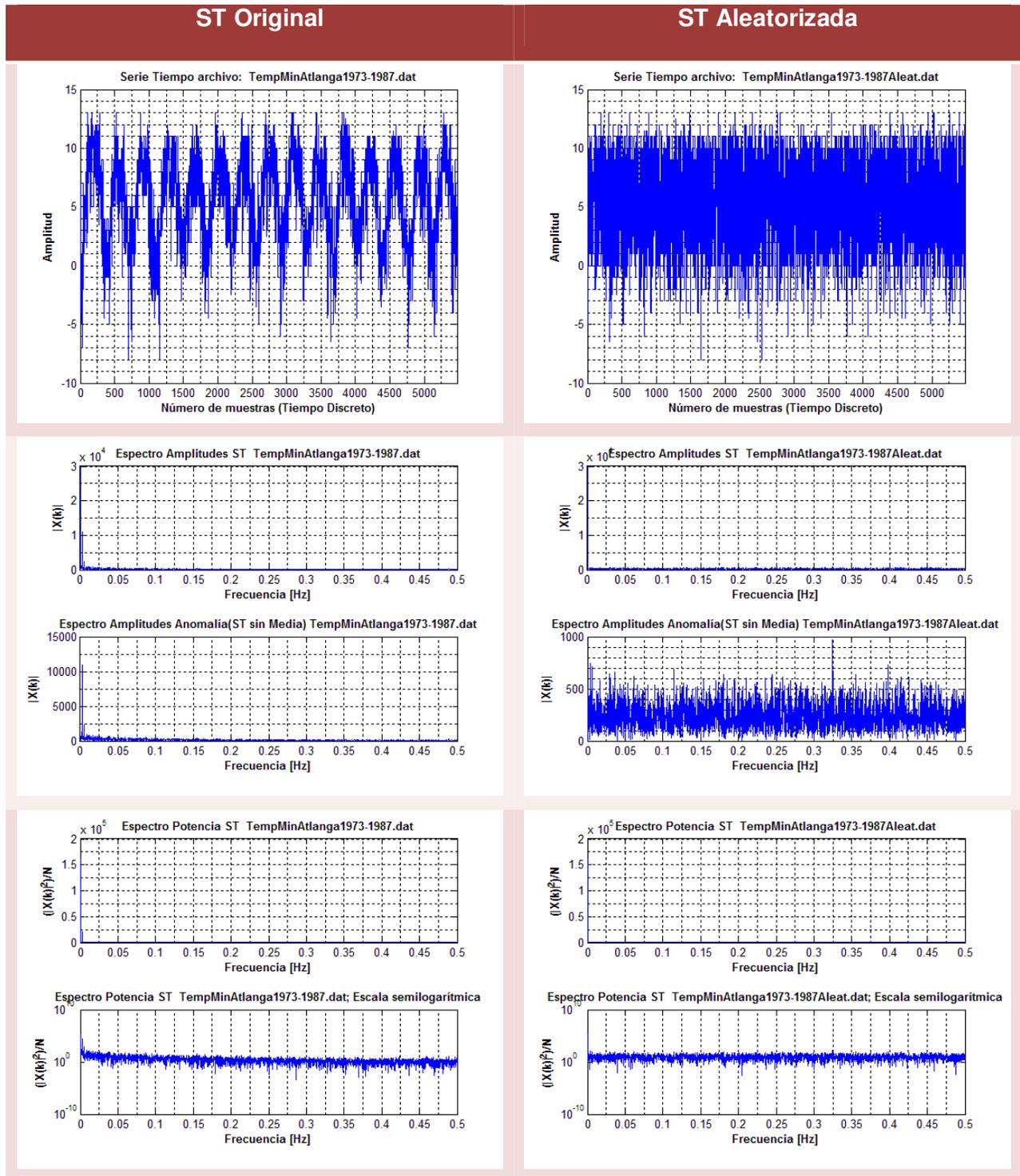


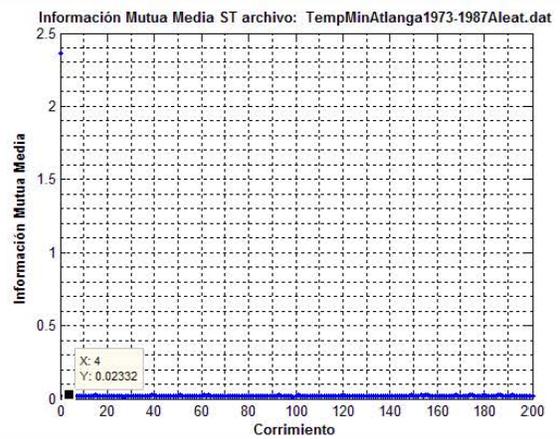
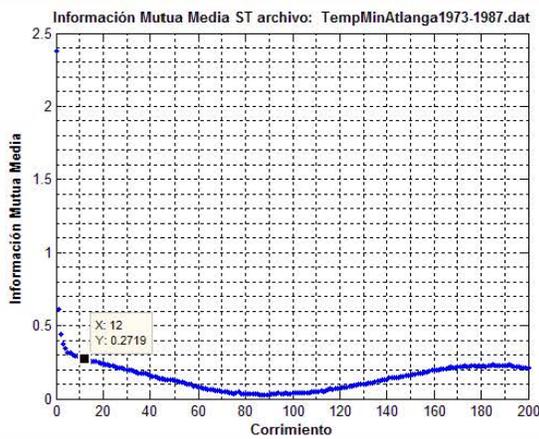
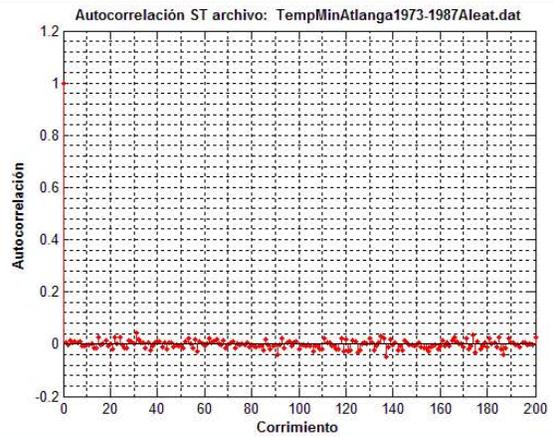
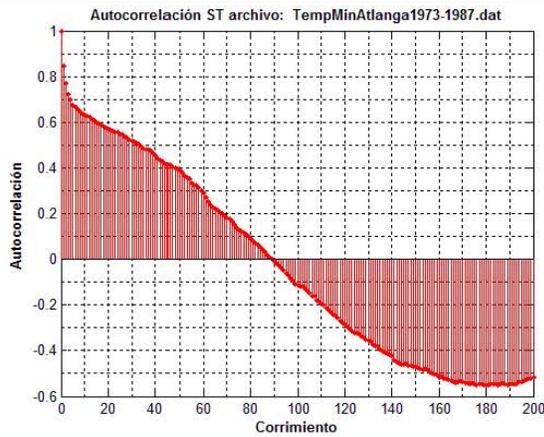
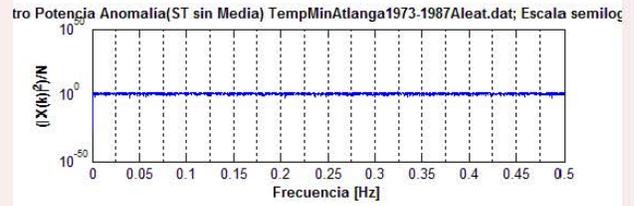
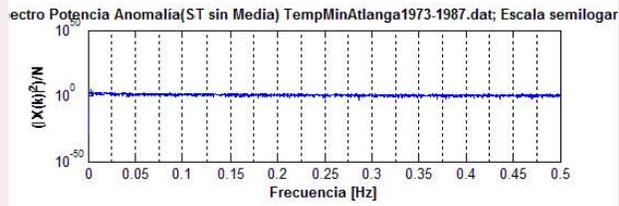
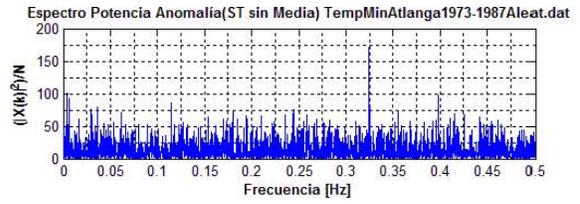
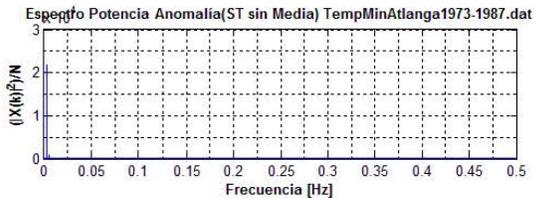


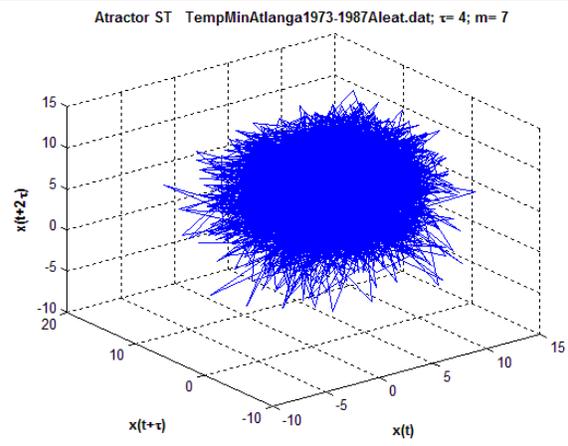
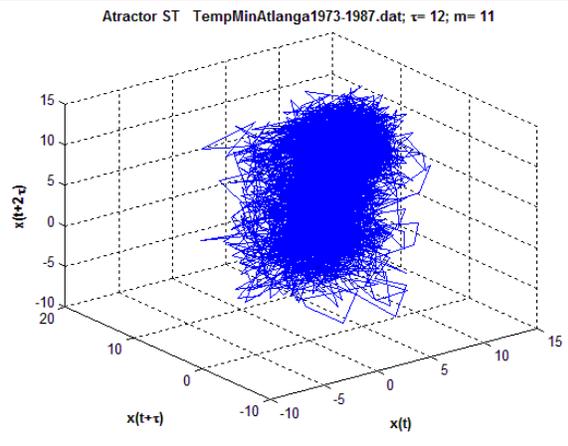
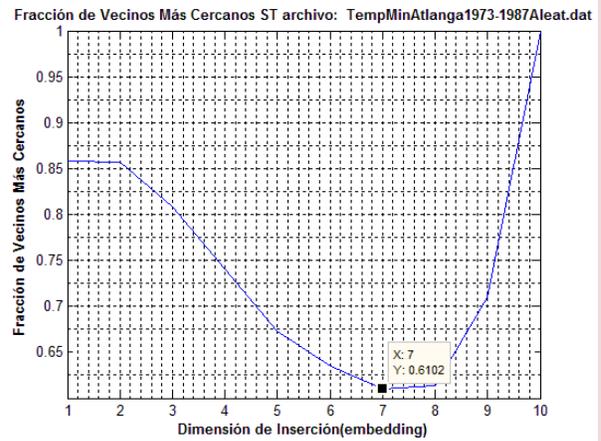
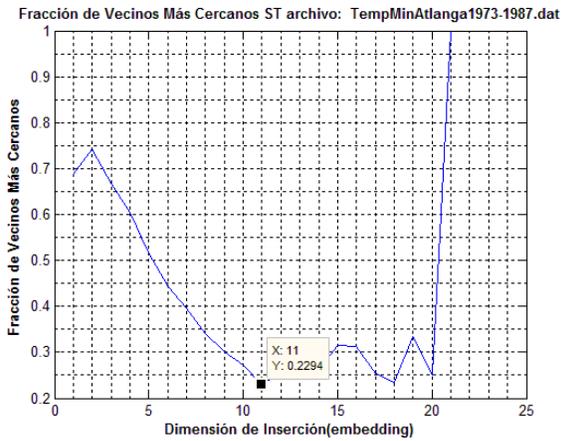


Experimento numérico para comparar el comportamiento de anomalías de las variables observadas precipitación y temperatura máxima, y estimar los valores de los parámetros correspondientes a su dinámica.

4.2.4. Experimentos Estación Atlanga P2: 1973-1987; Temperatura mínima



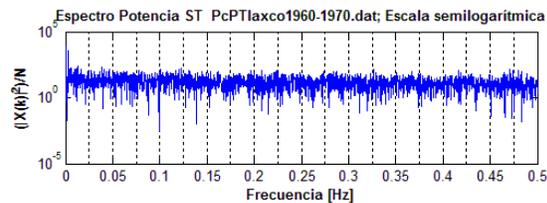
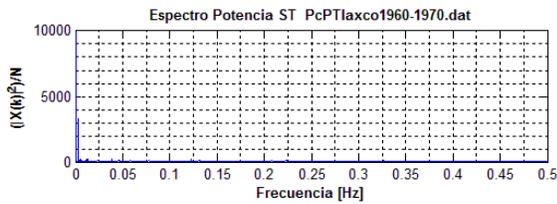
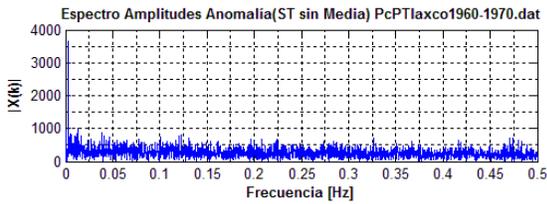
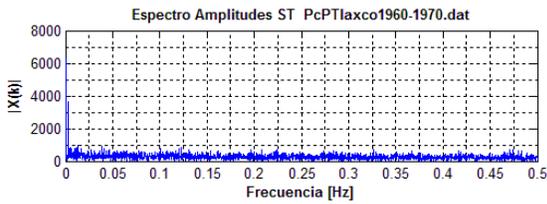
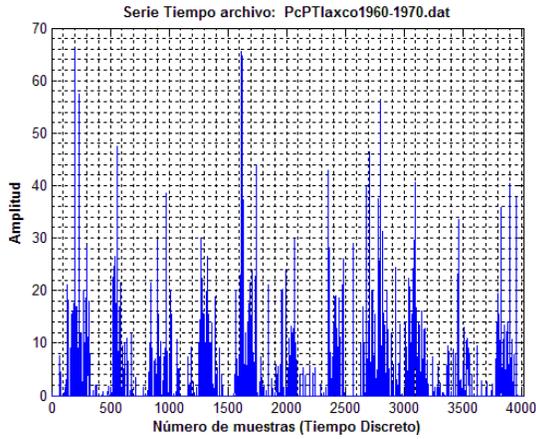




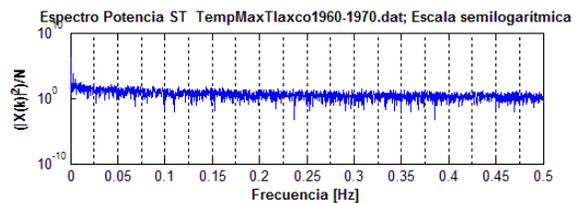
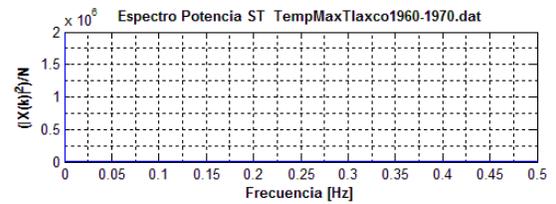
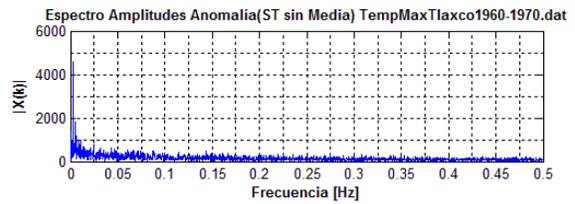
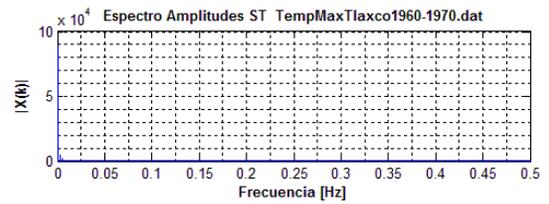
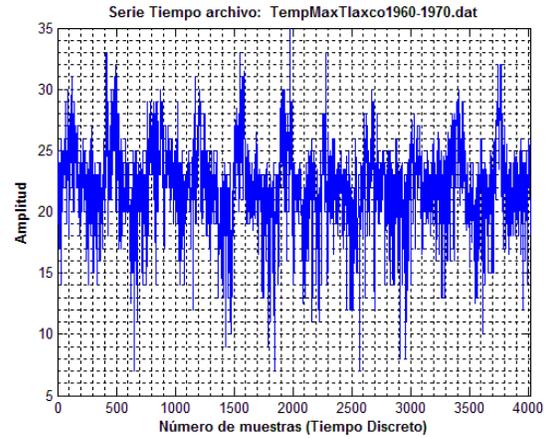
Experimento numérico para comparar el comportamiento de la anomalía de la variable observada temperatura mínima y su contraparte aleatoria y, exponer la presencia de una estructura interna en los datos, determinismo. Además, estimar los valores de los parámetros correspondientes a su dinámica.

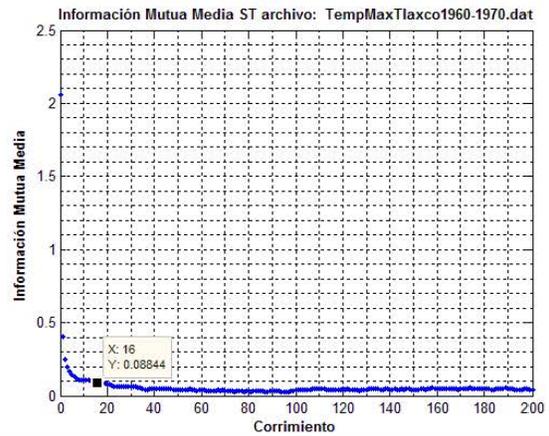
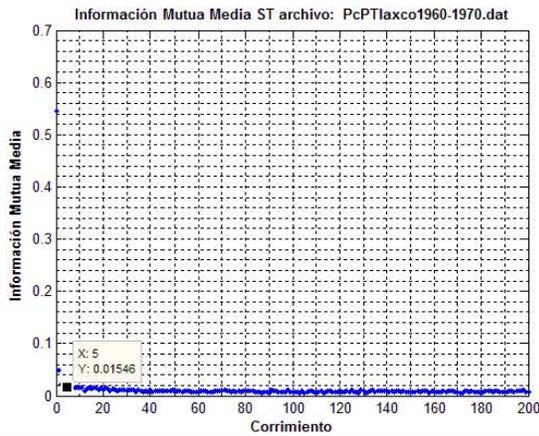
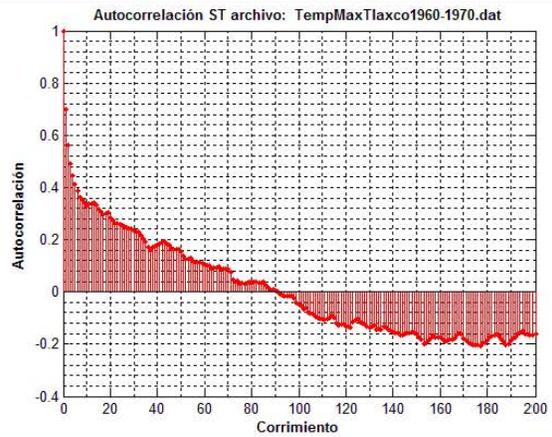
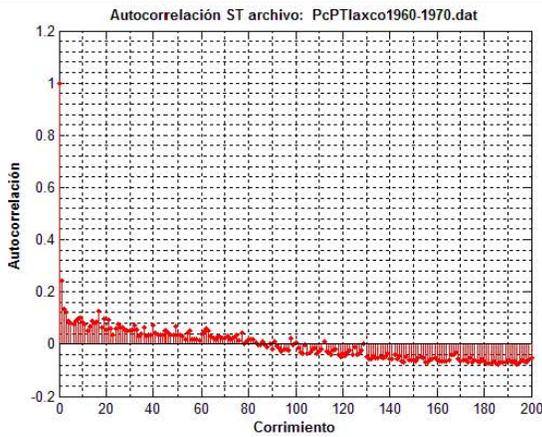
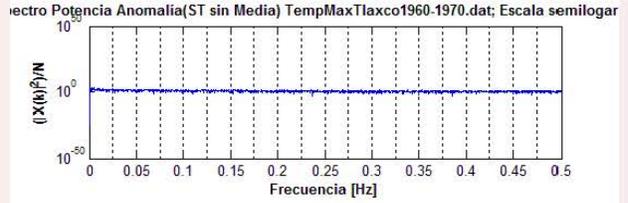
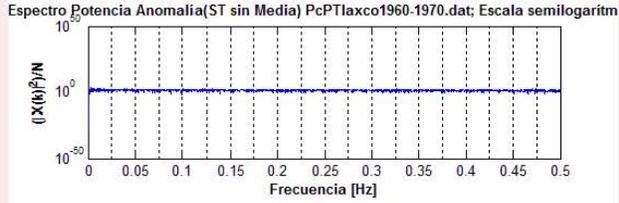
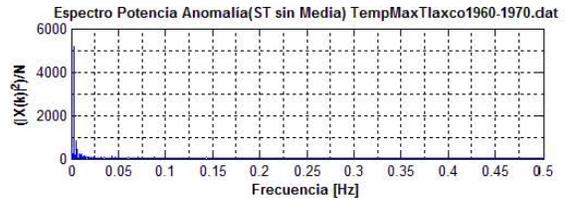
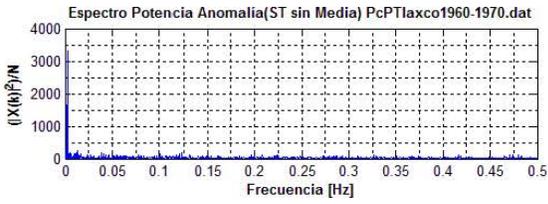
4.2.5. Experimentos Estación Tlaxco P1: 1960-1970

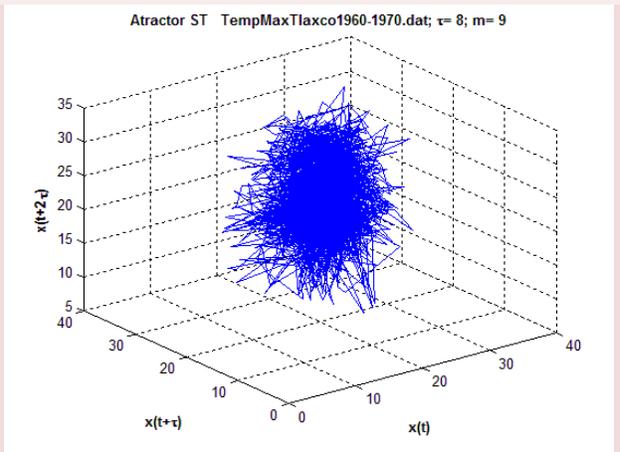
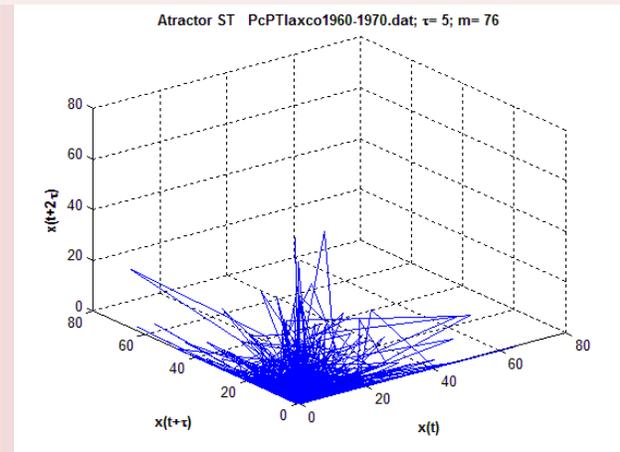
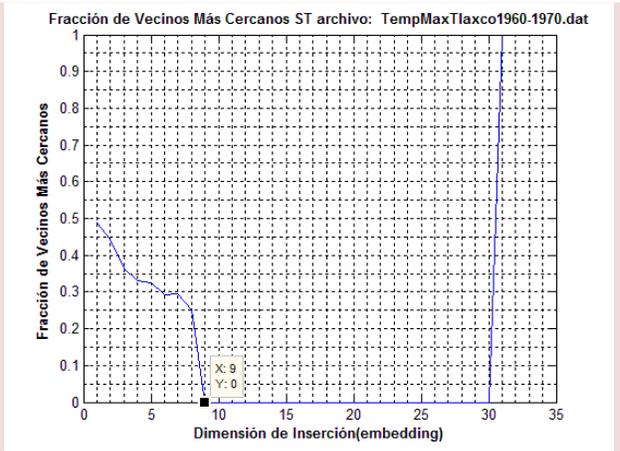
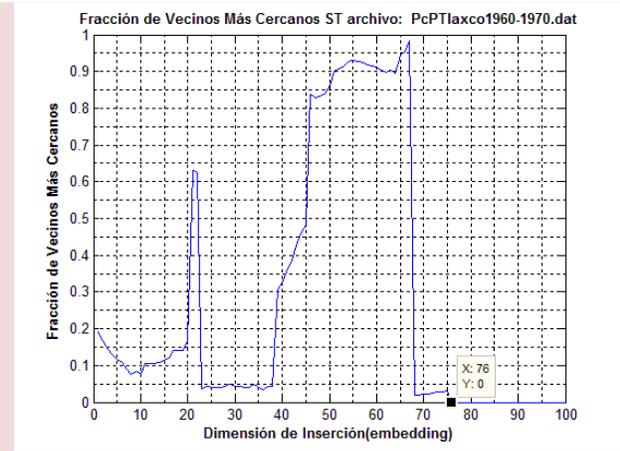
ST Precipitación



ST Temperatura Máxima

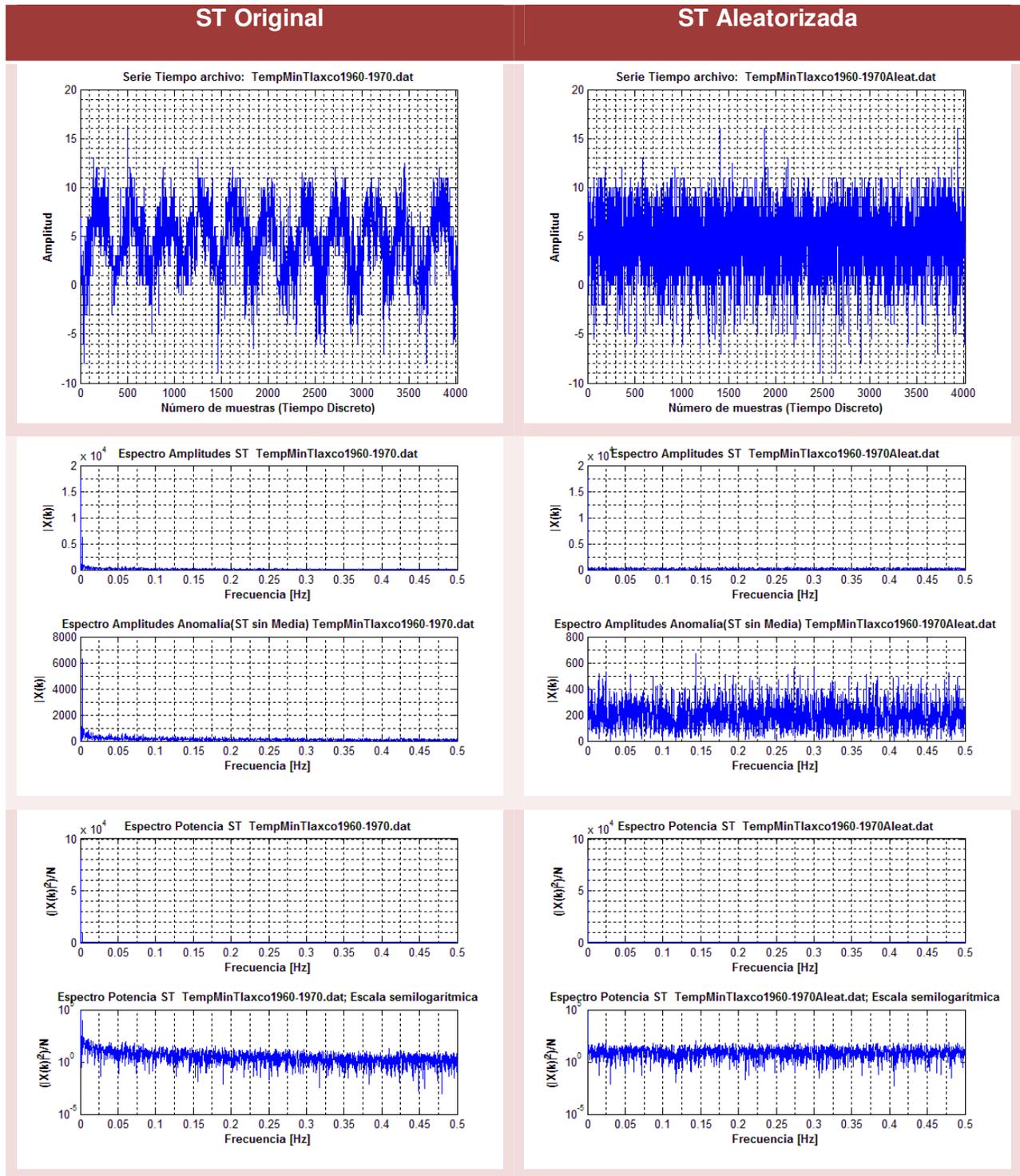


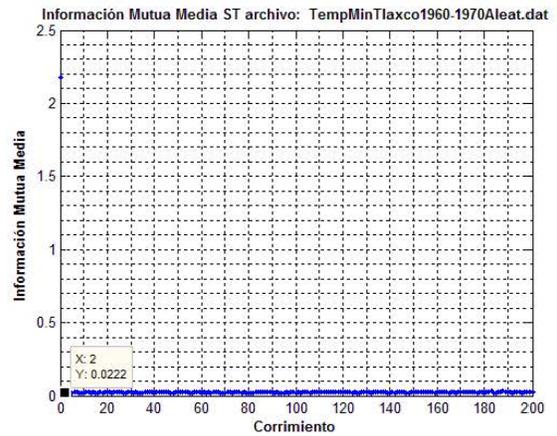
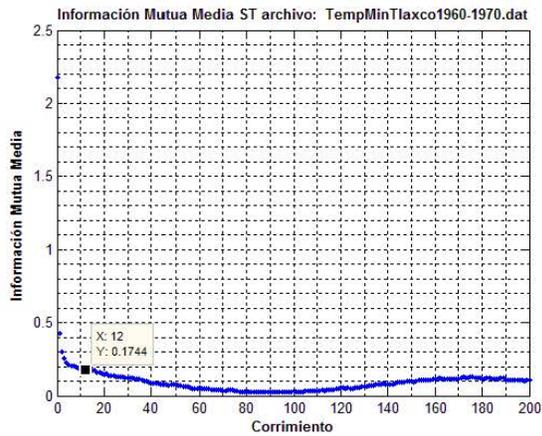
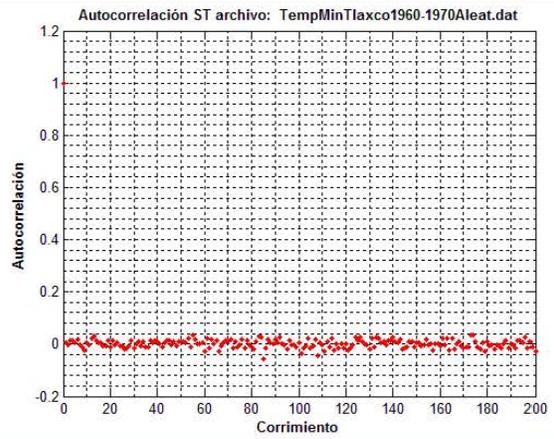
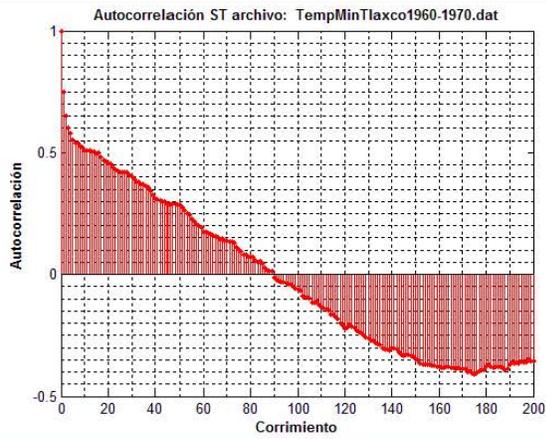
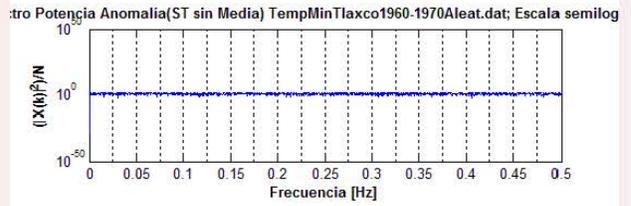
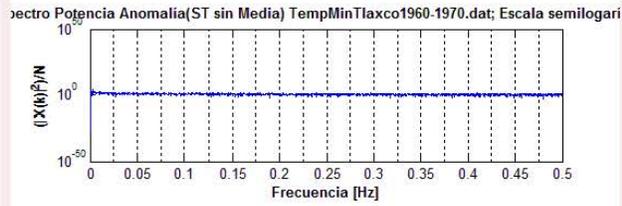
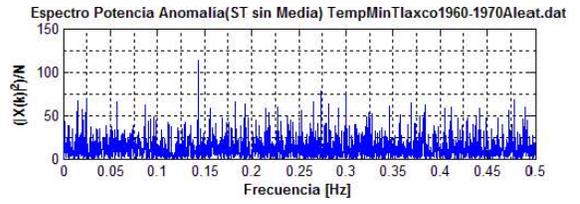
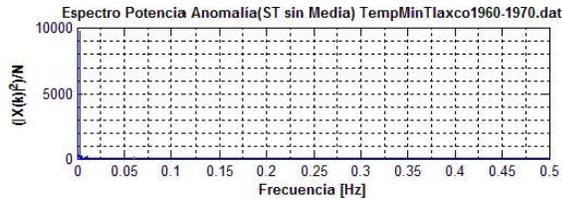


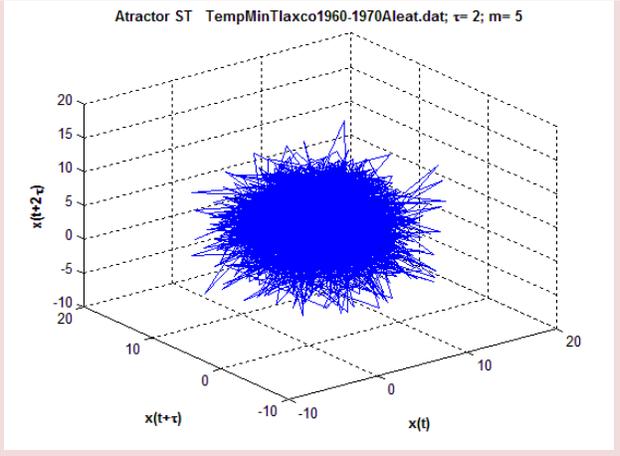
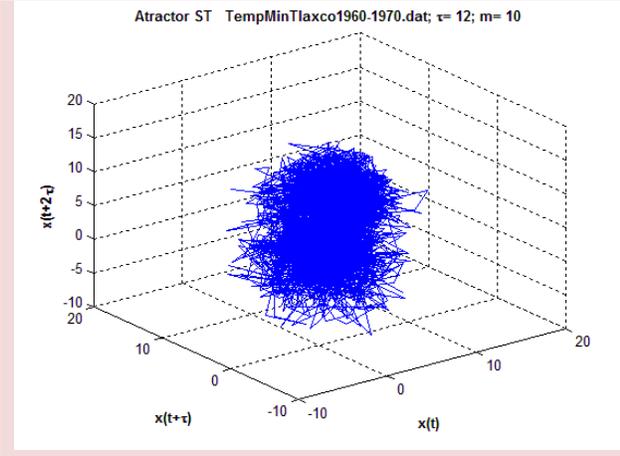
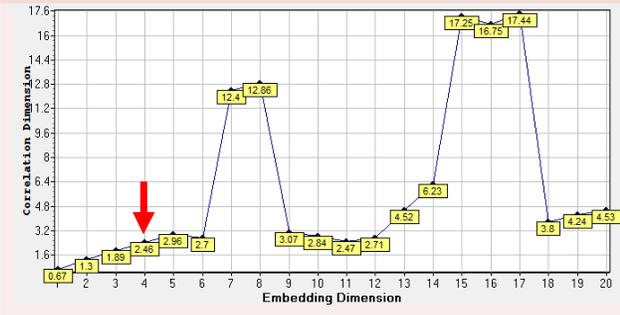
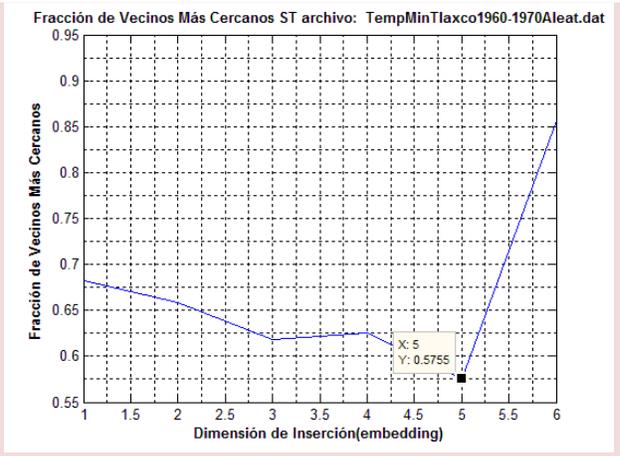
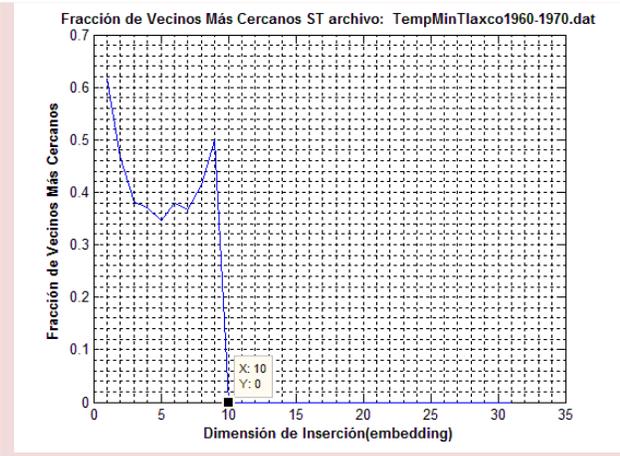


Experimento numérico para comparar el comportamiento de anomalías de las variables observadas precipitación y temperatura máxima y estimar los valores de los parámetros correspondientes a su dinámica.

4.2.6. Experimentos Estación Tlaxco P1: 1960-1970; Temperatura mínima





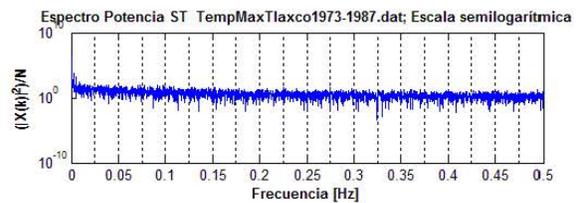
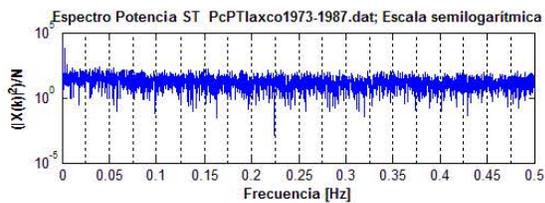
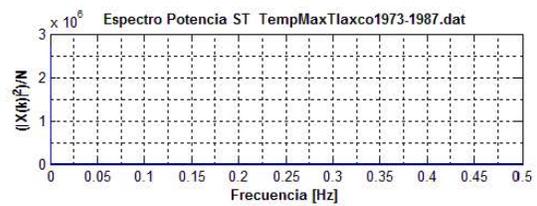
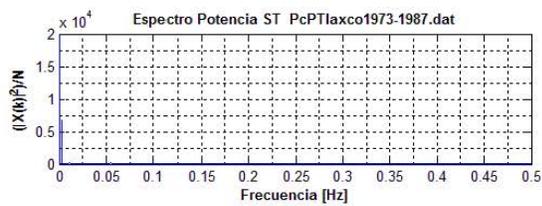
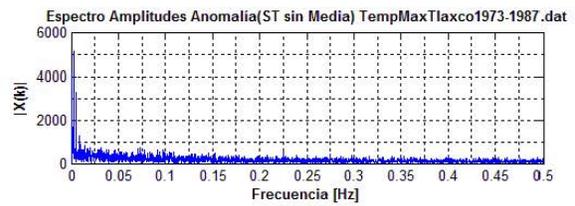
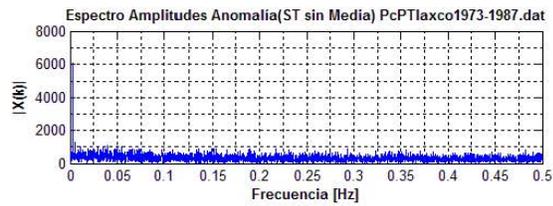
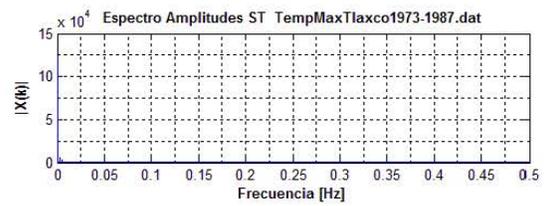
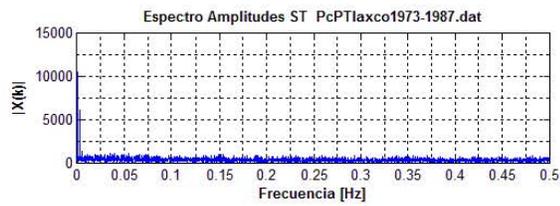
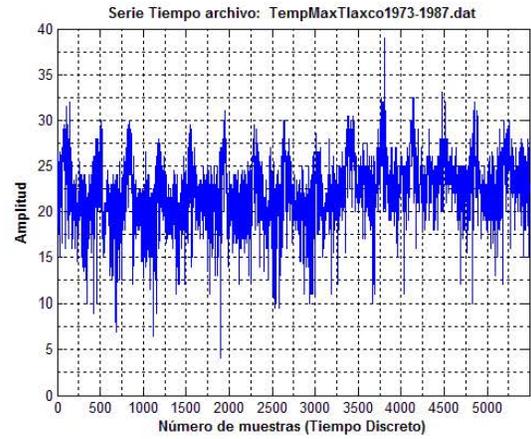
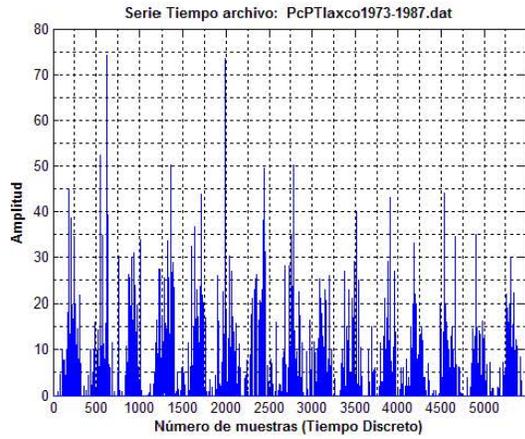


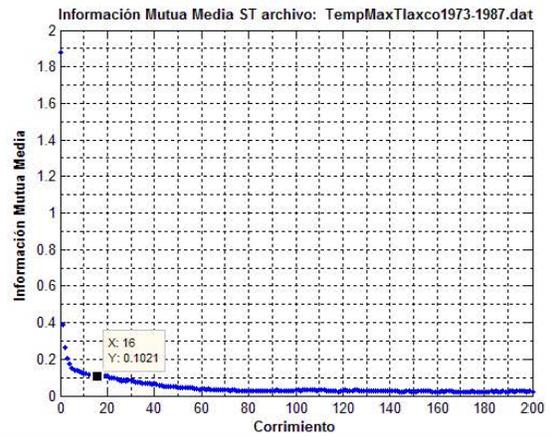
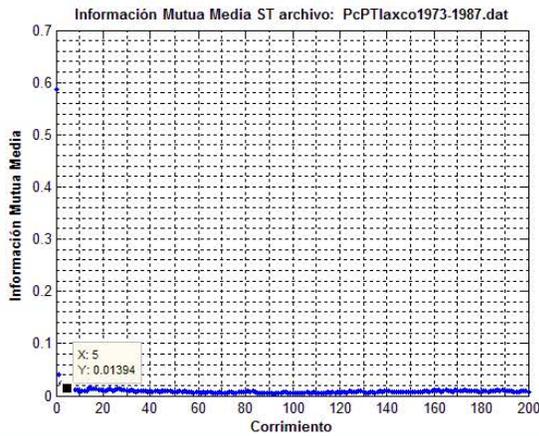
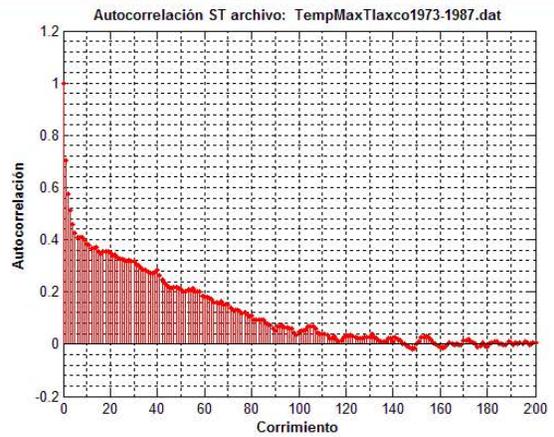
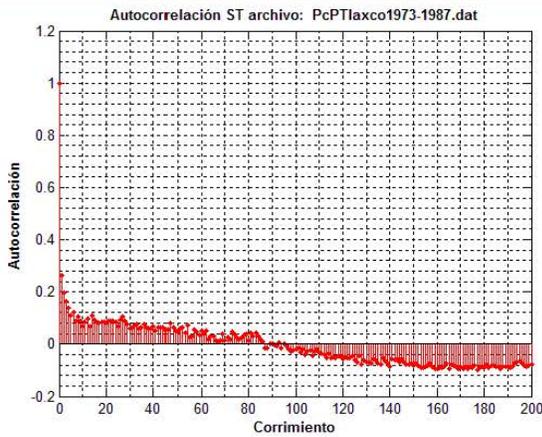
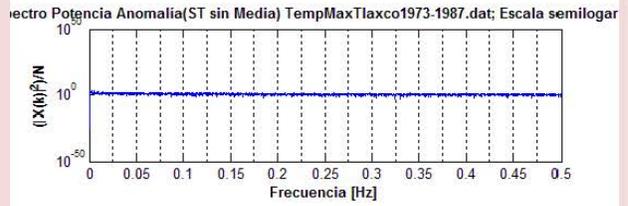
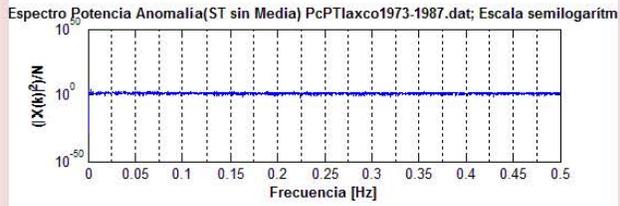
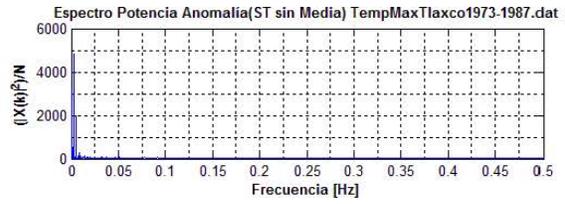
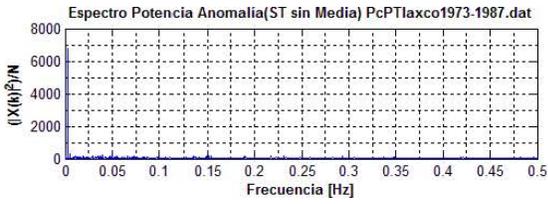
Experimento numérico para comparar el comportamiento de la anomalía de la variable observada temperatura mínima y su contraparte aleatoria y, exponer la presencia de una estructura interna en los datos, determinismo. Además, estimar los valores de los parámetros correspondientes a su dinámica.

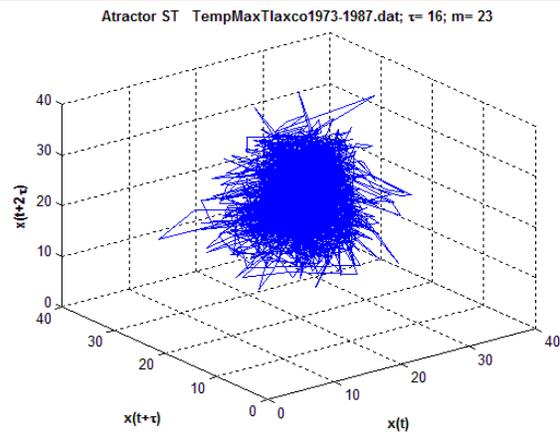
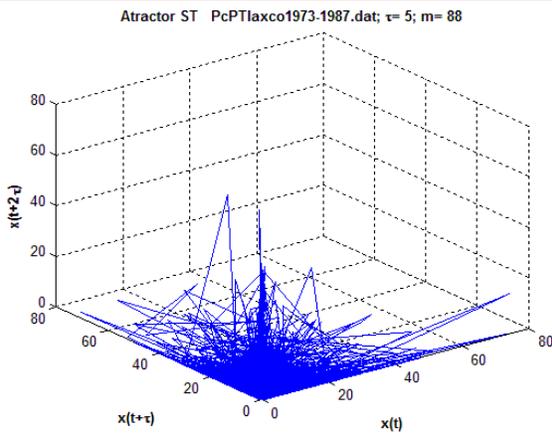
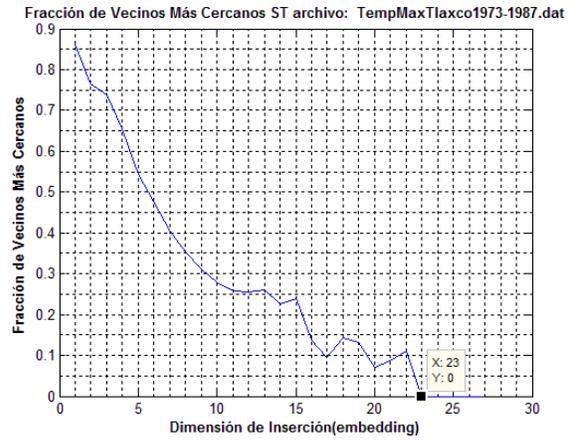
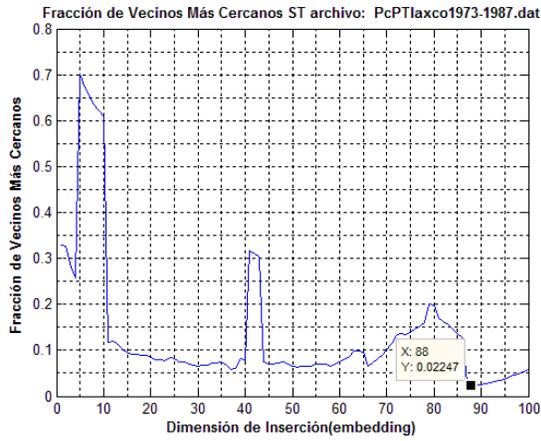
4.2.7. Experimentos Estación Tlaxco P2: 1973-1987

ST Precipitación

ST Temperatura Máxima

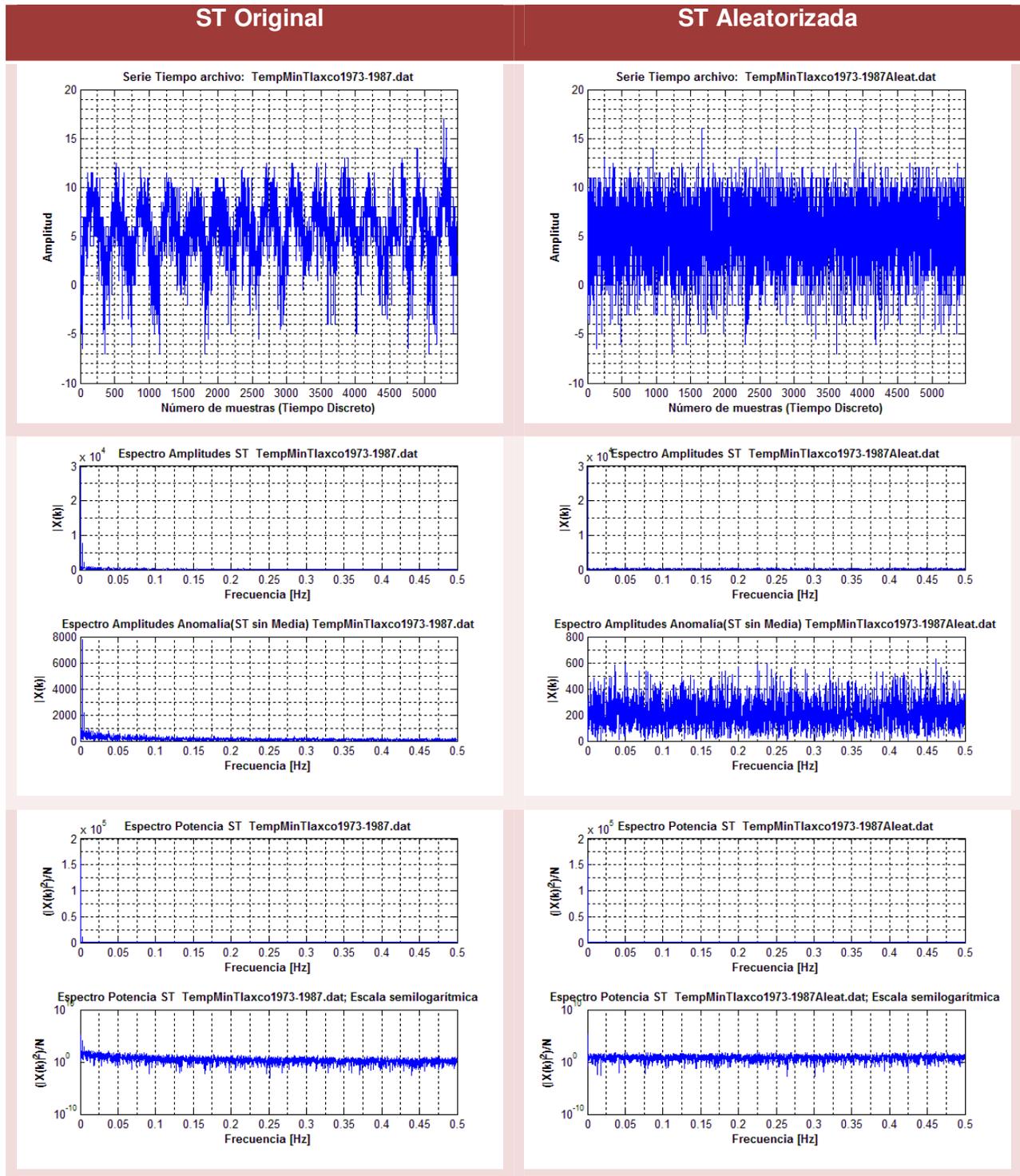


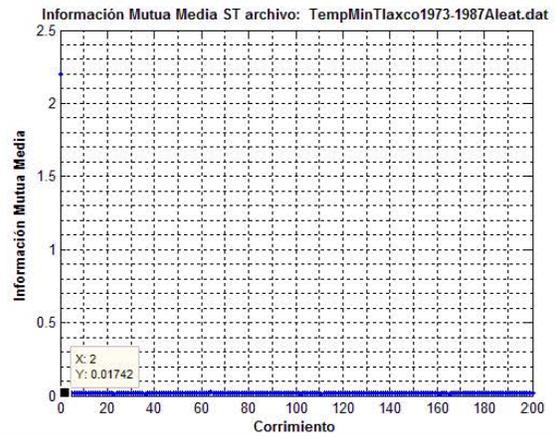
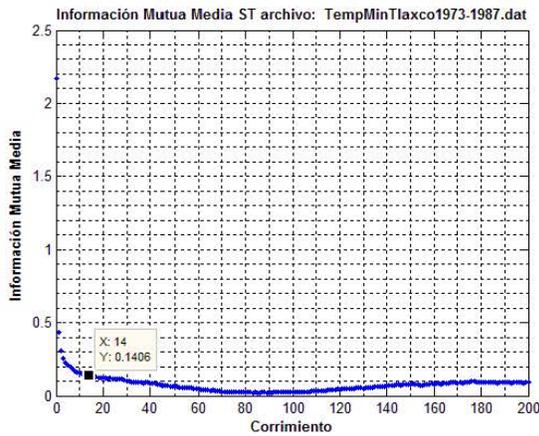
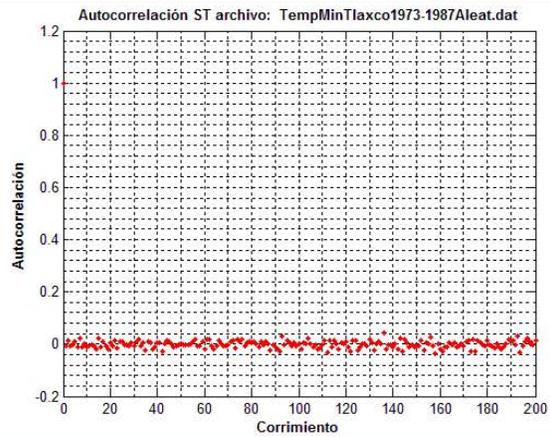
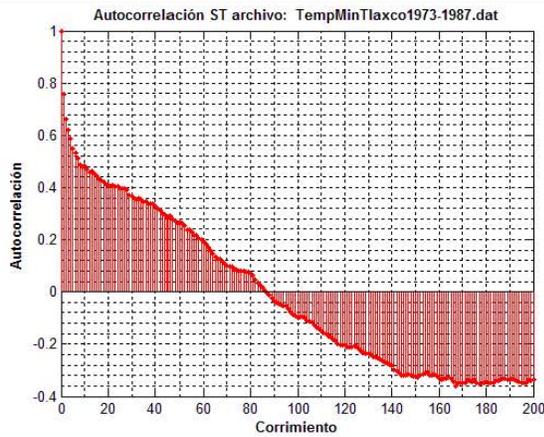
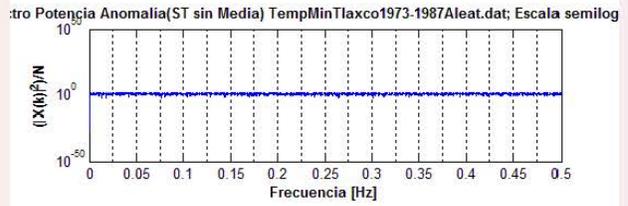
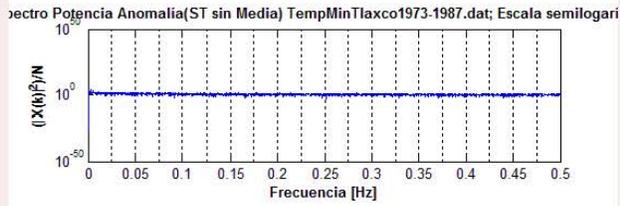
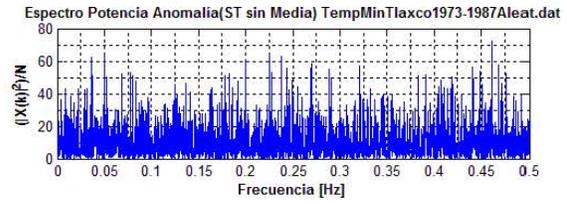
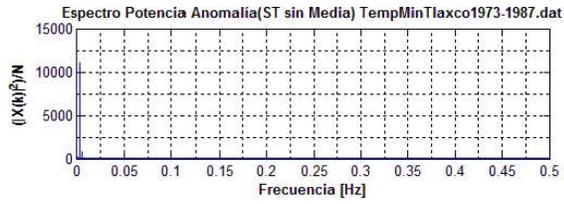


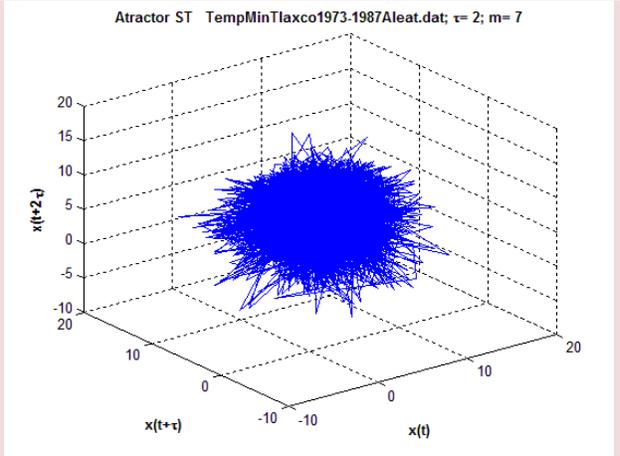
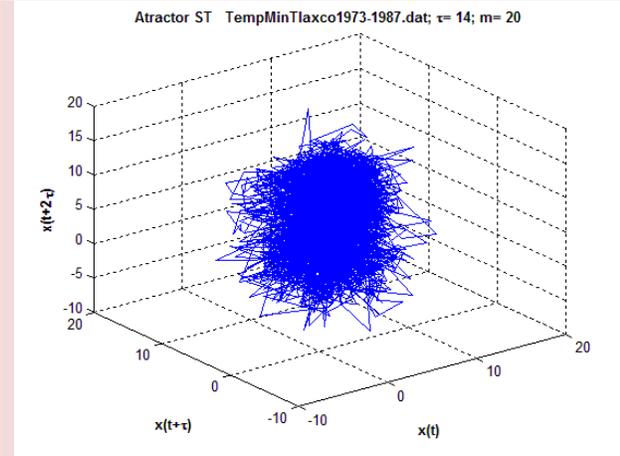
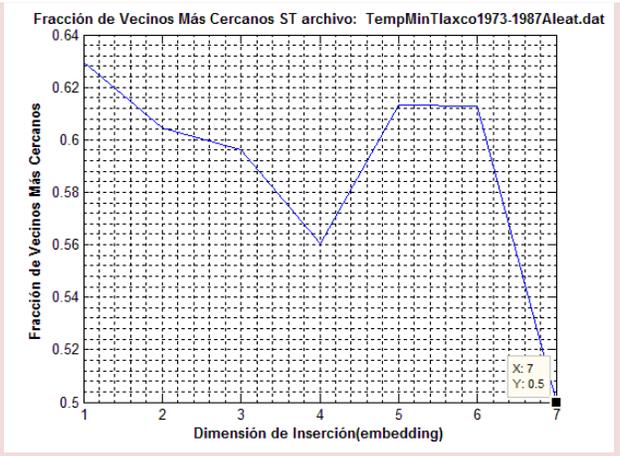
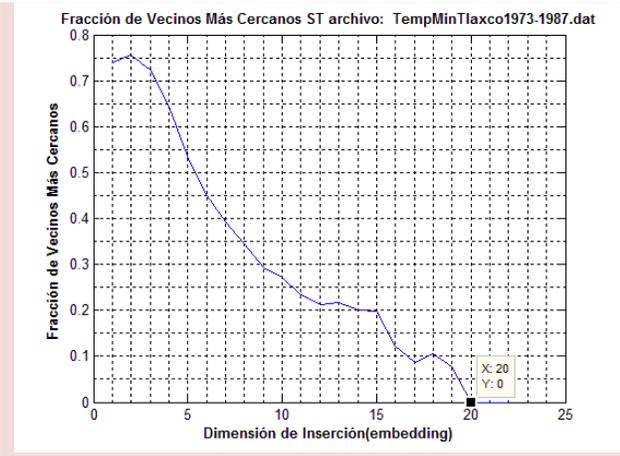


Experimento numérico para comparar el comportamiento de anomalías de las variables observadas precipitación y temperatura máxima, y estimar los valores de los parámetros correspondientes a su dinámica.

4.2.8. Experimentos Estación Tlaxco P2: 1973-1987; Temperatura mínima







Experimento numérico para comparar el comportamiento de la anomalía de la variable observada temperatura mínima y su contraparte aleatoria y, exponer la presencia de una estructura interna en los datos, determinismo. Además, estimar los valores de los parámetros correspondientes a su dinámica.

Capítulo 5. Análisis y Discusión de Resultados

5.1. Resultados

Como producto del proceso especificado en los dos capítulos anteriores, se obtuvieron tres tablas para los valores de los parámetros, los cuales fueron estimados mediante los diferentes métodos planteados, tanto para las series teóricas como para las observadas. En algunos casos, la fracción de vecinos más cercanos (**FVMC**) no decae a cero como lo señala la teoría, por lo que se consideró el valor de **m** para el que dicha fracción fuera **menor**.

5.1.1. ST teóricas: Modelo de Lorenz

ST	τ	m	D ₂
x Original	16	3	2.01
x Aleatorizada	5	ND; 4	8.92 - 8.95
y Original	15	3	2.12 - 2.16
y Aleatorizada	2	ND; 4	5.73 - 6.11
z Original	16	3	2.06 - 2.11
z Aleatorizada	4	ND; 5	7.33 - 7.55

Tabla 1. Parámetros estimados para las variables “x,y,z” del modelo de Lorenz. Las siglas **ND** hacen referencia al hecho de que la fracción de vecinos más cercanos (**FVMC**) no decae a cero, por lo que el valor **m** representado corresponde al valor para el cual dicha fracción es **menor**.

5.1.2. ST observadas diarias

Estación	Periodo	Anomalía ST	τ	m	D_2
Atlanga	P1	Pcp	5	ND; 30	No hay saturación
Atlanga	P1	Temp. Máx.	9	20	11.95 - 12.08
Atlanga	P1	Temp. Mín.	8	21	5.37 – 5.66
Atlanga	P1	Temp. Mín. Aleatorizada	2	ND; 6	2.61 - 3.06
Atlanga	P2	Pcp	6	80	No hay saturación
Atlanga	P2	Temp. Máx.	12	18	5.85 – 6.76
Atlanga	P2	Temp. Mín.	12	ND; 11	3.81
Atlanga	P2	Temp. Mín. Aleatorizada	4	ND; 7	5.49 – 6.23

Tabla 2. Parámetros estimados para las variables.

Estación	Periodo	Anomalía ST	τ	m	D_2
Tlaxco	P1	Pcp	5	76	No hay saturación
Tlaxco	P1	Temp. Máx.	8	9	5.86 - 5.95
Tlaxco	P1	Temp. Mín.	12	10	2.46 – 2.96
Tlaxco	P1	Temp. Mín. Aleatorizada	2	ND; 5	3.25 – 3.85
Tlaxco	P2	Pcp	5	88	No hay saturación
Tlaxco	P2	Temp. Máx.	16	23	6.77 – 6.99
Tlaxco	P2	Temp. Mín.	14	20	6.53 – 6.78
Tlaxco	P2	Temp. Mín. Aleatorizada	2	ND; 7	9.1 - 9.71

Tabla 3. Parámetros estimados para las variables.

5.2. Discusión de Resultados

5.2.1. ST teóricas: Modelo de Lorenz

Como era de esperarse, para las series teóricas-deterministas (variables “x,y,z”) se obtuvieron dimensiones no enteras, lo que indica que el sistema de ecuaciones de Lorenz tiene un límite de predictibilidad. Por otro lado, se observa que los comportamientos arrojados por los diferentes métodos aplicados a las tres ST, cambian significativamente de una variable a otra.

Así, para empezar, las gráficas de las tres ST lucen comportamientos intrincados y diferentes. Al comparar los espectros de amplitudes y potencia de la ST original y la anomalía se nota que éstos son muy parecidos en comportamiento, al respecto, los espectros de potencia presentan un decaimiento exponencial, así como una amplia gama de frecuencias. Respecto a la función de autocorrelación, puede observarse que para las variables “x” e “y”, ésta decae muy lentamente a cero, mientras que para la variable “z”, oscila de forma semejante a un movimiento amortiguado, por lo que decae a cero más rápidamente, sin embargo, para los tres casos se consideró el valor del corrimiento proporcionado por la función de información mutua media, ya que se comporta de forma muy similar para los tres, y lo más importante: los valores de τ son iguales a 16 para “x” y “z”, mientras que para “y” $\tau = 15$. Sobre la estimación de la dimensión contenedora para el atractor, en los tres casos, la fracción de vecinos más cercanos (**FVMC**) decayó a cero en el valor de **m** correspondiente a 3. En cuanto a la dimensión de correlación (D_2), para las tres variables se obtuvieron los valores: 2.01 para “x”, de 2.12 a 2.16 para “y” y 2.06 a 2.11 para “z”; sobre lo que debe resaltarse que si se redondea en todos los casos al entero mayor, se obtiene el valor de 3, lo cual hace referencia al hecho de que para modelar la dinámica del sistema se requiere de 3 ecuaciones o en su defecto, de 3 variables independientes, así, se comprueba que el atractor reconstruido proviene en efecto, de un sistema de 3 ecuaciones diferenciales ordinarias, por lo que obviamente se cumple la hipótesis central de la TSDNL, como era de imaginarse. Por último, la forma del atractor es la misma para las variables “x” e “y”, mientras que varía para “z”.

Respecto a las tres ST aleatorizadas, puede decirse que dan la apariencia de ser ruido aleatorio, ya que sus gráficas en el dominio del tiempo no presentan alguna estructura particular, y sus respectivos espectros presentan una amplia gama de frecuencias. En cuanto a la función de autocorrelación, en todos los casos la gráfica de ésta se observa casi sin puntos extremos que sean verdaderamente significativos, asemejándose más bien a una línea horizontal, lo cual sucede también para la función de información mutua media. En lo que corresponde a la estimación del valor de m , en ninguno de los casos la **FVMC** decayó a cero. Dadas las situaciones anteriores, se consideraron los valores de los parámetros que más se acercaran a los criterios estándar, es decir, para τ , se consideró el primer valor más pequeño (el primer mínimo de la función de información mutua media) y para m se usó el valor para el que la **FVMC** fuera **menor**, aunque no decayera a cero.

También, se calculó la dimensión de correlación (D_2) para los 3 casos de las ST aleatorizadas, y puede notarse que éstas son mayores que las respectivas dimensiones contenedoras m , lo cual en primera instancia es contradictorio con la teoría (Teorema de Takens). Por último, es posible apreciar que los tres atractores coinciden en que no poseen estructura definida alguna, a juzgar por sus respectivas gráficas, lo que sí logra notarse es que tienden a llenar la mayor cantidad de “espacio geométrico” posible.

5.2.2. ST observadas: Estación Atlanga

En este caso, las 3 variables presentan comportamientos variados, a pesar de esto, es posible afirmar que en general, los espectros de Fourier de las 3 anomalías usadas presentan una amplia gama de frecuencias, por lo que no es posible determinar pocas frecuencias dominantes que caractericen al sistema, hablando de los dos periodos estudiados.

Respecto a la función de autocorrelación, se observa que en los 3 casos y en ambos periodos considerados, se aprecia un comportamiento similar en cuanto a la forma que describen las gráficas, ya que éstas oscilan hasta llegar a cero y después se mantienen oscilando nuevamente pero sólo tomando números negativos, mientras que para la función de información mutua media, los

cambios que éstas describen son demasiado suaves, por lo que la identificación del primer mínimo se efectuó cuidadosamente; además, también se aprecia que las formas de las gráficas correspondientes a la misma anomalía, son similares, aunque diferentes al realizar la comparación entre anomalías del mismo periodo, es decir, entre precipitación, temperatura máxima y temperatura mínima, al igual que pasa con las ST teóricas mostradas anteriormente. En cuanto al valor del corrimiento τ , se observa que para el periodo P1 éste no es el mismo para las tres anomalías, sin embargo, para la TempMax y la TempMin son muy cercanos (9 y 8 respectivamente), mientras que el correspondiente a la Pcp es menor. Algo similar ocurre para los respectivos valores de m (20 y 21) correspondientes a las anomalías de TempMax y TempMin, además, se encontró que para la Pcp la **FVMC** no decayó a cero, por lo que se consideró un valor de 30. En lo que se refiere a la dimensión de correlación, se tiene una diferencia más marcada (mayor) entre las anomalías de temperatura, respecto los otros dos parámetros, y resalta el hecho de que la dimensión estimada para el atractor de la TempMax es mayor que la dimensión estimada para el atractor de la TempMin. El caso de la **precipitación** es especial, ya que la **dimensión del atractor no pudo obtenerse para ningún periodo ni estación**, esto debido a que, como puede apreciarse en todas las gráficas, no se produce saturación para algún valor particular, en otras palabras, el algoritmo usado no converge para el caso mencionado.

En cuanto al P2 para esta misma estación, ocurre algo similar a lo ya mencionado, teniéndose que en este caso, los valores de τ para la TempMax y la TempMin son iguales a 12 y para m son 18 y 11 respectivamente. Por su parte, las dimensiones de correlación para las anomalías de temperatura (5.85–6.76 y 3.81) no son tan dispares. Además, resalta la cuestión de que para este periodo la **FVMC** correspondiente a la Pcp sí decayó a cero en un valor bastante grande, $m = 80$.

También es evidente que las formas de los atractores se muestran diferentes entre anomalías, pero muy similares al realizar la comparación entre el mismo tipo de anomalía y diferente periodo.

Para el caso de las anomalías de TempMin aleatorizadas de ambos periodos, éstas presentan el mismo comportamiento que las ST teóricas aleatorizadas, por lo que sólo resta mencionar que la dimensión de correlación es menor que la dimensión contenedora m , como menciona la teoría para los dos casos.

5.2.3. ST observadas: Estación Tlaxco

Al igual que para la estación Atlanga, las 3 anomalías presentan comportamientos variados, a pesar de esto, también es posible afirmar que en general, los espectros de Fourier de las 3 anomalías usadas presentan una amplia gama de frecuencias, por lo que no es posible determinar pocas frecuencias dominantes que caractericen al sistema, hablando de ambos periodos.

Respecto a las funciones de autocorrelación, y de información mutua media, los comportamientos también corresponden de forma similar a lo descrito para la estación Atlanga.

Los valores estimados de los parámetros quedaron de la siguiente forma:

Para el P1, los valores de τ para las anomalías fueron 5, 8 y 12; los de m 76, 9 y 10 (Pcp, TempMax y TempMin respectivamente). En cuanto a las dimensiones de correlación, ésta resultó mayor para el atractor de la TempMax (5.86-5.95) que para el atractor de la TempMin (2.46–2.96).

Análogamente, para el P2, los valores de τ para las anomalías fueron 5,16 y 14; los de m 88, 23 y 20 (Pcp, TempMax y TempMin respectivamente). En cuanto a las dimensiones de correlación, ésta resultó ser ligeramente mayor para el atractor de la TempMax (6.77–6.99) que para el atractor de la TempMin (6.53–6.78), sin embargo, en términos prácticos, si se redondea al entero mayor (7 en otras palabras), dicha diferencia es nula.

También es evidente que las formas de los atractores se muestran diferentes entre anomalías, pero muy similares al realizar la comparación entre el mismo tipo de anomalía, diferente periodo y estación.

Para el caso de las anomalías de TempMin aleatorizadas de ambos periodos, éstas presentan también el mismo comportamiento que las ST teóricas

aleatorizadas, aunque para el P2, la D_2 es mayor que la dimensión contenedora m , tal y como sucede con las ST teóricas que se aleatorizaron.

5.3. Interpretación global

De acuerdo a los resultados obtenidos, es posible distinguir varias cuestiones interesantes:

Al aleatorizar las ST teóricas-deterministas, se obtuvieron ST que representan el extremo opuesto a las primeras, con lo cual puede establecerse la naturaleza de las ST observadas al aplicarles los métodos de análisis no lineal: éstas ST presentan un comportamiento intermedio entre el determinismo y la aleatoriedad, ya que por un lado, se comprobó que la dinámica del sistema describe una cierta estructura particular (descrita por el atractor) y por otro, las mediciones siempre poseen algún grado de ruido, ya que no es posible en la práctica removerlo por completo.

Los parámetros estimados para las anomalías de las ST teóricas, no son idénticos para el caso del corrimiento τ y la D_2 , sin embargo, para éste último la diferencia se vuelve despreciable al considerar lo mencionado en la sección anterior, además, para el primer parámetro la diferencia no es muy grande (una unidad).

En los tres casos, las D_2 correspondientes a las ST teóricas aleatorizadas son mayores a sus correspondientes valores de m , lo cual contradice la teoría, no obstante, esto es explicable en los términos siguientes: como se mencionó en su momento, la **FVMC** no decae a cero en ninguno de los casos, por lo que se consideró el valor de m para el que dicha fracción fuera menor; la razón por la que la **FVMC** no toma nunca un valor de cero es que la ST en cuestión no posee una estructura interna definida, porque posee una naturaleza pseudoaleatoria, por lo que cuando el algoritmo busca vecinos cercanos en dimensiones más altas, nunca permanecen los mismos puntos como vecinos al pasar de una dimensión a otra.

En cuanto a los datos observados para ambas estaciones climatológicas, se encontró que al igual que las ST teóricas, las anomalías de las ST observadas,

presentan la cuestión de que los respectivos parámetros estimados no son idénticos, aunque se observa que son muy parecidos, como sucede para las series teóricas. Por un lado, esto se explica mediante el hecho de que el modelo de Lorenz al ser teórico, es un modelo simplificado, lo que implica que no incorpora procesos adicionales que interactúan e influyen constantemente en la dinámica del sistema, tal y como sucede en la realidad, ésta es la razón por la que difieren los valores de los parámetros estimados de una anomalía respecto a otra, a pesar de corresponder a la misma región.

Por otro lado, de dichos parámetros estimados, el más importante desde un punto de vista práctico es la dimensión de correlación, ya que es posible atribuirle una explicación física a su valor: el hecho de que la D_2 del atractor de la TempMax sea mayor a la D_2 del atractor de la TempMin se debe a que la primera refleja una mayor cantidad de procesos activos, dado que durante el día, la luz solar proporciona la energía necesaria para que los procesos atmosféricos se encuentren en mayor movimiento y constante cambio, mientras que la segunda implica que la atmósfera se encuentra en mayor reposo, lo que se traduce en una cantidad menor de procesos activados.

Respecto al hecho de que no haya sido posible determinar la D_2 para ninguna de las anomalías observadas de precipitación, la explicación es que la Pcp es una variable complicada de tratar, ya que la existencia de una gran cantidad de ceros en los registros, describe cambios abruptos y/o repentinos en la ST observada, en otras palabras, esto quiere decir que la intermitencia está presente en el sistema. Dentro del contexto de los sistemas dinámicos, la intermitencia se refiere a pasar de comportamientos periódicos a comportamientos irregulares y viceversa, además de que ésta puede ser abrupta o modulada.

Capítulo 6. Conclusiones y Trabajos a Futuro

6.1. Conclusiones

Este trabajo expone de forma clara y concreta los métodos y las herramientas elementales usadas en el Análisis No Lineal de Series de Tiempo (ANST), además, aplica dichos conceptos tanto a series teóricas como observadas.

El enfoque de Minería de Datos aplicado en este trabajo corresponde al hecho de que en la **MD**, se **coleccionan** los **datos** y se espera **extraer hipótesis** de **ellos**, en este caso, una de las hipótesis generadas es: afirmar la posibilidad de incorporar el ANST de forma complementaria, en los esquemas de asimilación de datos¹¹ atmosféricos usados en conjunto con un modelo de pronóstico, los cuales conforman a su vez los sistemas de pronóstico. Por otra parte, es posible afirmar que cada una de las anomalías observadas nos describe en cierto sentido su propia naturaleza, ya que por ejemplo, al no poder determinar la D_2 para ninguno de los atractores de las anomalías de Pcp observadas, se está exaltando el hecho de que el comportamiento de la variable precipitación es bastante complejo y depende de variados factores que interactúan de forma no lineal, lo cual da cuenta del hecho de que en los propios modelos de predicción numérica del tiempo, la lluvia no ha podido ser modelada o incluso comprendida del todo, en consecuencia, debe aplicársele también un tratamiento previo como la descomposición en pentadas, por ejemplo, para reducir el efecto de la intermitencia, al tratar de usar los métodos de ANST para su caracterización.

Respecto a las variables restantes, como se observó, tanto la temperatura máxima como la mínima ofrecen una respuesta bastante aceptable para ser caracterizadas por los métodos de ANST. Por otra parte, otra hipótesis generada como producto del enfoque exploratorio de MD considerado sobre los datos es: como se mencionó, las D_2 de los atractores de la TempMax fueron

¹¹El término asimilación de datos se refiere al proceso por medio del cual se combina información meteorológica como datos observados, un modelo de pronóstico, un modelo de observaciones e información *a priori* con el fin de estimar el estado más probable y consistente de la atmósfera, para después efectuar un ciclo de pronóstico.

mayores que las correspondientes para los atractores de la TemMin, lo cual como se discutió se explica en términos de la cantidad de procesos activos que contribuyen al cambio de las variables que describen a la atmósfera, esto nos lleva a la cuestión de que entonces, si se quiere mejorar los pronósticos del tiempo específicamente, sería necesario incluso, considerar el uso de diferentes modelos numéricos para las diferentes horas del día (para diferentes escalas temporales), de acuerdo a la cantidad de variables o ecuaciones requeridas para modelar la dinámica de la región estudiada.

Como quedó expuesto en capítulo 2, los métodos de ANST poseen similitudes con sus contrapartes lineales, algo que no se trató en el presente trabajo es la reducción de ruido en los datos, debido en parte a la extensión de la presente y por otra parte, porque deseaba estudiarse el desempeño de los métodos de ANST con datos a los que sólo se les aplicó un control de calidad convencional, de acuerdo a los estándares internacionales.

Como puede apreciarse, los métodos de ANST tienen bastante potencial, sólo requieren ser más estudiados en casos prácticos que impliquen ST observadas.

6.2. Aportaciones

El presente trabajo contribuye a la difusión de los métodos de ANST mediante una exposición clara de los conceptos teóricos más básicos en los que se fundamenta dicho análisis. A su vez, expone la pertinencia de usar dichos métodos, haciendo énfasis tanto en el grado de linealidad como el de determinismo de las ST utilizadas.

Para las ST teóricas (modelo de Lorenz) se demostró que los parámetros más elementales estimados mediante los métodos de ANST, no son idénticos, a pesar de que las 3 ST deberían de presentar la misma dinámica, dado el hecho de que provienen de un modelo teórico-determinista. Lo mismo quedó demostrado para las anomalías observadas, sin embargo, en este caso era más factible que éste fuera el caso, debido a su naturaleza.

Se concluyó que las variables que mejor respondieron a los métodos aplicados fueron la temperatura máxima y la temperatura mínima, sin embargo, es

posible afirmar que de las dos, la segunda ofrece una mejor respuesta para ser caracterizada por dichos métodos.

6.3. Trabajos a Futuro

A pesar de que los métodos de análisis de series de tiempo aún se encuentran en una fase de experimentación y refinamiento, están cobrando bastante importancia en la caracterización de series de tiempo, las cuales reflejan la complejidad del sistema al que representan. Así, queda mucho camino por recorrer en las ciencias de la atmósfera y en general en el ámbito científico, al respecto del potencial que tienen los métodos de ANST. Por esta razón sería muy recomendable efectuar el análisis de los mismos datos aplicando muchos más algoritmos correspondientes al análisis no lineal, tales como la reducción no lineal de ruido, la creación de *surrogates*¹² de las ST, el cálculo del espectro de Lyapunov, el cálculo de entropías, el uso de los gráficos de recurrencia (RP) así como la predicción no lineal; con lo que se caracterizarían las ST por completo desde el enfoque no lineal.

Por otro lado, sería interesante incluir en el análisis todas las contrapartes lineales de los métodos y así efectuar una comparación total de las respectivas técnicas, con el fin de establecer los alcances y las limitaciones de las mismas. También, sería muy benéfico complementar el análisis de la base de datos mediante las técnicas de Minería de Datos, con el fin de encontrar patrones, establecer correlaciones y modelos descriptivos/predictivos de las observaciones disponibles, con los cuales pudieran establecerse teleconexiones entre las estaciones y posiblemente también entre datos extraídos de boyas marinas.

¹² En español, dicho término aún no ha sido traducido de manera formal, no obstante, literalmente significa “sustitutos”. Esto refiere a la creación de series de tiempo “sustitutas”, las cuales posean la misma distribución de probabilidad, así como la misma función de autocorrelación, las cuales se emplean para validar alguna hipótesis concebida, acerca del proceso que generó la ST, siendo ésta generalmente que los datos bajo estudio han sido generados por un proceso ARMA (*Autoregressive Moving Average Process*, o Proceso Autoregresivo de Media Móvil).

Bibliografía y Referencias

- Abarbanel H. D. I., Brown R., Sidorowich J. J., and Tsimring L. Sh., The analysis of observed chaotic data in physical systems. *Rev M Phys*, 65:1331-1392, 1993.
- Abarbanel H. D. I., Rabinovich M I., Suschik M. M., Introduction to Nonlinear Dynamics for physicists, World Scientific Lecture Notes in Physics 53 (1993).
- Abarbanel H. D. I., Analysis of observed chaotic data, Springer-Verlag, 1995.
- Addison, P. S., Fractals and Chaos: an illustrated course, IOP Publishing Ltd 1997.
- Alligood, K. T., Sauer T.D., Yorke J. A., Chaos: an introduction to dynamical systems, Springer-Verlag New York Inc., 1996, 603 pp.
- Anishchenko V. S., Astakhov V., Neiman A., Vadivasova T., Schimansky-Geier L., Nonlinear Dynamics of Chaotic and Stochastic Systems: Tutorial and Modern Developments, Springer-Verlag Berlin Heidelberg 2007.
- Aubin D., Dahan-Dalmedico A., Writing the History of Dynamical Systems and Chaos: *Longue Durée* and Revolution, Disciplines and Cultures, *Historia Mathematica* 29 (2002), 273–339, 2002 Elsevier Science.
- Claerbout J., Earth soundings analysis: processing versus inversion, Stanford University, 2004.
- Chatfield C., The analysis of time series: an introduction, CHAPMAN & HALL/CRC, 1995.
- Casillas-Lavín Tomás, Tesis de licenciatura, Análisis de series temporales caóticas para la detección de dinámica no lineal: sistemas tiempo/clima, Facultad de Ciencias, UNAM, 2004.
- Cooray T.M.J.A., Applied time series: analysis and forecasting, Alpha Science International Ltd., 2008.
- Devaney, R. L., A first course in chaotic dynamical systems: theory and experiment, Addison Wesley Publishing Company, 1992, 302 pp.

- Diccionario de la lengua española-vigésima segunda edición
<<http://buscon.rae.es/drae/>>[consulta: 6-09-11].
- Donner et al., Lecture Notes in Earth Sciences, Springer-Verlag Berlin Heidelberg, 2008.
- García Carrillo Rosalba, Técnicas de programación, McGraw Hill, 2003.
- Goldberger AL. Nonlinear Dynamics, Fractals, and Chaos Theory: Implications for neuroautonomic heart rate control in health and disease [Biblioteca Virtual de la Salud] < <http://www.physionet.org/tutorials/ndc/>> [consulta: 6-09-11].
- Hegger R., Kantz H., Schreiber, T., Practical implementation of nonlinear time series methods: The TISEAN package, CHAOS 9, 413 (1999).
- Heij C., Ran A., Van Schagen F., Introduction to mathematical systems theory: linear systems, identification and control, Birkhäuser Verlag 2007.
- Herbert DE. Chaos and the changing nature of science and medicine: An introduction [Biblioteca Virtual de la Salud] <http://www.medphysics.wisc.edu/~jrc/rev5.htm> [consulta: 6-09-11].
- Hernández-Arriaga Roberto, Tesis de licenciatura, determinación de la dimensión del atractor y el límite de predictibilidad de una serie de tiempo de horas de calor, Facultad de Ciencias, UNAM, 1996.
- Hernández, O. J., Ramírez, Q. Ma. J. y Ferri, R. C. Introducción a la minería de datos, Pearson, Última reimpresión 2008, 656 Págs.
- Kantz H., Schreiber T., Nonlinear time series analysis. Cambridge, University Press, Cambridge (1997).
- Kostelich, E. J., The analysis of chaotic time-series data, Systems & Control Letters 31 (1997) 313-319.
- Lathi B.P., Signal processing & Linear Systems, Berkeley-Cambridge Press, 1998.
- Lauterborn W., Kurz T. Parlitz U., Experimental Nonlinear Physics, The J. Franklin Inst. Vol. 334B, No. 5/6, pp. 865-907, 1997.
- Madrigal-Loyola ,S.C., Teoría de sistemas
<<http://www.tuobra.unam.mx/obrasPDF/publicadas/010820192601.html>> [consulta: 6-09-11].

- Mandelbrot B. B., *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1982.
- Marwan N., Trauth M. H., Vuille M., Kurths J., Comparing modern and Pleistocene ENSO-like influences in NW Argentina using nonlinear time series analysis methods, *Climate Dynamics* 21: 317–326, (2003).
- Morales-Acoltzi., Alva-Pacheco J.C., Peña-Maciel D., Bustamante- García A.R., Monroy-Martínez J. D. y Bernal-Morales R., 2011, Experimentos numéricos de predicción del tiempo/clima: un enfoque no lineal, *Memorias Congreso UGM 2011, Puerto Vallarta, Jal., México.*
- Morales-Acoltzi T., Bernal-Morales R., Suárez-Sánchez J., Hernández-Rodríguez M. de L., Magaña-Rueda V.O., Herrera-Cortés S., Gómez Leticia, Ordoñez-Díaz B. J. , Barrera-Bassols N., Cuevas-Sánchez A., Variabilidad y cambio climático de la lluvia en tres sitios en el estado de Tlaxcala, Cap. del libro *Foro del agua para enfrentar el cambio climático*, pp. 23, 2012
- Munakata T., Ditto W., *Principles and Applications of Chaotic Systems*, *Communications of the ACM*, November 1995/Vol. 38, No. 11.
- Pichín Quesada M de J, Fariñas Salas AO, Miyares Quintana SM *Los sistemas vivos y las ciencias de las complejidades. Relación entre soma y red biológica* [artículo en línea]. *MEDISAN* 2004;8(3).
<http://bvs.sld.cu/revistas/san/vol8_3_04/san07304.htm> [consulta: 6-09-11].
- Pryor R. G.L., Bright J. E.H., *Applying Chaos Theory to Careers: Attraction and attractors*, *Journal of Vocational Behavior* 71 (2007) 375–400.
- Richter M., Schreiber T., *Phase space embedding of electrocardiograms*, *Phys. Rev. E* 58, 6392 (1998).
- Sánchez-González, P. 2005. Descubrimiento de conocimiento de la tuberculosis en el estado de Tlaxcala, mediante técnicas de minería de datos. Tesis de Maestría. Instituto Tecnológico de Apizaco. México, TLAX. 106 pp.
- Schreiber T., Schmitz A., *Surrogate time series*, *Physica D* 142 346 (2000).

- Schreiber T., Interdisciplinary application of nonlinear time series methods, Phys. Reports 308, 1 (1999).
- Small, M., Applied nonlinear time series analysis: applications in physics, physiology and finance, World Scientific Publishing, 2005.
- Smith S., The scientist and engineer's guide to digital signal processing, California Technical Publishing, Second Edition, 1999.
- Strogatz, S. H., Nonlinear Dynamics and Chaos. Perseus Publishing, 2000.
- Tomás Morales-Acoltzi, Mauro Valdés-Barrón, Alfonso Medel-Rojas, Manuel González-Pérez, Guillermo de la Rosa-Orea y Rogelio Bernal-Morales, Minería de datos y radiación solar global, UNAM, México D.F., pp. 125-128, Vehículos aeroespaciales: aerospace vehicles, Ciencia y tecnología aeroespacial, Soc. Mexicana de Ciencia y tecnología aeroespacial, Puebla, México, SOMECYTA, 2011, ISBN: 978-607-95625-0-2.
- Trauth, M. H., Matlab Recipes for Earth Sciences, 3rd edition, Springer-Verlag Berlin Heidelberg, 2010.
- Turcotte, D. L., Fractals and Chaos in Geology and Geophysics, 2nd Edition, Cornell University, New York, 1997.
- X. Zhang, Yang F., Climate Research Branch Environment Canada Downsview, Ontario Canada, RClimDex (1.0): Manual del Usuario, Versión en español por José Luis Santos, Centro Internacional Para la Investigación del Fenómeno de El Niño CIIFEN, Septiembre 20, 2004.

Anexo 1. Modelo de Lorenz

En 1963, Edward Lorenz publicó su trabajo titulado “flujo determinístico no periódico”, el cual detalla el comportamiento de un modelo matemático simplificado de ecuaciones diferenciales ordinarias que representa el funcionamiento de la atmósfera. Lorenz mostró cómo un modelo matemático relativamente simple (uno que no posee aleatoriedad asociada) era capaz de producir comportamiento aparentemente impredecible, llamado después caos. Dicho sistema de ecuaciones acopladas representa un modelo 2D de convección térmica, conocido como convección de Rayleigh–Benard. El diagrama (Figura 1) muestra la convección de Rayleigh–Benard entre dos placas horizontales. La placa inferior se encuentra a una temperatura T_I que es mayor que la de la placa superior T_S . Para pequeñas diferencias entre las dos temperaturas, el calor es conducido a través del fluido estacionario entre las placas. Sin embargo, cuando $T_I - T_S$ se vuelve lo suficiente grande, las fuerzas de empuje dentro del fluido calentado superan la viscosidad del fluido y un patrón de vórtices estacionarios recirculando en contrarrotación se establece entre las placas. Así, Lorenz notó mediante este simple modelo que diferencias muy pequeñas en las condiciones iniciales ocasionaban enormes diferencias en el comportamiento final. Por lo que concluyó que si este tipo de comportamiento era posible en un sistema dinámico tan simple, entonces también podía ser posible en sistemas físicos mucho más complejos.

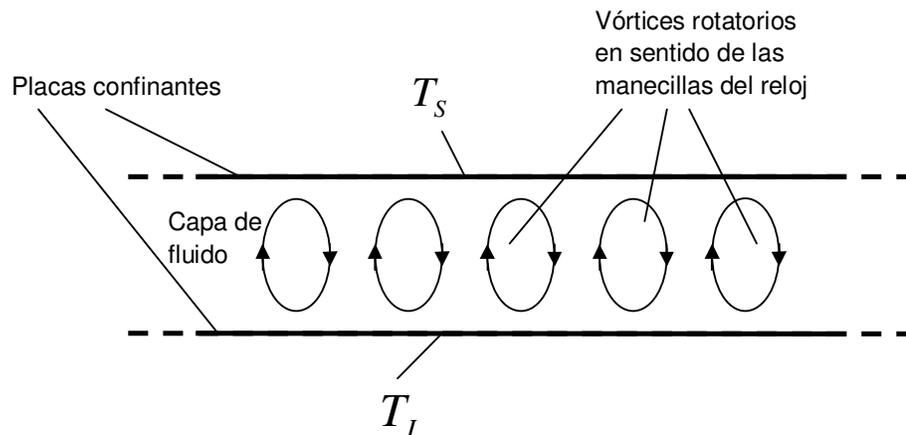


Figura 1. Celdas de convección en un fluido

Las ecuaciones de Lorenz son:

$$\begin{aligned}\dot{x} &= -\sigma(x - y) \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz\end{aligned}$$

Donde:

σ → Es el número de Prandtl, el cual relaciona las pérdidas de energía dentro del fluido debidas a la viscosidad con aquellas pérdidas debidas conducción térmica.

r → Corresponde a la medida adimensional de la diferencia de temperatura entre las placas, conocida como el número de Rayleigh.

b → Hace referencia a la razón existente entre la altura vertical de la capa de fluido y la extensión horizontal de los vórtices convectivos dentro de ésta.

Es importante notar que este sistema tiene dos términos no lineales: “ xy ” y “ xz ”, ambos exhiben movimiento periódico y caótico, dependiendo de los valores de los parámetros de control σ , r y b . Además, hay que señalar que “ x ” representa el gradiente (la variación) de giro convectivo, “ y ” la variación horizontal de la temperatura y “ z ” la variación vertical de temperatura; y no coordenadas espaciales, así también son las respectivas soluciones de las ecuaciones del modelo de Lorenz.