Una visión estadística no-lineal de El Niño; Simulación y posible pronóstico.

Walter Ritter Ortíz *
Pedro A. Mosiño *
Ricardo Klimek Gamas **

Resumen

La dinámica de fluidos, la meteorología y la ecología son sistemas no-lineales que varían drásticamente ante cambios mínimos en sus componentes, donde el orden y la formas con creados no mediante complejos controles sino por la presencia y guía de unos pocos principios y fórmulas. La no-linealidad será siempre el rasgo característico de la evolución de los fenómenos naturales como el fenómeno de "El Niño"; donde los sistemas de no-equilibrio, tanto del estado del tiempo, como de los sistemas ecológicos; se dan sólo como variantes de sistemas complejos que va surgiendo del flujo constante de la energía solar en la biosfera, viéndoseles desenvolverse a través de múltiples bifurcaciones, donde van intercalándose largos períodos de estabilidad con oscilaciones aparentemente azarosas en épocas de inestabilidad. Donde el estar más alejados del equilibrio termodinámico, más sensibilidad de respuesta se manifiesta al cambio de sus estructuras y más sofisticados serán también los ciclos y procesos de retroalimentación que los mantienen. Las bifurcacioes catastróficas nos dan apariciones y desapariciones súbitas de atractores estáticos, periódicos ó caóticos y son la clase de transformaciones que sustentan la evolución de sistemas que van desde los átomos, hasta especies ó sistemas climáticos, ecológicos y sociedades. El propósito del presente trabajo es detectar procesos determinísticos ó caóticos en datos aparentemente aleatorios. Cuando se prueba que existe "caos", se dan herramientas (Distribuciones probabilística, análisis espectral, exponentes de Hurst y Lyapunov y medidas de dimensión fractal) que nos permiten determinar propiedades de las ecuaciones representativas de pronóstico de redes "neuronales", predicción no-lineal y descomposición de valores singulares.

Abstract

The dynamics of fluids, the meteorology, and the ecology are non-linear systems which vary drastically in the presence of minimal changes in their components, where the order and the form are created and influenced by very few principles and formulas. The non-linearity is always a key characteristic in the evolution of natural phenomena like "El Niño". Systems in non-equilibrium such as ecological systems are only present as variants of complex systems which appear from the constant flow of solar energy in the biosphere. These alternate between large periods of stability with oscillations which are apparently random in periods of instability ; the more removed the thermodynamic equilibrium, the more sensitive is the response that occurs in the change of its structures, and the more sophisticated are the cycles and processes of feedback that maintain them. The catastrophic bifurcations give us sudden appearances and disappearances of statistical, periodic, and chaotic attractors, and they are the class of transformations that sustain the evolution of systems from the level of atoms to species or climatic, ecological, and social systems. The objective of this study is to detect the determining or chaotic processes in data that seem random. When it is proven that 'chaos' exits, it gives us the means to determine the properties of equations that represent the behaviors of the phenomenon in question, including prognostic methods of neural systems, non-linear prediction, and the decomposition of particular values.

Introducción

Para describir trayectorias de sistemas sencillos que consisten en un número muy pequeño de cuerpos en movimiento, la física de Newton sigue siendo de gran utilidad ; pero este determinismo a menudo es irrelevante en el mundo real, donde es imposible poseer un perfecto conocimiento de la forma en que se comporta un sistema complejo, que comprende muchos cuerpos. Desafortunadamente en una abrumadora mayoría de casos, los sistemas meteorológicos

* *Centro de Ciencias de la Atmósfera, UNAM*
** *Universidad del Mar (UMAR), Puerto Angel, Oax*

como el sistema océano-atmosfesta del fenómeno de “El Niño”, muestran que la más leve incertidumbre en la descripción de sus condiciones iniciales conducirán irremediablemente a futuros muy distintos, y por lo tanto hay que desarrollar modelos especiales para estos fines, que son situaciones clásicas de un sistema de dinámica no-lineal. Es decir que mientras la casualidad lineal puede funcionar suficientemente bien para sistemas limitados, mecánicos y aislados, en general se necesita algo más delicado y complejo para describir la extensa riqueza no-lineal de la naturaleza. Mientras que la solución de ecuaciones lineales nos permiten generalizaciones, las no-lineales tienden a ser individualizadas, con términos que a través de la retroalimentación se multiplican repetidamente, mostrando recurrencias, turbulencia, rupturas y rizos de todas clases. Poincaré (1889) observó que los sistemas cuasi-períodicos son predominantes en la naturaleza al extender la mecánica de Newton a tres o más cuerpos, encontró el potencial para la no-linealidad, la inestabilidad y el caos. Al utilizar un mayor rigor de análisis para los sistemas no-lineales, empezamos a entender nuevas formas de interpretar sus fluctuaciones y cambios ; donde el orden y la forma son creados no mediante complejos controles sino por la presencia y guía de unos pocos principios y fórmulas. Esta nueva dinámica demuestra que todo movimiento y cambio surge de una ley del todo y que los patrones y sucesos de la naturaleza son la expresión de la unidad fundamental de su forma. Así cuando diferentes partes de un fluido viajan a diferentes velocidades, para velocidades bajas se presenta un movimiento uniforme, pero conforme aumentamos de velocidad las inestabilidades no se hacen esperar a través de una serie de rotaciones internas, con creciente complejidad y oscilaciones en diferentes frecuencias, las cuales terminan de forma aleatoria.

Con la primera inestabilidad pasamos de un atractor de punto fijo a uno de ciclo límite y posteriormente a un atractor “toroidal” con un creciente número de dimensiones, ingresando en espacios de dimensión fraccional, donde las fluctuaciones aleatorias se anudan en la forma de un atractor extraño. Los modelos de sistemas dinámicos podrían ser universales, es decir, no específicos para ejemplos individuales, sino representativos de clases enteras de sistemas, poniendo de manifiesto no solo la aparición del caos, sino también la manera en que éste puede ser creado. El comportamiento de todo el sistema depende esencialmente de su complejidad y el comportamiento individual depende del todo. En la evolución de estos sistemas dinámicos se ha observado que es el desequilibrio la condición necesaria para el crecimiento del sistema. Estos sistemas usan el desequilibrio para evitar el deterioro y disipan energía a fin de encontrar nuevas formas de organización; es decir que actúan como sistemas autoorganizativos. La naturaleza en su manifestación contiene diversos grados de orden cuya jerarquía nos toca descifrar, ya que la observación y manifestación del azar puede ser sólo nuestra incapacidad para comprender otros grados superiores de orden.

Métodos: Propiedades básicas y observaciones generales.

El comportamiento de los sistemas complejos se investiga experimentalmente observando una variable que se considera “de estado”, a lo largo de un tiempo determinado. Pero una serie temporal de una variable única pudiera parecer que posee un contenido muy limitado de información; estando limitada a una perspectiva “unidimensional” del sistema. Sin embargo, una serie temporal contiene informaciones mucho más ricas : llevando el sello de todas las demás variables que participan en la dinámica y permite identificar, independientemente de cada modelo, algunas propiedades clave del sistema subyacente. Donde una secuencia de dichos estados define una curva llamada “trayectoria en el espacio de fases”. Posteriormente el sistema alcanza un régimen permanente, quedando reflejado en la convergencia de familias enteras de trayectorias del espacio de fases hacia un conjunto inferior del espacio de fases. Esta conjunto inferior invariante será el “atractor”. Donde su dimensión brinda información valiosa acerca de la dinámica del sistema. Así para la dimensión $d=1$, tendremos oscilaciones periódicas autoexcitadas ; si $d=2$, tendremos oscilaciones cuasiperiódicas con dos frecuencias inconmensurables ; si d es entero y mayor de 2, entonces el sistema presentará previsiblemente un comportamiento de oscilación caótico que presentará una gran dependencia respecto a las condiciones iniciales y una imposibilidad de calcular intrinseca. Poincaré (1889) en su estudio sobre la existencia de soluciones periódicas para la ecuaciones diferenciales, demuestra cómo obtener tales soluciones desarrollando la variable involucrada como una serie infinita, en donde cada término es una función periódica del tiempo. Existiendo series que satisfacen las ecuaciones formalmente y cuyos coeficientes son periódicos. Al demostrar la existencia de soluciones periódicas lo que implica la convergencia de la serie, la unicidad de las soluciones significa que el movimiento es periódico. La existencia de soluciones periódicas depende de las propiedades topológicas de la relación entre la posición inicial de un punto y su posición después de transcurrido un período. La existencia de la sección de Poincaré puede forzar por motivos topológicos, a que tenga lugar una solución periódica, desechando aspectos confusos y simplificando la observación de la dinámica del sistema; donde basta con mirar unos pocos estados iniciales y seguir la evolución de cada uno de ellos hasta que con el choque de su regreso se obtiene una solución periódica. Empleando una sección de Poincaré es posible detectar este movimiento periódico. Para que exista periodicidad, la curva debe volver a la sección exactamente en el punto de partida. Cuando se intersectan las trayectorias de un sistema dinámico bidimensional lo hacen en puntos singulares, clasificándose como de “silla de montar” y de “nodo”. Cuando sucede la misma cosa en una sección transversal bidimensional de una hoja, la intersección puede volver a ser o un nodo, etc., pero ahora existe una segunda posibilidad : la intersección en un punto “no singular”. La curva de Lorenz (1963) es una especie de hermana pobre de la sección de Poincaré. En vez de representar una variable en períodos de tiempo regulares, los intervalos de tiempo son irregulares, aunque existe un ritmo subyacente definido. Se puede predecir algo de la dinámica del sistema pero sólo a corto plazo, ya que para obtener una predicción a largo plazo, los errores crecen tan rápido que las predicciones se hacen absurdas. Toda la dificultad en la predicción del tiempo radica en que el tiempo no es periódico.

El objetivo del análisis experimental de la información es encontrar un patrón o estructura que modele los datos. La estructura misma puede no ser simple y obvia, pero una vez encontrada, la información original adquiere una nueva dimensión de simplicidad. Analizaremos información en series de tiempo para una sola variable (temperatura oceánica) con las consideraciones de que la presencia de un atractor extraño, puede llevarnos a series de tiempo complicadas y de que dicho fenómeno puede ser descrito por ecuaciones no-lineales simples. La presencia de un atractor extraño, puede muy bien ser revelado de gráficas de fase-espacial. Si los datos son dados en intervalos cortos de tiempo y no tienen un componente de ruido significativo, entonces los gráficos de fases-espaciales pueden ser construidos con componentes de la información y sus derivadas y entonces toda la información puede ser representada como una trayectoria en esta fase espacial. El comportamiento de los sistemas complejos se investiga experimentalmente observando una variable, que se considera “de estado”, a lo largo de un intervalo de tiempo determinado. El procedimiento será el de considerar fases espaciales de dimensión cada vez mayor, hasta que un incremento más no cambie la topología de la estructura observable. Una vez que el atractor extraño ha sido identificado, puede ser cuantificado al calcular varias medidas de su dimensión y su exponente de Lyapunov. La dimensión es una medida de la complejidad de la trayectoria de la fase-espacial, mientras que el exponente de Lyapunov es una medida de la sensibilidad a las condiciones iniciales. Para tener una medida exacta de la dimensión de un atractor, es teóricamente necesario envolverlo en un espacio de dimensión de al menos 2d+1, donde d es la mínima dimensión integral conteniendo al atractor. Hay varias formas de definir la dimensión de un atractor y la más simple es la dimensión de capacidad, que describe la geometría de la dimensión de el atractor sin considerar qué tan frecuente la trayectoria visita las localidades sobre el atractor. La dimensión de correlación es una frontera inferior de la dimensión de capacidad, pero en la mayoría de los casos se aproximan sus valores. El mayor obstáculo en el cálculo de la información fractal y topológica es la insuficiencia de los datos. Sin embargo, las ecuaciones del modelo

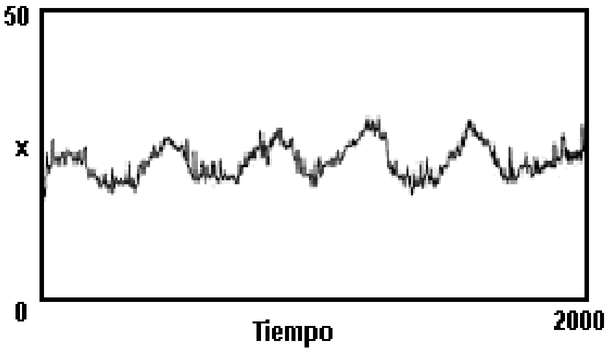

Fig.1 Valores de la temperatura del oceano frente a la costa del Peru, obtenidos a un metro de profundidad por medio de boyas, periodo 1988-1994 (2 000 datos).

derivadas de la descomposición de valores singulares puede usarse para generar los datos necesarios para calcular la dimensión y otras propiedades del atractor, obteniendo una simplificación significativa para cuando las ecuaciones reflejan razonablemente la topología de la información. Podemos decir que las metodologías de análisis de la información de sistemas no-lineales, están aún muy lejos de ser una ciencia exacta, por lo que proponemos ciertos procedimientos generales a seguir, para encontrar patrones o estructuras que modelen los datos y que nos puedan ayudar a descubrir fenómenos no-lineales en la información que manejemos. Utilizaremos varios tipos de análisis estadístico con este propósito analizando 6 años de información (2000 datos) para los años de (1988 - 1994) de temperatura oceánica, fig. (1), medida en una boya a un metro de profundidad en la zona ecuatorial denominada "niño 1", a fin de vislumbrar alguna manifestación de las ya señaladas en los capítulos introductorios. Se construyen "Espacios de fase multidimensional", "mapas de retorno" y "figuras de Poincaré", donde la gráfica de los datos nos permite presentar la información en diferentes formas incluyendo el graficado de cada valor contra su inmediato predecesor y gráficas de las diferentes derivadas de la información contra los valores originales para revelar la topología de la solución. Debemos esperar que los sistemas "periódicos" exhibirán figuras cerradas, donde para situaciones más complicadas, estas se darán en regiones de dos o tres dimensiones y con estructuras poco definidas. Este mismo tipo de estructuras borrosas se darán también cuando el sistema sea casi periódico para tiempos de largo plazo. Después de esto, la información caótica aparecerá en la forma de un atractor con estructura de dimensión fraccional (fractal) menor de dos dimensiones, con repetidos estiramientos y plegamientos de las trayectorias, causando que los puntos vecinos se separen pero cuando los datos estén dominados por el ruido, estos no presentaran estructura alguna. En la fig. (2), podemos encontrar similitudes tipológicas de nuestros datos de temperatura con los del movimiento Browniano. En pruebas probabilísticas de conglomerados, para observar diferencias entre ambos grupos de datos, el "ruido blanco" del movimiento Browniano llena uniformemente las figuras y los datos caóticos de nuestra información producen aglomeraciones localizables regionalmente fig. (3), señalando claras diferencias entre los dos grupos de datos. Posteriormente las "pruebas de análisis espectral" que se realizan a través de "transformaciones rápidas de Fourier", sobre los datos y donde se presenta su "amplitud media cuadrática (potencia espectral) como función de la frecuencia sobre escalas lineales, log-lineales o loglog; debemos esperar que los espectros de potencia que son líneas rectas sobre la escala lg-lineal sean buenos candidatos de representación caótica y de que los "Periodigramas acumulativos" (integral de la potencia espectral sobre la frecuencia) deban seguir una línea de 45° si el espectro de potencias es liso indicando "ruido blanco". En la fig. (4) se muestra el periodigrama de nuestra información a analizar y la del movimiento Browniamo (línea de 45°), lo que nos da las diferencias bien definidas de algo que por falta de datos parecía similar en los espacios de fase. El método de la entropía máxima muestra una alta eficiencia en localizar los picos de perioricidad en las frecuencias de la información que de otra manera pasarían simplemente como ruido. Sin embargo presentan el peligro, para cuando no existan datos suficientes, de que podrían presentarse situaciones de correlación "espuria". Este método también es utilizado para eliminar tendencias (desestacionalizar) en la información y predecir los siguientes valores en la serie de tiempo, mostrando diferencias entre lo observado y lo pronosticado fig. (5). El exponente de Lyapunov es una medida de la razón en que trayectorias vecinas tienden a divergir

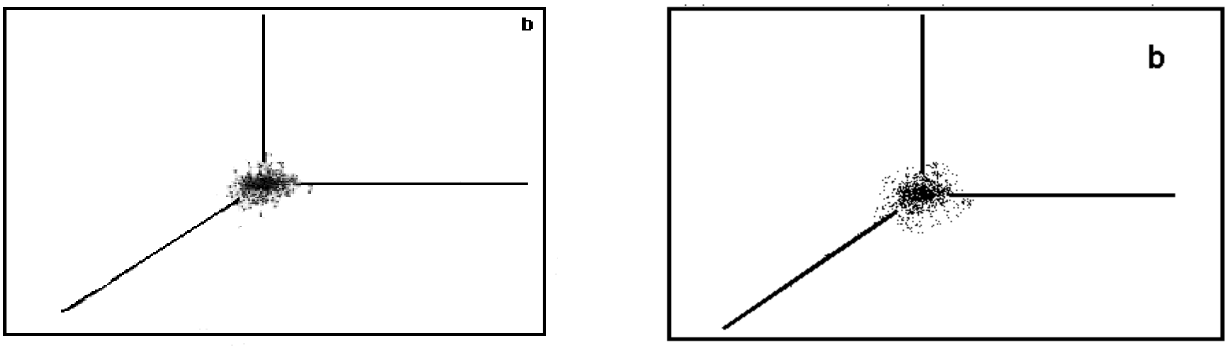

Fig.2 Diagramas tridimensionales de fase espacial para datos del movimiento Browniano (a) y de la temperatura del océano (b).

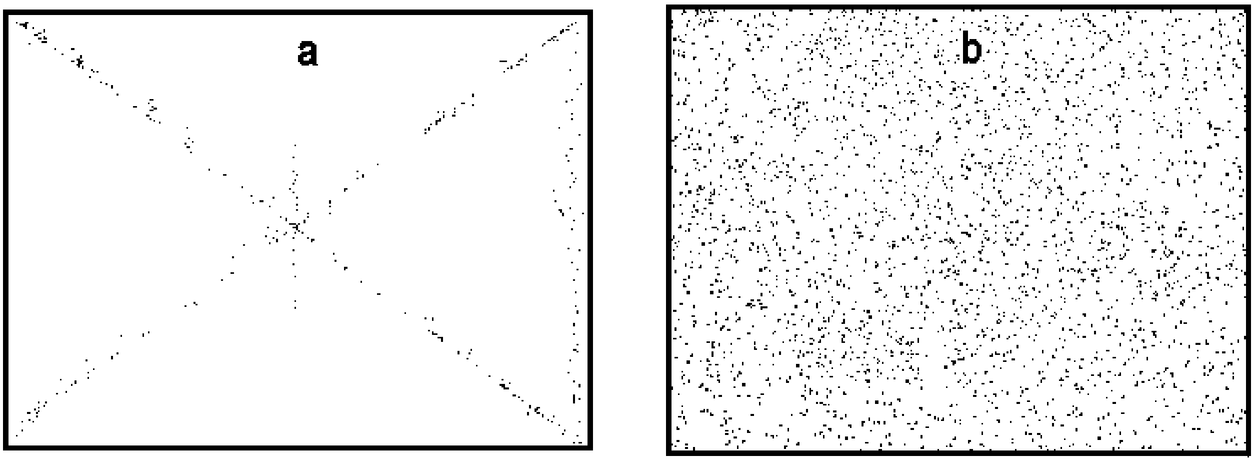

Fig. 3 Distribución probabilística de conglomerados de datos de temperatura del océano (a) y del movimiento Browiano (b); derivados de consideraciones de Ecología Cuantitativa.

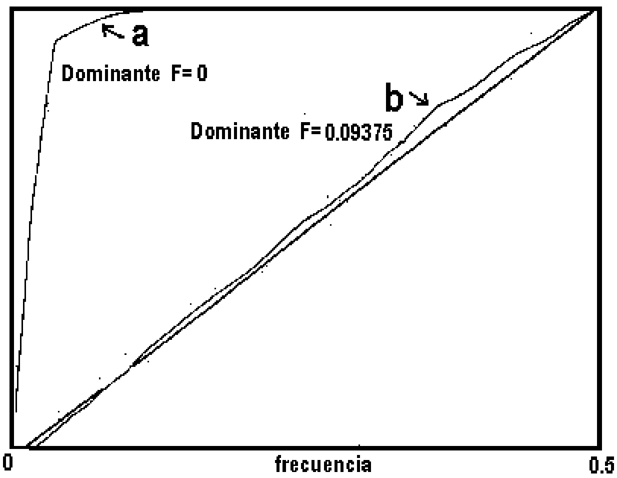

Fig. 4 Periodigrama acumulativo integral de la potencia espectral sobre la frecuencia) para los datos de la temperatura del océano (a) y el espectro de potencia del movimiento Browniano (b) siguiendo una línea a 45^0, indicando ruido.

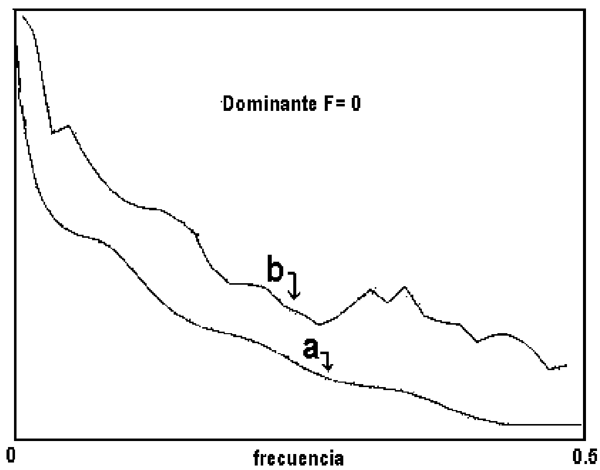

Fig. 5 Periodigrama por el método de frecuencias dominantes o entropía maxima (a) y de transformadas rápidas de Fourier (b).

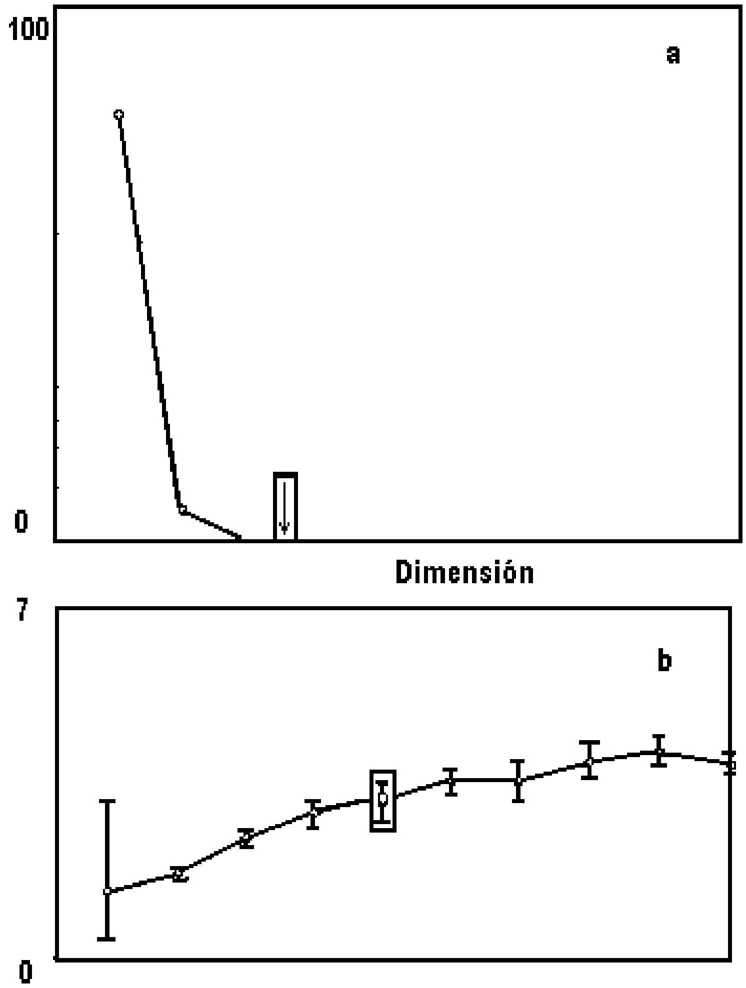

Fig. 6 Gráficas para el cálculo de la dimensión de correlación, (a) Porcentaje de falsos vecinos que convergen a cero en la mínima dimensión envolvente, (b) Dimensión de correlación calculada en función de la mitad de la dimensión envolvente (3,435+0.339)

en el espacio de fases, donde órbitas caóticas tienen al menos un exponente de Lyapunov positivo y serán todos negativos en el caso de que sean periódicas, o cero en la proximidad de una bifucación. Debemos determinar la "dimensión envolvente" apropiada para reconstruir el atractor con la "dimensión de correlación", donde se busca la saturación y conforme la envolvente vaya aumentando, ocurrirá la saturación en la correlación en un valor no mayor a la mitad de la envolvente. La dimensión de un atractor se dice que está bien definida si existe una región de saturación plana en la gráfica de la "dimensión de capacidad" contra el logaritmo de la dimensión lineal normalizada. Dimensiones mayores de 5 significan que los datos son esencialmente aleatorios. Identificada la "envolvente" se procede a cuantificar el atractor, donde las dimensiones de correlación y de capacidad describirán su grado de complejidad fig. (6). La entropía es la suma de los exponentes positivos de Lyapunov y su reciproco es aproximadamente el tiempo en el cual predicciones significativas son posibles de realizar. La función de correlación nos da una medida de qué tan dependiente son los datos de la información con respecto de sus vecinos temporales. Las eigenfunciones dominantes de la matriz de correlación son usadas para formar las funciones básicas para la modelación no-lineal de los datos. Si los datos consisten de componentes caóticas y aleatorias este procedimiento debe ayudar a extraer el caos de la aleatoriadad de la información. El número de eigenvalores significativos, es también una medida de la complejidad del sistema fig. (8 y 9).

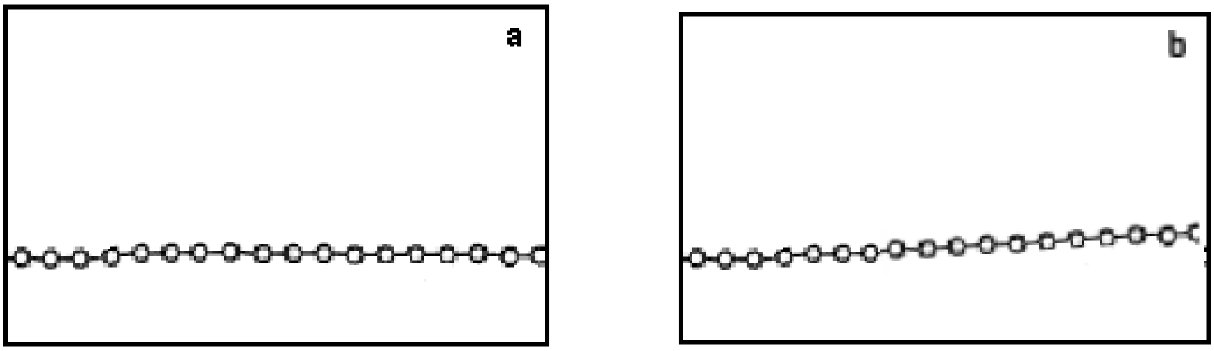

Fig. 7 Gráficas de valores pronosticados de la temperatura del océano por el método de redes nueronales (a) y de la predicción no lineal (b). para 18 días.

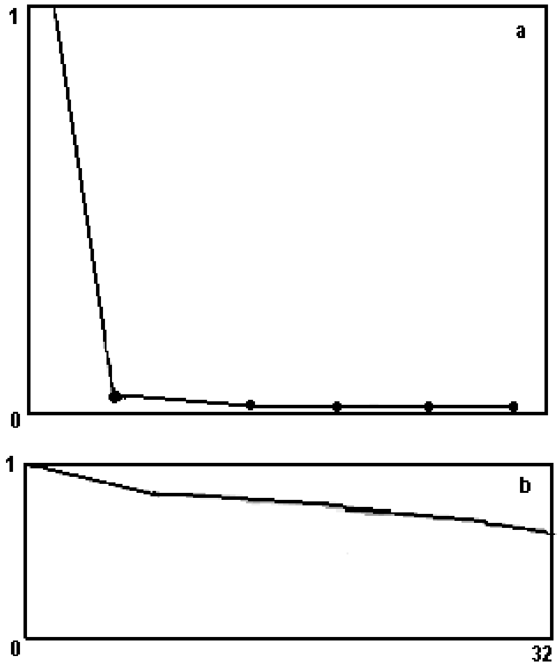

Fig. 8 a) Gráfica de los valores de los eingenvectores (de mayor a menor) derivados de la matriz de correlación o del análisis de componentes principales. (b) Gráfica de la función de correlación. Da una medida de que tan dependientes son los datos de sus vecinos temporales. El valor de (62.573) en que la función de correlación cae primero a 1/e, y el tiempo de correlación del orden inverso del exponente de Lyapunov, donde si los datos aleatorios no tienen correlación, rápido caen sus valores a cero.

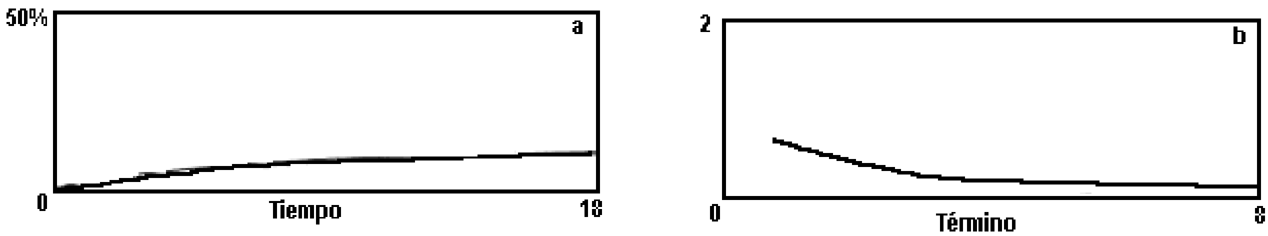

Fig. 9 a) Error de predicción promedio del método de predicción no lineal como porcentaje del máximo error posible. Para datos caóticos el error crece exponencialmente saturándose en el cociente: desviación estandar/rango. b) Pesos de los n coeficientes de la predicción neuronal donde cada termino en términos de pesos determinados por el mejor ajuste a los datos existentes. La red aprende al variar los pesos para minimizar el error. Da una medida burda del determinismo existente en los datos ya que ni un sistema caótico pierde información conforme avanza el tiempo.

Métodos de pronóstico

a. MÉTODO NEURONAL.- El método neuronal nos provee de medios para reducir el ruido en la información, produciendo datos adicionales y pronosticando términos a corto plazo de una serie de tiempo determinística; donde cada término en la serie de tiempo es considerada que esta dada por la superposición de los previos N términos, con "pesos" determinados por el mejor ajuste a los datos. El método neuronal aprende al ir variando los pesos para minimizar el error, emulando la forma como el cerebro funciona. La figura (7) muestra la solución a corto plazo después del final de los datos, basada sobre los pesos óptimizados.

b. PREDICCION NO-LINEAL.- La predicción nolineal nos da una medida del grado de deterministmo en la información, remueve el ruido y permite predicciones a corto plazo fig. (7).

Resultados

Los resultados del presente estudio, considerados como más sobresalientes, son los siguientes :

a.- La estructura topológica de la información analizada presenta cierta similitud con la del movimiento Browniano, sin embargo al aplicar procedimientos probabilísticos derivados de metodologías de ecología cuantitativa y e análisis espectral, las diferencias aparecen de forma bastante clara.

b.- Se calcula la dimensión del atractor de la información analizada, resultando del orden de 3.5 lo que delata una alta complejidad del sistema (el atractor de Lorenz toma valores apenas de 2.1) , lo cual es de esperarse ya que a un metro de profundidad la interación océano-atmósfera es bastante relevante.

c.- Al calcular el número necesario de datos para obtener un atractor bien definido y no confundirse con manifestaciones de posibles situaciones de periodicidad de largo plazo, se obtuvieron valores próximos a 5000 datos, lo que implica la obtención e incorporación de más del doble de la información existente.

d.- La aplicación de métodos de pronóstico a corto plazo (18 días) por las técnicas de máxima entropía, redes neuronales y métodos no-lineales, coinciden entre sí al presentar durante estos 18 días situaciones acertadas clásicas de persistencia fig. (7). Falta por probar sus respuestas a situaciones de cambio repentino anual de gran intensidad como pudiera ser la aparición súbita de este ultimo fenómeno de el niño (1997 - 1998).

Propuestas y estrategias de análisis en sistemas dinámicos no-lineales (fenómeno de "El Niño").

En la previa manipulación de los datos es aconsejable realizar las siguientes operaciones para la determinación de la posible falta de ESTACIONARIDAD (tendencia general de comportamiento en la información, que es parte de la dinámica del sistema pero no es lo relevante).

1.- Buscar por "Estructuras Obvias" en la información graficada como :

a.- En las gráficas de fase-espacial.
b.- Mapas de retorno.
c.- Figuras de Poncaré.

2.- Buscar frecuencias sobresalientes en ;

a.- Distribuciones de probabilidad.
b.- Ajustes polinomiales de bajo orden.

3.- Utilizar el Análisis de Potencias Expectrales (por series de Fourier) y Frecuencias Dominantes.

4.- Utilizar la hipótesis de que la información representa un "Atractor (Extraño)" u otro sistema caótico (posiblemente en una fase espacia de alta dimensión).

a.- Determinar la "dimensión envolvente" apropiada para reconstruir el atractor con la "dimensión de correlación", donde se busca la saturación y conforme la primera dimensión vaya aumentando, ocurrirá la saturación en la segunda, en un valor no mayor a la mitad de la primera.

b.- La pendiente de la suma de correlación, cuando está propiamente bien definida la envolvente, debe exhibir una clara manifestación de meseta y si no es posible encontrarla es que el sistema es caótico y de dimensión alta, ó está contaminado por ruido.

c.- Si se logra identificar una envolvente apropiada podemos proceder a cuantificar el atractor, donde la “Dimensión de Correlación” y la “Dimensión de Capacidad”, describirán su grado de complejidad.

d.- El “Exponente de Lyapunov” nos dará una medida de la sensibilidad del sistema a sus condiciones iniciales.

5.- PRUEBA DE VERIFICACIÓN FINAL.- Bajo cualquier conclusión anteriormente alcanzada, se debe siempre de hacer uso de esta prueba con el siguiente procedimiento.

a.- Utilizar la transformada de Fourier.
b.- Aleatorizar las fase.
c.- Utilizar la transformada inversa de Fourier.

Como RESULTADO, se obtiene una nueva serie de Fourier, con las mismas propiedades espectrales como el original, pero con el determinismo removido.

Recomendación

Si se encuentra estadísticamente significativa la dimensión del atractor (con un valor menor de 5.0) y un exponente de Lyapunov positivo, vale la pena analizar la información utilizando el método de “DESCOMPOSICIÓN DE VALORES”.

Agradecimientos

Se agradece la ayuda prestada por el M. en C. Alfonso Estrada, durante la realización de este trabajo, así como la de Ma. Esther Grijalva por preparar la versión final.

Bibliografía

Lorenz E. N. 1963. Deterministic nonperiodic flow. *J. Atm.* Sci., 20:130-141

Poincaré H. 1899. Les méthodes nouvelles de la mecánique séleste. Gauthier-Villars, Paris.