TIME SERIES FORECASTS VIA WAVELETS: AN APPLICATION TO CAR SALES IN THE SPANISH MARKET[1]

BY

MIGUEL A. ARINO

IESE
UNIVERSIDAD DE NAVARRA
Avda. Pearson 21
08034 Barcelona
SPAIN

aarino@iese.es

Abstract: In this paper we propose to apply wavelet theory in forecasting economic time series. The method consists in decomposing the series into its long-term trend and its seasonal component according to the shape of the scalogram of the discrete wavelet transform of the series. Each component is then extended to provide a forecast of the total series. The method is applied to the data set of monthly car sales in the Spanish market and the results are compared with the Box-Jenkins forecast of the series.

Key words: Forecasts, Arima models, Wavelets.

[1]This article has been prepared while the author was visiting the Institute of Statistics and Decision Sciences at Duke University. The author acknowledges the invaluable help of Brani Vidakovic during the research on wavelet methodology.

1. Introduction.

Our purpose in this paper is to present a methodology for forecasting univariate time series. This methodology combines standard forecasting techniques with "wavelet methodology" described in this paper. The recently developed wavelet theory has proven to be a useful tool in the analysis of some problems in engineering and related fields. However, the potential of this theory -which will be outlined in the next section- for analyzing economic problems has not been fully exploited yet. This paper presents one of its many possible applications in this field. As an example we will apply the methodology to forecast car sales in the Spanish market and compare the results with those given by standard forecasting techniques. Roughly speaking, using wavelets, we decompose a time series (x_t) in its trend (y_t) and its seasonal component (z_t). The way of decomposing the series is explained in Arino and Vidakovic (1995) and is outlined later in this paper. Then we apply standard forecasting techniques to each component to obtain a forecast of the original series (x_t).

This paper is organized as follows: After this introduction, a brief summary of wavelets is presented together with the discrete wavelet transform (DWT) and the application of wavelet analysis to decompose a time series into its long-term trend and its seasonal component. In Section 3 we apply the wavelet analysis described in the previous section to the data set of monthly car sales in the Spanish market. The seasonal component and the long-term trend of the series are found and each one of them is forecasted ultimately providing a forecast of the total series. In the last section we present our conclusions.

2. Wavelets.

In this section we first present wavelets briefly. Then we review the discrete wavelet transform, which is the wavelet counterpart to the discrete Fourier transform. Finally we show the splitting of a time series into cyclical components by using wavelet analysis. As in Fourier analysis, there are continuous and discrete versions of wavelet analysis. Although for our purposes it is sufficient to deal with the discrete version, we provide a general summary of wavelets for decompose continuous functions, as well. Since we will be dealing with discrete data sets, our focus will be on the discrete wavelet transform.

2.1. An introduction to wavelets.

The wavelet series representation of a square-integrable function f is in some way, similar to the Fourier series representation. The Fourier series representation of a function takes as orthonormal basis for its expansion the simple harmonics $\{e^{inx}\}$, while for the wavelet representation different sets of orthonormal bases $\{\psi_{jk}(x)\}$ satisfying some conditions are used.

If f is a square-integrable function defined on the interval $[0, 2\pi]$, the Fourier series representation of f is

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \, e^{inx}$$

where the constants c_n are defined as

$$c_n = 1/2\pi \int_0^{2\pi} f(x) \, e^{inx} dx.$$

A characteristic of this Fourier series representation is that the orthonormal basis $\{w_n\}$ on which the function f is expanded, is generated by dilation of a single function $w(x) = e^{ix}$, that is $w_n(x) = w(nx)$.

For the wavelet series representation of a function, we expand that function in terms of some orthonormal base $\{\psi_n\}$ different from the trigonometric base. Each one of these possible bases is generated by translations and dilations of a single wavelet "mother" as the trigonometric basis is generated by dilations of the function e^{ix}.

Given a mother wavelet mother ψ, the wavelet ψ_{jk} is obtained from ψ by dilating by the factor 2^j and translating by $2^{-j}k$ as follows:

$$\psi_{jk}(x) = 2^{j/2} \, \psi(2^j x - k).$$

The coefficient $2^{j/2}$ is a scaling factor. Under suitable conditions on ψ (see Chui, 1992 and Meyer, 1992) the set $\{\psi_{jk}\}$ forms an orthonormal basis for the space of the square-integrable functions. The wavelet series representation of a function f is

$$f(x) = \sum_{j,k} f_{jk} \, \psi_{jk}(x)$$

where f_{jk} are the wavelet coefficients of f with respect to the basis $\{\psi_{jk}\}$, which can be found by

$$f_{jk} = \int_{-\infty}^{\infty} f(x)\, \psi_{jk}(x)\, dx.$$

The most simple example of wavelet basis is the Haar basis generated by the “mother” wavelet ψ defined by

$$\psi(x) = \begin{cases} 1 & \text{for } 0 \le x < 1/2 \\ -1 & \text{for } 1/2 \le x < 1 \\ 0 & \text{otherwise,} \end{cases}$$

but for many applications, the wavelet basis used is a family of continuous functions. In fact, the most used wavelet bases are the Daubechies’ bases whose “mother” functions have no closed analytical form (for a description of this family of bases, see Daubachies, 1992 and Chui, 1992). All the wavelet decompositions in this paper will be made with respect to the DAUB#8 basis.

There are several advantages to using wavelet functions instead of the simple harmonics of the kind e^{inx} to represent a function. The wavelets in “level j” $\psi_{j\cdot}$, detect components of a specific frequency in the function to be analyzed.While the representation through Fourier analysis of very simple local events in a function requires many terms of the kind Ae^{inx}, with few wavelets ψ_{jk} (within each level j, and for different values of k) local events of a function can be easily detected. This is due to the locality of wavelet bases functions

2.2. Discrete Wavelet Transform.

From a data analysis point of view, wavelets provide a representation of functions generated by data sets. Given $\mathbf{y} = (y_0, \ldots, y_{2^n-1})$ a data vector of size 2^n, $\mathbf{y}$ can be considered as a function f on [0,1) defined by

$$f(x) = y_k \qquad \text{for } x \in [k/2^n, (k+1)/2^n).$$

This function f is square integrable and its wavelet decomposition has the form

$$f(x) = c_{00}\, \varphi(x) + \sum_j \sum_k d_{jk}\, \psi_{jk}(x)$$

where φ is a (possibly periodized) scaling function.

Given a sequence $\mathbf{a} = (a_0, \ldots, a_{N-1})$, the discrete Fourier transform of $\mathbf{a}$ is the sequence $\mathbf{b} = (b_0, .$

$. . , b_{N-1}$) defined by

$$b_j = \sum_{t=0}^{N-1} a_t \, e^{-i(2\pi j/N)t} \qquad j = 0, . . . , N-1.$$

This sequence **b** is a linear transformation of **a**, which allows us to look at the data set in the frequency domain instead of the time domain. **b** is the set of Fourier coefficients of the discrete Fourier transform of **a**. In a similar way, given a wavelet basis $\psi(_{jk}$), the discrete wavelet transformation of a data vector **x** of dimension $N=2^n$ is another data vector **d** of the same dimension as **x** which allows us to look at the data set **x** in local and frequency domain instead of in the time domain. The vector **d** is the set of coefficients in the wavelet series representation of **x** with respect to the wavelet basis chosen. This transformation is linear and orthogonal, and can be described by a NxN orthogonal matrix W (details about DWT can be found in Nason and Silverman, 1994; Vidakovic and Muller, 1994; Strang, 1993; or Arino and Vidakovic, 1995).

Since the elements of **d** are the wavelet coefficients of **x** with respect to the basis (ψ_{jk}), **d** can be written as

$$\mathbf{d} = (c_{00}, d_{00}, d_{10}, d_{11}, d_{20}, . . . , d_{n-1,2^{n-1}-1}),$$

where c_{00} is the coefficient of the scaling function φ. Apart from c_{00} there is one 0-level coefficient d_{00}, two 1-level coefficients d_{10} and d_{11}, and in general there are 2^j j-level coefficients d_{j0}, d_{j1}, . . . and $d_{j,2^j-1}$. The last level of coefficients is the (n-1)-level.

Finally we define the scalogram of **d**, which is the wavelet counterpart of the periodogram in Fourier analysis. If **d** is the vector of coefficients of the discrete wavelet transformation of **x** (i.e. **d** = W **x**), the energy of **d** at level j is defined as

$$E(j) = \sum_{k=0}^{2^j-1} d_{jk}^2 \qquad \text{for } j = 0, . . . , n-1$$

The scalogram of **d** is the vector of energies

$$(c_{00}^2 , E(0), E(1), . . . , E(n-1)).$$

The scalogram of the discrete wavelet transform of a time series is used to decompose the series into cycles of different frequencies.

2.3. Decomposition of a time series into cycles.

Let $\mathbf{x} = (x_t)$ be a data set of size 2^n. Our objective is to decompose $\mathbf{x}$ into two data sets $\mathbf{y} = (y_t)$ and $\mathbf{z} = (z_t)$ of the same dimension as $\mathbf{x}$, such that $\mathbf{x} = \mathbf{y} + \mathbf{z}$, where each one of them reflects fluctuations of (x_t) at different frequencies. We present here a brief summary of the methodology used. The details can be found in Arino and Vidakovic (1995).

If at a low level j of the wavelet decomposition of $\mathbf{x}$, most of the coefficients d_{jk} are "big", it means that a long-period, low-frequency cyclic component is present in $\mathbf{x}$. The length of the period need not be constant, because some particular d_{jk} of that level can be "small". On the contrary, if at a high-level j, most of the coefficients d_{jk} are "big", a short-period, high-frequency cyclic component is present in $\mathbf{x}$.

In the analysis of economic time series, it is usual to find a 12-month period seasonal component and a long-term trend. Thus, it is reasonable to find two peaks in the scalogram of the wavelet coefficients $\mathbf{d}$ of an economic data set $\mathbf{x}$. Splitting $\mathbf{d}$ into $\mathbf{d}^{(1)}$ and $\mathbf{d}^{(2)}$ as explained below, and applying to the split parts the inverse wavelet transform W^{-1}, the two components of $\mathbf{x} = (x_t)$ can be identified. More specifically, let us assume that in the scalogram of $\mathbf{d}$, two peaks at levels $j_1 < j_2$ are present. We divide the set {0, 1,2, . . . , n-1} of levels into two subsets A = {0, 1, . . . , j} and B = {j+1, . . . , n-1} in such a way that levels in set A are around j_1, the level of the first peak, and levels in set B are around j_2, the level of the second peak. Then, $\mathbf{d}^{(1)}$ and $\mathbf{d}^{(2)}$ are defined as follows:

$$\mathbf{d}^{(1)} = (c_{00}, d_{00}, d_{10}, d_{11}, \ldots, d_{j0}, \ldots, d_{j,2^j-1}, 0, \ldots 0)$$

$$\mathbf{d}^{(2)} = (0, \ldots, d_{j+1,1}, \ldots, d_{j+1,2^{j+1}-1}, \ldots, d_{n-1,2^{n-1}-1}),$$

and the cycles at which $\mathbf{x}$ is decomposed are

$$\mathbf{y} = W^{-1}\,\mathbf{d}^{(1)}$$

and

$$\mathbf{z} = W^{-1}\,\mathbf{d}^{(2)}.$$

Normally, $\mathbf{y}$ represents a long-term trend of the economic data set. This component should be somehow related to long-term trends of other economic magnitudes, all of them defining the business cycle of an economy. $\mathbf{z}$ usually represents the 12-month seasonal behavior of the economic time series. It should be easier to forecast $\mathbf{y}$ and $\mathbf{z}$ separately, than the whole series $\mathbf{x}$, since **x** presents a mixed cyclical behavior while **y** and **z** present, by construction, its specific periodicity: **z** should have a 12-month periodicity and **y** should be quite smooth, as it reflects the long-term trend of the series **x**.

3. Analyzing car sales in the Spanish market.

In this section we apply the methodology described in Section 2 to model and forecast car sales in the Spanish market. The data set **x** = (x_t) is the monthly car sales in Spain from January 1974 to December 1994. There are 252 observations available. The first 240 observations are used to build the model, the last 12 observations (monthly car sales during 1994) to check the forecast. A graph of the car sales is shown in Figure 1. The data set is listed in Appendix 1. We first present the Box-Jenkins model and then the wavelet model.

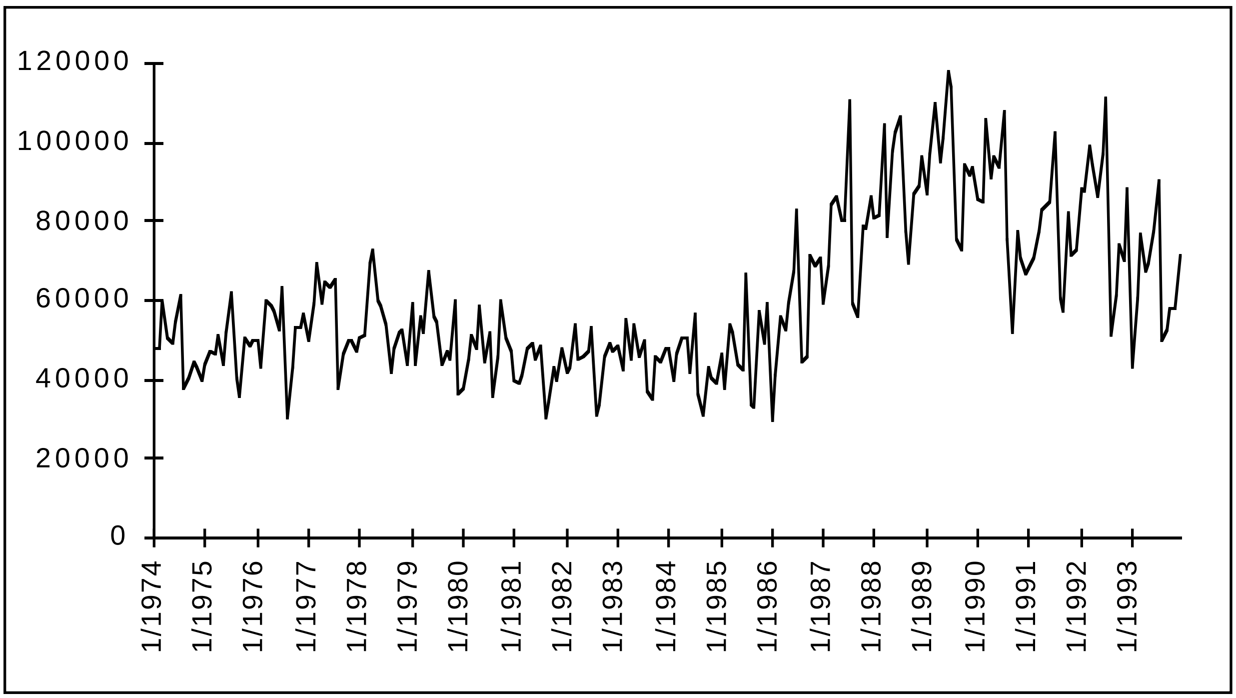

Figure 1. Monthly car sales in the Spanish market.

3.1. Box-Jenkins model.

Two Box-Jenkins models explain quite well the behavior of our time series. The first one is (as in many economic time series) an ARIMA(0,1,1)x$(0,1,1)_{12}$. The second one is an ARIMA(2,1,0)x$(0,1,1)_{12}$. Both models have robust coefficients, and their within the sample root mean square of errors are 7,612 and 7,617 respectively. Likewise, both of them give similar forecasts. The first model will be used since it is more parsimonious. The model is

$$\nabla_{12}\nabla_1 \, x_t = (1 - \underset{(12.59)}{0.64} B)(1 - \underset{(14.15)}{0.68}B^{12})\, a_t \qquad (1)$$

where operators B^k and ∇_k are defined by $B^k(x_t) = (x_{t-k})$ and $\nabla_k = 1 - B^k$. In parenthesis are presented the t-statistics of each coefficient. Monthly forecasts according to model (1) together with their standard error and actual sales are presented in Table 1. The out of the sample root mean square error (RMSE) is 16,963.9. Figure 2 shows a graph of the sales for the period 1992-1994 and the forecasts for 1994 with the 95% confidence interval.

Good references on Box-Jenkins models and standard forecasting techniques are in Box and Jenkins (1976), Pankratz (1991) and Granger and Newbold (1986).

	Forecasts	Stand. Error	Actual Sales	Error
Jan-94	54260	7612.32	54977	717
Feb-94	61128	8087.02	64490	3362
Mar-94	75276	8535.37	88819	13543
Apr-94	67910	8961.31	72145	4235
May-94	69004	9367.90	89907	20903
Jun-94	74912	9757.57	98017	23105
Jul-94	88028	10132.26	104812	16784
Aug-94	45174	10493.58	65899	20725
Sep-94	44534	10842.87	65201	20667
Oct-94	59786	11181.25	71931	12145
Nov-94	56174	11509.69	68761	12587
Dec-94	65158	11829.01	93883	28725
				RMSE = 16,963.9

Table 1. Results of model (1)

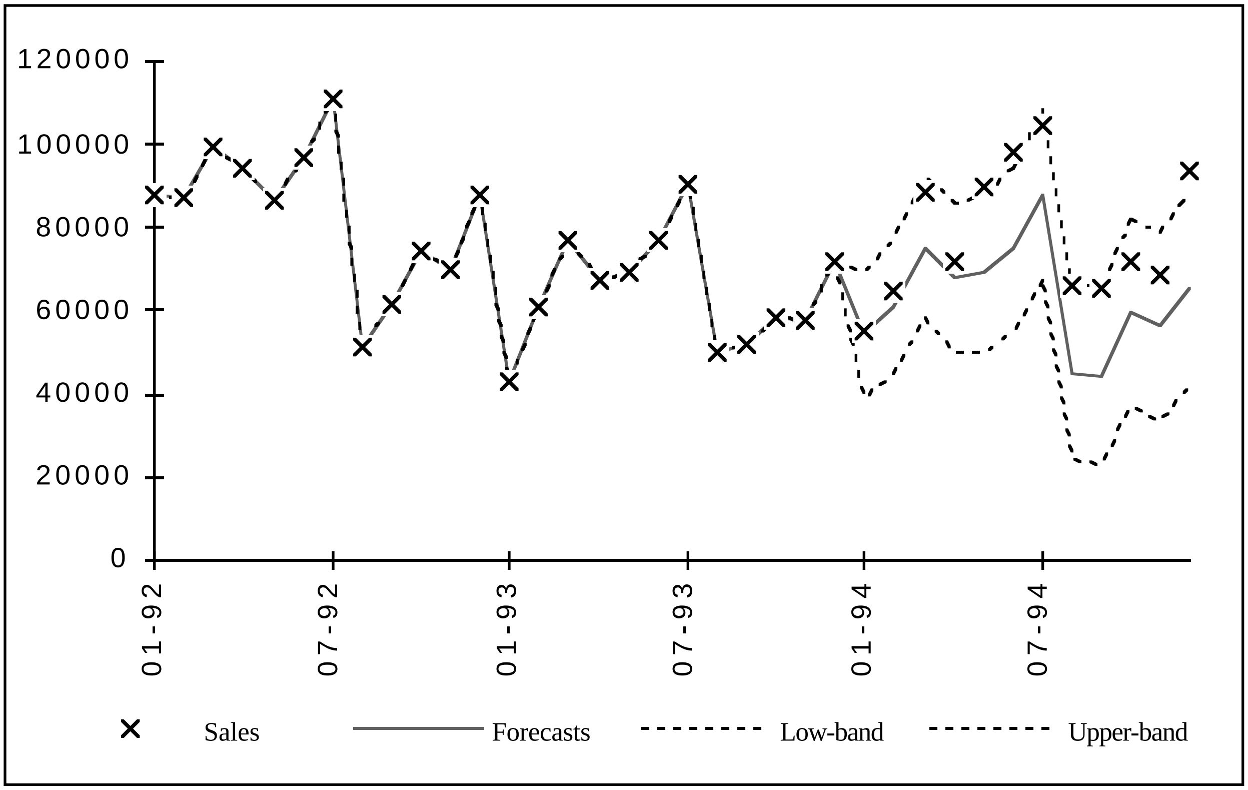

Figure 2. Actual car sales for 1992-1994 and forecasts for 1994 with the 95% confidence interval according to model (1).

We compare these forecasts with the forecasts made after decomposing the data set with the wavelet methodology. Of course, better forecasts could be made if we include in our model exogenous variables and structural models. However, the main purpose of this paper is not to forecast car sales, but to present a new forecasting technique and compare its performance with that of traditional methods.

3.2 Wavelet model.

To forecast car sales for 1994 we first need to decompose the series in its long-term trend and seasonal component as explained in section 2. To that purpose we apply the discrete wavelet transform to $\mathbf{x}=(x_t)$. Since the size of the data set is conveniently a power of two, and we only have 240 data points, we add at the end of the series the 16 point forecasts given by the Box-Jenkins model (1)[2]. Then we subtract 60,603 to each x_t to make zero the mean of this series. For simplicity, this new zero-mean series will be called $\mathbf{x} = (x_t)$ also. Then we apply the discrete wavelet transform to $\mathbf{x}$ to obtain the series $\mathbf{d}$ of its wavelet coefficients[3]. These coefficients are

[2]The first twelve forecasts are listed in Table 1 and the last four are 50,564; 57,432; 71,580 and 64,214.

listed in Appendix 2. As mentioned before, the wavelet basis used is the DAUB#8 basis. A graph of the scalogram of **d** is presented in Figure 3. (The first 0-level of **d** corresponding to c_{00} is not presented in the graph).

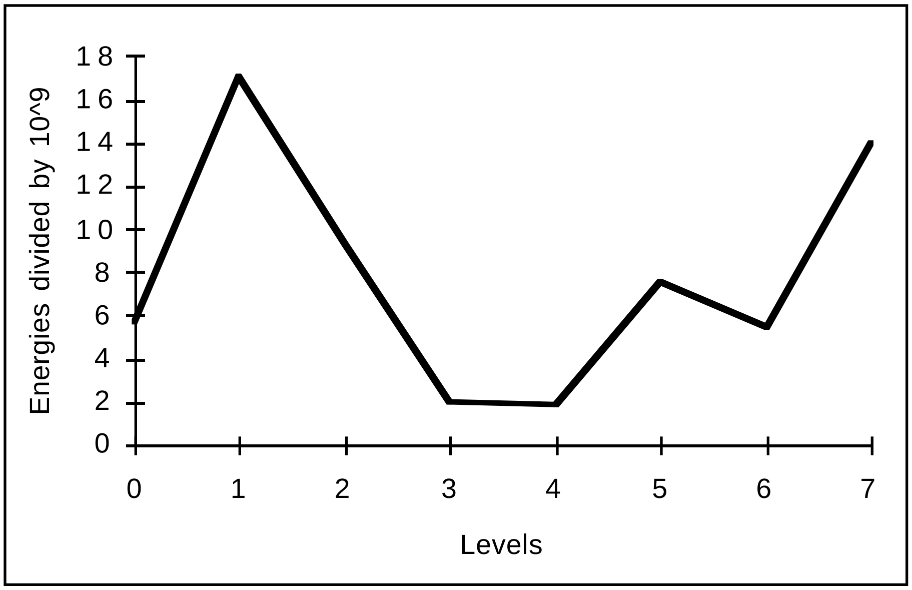

Figure 3. Scalogram of the wavelet coefficients of **x**.

Two major peaks at levels 1 and 7 are present in the scalogram. Therefore, as explained in Section 2, we split **d** into $\mathbf{d^{(1)}}$ and $\mathbf{d^{(2)}}$. $\mathbf{d^{(1)}}$ takes coefficients of **d** in levels 0, 1, 2 and 3 padded by zeros, and $\mathbf{d^{(2)}}$ takes coefficients of d in levels 4, 5, 6 and 7, padded by zeros as well:

$$\begin{aligned} \mathbf{d^{(1)}} &= (c_{00}, d_{00}, d_{10}, d_{11}, d_{20}, \ldots, d_{30}, \ldots d_{3,2^3-1}, 0, \ldots, 0) \\ \mathbf{d^{(2)}} &= (0, 0, \ldots 0, d_{40}, \ldots, d_{4,2^4-1}, d_{50}, \ldots, d_{7,2^7-1}). \end{aligned} \tag{2}$$

Applying the inverse DWT we obtain

$$\mathbf{y} = (y_t) = W^{-1}\ \mathbf{d^{(1)}}$$

and

$$\mathbf{z} = (z_t) = W^{-1}\ \mathbf{d^{(2)}}.$$

---

[3] All these computations have been made with the statistical package Splus and the wavelet package Wavetresh. Wavetresh is a noncommercial software which can be obtained via internet from lib.stat.cmu.edu or hensa.unix.ac.uk. Nason and Silverman (1994) describe the use of this software. The data set employed and all the outputs from the software necessary to reproduce and check the results are listed in Appendix 1 to 3. The dats can also be obtained electronically upon request from the author.

After adding 60,603 (the average subtracted to **x**) to each y_t, we obtain the components of **x**. The data points are listed in Appendix 3. Due to boundary effects in the wavelet decomposition and reconstruction algorithm, the decomposition does not work well for the first data points nor for the last ones. For that reason, we delete the first 36 data points in **x**, **y** and **z**, and, of course, the last 16 which had been artificially added. In this way we have the series of monthly car sales in the Spanish market and its components from January 1977 to December 1993. We will work with these series from now on. A graph of **x** and its components **y** and **z** is shown in Figure 4.

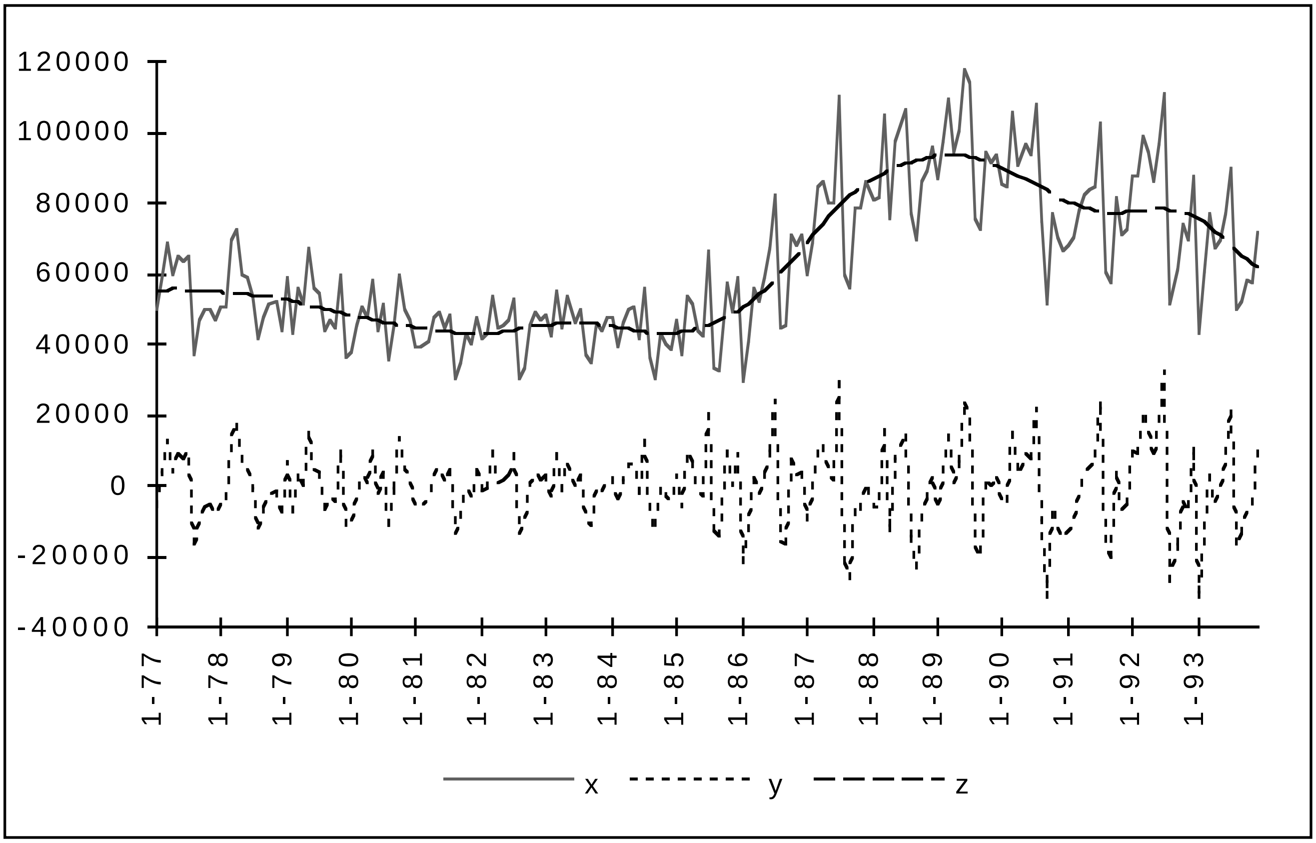

Figure 4. Series **x** and its long-term (**z**) and seasonal (**y**) components from January 1977 until December 1993.

*3.2.1. Forecasting **y**.*

y is a long-term trend time series from which seasonality has been removed. Thus, it is expected that no seasonal component will appear in the model specified for **y**. In fact, the best ARIMA model for **y** is an ARIMA(1,3,0):

$$(1 - \underset{(46.63)}{0.957}\, B)\, \nabla^3 \, z_t = a_t. \tag{3}$$

The within the sample root mean square error in this model is 6.66. The coefficient of the model is, although close, smaller than 1. Since the algorithm that estimates the parameter of the model converges we consider this model valid. Except for k = 6, 12, 18 ..., the autocorrelations function of the errors r(k) can be considered as 0. For k = 6, 12, 18 ..., r(k) are more problematic. Since there is no reason to belive that a six-month seasonality should be present in the series no further steps will be taken by now to improve the model. An alternative model will be explored in Section 4. In table 2 we present the forecasts of **y** for the twelve months of 1994 with their standard errors.

3.2.2. Forecasting ***z****.*

z is a time series which presents a seasonality of 12 months as expected. The best ARIMA model which describes its behavior is an ARIMA$(0,1,1)_{12}$:

$$\nabla_{12}\, z_t = (1 - \underset{(13.46)}{0.701}\, B^{12})\, a_t \tag{4}$$

The within the sample root mean square error in this model is 7,311. The autocorrelation function of the errors indicates that they can be considered as white noise. Forecasts of **z** for the twelve months of 1994 together with their standard errors are shown in Table 2.

This series seems to describe quite well the typical seasonality of car sales in Spain. The average of the series is almost zero, indicating that **z** captures seasonal deviations of the series **x** from its general trend: months of June and most of July exhibit over-average car sales while August and September exhibit under-average sales. Accordingly, the average of forecasts of **z** for the twelve months of 1994 is almost zero also. In addition, other forecast techniques, like classical decomposition of time series, offer similar results.

3.2.3. Forecasting ***x****.*

Forecasts of **x** for the twelve months of 1994 are the sum of forecasts of **y** and **z**. They are shown in Table 2 together with the standard errors and the actual values of **x**. It can be seen that the out of the sample RMSE of the Box-Jenkins forecast is 16,964 while for the wavelet forecast it is 12,895. Moreover, in each month, the standard error of the wavelet forecast is smaller than that of the Box-Jenkins forecast. Thus for this particular example, the wavelet method performs better than the Box-Jenkins model. In addition, the forecasts of the Box-Jenkins model underestimates the actual sales as can be seen in Figure 2 and Table 1 while the wavelet forecasts are less biased. Figure 5 shows a graph of actual car sales for the period 1992-94 and the forecasts given by the wavelet method with the 95% confidence interval.

	y Fore.	**y** Stand. Err.	**z** Fore.	**z** Stand. Err.	**x** Fore.	**x** Stand. Err.	Actual Data	Error
Jan-94	61355	6.67	-11495	7310.95	49861	7310.96	54977	5116.29
Feb-94	61154	27.21	-4596	7310.95	56558	7311.00	64490	7932.44
Mar-94	61411	70.69	9868	7310.95	71279	7311.30	88819	17540.27
Apr-94	62194	147.19	2844	7310.95	65039	7312.43	72145	7106.33
May-94	63567	267.66	4468	7310.95	68034	7315.85	89907	21872.61
Jun-94	65590	443.74	10612	7310.95	76202	7324.41	98017	21814.67
Jul-94	68322	687.69	24101	7310.95	92423	7343.23	104812	12388.89
Aug-94	71819	1012.23	-17916	7310.95	53903	7380.69	65899	11996.18
Sep-94	76135	1430.50	-18251	7310.95	57884	7449.59	65201	7317.33
Oct-94	81320	1955.97	-2292	7310.95	79028	7568.08	71931	-7097.38
Nov-94	87424	2602.39	-5449	7310.95	81975	7760.31	68761	-13213.88
Dec-94	94494	3383.75	3643	7310.95	98138	8056.04	93883	-4254.51
							RMSE=	12,895.6

Table 2. Results of "wavelet forecast" (models (3) and (4)).

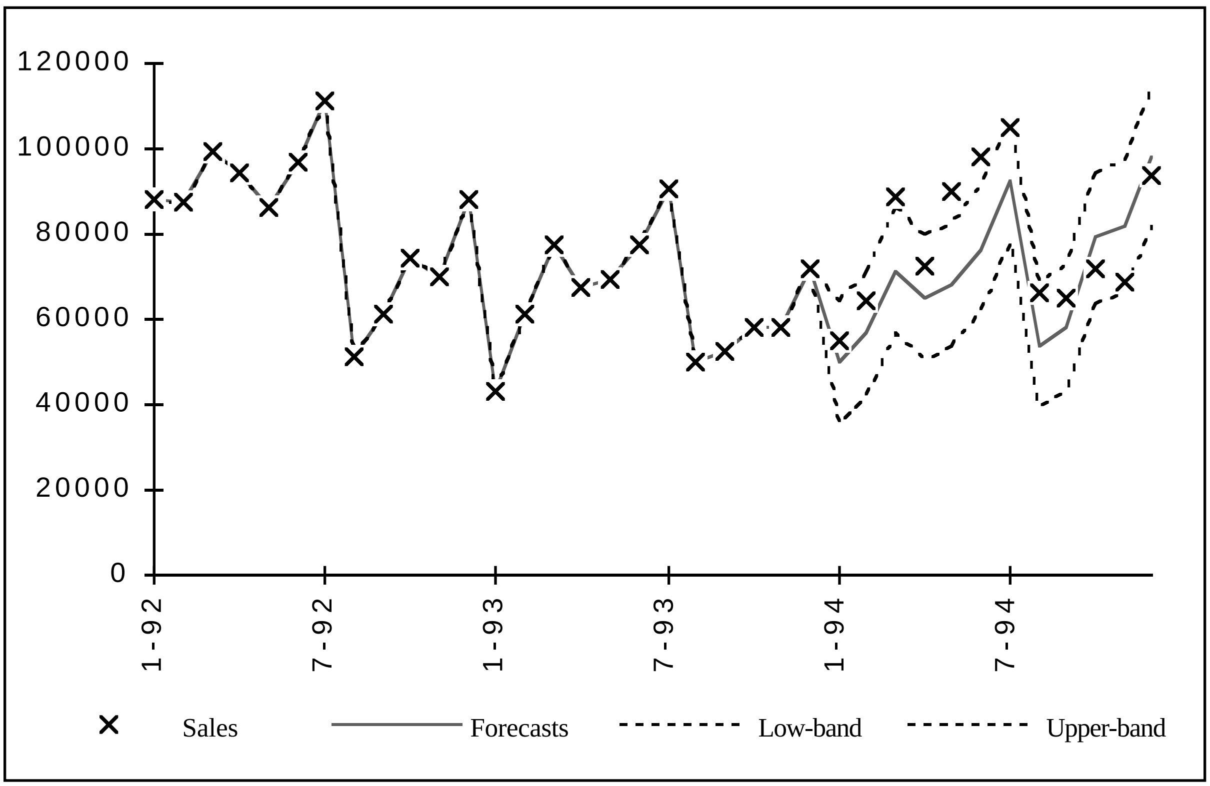

Figure 5. Actual car sales for 1992-1994 and forecasts for 1994 with the 95% confidence interval according to the “wavelet” method (models (3) and (4)).

With just one example it can not be concluded that wavelet methodology performs better than the standard techniques. Further studies will have to be made to understand the conditions under which this method is the appropriate one to use. However, the usefulness of wavelet theory in the analysis of time series, has to be better understood, and by now can not be disregarded.

4. Further considerations.

An alternative analysis of series **y** can be made to provide another global forecast for series **x**. The ARIMA model for **y** would be an ARIMA(1,4,0)x$(1,0,0)_{12}$:

$$\underset{(5.12)}{(1 - 0.356\, B)}\ \underset{(18.25)}{(1-0.8\, B^{12})}\ \nabla^4 y_t = a_t. \qquad (5)$$

The within the sample RMSE of this model is 4.1. For this model the errors can be considered as white noise. In addition the roots of the polynomial associated to the model are far away from the unit circle. By fitting this model one recognizes that a little part of the twelve months seasonal cycle in **x** remains in **y**. Had level j = 3 of the wavelet decomposition of series **x** been joined to $\mathbf{d}^{(2)}$ instead of $\mathbf{d}^{(1)}$ in (2), perhaps the total seasonality would have been removed from **y**. However we are inclined to think that it would have added noise to series **z** rather than transfering the seasonality.

Table 3 presents forecasts of **y** for the twelve months of 1994 according to model (5) and the corresponding forecasts for series **x**. To make the table self-contained we reproduce in it the forecasts of **z** given in Table 2. All forecasts are presented with their corresponding standard error.

	y Fore.	**y** Stand. Err.	**z** Fore.	**z** Stand. Err.	**x** Fore.	**x** Stand. Err.	Actual Data	Error
Jan-94	61365	4.10	-11495	7310.95	49870	7310.95	54977	5106.80
Feb-94	61195	18.32	-4596	7310.95	56599	7310.98	64490	7890.83
Mar-94	61533	50.75	9868	7310.95	71400	7311.13	88819	17418.90
Apr-94	62482	111.04	2844	7310.95	65326	7311.80	72145	6818.55
May-94	64149	210.19	4468	7310.95	68617	7313.97	89907	21290.31
Jun-94	66637	360.41	10612	7310.95	77250	7319.83	98017	20767.43
Jul-94	70051	575.00	24101	7310.95	94152	7333.53	104812	10659.58
Aug-94	74493	868.25	-17916	7310.95	56577	7362.33	65899	9321.87
Sep-94	80076	1255.40	-18251	7310.95	61825	7417.96	65201	3376.10
Oct-94	86934	1752.57	-2292	7310.95	84642	7518.08	71931	-12711.36
Nov-94	95208	2376.69	-5449	7310.95	89759	7687.57	68761	-20997.58
Dec-94	105045	3145.51	3643	7310.95	108688	7958.91	93883	-14805.45
							RMSE=	14,013.1

Table 3. Results of "wavelet forecasting" (models (4) and (5)).

It can be seen that the out of the sample RMSE for the forecasts of x in this model is 14,013.1, improving that of the Box-Jenkins model (1) by 21% ((16,963/14,013)-1). However, the wavelet model presented in Section 3, is still more accurate. Its out of the sample RMSE improves that of the Box-Jenkins model (1) by 31.5%.

If one had to chose between models (3) or (5) to describe the behavior of series y model (3) would be chosen for several reasons. The most important being that that is not a bad model in the sense that all the statistics to assess the goodness of a model are in its favor with the exception of some values of the autocorrelation function of the errors. In addition it is a simple model in which the twelve months seasonality does not appear as should be expected. Finally, its forecast performance is better than model (5), although this is not a test that qualifies a model; its is rather a consequence of a good model. We should repeat at this stage that the main purpose of this paper was to present a new method to forecast time series, by forecasting different frequencies separately.The data set was just exemplary. The method has been applied to the data set of car sales in Spain to show its possibilities.

5. Conclusion.

We have presented an application of the recently developed “wavelet” theory to analyze and forecast economic time series. The method consists of applying a discrete wavelet transform to a specific data set and splitting its wavelet coefficients into two parts. Each part, after the inverse wavelet transform is applied to each one, captures the seasonal movements of the series and the longer-term changes respectively. In this way, the time series is decomposed into two components. One of the components represents the long-term trend of the series and the other the one year periodic fluctuations along the trend, according to the seasonal nature of the series. Applying standard forecasting techniques to each component, a forecast of the total series can be obtained. The methodology has been applied to the data set of monthly car sales in the Spanish market, where succesful forecasts have been obtained. Although one example is not enough to assess the performance of the method, it indicates that there is reason for further research of applications of wavelet theory to analyze time series.

References

[1] Arino, M. A. and Vidakovic B. (1995). On wavelet scalograms and their applications in economic time series, Discussion Paper **95-21**, ISDS, Duke University.

[2] Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control,* rev. ed., San Francisco: Holden Day

[3] Chui, K. (1992). *An Introduction to Wavelets,* Academic Press.

[4] Daubechies, I. (1992). *Ten Lectures on Wavelets*, SIAM, CBMS-NST Conference Series **61**.

[5] Granger, C.W.J. and Newbold, P. (1986).*Forecasting Economic Time Series*, 2nd Ed., San Diego: Academic Press.

[6] Meyer, Y. (1992). Wavelets and Operators, Cambridge University Press.

[7] Nason, G. P. and Silverman, B.W. (1994). The Discrete Wavelet Transform in S, *Journal of Computational and Graphic Statistics*, **3**, 163-191

[8] Pankratz, A. (1991). *Forecasting with Dynamic Regression Models*, New York: Wiley Interscience.

[9] Strang, G. (1993). Wavelet transforms versus Fourier transforms, *Bulletin (New Series) of the American Mathematical Society*, **28**(2), 288-305.

[10] Vidakovic, B. and Muller, P. (1994). Wavelet for kids. A tutorial introduction, Discussion Paper **94-13**, ISDS, Duke University.

Appendix 1

Monthly car sales in the Spanish Market

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1974	47760	47654	59778	50254	49308	54737	61003	37653	40370	44303	42969	39934
1975	43948	47214	46394	51488	43658	52136	61981	40590	35968	50693	48517	49601
1976	49761	42762	59650	58365	56900	52200	63506	30258	43093	53449	53330	56403
1977	49976	59221	69055	59554	64926	63592	65377	37467	46702	49938	50059	46992
1978	50636	50962	69456	72807	60188	58801	53716	41416	48016	51817	52226	43992
1979	59374	43428	56098	51679	67079	55960	54722	43834	47005	44955	59949	36569
1980	37905	45194	50833	47518	58254	44363	51821	35478	45684	59943	50357	46799
1981	39502	39137	40775	47581	49368	44997	48687	30494	34648	42910	39933	47684
1982	41873	43091	53729	44875	45669	46931	53236	30694	33788	45843	48932	47072
1983	48602	42331	55121	45002	53608	45945	49821	36967	35288	46017	44190	47544
1984	47643	39716	46422	50404	50643	41744	56412	36192	30697	43257	40266	38833
1985	46744	37549	54185	51643	43781	42720	66489	33391	32927	57521	49047	59055
1986	29330	41088	55732	52388	59408	67537	82716	44394	45734	71291	68449	70984
1987	59510	68650	84380	86095	80288	79878	110543	59501	55764	78690	78445	86520
1988	81012	81713	104756	75912	97868	102419	106741	77493	69431	86576	89089	96210
1989	86803	97215	109727	94817	100780	117883	113877	75610	72586	94609	91543	93923
1990	85485	84873	105810	91076	96564	93946	107865	75211	51849	77259	70702	66374
1991	67998	70581	77769	82803	84268	84648	102518	60327	57351	81918	71273	72607
1992	87950	87520	99161	94201	86357	96992	111324	51450	61320	74330	69783	88066
1993	43150	60892	77093	67525	69440	77285	90288	49920	52187	58109	57913	71659
1994	54977	64490	88819	72145	89907	98017	104812	65899	65201	71931	68761	93883

Appendix 2

c_{00}

146876.3

d_{00}

-75299.6

$d_{1,k}$

-14442.6-129874.1

$d_{2,k}$

-18870.0 -10021.9 -48167.9 80346.3

$d_{3,k}$

10154.8 -7099.4 3794.8 -13815.8 20387.0 -15135.3 31567.9 5647.5

$d_{4,k}$

3625.9 -12377.5 -1741.4 7412.5 7797.3 4673.1 -1782.4 -3022.4 -5690.8 1058.2 822.4 -5091.9 8288.8
-25129.9 28434.8 -4927.9

$d_{5,k}$

-60.9 9087.8 -5984.8 -14322.4 12416.4 -814.6 -28808.5 17825.1 -9644.1 -4861.6 6637.0 -403.7 -12686.0
6366.5 5320.9 -10420.7 12989.8 -4325.5 -2734.7 20635.3 410.0 -20288.5 25842.7 -2609.5 -18335.2 34033.1
-13653.5 -24782.1 25008.2 -16042.2 -13497.2 18680.0

$d_{6,k}$

989.4	-158.6	4558.0	-2456.6	2954.8	10589.8	4858.3	-6183.8	4704.8	2885.5	-6502.7	5110.9	9299.0
-1614.3	-5807.0	1854.9	-10485.9	3823.8	6509.9	-9499.6	-2674.7	10514.0	-12139.9	9085.4	292.2	-1499.8
7068.2	3126.6	-4296.9	6722.0	4550.2	-2253.5	8269.4	3791.2	5241.9	11494.0	13287.0	-13122.6	20695.0
9406.7	2159.7	19700.0	-6882.2	-2649.3	18082.0	-6954.3	-7085.5	18488.7	4884.3	1169.9	16556.9	-2304.2
-4632.8	8919.0	-8424.7	2618.5	22040.9	13199.0	-4957.4	21926.0	-1584.3	-1307.2	17730.6	1064.9	

$d_{7,k}$

1054.2	3456.8	5651.1	-2632.7	13726.3	-7281.8	5196.7	-3041.6	-3825.8	-562.0	10593.7	-13363.9	5524.6
5156.6	2678.0	5091.6	18100.2	-9122.2	545.1	-3404.4	9967.5	-469.2	14997.0	-5131.8	4786.2	3157.3
-962.5	-3597.1	7034.0	-1733.1	9288.9	10516.1	5270.1	6992.4	1886.4	6159.4	12105.6	-7098.7	9892.5
7503.7	7365.7	-8824.0	872.2	-2462.0	-281.1	4579.5	7288.4	-6831.8	-2559.2	1989.6	6154.7	-304.5
11923.3	-8865.5	6511.6	3915.2	8794.0	3614.8	3846.0	-7520.6	2029.6	3994.8	-1215.4	9515.3	6539.8
-10851.5	7457.9	7441.3	212.5	5861.5	17293.2	-21663.5	-5448.3	-7155.2	10558.8	-184.6	19461.4	-20655.3
4211.0	-5393.2	4549.4	7153.4	25420.1	-23242.2	5611.2	3167.7	22754.0	-3274.8	12045.2	-15154.1	1988.9
-4253.6	12284.1	-10277.3	18086.2	-19552.1	5365.3	1501.8	13148.3	2202.1	9544.2	-23569.3	11656.4	-4032.1
1506.8	6530.7	18327.3	-22925.1	11139.1	2899.8	949.9	-679.3	32848.7	-19942.1	-12634.3	-10267.1	13881.6
-3255.6	22508.2	-12797.5	-4943.7	-1639.1	7768.0	-2.7	21326.4	-17971.5	-647.6	-2807.9		

Apendix 3

a) Series **y**

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1974	48200.4	50176.7	52546.5	53798.4	52688.7	50512.7	48297.1	46551.3	46095.7	46300.0	46519.7	46729.1
1975	46610.7	46470.8	46609.8	46779.0	46964.1	47093.1	47056.3	47069.6	47242.7	47514.9	47910.3	48400.1
1976	48938.1	49580.2	50322.4	51077.9	51807.1	52494.1	53115.6	53687.6	54216.7	54662.9	55013.0	55281.1
1977	55470.0	55593.7	55671.2	55708.5	55715.8	55705.9	55679.8	55627.7	55545.3	55440.6	55318.2	55182.6
1978	55039.7	54887.3	54728.5	54577.8	54440.2	54308.0	54170.9	54009.1	53807.1	53569.5	53300.0	53003.5
1979	52688.1	52350.7	51994.1	51632.2	51272.3	50917.2	50567.0	50210.8	49842.4	49470.9	49102.4	48742.8
1980	48399.3	48068.4	47748.8	47445.0	47156.7	46882.2	46617.9	46353.7	46083.1	45807.7	45529.7	45254.8
1981	44988.9	44730.9	44483.2	44252.2	44041.7	43855.6	43695.9	43558.8	43443.4	43355.0	43297.8	43276.5
1982	43295.7	43353.6	43449.1	43582.3	43750.7	43950.8	44177.0	44419.0	44668.1	44919.6	45168.4	45411.4
1983	45644.5	45857.2	46040.1	46185.0	46283.8	46333.3	46331.2	46275.1	46165.8	46006.6	45804.4	45571.3
1984	45318.4	45051.1	44775.1	44493.2	44210.4	43941.7	43700.8	43500.2	43353.3	43266.6	43246.5	43301.3
1985	43434.6	43647.1	43935.2	44288.4	44699.5	45171.8	45710.2	46325.2	47025.9	47809.0	48674.6	49627.9
1987	69071.1	70913.5	72740.4	74523.6	76240.5	77878.8	79429.8	80895.2	82276.8	83566.5	84763.8	85876.8
1988	86908.8	87863.1	88739.2	89525.1	90216.3	90824.7	91360.4	91835.9	92263.4	92640.9	92968.1	93248.9
1989	93480.3	93657.6	93771.0	93799.0	93727.2	93556.7	93291.8	92946.2	92533.0	92048.0	91494.2	90882.4
1990	90218.3	89511.4	88767.3	87978.8	87146.7	86285.3	85407.9	84532.7	83677.3	82846.1	82046.2	81288.8
1991	80581.5	79934.9	79357.6	78850.6	78417.0	78061.5	77786.9	77596.6	77490.8	77461.4	77497.9	77587.3
1992	77714.1	77866.4	78029.9	78187.5	78320.1	78402.0	78405.8	78303.5	78070.6	77692.5	77159.9	76470.1
1993	75626.3	74631.5	73499.5	72257.5	70932.5	69548.9	68133.4	66708.3	65309.5	64001.1	62853.5	61946.3
1994	61337.0	61021.1	60968.5	61112.5	61393.5	61824.3	62290.0	62463.9	62497.3	63359.5	64397.2	62094.8
1995	57757.6	61112.7	69146.9	56956.4								

b) Series **z**

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1974	-440.4	-2522.7	7231.4	-3544.4	-3380.8	4224.3	12705.9	-8898.3	-5725.7	-1997.0	-3550.7	-6795.1
1975	-2662.7	743.2	-215.8	4709.0	-3306.1	5042.9	14924.7	-6479.6	-11274.7	3178.1	606.7	1200.9
1976	822.9	-6818.2	9327.6	7287.1	5092.9	-294.1	10390.4	-23429.6	-11123.7	-1213.9	-1683.0	1121.9
1977	-5494.0	3627.3	13383.8	3845.5	9210.2	7886.1	9697.2	-18160.7	-8843.3	-5502.6	-5259.2	-8190.6
1978	-4403.8	-3925.3	14727.5	18229.2	5747.8	4493.0	-454.9	-12593.1	-5791.1	-1752.5	-1074.0	-9011.5
1979	6685.9	-8922.7	4103.9	46.8	15806.7	5042.8	4155.0	-6376.8	-2837.4	-4515.9	10846.6	-12173.8
1980	-10494.3	-2874.4	3084.2	73.0	11097.3	-2519.2	5203.1	-10875.8	-399.1	14135.3	4827.3	1544.2
1981	-5486.9	-5593.9	-3708.2	3328.8	5326.3	1141.4	4991.1	-13064.8	-8795.4	-445.0	-3364.8	4407.5
1982	-1422.7	-262.6	10279.9	1292.7	1918.3	2980.2	9059.0	-13725.0	-10880.1	923.4	3763.6	1660.6
1983	2957.5	-3526.2	9080.9	-1183.0	7324.2	-388.3	3489.8	-9308.1	-10877.8	10.3	-1614.4	1972.7
1984	2324.6	-5335.1	1646.9	5910.8	6432.6	-2197.7	12711.2	-7308.2	-12656.3	-9.6	-2980.5	-4468.3
1985	3309.4	-6098.1	10249.7	7354.6	-918.5	-2451.8	20778.8	-12934.2	-14098.9	9712.0	372.4	9427.1
1986	-21339.2	-10713.0	2710.1	-1932.9	3716.9	10402.7	24065.6	-15847.8	-16174.1	7653.8	3032.5	3750.5
1987	-9561.1	-2263.5	11639.6	11571.4	4047.5	1999.2	31113.2	-21394.2	-26512.8	-4876.5	-6318.8	643.2
1988	-5896.8	-6150.1	16016.8	-13613.1	7651.7	11594.3	15380.6	-14342.9	-22832.4	-6064.9	-3879.1	2961.1
1989	-6677.3	3557.4	15956.0	1018.0	7052.8	24326.2	20585.2	-17336.2	-19947.0	2561.0	48.8	3040.6
1990	-4733.3	-4638.4	17042.7	3097.2	9417.3	7660.7	22457.1	-9321.7	-31828.3	-5587.1	-11344.2	-14914.8
1991	-12583.5	-9354.0	-1588.6	3952.4	5851.0	6586.5	24731.1	-17269.6	-20139.8	4456.6	-6224.9	-4980.3
1992	10235.9	9653.6	21131.1	16013.5	8036.9	18590.0	32918.2	-26853.5	-16750.6	-3362.5	-7376.9	11595.9
1993	-32476.3	-13739.5	3593.5	-4732.5	-1492.5	7736.1	22154.6	-16788.3	-13122.5	-5892.1	-4940.6	9712.7
1994	-7077.4	106.8	14307.7	6797.7	7610.4	13087.8	25738.4	-17289.5	-17962.8	-3573.3	-8223.4	3063.6
1995	-7193.7	-3680.5	2433.5	7258.1								