

MEDIDAS DE DIMENSIÓN E INFORMACIÓN COMO ESTADÍSTICAS DISCRIMINANTES PARA LA PRUEBA DE HIPÓTESIS

Dimension and Information Measurements as Discriminating Statistics for Hypothesis Testing

RESUMEN

El método de los datos sustitutos es utilizado para la detección de no linealidad en series temporales. Los resultados del método dependen de la estadística discriminante utilizada para realizar la prueba de hipótesis. Comúnmente se utiliza la dimensión de correlación para este fin, la cual depende de una reconstrucción óptima del atractor. Desafortunadamente, en muchas ocasiones esto no es posible. En este documento se utilizan diferentes estadísticas no lineales que no dependen de un atractor para realizar prueba de hipótesis. Los resultados se comparan con los obtenidos utilizando la dimensión de correlación. Se observa que la complejidad de Lempel-Ziv es una buena estadística para la prueba de hipótesis.

PALABRAS CLAVES: análisis de series de tiempo, complejidad de Lempel-Ziv, datos sustitutos pseudo periódicos, información mutua promedio, método de los datos sustitutos, métodos de Monte Carlo, prueba de hipótesis.

ABSTRACT

Surrogate data method is used to detect nonlinearity in time series. The method results depend on the discriminating statistic used to perform the hypothesis testing. Correlation dimension is commonly employed for this purpose, which depends upon the optimal reconstruction of the attractor. Unfortunately, in many cases this is not possible. In this article several nonlinear statistics that do not depend upon an attractor are used to do hypothesis testing. The results are compared with the objectives using correlation dimension. It is noted that Lempel-Ziv complexity is a good statistic for hypothesis testing.

KEYWORDS: time series analysis, Lempel-Ziv complexity, pseudo periodic surrogates, average mutual information, surrogate data method, Monte Carlo methods, hypothesis testing.

1. INTRODUCCIÓN

Existen dos razones que motivan a un investigador a aplicar métodos de análisis no lineal a una serie de datos, la primera es que se han utilizado una serie de métodos lineales y aún existen estructuras desconocidas en la serie de datos. La segunda es por un conocimiento a priori de la no linealidad del sistema dinámico que generó la serie de datos. Sin embargo, a pesar de las características no lineales del sistema, es imposible afirmar que la señal específica que se mide refleje este comportamiento. A fin de cuentas no se desea que el análisis realizado sobre la serie temporal refleje los prejuicios del investigador sobre el sistema dinámico subyacente. Por lo tanto, previo a la aplicación de cualquier método de análisis no lineal a una serie de datos es necesario descartar primero que esta haya sido generada por algún sistema dinámico sencillo. Para alcanzar este objetivo Theiler et al. [1] basados en el método de prueba estadística de hipótesis de Monte-Carlo desarrollaron el método de los datos sustitutos, el cual se emplea para clarificar declaraciones sobre la presencia de no linealidad en series temporales.

El procedimiento para establecer una prueba estadística de hipótesis se puede resumir de la siguiente manera: i) se selecciona una hipótesis nula; ii) se generan un conjunto de realizaciones simuladas (o sustitutos) que satisfagan la hipótesis y sean similares a los datos originales; iii) se elige una estadística o una batería de estadísticas (para pruebas de no linealidad es preferible usar estadísticas no lineales); iv) se calcula la distribución de frecuencias de la estadística para las distintas realizaciones y; v) con la guía de la distribución de frecuencias, se elige un criterio para determinar si se rechaza o no la hipótesis [1-4]. Con las hipótesis siempre se avanza de modelos simples a modelos más generales y complejos.

El método de los datos sustitutos provee una forma rigurosa para aplicar pruebas estadísticas de hipótesis a series temporales experimentales. Tres hipótesis nulas, conocidas como hipótesis estándar fueron originalmente propuestas por Theiler et al. [1] en 1992: 0) ruido independiente e idénticamente distribuido (i.i.d); 1) ruido filtrado linealmente; y 2) transformación no lineal estática monótona del ruido filtrado linealmente. En la

DIEGO L. GUARÍN

Ingeniero Físico
Estudiante Maestría Ing. Eléctrica
Grupo de Investigación en Control e Instrumentación
Universidad Tecnológica de Pereira
dlguarin@gmail.com

ÁLVARO Á. OROZCO

Ingeniero Electricista, Ph. D.
Docente Titular
Grupo de Investigación en Control e Instrumentación (Director)
Universidad Tecnológica de Pereira
aaog@utp.edu.co

JUAN SEBASTIÁN HURTADO J

Ingeniero Electricista
Universidad Tecnológica de Pereira
jshurtado@utp.edu.co

actualidad se han generado otra serie de hipótesis conocidas como hipótesis no lineales, una de ellas es que los datos son generados por un sistema dinámico pseudo periódico (los sistemas pseudo periódicos son usualmente confundidos con sistemas caóticos) [3].

A continuación se mencionan los algoritmos por medio de los cuales es posible generar sustitutos que satisfagan cada hipótesis y sean similares a los datos. Igualmente, en este artículo se realizaron prueba de hipótesis en tres series temporales reales: I) valor diario promedio de la temperatura; II) promedio mensual de manchas solares; e III) intensidad de un laser NH3. El presente documento es novedoso debido a que se comparan los resultados obtenidos con tres estadísticas no lineales, dos de las cuales no requieren la reconstrucción de un atractor, lo cual en muchas situaciones reales puede ser problemático [2].

2. MATERIALES Y MÉTODOS

2.1. Materiales

Dado que la finalidad del presente artículo no es obtener conclusiones definitivas sobre algún tipo de señal en particular, sino, presentar una metodología y las posibles conclusiones que se pueden obtener al aplicarla, entonces se decidió utilizar un conjunto de señales que provienen de diversos sistemas dinámicos. Primero se analiza la variación de la temperatura promedio diaria en *Madison (WI, USA)*, la serie cuenta con 3652 registros y fue obtenida de <http://sprott.physics.wisc.edu/chaosta/data/msntemp.dat>. La segunda serie es el promedio mensual de manchas solares, la cual cuenta con 3127 registros y fue obtenida de ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_NUMBERS/MONTHLY.PLT. Finalmente se analiza la intensidad de las pulsaciones de un laser infrarrojo-lejano NH3, la serie cuenta con 9093 registros y fue obtenida de <http://www.psych.stanford.edu/~andreas/TimeSeries/SantaFe/A.cont>.

2.2 Métodos

A continuación se introducen los algoritmos por medio de los cuales se pueden generar sustitutos a partir de los datos que sean consistentes con cada una de las hipótesis. Se presenta brevemente el método estándar de prueba estadística de hipótesis introducido por Theiler et al. [1] y ampliamente estudiado en la literatura [4,6-7]. Posteriormente se describe el sistema de generación de sustitutos para señales pseudo periódicas, originalmente introducido por M. Small et al. [3] y ampliamente utilizado en la actualidad [2,8-11].

2.2.1. Algoritmos Estándar

Los tres algoritmos originalmente propuestos por Theiler et al. [1] para generar sustitutos que satisfagan las tres hipótesis lineales son el algoritmo 0, el algoritmo 1 y el

algoritmo 2 (AAFT); éstos se encuentran explicados de forma detallada en Guarín et al. [5].

2.2.2. Algoritmo PPs

El algoritmo PPs genera sustitutos que preservan las características deterministas secundarias (como la estructura periódica de los datos) pero destruye las estructuras periódicas finas (como el caos determinista). M. Small y C.K. Tse [12] aplicaron el algoritmo a varias series artificiales y experimentales pudiendo discriminar entre caos y una órbita periódica ruidosa. Uno de los resultados destacables de esta publicación es que usan exitosamente la información mutua promedio como estadística. Para obtener los sustitutos se debe seguir el siguiente procedimiento:

Sea $\{z_t\}_{t=1}^N$ una serie temporal escalar de N elementos. Para una dimensión de embebimiento (d_E) y un tiempo de retardo óptimo (τ) se reconstruye el atractor como se muestra en la ecuación (1) de acuerdo al teorema de embebimiento de Takens [13]:

$$y_t = [z_t, z_{t+\tau}, \dots, z_{t+(d_E-1)\tau}] \quad (1)$$

Para $t = 1, 2, \dots, N - (d_E - 1)\tau$. Por conveniencia se define $W = N - (d_E - 1)\tau$, por lo tanto la serie embebida se puede escribir como $\{y_t\}_{t=1}^W$. Para generar los sustitutos se utiliza el siguiente procedimiento:

1. Se escoge de manera aleatoria una condición inicial s_1 , donde $s_1 \in \{y_t | t = 1, \dots, W\}$. Se hace $i = 1$.
2. Se calcula $d_j = \|s_i - y_j\|$ entre el estado actual i y todos los puntos de atractor.
3. Se asigna un peso $\omega_j = \exp \frac{d_j}{\rho}$ a cada punto y una probabilidad $p_j = \omega_j / \sum_k \omega_k$.
4. Se escoge un vecino s_i de acuerdo con la distribución de probabilidad implícita. Es decir, la probabilidad de elegir y_j es p_j .
5. Se hace $s_{i+1} = y_j$.
6. Se repite el proceso desde el paso 2 hasta que $i = N$.

La serie de tiempo vectorial $\{s_t\}_{t=1}^N$ representa una trayectoria estocástica del atractor aproximado por $\{y_t\}_{t=1}^W$ con la misma dinámica subyacente que la serie original pero contaminada con ruido. La serie de datos sustitutos generados con el algoritmo PPs para $\{z_t\}_{t=1}^N$ es $\{(s_t)_1\}_{t=1}^N = \{(s_1)_1, (s_2)_1, \dots, (s_t)_1, \dots, (s_N)_1\}$.

El algoritmo anterior depende de tres parámetros: la dimensión de embebimiento d_E , el tiempo óptimo de retardo τ y el radio del ruido ρ . La selección de los parámetros d_E y τ está ampliamente estudiada [14]. Con respecto al radio del ruido debe tenerse en consideración que si este es muy grande entonces el algoritmo introducirá mucha aleatoriedad y los sustitutos no se parecerán a los datos. Por el contrario si ρ es muy pequeño no se introducirá suficiente aleatoriedad y los sustitutos poseerán determinismo no periódico. El valor ρ

se elige de tal manera que la cantidad de secuencias que posean dos o más valores idénticos en los sustitutos y la serie original se maximicen [8]. El hecho que la estadística de los datos no es típica de la estadística de los sustitutos generados con el algoritmo PPs indica la presencia de una dinámica determinista no trivial en los datos.

Finalmente para aplicar estos sustitutos es necesario escoger una estadística adecuada. En su artículo original, Theiler et al. [1] utilizaron la dimensión de correlación, el máximo exponente de Lyapunov y el error de predicción no lineal; una característica común de estas tres estadísticas es que requieren la reconstrucción de un atractor a partir de la serie de datos, lo cual no siempre es posible [2]. Small y Judd demostraron que la dimensión de correlación es una estadística esencial [15], lo cual la hace ideal para ser utilizada en pruebas de hipótesis cuando se utilizan realizaciones limitadas o no limitadas.

Por este motivo, en el presente estudio se utiliza la dimensión de correlación calculada con el algoritmo de Judd [16] como estadística para el rechazo o aceptación de las hipótesis, pero además se utilizan la complejidad de Lempel – Ziv [2] y la información mutua promedio [14] para caracterizar la dinámica de cada serie temporal, la ventaja que estas estadísticas tienen sobre la dimensión de correlación es que no se requiere la construcción de un atractor.

3. RESULTADOS Y DISCUSIÓN

Antes de realizar una prueba estadística de hipótesis nulas es necesario definir el número total de sustitutos que se van a generar. Según Theiler y Prichard [4] el número total de sustitutos debe ser $M = (2/\alpha - 1)$, α es la probabilidad residual de falsos rechazos, y está relacionada con el nivel de significancia de la prueba por medio de $(1-\alpha) \times 100\%$. Entonces por ejemplo si se generan M sustitutos, el número total de series sería $M + 1 = 2/\alpha$, de este manera la probabilidad que se rechace la hipótesis cuando esta es cierta es $\alpha \times 100\%$ (en este caso rechazar la hipótesis se refiere a que la estadística de los datos es mayor o menor que el conjunto de estadísticas de los sustitutos, este tipo de prueba se conoce como *prueba de dos lados*). Para la presente investigación se trabajó con un nivel de significancia del 95%, es decir $\alpha=0,05$ y $M=39$ para cada algoritmo.

3.1. Variación de la temperatura

A continuación se analiza la temperatura promedio diaria en *Madison (WI, USA)*. En la figura 1 se muestra la señal original y uno de los sustitutos generados con cada uno de los algoritmos, es notable la similitud que tienen los sustitutos generados con el algoritmo 2. En la figura 2 se muestran los resultados obtenidos para $d_E=7$ y $\tau=81$, se

realizaron pruebas para $d_E=6-12$ y se obtuvieron resultados similares.

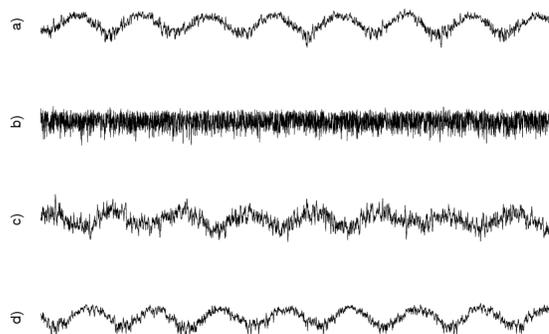


Figura 1. a) Temperatura promedio diaria. b) c) d) Sustitutos generados con los algoritmos 0,1 y 2 respectivamente.

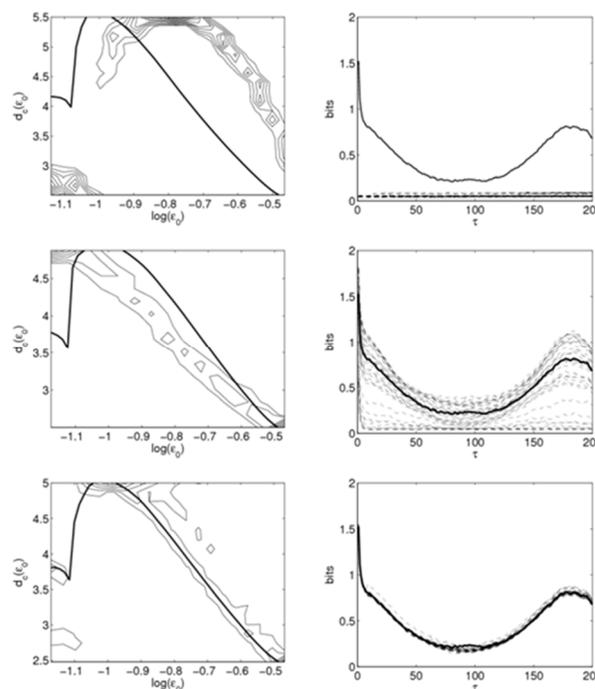


Fig 2. En la primera columna se presenta la curva de dimensión de correlación para los datos y los sustitutos calculados con algoritmo de Judd de la temperatura promedio diaria, para los sustitutos se presentan como una figura de contornos de la distribución de probabilidad. En la segunda columna se presenta la curva de la AMI para las 40 señales, para los datos se muestra como una línea gruesa. La primera fila es el algoritmo 0 y la tercera el algoritmo 2 (iAAFT).

	C_{min}	C_{max}
Alg0	1.0240	1.0597
Alg1	0.3508	0.8393
Alg2	0.2852	0.3727

Tabla 1. Valores máximos y mínimos obtenidos al calcular la complejidad de Lempel-Ziv a los sustitutos generados con los algoritmo 0, 1 y 2 para la temperatura promedio diaria.

En la primera columna de la figura 2 se tienen las curvas de dimensión de correlación para cada uno de los algoritmos, para los 39 sustitutos se presentan las distribuciones de probabilidad como figuras de contorno mientras que para los datos se presenta la curva como una línea gruesa. En la columna derecha se tienen las curvas de AMI. Para la serie original la complejidad de Lempel-Ziv es $C = 0.3500$, para los sustitutos generados con los diferentes algoritmos se obtuvieron los resultados mostrados en la tabla 1.

La hipótesis 0 se rechaza con las tres estadísticas, es notable que la complejidad de Lempel-Ziv es $C \approx 1$ para todos los sustitutos generados con el algoritmo 0, lo cual concuerda con ruido i.i.d, también se observa que la AMI para los sustitutos es una línea muy cercana a 0, la hipótesis 1 fue rechazada al utilizar la dimensión de correlación como estadística, lo mismo ocurre cuando se calcula la complejidad de Lempel-Ziv. No ocurre lo mismo con las curvas de AMI, las cuales tienen formas desde las encontradas para el caso de ruido i.i.d hasta curvas similares a la de los datos. No se puede rechazar la hipótesis 2 con ninguna de las estadísticas utilizadas, no hay una notable diferencia entre el comportamiento de las curvas de AMI en este caso, además la curva de dimensión de correlación para todos los sustitutos es similar a la obtenida con los datos. La complejidad de Lempel-Ziv está distribuida en un rango pequeño y el valor calculado para los datos está contenido dentro de dicho rango.

3.2. Manchas solares

A continuación se analiza una señal del promedio mensual de manchas solares, el análisis de este tipo de señales es importante debido a la influencia que tienen las manchas solares sobre la calidad en las telecomunicaciones, por este y muchos otros motivos los científicos se han tomado la tarea de medir las manchas solares desde 1749, y esta actividad aún se realiza. La serie analizada cuenta con mediciones realizadas desde enero de 1794 hasta julio de 2009. En la figura 3 se observa la señal original y los sustitutos generados con los 4 algoritmos, para el algoritmo PPs se utilizó $d_E=6$, $\tau=35$ y $\rho=0.1$. En la figura 4 se muestran los resultados obtenidos para $d_E=7$ y $\tau=35$, se realizaron pruebas para $d_E=5-9$ y se obtuvieron resultados similares. Para la serie original la complejidad de Lempel-Ziv es $C = 0.4083$, para los sustitutos generados con los diferentes algoritmos se obtuvieron los resultados mostrados en la tabla 2.

El sustituto generado con el algoritmo PPs (figura 3.e)) es bastante similar a la señal original. Los demás sustitutos no son similares a la serie, lo cual hace pensar que la hipótesis correspondiente será rechazada. En la figura 4 se observa que la dimensión de correlación de los sustitutos generados con los 3 algoritmos estándar no

es consistente con la dimensión de correlación de los datos, lo mismo ocurre con las curvas de AMI y con los valores de la complejidad de Lempel-Ziv. Pero para los sustitutos generados con el algoritmo PPs lo anterior no

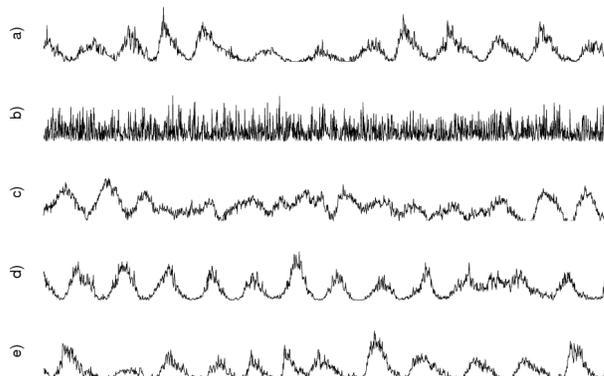


Figura 3. a) Promedio mensual de manchas solares. b) c) d) e) Sustitutos generados con los algoritmos 0, 1, 2 y PPs.

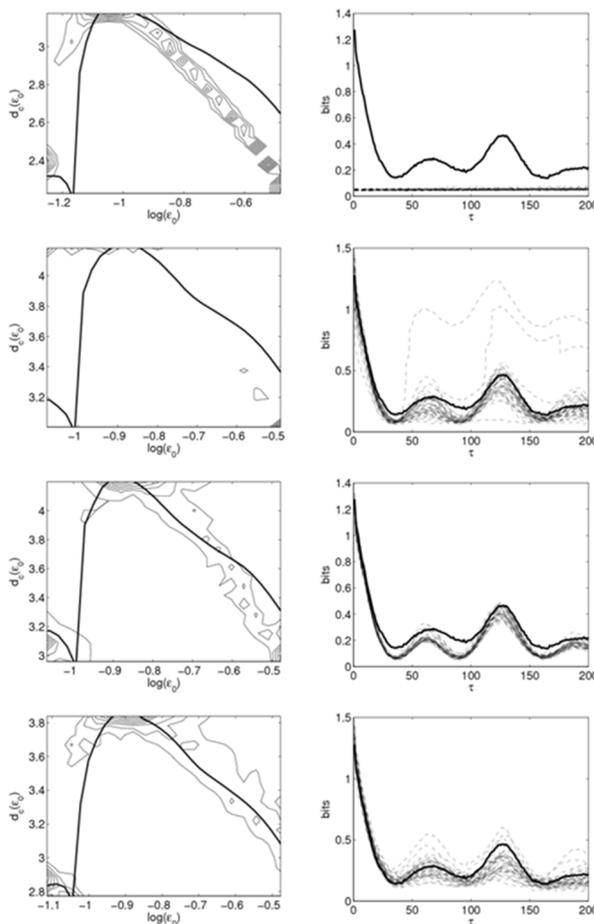


Figura 4. En la primera columna se presenta la curva de dimensión de correlación para los datos y los sustitutos calculados con algoritmo de Judd del promedio mensual de manchas solares, para los sustitutos se presentan como una figura de contornos de la distribución de probabilidad. En la segunda columna se presenta la curva de la AMI para las 40 señales, para los datos se muestra como una línea gruesa. La primera fila es el algoritmo 0 y la cuarta el algoritmo PPs.

	C_{min}	C_{max}
Alg0	1.0282	1.0616
Alg1	0.4321	0.6125
Alg2	0.4157	0.5382
AlgPPs	0.3742	0.4123

Tabla 2. Valores máximos y mínimos obtenidos al calcular la complejidad de Lempel-Ziv a los sustitutos generados con los algoritmo 0, 1, 2 y PPs para el valor promedio de las manchas solares.

ocurre, es decir que no es posible rechazar la hipótesis nula con ninguna de las estadísticas, esto concuerda con la literatura [3,16].

3.3. Intensidad de un laser de NH3

A continuación se analiza la intensidad de las pulsaciones de un laser infrarrojo-lejano NH3. Para el algoritmo PPs se utilizó $d_E=7$, $\tau=2$ y $\rho=0,5$. En la figura 5 se muestran los resultados obtenidos para $d_E=7$ y $\tau=2$, se realizaron pruebas para $d_E=6-9$ y se obtuvieron resultados similares. Para la serie original la complejidad de Lempel-Ziv es $C = 0.2169$, para los sustitutos generados con los diferentes algoritmos se obtuvieron los resultados mostrados en la tabla 3.

En la figura 6 se observa la dimensión de correlación tanto para la serie original como para los 39 sustitutos, se nota la marcada diferencia entre el valor de $d_c(\epsilon_0)$ para los datos y los sustitutos, por lo tanto se rechaza la hipótesis. También se observa que por medio de las curvas de AMI es posible rechazar las hipótesis. Algo similar ocurre para los 3 algoritmos estándar, es decir, fue posible rechazar las 4 hipótesis. Lo mismo se puede verificar utilizando la medida de complejidad de Lempel-Ziv, el valor calculado para la serie original es menor que el calculado para los sustitutos generados con los 4 algoritmos. Debido a esto se procede a calcular el máximo exponente de Lyapunov, y se obtiene que para esta serie $\lambda_{max} = 0,08$. Es por lo tanto posible asegurar con cierta certeza que como los datos no son generados por un proceso lineal estocástico (o una transformación no lineal de éste) y dado que el máximo exponente de Lyapunov es positivo, entonces los datos son generados por un sistema dinámico caótico, lo cual concuerda con la literatura [17].

El uso de la información mutua promedio como estadística para prueba de hipótesis no es común, de hecho sólo se pudo encontrar una referencia donde se usa [12], en este artículo se usa esta estadística no lineal con éxito, algo que es importante resaltar es que la AMI por su definición depende del orden de los datos, ya que esta mide de manera no lineal la correlación entre un valor $z(n)$ y un valor $z(n + \tau)$, $\forall n$. Por lo tanto esta no es una buena estadística para el algoritmo 0, pero como se pudo observar cuando se tiene ruido i.i.d la curva de AMI siempre tiene un mismo comportamiento, es una línea

cerca a cero, por lo tanto se puede usar para diferenciar entre ruido i.i.d (los sustitutos generados con el algoritmo 0) y otro tipo de procesos. Se pudo observar que la curva de AMI para los sustitutos siempre (o casi siempre) se parece a la curva de AMI que se obtuvo para los datos originales (a excepción de los sustitutos generados con el algoritmo 0), esto a pesar que cada conjunto de sustitutos originales (a excepción de los sustitutos generados con el algoritmo 0) no siempre era satisfecha por el sistema dinámico que generó los datos. Gracias a esto se puede afirmar que la forma que tiene la curva de AMI no depende del proceso subyacente que generó los datos como se ha llegado a suponer [18].

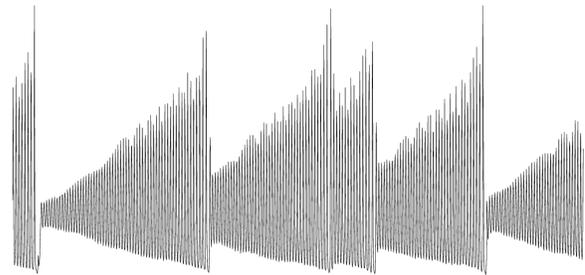


Figura 5. Intensidad de un laser NH3

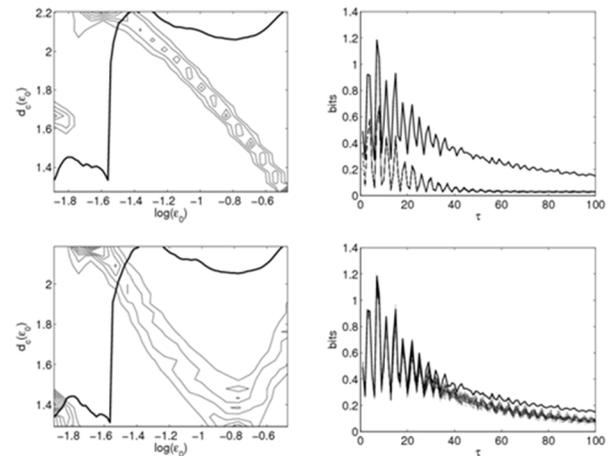


Figura 6. En la primera columna se presenta la curva de dimensión de correlación para los datos y los sustitutos calculados con algoritmo de Judd del promedio mensual de manchas solares, para los sustitutos se presentan como una figura de contornos de la distribución de probabilidad. En la segunda columna se presenta la curva de la AMI para las 40 señales, para los datos se muestra como una línea gruesa. La primera fila es el algoritmo 2 (iAAFT) y la segunda el algoritmo PPs.

	C_{min}	C_{max}
Alg0	1.0255	1.0430
Alg1	0.6479	0.7217
Alg2	0.5626	0.5930
AlgPPs	0.3283	0.3818

Tabla 3. Valores máximos y mínimos obtenidos al calcular la complejidad de Lempel-Ziv a los sustitutos generados con los algoritmo 0, 1, 2 y PPs para la intensidad del laser NH3.

4. CONCLUSIONES

El método de los datos sustitutos permite por medio del uso de estadística no lineales validar o rechazar declaraciones de no linealidad en series temporales, en este documento se revisaron las tres hipótesis clásicas las cuales permiten diferenciar entre sistemas lineales estocásticos, transformaciones no lineales de sistemas lineales estocásticos y sistemas no lineales, se mostró también una hipótesis desarrollada recientemente la cual permite discriminar entre sistemas caóticos y sistemas periódicos contaminados con ruido.

Durante las pruebas experimentales se utilizaron tres estadísticas (dimensión de correlación calculada con el algoritmo de Judd, la curva de AMI y la complejidad de Lempel-Ziv), se encontró que con un uso apropiado de la complejidad de Lempel-Ziv es posible llegar a las mismas conclusiones que se llegaron con la curva de dimensión de correlación (la conclusión obtenida por medio de la dimensión de correlación se usa como referencia dado el gran número de resultados positivos obtenidos en la diversa literatura, y los hallazgos encontrados en [15]), pero se tiene la ventaja que no se requiere la construcción de un atractor lo cual no siempre es posible. Se encontró además que la curva de AMI no depende del proceso subyacente que generó los datos, esto se observa debido a que a pesar de que la serie original y los sustitutos satisfacen hipótesis diferentes (comprobado con las demás estadísticas), la curva del AMI es muy similar en ambos casos, esto ocurre particularmente para los sustitutos generados con el algoritmo 2 y el algoritmo PPS.

De las bases de datos se puede decir que la media diaria de temperatura se comporta como una transformación monótona no lineal de ruido linealmente correlacionado, es decir el proceso que genera los datos es un proceso estocástico linealmente correlacionado, sin embargo hay una función de observación $h()$ que es no lineal, por lo tanto la serie es no lineal pero debido al proceso de observación. En cuando al promedio mensual de manchas solares se pudo probar que es una realización de una órbita periódica contaminada por ruido, es decir el sistema dinámico es no lineal pero no es caótico. Para la intensidad del Laser NH3 se rechazaron las 4 hipótesis, y dado que está establecido que el máximo exponente de Lyapunov de esta serie es positivo, se puede afirmar que el sistema dinámico es caótico.

5. AGRADECIMIENTOS

D. L. Guarín es patrocinado por la Universidad Tecnológica de Pereira y COLCIENCIAS, programa “Jóvenes investigadores e innovadores 2011”.

6. BIBLIOGRAFÍA

- [1] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank, and J. D. Farmer, “Testing for nonlinearity in time series: The method of surrogate data”, *Physica D*, vol. 58, pp.77–94, 1992.
- [2] M. Small, *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*. Singapore: World Scientific, 2005.
- [3] M. Small, D. Yu, and R. Harrison, “Surrogate test for pseudoperiodic time series data”, *Physical Review Letters*, vol. 87, no. 18, pp. 1–4, 2001.
- [4] J. Theiler and D. Prichard, “Constrained-realization Monte-Carlo method for hypothesis testing”, *Physica D*, vol. 94, pp. 221–235, 1995.
- [5] D. L. Guarín, C. H. Rodríguez and Á. Á. Orozco, “Pruebas de no linealidad: el método de los datos sustitutos”, *Scientia et Technica*, vol. 44, pp. 292–297, 2010.
- [6] T. Schreiber and A. Schmitz, “Surrogate time series”, *Physica D*, vol. 142, pp. 346–382, 2000.
- [7] A. I. Mees, Ed., *Nonlinear Dynamics and Statistics*, Boston: Birkhauser, 2001.
- [8] M. Small and C. K. Tse, “Detecting determinism in time series the method of surrogate data”, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, no. 5, pp. 663–672, 2003.
- [9] M. D. Nurujjaman, “Nonlinear dynamics experiments in glow discharge plasma”. Ph.D. dissertation, Saha Institute of Nuclear Physics, Kolkata, India, 2009.
- [10] C. W. Wang, Ed., *Nonlinear Phenomena Research Perspectives*, New York: Nova Science, 2007.
- [11] Y. Zhao, J. Sun, and M. Small, “Evidence consistent with deterministic chaos in human cardiac data: surrogate and nonlinear dynamical modeling”, *International Journal of Bifurcation y Chaos*, vol. 18, no. 1, pp. 141 – 160, 2008.
- [12] M. Small and C. Tse, “Applying the method of surrogate data to cyclic time series”, *Physica D*, vol. 164, pp. 187–201, 2002.
- [13] F. Takens, “Detecting strange attractors in turbulence”, *Lecture Notes in Math*, vol. 898, pp. 366–381, 1981.
- [14] H. D. I. Abarbanel, *Analysis of observed chaotic data*, 1st Ed., Springer, 1996.
- [15] M. Small and K. Judd, “Correlation dimension: A pivotal statistic for non-constrained realizations of composite hypotheses in surrogate data analysis”, *Physica D*, vol. 120, pp. 386–400, 1998.
- [16] K. Judd, “Estimating dimension from small samples”, *Physica D*, vol. 71, no. 509–518, 1994.
- [17] H. D. I. Abarbanel, Z. Gills, C. Liu and R. Roy, “Nonlinear-time-series analysis of chaotic laser dynamics”, *Phys. Rev. A*, vol. 53, pp. 440 – 453, 1996.
- [18] N. Okamura, H. Okamoto and S. Tomii, “A New Classification of Strange Attractors of Chaos from Mutual Information”, in *Chaos Memorial Symposium in Asuka : selected papers dedicated to professor Yoshisuke Ueda on the occasion of his 60th birthday*, pp 79 – 84, 1997.