

Estudio comparativo de métodos no-lineales para el análisis de series temporales caóticas

H.D. NAVONE y H.A. CECCATTO*

IFIR (UNIVERSIDAD NACIONAL DE ROSARIO - CONICET)
BV. 27 DE FEBRERO 210 BIS - (2000) ROSARIO - ARGENTINA
e-mail: ceccatto@ifir.edu.ar

La complejidad de los sistemas naturales hace muchas veces dificultosa la construcción de modelos teóricos adecuados que describan su evolución dinámica. Es por ello que muchos de los esfuerzos actuales están dirigidos al desarrollo de metodologías que permitan predecir el comportamiento futuro de un sistema a partir de su historia. En el caso de sistemas caóticos, para esta reconstrucción dinámica es necesario recurrir a métodos no lineales de análisis de series temporales. En este trabajo se comparan las respuestas de algunos de estos métodos en la predicción de series caóticas generadas numéricamente. En particular se considera la situación en que los datos disponibles son escasos, lo que generalmente ocurre en las aplicaciones concretas.

The complexity of real-world systems often makes impossible to model their dynamical evolution. For this reason, currently many efforts are devoted to predict a complex system's future behavior from its database record. In particular, for chaotic systems this dynamical reconstruction requires nonlinear methods of time-series analysis. In this work we compare the performances of some of these methods on chaotic series generated numerically. We consider the particular situation of scarce data, which is frequently faced in practical applications.

En este trabajo comparamos las performances de distintos algoritmos no-lineales en la predicción de series temporales caóticas en el caso en que los datos disponibles son escasos, lo que generalmente ocurre en la práctica. En particular, consideraremos dos métodos que han sido reconocidos como universales, simples y confiables. El primero constituye un ajuste global del atractor caótico utilizando redes neuronales^{1,2}; el segundo, conocido como Wimplex³, tiene un carácter local y pertenece a la clase de métodos que hacen uso de situaciones similares a la actual en la historia del sistema para predecir el comportamiento futuro del mismo ("instance-based methods").

Consideraremos además la introducción de correcciones lineales al algoritmo Wimplex, en un esquema que llamaremos Aproximación de Hiperplano Local (AHL). De acuerdo con este esquema, las predicciones del Wimplex serán corregidas ajustando un hiperplano en las proximidades del estado actual del sistema.

Los ejemplos sobre los que aplicaremos los algoritmos arriba mencionados son las series caóticas correspondientes al mapa logístico,

$$x_{t+1} = r x_t (1 - x_t) \quad (1)$$

con $r = 3.8$, y al sistema de Lorenz,

$$\begin{aligned} \frac{dx}{dt} &= -ax + ay \\ \frac{dy}{dt} &= -xz + bx - y \\ \frac{dz}{dt} &= xy - cz \end{aligned} \quad (2)$$

con $a = 16.0$, $b = 120.1$, y $c = 4.0$. La elección de constantes es la misma que en [1]. Las series así generadas constituyen un conjunto de datos "limpios", donde los distintos métodos pueden ser chequeados sin las complicaciones propias de

* Autor a quién debe dirigirse la correspondencia.

la observación real de un sistema (errores experimentales, ruido, etc.).

El problema a resolver es la reconstrucción de la dinámica intrínseca de las series caóticas generadas por (1) y (2):

$$x_t = F(\mathbf{X}_t) \quad (3)$$

con $\mathbf{X}_t = (x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-d\tau})$ un vector en el espacio de pseudofases donde la reconstrucción se lleva a cabo⁴. Tanto la dimensión d de este espacio como la función F deben ser obtenidas del registro histórico del comportamiento del sistema. Un método sencillo para realizar dicha tarea es el llamado Wimplex ($W(d,n)$)³. Según este método, uno busca los n vectores \mathbf{X}_i ($i=1,n$) en los datos que sean más próximos al actual, \mathbf{X}_b , y predice

$$x_t = \bar{x}_i \equiv \sum_{i=1}^n x_i P(\mathbf{X}_t - \mathbf{X}_i)$$

donde $x_i = F(\mathbf{X}_i)$. La función de peso se toma usualmente $P(\mathbf{X}_t - \mathbf{X}_i) \propto 1/d(\mathbf{X}_t - \mathbf{X}_i)$, con $d(\mathbf{X}_t - \mathbf{X}_i)$ la distancia euclídea entre \mathbf{X}_t y \mathbf{X}_i .

Otro método frecuentemente utilizado para reconstruir series temporales caóticas está basado en las estructuras computacionales conocidas como redes neuronales^{1,2} (ver [5] para una introducción general al tema). En este trabajo usaremos redes del tipo "feedforward", con d neuronas en la capa de entrada, n neuronas en la capa oculta, y una neurona en la capa de salida ($NN(d,n)$). Las interconexiones entre neuronas -pesos- se eligen de manera que la red "aprenda" a relacionar entradas y salidas de acuerdo con la función (3), para lo cual se hace uso de los ejemplos contenidos en la base de datos (entrenamiento por retropropagación de errores⁵).

Además de los métodos mencionados consideraremos una corrección lineal al algoritmo Wimplex que llamamos Aproximación de Hiperplano Local (AHL(d)). En la implementación del Wimplex discutida arriba, si los vectores en el registro histórico cubren razonablemente bien el atractor se puede aproximar

$$x_t = F(\mathbf{X}_t) \cong F(\mathbf{X}_i) + \mathbf{A}_i \bullet (\mathbf{X}_t - \mathbf{X}_i) \quad (4)$$

con $\mathbf{A}_i = \nabla F(\mathbf{X}_i)$. Multiplicando estas ecuaciones por alguna función peso preestablecida $P(\mathbf{X}_t - \mathbf{X}_i)$, y sumando todas ellas se obtiene la predicción

$$x_t = F(\mathbf{X}_t) \cong \bar{x}_i + \mathbf{A}_i \bullet (\mathbf{X}_t - \bar{\mathbf{X}}_i) \quad (5)$$

Para hallar las d componentes de \mathbf{A}_i basta con eliminar la n -ésima ecuación en (4) y elegir $n-1=d$, lo que reduce el problema a resolver un sistema lineal cuadrado. De esta manera, la estrategia que lleva a (5) equivale a predecir x_i como el valor correspondiente a \mathbf{X}_i sobre el plano local óptimo que contiene al punto $(\bar{\mathbf{X}}_i, \bar{x}_i)$.

En la práctica generalmente los datos son escasos y la aproximación (4) no siempre es válida. En tal caso las correcciones lineales son algunas veces demasiado grandes, y es preferible retener las predicciones del Wimplex. Hemos establecido una cota superior,

$$|\mathbf{A}_i \bullet (\mathbf{X}_t - \bar{\mathbf{X}}_i)| \leq \alpha x_{max}$$

para discriminar en que caso deben retenerse las correcciones lineales. Esta cota se expresa como una fracción α del valor más grande x_{max} en el registro (el valor de α se elige en la forma explicada más abajo en las aplicaciones).

A fin de comparar las performances de los distintos métodos en predecir el comportamiento futuro de las series caóticas (1) y (2) hemos seguido los siguientes pasos: i) generamos $2N$ iteraciones, ii) dividimos cada registro en dos conjuntos, de manera de utilizar las primeras N iteraciones como datos y la segunda mitad como intervalo de validación, y iii) usando la información contenida en el conjunto de datos realizamos predicciones a m pasos de los registros en el intervalo de validación. Estimamos la habilidad de cada método en predecir el comportamiento futuro de las series por medio de la varianza relativa promedio

$$ARV(m, S) = \frac{\sum_{i \in S} (x_i - f_i(m))^2}{\sum_{i \in S} (x_i - \bar{x})^2}$$

En esta expresión S puede ser alternativamente el conjunto de datos ($i=1,N$) o de validación ($i=N+1,2N$), $f_i(m)$ es la predicción a m pasos de x_i , y \bar{x} es el valor medio de x en S . Una predicción perfecta corresponde a $ARV=0$, mientras que una predicción constante e igual al promedio de la serie da $ARV=1$. Una vez obtenidos los resultados según los pasos descriptos, se repitieron todos los cálculos partiendo de un registro inicial más largo conteniendo $4N$ iteraciones.

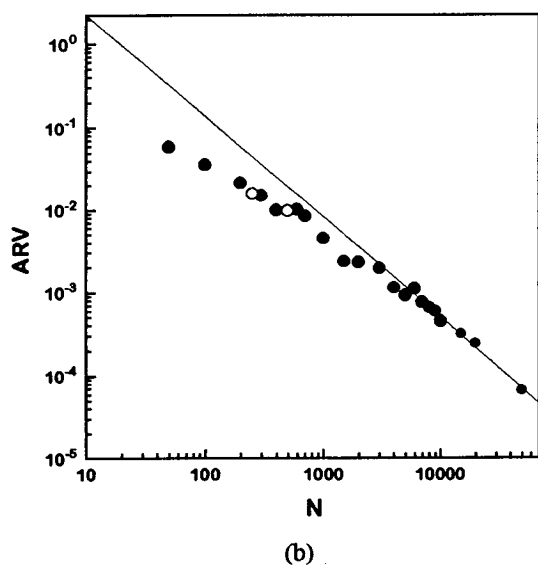
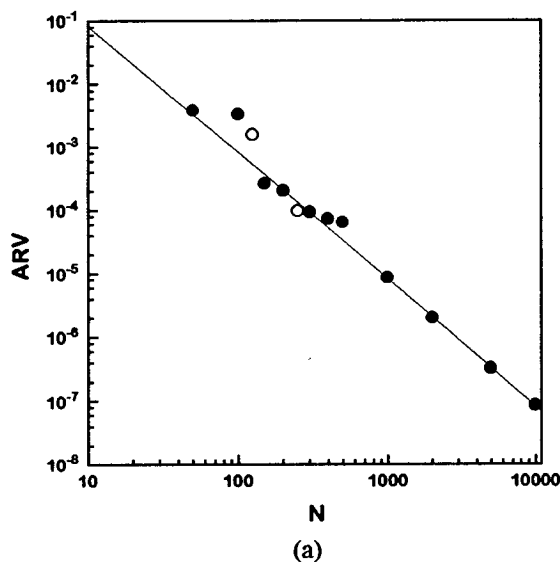


Figura 1. Error relativo ARV para predicciones a un paso como función del número de datos N , usando el método $Wimplex(d,n)$. a) Mapa Logístico con $d=1, n=2$; b) Sistema de Lorenz con $d=4, n=1$. Los círculos corresponden a los valores de N usados en este trabajo

Como fuera establecido anteriormente, nuestro interés es comparar las performances de los distintos métodos cuando los datos son escasos. Para caracterizar dicha situación, estudiamos el comportamiento de ARV correspondiente al método $Wimplex$ como función del número de datos N en el registro histórico (Fig. 1). Para un gran número de datos el comportamiento esperado es $\log ARV \sim \log N^6$; en el estudio que sigue usaremos valores de N para los cuales el error se aparta de esta ley de escala, indicando que los puntos no cubren uniformemente el atractor.

Presentamos primero los resultados para el mapa logístico (1). Inicialmente tomamos $N=125$, y predijimos el conjunto de validación utilizando los distintos métodos discutidos arriba. En este caso los errores de predicción siempre decrecen con α (Fig. 2), de manera que mantuvimos todas las correcciones lineales AHL.

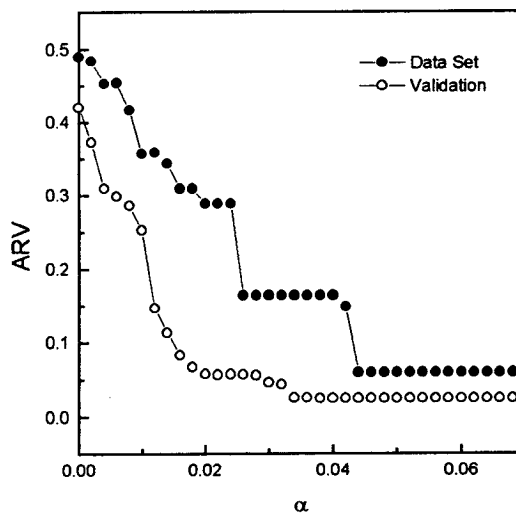


Figura 2. Error relativo ARV en predicciones a 8 pasos para el mapa logístico, como función del tamaño máximo permitido de las correcciones lineales. Resultados correspondientes a $N=125$ datos.

En la Fig. 3 presentamos los mejores resultados obtenidos, los que requieren jugar con los parámetros n y d en los distintos métodos para optimizar su elección. De acuerdo con estos resultados, la AHL resulta claramente el método más eficiente. Duplicando el largo del registro se obtienen mejores resultados, pero no hay cambios cualitativos con respecto a lo mostrado en las Figs. 2 y 3.

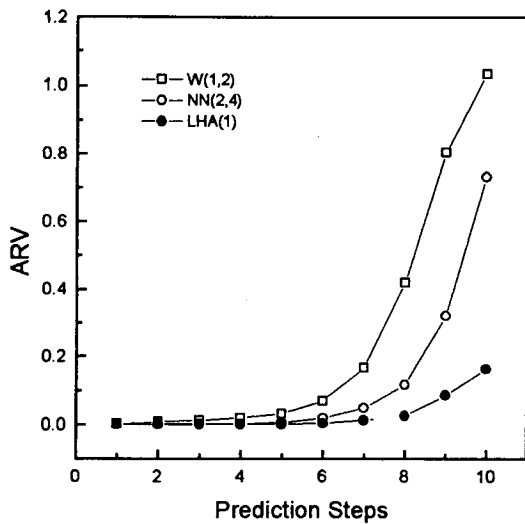


Figura 3. Error relativo ARV en el intervalo de validación para el mapa logístico, como función del número de pasos de predicción. Resultados correspondientes a $N=125$ datos.

Para el sistema de Lorenz (2) elegimos $N=250$, y los correspondientes resultados se muestran en las Figs. 4 y 5.

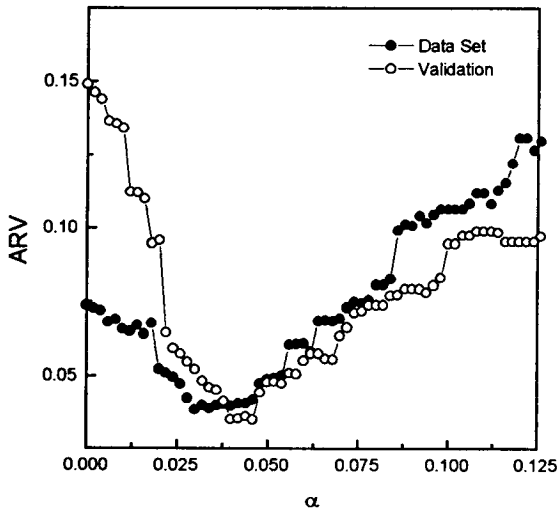


Figura 4. Error relativo ARV en predicciones a 5 pasos para el sistema de Lorenz, como función del tamaño máximo permitido de las correcciones lineales. Resultados correspondientes a $N=250$ datos.

Debido a la estructura del atractor de Lorenz este número de datos resulta comparativamente más escaso que los usados para el mapa logístico. Como consecuencia, hay un valor óptimo de α que puede estimarse hallando el mínimo error en la predicción del propio conjunto de datos (Fig. 4). En este caso

nuevamente la AHL resulta el mejor método de predicción.

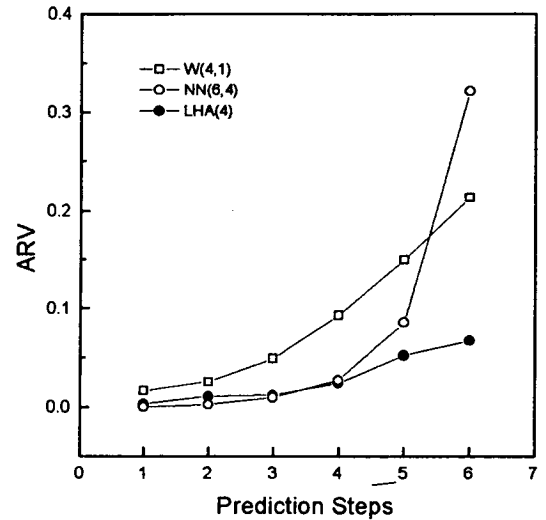


Figura 5. Error relativo ARV en el intervalo de validación para el sistema de Lorenz, como función del número de pasos de predicción. Resultados correspondientes a $N=250$ datos.

En conclusión, hemos propuesto la AHL como un algoritmo simple para predecir series temporales caóticas. Mostramos que es muy eficiente, aún en el caso de registros cortos, superando otros métodos más complejos frecuentemente usados en la literatura. La aplicación de la AHL a fenómenos reales requiere previamente la investigación de los efectos del ruido sobre dicho algoritmo.

Agradecimientos

HAC agradece a la Fundación Antorchas por el financiamiento parcial de este trabajo.

Referencias

- 1 - Elsner J.B. and Tsonis A.A., *Bull. Amer. Meteor.* **73**, 49 (1992); *Corrigendum Ibid* **74**, 243 (1993).
- 2 - Weigend A.S., Huberman B.A. and Rumelhart D.E., *Int. J. Neural Systems* **1**, 193 (1990).
- 3 - Fowler A.D. and Roach D.E., *Computers & Geosciences* **19**, 41 (1993).
- 4 - West B.J., "An Essay on the Importance of Being Nonlinear", Springer, Berlin (1985).
- 5 - Rumelhart D.E., McClelland, and the PDP Research Group, "Parallel Distributed Processing", Cambridge, MIT Press (1986).
- 6 - Farmer J.D. and Sidorowich J.J., *Phys. Rev. Lett.* **59**, 845 (1987).