

UTILIZACION DE LAS REDES NEURONALES ARTIFICIALES EN EL ANALISIS Y PREDICCIÓN DEL IPC REGIONAL

Joaquín Antonio PACHECO BONROSTRO

E.U.E. Empresariales

Departamento de Economía

C/Francisco Vitoria s/n

Burgos 09006

Cristina Rocío DELGADO SERNA

Licenciada en C.Económicas y Empresariales

S.A.A.T. (Sociedad Anónima de Aduanas y Transportes)

Aduana Interior. Despacho nº1. Villafraía

Burgos 09192

Resumen.- En este trabajo se pone a prueba la capacidad predictiva de una red neuronal artificial, concretamente el modelo denominado Perceptrón Multicapa con el algoritmo de Propagación hacia atrás, en el análisis del IPC, tanto el general como el de los diversos sectores, de Castilla y León. Los datos mensuales son aportados por el Instituto Nacional de Estadística. Se comparan los resultados con los obtenidos utilizando la conocidísima metodología de Box-Jenkins y los modelos de ARIMA univariantes. Los resultados son satisfactorios y en cualquier caso ponen de manifiesto una vez más la *flexibilidad* de las redes neuronales frente a la *rigidez* de algunos de los métodos estadísticos y econométricos más clásicos.

Palabras clave: Perceptrón Multicapa, modelos ARIMA, Predicción, Análisis del IPC Regional

1.- INTRODUCCION: MODELOS ARIMA Y REDES NEURONALES

1.1.- Los Modelos univariantes en el Análisis de la Inflación

Como comentábamos en un trabajo anterior, (PACHECO y CALZADA, (1.990)) los métodos cuantitativos son importantes en el análisis de la coyuntura económica en dos sentidos fundamentalmente: en obtener modelos estadísticos-econométricos que expliquen la generación de datos y en aplicar procedimientos estadísticos que permitan extraer de una determinada serie temporal una señal de interés -aspecto esencial del fenómeno- no observable como puede ser la tendencia.

Para poder discutir la naturaleza de estos modelos conviene considerar que cada fenómeno económico individual aparece relacionado con otros muchos fenómenos, es decir, los fenómenos

económicos no están aislados unos de otros, sino que se determinan conjuntamente. En consecuencia se necesitaría un modelo capaz de explicar el comportamiento de todas sus variables, incluso en el caso que solo queramos estudiar una sola variable, como es nuestro caso, necesitaríamos considerar la evolución de las disponibilidades líquidas de dinero, los tipos de interés, evolución de la renta,... para estudiar el comportamiento de la inflación.

Cuando se dispone de un modelo teórico adecuado, de información estadística suficiente y de una previsión precisa de los factores externos o variables explicativas es preferible utilizar un modelo completo del sistema que se estudia, que será multicuacional y multivariante. No obstante en muchas ocasiones no se da esta circunstancia siendo necesario recurrir a un análisis de carácter univariante.

Una solución, por tanto, es el análisis del fenómeno en sí mismo, considerando exclusivamente aquellas variables concreta que pretendemos explicar. Así partiendo de los datos de la serie, se trata de obtener el modelo cuantitativo que explique su comportamiento. Naturalmente desde un punto de vista teórico constituye una aproximación al fenómeno observado, peor que en el caso anterior, pero puede ser muy atractivo al requerir menos información, ser de más fácil elaboración y permitir menos ambigüedades.

Por otra parte cualquier modelo univariante -por ejemplo los ARIMA-, tienen plena justificación económica. En efecto suponiendo que las variables exógenas que entran en el modelo econométrico vienen determinadas por modelos ARIMA, se puede demostrar que las variables endógenas siguen también modelos ARIMA. Por tanto, estos modelos son válidos para explicar el comportamiento de variables como el I.P.C. En ellos, al considerar la serie de datos históricos se esta considerando las variables causales de los precios, aunque se hace con retraso. Esto último hace que los modelos univariantes sean ineficientes en el sentido de que no incorpora el efecto de las innovaciones recientes de las variables explicativas del modelo econométrico, ya que estas innovaciones pueden necesitar un periodo de tiempo para incorporarse plenamente en los datos históricos de los precios. Sin embargo al ser más fáciles de construir se garantiza una mejor aproximación más correcta al modelo real y la posibilidad de obtener predicciones con errores de media nula es mayor.

1.2.- Las Redes Neuronales

La computación programada, basada en reglas de decisión y algoritmos codificados en programas que se ejecutan en ordenadores digitales, ha dominado el procesamiento de la información en las últimas décadas. El avance de la Electrónica y la Computación ha permitido hoy en día podamos disponer sobre nuestra mesa de trabajo ordenadores personales de gran potencia. Pese a ello, existen tareas que ni los grandes supercomputadores son capaces de resolver, o lo hacen de un modo poco eficiente, mientras que el cerebro lo viene haciendo desde hace millones de años con suma facilidad y eficacia.

Según se comenta en el texto de FREEMAN y SKAPURA, (1.993), Todavía hay mucha gente que se sorprende con el hecho de que si bien los ordenadores convencionales se han aplicado de

forma muy prospera a una gran variedad de áreas, y con una velocidad de varias operaciones simples en unos nanosegundos (10^{-9} seg.) no pueden solucionar la tarea simple de reconocer una cara humana. Esto puede realizarlo fácilmente un niño, aunque con la velocidad de la neurona humana en la recepción visual sea relativamente lenta (varias operaciones simples en unos milisegundos). Dado que la respuesta del proceso completo se puede realizar en unos pocos milisegundos, el modelo de conexiones de unidades neuronales parece ser el único camino para lograr este tiempo de respuesta.

Por esta razón los científicos han vuelto la vista hacia el cerebro tratando de estudiarlo desde el punto de vista de la computación. Las diferencias que separan el cerebro y el ordenador son enormes como se refleja en el cuadro siguiente extraído del trabajo de SERRANO-CINCA, (1.995):

	Cerebro	Ordenador
Velocidad de Procesamiento	10^{-3} s.	10^9 s.
Modo de Procesamiento	Paralelo	Serie
Número de procesadores	10^{11}	Pocos
Tipo de control de proceso	Democrático	Dictatorial
Conexiones	10000 por procesador	Pocas
Almacenamiento del conocimiento	Distribuido	En posiciones precisas
Tolerancia a fallos	Amplia	Poca o nula

Comparación del cerebro con un ordenador convencional

En este proceso del pensamiento científico surgen las redes neuronales artificiales, con la idea de tomar las características esenciales de la estructura neuronal del cerebro para crear sistemas que lo minimicen en parte, mediante sistemas electrónicos o mediante simulación por ordenador, aprovechando sus propiedades de cálculo. Estos sistemas están compuestos por multitud de procesadores simples y pueden adquirir, almacenar y utilizar conocimiento experimental, obtenido a partir de ejemplos.

Por otro lado, algunas de las redes neuronales existentes han mostrado potencial para adquirir alguna de las funciones del cerebro humano como la asociación, categorización, generalización y adaptación. Las redes neuronales tienen la capacidad asociativa de recordar por entero una información compleja usando una parte de ella. Usando la capacidad auto-organizativa pueden adquirir y organizar conocimientos mediante el aprendizaje en respuesta a estímulos externos. Además, se usan para la clasificación de patrones y una vez pasado el proceso de aprendizaje son capaces de la generalización.

De esta forma encontramos trabajos en los que se compara las redes neuronales con métodos estadísticos multivariantes usados para problemas de reconocimiento de patrones y/o clasificación; así en el trabajo de SERRANO-CINCA, (1.995), se compara la estructura neuronal denominada *Mapas Autoorganizados* con el Escalado Multidimensional y Análisis de

Conglomerados, la *Red de Propagación hacia atrás* con el Análisis Discriminante y el Análisis Factorial. También en el trabajo de GUILLEN y SOLDEVILLA (1.996) se compara esta estructura neuronal con el Análisis Discriminante y con la Regresión Logística,...

Pero el potencial de las redes neuronales va más allá del reconocimiento de patrones y para algunos autores son una manera alternativa de enfocar los problemas de predicción, tanto en modelos económicos univariantes como multivariantes. Trabajos donde se hacen comparaciones con modelos ARIMA son el de OTERO y TRUJILLO, (1.993), DE LA FUENTE y otros, (1.996). En los trabajos de ARAGON y GARCIA, (1.996a) y (1.996b), y GARCIA y PACHECO, (1.996) se usan diferentes estructuras neuronales para la predicción del IBEX-35.

En el presente trabajo pretendemos predecir la serie temporal del IPC de Castilla y León comparando la capacidad predictiva de una red neuronal con la de un modelo ARIMA univariante. Los datos que se utilizarán son los valores mensuales del IPC facilitados por el Instituto Nacional de Estadística. Para estimar el modelo ARIMA y entrenar la red neuronal se usarán los datos que van de Enero de 1.978 a Diciembre de 1.993. Los datos de los periodos siguientes, (años 1.993 en adelante), se usarán para contrastar y comparar la capacidad predictiva de ambas metodologías.

La arquitectura neuronal utilizada en este trabajo es la conocida como *Perceptrón Multicapa*, o también conocida como *Red de Propagación hacia atrás* por la forma en que se realiza su proceso de aprendizaje.

En los apartados siguientes se explican brevemente ambas metodologías modelos ARIMA y Perceptrón Multicapa. Con los modelos y estructuras que se obtienen se realizan y comparan las predicciones.

2.- LOS MODELOS ARIMA. METODOLOGIA DE BOX-JENKINS

Recordemos que sea una serie temporal Y_t se dice que sigue un modelo ARIMA de orden $(p,d,q) \times (P,D,Q)_s$, si responde a la siguiente ecuación:

$$F(B) \cdot (1-B)^d \cdot Y_t = c + G(B) \cdot e_t,$$

siendo

e_t un proceso de 'ruido blanco' o error aleatorio,

$F(B) = (1 - \beta_1 B - \dots - \beta_p B^p)$, (polinomio autoregresivo),

$G(B) = (1 - \mu_1 B - \dots - \mu_q B^q)$, (polinomio de medias móviles).

$F_s(B) = (1 - \gamma_1 B - \dots - \gamma_p B^{p \cdot s})$, (polinomio autoregresivo estacional)

$G_s(B) = (1 - \rho_1 B - \dots - \rho_q B^{q \cdot s})$, (polinomio de medias móviles estacional)

B el 'operador retardo', es decir, $BY_t = Y_{t-1}$,

c una constante.

s longitud de la estacionalidad.

Siguiendo el proceso propuesto por BOX y JENKINS, (1.976), se va a proponer un modelo, se va a estimar (tanto los órdenes de los polinomios como el valor de los parámetros) y se va a validar. Para ello se hará uso de las funciones de autocorrelación (FAC) y autocorrelación parcial (FACP). Una explicación mas exhaustiva de este proceso se puede encontrar en el texto de URIEL, (1.985). Una vez obtenido el modelo se usa este para realizar predicciones futuras.

En la actualidad existen varios paquetes estadísticos que incorporan módulos que utilizan la metodología de BOX-JENKINS para series univariantes, como BMDP, SPSS, STATGRAPHICS, TSP, En nuestro caso se ha usado el BMDP (módulo 2T), que por otra parte, es uno de los pocos que incorporan función de transferencia.

Utilizando la metodología anterior se ha estimado y validado el siguiente modelo, para la serie Y_t del IPC:

$$(1 - \beta_1 B^6 - \dots - \beta_2 B^{12}) \cdot (1 - \beta_3 B^{12}) \cdot (1 - B) \cdot Y_t = e_t$$

Con las siguientes estimaciones de los parámetros:

Parámetro	Valor estimado	St.Error	T-Radio
β_1	0.1152	0.0417	2.76
β_2	0.8284	0.0433	19.31
β_3	-0.3876	0.0688	-5.72

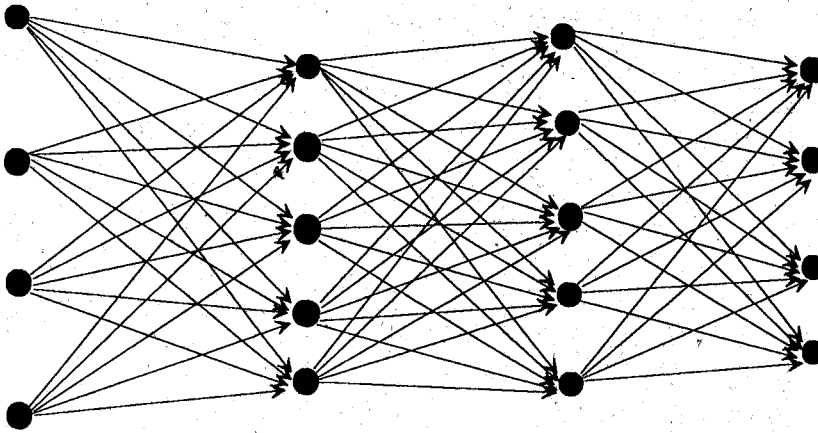
3.- EL PERCEPTRON MULTICAPA

Las redes neuronales artificiales son modelos matemáticos simplificados de las redes de neuronas que se encuentran en el cerebro humano. Actualmente se esta trabajando con diferentes modelos o estructuras de redes neuronales, que se inspiran en menor o mayor medida, en el funcionamiento de alguna parte del cerebro: Mapas Autoorganizativos, Red de Contrapropagación, Adeline y Madeline, ... o incluso en procesos fisicos como La Memoria de Hopfield continua, la Maquina de Boltzmann,...

En nuestro caso vamos a utilizar el modelo denominado Perceptrón Multicapa o también conocido como Red de Propagación hacia atrás. Es, probablemente, el modelo de red neuronal más extendido aunque, curiosamente, el menos inspirado en el funcionamiento del cerebro, según comentan FREEMAN y SKAPURA, (1.994).

El perceptrón multicapa es un conjunto de "neuronas" o unidades de procesamiento distribuidas en un conjunto de capas, de forma que cada neurona de una capa están conectadas con las neuronas de la capa siguiente para mandarlas información. Aunque el número de capas es variable

siempre existe una capa de entrada, una capa de salida que da la respuesta, y al menos una capa intermedia para procesar la información.



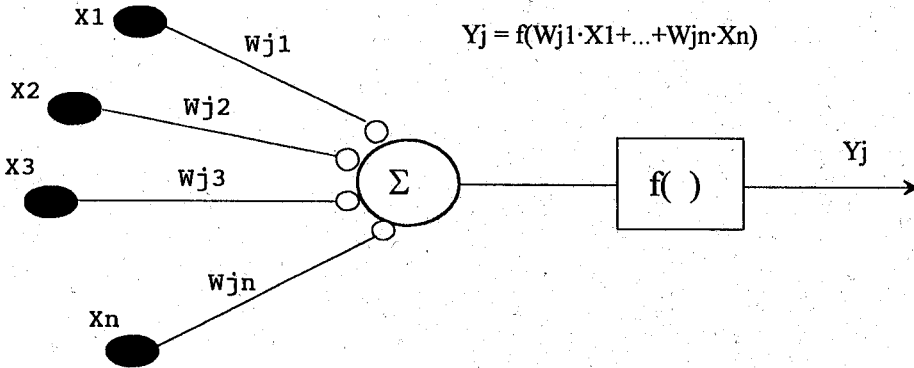
Estructura del perceptrón multicapa con dos capas intermedias

Cada neurona o unidad de procesamiento recibe el conjunto de señales que le envían las neuronas de la capa anterior, la suma de estas señales constituyen el valor de entrada en dicha neurona. A este valor de entrada se le aplica una función de transferencia obteniéndose el valor o señal de salida que envía a las neuronas de la capa siguiente. Las funciones de transferencia más usadas suelen ser de tipo *lineal*, *umbral* ('o todo o nada'), o *sigmoideal*. Este es el tipo que vamos a utilizar, es decir

$$f(x) = \frac{1}{1 + e^{-x}}$$

En algunos modelos se combinan dos de estos tipos de funciones. Por ejemplo, cuando se aplica la función de transferencia sigmoideal si la entrada supera un determinado valor.

Por otra parte cada conexión entre neuronas lleva asociado un valor denominado "peso sináptico", de tal forma que la señal que envía la neurona de partida se ve multiplicando o ponderado por dicho valor.



Modo de funcionamiento de cada neurona o unidad de procesamiento

La forma de trabajar con el Perceptrón multicapa consiste primeramente en diseñar la red, más concretamente, determinar el número de capas y el número de neuronas en cada capa. A continuación se entrena la red para determinar el valor de los pesos sinápticos.

En nuestro caso en cuanto el número de capas se ha optado por trabajar con una sola capa oculta, parece ser la opción más comúnmente usada en los trabajos de este tipo anteriormente mencionados. En cuanto a la distribución del número de neuronas de las diferentes capas, se utilizan k en la capa de entrada, $2 \cdot k + 1$ en la capa oculta y 1 en la capa de salida.

Al utilizar la red para la previsión de una serie temporal basándonos en datos históricos el número k de neuronas de entrada coincidirá con el número de datos anteriores utilizados en dicha predicción. Al tratarse de una serie univariante el número de valores a predecir y por consiguiente el número de neuronas de salida es uno. Hemos elegido el valor de $2 \cdot k + 1$ para el número de neuronas de la capa intermedia aprovechando los resultados reflejados en el trabajo de FUENTE y otros, (1.996).

En la fase de entrenamiento o aprendizaje se presenta un conjunto de casos o vectores de entrenamiento. Cada uno de estos vectores es leído por la red produciéndose una salida. Se compara el valor (o valores) de la capa salida con el valor esperado calculándose el error producido. A continuación se van modificando los pesos sinápticos de los enlaces "hacia atrás" (comenzando por los que finalizan en la capa de salida), de forma que se reduzca el error producido.

Se vuelven a presentar todos los casos, y se vuelve a calcular el error y modificar los pesos. El proceso finaliza hasta que el error se reduzca a niveles aceptables.

Desde hace algún tiempo están apareciendo programas incluso paquetes que utilizan esta este tipo red, NEUROSHEEL, ANSIM,... En nuestro caso hemos optado por programar la simulación

de nuestra propia Red utilizando el compilador. BORLAND PASCAL (ver. 7.0), para lo cual hemos contado con la valiosa ayuda de nuestro compañero ALBERTO ARAGON de la E.U.E. Empresariales de Burgos.

Se ha configurado una red con 14 neuronas en la capa de entrada, 29 en la oculta y una en la salida. Como valores de entrada se toman

$$Z_s, \text{ para } s = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 18, 24;$$

siendo

$$Z_s = Y_s - Y_{s-1};$$

y como valor de salida Z_t .

El número de vectores de entrenamiento es de 167, del 26 al 192. El número de iteraciones que usa la red simulada para entrenar la red es de 6460.

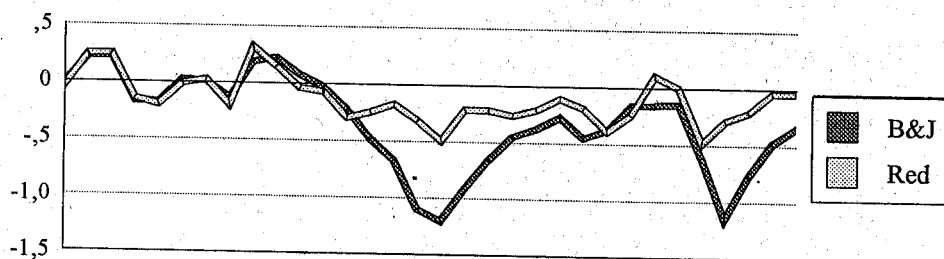
4.- RESULTADOS Y CONCLUSIONES

A continuación se muestra una tabla que resume los resultados de las predicciones de los 32 valores mensuales siguientes, (de Enero de 1.994 a Agosto de 1.996):

Periodo	Valor Real	Box-Jenkins	Error - (B-J)	Red Neuronal	Error - (Red)
193	108.4	108.3880	-0.01196	108.3631	-0.0369
194	108.3	108.5243	0.22428	108.5429	0.2429
195	108.6	108.8312	0.23118	108.8506	0.2506
196	109.2	109.0325	-0.1675	109.0572	-0.1428
197	109.4	109.2225	-0.17749	109.1995	-0.2005
198	109.5	109.5249	0.02486	109.4892	-0.0108
199	110.0	110.0136	0.01363	110.0253	0.0253
200	110.9	110.7641	-0.13593	110.6995	-0.2005
201	111.1	111.2835	0.18353	111.4202	0.3202
202	111.3	111.5242	0.22417	111.4495	0.1495
203	111.5	111.5766	0.07661	111.4592	-0.0408
204	111.9	111.8706	-0.02938	111.8302	-0.0698
205	113.0	112.7719	-0.22813	112.7023	-0.2977
206	113.4	112.9158	-0.48415	113.1529	-0.2471
207	113.9	113.2312	-0.66881	113.7255	-0.1745
208	114.6	113.4994	-1.10059	114.2822	-0.3178

Periodo	Valor Real	Box-Jenkins	Error - (B-J)	Red Neuronal	Error - (Red)
209	114.9	113.6951	-1.20489	114.4059	-0.4941
210	114.9	113.9835	-0.91655	114.6847	-0.2153
211	115.1	114.4406	-0.65942	114.8810	-0.2190
212	115.5	115.049	-0.45099	115.2357	-0.2643
213	115.9	115.5191	-0.38092	115.6879	-0.2121
214	116.1	115.8271	-0.27293	115.9890	-0.1110
215	116.4	115.9463	-0.45367	116.2085	-0.1915
216	116.7	116.3172	-0.38284	116.3052	-0.3948
217	117.3	117.1364	-0.16357	117.0521	-0.2479
218	117.5	117.3444	-0.15558	117.6025	0.1025
219	117.8	117.6566	-0.14343	117.8001	0.0001
220	118.5	117.8829	-0.61705	118.0125	-0.4875
221	119.2	118.0453	-1.15473	118.9121	-0.2879
222	119.1	118.3233	-0.77673	118.8872	-0.2128
223	119.3	118.8123	-0.48766	119.2586	-0.0414
224	119.7	119.3512	-0.34876	119.6582	-0.0418

La evolución de los errores en la predicción se muestra más claramente en este gráfico:



Errores en la predicción del modelo ARIMA (en negro) y del Perceptrón Multicapa

A la vista de estos resultados parece claro concluir que la Red Neuronal se comporta ligeramente mejor, globalmente hablando, que el modelo ARIMA. En cualquier caso la evolución de las curvas de errores que producen ambos métodos son muy parecidos y casi siempre en el mismo sentido.

Es interesante comprobar que este comportamiento ligeramente mejor de La red Neuronal se produce en la predicción de una serie con una clara componente estacional, en donde se podría esperar que los modelos ARIMA predijeran algo mejor a tenor de los resultados obtenidos en el trabajo de DE LA FUENTE y otros, (1.995). En cualquier caso, también hay que señalar que el diseño de la Red Neuronal y la elección de los datos que configurarían los vectores de entrada se ha hecho posteriormente a la estimación del modelo ARIMA, y utilizando los resultados obtenidos por este.

Quizás por esto último podamos concluir que independientemente de que metodología sea mejor, no podemos dejar totalmente de lado ninguna. Incluso, como en este caso, se puede usar como estrategia combinar ambas.

5.- REFERENCIAS Y BIBLIOGRAFIA

ARAGON,A Y GARCIA,A.

"Predicción del Ibex-35 a muy corto plazo, utilizando la metodología de las Redes Neuronales".

Actas del X Congreso de ASEPELT-ESPAÑA. Albacete, Junio de 1.996a.

ARAGON,A Y GARCIA,A.

"Prediction of the Ibex-35 index by usig Neural Network methodology".

Proceedings of the ITHURS 96. León-Spain. 1.996b.

BOX,G. and JENKINS,G.

"Time Series Analisis: Forecasting and Control".

2nd. Ed. San Francisco: Holden Day. 1.976.

DE LA FUENTE,D., PINO,R., SUAREZ,C. y MAYO,J.L.

"Análisis comparativo de los métodos de previsión univariante, Box-Jenkins, Redes Neuronales y Espacios de Estado".

Estudios de Economía Aplicada (ASEPELT-ESPAÑA), n. 5, pp. 5-32. Proyecto Sur. 1.996.

FREEMAN,J and SKAPURA,D.

"Redes Neuronales: Algoritmos, Aplicaciones y Técnicas de Programación".

Addison-Wesley/Diaz de Santos. Madrid. 1.993.

GARCIA,A. y PACHECO,J.

"Prediction of the Ibex-35, using the Boltzmann Machine".

Proceedings of the ITHURS 96. León-Spain. 1.996.

GUILLEN,M. y SOLDEVILLA,C.

"On the Perfomance of the BackPropagation Networks in Econometric Analysis".

Proceedings of the ITHURS 96. León-Spain. 1.996.

OTERO, J.M. y TRUJILLO, F.

"*Predicción Multivariante y Multiperiodo de una Serie Temporal mediante una Red Neuronal*".

Estadística Española. Vol. 35, n. 133, pgs. 345-375, 1.993.

PACHECO, J. y CALZADA, J.M.

"*Análisis Cuantitativo del IPC en Castilla y León. Una comparación con el estado español*".

Actas del 2º Congreso de Economía Regional de Castilla y León. León, Noviembre de 1.990.

SERRANO-CINCA, C.

"*Las Redes Neuronales Artificiales en el Análisis de la Información Contable*".

Tesis Doctoral. Dpto. de Contabilidad y Finanzas de la Universidad de Zaragoza. (1.995).

URIEL, E.

"*Análisis de Series Temporales. Modelos ARIMA*".

Colección *Abaco*. Paraninfo. 1.985.