

RECONSTRUCCIÓN DE DINÁMICAS CAÓTICAS Y/O RUIDOSAS USANDO REDES NEURONALES

H. D. Navone, H. A. Ceccatto

*Instituto de Física Rosario, Universidad Nacional de Rosario y CONICET,
Boulevard. 27 de Febrero 210 Bis, (2000) Rosario, República Argentina.*

Se muestra que las redes neuronales son capaces de aprender la ley dinámica de series temporales caóticas y/o ruidosas a partir de un número limitado de observaciones. La capacidad predictiva de las mismas supera la de los métodos convencionales tales como el algoritmo no-lineal wimplex, o los modelos autorregresivos lineales. Como ejemplos de aplicación a series caóticas se reconstruyen las dinámicas del mapa logístico -el cual presenta una autocorrelación despreciable- y del mapa de Lorenz -que posee una significativa autocorrelación-. A diferencia de lo sugerido recientemente en la literatura, en ambos casos las predicciones a corto plazo de las redes con arquitecturas adecuadas son mejores que las de los métodos mencionados. Finalmente, se presentan los resultados de la aplicación de redes neuronales en el análisis de series temporales reales ruidosas.

Recientemente se han desarrollado distintas técnicas de predicción de series temporales basadas en el carácter determinístico de sistemas complejos¹⁻³. Tales métodos tratan de capturar la estructura de corto alcance de las series temporales, expresando el valor presente x_t como una función de d valores previos de la propia serie,

$$x_t = F(x_{t-1}, x_{t-2}, \dots, x_{t-d})$$

Los vectores $(x_{t-1}, x_{t-2}, \dots, x_{t-d})$ definen un espacio de (pseudo-) fases d -dimensional. El problema de la predicción -extrapolación temporal- es entonces reformulado como una interpolación en dicho espacio. Elsner y Tsonis³ han presentado una breve reseña de algunas de estas nuevas metodologías. En particular, enfatizaron el éxito de las redes neuronales en la realización de predicciones a corto plazo, y aún en la distinción entre caos y ruido⁴. Usando datos generados computacionalmente y observaciones experimentales, mostraron las ventajas de esta técnica sobre los modelos autorregresivos lineales (AR) en el análisis de series temporales. Sin embargo, más recientemente dichos autores reconsideraron sus resultados, concluyendo que las redes neuronales presentan una mejor performance que los modelos AR sólo para series temporales con baja autocorrelación⁵.

En este trabajo mostramos que las redes neuronales entrenadas apropiadamente superan a los modelos AR lineales en la predicción de series temporales con comportamiento caótico, independientemente del grado de autocorrelación. Más aún, mostramos que la capacidad predictiva de esta técnica resulta superior a la del algoritmo no-lineal wimplex (d)⁶, el cual es una simplificación propuesta recientemente del método desarrollado por Sugihara y

May⁷. En primer lugar comparamos los métodos arriba mencionados aplicándolos a dos conjuntos de datos generados computacionalmente: la ecuación logística -que presenta una autocorrelación despreciable-, y el mapa de Lorenz -el cual posee una considerable autocorrelación- (Fig. 1). En la última parte de este trabajo comentamos brevemente los resultados obtenidos para series temporales reales.

La ecuación logística está dada por

$$x_{t+1} = ax_t(1-x_t)$$

A los efectos de comparar nuestros resultados con los de Elsner y Tsonis, al igual que ellos⁵ se tomó $a = 3.8$ y $x_0 = 0.23489$. Se generaron 1000 valores iterando la ecuación anterior, empleándose los primeros 100 valores para obtener los modelos AR y wimplex óptimos (se usó un pequeño número de iteraciones ya que en la mayoría de las aplicaciones reales la cantidad de datos es escasa). El modelo AR óptimo resultó ser de orden 6 (AR(6)), mientras que el mejor wimplex correspondió a un espacio de pseudofase con $d = 1$. Los 900 valores no usados en el ajuste de estos modelos fueron predichos por ambos métodos. A los efectos de evaluar la performance de los mismos se calculó el coeficiente de correlación (CC) entre los valores predichos y actuales (objetivos), y la variancia relativa promedio (ARV), definida como

$$ARV = \frac{\sum_{k \in S} (\text{objetivo}_k - \text{predicción}_k)^2}{\sum_{k \in S} (\text{objetivo}_k - \text{media})^2}$$

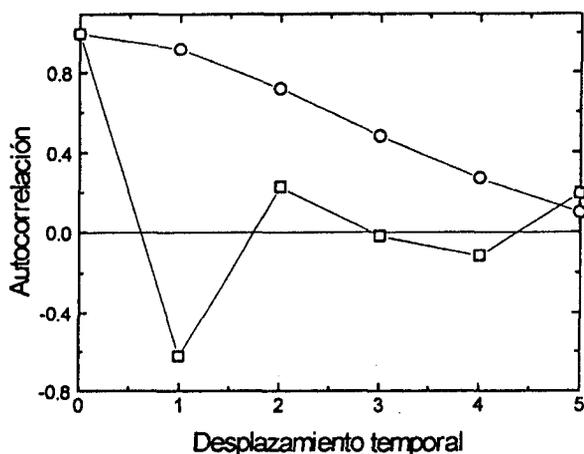
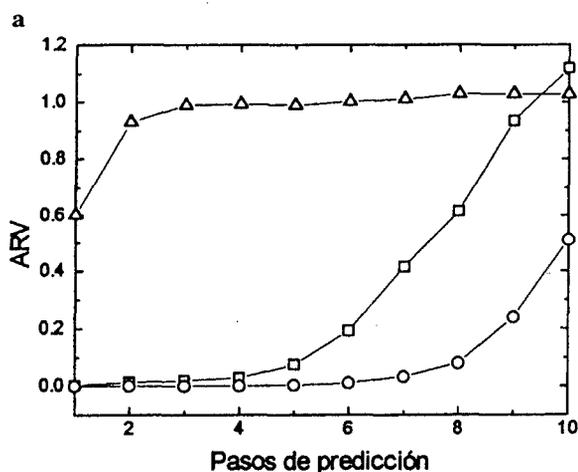


Fig. 1: Autocorrelación correspondiente a la ecuación logística (cuadrados) y al mapa de Lorenz (círculos).

donde S es el conjunto de 900 valores y "media" es su promedio. Un valor 1 para ARV significa que el método sólo es capaz de predecir el promedio de la serie temporal. Estos dos índices se grafican en la Fig. 2a,b. Se observa que el algoritmo no-lineal wimplex supera claramente al modelo AR. En estas figuras también se presentan los resultados obtenidos usando una red de arquitectura 2:4:1 (dos unidades de entrada, 4 unidades ocultas, y una unidad de salida). Después de varias miles de épocas de entrenamiento, la red neuronal es capaz de capturar la dinámica de corto alcance de la serie temporal y superar a los otros métodos. En particular, la red produce resultados con errores despreciables hasta siete pasos en el futuro, y luego muestra un rápido deterioro en su habilidad predictiva debido a la



naturaleza caótica de la serie temporal. Es interesante destacar que los resultados no cambian considerablemente para otras arquitecturas tales como 2:3:1 y 3:3:1. Incrementando el número de datos de 100 a 500 para el ajuste de los modelos, la performance de la red neuronal y del modelo AR no cambian demasiado, mientras, como era de esperar, el algoritmo wimplex mejora considerablemente alcanzando la performance de la red neuronal.

En el caso del mapa de Lorenz,

$$\frac{dx}{dt} = -ax + ay$$

$$\frac{dy}{dt} = -xz + bx - y$$

$$\frac{dz}{dt} = xy - cz$$

tomamos $a = 16.0$, $b = 120.1$ y $c = 4.0$, que al igual que antes corresponden a los valores usados por Elsner y Tsonis³. Integrando el sistema de ecuaciones mediante el método de Runge-Kutta de cuarto orden con un paso temporal de 0.03, se generaron nuevamente 1000 valores. En este caso la serie temporal presenta una considerable autocorrelación, como se muestra en la Fig. 1. Nuevamente los primeros 100 valores se usaron para ajustar los modelos AR y wimplex óptimos, los cuales en este caso resultaron ser el AR(12) y wimplex(6). Sus performances se muestran en las Figs. 3a,b. En las mismas figuras se grafican los resultados obtenidos con la red 6:4:1, cuya performance supera a la de los otros métodos. Es interesante destacar que debido a la mayor autocorrelación, la capacidad predictiva del modelo lineal AR(12) es mejor que en el caso de la ecuación

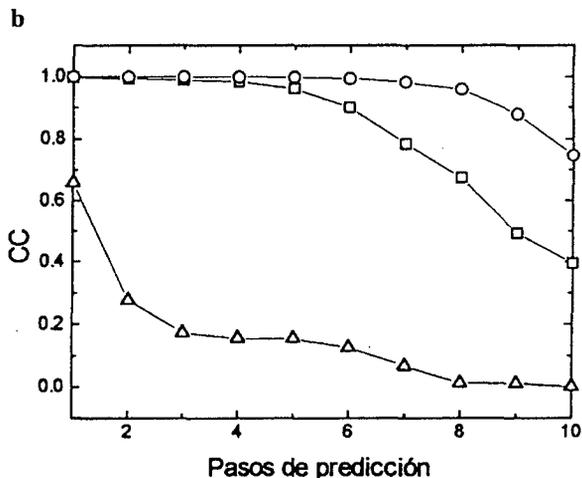


Fig. 2: a) Variación relativa promedio para la ecuación logística y los diferentes modelos predictivos: AR(6) (triángulos), wimplex(1) (cuadrados), y red neuronal 2:4:1 (círculos), y b) Coeficientes de correlación entre los valores observados y predichos.

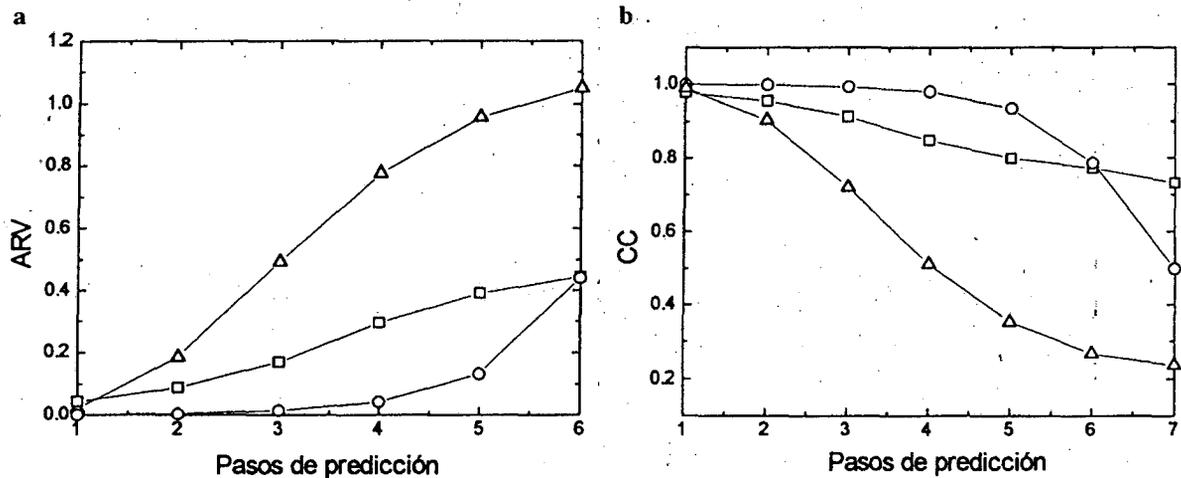


Fig. 3. a) Variancia relativa promedio para el mapa de Lorenz y los diferentes métodos predictivos: AR(12) (triángulos), wimplex(6) (cuadrados), y red neuronal 6:4:1 (círculos), y b) Coeficientes de correlación entre los valores observados y predichos.

logística. Estos resultados no dependen significativamente de la arquitectura de la red. Por otro lado, para un mayor número de datos de ajuste (500) el modelo wimplex(6) mejora su performance y los coeficientes CC y ARV correspondientes se acercan a los de la red neuronal.

A continuación se discuten los resultados obtenidos en el análisis de series temporales reales. En particular, se consideraron las predicciones de los diferentes métodos correspondientes a las series temporales de: i) la actividad solar (medida a través del número relativo de Wolf)^{8,9}, y ii) el total de lluvia producida durante los monsoones de verano en la India¹⁰. Ambos registros son usualmente considerados como bancos de prueba en la literatura sobre series temporales. Si bien en estos casos la dinámica intrínseca puede ser no-caótica, las periodicidades en los registros no son evidentes y están ocultas por ruido. Se encontró que en ambos casos las redes neuronales presentan una capacidad predictiva mucho mayor que los algoritmos AR y wimplex. Las performances de los distintos métodos en la predicción de la actividad solar durante el período 1870-1889 (en el cual los métodos convencionales presentan una alta capacidad predictiva) son:

AR(3)	ARV=0.68	CC=0.92
Wimplex(12)	ARV=0.40	CC=0.92
Red Neuronal 12:3:1	ARV=0.09	CC=0.98

Para la predicción de precipitaciones producidas por los monsoones de verano en la India se consideró el período 1969-1984, el cual presenta también una gran predictabilidad en el análisis con métodos convencionales. En este caso los resultados obtenidos son los

siguientes:

AR(7)	ARV=0.94	CC=0.25
Wimplex(7)	ARV=0.68	CC=0.72
Red Neuronal 7:4:1	ARV=0.25	CC=0.87

A partir de estos resultados se puede ver que las redes neuronales son particularmente adecuadas para el estudio de series temporales ruidosas. Los detalles de estos estudios pueden hallarse en⁹⁻¹⁰. Finalmente, es importante remarcar que la metodología basada en redes neuronales es susceptible de posteriores mejoras a partir de la modificación de las arquitecturas empleadas, mayores períodos de entrenamiento, etc.

REFERENCIAS

1. J. D. Farmer y J. J. Sidorowich, Phys. Rev. Lett. 59, 845-848 (1987)
2. J. D. Farmer y J. J. Sidorowich, Los Alamos National Laboratory, LA-UR-88-901 (1988).
3. J. B. Elsner y A. A. Tsonis, Bulletin American Meteorological Society 73, 49-60 (1992).
4. J. B. Elsner, J. Phys. A: Math. Gen. 25, 843-850 (1992).
5. J. B. Elsner y A. A. Tsonis, Bulletin American Meteorological Society 74, 243 (1993).
6. A. D. Fowler y D. E. Roach, Computers & Geosciences 19, 41-52 (1993).
7. G. Sugihara y R. M. May, Nature 344, 734-741 (1990).
8. A. S. Weigend, B. A. Huberman y David E. Rumelhart, International Journal of Neural Systems 1, 193-209 (1990).
9. R. A. Calvo, H. A. Ceccatto y R. D. Piacentini, The Astrophysical Journal (en prensa).
10. H. D. Navone y H. A. Ceccatto, Climate Dynamics (en prensa).