

PREDICCIÓN NO-LINEAL DE TIPOS DE CAMBIO: ALGORITMOS GENÉTICOS, REDES NEURONALES Y FUSIÓN DE DATOS

MARCOS ALVAREZ-DIAZ*

Department of Applied Economics, University of Vigo

Lagoas- Marcosende Vigo, Spain

and

ALBERTO ALVAREZ

ISME-DSEA Department of Electrical Engineering, University of Pisa

Via Diotisalve 2,56100 Pisa, Italy

ABSTRACT

It is widely proved the existence of non-linear deterministic structures in the exchange rates dynamic. In this work we intend to exploit these non-linear structures using forecasting methods such as Genetic Algorithm and Neural Networks in the specific case of the Yen/\$ and British Pound/\$ exchange rates. We also employ a novel perspective, called Data Fusion, based on the combination of the obtained results by the non-linear methods to verify if it exists a synergic effect which permits a predictive improvement. The analysis is performed considering both the point prediction and the devaluation or appreciation anticipation.

Keywords: Data Fusion, Genetic Algorithms, Neural Networks, Exchange Rates Forecasting

JEL: C14, C53, G14

* Departamento de Economía Aplicada, Universidad de Vigo Lagoas- Marcosende s/n, 36200 Vigo. e-mail mad@uvigo.es .

1-. Introducción

Debido a su importancia en la actividad económica internacional, la anticipación de la dinámica de los tipos de cambio ha sido un objetivo recurrente en el ámbito financiero durante las últimas décadas. Sin embargo, a pesar del esfuerzo realizado y constatado en numerosos trabajos empíricos, no se han obtenido claras evidencias de predictibilidad. En la literatura se han desarrollado varios modelos teóricos caracterizados por asumir una perspectiva lineal y estocástica¹. Se considera que la evolución temporal de los tipos de cambio puede ser explicada por una dinámica lineal perturbada por shocks exógenos, aleatorios e impredecibles. La verificación empírica de estos modelos muestra muy a menudo signos incorrectos, baja significatividad estadística de los parámetros estimados y un escaso poder predictivo. Respecto al análisis predictivo, en la conocida competición realizada por Meese y Rogoff (1983) se comprobó cómo la gran mayoría de los modelos multivariantes y univariantes no eran capaces de mejorar las predicciones “out-of-sample” de un simple paseo aleatorio. Este resultado corrobora la hipótesis ampliamente asumida de eficiencia en los mercados cambiarios².

Recientes resultados empíricos y teóricos parecen soportar la idea de la existencia de un comportamiento determinista no-lineal en la evolución mostrada por los tipos de cambio (Hsieh, 1989; Brooks, 1996). La presencia de estas dinámicas no-lineales podría implicar la posibilidad de explotarlas para realizar predicciones más

¹ Para una revisión de la literatura teórica y empírica sobre la predicción de tipos de cambio se recomienda la lectura del libro de Baille y McMahon (1989).

² En su versión débil, la asunción de esta hipótesis supone la imposibilidad de predecir rentabilidades futuras en función de sus valores presentes y pasados (Fama, 1970).

precisas que aquellas otras proporcionadas por un modelo estocástico lineal y, en concreto, por el modelo referente en la literatura: el paseo aleatorio. En este sentido, autores como Fernández-Rodríguez y Sosvilla-Rivero (1998) proveen evidencias en favor de la predicción no-lineal de los tipos de cambio.

Los trabajos de Takens (1981) y Casdagli (1989), entre otros, han establecido la metodología necesaria para la modelización no-lineal de series temporales. En concreto, el Teorema de Takens establece que, dada una serie temporal determinista $\{x_t\}_{t=1}^N$, existe una función $F : \mathfrak{R}^m \rightarrow \mathfrak{R}$ tal que

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

donde τ es el factor de retardo y m la dimensión de incrustación. Por tanto, el Teorema garantiza la posibilidad de anticipar la dinámica futura de una serie temporal determinista considerando únicamente sus valores pasados. El problema a resolver consiste en encontrar una buena representación o aproximación funcional de $F(\cdot)$.

Gracias a los avances en el campo informático se han desarrollado potentes y sofisticadas técnicas no-lineales de predicción para la aproximación funcional de $F(\cdot)$, como pueden ser el método de ocurrencias análogas, las redes neuronales y los algoritmos genéticos. La econometría aplicada ha incorporado toda esta serie de técnicas para la predicción de diferentes tipos de cambio. En un primer momento, las aproximaciones más empleadas fueron aquellas basadas en generalizaciones del método de ocurrencias análogas. Por ejemplo, Diebold y Nason (1990) aplicaron el método de regresión localmente ponderada, Bajo, Fernández y Sosvilla (1992) emplearon

predictores baricéntricos y Lisi y Medio (1997) una regresión local. En los últimos años se observa una mayor intensidad en el uso de redes neuronales (Kuan and Liu, 1995; Tenti, 1996; Yao et al., 1997; Zang and Hu, 1998; Hu et al., 1999; Yao and Tan, 2000; Walzack, 2001). De forma más reciente se ha empleado programación genética para la predicción de tipos de cambio (Álvarez-Díaz y Álvarez, 2002).

Un aspecto interesante y escasamente considerado en el ejercicio predictivo consiste en explotar los efectos sinérgicos que pudieran existir entre diferentes métodos predictivos. Esta posibilidad de análisis, denominada fusión de datos o composición de predicciones, permite obtener una aproximación a la dinámica subyacente en los datos a partir de una combinación de las predicciones obtenidas con métodos individuales. Se han argumentado 2 razones básicas por las cuales se recomienda el uso de la fusión de datos. En primer lugar, diferentes métodos de predicción poseen diferentes ventajas y reflejan diferente información. Por este motivo, es de esperar que una combinación de predicciones permita aprovechar las ventajas de cada método y mejorar los resultados en términos predictivos. En segundo lugar, la composición permite reducir los sesgos que pudieran existir empleando sólo un método predictivo.

A la hora de llevar a cabo un ejercicio predictivo basado en la fusión de datos, es preciso definir una forma de combinar predicciones. El procedimiento habitual consiste en emplear una regresión para asignar pesos a cada una de las predicciones obtenidas por cada método individual (Kwok and Lubecke, 1990). Recientemente, asumiendo la limitación de una perspectiva lineal, se ha empleado un método de combinación basado en redes neuronales (Lubecke et al., 1998). Sin embargo, la adopción de una red neuronal requiere el tedioso proceso de especificar una arquitectura previa.

En este trabajo empleamos un algoritmo genético para combinar las predicciones obtenidas por 2 métodos individuales (en particular, un algoritmo genético y una red neuronal). El objetivo perseguido se centra en comprobar si, empleando tanto los métodos considerados individualmente o en combinación, es posible aproximar con precisión la función $F(\cdot)$ en el caso específico del tipo de cambio semanal del yen y la libra respecto al dólar estadounidense.

El trabajo se estructura en 5 secciones. Después de esta sección introductoria se presentan los métodos predictivos empleados. A continuación, en la sección 3, se comentan ciertos aspectos del ejercicio predictivo a desarrollar, se describen brevemente las series y se analizan las regularidades empíricas observadas. En la sección 4 se muestran los resultados obtenidos por cada método en términos de predicción puntual y del porcentaje de aciertos de signo (tanto a uno como a diferentes horizontes temporales). Por último se finaliza con una sección dedicada a conclusiones.

2-. Métodos No-lineales de Predicción

Redes Neuronales

Este método no-lineal, inspirado en los estudios sobre el funcionamiento del cerebro y del sistema nervioso, ha sido empleado para resolver numerosos problemas económicos y financieros destacando, en gran medida, su aplicación en la predicción de series temporales. En la literatura especializada se distinguen diferentes tipos de redes neuronales (Gately, 1996), si bien la feedforward multilayer network con un algoritmo de aprendizaje basado en la técnica backpropagation (Rumelhart y McClelland, 1986) es la red más popular en economía y finanzas (Wong, 1995; Yao, Li and Tan, 1997). La

principal ventaja de las redes neuronales consiste en su elevada capacidad para detectar y explotar la no-linealidad existente en los datos, aun en condiciones donde existen datos incompletos o la presencia de ruido es importante. En nuestra aplicación empírica empleamos una feedforward network con 3 niveles cuya formulación estadística puede ser expresada como

$$\hat{x}_t = \Phi \left(\beta_0 + \sum_{h=1}^H \beta_h \cdot \Psi_h \left(\alpha_0 + \sum_{j=1}^J \alpha_{hj} \cdot x_{t-j} \right) \right) \quad (2)$$

donde \hat{x}_t es el output del modelo. Las funciones $\Psi(\cdot)$ y $\Phi(\cdot)$ se denominan función de transferencia del *nivel oculto* y del *nivel output*, respectivamente. La red presenta J inputs (retardados de la variable x_t), H unidades de proceso (neuronas) en el nivel oculto y 1 output. Inicialmente, los pesos α_{hj} y β_h son determinados aleatoriamente adoptando valores dentro de un determinado rango. Por medio de un proceso iterativo de aprendizaje basado en la técnica backpropagation, los valores de estos pesos son modificados de forma que la diferencia entre el valor real y el valor estimado (output de la red neuronal) sea mínimo. Ha sido demostrado por numerosos trabajos teóricos que este tipo de red, con un único nivel oculto y con un número suficientemente elevado de unidades, es capaz de aproximar cualquier función no-lineal con un determinado grado de precisión (Cybenko, 1989; White, 1990). Su empleo ha sido mayoritario en las aplicaciones empíricas sobre predicción de series financieras y, en particular, de tipos de cambio (Hu et al., 1999).

Además de la complejidad existente en los datos, el éxito predictivo de una red depende en gran medida de la correcta determinación de su arquitectura. Por tanto, es

preciso especificar de forma óptima el número de inputs (J) y el número de unidades de proceso en el nivel oculto (H) así como seleccionar la estructura de las funciones de transferencia ($\Psi(\cdot)$ y $\Phi(\cdot)$). Por ejemplo, un número demasiado elevado de H puede originar problemas de overfitting y, en consecuencia, una ausencia de generalización. Por otro lado, con un número insuficiente de unidades de proceso la red puede perder capacidad predictiva al no explotar plenamente la no-linealidad existente en los datos. En la literatura es posible encontrar reglas para definir el número de inputs y de unidades de proceso en el nivel oculto pero ninguna es perfecta ni se ha asumido de forma generalizada (Yao, Tan and Poh, 1999). Una recomendación muy común y ampliamente empleada consiste en determinar J y H por medio de un proceso de “prueba y error”. En nuestro trabajo asumimos esta recomendación determinando aquellos valores que reportaban el mínimo error en un subconjunto muestral destinado exclusivamente a este propósito (conjunto de selección).

Respecto a las funciones de transferencia, se suele definir una estructura no-lineal en el nivel oculto ($\Psi(\cdot)$) y lineal en el nivel output ($\Phi(\cdot)$) (Qui, 1999). Nuestra propia experiencia en anteriores trabajos sobre predicción financiera nos ha llevado a considerar la función tangente hiperbólica. Además, esta estructura funcional se ha citado como la más común en la predicción de series temporales (Chapman, 1994).

Otro problema a resolver consistió en la excesiva variabilidad de los resultados ante diferentes pesos iniciales (Racine, 2000). La solución propuesta consistió en considerar óptima aquella arquitectura que optimizase el criterio de ajuste en el conjunto de selección después de operar la red con un determinado número de modificaciones de los pesos iniciales (Hu et al, 1999).

En definitiva, siguiendo las recomendaciones existentes en la literatura y nuestra propia experiencia, se aplicó una feedforward backpropagation network con 3 layers para predecir la dinámica del tipo de cambio libra y yen respecto al dólar estadounidense. En cuanto a su arquitectura, el número de inputs y de unidades de proceso fueron determinados por un proceso de prueba y error. Por su parte, se consideró una función de transferencia tangente hiperbólica para el nivel oculto y lineal para el nivel output.

Algoritmos Genéticos

De forma reciente se ha incorporado al análisis un novedoso procedimiento basado en las teorías darwinianas de selección natural y supervivencia. Estos procedimientos, denominados algoritmos genéticos o evolutivos, han sido desarrollados por Holland (1975) y difundidos por Goldberg (1989) y Koza (1992). Los algoritmos genéticos (AG) ya han demostrado su robustez para aproximar explícitamente la forma funcional $F(\cdot)$ en el análisis no-lineal de series temporales (Szpiro, 1997; Yadavalli et al., 1999; Álvarez et al., 2001). Su interés ha crecido intensamente en los últimos años entre los investigadores de múltiples disciplinas incluida la economía (Koza, 1995; Szpiro, 1997; Beenstock and Szpiro, 2002) y las finanzas (Neely et al., 1997; Allen and Karjalainen, 1999; Fyfe et al., 1999; Kaboudan, 2000).

Un AG permite encontrar una solución óptima a partir de la evolución de una población inicial de soluciones alternativas. Supone un proceso de búsqueda eficiente a través de un amplio número de posibilidades con el objetivo de hallar una expresión matemática que, dentro del espacio de posibles soluciones, mejor describa la dinámica presentada por una serie temporal. Esta técnica presenta toda una serie de ventajas

frente a las tradicionales técnicas no-lineales de predicción. En primer lugar, a diferencia de las redes neuronales y de los métodos basados en ocurrencias análogas, los AG obtienen explícitamente una ecuación matemática como aproximación a la verdadera dinámica de la serie. Además, se presentan como un método predictivo más flexible que las redes neuronales ya que no requieren la costosa especificación de una arquitectura previa.

En este trabajo se ha empleado un algoritmo genético para la predicción de series temporales denominado DARWIN (Álvarez et al., 2001; Álvarez-Díaz y Álvarez, 2002)). Su funcionamiento se basa en la simulación en un ordenador del proceso evolutivo observado en la Naturaleza y que puede ser explicado por medio de una serie de pasos iterativos. Para empezar, en el primer paso se generan de forma aleatoria una población inicial de N ecuaciones matemáticas por medio de una sencilla combinación aleatoria de operadores y operandos de la forma

$$S_j = ((A \otimes B) \otimes (C \otimes D)) \quad \forall 1 \leq j \leq N \quad (3)$$

en donde A , B , C y D son los argumentos (genes operandos) y el símbolo \otimes representa a los operadores matemáticos (genes operadores). Los argumentos considerados pueden ser números reales pertenecientes a un intervalo (coeficientes de las ecuaciones) o bien valores retardados de la variable $(x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-m\tau})$. Por su parte, los operadores matemáticos (\otimes) empleados serán la suma (+), resta (-), multiplicación (\cdot) y división (\div), este último operador estará protegido para evitar cocientes entre 0 o números muy pequeños. También cabe la posibilidad de incluir otros operadores matemáticos (como

el logaritmo o los trigonométricos, por ejemplo) pero a costa de incrementar la complejidad en el proceso de optimización funcional. Además, trabajos previos llevados a cabo con DARWIN (Álvarez et al., 2001) así como con otros AG (Szpiro, 1997; Yadavalli et al., 1999) han demostrado la posibilidad de describir dinámicas complejas por medio de expresiones matemáticas construidas simplemente con los operadores aritméticos.

En el segundo paso, una vez determinada la población inicial, comienza el proceso evolutivo seleccionando aquellas ecuaciones que presentan una mayor fortaleza según algún criterio de ajuste. Para el caso de predicción puntual se ha determinado como criterio de fortaleza el *Error Cuadrático Medio Normalizado* definido como

$$ECMN_j = \frac{1}{Var(x_t)} \cdot \frac{\sum_{t=m+1}^M [x_t - \hat{x}_t]^2}{M} \quad (3)$$

donde $ECMN_j$ es el Error Cuadrático Medio Normalizado presentado por la ecuación j -ésima ($\forall 1 \leq j \leq N$), $Var(x_t)$ es la varianza de la serie temporal, \hat{x}_t el valor predicho y M el número total de observaciones destinadas a entrenar el AG. Una generalización de esta expresión ha sido empleada como criterio de fortaleza en los AG de Szpiro (1997) y Yadavalli et al. (1999).

A continuación, todas las ecuaciones de la población inicial son clasificadas en orden decreciente según su $ECMN_j$. Las ecuaciones con valores muy elevados del ECMN son aniquiladas mientras que, por el contrario, las que poseen un valor más

pequeño tienen una mayor probabilidad de supervivencia constituyendo la base de la siguiente generación.

Las ecuaciones supervivientes al proceso de selección son empleadas para generar a los individuos de una nueva generación (proceso de reproducción). Para ello se les aplican los denominados operadores genéticos: Clonación, Cruzamiento y Mutación. Con la clonación las mejores ecuaciones son copiadas exactamente a la siguiente generación. Por su parte, con el operador cruzamiento se seleccionan parejas de ecuaciones con pequeños valores de $ECMN_j$ para intercambiar partes de sus argumentos y operadores matemáticos. Por último, la mutación implica el reemplazamiento aleatorio de algún operador o argumento en un porcentaje reducido de ecuaciones.

En definitiva, la nueva población generada a partir de la población inicial estará constituida por individuos clonados, mutados y cruzados. A partir de este momento el proceso evolucionará repitiendo los pasos de selección y reproducción de forma iterativa en busca de aquella expresión matemática que menor $ECMN$ presente. Después de un número determinado de generaciones del proceso evolutivo, la iteración terminará y el AG ofrecerá explícitamente como resultado una ecuación matemática que considera óptima para representar la verdadera dinámica de la serie temporal³.

³ Para nuestro problema específico de predicción, se configuró el AG de tal manera que el número máximo de argumentos y operadores permitido era 20, cada generación estaba constituida por una población máxima de 120 ecuaciones y, en cada caso, se consideró un máximo de 5000 generaciones.

Para el propósito de la fusión de datos, DARWIN fue re-programado para permitir una perspectiva de análisis multivariante⁴. Por tanto, a partir de los resultados obtenidos por el AG (\hat{G}) y la red neuronal (\hat{N}) para predecir una series temporal $\{x_t\}_{t=1}^T$, se empleó DARWIN para encontrar aquella expresión matemática

$$\hat{x}_t = H(\hat{G}, \hat{N}) \quad (4)$$

que permitiese una combinación óptima de las predicciones en términos de un determinado criterio de ajuste (por ejemplo, el ECMN).

3-. Comentarios y Regularidades Empíricas Observadas en los Datos

La base de datos empleada ha sido obtenida del Pacific Exchange Rate Service (Universidad de British Columbia) y está constituida por datos semanales del tipo de cambio libra y yen respecto al dólar estadounidense. La muestra finalmente seleccionada abarca un periodo comprendido entre la primera semana de 1973 hasta la última semana de Julio de 2002 (1542 observaciones en total). La elección de una periodicidad semanal se justifica para minimizar los sesgos originados por el efecto día de la semana (Lo and Mackinlay, 1988; LeBaron, 1993) y el efecto fin de semana (Zhang and Hu, 1998). Además, se asume que la periodicidad semanal contiene suficiente información como para capturar la dinámica mostrada por los tipos de cambio (Yao and Tan, 2000).

Como es habitual en la mayoría de las aplicaciones empíricas, se considera para el análisis la diferencia del logaritmo de los tipos de cambio

$$r_t = \log(y_t) - \log(y_{t-1}) \quad (5)$$

⁴ Al contrario de un análisis univariante, ahora DARWIN permite el buscar relaciones funcionales entre 2 o más series temporales.

donde y_t es el tipo de cambio objeto de análisis, $\log(y_t)$ es su transformación logarítmica y r_t se considera su rendimiento. Esta transformación ha llegado a ser estándar en el análisis financiero ya que permite obtener una serie estacionaria, puede ser interpretada como una rentabilidad y, además, se presenta como una variable de mayor interés para los operadores financieros (Brooks, 1996). No obstante, también se reconoce la posible amplificación del ruido existente en la serie (Soofi and Cao, 1999).

Siguiendo las recomendaciones existentes en la literatura (Yao and Tan, 2000), se ha dividido la muestra disponible en 3 subconjuntos: Entrenamiento, Selección y, finalmente, “Out-of-Sample”. El subconjunto de entrenamiento, compuesto por las 1080 primeras observaciones, está reservado para el entrenamiento de la red neuronal y la evolución del algoritmo genético⁵. El subconjunto de selección, integrado por las 306 siguientes observaciones, permite seleccionar la dimensión de incrustación para cada uno de los métodos y, en particular, determinar la arquitectura de la red neuronal. Por su parte, para el subconjunto “out-of-sample” se reservan las últimas 154 observaciones. Esta última submuestra, que abarca aproximadamente 3 años, permitirá verificar la validez y consistencia predictiva de los métodos propuestos.

Para el caso de la predicción puntual, tal y como recomienda Casdagli (1989), la medida de ajuste considerada fue el ECMN definido en la ecuación (3). Este criterio para evaluar la capacidad predictiva de los métodos considerados ha sido empleado en multitud de trabajos dedicados a la predicción de series financieras (Elms, 1994; Yao, Tan and Poh, 1999; Yao and Tan, 2000; Tenti, 1996). Se basa en la comparación de los

⁵ En el caso del AG, el conjunto de entrenamiento se reserva para su evolución tanto en el caso de la predicción individual como en la fusión de predicciones.

errores del método no-lineal respecto a los errores que se obtendrían si se considerase la media de la serie como predictor. De esta forma, un valor del ECMN mayor/igual/menor que uno implicaría una capacidad predictiva peor/igual/mejor que emplear la media como predictor. Por otro lado, para la predicción de una apreciación o depreciación de empleo como criterio el porcentaje de signos correctamente predichos.

La tabla 1 presenta los principales estadísticos descriptivos para cada una de las series. El análisis descriptivo de los datos permite observar ciertas regularidades empíricas comunes en la mayoría de las rentabilidades financieras (Guarda and Salmon, 1993). Para empezar, la media es aproximadamente igual a cero, la función de distribución empírica es no normal presentando leptocurtosis y asimetría y, analizando la gráfica 1, se puede comprobar la baja correlación serial existente. Esta última característica indicaría una ausencia de estructura en el nivel medio verificándose, desde una perspectiva lineal, la Hipótesis de Mercados Eficientes y, en consecuencia, la impredecibilidad de las series analizadas. Sin embargo, como apunta Hsieh (1989), las oscilaciones observadas en los tipos de cambio pueden presentar dependencias no-lineales a pesar de ser linealmente incorreladas⁶. En este caso, la aplicación de métodos no-lineales podrían encontrar y explotar estructuras potencialmente predecibles ocultas para las técnicas lineales.

⁶ Para detectar posibles estructuras no-lineales en nuestros datos se ha aplicado el test BDS (Brock et al., 1996) y el test de rachas. Ambos estadísticos rechazan de forma altamente significativa la hipótesis de independencia verificando la presencia de comportamientos no-lineales en las series analizadas.

4-. Resultados

Predicción Puntual

En este apartado se analiza la habilidad del algoritmo genético, de la red neuronal así como de la composición de predicciones para aproximar de forma puntual la evolución de los tipos de cambio objeto de análisis.

Considerando la predicción a un periodo, en las gráficas 1 y 2 se presentan la sensibilidad de los métodos no-lineales ante diferentes dimensiones de incrustación (m) en términos del ECMN obtenido en el periodo de selección. Para ambas series, se verifica cierta estabilidad predictiva ante variaciones en la dimensión de incrustación. A pesar de esta estabilidad y siguiendo las recomendaciones de Casdagli (1989), se seleccionó aquel valor de m que reportó el menor ECMN en el periodo de selección.

En las tablas 2 y 3 se muestran los resultados obtenidos por las técnicas predictivas empleadas así como las dimensiones de incrustación finalmente seleccionadas. Para el caso Yen/\$, se comprueba cómo el AG y la red neuronal presentan unos ECMN “out-of-sample” similares y menores que 1 (0.9313 y 0.9329 para el AG y la red, respectivamente). Estos resultados parecen indicar una leve capacidad predictiva. Por su parte, con la fusión de datos se obtiene un ECMN más bajo (0.9233) verificando, de esta manera, la existencia de una cierta sinergia predictiva al combinar las predicciones de los métodos. Para la libra/\$ se observan unos resultados similares empleando el AG, la red o la fusión de datos (0.919, 0.9261 y 0.9189, respectivamente). En este caso, la combinación de predicciones no permite una mejora respecto a los métodos individuales.

Nuestros resultados a un periodo se asemejan a los obtenidos por Tenti (1996) cuando analizó la habilidad de una serie de recurrents neural networks para predecir la evolución del tipo de cambio Deutsche Mark/\$. Además, parecen corroborar las

conclusiones obtenidas por Diebold y Nason (1990). Estos autores consideraron que la presencia de estructuras deterministas no-lineales en la evolución de los tipos de cambio no pueden ser explotadas significativamente en términos predictivos. Sin embargo, plantean la posibilidad de la predicción a diferentes periodos de predicción. En su aplicación empírica, los propios autores no encontraron evidencias significativas en sus predicciones a 4, 8 y 12 periodos para 10 de los principales tipos de cambio.

En la gráfica 3 y 4 se presentan los resultados obtenidos considerando diferentes periodos predictivos. El empleo de los métodos propuestos para los distintos tipos de cambio permite comprobar cómo a un periodo se obtienen las predicciones más precisas y, a partir de la predicción a 2 periodos, el ECMN “out-of-sample” aumenta y fluctúa en torno a uno. En base a la metodología desarrollada por Sugihara and May (1990), este hecho parece corroborar la presencia de una estructura determinista en la dinámica de los tipos de cambios levemente predecible, aunque sólo a muy corto plazo.

Porcentaje de Aciertos de Signo

Una aproximación complementaria a la predicción puntual consiste en anticipar la dirección futura que adoptará el tipo de cambio. El objetivo no se centra en conocer el valor exacto de la evolución en un momento determinado sino que se pretende predecir si el tipo de cambio sufrirá una apreciación o depreciación en el futuro. Desde un punto de vista práctico hay cierto interés en predecir la dirección de signo antes que el valor puntual de la variable ya que, en el mercado cambiario, errores predictivos muy pequeños pero en la dirección equivocada pueden originar importantes pérdidas de capital (Tenti, 1996; Lisi and Medio, 1997). En la mayoría de los trabajos empíricos se obtuvieron unos porcentajes de aciertos que permitían rechazar la aleatoriedad en la

dinámica de los tipos de cambio (Walzack, 2001). Sin embargo, se puede constatar las serias dificultades para superar el umbral del 60% de aciertos en las predicciones a un periodo.

La aplicación de los diferentes métodos propuestos al problema específico de la predicción del signo nos ha llevado a considerar como criterio de ajuste la expresión

$$SR = \frac{\sum_{t=m+1}^M \theta[r_t \cdot \hat{r}_t > 0]}{M} \quad (6)$$

donde SR es el ratio de aciertos de signo (Success Ratio), r_t es la rentabilidad observada, \hat{r}_t es la rentabilidad predicha, $\theta(\cdot)$ es la función Heaviside ($\theta(\cdot) = 1$ si $r_t \cdot \hat{r}_t > 0$ y $\theta(\cdot) = 0$ si $r_t \cdot \hat{r}_t < 0$) y M es el número total de observaciones de la submuestra empleada. Además, a la hora de combinar las predicciones, esta expresión ha sido utilizada como criterio de fortaleza del AG en sustitución de la definida en la ecuación (3). Por tanto, para el caso de la combinación de predicciones, el AG evaluará y seleccionará las ecuaciones en función de la expresión (5).

Como en el caso de la predicción puntual a un periodo, en las gráficas 6 y 7 se muestra la baja sensibilidad del porcentaje de aciertos de signos predichos correctamente respecto a la dimensión de incrustación. Asimismo, siguiendo la misma metodología empleada con la predicción puntual, se seleccionó aquel valor que reportaba el mayor porcentaje en el conjunto de selección.

En las tablas 4 y 5 se presentan las dimensiones finalmente seleccionada y el porcentaje de aciertos de signo obtenido por cada método para cada uno de los tipos de cambio analizados. También se presentan los resultados derivados de la aplicación del

test no-paramétrico propuesto por Pesaran y Timmermann (1992). En el caso de las predicciones out-of-sample, el test verifica si los porcentajes de aciertos obtenidos por los diferentes métodos difieren significativamente de aquellos que se obtendrían si r_t y \hat{r}_t fuesen independientes. El test, bajo la hipótesis nula de independencia, se distribuye según una Normal estandarizada. Por tanto, los valores críticos con un nivel de significatividad del 1%, 5% y 10% corresponden a 2.33, 1.645 y 1.282, respectivamente.

Para la dinámica del tipo de cambio Yen/\$, tabla 4, el AG y la red neuronal obtienen unos porcentajes de acierto out-of-sample en torno al 60%. Sin embargo, no se constata la existencia de un efecto sinérgico cuando se combinan las predicciones alcanzadas por ambos métodos. Analizando los valores mostrados por el test de Pesaran-Timmermann, existen argumentos estadísticos para rechazar la hipótesis de independencia entre los valores reales del tipo de cambio Yen/\$ y los valores predichos por los diferentes métodos no-lineales empleados.

Para el caso del tipo de cambio Libra/\$, se observa una mayor habilidad predictiva out-of-sample del AG respecto a la red neuronal (58.06% frente al 56.13%, respectivamente). Por otro lado, al contrario del caso anterior, la fusión de datos obtiene un aumento en el porcentaje de acierto (59.35%). La aplicación del test de Pesaran-Timmermann permite verificar la significatividad estadística de las predicciones ofrecidas por los métodos.

En resumen, el análisis de la capacidad predictiva out-of-sample por medio de un AG, una red neuronal y una combinación de los resultados de ambos métodos, ha

permitido conseguir, en todos los casos y para todos los tipos de cambio estudiados, unos porcentajes de acierto de signo significativamente superiores a aquellos que se obtendría empleando un paseo aleatorio o, de forma análoga, considerando que r_t sigue un proceso aleatorio (porcentaje de aciertos en torno al 50%).

En las gráficas 8 y 9 se muestran los resultados obtenidos para diferentes periodos de predicción. De este análisis cabría esperar, a partir del segundo periodo, una fluctuación del porcentaje de aciertos en torno al valor del 50%. Sin embargo, como se puede comprobar, existe una elevada capacidad por parte de ciertos métodos para predecir significativamente a 4 periodos. De esta forma, para el caso del Yen/\$, la fusión de datos pero, sobretudo, el AG permite alcanzar unos elevados porcentajes de acierto. Observando la gráfica correspondiente a la libra/\$, se comprueba cómo esta habilidad le corresponde a la fusión de datos y a la red neuronal. Una explicación a este hallazgo pudiera deberse a la existencia de estructuras estacionales no-lineales que permiten anticipar el signo de la evolución de los tipos de cambio únicamente a 4 periodos.

5-. Conclusiones

En este trabajo se ha aplicado un AG y una red neuronal para predecir la evolución de los tipos de cambio Yen y libra respecto al dólar estadounidense. Asimismo, empleando de nuevo un AG, se han combinado las predicciones obtenidas por ambos métodos para verificar la posible existencia de efectos sinérgicos (fusión de datos). El análisis se llevó a cabo adoptando como criterio de ajuste el ECMN en el caso de la predicción puntual así como el porcentaje de aciertos en la predicción de una apreciación o depreciación.

En general, los resultados obtenidos no muestran grandes diferencias respecto a los ya obtenidos en la literatura sobre predicción no-lineal de tipos de cambio (Tenti, 1996; Walzack, 2001). Considerando la predicción puntual, se observa una débil capacidad predictiva a un periodo por parte de todos los métodos empleados. Esta capacidad se pierde cuando se realizan predicciones a un horizonte temporal mayor. Respecto a la predicción de signo a un periodo, los métodos empleados obtienen unos porcentajes de aciertos estadísticamente distintos a los que se obtendrían si las series analizadas fuesen aleatorias. Sin embargo, tal y como se constata también en la literatura, los porcentajes obtenidos no superan el umbral del 60%. Cuando se realizan predicciones de signo a diferentes periodos, se observa una habilidad generalizada por todos los métodos para anticipar significativamente la dirección del signo de las series analizadas a 4 periodos. Por otro lado, con nuestro trabajo también se ha constatado cómo la fusión de datos sólo permite, en el mejor de los casos, pequeñas mejoras respecto a los métodos individuales, tanto considerando predicción puntual como de signo.

En conclusión, a pesar de aplicar cada vez herramientas más sofisticadas y potentes, todavía no se ha conseguido obtener grandes mejoras en la predicción de tipos de cambio. En la literatura se han propuesto varias explicaciones. En primer lugar, es posible que existan débiles estructuras no-lineales en los tipos de cambio que no reportan importantes mejoras predictivas (Diebold and Nason, 1990). Otra posible explicación se basa en admitir la presencia de un componente no-lineal predecible pero todavía sería necesario una mejora en el diseño y desarrollo de las técnicas predictivas. En este sentido, Franses and Griensven (1998) consideran que una posible causa de la

pobreza predictiva observada pudiera deberse a una selección errónea de los inputs de los modelos no-lineales. En vez de utilizar retardos de la variable, sugieren la posibilidad de emplear variables basadas en el análisis técnico. Por otro lado, también se ha propuesto el estudio y desarrollo de nuevos métodos que permitan una aproximación más precisa a la verdadera pero desconocida dinámica de la serie. En este sentido, futuras líneas de investigación podrían centrarse en los denominados métodos híbridos (por ejemplo, utilizar redes evolutivas (Alvarez, 2002)). Una última explicación, planteada por Stengos (1996), incide en la existencia de estructuras deterministas predecibles en las series financieras aunque, debido a su elevada complejidad, un análisis predictivo preciso exigiría un número de observaciones extremadamente elevado.

References

- Allen, F. and Karjalainen, R. (1999), "Using Genetic Algorithms to Find Technical Trading Rules", *Journal of Financial Economics*, 51, pp. 245-271.
- Álvarez A., A. Orfila y J. Tintore (2001) DARWIN- an evolutionary program for nonlinear modeling of chaotic time series, *Computer Physics Communications*, 136, 334-349.
- Álvarez A. (2002) A neural network with evolutionary neurons, *Neural Processing Letters*, 16, pp. 43-52.
- Álvarez-Díaz, M. and A. Álvarez (2002) Forecasting exchange rates using genetic algorithms, *Applied Economic Letters*, forthcoming.
- Baille R. and P. McMahon (1989) *The foreign exchange market: Theory and evidence*. Cambridge University Press.
- Bajo O., F. Fernández and S. Sosvilla (1992) Chaotic Behaviour Exchange-Rate Series. First Results for the Peseta-U.S. Dollar Case, *Economics Letters* 39, pp. 207-211.
- Beenstock M. and G. Szpiro (2002) Specification search in nonlinear time-series models using genetic algorithms. *Journal of Economic Dynamic and Control*, 26, pp. 811-835.

Brock W., D. Dechert and J. Scheinkman (1997) A test for independence based on the correlation dimension. *Econometric Review*, 15, 3, pp. 197-235.

Brooks C. (1996) Testing for non-linearity in daily sterling exchange rates. *Applied Financial Economics*, 6, pp. 307-317.

Casdagli M. (1989) Nonlinear prediction of chaotic time series, *Physica D*, 35, 335-356.

Chapman J. A. (1994) Stock Market reading systems through neural networks: developing a model, *International Journal Applied Expert Systems* 2 (2), 1994, 88-100.

Cybenko, G. (1989) Approximation by superposition of a sigmoidal function, *Mathematics of Control, Signals and Systems*, 2, pp. 303-314.

Diebold, F. X. and J. A. Nason (1990) Nonparametric exchange rate prediction?, *Journal of International Economics*, 28, pp. 315-332.

Elms, D. (1994) Forecasting in financial markets, in J. Creedy and V. L. Martin (eds.), *Chaos and Non-Linear Models in Economics. Theory and Applications*, pp. 169-186.

Fama, E.F. (1970) Efficient capital markets: A review of theory and empirical works. *Journal of Finance*, 25. pp. 383-417.

Fernández-Rodríguez F. and S. Sosvilla-Rivero (1998) Testing nonlinear forecastability in time series: theory and evidence from EMS, *Economic Letters* 59, 49-63.

Franses P. H. and K. Griensven (1998) Forecasting exchange rates using neural networks for technical trading rules. *Studies in Non-linear Dynamics and Econometric*, 2, 4, pp. 109-114.

Fyfe, C., J. P. Marney and H. F. E. Tarbert (1999) Technical analysis versus market efficiency- a genetic programming approach, *Applied Financial Economics*, 9, pp. 183-191.

Gately, E. (1996) *Neural networks for financial forecasting*, Wiley Trader's Advantage, John Wiley & Sons.

Goldberg D. E. (1989) *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley.

Guarda P. and M. Salmon (1993) Dinámica no lineal y tipo de cambio. *Cuadernos Económicos del I.C.E.*, 53, pp. 173-219.

Holland J. H. (1975) *Adaptation in natural and artificial systems*, Ann Arbor. The University of Michigan Press.

Hsieh, D. A. (1989) Testing for nonlinear dependence in daily foreign exchange rates, *Journal of Business*, 62, pp. 329-368.

Hu M. Y., G. Zhang, C. X. Jiang and B. E. Patuwo (1999) A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting. *Decision Sciences*, 30, 1, 197-216.

Kaboudan, M. A. (2000) "Genetic Programming Prediction of Stock Prices", *Computational Economics*, 16, pp. 207-236.

Koza J. R. (1992) *Genetic programming: On the programming of computers by means of natural selection*", The MIT Press, Cambridge.

Koza J. R. (1995) *Genetic Programming for Economic Modeling*, en *Intelligent Systems and Business*, eds. S. Goonatilake and P. Treleaven. John Wiley & Sons, pp. 251-269.

Kuan C. M. and T. Liu (1995) Forecasting exchange rates using feedforward and recurrent neural network. *Journal of Applied Econometrics*, 10, 4, 347-364.

Kwok C. C. and T. H. Lubecke (1990) Improving the correctness of foreign exchange forecasts through composite forecasting. *Management International Review*, 4, 30, 331-352.

LeBaron B. (1993) Forecast improvements using a volatility index, in *Nonlinear Dynamics, Chaos and Econometrics*. Eds. M. H. Pesaran y S. M. Potter. John Wiley & Sons, pp.129-142.

Lisi F. and A. Medio (1997) Is a random walk the best exchange rate predictor?, *International Journal of Forecasting*, 13, 255-267.

Lo W. A. and A. C. MacKinlay (1998) Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, 1, 1, pp. 41-66.

Lubecke T., K. D. Nam, R. E. Markland and C. C. Kwok (1998) Combining foreign exchange rate using neural network, *Global Finance Journal*, 9, 1, pp. 5-27.

Meese, R and K. Rogoff (1983) Empirical exchange rate models of the 1970's: Do they fit out of sample?, *Journal of International Economics*, 14, pp. 3-24.

Neely C., P. Weller and R. Dittmar (1997) Is technical analysis in the foreign exchange market profitable? A genetic programming approach. *Journal of Financial and Quantitative Analysis*, 32, 4, pp. 405-426.

Pesaran M. H. and A. Timmermann (1992) A simple nonparametric test of predictive performance, *Journal of Business & Economic Statistics*, 10, 4, 461-465.

Qui, M. (1999) Nonlinear predictability of stock returns using financial and economic variables, *Journal of Business & Economic Statistics*, 17, 4, pp. 419-429.

Racine, J. On the nonlinear predictability of stocks returns using financial and economic variables, *Journal of Business & Economic Statistics*, 9, 3, pp. 380-382.

Rumelhart D. E. and J. L. McClelland (1986) Parallel distributed processing: Explorations in the micro-structure of Cognition, Vol. 1, 318-362. The MIT Press.

Soofi A. S. and L. Cao (1999) Nonlinear deterministic Forecasting of daily Peseta-Dollar Exchange Rate, *Economic Letters* 62, 175-178.

Stengos, T. (1996) Nonparametric forecasts of gold rates of return, in Barnett W., Kirman A. y Salmon M. (eds.), *Nonlinear Dynamics and Economics*, pp. 393-406.

Sugihara G. and R. M. May (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, *Nature*, 344, pp. 734-741.

Szpiro, G. G. (1997a) Forecasting chaotic time series with genetic algorithm, *Physical Review E*, 55, 3, pp. 2557-2568.

Szpiro G. G. (1997b) A search for hidden relationships: Data mining with genetic algorithms. *Computational Economics*, 10, pp. 267-277.

Takens, F. (1981) "Detecting strange attractors in turbulence", in D. A. Rand and L. S. Young (eds.), *Dynamical Systems and Turbulence*. Springer-Verlag, pp. 366-381.

Tenti, P. (1996) "Forecasting foreign exchange rates using recurrent neural networks", *Applied Artificial Intelligence*, 10, pp. 567-581.

Walzack S. (2001) An empirical analysis of data requirements for financial forecasting with neural networks, *Journal of Management Information Systems*, 17, 4, 203-222.

White H. (1990) Connectionist nonparametric regression multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*, 3, 535-549.

Wong, B. K. (1995) A bibliography of neural network business application research: 1988-September 1994, *Expert Systems* 12, 3.

Yadavalli, V. K, R. K. Dahule, S. S. Tambe and B. D. Kulkarni (1999) Obtaining functional form for chaotic time series evolution using genetic algorithm, *American Institute of Physics*, 9, 3, pp. 789-794.

Yao J., Y. Li and C. L. Tan (1997) Forecasting the exchange rates of CHF vs USD using neural networks, *Journal of Computational Intelligence in Finance*, 5, 2, pp. 7-13.

Yao J., C. L. Tan and H. L. Poh (1999) "Neural networks for technical analysis: A study on KLCP", *International Journal of Theoretical and Applied Finance*, 2, 2, pp. 221-241.

Yao, J. and C. L. Tan (2000) A case study on using neural networks to perform technical forecasting of Forex, *Neurocomputing* 34, pp. 79-98.

Zhang G. and M. Y. Hu (1998) Neural network forecasting of the British Pound/US Dollar exchange rate. *Omega*, 26, 4, pp 495-506.

Table 1. Descriptive Statistics

STATISTICS	EXCHANGE RATES	
	Yen/\$	B. Pound/\$
Minimum	-0.119895	-0.059498
Maximum	0.061198	0.080250
Mean	-0.000586	0.000260
Std. Dev.	0.012697	0.011762
skewness	-1.128799	0.292725
kurtosis	7.695171	3.629172

Figure 1. Autocorrelations

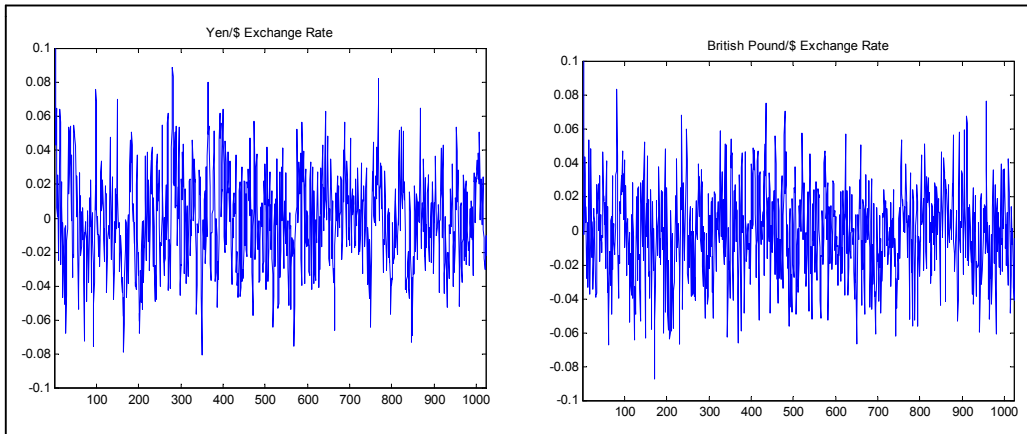


Figure 2. Selection of the Embedding Dimension Yen/\$ Exchange Rate.

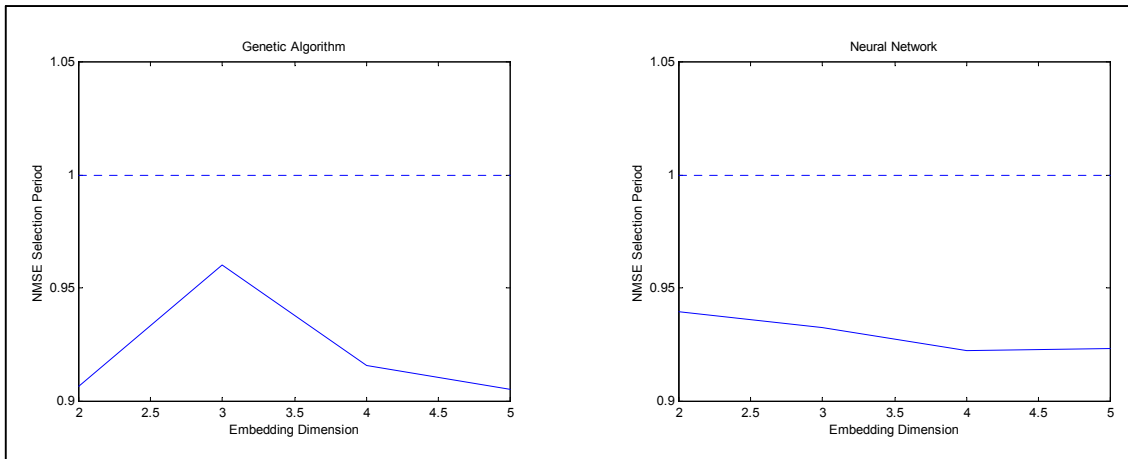


Figure 3. Selection of the Embedding Dimension British Pound/\$ Exchange Rate.

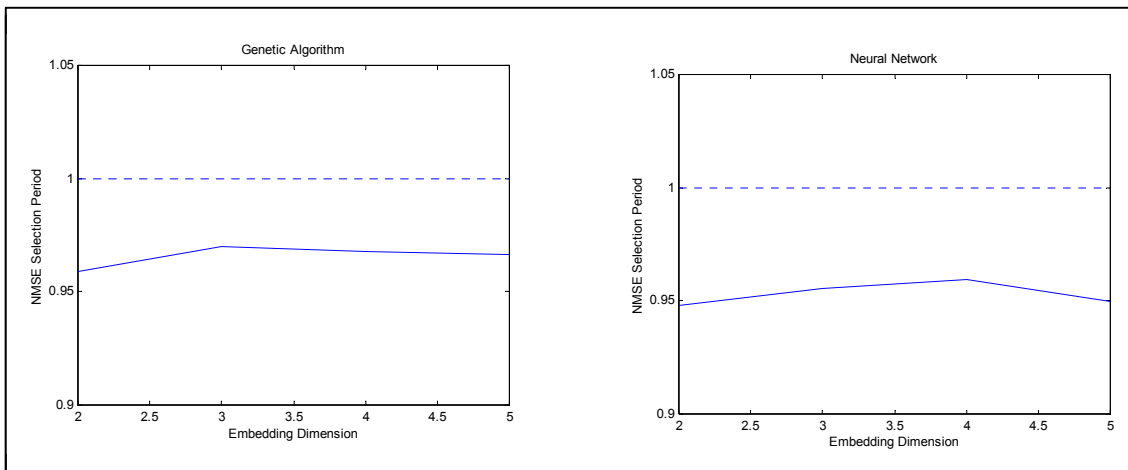


Table 2. Yen/\$ Exchange Rate Results One Period Ahead.

GENETIC ALGORITHM				
Embedding Dimension	Normalized Mean Squared Error			Equation
	Training	Selection	Out-of-sample	
5	0.9153	0.9051	0.9313	$\hat{r}_t = \left\{ \frac{r_{t-1}}{r_{t-1} + 3.9} - r_{t-4}^2 \right\}$

NEURAL NETWORK				
Embedding Dimension	Number Hidden Units	Normalized Mean Squared Error		
		Training	Selection	Out-of-sample
4	4	0.9198	0.9225	0.9329

DATA FUSION				
Normalized Mean Squared Error			Equation	
Training	Selection	Out-of-sample	$\hat{r}_t = \left\{ \hat{G}_t + \frac{\hat{N}_t}{7.49 + [\hat{G}_t \cdot (\hat{G}_t + \hat{N}_t) + \hat{G}_t - \hat{N}_t] \cdot \left[1.49 - \frac{3.53}{\hat{N}_t} \right]} \right\}$	
0.8951	0.9051	0.9233		

where \hat{G} and \hat{N} are the predictions obtained by the genetic algorithm and neural network, respectively.

Table 4. British Pound/\$ Exchange Rate Results One Period Ahead.

GENETIC ALGORITHM				
Embedding Dimension	Normalized Mean Squared Error			Equation
	Training	Selection	Out-of-sample	
2	0.9393	0.9591	0.9190	$\hat{r}_t = \left\{ \frac{3.74 \cdot r_{t-2}^2}{16.9664 \cdot r_{t-1} + r_{t-2}} \right\}$

NEURAL NETWORK				
Embedding Dimension	Number Hidden Units	Normalized Mean Squared Error		
		Training	Selection	Out-of-sample
2	1	0.9536	0.9479	0.9261

DATA FUSION				
Normalized Mean Squared Error			Equation	
Training	Selection	Out-of-sample	$\hat{r}_t = \left\{ \frac{\hat{N}_t \cdot \hat{G}_t}{\hat{N}_t - \hat{N}_t \cdot \hat{G}_t} \right\}$	
0.9333	0.9591	0.9189		

* where \hat{G} and \hat{N} are the predictions obtained by the genetic algorithm and neural network, respectively.

Figure 4. Punctual Prediction to Different Periods: Yen/\$ Exchange Rate

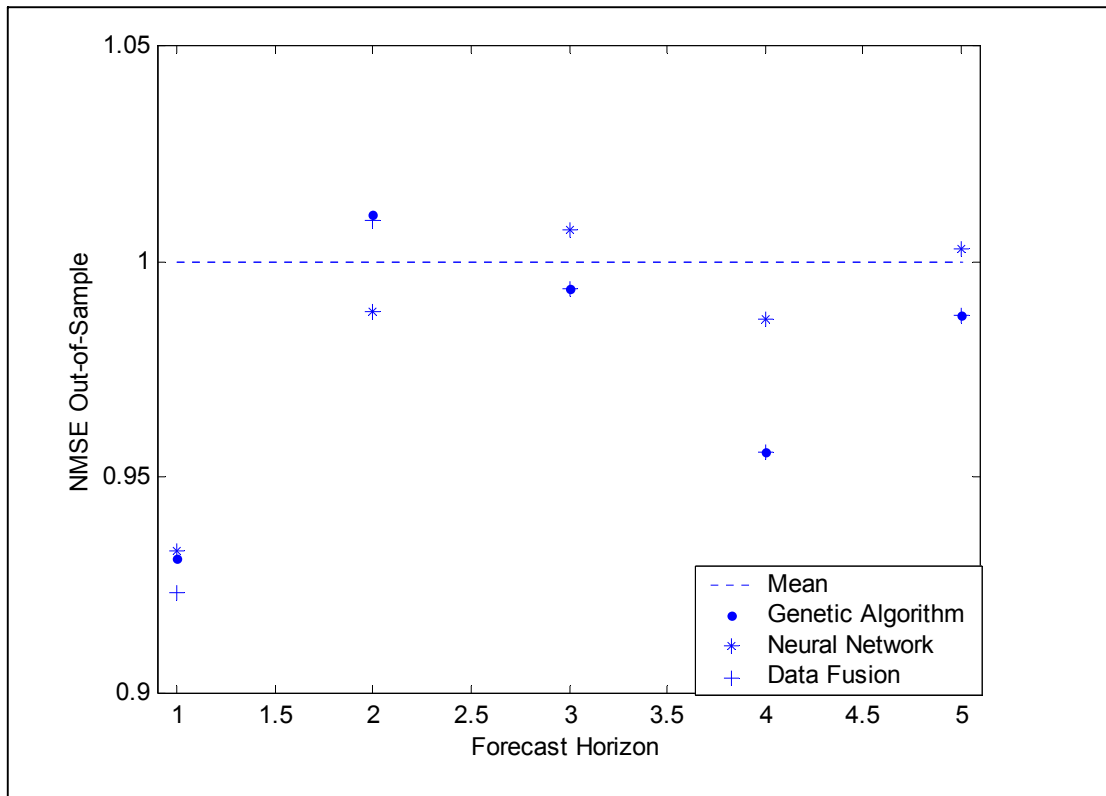


Figure 5. Punctual Prediction to Different Periods: British Pound/\$ Exchange Rate

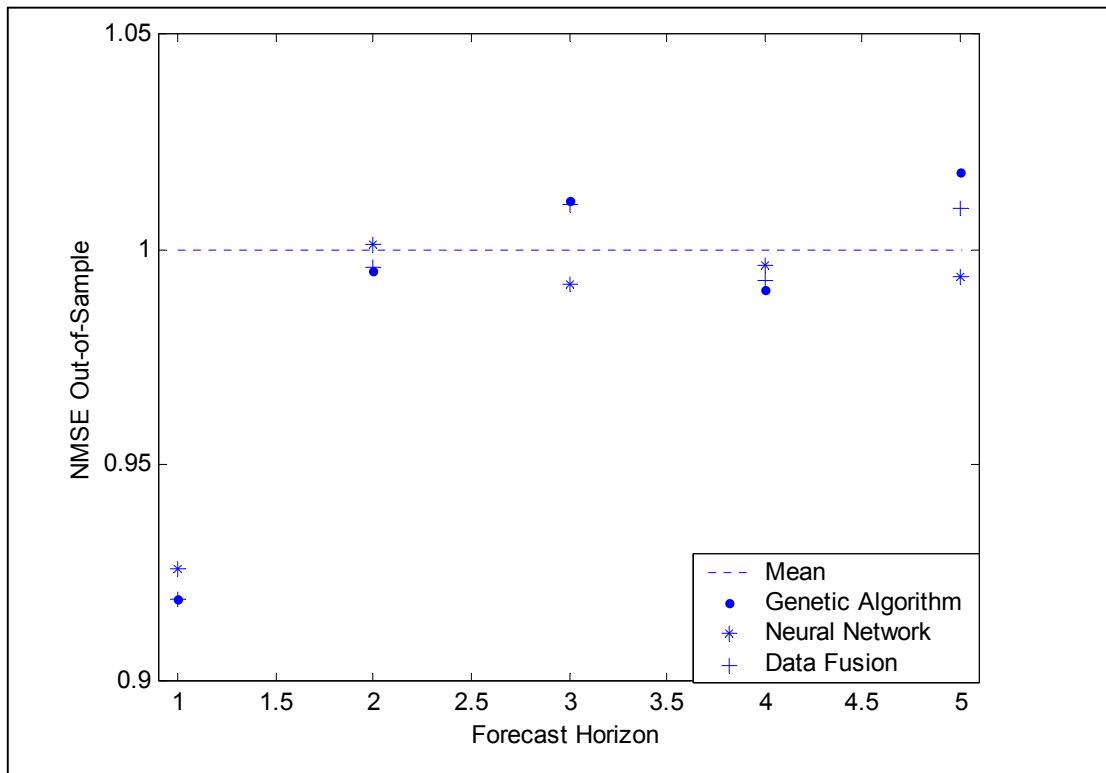


Figure 6. Selection of the Embedding Dimension Yen/\$ Exchange Rate.

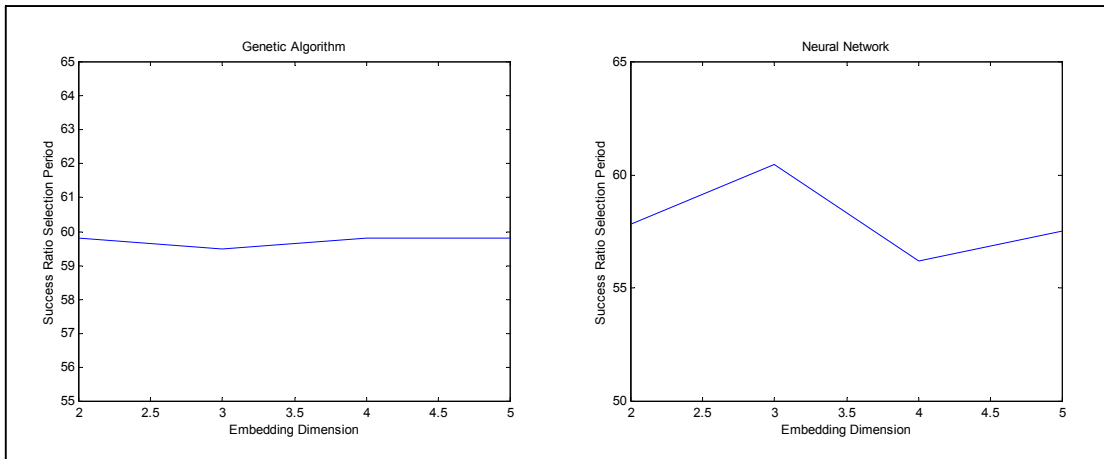


Figure 7. Selection of the Embedding Dimension British Pound/\$ Exchange Rate.

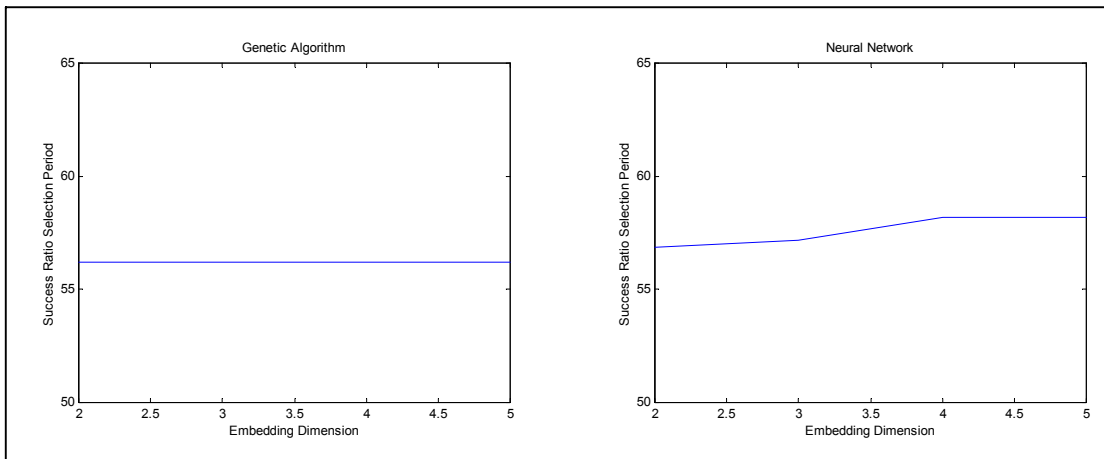


Table 4. Success Ratio One Period Ahead: Yen/\$ Exchange Rate

GENETIC ALGORITHM					
Embedding Dimension	Success Ratio			DA Out-of-Sample	Equation
	Training	Selection	Out-of-sample		
2	62.02	59.8	59.35	2.2763**	$\hat{r}_t = \left\{ \frac{r_{t-1} \cdot (r_{t-1}^2 + r_{t-2})}{3.56 \cdot (r_{t-1}^2 + r_{t-2}) + 9.15 \cdot r_{t-1} \cdot r_{t-2}} \right\}$

** Significance at the 5%.

NEURAL NETWORK					
Embedding Dimension	Number Hidden Units	Success Ratio			DA Out-of-Sample
		Training	Selection	Out-of-sample	
3	5	60.76	60.46	60.00	2.5570***

*** Significance at the 1%.

DATA FUSION				
Training	Success Ratio		DA Out-of-Sample	Equation
	Selection	Out-of-sample		
63.08	61.44	58.71	2.2637**	$\hat{r}_t = \left\{ \hat{N}_t + \frac{\hat{G}_t \cdot (2.94 \cdot \hat{G}_t - 1.03 \cdot \hat{N}_t)}{8.7508 \cdot \hat{N}_t - 21.6384 \cdot \hat{G}_t} \right\}$

** Significance at the 5%. \hat{G} and \hat{N} are the predictions obtained by the genetic algorithm and neural network, respectively.

Table 5. Success Ratio One Period Ahead: British Pound/\$ Exchange Rate

GENETIC ALGORITHM					
Embedding Dimension	Success Ratio			DA Out-of-Sample	Equation
	Training	Selection	Out-of-sample		
2	62.43	56.21	58.06	2.0152**	$\hat{r}_t = \left\{ \frac{3.74 \cdot r_{t-1}^2}{16.9664 \cdot r_{t-1} + r_{t-2}} \right\}$

* Significance at the 5%.

NEURAL NETWORK					
Embedding Dimension	Number Hidden Units	Success Ratio			DA Out-of-Sample
		Training	Selection	Out-of-sample	
4	3	58.77	58.17	56.13	1.5603*

* Significance at the 10%.

DATA FUSION				
Training	Success Ratio		DA Out of Sample	Equation
	Selection	Out-of-sample		
63.23	56.54	59.35	2.3358*	$\hat{r}_t = \left\{ \hat{G}_t + 0.2722 \cdot \hat{N}_t \cdot [59.937 \cdot \hat{N}_t \cdot (\hat{N}_t + 5.02) + \hat{G}_t] \right\}$

* Significance at the 1%. \hat{G} and \hat{N} are the predictions obtained by the genetic algorithm and neural network, respectively.

Figure 8. Success Ratio to Different Periods: Yen/\$ Exchange Rate.

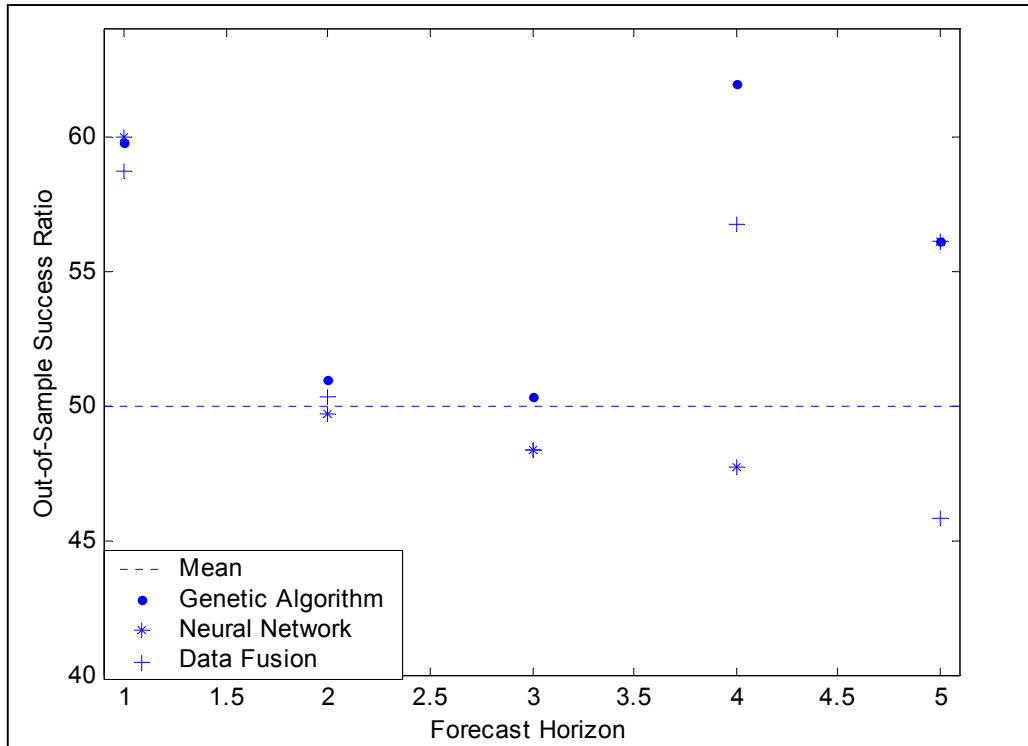


Figure 9. Success Ratio to Different Periods: British Pound/\$ Exchange Rate.

