# NONLINEAR FORECASTING IN ECONOMICS: A COMPARISON BETWEEN COMPREHENSION APPROACH VERSUS LEARNING APPROACH. AN APPLICATION TO SPANISH TIME SERIES

Elena Olmedo Juan M. Valderas Ricardo Gimeno Lorenzo Escot

FUNDACIÓN DE LAS CAJAS DE AHORROS DOCUMENTO DE TRABAJO Nº 2042005 De conformidad con la base quinta de la convocatoria del Programa de Estímulo a la Investigación, este trabajo ha sido sometido a evaluación externa anónima de especialistas cualificados a fin de contrastar su nivel técnico.

ISBN: 84-89116-07-5

La serie **DOCUMENTOS DE TRABAJO** incluye avances y resultados de investigaciones dentro de los programas de la Fundación de las Cajas de Ahorros.

Las opiniones son responsabilidad de los autores.

# Nonlinear Forecasting in Economics: a comparison between Comprehension Approach versus Learning Approach. An Application to Spanish Time Series

Elena Olmedo<sup>1</sup>, Juan M. Valderas<sup>1</sup>, Ricardo Gimeno<sup>2</sup> and Lorenzo Escot<sup>3</sup>

#### **Summary**

In this paper alternative non-parametric forecasting techniques are analysed, making emphasis in the difference between the reconstruction and the learning approach. The first one is based on Takens Theorem that recovers unknown dynamic properties of the system; it is appropriate in deterministic systems. The second one is a powerful instrument in noisy systems. These techniques are applied to the forecasting of Spanish unemployment and the Industrial Production Index using one-to-one forecasting and comparing the results with the one obtained using nonlinear models, specifically TAR model.

**JEL Classification:** B41, C14, C32, C45, C51, C53

**Keywords:** Forecasting, reconstruction, neural networks, Takens Theorem, TAR model.

<sup>&</sup>lt;sup>1</sup> Dep. Economía Aplicada I. Universidad de Sevilla. Fac. Ciencias Económicas y Empresariales, Avda. Ramón y Cajal 1, 41018 Sevilla, Spain. olmedo@us.es

<sup>&</sup>lt;sup>2</sup> Dep. Métodos Cuantitativos. Universidad Pontificia Comillas. C/Alberto Aguilera 23, 28015 Madrid, Spain. rgimeno@upco.es

<sup>&</sup>lt;sup>3</sup> Dep. Economía Aplicada III. Universidad Complutense de Madrid. Fac. Ciencias Económicas y Empresariales, Campus de Somosaguas, 28223 Madrid, Spain. escot@ccee.ucm.es

#### **Forecasting in Economics**

One of the main purposes of science is forecasting, that is, the knowledge of future evolution of analysed variables. Models, which constitute simplified depictions of reality, are usually used to make predictions. In Economics, we use the data from time series to get knowledge about the model which can generate the observed behaviour, and after that, we use the model to make forecasting. In traditional time series analysis, a common approach involves linear models which conceive time series as outcomes of a linear stochastic process. The "forecasting paradox" is a common feature of these models, in the sense that we do not usually get good forecasting in spite the goodness of fit.

Nowadays, we can speak about three different approaches in the analysis and forecasting of time series. Reconstruction methodology is very useful if we are dealing with a low dimensional deterministic system, because we are recovering the topology and the dynamic of the original unknown system, so the reconstructed system will probably allow us to make better forecasting than those achieved with traditional techniques. On the other hand, if we are working with a complex system (of high dimension or with a high level of noise) it seems unquestionable nowadays the effectiveness achieved in forecasting shown by the learning approach based upon neural networks. However, we can not forget the impulse experimented by nonlinear modelling after the re-discovering of chaos theory and the property of sensible dependence of initial conditions, which allows complex behaviour in simple deterministic systems.

#### Thus, we can speak about:

- The comprehension approach, by means of reconstruction theorem, which is suitable when the unknown underlying system has a low dimension because, in this case, the rebuilt surface is simple enough to be fitted quite approximately and hence we will be able to use the reconstruction for the task of prediction.
- The learning approach, by virtue of neural networks, which it is more adequate when the sistem is not simple enough to be reconstructed.
- The nonlinear models, which applying parametric techniques of fitting (nonlinear) functions to data.

The first and the second are forecasting techniques inside which what some authors call non-parametric forecasting because no data generator model is established a priori. Nieto de Alba (Nieto de Alba, 1998) states that the key idea which makes a difference between "chaotic econometrics" and "standard econometrics" is the existence of inference techniques which do not rely on the knowledge of a data generator model, it is nothing else but non-parametric inference. In this paper, we are trying to analyse the behaviour of these techniques in economic time series, and compare the results obtained applying nonlinear modelling.

First in this paper we are going to briefly describe the methods we are going to use and latter we will use them to forecast Spanish unemployment evolution.

#### The problem of prediction

Casdagli (Casdagli, 1989) works with the problem of nonlinear forecasting as it was an "inverse problem" within the theory of dynamic sytems. The usual problem in this field is, having knoledge of the model, to describe its evolution as time passes by; the inverse problem consists of given the evolution of the system, stated by collected data, building a model which may have generated them. If we set up a mapping, usually nonlinear, which may have generated those data set, we might have a predictive model.

To identify this mapping, we generally split up data in two groups; the first group, containing the training data, is used to make an interpolation or approximation to the unknown mapping and the second group is used to evaluate the goodness of fit in terms of the prediction errors.

Let  $f:R^m \to R^m$  a smooth mapping which owns an strange attractor  $\Lambda$  with a natural invariant ergodic measure, so  $x_n = f^n \big( x_0 \big)$   $1 \le n < \infty$ , is a typical sequence of iterations inside of the attractor. An inverse problem tries to build a smooth mapping  $\tilde{f}_\infty: R^m \to R^m$  so that  $x_{n+1} = \tilde{f}_\infty \big( x_n \big)$ ,  $1 \le n < \infty$ . Since these points are dense in the attractor and the mapping is smooth, there exists a unique solution to the inverse problem, and hence this mapping is a perfect predictive model inside the strange attractor.

When we have a finite number of iterations  $x_n$ ,  $1 \le n \le N$ , our problem consists in locate a smooth mapping  $\tilde{f}_N : R^m \to R^m$  such that  $x_{n+1} = \tilde{f}_N\left(x_n\right)$ ,  $1 \le n \le N-1$ . If m=1 we can pick up whatever smooth mapping whose graphic intercepts the points  $\left(x_n, x_{n+1}\right)$ ,  $1 \le n \le N-1$ . This mapping would be our predictor mapping. In order to select the best predictor mapping, we use the prediction error or cost function of  $\tilde{f}_N$ .

To verify the predictor mapping working with real data, data set is split up in two groups: the training data to fit the predictor mapping and the verification data to test the achieved predictor mapping.

## Forecasting techniques in the comprehension approach by the reconstruction of the space state. Local methods

This techniques are based in the Takens Reconstruction Theorem, which basic ideas we are going to expose. We suppose that the evolution of our dynamic system is consequence of the dynamic of a continuous mapping F which sets a relationship between the present and the future state of the system, that is,  $x_n = F\left(x_{n-1}\right)$ . So we want to estimate  $\hat{F}$  using the data from the time series. But these data are not directly the different states of the underlying system, but their measurements using an unknown mapping h. This is the reason why the Takens Reconstruction Theorem is neccessary, in order to work with a reconstructed system that keeps the dynamic of the original. Takens stated that, under certain conditions, there is a dipheomorphism between original dynamics of the unknown dynamic system (F) and dynamics of the reconstructed system  $(\Psi)$ . So that the the forecasting problem is to find  $\hat{\psi}_N$  such that

$$a_m(t+1) = \hat{\psi}_N(a_m(t)) \quad 1 \le t \le N-m$$

where  $a_m(t) = (x(t), x(t-\tau), ..., x(t-(m-1)\tau))$  are called *m*-histories, being *m* the embedding dimension. For embedding dimension greater than one, we will consider orthogonal projections  $\pi_i \hat{\psi}_N : R^m \to R$  i = 1, ..., m to get the predicted value. That is (Wan, 1994), theorically, there is a nonlinear autoregression of the kind

$$x_{(n+m)\tau} = f(x_{(n+m-1)\tau}, ..., x_{n\tau})$$

So we try to estimate this  $\tilde{f}$ . To test if our prediction is good, we compute the predictor error starting with a low embedding dimension and repeat calculations until that we achieve an error level small enough. At last, we must realise that there are various interpolation techniques, we will choose the one who gives us the lowest error but do not increase the required calculations.

We are going to apply near neighbours methods to get this estimation (see Loren, 1963, Farmer and Sidorowich, 1987 and 1998 and Kantz and Schreiber, 1994). These methods are used when the observed dynamic is so complex so as to fit a local predictor function. Local models can be seen like the linearization of the lobal nonlinear dynamic considering points of the neighbourhood of the one we want to forecast. To forecast x(t+T) using near neighbours method, we first

impose a metric in the state space that is denoted by  $\|\cdot\|$  and then we locate a number k of m-histories, the closest to  $a_m(t)$ , that is, k m-histories  $a_m(t')$  for t' < t with the lowest distance  $\|a_m(t) - a_m(t')\|$ . Then a local predictor is built, taking every neighbour m-history as a point in the origin set, and the respective x(t'+T) as a point in the image set. We look for a a mapping that fits the k points  $\left(a_m(t'), x(t'+T)\right)$ . The simplest approach is to consider only the closest neighbour. In this particular case,  $\hat{x}(t+T) = x(t'+T)$ . In the general case, a linear mapping (arithmetic mean, weighted mean, linear regression) is fitted to the points  $\left(a_m(t'), a_m(t'+T)\right)$  by minimum square errors. This polinomy will be the predictor mapping.

### Forecasting by learning with neural networks. Global methods

When the system is high-dimensional or noisy, the reconstruction approach do not provide good forecasts, and the learning approach with neural networks is more suitable because, due to its generality, a neural network can replicate whatever nonlinear mapping behaviour. Neural networks (see, for example, Nychka et al., 1992, Jungeilges, 1996 and White, 1989) are a kind of non-linear models inspired in the brain's neural architecture. They were developed in field of Artificial Intelligence as an attempt of modelling the learning ability of biological systems by brain's structure reproduction. Neural networks illustrate the idea of understanding the learning process as a recursive statistical process. They are input-output models because from an input vector x, the neural network generates an output vector y = g(x), where the mapping g is determined by neural structure. Neural networks are then global methods, because they try to find an estimation of the function that governs the global dynamics of the system.

Similar to the techniques analysed in the previous section, neural networks make an effort to solve the so-called "inverse problem" of fitting a non-linear mapping using the successive values of an observed time series. We can sign out two major advantages of these techniques respect to comprehension techniques:

 Recent developments in neural networks literature give support to the theoretical backgrounds about the universality of feedforward networks, which will be defined later, to approximate mappings because feedforward neural networks with an arbitrary number of neurons can

- approximate both uniformly continuous function and the derivative of an arbitrary function.
- When fitting, the number of parameters increases linearly with the order of approximation, whereas in most of the models the growth of parameters with the order of approximation is exponential.

A neural network is an interconnected structure, composed by different layers; each layer has a number of neurons connected with neurons in the next layer, so that each neuron in a layer feeds every neuron in the next layer. Every neuron generates an output according to a mapping, called activation mapping, which is a weighted linear combination of the inputs received from neurons in the former layer. These weights are fitted with respect to a specified learning algorithm which tries to minimize the errors, discrepancy between the input vector and the output of the network, computed in an objective function.

The simplest case is a neural network without hidden layers: it consists of an output unit which computes a weighted linear combination of d inputs:  $out^{(t)} = \sum_{i=1}^d \omega_i x_i^{(t)}$  where the superindex t denotes an specific patron,  $x_i^{(t)}$  is the value of  $i^{th}$ -esime patron input and  $\omega_i$  stands for the weight of the  $i^{th}$ -esime input to produce the output. Given this relationship input-output, the main task of learning refers to find the way to change the weights so that the output generated by the neural network stands close to the desired output which is usually named target<sup>(t)</sup>. The proximity or discrepancy between both is expressed through a cost function, for instance the mean square error  $E^{(t)} = \left(out^{(t)} - target^{(t)}\right)^2$ . A learning

The key feature of neural networks, responsible of their power and popularity, is the possibility of inserting within the architecture of the network, one or more non-linear hidden layers between the input and output layers. These non-linearities make possible the interaction among inputs and hence it is possible to fit mappings that are more complex.

The simplest non-linear neural network consists of just one hidden layer and it is composed of the next components:

d inputs

algorithm is used to fit the weightns.

- inputs are connected to a non-linear hidden layer of units
- hidden units are connected to the output linear unit
- output and hidden units have adjustable biases b
- weights can be positive, negative or null

The answer of a single unit is called the *activation value* or simply *activation*. A common choice for the non-linear activation function of hidden units is a composition of two operators, the mapping trough an activation function of a

linear function. First, the incoming inputs to the hidden unit h are linearly connected and a bias  $b_h$  is added

$$\xi_{h}^{(t)} = \sum_{i=1}^{d} \omega_{hi} x_{i}^{(t)} + b_{h}$$

After this, the output is generated mapping  $\xi_h^{(r)}$  through a transference function or activation. Generally, the kind of activation functions most commonly used for practical purposes are *sigmoidal functions* which, in addition to its similarity in bold strokes to the biological behaviour of a neuron, generate, whatever the input may be, a bounded output. Moreover, sigmoidal functions are continuous and differentiable, which is more operational for empirical purposes. Logistic function is commonly used in practical applications:

$$S\left(\xi_h^{(t)}\right) = \frac{1}{1 - e^{-a\xi_h^{(t)}}}$$

where the slope a establishes the smoothness of the answer.

When we try to fit the weights when the desired outputs of the hidden units are unknown, we make use of the chain rule to solve this problem. Weights are fitted giving small steps towards the direction signed out for a negative value of the gradient. The error of the weights between the hidden layer h and the output layer for a given patron is now  $(\operatorname{out}^{(t)}\text{-target}^{(t)}) \times S'\left(\xi_h^{(t)}\right)$ . When we are dealing with weights not connected in a straight way with the output, the error is computed recursively in terms of the errors of the units which are directly connected to the hidden unit and the weights of these connections.

Among the various types of neural networks, the most commonly used in the field of time series forecasting are *feedforward neural networks*. The network structure is given in Figure 1, where a scheme of a feedforward neural network with one hidden layer is shown. We will call this *single-layer feedforward network* (SLFFN). The input values are  $x_1$  y  $x_2$ .

Inputs are collected in the two input units which just simply send the inputs to the hidden units  $u_i$ . Every connection, signalled in the scheme with arrows, makes a linear mapping of inputs according to the intensity factor or weight. The output of the hidden unit consists in a non-linear mapping of the linearly mapped input according to an activation function. This activation function is the same for every unit, but each of them has its own bias  $\gamma_{i0}$  which is equivalent to an external input or simply the intrinsic neural activity level. The activation function  $\psi(u)$  is

generally a sigmoid<sup>1</sup>, with two asymptotes,  $\psi(u)$  tends to 0 when  $u \to -\infty$  and tends to 1 when  $u \to \infty$ .

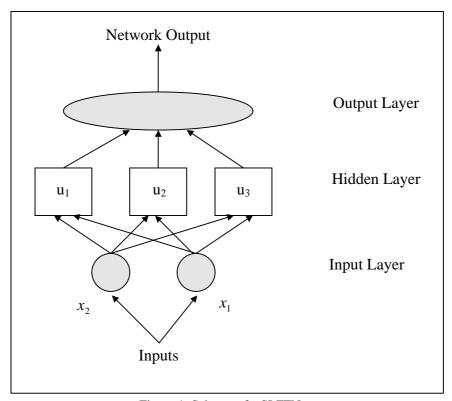


Figure 1. Scheme of a SLFFN

The outputs of the hidden layer go to the output layer where a linear mapping of the total input is undertaken. In a feedforward neural network, at instant t, with n inputs  $x_t = (x_{1,t},...,x_{n,t})$  and q units in the hidden layer, the output of this network is

$$o_{t} = \Phi\left(\beta_{0} + \sum_{i=1}^{q} \beta_{i} \Psi\left(\gamma_{i0} + \sum_{j=1}^{n} \gamma_{ij} x_{j,t}\right)\right) \equiv f\left(x;\theta\right)$$

where  $\theta = (\beta_0,...,\beta_q,\gamma_1,...,\gamma_q)$  and  $\gamma_j = (\gamma_{j0},...,\gamma_{jn})$  are the parameters to be fitted and functions  $\Psi$  and  $\Phi$  (in the previous example  $\Phi$  was supposed to be

the identity function) are known activation functions,  $o_t$  can be considered a target estimator,  $y_t$  and  $x_t$  is the inputs vector.

So, given a sequence of inputs x(k), the neural networks produces a sequence of outputs O(W,x(k)), where W stands for the set of operating coefficients. These parameters are adjustable, so we speak about adaptative memory. This "learning ability" may improve forecasting, that is what we mean when we say that there is some learning ability. Mozer (Mozer, 1994) quoted that there are several algorithms which operate in that way; we point out "backforwarding algorithm" which is explained more thoroughly by Wan (Wan, 1994) and that basically gives small steps according to the parameter  $\eta$ , the learning rate, in the direction where the gradient  $\nabla f$  decreases more quickly, that is subtracting the local gradient. The new weights so attained are expresses in terms of the previous ones according to the next expression:

$$\hat{\boldsymbol{\theta}}_{t+1} = \hat{\boldsymbol{\theta}}_{t} + \eta \nabla f\left(\boldsymbol{x}_{t}, \hat{\boldsymbol{\theta}}_{t}\right) \cdot \left[\boldsymbol{y}_{t} - f\left(\boldsymbol{x}_{t}, \hat{\boldsymbol{\theta}}_{t}\right)\right]$$

where  $y_t$  is the desired output. So the weights fitted are those which minimize the cost function, generally defined as the mean square error between the desired output and the net output using the weights.

The proceeding with this is algorithm is as follows: let us suppose that we want to estimate y(k) in the time series, the known input in the neural network is y(k-1), and the output is given by

$$\hat{y}(k) = o(y(k-1))$$

In the training session, the square errors  $e(k)^2 = (y(k) - \hat{y}(k))^2$  are minimized according to "backforwarding algorithm", knowing that both the input and the desired outcome, which we had denoted as d(k) are given by the training data. This is what is called open-loop learning because the network output is not an input itself.

#### TAR model

TAR model was enunciated by Tong [Tong, 1993] and has been used frequently in several different disciplines always associated with the existence of saturation

levels, which are known are thresholds. The basic idea behind this model is simply local approximation, that is, the introduction of differente regimes using these thresholds so that we can analyse a complex system desomposing it in different linear and simple systems.

So the introduction of nonlinearity in this models is using local approximations, introducing different kind of behaviours depending on the belonging to different situations. Concretely, the simplest threshold autorregresive model distinguís two autorregresive regimes with one unique lag depending on the past values of the variable compared with the threshold value. Formally this model is:

$$x_{t} = \begin{cases} \beta_{1}^{1} + \beta_{2}^{1} x_{t-1} + \varepsilon_{t} & \text{if } x_{t-1} < \text{threshold} \\ \beta_{1}^{2} + \beta_{2}^{2} x_{t-1} + \varepsilon_{t} & \text{if } x_{t-1} \ge \text{threshold} \end{cases}$$

where  $\varepsilon_t$  are random shocks with usual behaviour.

Note that this threshold model can be considered as an extended usual linear model, so that can be estimated using OLS.

## Application to Spanish economic time series forecasting

We have selected two different cojunctural time series. The first one is the monthly evolution of Spanish unemplyment, data elaboreted by the INEM<sup>2</sup>, from January 1964 to March 2001, 447 observations. And the second one is the Industry Production Index from January 1975 to August 2004, 356 observations. Both time series have enaugh data to apply the prediction methods explained above. These techniques need that data must be stationary. In unemployment data we have taken logarithm and applied the Hodrick-Prescott filter to finally test the existence of unit roots. In this sense we use the usual Dickey-Fuller test and the Franses test to verify the existence of monthly unit roots. In IPI data after taken logarithm and taken unseasonal component, we applied the Hodrick-Prescott filter. The results are in Table 1 and Table 2.

Dickey-Fuller test			
ADF Statistic	Unemployment	-10,5414	
ADI Statistic	IPI	-5,7618	
1% p-value		-2,5703	
5% p-value		-1,9402	

T-1-1-1 D'-1 F-11	1 11 36 77	
10% p-value	-1.616	

Table 1. Dickey-Fuller unit root test (tabulated by MacKinnon)

Franses Test					
	Unem	IPI			
Null Hypothesis	t or F statistic	Filter to apply	t or F statistic		
$\pi_1 = 0$	-8,434395		-3.5110203*		
$\pi_2 = 0$	-2,422770	(1+L)	-11.325054*		
$\pi_3 = \pi_4 = 0$	9,425708		26.490689*		
$\pi_5 = \pi_6 = 0$	7,032434		96.654696*		
$\pi_7 = \pi_8 = 0$	7,899472		28.463917*		
$\pi_9 = \pi_{10} = 0$	12,14589		36.651973*		
$\pi_{11} = \pi_{12} = 0$	7,622910		22.144383*		

Table 2. Franses Test of mensual stationarity

We used the following methodology: we split the data in two sets, the first with 350 observations and the second with 50 and we used the first set to find the neighbours in the second. Nevertheless, since in the first part we worked with one-to-one predictions, we forecasted just one-step forward, once the value was predicted, the actual value joins the first group so that it can be used to find the neighbours in the next value.

We have computed forecasting error for embedding dimensions from two to ten and for a different number of neighbours. The minimum error achieved will give us the adequate dimension to use in the forecasting and the number of neighbours to be used.

We have taken into consideration a number of neighbours ranging from one to one hundred, and the embedding dimension ranging from one to ten. We found out that the mean square error was minimum for embedding dimensions from six to eight and smaller for an average number of neighbours. More precisely, the minimum for the mean square error as well as the minimum for the proportion of the mean square error respect to the variance, is achieved by an embedding dimension of seven with fifteen neighbours. Consequently we cannot talk of a global model and the presence of non-linearity is confirmed.

The Table 3 shows the comparative results of the MSE for the best choice of the three methods with respect to the variance of the real series that is forecasted.

	Ur	nemployme	nt		IPI	
FORECASTING METHOD	Embedding	Number of	MSE/var	Embedding	Number of	MSE/var

		dimension	neighbours		dimension	neighbours	
Camanahanaian	Mean	6	10	0,0652	2	10	0,1135
Comprehension approach	Weighted mean	9	5	0,1243	6	5	0,3596
арргоасп	Linear regression	9	20	0,0425	8	20	0,0196
Learning approach		9		0.9749			-

Table 3. Minimum MSE/var of one-to-one predictions

0.2767

0.1177

TAR model

As we can see in the table above, the better predictions are achieved using liner regresion in the comprehension approach, followed by the estimation of TAR model. Nonetheless, in general the results are better using the reconstruction approach because of the good estimation of turning points. We show the forecasted and real time series in both cases using both methods in both time series.

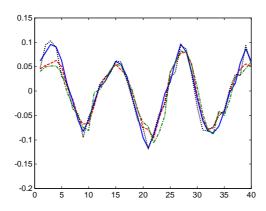


Figure 2. Reconstruction methods. Unemployment time series

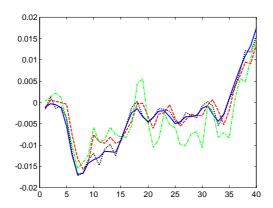


Figure 3. Reconstruction methods. IPI time series

In both figures the real time series is in blue, the forecasted time series using mean is in red, the forecasted time series using weighted mean is in green and using linear regression is in blak.

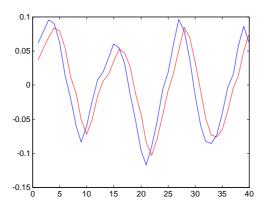


Figure 4. TAR model estimation. Unemployment time series

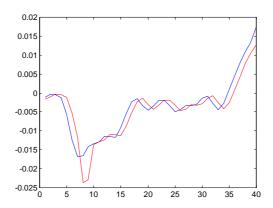


Figure 5. TAR model estimation. IPI time series

In both figures original time series is depicted in blue and forecasted values in red. We can see the goodness of the prediction, but we also note that the forecasted values seems to be lagged with respect the real ones. In other words, the model does not adjust the turning points.

#### **Conclusions**

Hence, we can conclude that, the analysis of the behaviour of the different techniques in the reconstruction approach shows that these techniques provide, in general, good predictions, except in the case of neural nets. The better results are achieved in both time series using reconstruction methods, particularly the linear regression of twenty neighbours. In this sense, it is very important the captation of turning points one period before.

#### References

CASDAGLI, M. (1989). Nonlinear prediction of chaotic time series. *Physica D* 35: 35-56

FARMER, J.D. and SIDOROWICH, J.J. (1987). Predicting chaotic time series. *Physical Review Letters* 59: 845-848

JUNGEILGES, J.A. (1996). Operational characteristics of White's test. In *Nonlinear dynamics and economics*, BARNETT, W.A. AND KIRMAN, A.P. and SALMON, M. (eds). Cambridge University Press: Cambridge

KANTZ, H. and SCHREIBER, T. (1994). Nonlinear time series analysis. Cambridge University Press: Cambridge

LORENZ, E.N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences* 20, 2: 130-141

MOZER, M.C. (1994). Neural net architectures for temporal sequence processing. In *Time series prediction: forecasting the future and understanding the past*. WEIGEND, A.S. and GERNSHENFELD, N.A. (eds.) Addison-Wesley: Reading, MA

NIETO DE ALBA, U. (1998). Historia del tiempo en economía. McGraw Hill: Madrid

NYCHKA, D., ELLNER, S., GALLANT, A.R. and MCCAFFREY, D. (1992). Finding chaos in noisy systems. *Journal of the Royal Statistical Society B* 54, 2: 399-426

WAN, E. (1994). Time series prediction by using a connectionist network with internal delay lines. In *Time series prediction: forecasting the future and understanding the past*. WEIGEND, A.S. and GERSHENFELD, N.A. (eds.) Addison-Wesley: Reading, MA

WHITE, A. (1989). Some asymptotic results for learning in single hidden-layer feedforward network models. *Journal of the American Statistical Association* 84, 408: 1003-1013

<sup>1</sup> As we stated before, this kind of functions is chosen because it reflects in bold strikes the threshold properties of real neurons. Moreover, because sigmoidal functions are bounded, hence the errors are bounded

<sup>2</sup> INEM stands for the "Instituto Nacional de Empleo" meaning National Institute for the Employment

15

#### FUNDACIÓN DE LAS CAJAS DE AHORROS

#### **DOCUMENTOS DE TRABAJO**

#### Últimos números publicados

159/2000	Participación privada en la construcción y explotación de carreteras de peaje Ginés de Rus, Manuel Romero y Lourdes Trujillo
160/2000	Errores y posibles soluciones en la aplicación del <i>Value at Risk</i> Mariano González Sánchez
161/2000	Tax neutrality on saving assets. The spahish case before and after the tax reform Cristina Ruza y de Paz-Curbera
162/2000	Private rates of return to human capital in Spain: new evidence F. Barceinas, J. Oliver-Alonso, J.L. Raymond y J.L. Roig-Sabaté
163/2000	El control interno del riesgo. Una propuesta de sistema de límites riesgo neutral Mariano González Sánchez
164/2001	La evolución de las políticas de gasto de las Administraciones Públicas en los años 90 Alfonso Utrilla de la Hoz y Carmen Pérez Esparrells
165/2001	Bank cost efficiency and output specification Emili Tortosa-Ausina
166/2001	Recent trends in Spanish income distribution: A robust picture of falling income inequality Josep Oliver-Alonso, Xavier Ramos y José Luis Raymond-Bara
167/2001	Efectos redistributivos y sobre el bienestar social del tratamiento de las cargas familiares en el nuevo IRPF Nuria Badenes Plá, Julio López Laborda, Jorge Onrubia Fernández
168/2001	The Effects of Bank Debt on Financial Structure of Small and Medium Firms in some European Countries Mónica Melle-Hernández
169/2001	La política de cohesión de la UE ampliada: la perspectiva de España Ismael Sanz Labrador
170/2002	Riesgo de liquidez de Mercado Mariano González Sánchez
171/2002	Los costes de administración para el afiliado en los sistemas de pensiones basados en cuentas de capitalización individual: medida y comparación internacional.  José Enrique Devesa Carpio, Rosa Rodríguez Barrera, Carlos Vidal Meliá
172/2002	La encuesta continua de presupuestos familiares (1985-1996): descripción, representatividad y propuestas de metodología para la explotación de la información de los ingresos y el gasto. Llorenc Pou, Joaquín Alegre
173/2002	Modelos paramétricos y no paramétricos en problemas de concesión de tarjetas de credito. Rosa Puertas, María Bonilla, Ignacio Olmeda

174/2002	Mercado único, comercio intra-industrial y costes de ajuste en las manufacturas españolas. José Vicente Blanes Cristóbal
175/2003	La Administración tributaria en España. Un análisis de la gestión a través de los ingresos y de los gastos. Juan de Dios Jiménez Aguilera, Pedro Enrique Barrilao González
176/2003	The Falling Share of Cash Payments in Spain. Santiago Carbó Valverde, Rafael López del Paso, David B. Humphrey Publicado en "Moneda y Crédito" nº 217, pags. 167-189.
177/2003	Effects of ATMs and Electronic Payments on Banking Costs: The Spanish Case. Santiago Carbó Valverde, Rafael López del Paso, David B. Humphrey
178/2003	Factors explaining the interest margin in the banking sectors of the European Union. Joaquín Maudos y Juan Fernández Guevara
179/2003	Los planes de stock options para directivos y consejeros y su valoración por el mercado de valores en España. Mónica Melle Hernández
180/2003	Ownership and Performance in Europe and US Banking – A comparison of Commercial, Cooperative & Savings Banks. Yener Altunbas, Santiago Carbó y Phil Molyneux
181/2003	The Euro effect on the integration of the European stock markets. Mónica Melle Hernández
182/2004	In search of complementarity in the innovation strategy: international R&D and external knowledge acquisition. Bruno Cassiman, Reinhilde Veugelers
183/2004	Fijación de precios en el sector público: una aplicación para el servicio municipal de suministro de agua.  Mª Ángeles García Valiñas
184/2004	Estimación de la economía sumergida es España: un modelo estructural de variables latentes. Ángel Alañón Pardo, Miguel Gómez de Antonio
185/2004	Causas políticas y consecuencias sociales de la corrupción. Joan Oriol Prats Cabrera
186/2004	Loan bankers' decisions and sensitivity to the audit report using the belief revision model. Andrés Guiral Contreras and José A. Gonzalo Angulo
187/2004	El modelo de Black, Derman y Toy en la práctica. Aplicación al mercado español. Marta Tolentino García-Abadillo y Antonio Díaz Pérez
188/2004	Does market competition make banks perform well?. Mónica Melle
189/2004	Efficiency differences among banks: external, technical, internal, and managerial Santiago Carbó Valverde, David B. Humphrey y Rafael López del Paso

190/2004	Una aproximación al análisis de los costes de la esquizofrenia en españa: los modelos jerárquicos bayesianos F. J. Vázquez-Polo, M. A. Negrín, J. M. Cavasés, E. Sánchez y grupo RIRAG
191/2004	Environmental proactivity and business performance: an empirical analysis Javier González-Benito y Óscar González-Benito
192/2004	Economic risk to beneficiaries in notional defined contribution accounts (NDCs) Carlos Vidal-Meliá, Inmaculada Domínguez-Fabian y José Enrique Devesa-Carpio
193/2004	Sources of efficiency gains in port reform: non parametric malmquist decomposition tfp in- dex for Mexico Antonio Estache, Beatriz Tovar de la Fé y Lourdes Trujillo
194/2004	Persistencia de resultados en los fondos de inversión españoles Alfredo Ciriaco Fernández y Rafael Santamaría Aquilué
195/2005	El modelo de revisión de creencias como aproximación psicológica a la formación del juicio del auditor sobre la gestión continuada Andrés Guiral Contreras y Francisco Esteso Sánchez
196/2005	La nueva financiación sanitaria en España: descentralización y prospectiva David Cantarero Prieto
197/2005	A cointegration analysis of the Long-Run supply response of Spanish agriculture to the common agricultural policy José A. Mendez, Ricardo Mora y Carlos San Juan
198/2005	¿Refleja la estructura temporal de los tipos de interés del mercado español preferencia por la li- quidez? Magdalena Massot Perelló y Juan M. Nave
199/2005	Análisis de impacto de los Fondos Estructurales Europeos recibidos por una economía regional: Un enfoque a través de Matrices de Contabilidad Social M. Carmen Lima y M. Alejandro Cardenete
200/2005	Does the development of non-cash payments affect monetary policy transmission? Santiago Carbó Valverde y Rafael López del Paso
201/2005	Firm and time varying technical and allocative efficiency: an application for port cargo handling firms  Ana Rodríguez-Álvarez, Beatriz Tovar de la Fe y Lourdes Trujillo
202/2005	Contractual complexity in strategic alliances Jeffrey J. Reuer y Africa Ariño
203/2005	Factores determinantes de la evolución del empleo en las empresas adquiridas por opa Nuria Alcalde Fradejas y Inés Pérez-Soba Aguilar
204/2005	Nonlinear Forecasting in Economics: a comparison between Comprehension Approach versus Learning Approach. An Application to Spanish Time Series Elena Olmedo, Juan M. Valderas, Ricardo Gimeno and Lorenzo Escot